INVESTIGATION OF HPV PATHOGENESIS IN THE HEAD AND NECK USING NOVEL TONSIL MODELS

By

DHANANJAY EVANS

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF BIRMINGHAM

Institute of Cancer and Genomic Sciences
College of Medical and Dental Sciences
University of Birmingham
April 2022

Abstract

Human papillomavirus type 16 (HPV16) is present in >90% of HPV-driven head and neck squamous cell carcinomas (HNSCC) whereas HPV18, the second most common high-risk HPV type, is rarely detected. To understand HPV16 pathogenesis in the head and neck, we developed isogenic HPV16 and HPV18 episome-containing primary human tonsil keratinocytes models, which accurately recapitulate the molecular phenotype of the tonsils, the primary site of HPV-driven HNSCC. Differentiation in organotypic raft culture showed evidence of productive infection including viral DNA amplification and expression of viral late proteins. Long-read Nanopore RNA-sequencing was used to characterise and compare HPV16 and HPV18 transcriptome organisation in undifferentiated cells. These studies demonstrated higher activity of the major viral early promoter in HPV16 cells, contributing to higher expression of HPV16 E7 oncogenes. Further, HPV18 cells showed increased levels of transcripts encoding E8^E2, a repressor of early viral transcription, suggesting that HPV18 elicits greater E8^E2-mediated regulatory control of the early promoter than HPV16. Finally, short-read Illumina sequencing was performed to investigate virus-mediated transcriptional reprogramming of host transcriptomes. Here, HPV16 was shown to establish a more immunosuppressive environment than HPV18 by strongly downregulating activities of central transcriptional immune mediators including NF-kB, IFN-α, and IFN-γ, impairing expression of critical antiviral effector genes like CD40, CCL2, and IL1B. This could provide an essential survival advantage for HPV16-infected cells in the tonsils.

Acknowledgements

I would like to acknowledge the following people for their support over this last four years. First of all, thanks to my supervisors Dr Sally Roberts and Professor Jo Parish for all their guidance, knowledge and valuable suggestions that helped shape this project and thesis. As well, thank you to the Medical Research Council for all the funding of this research, along with additional funding support that helped me through several lockdowns.

I am grateful to all past and present members of the 'HPV supergroup' for their individual teachings, practical support, and encouragement. This crucially includes Megan Burley, Jack Ferguson, Kamini Magon, Karen Campos-León, Paul McCormack, Boris Noyvert, and Gosia Wiench, each being incredible scientists that I have thoroughly enjoyed working alongside. As well, I would like to say a big thank you to Megan for helping me through my Western blot woes but more importantly for being a great friend. My time in Birmingham would not have been the same if not for all the deep chats and laughs I have shared with Megan, Kamini, Jack, Amalia, Karen, Paul, Shabana, and James.

I would like to express my deepest gratitude to my fiancée, Ruth, for her patience and support throughout this PhD, for moving to Birmingham, and for always looking after me. Finally, I would like to thank all my family in Sheffield and Telford, for their encouragement and guidance, and above all, my parents, for always keeping my head held high - I dedicate this thesis to them.

"*And by and by Christopher Robin came to the end of things, and he was silent, and he sat there, looking out over the world, just wishing it wouldn't stop*"

–A. A. Milne, The House at Pooh Corner

Table of contents

List of figures

List of tables

List of abbreviations

AKNA	AT-hook transcription factor
ALTER	Agitated low temperature epitope retrieval
ASR	Age-standardized incidence rates
BrdU	5-Bromo-2'-deoxyuridine
BSA	Bovine serum albumin
CC	Cervical carcinoma
CD40L	CD40 ligand
CDK	Cyclin dependent kinase
CIN	Cervical intraepithelial neoplasia
Ct	Threshold cycle
CTCF	CCCTC-binding factor
DAPI	4',6-diamidino-2-phenylindole
DDR	DNA damage response
DEPTOR	DEP domain-containing mTOR-interacting protein
DMEM	Dulbecco's modified Eagle's medium
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	EGF receptor
EMT	Epithelial mesenchymal transition
FBS	Foetal bovine serum
FCS	Foetal calf serum
GSEA	Gene set enrichment analysis
H&E	Haematoxylin and eosin
HBRC	Human Biomaterials Resource Centre
HFK	Human foreskin keratinocyte
HFK-18	HFK maintaining HPV18 genomes
HINGS	Heat inactivated goat serum
hnRNP	Heterogeneous ribonucleoproteins
HP	High passage
HPV	Human papillomavirus
HPV-OPSCC	HPV-associated oropharyngeal squamous cell carcinoma
HPV16-HTK	Human tonsil keratinocytes transfected with HPV16
HPV16/18-HTK	Human tonsil keratinocytes transfected with either HPV16 or HPV18
HPV18-HFK	Human foreskin keratinocytes transfected with HPV18
HPV18-HTK	Human tonsil keratinocytes transfected with HPV18
HR-HPV	High Risk HPV
HSIL	High-grade squamous intraepithelial lesions
HTK	Human tonsil keratinocytes

IF	Immunofluorescence
IFN-α	Interferon Alpha
IFN-β	Interferon Beta
IFN-γ	Interferon Gamma
IFNAR	IFN-α/β receptor
IGV	Integrated Genomics Viewer
IL-1R	IL-1 family receptors
IL18R	IL18 receptor
IRF	Interferon regulatory factor
ISG	Interferon stimulated gene
J2-fibroblasts	NIH-3T3 J2 mouse fibroblasts
K1	Keratin 1
LCR	Long control region
LP	Low passage
LSIL	Low-grade intraepithelial lesions
MCM	Minichromosome maintenance protein complex
MHC-I	Major histocompatibility complex I
MHC-II	Major histocompatibility complex II
mRNA	Messenger RNA
MSigDB	Molecular signatures database
NF-kB	Nuclear Factor Kappa B
NIKS	Normal immortalised keratinocytes
NKT	Natural killer cell
NS	Non-significant
OPSCC	Oropharyngeal squamous cell carcinoma
ORF	Open reading frame
pAE	Early polyadenylation signal
pAL	Late polyadenylation signal
PAMP	Pathogen-associated molecular patterns
PBM	PDZ binding motif
PBS	Phosphate-buffered saline
PC1	First principal component
PC2	Second principal component
PCA	Principle component analysis
PCDF	Polyvinylidene fluoride
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PD	Population doublings
pE	Early viral promoter
pE8	E8 promoter
PI3K	Phosphoinositide 3-kinase
pL	Late viral promoter
PRR	Pattern recognition receptor
q-value	Adjusted p-value
qPCR	Quantitative polymerase chain reaction

Rb	Retinoblastoma
RCB	Reconstitution buffer
RIN	RNA integrity numbers
RNA-Seq	RNA sequencing
RPM	Reads per million
SA	Splice acceptor
SCC	Squamous cell carcinoma
SCC	saline-sodium citrate
SD	Splice donor
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SFM	Serum-free keratinocyte media
SRSF	Serine-arginine rich splicing factor
SSC	Saline-sodium citrate
STAT	Signal transducer and activator of transcription
TBE	Tris-Borate-EDTA
TBS	Tris-buffered saline
TBS-T	Tris-buffered saline containing 0.1% Tween 0.20
TE	Tris-EDTA
TEM	Tetraspanin enriched microdomains
TLR	Toll-like receptor
TNF-α	Tumour necrosis factor alpha
TNFR	Tumour necrosis factor receptor
TRAF	TNFR-associated factor
TSS	Transcriptional start site
U2OS	Human osteosarcoma-derived cells
ULB	Urea lysis buffer
UV	Ultraviolet

Chapter 1: Introduction

1.1 Human papillomavirus

1.1.1 Human papillomavirus

Human papillomaviruses (HPV) are non-enveloped, circular, double-stranded DNA viruses of the *Papillomaviridae* family, containing small genomes at around 8 kilobases (kb) size. HPV are common epitheliotropic viruses that selectively infect basal keratinocytes of stratified and mucosal epithelium of the skin, anogenital and oropharyngeal-laryngeal tracts. The *Papillomaviridae* family are thought to have evolved alongside humans and other animals for approximately 350 million years, resulting in a genomically and phenotypically diverse group of viruses [1, 2]. Over 250 different papillomaviruses are now recognised (https://pave.niaid.nih.gov), infecting a diverse range of animals, including mammalian, reptilian, and avian species. Over 200 types of HPV have been fully sequenced, 13 of which have been identified as type 1 carcinogens [3]. These are classed as high-risk and are transmitted sexually via anogenital or oral-genital transmission. High risk-HPVs (HR-HPVs) are the most prevalent sexually transmitted infection in the UK that are mainly linked to cancers of the cervix, vulva, vagina, penis, anus, and oropharynx [4]. These contribute to a substantial proportion of the global cancer burden and are thought to account for approximately 4.5% of all global cancer cases [5].

1.1.2 HPV classification

In humans, HPV are classified into one of five genera (alpha, beta, gamma, mu and nu), which are assorted based on the nucleotide sequence of the L1 open reading frame (ORF), as well as epithelial tissue-type tropisms, life cycle features, and carcinogenic potential (see Figure 1-1) [2, 6, 7]. The majority of HPV belong to alpha, beta, and gamma genera. Alpha HPVs infect both cutaneous and mucosal epithelia, whereas beta, gamma, mu, and nu HPV-types primarily infect cutaneous epithelia. While HPV of these latter four genera, namely beta-HPVs, have been infrequently associated with skin cancer development in immunocompromised individuals, infections by these genera generally cause benign cutaneous lesions and warts and are not typically associated with disease progression [8-11].

Alpha HPVs are the most characterised HPV types due to their association with cancer. These are subdivided into low- and high-risk types, comprising over 65 HPV genotypes in total (Figure 1-1). The low-risk alpha HPV are generally associated with benign lesions and warts in anogenital regions and are not commonly associated with carcinogenesis. These include HPV6 and HPV11, which are attributed to a high proportion of benign genital warts and also infrequently cause recurrent respiratory papillomatosis, a rare and difficult to treat disease in which HPV drive production of wart-like growths that line and restrict the larynx [12-15]. HR-HPV have high oncogenic potential, being considerable drivers of squamous cell carcinoma (SCC) and adenocarcinomas and include HR-HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68. Of these, HPV16, HPV18, and HPV31 have been amongst the most well researched types due to their high prevalence rates in both anogenital and head and neck cancers, where HPV16 and HPV18 contribute to 72.4% of HPV-driven cancers [5]. These

differences in oncogenic potential between high- and low-risk subtypes are largely due to alternate affinity of viral E6 and E7 proteins to the tumour suppressor proteins p53 and Rb, which centrally mediate senescence and apoptosis, as later discussed in section 1.3.3 [16-18].

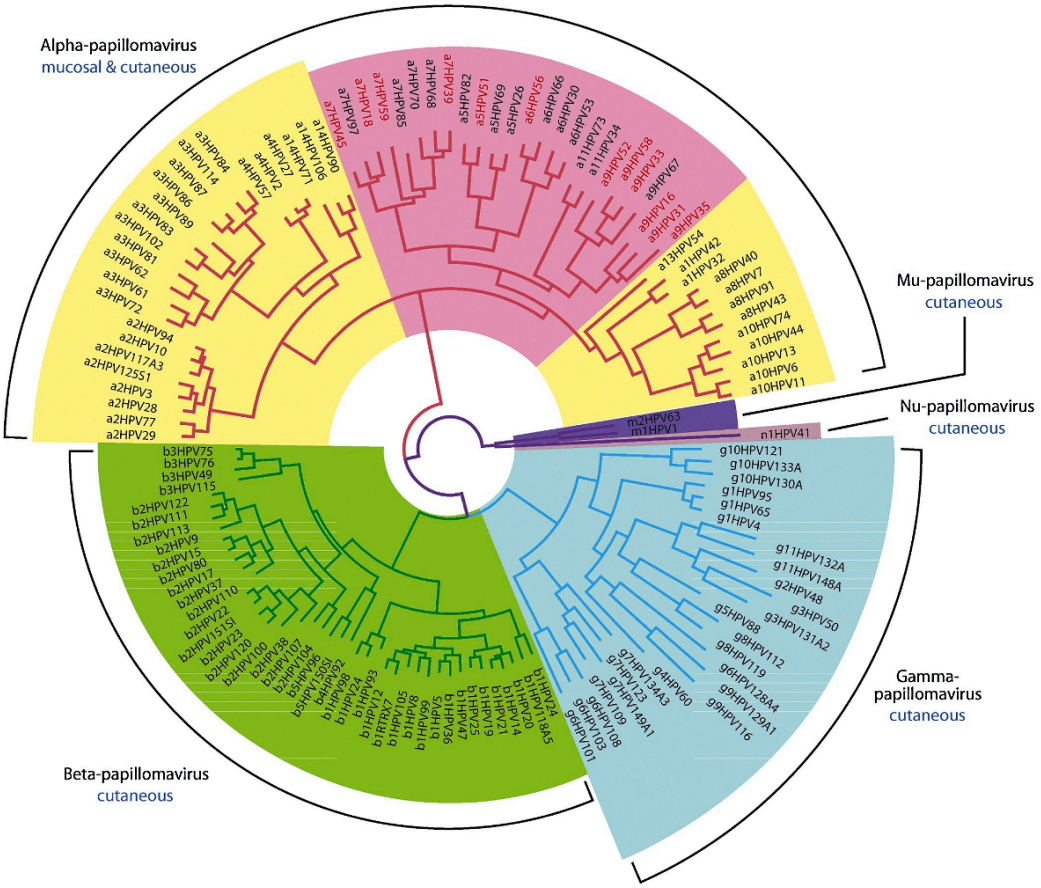


Figure 1-1: Phylogenetic categorisation of HPV types

Members of each genera share at least 60% DNA sequence similarity in the L1 open reading frame. These form the five major HPV genera alpha (yellow and pink), beta (green), gamma (blue), mu (purple) and nu (burgundy). The alpha HPVs are further subdivided into high-risk (pink) and low-risk (yellow) types. The alpha, beta and gamma lineages comprise the three largest categories. To form a new HPV type, HPV genomes must demonstrate at least 10% dissimilarity from other genera members. HPV subtypes are defined when sharing between 2 – 10% L1 ORF sequence similarity, and variants will have no less than 2% dissimilarity [19, 20]. Image taken from Egawa and Doorbar [21].

1.1.3 HPV and cancer

HR-HPV are the most common sexually transmissible viral infections in the world, affecting the overwhelming majority of unvaccinated, sexually active individuals at some point in their lifetimes [22]. The asymptomatic nature of HR-HPV infection along with high infectability and long-lasting infections all underlie high levels of transmission. Infections normally last from a few months up to 2 years following transfection and will usually resolve spontaneously without developing any pathologies [23-25]. However, a small proportion of infected individuals will maintain longer lasting, persistent infections that may develop into low-grade neoplasia, which can further progress into high-grade, or in a select number of cases, invasive cancer. However, HPV-associated malignancies are thought to take decades to develop following initial infection [26, 27].

While the vast majority of HR-HPV infections do not result in malignant disease, HPV-driven cancers still account for a high proportion of the global cancer burden, largely due to high infection rates. According to data collected by GLOBOCAN for 2012 (http://globocan.iarc.fr/), which was evaluated by de Martel, Plummer [5], HPV were estimated to contribute to 4.5% of all worldwide cancers, currently accounting for around 690,000 new cases each year [28]. This was predominately attributed to women, constituting 8.6% of all female cancers as opposed to just 0.8% of all male cancers. This was largely skewed due to high incidence rates in cervical tissues, which are particularly susceptible to HPV-driven disease progression and make up approximately 80% of all HPV-associated cancers [5, 28].

1.1.4 The epidemiology of HPV-HNSCC

The number of head and neck squamous cell carcinomas (HNSCC) totalled 890,000 new cases worldwide along with 507,000 deaths in 2017 [29]. HNSCCs develop in the mucosal surfaces of the upper digestive tract consisting of the oral cavity, pharynx, larynx, and salivary glands. Significant risk factors for HNSCC development include excessive smoking, alcohol consumption, and genetic predisposition, as reviewed by Johnson, Burtness [30]. HR-HPV have also been widely linked to HNSCC development, particularly in the oropharynx, oral cavity, and larynx (Figure 1-2a), which according to data from the GLOBACAN 2018 report were attributed to 52,000 global cases, up from 37,200 in 2012 [5, 28]. As such, HPV-associated HNSCCs (HPV-HNSCC) are now recognised as a significant and distinct subclass of HNSCCs. Of these 52,000 cases, 42,000 originated in the oropharynx, 5,900 in the oral cavity, and 4100 in the larynx. Numerous studies have now specified the oropharynx as the primary region of HPV-HNSCC, where HPV-associated oropharyngeal SCC (HPV-OPSCC) commonly originate within specialised reticulated epithelium of the palatine tonsils and base of tongue [4, 31-35].

As the number of people smoking has been gradually decreasing over the last three decades, so have the incidences of HPV-negative HNSCC. Conversely, incidences of HPV-HNSCC have become increasingly prevalent each year [5, 28, 36, 37]. This is especially true for more developed regions such as the UK, Europe, and particularly in the US, where HPV-OPSCC have now become more common than HPV-negative OPSCC [31, 38-41]. In the US, HPV-associated HNSCC numbers have now also overtaken cervical cancers [42], an epidemiological landmark that is also expected to occur, in time, in other developed regions [33, 38-41]. In the UK specifically, the incidence of HPV-OPSCC has tripled in men and doubled in women between

1995 and 2011 with an overall 100% increase in yearly cases [31]. Further, meta-analysis of just under 4,500 individuals suggests that around 4.5% of individuals have either an oral or oropharyngeal HPV infection at any given time [43]. Given these significant increases in incidences, HPV-associated OPSCCs are now being described as epidemic in Europe and the US [32, 40, 44, 45].

In direct contrast to cervical cancers, cases of HPV-associated OPSCCs are much more prominent in developed regions of the world such as Europe and the United States than in developing areas (Figure 1-2b) [28, 33, 39, 46]. It is not clear why this is the case, however differences in oral sexual behaviour could be a contributing factor. In addition, in the US, HPV-associated HNSCC has been shown to be more predominant in white as opposed to non-white populations, which may suggest differential genomic or environmental factors [47]. Interestingly, HPV-HNSCC is also more prevalent in men than in women, where case rates have been reported at between 2- to 6-fold higher in men [48-50]. This could be due to a combination of factors, including differences in sexual transmission routes, lifestyle behaviours, hormonal factors, and most notably, genetic/epigenetic variability [51-54].

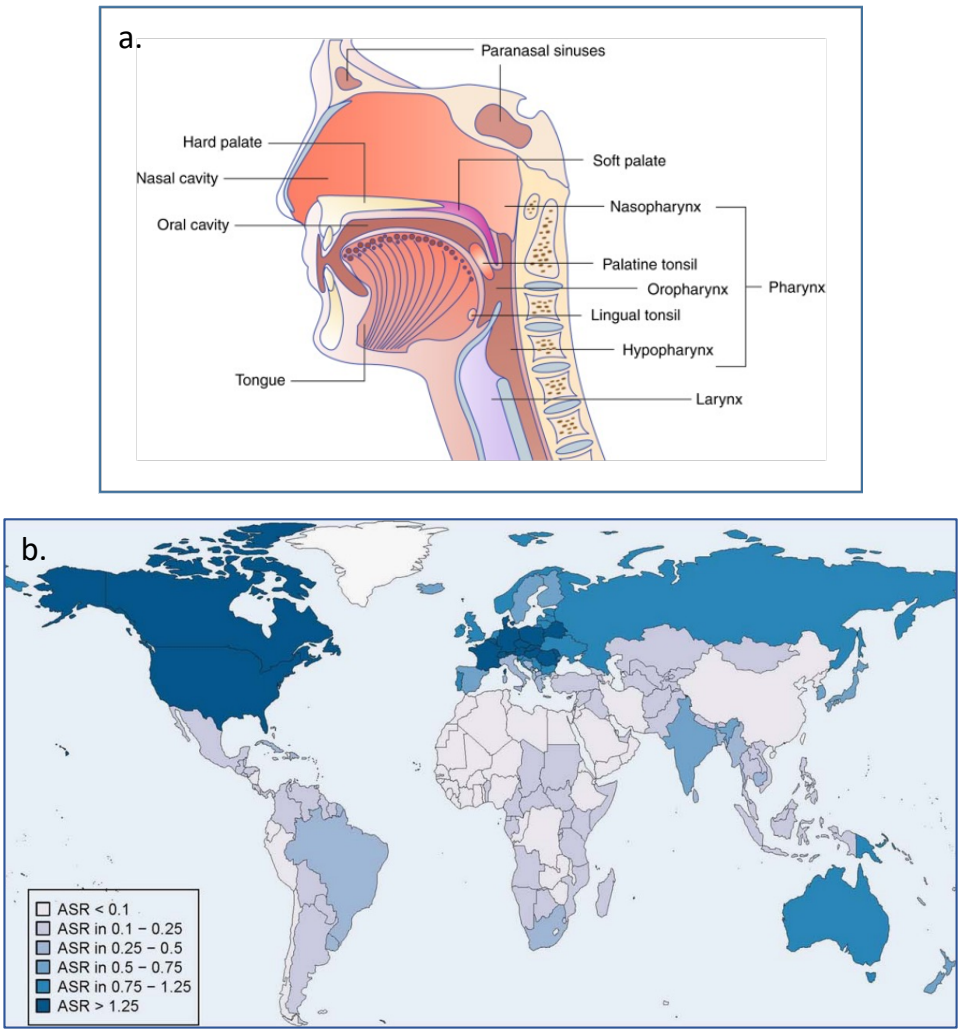


Figure 1-2: Anatomical and global locality of HPV-HNSCCs

a. A schematic representation of head and neck subsites, taken from Sabatini and Chiocca [55]. HPV infections and HPV cancers in the head and neck most commonly originate in oropharyngeal subsites, particularly within specialised reticulated epithelium of the palatine tonsils and base of tongue [34, 35, 56]. HPV-HNSCC also originate, to a far lesser extent, in the oral cavity and larynx [28]. **b.** Estimated global HPV-associated head and neck cancer cases. Produced by de Martel, Plummer [5]. Demonstrates age-standardized incidence rates (ASR), per 100,000, of HPV-attributable HNSCC by country. Findings are based on GLOBOCAN 2012 data for global head and neck cancer cases, which was adjusted to the proportion of HNSCC containing HPV E6 and E7 mRNA. Here, HPV-HNSCC is demonstrated to be most frequent in continents such as North America and Europe.

1.1.5 Presentation, prognosis, and treatments of HPV-HNSCC

Initial symptoms of both HPV-positive and HPV-negative HNSCCs commonly include sore throats and neck masses [57]. However, these tumour types normally present at different anatomical locations. Whilst the vast majority (>80%) of HPV-HNSCCs appear in the oropharynx, typically within the tonsil crypts [4, 31-35], HPV-negative HNSCCs are much less prevalent in the oropharynx, accounting for <30% of cases [58]. Further differences in histopathological appearance are frequently observed between HPV-positive and HPV-negative OPSCC, where HPV-containing tumours consist of non-keratinizing and poorly differentiating cells with a more basaloid-like phenotype, and HPV-negative OPSCC consist of normally keratinizing cells [59-61]. Distinguishable molecular characteristics also exist between HPV-positive and HPV-negative OPSCC (further discussed in section 1.4.7), which are established during diagnosis to differentiate the two cancers. One commonly used differential marker is p16, which is ordinarily overexpressed in HPV-OPSCC (see section 1.3.3), but not in HPV-negative OPSCCs, which can be detected by immunohistochemical staining of a tumour biopsy.

Despite distinct molecular differences between HPV-OPSCC and HPV-negative OPSCC, treatment guidelines remain similar between these alternative cancer subtypes. In the UK, current treatment strategies normally consist of surgical excision of tumours, often followed by primary radiotherapy and in cases of advanced disease, a combination of chemotherapy and radiotherapy [62]. With such conventional treatments, two-year progression-free survival with HPV-OPSCC is around 75 – 86%. These are associated with a much more favourable prognosis than HPV-negative OPSCC, where two-year progression-free survival lies around

53% [63, 64]. This is largely due to increased sensitivity and responsiveness of HPV-OPSCC to radiotherapy and chemotherapy, along with a lower mutational burden [64, 65]. Despite this more favourable therapeutic response, disease recurrence occurs in around 10 – 25% of HPV-OPSCCs, usually taking place around two to five years after treatment, something that may be associated with mutations of *PIK3CA* [66, 67].

Clinical trials are currently underway to find more targeted and less invasive HPV-OPSCC treatment strategies, which aim to reduce treatment morbidity and improve survival in cases with recurrent disease. This includes trials such as the EVADER trial (NCT03822897), which aims to determine minimal dose radiotherapy and chemotherapy that still effectively treats early stage OPSCC. Trials for more targeted treatments include epidermal growth factor (EGFR)-inhibitors such as Cetuximab, where EGFR is normally overexpressed in HPV-HNSCCs as well as in a significant subset of HPV-negative HNSCCs [68-70].

Finally, a number of immunotherapies are in development which aim to stimulate clearance of cancerous HPV-containing cells either by vaccine priming of the immune response against viral antigens such as E6 and E7 [71], or via targeted reactivation of innate immune responses which have been impaired by the HPV oncoproteins E5, E6 and E7 (see section 1.4.5). Prospective OPSCC-targeting immunotherapies have been recently reviewed by Frazer and Chandra [72] and Lechner, Liu [73]. However, one exciting immunotherapy target for HNSCC is the PD-1/PD-L1 checkpoint axis. PD-L1 is an immune dampening antagonist that is secreted by epithelial cells. This binds to PD-1 receptors on T-cells to prevent autoimmunity [74, 75], though these immune privileges are thought to favour HPV infection and carcinogenesis (see

section 1.4.5) [76, 77]. Two drugs, pembrolizumab and nivolumab, which are PD-1/PD-L1 checkpoint inhibitors, have been approved by the FDA for the treatment of recurrent or metastatic HNSCCs, demonstrating superior survival and lower toxicity than chemotherapy [78-81]. As such, pembrolizumab, either on its own or in combination with chemotherapy, has now been approved as a first-line treatment for patients with recurrent or metastatic HNSCC [78-81].

1.1.6 HR-HPV targeting vaccines

Successful vaccine and smear testing programmes have had a huge impact towards cervical cancer prevention in higher-income countries [5, 28]. Very recently, a real-world UK-based study looking into the overall effectiveness of the bivalent vaccine, Cervarix®, comprising HPV16 and HPV18 L1 virus-like particles, which was routinely given to teenage girls between 2008 - 2010, showed an almost 90% reduction in cervical cancer cases when compared with an unvaccinated cohort [82]. Three effective HPV vaccines are now commonly implemented around the world: Gardasil®, Cervarix®, and Gardasil® 9. These each provide effective protection against HPV16 and 18, and in some cases, other HR-HPV types 6, 11, 31, 33, 45, 52, and 58, which altogether protect against 90% of cancer-causing virus types.

These vaccines have also been shown to provide good protection against oral HPV infection, which will likely also significantly reduce the risk of HPV-mediated progression in the head and neck [83]. However, many countries, including those in more developed regions, currently only vaccinate girls. In the UK, HPV vaccines have only recently (from 2019) been extended to boys. Further, vaccine uptake in some developed countries such as the US is low [83]. It is

therefore likely that HR-HPVs will remain responsible for a significant proportion of cancers, for many decades to come. This is particularly true for developing countries that do not currently have as widely available vaccine programmes [5, 28, 83].

1.2 HR-HPV genome organisation

HR-HPV genomes typically contain eight open reading frames (ORF), in addition to a long control region (LCR), containing the viral origin of replication, transcriptional promoter sequences and cis-elements that regulate viral transcription initiation and polyadenylation (Figure 1-3) [84-86]. Viral gene expression is a highly regulated program that is adjusted in a temporal manner closely following the epithelial differentiation programme. This ensures that production of specific viral proteins occurs as they are required at different stages of the viral life cycle. Differential expression of early and late viral genes is largely attributed to regulation of two major (early and late) viral promoters and polyadenylation signals [87-89]. Studies of HPV genome organisation and transcription have so far primarily focussed on HPV16, HPV18 and HPV31, mostly in anogenital models, lesions, and cancers, as well as in epithelial osteosarcoma derived U2OS cells, but is so far lacking in oropharyngeal regions.

1.2.1 Differentiation-dependent HPV promoter activity and polyadenylation

During epithelial migration from basal to superficial epithelium, viral gene expression is broken down into three parts. Firstly, in basal keratinocytes and lower epithelial layers, transcription of early viral transcripts starts at the early promoter (pE), located in the LCR at position (p)97 of the HPV16 genome and p102 of the HPV18 genome, and ends at the early

polyadenylation (pA^E) site (Figure 1-3). This primarily drives transcription of the oncogenes, E6 and E7 and, to a lesser extent, transcription of E1, E2, E4, E5, and E8, as is later illustrated in Figure 1-6 [90, 91]. Secondly, in mid and upper intermediate layers, differentiation-dependent activation of the late viral promoter (pL), located in the E7 ORF at around p670 (HPV16) and p811 (HPV18), along with continued use of the pA^E site, significantly upregulates expression of E1, E2, E4 and E5, which are required for viral genome amplification [87, 92, 93]. During the final stages of cellular differentiation in the upper epithelial layers, exclusive usage of the pL, along with termination at late polyadenylation (pA^L) sites, located in the LCR, promote transcription of the late viral transcripts, L1 and L2, as well as E4 [94, 95]. This requires read-through of the early polyadenylation site [90].

Three other viral promoter regions have also been described; the E8 promoter (pE8), located around p1301 (HPV16) and p1357 (HPV18), which initiates transcription of the splice product, E8^E2, as well as two other lesser defined promoters, p500/p520 and p3000/p3500 [90, 91, 96, 97]. Transcripts starting at the p500/p520 promoter are predicted to encode E7, E5 and E1^E4 in HPV16- and HPV18-containing cells [90, 91, 96, 98, 99], and transcripts beginning at the p3000 promoter are considered to encode E5, as recently demonstrated in the Parish lab in HPV18-transfected HFK models [98]. Here, activity of p520 and p3000 promoters was shown to increase following cellular differentiation but remained at relatively low levels in comparison to pE and pL.

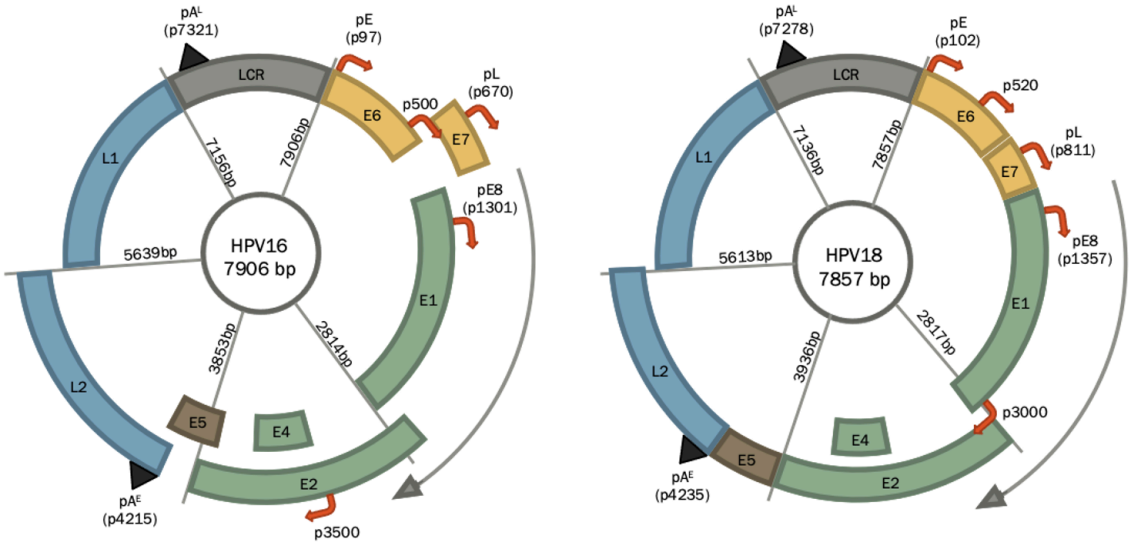


Figure 1-3: HPV16 and HPV18 genome organisation

HPV16 and HPV18 genomes are made up of double stranded circular DNA. The diagram depicts the open reading frames of all non-spliced early and late viral genes that make up each viral genome. Viral promoters are labelled with red arrows at the start of each promoter region, consisting of the early promoter (pE), the late promoter (pL), the E8 promoter (pE8), the p500/p520 promoter and the p3000/p3500 promoter. The start of the early (pA^E) and late (pA^L) polyadenylation site regions are marked with black arrowheads.

1.2.2 Alternative splicing of the HPV genome

Splicing is the function by which noncoding introns are removed from pre-mRNA transcripts and exons are ligated together to form protein-encoding mRNA, a process that is catalysed by macromolecular ribonucleoprotein (RNP) complexes known as spliceosomes. Here, the spliceosome recognises splice donor sites, located on the 5' end of a splice junction, and splice acceptor sites, located on the 3' end. These 5' and 3' splice sites contain sequences specific to

components of the spliceosome, which consists of five small nuclear RNPs (snRNPs), U1, U2, U4, U5 and U6 [100, 101].

Splicing efficiency is controlled principally by *cis*-acting regulatory sequences located within exons and introns and trans-acting proteins [102, 103]. Here, exons contain both exonic sequence enhancers (ESE) and exonic sequence silencers (ESS). Similarly, introns contain intronic sequence enhancers (ISE) and silencers (ISS) [104]. Sequence enhancers are able to bind with the highly conserved, host-cell, serine-arginine rich protein (SR protein) family members such as the serine-arginine-rich splicing factors (SFRS) proteins 1 – 7 and 9, which promote splicing by several methods including assisting and stabilising early spliceosome complex formation [105, 106]. Conversely, sequence silencers interact with members of the heterogeneous nuclear RNP (hnRNP) family which inhibit splicing, largely via antagonising the activities of SR proteins [107].

Whilst the process of constitutive splicing removes all intronic mRNA from a pre-mRNA transcript, a second process of 'alternative splicing' involves the selection of different combinations of splice sites within a transcript, resulting in additional protein diversity [108]. For HPV, alternative splicing allows for expression of numerous viral transcripts encoding a diverse range of proteins required for viral life cycle initiation, maintenance, and completion, all originating from a single compact genome [91, 96, 97, 109]. Due to this combination of alternative promoter usage, polyadenylation and splicing, 20 - 30 variable polycistronic mRNA transcripts have been described for HPV16 and HPV18, each encoding functionally distinct proteins (https://pave.niaid.nih.gov/#explore/transcript_maps) [85, 91, 96, 97, 109, 110].

The transcriptional arrangements of HPV16 mRNA species are illustrated in Figure 1-4. These demonstrate the polycistronic nature of HPV transcripts, which can encode several functionally distinct proteins. HPV mRNA transcripts contain numerous translation initiation codons (AUG), which are contained within translation promoting Kozak consensus sequences. Translation occurs most efficiently from the first AUG following upstream ribosomal attachment at the 5' end of an mRNA transcript. However, efficiency is further defined by the strength of the Kozak sequence [111]. A strong Kozak sequence can effectively impede scanning ribosomes from accessing start codon sequences at downstream ORFs. To overcome this, differential promoter usage and alternative splicing ensure the translation of a specific viral protein by generating mRNA species where a required viral gene is either displayed as the most 5' ORF, or is preceded by a weak Kozak sequence that may permit ribosomes to scan upstream AUG, in a process termed 'leaky scanning' [103, 112]. In addition, the regulated processes of ribosomal shunting, or translation-reinitiation may allow ribosomes to skip upstream AUG and initiate downstream translation [113-115].

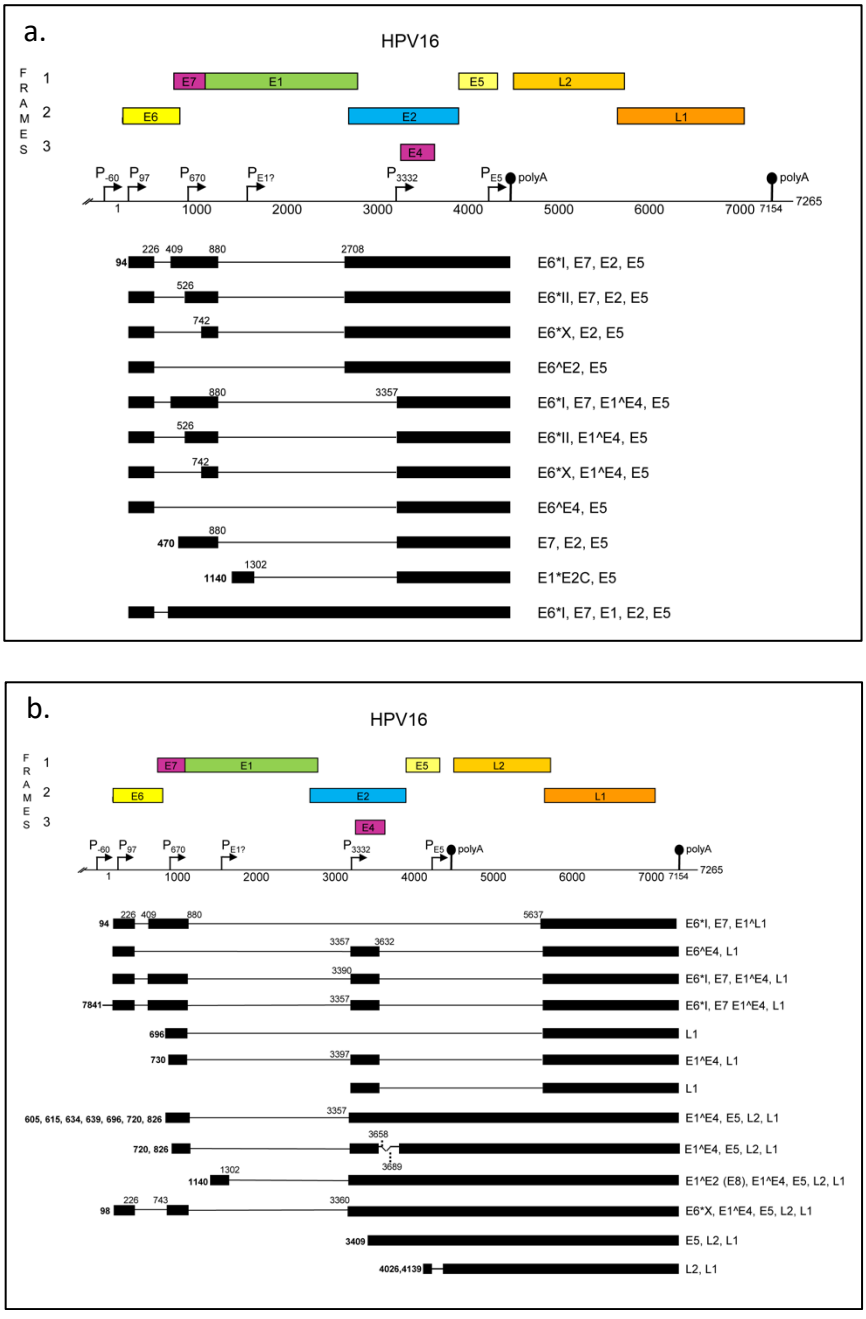


Figure 1-4: Schematic of early and late HPV16 mRNA transcripts

Taken from https://pave.niaid.nih.gov. **a**. shows HPV16 transcripts ending at the early polyadenylation site and **b**. shows transcripts ending at the late polyadenylation site. The top of each schematic depicts the linearized HPV16 genome, where open reading frames (ORF) are represented by different colours. The relative nucleotide positions of the viral early and late promoter regions and polyadenylation sites are shown just below the linearized genome. Viral mRNA transcripts demonstrate transcriptional start sites, splice donor sites, splice acceptor sites, and transcript termination sites. The predicted coding potential for each transcript is shown on the right.

1.2.3 Regulation of HPV gene transcription and alternative splicing

The fine-tuning of viral gene expression during epithelial differentiation is crucial for successful life cycle completion. This is regulated at both transcriptional and post-transcriptional levels via modulation of viral promoter usage, alternative polyadenylation, alternative splicing, mRNA stability, and translation [85, 116]. A number of virus and host elements play important and cooperative roles in modulating these transcriptional elements in a differentiation-dependent manner. These crucially include the cellular splice enhancing SR proteins and splice repressing hnRNPs, which as described in section 1.3.2, respectively stimulate, and repress RNA binding elements located on the viral genome [85, 106]. Production and activity of SR proteins and hnRNPs are up/downregulated throughout cellular differentiation in a manner that supports early and late viral life cycle gene expression [117]. This is largely regulated by HPV E2 proteins, which can both transactivate and recruit SRSF proteins to specific nucleotide sequences in the viral genome [87, 118-120].

HPV E2 plays many important roles in the viral life cycle but is most commonly recognised as a central mediator of viral gene expression, as well as of DNA replication, and viral genome maintenance via segregation and tethering of the viral genome to host chromatin (further discussed in section 1.3.2) [121]. E2 proteins are made up of three functional domains; a C-terminal DNA-binding domain, an N-terminal transactivation domain, and a central hinge region that provides a flexible link between these two terminals [121, 122]. These allow E2 to interact with and regulate multiple corresponding binding sites located in the long control region (LCR) of the viral genome, supporting regulation of the early viral promoter and viral origin of replication (see

Figure 1-5) [123-125].

E2 plays both activating and repressive roles in viral gene expression, in a differentiation-dependent manner according to E2 protein levels and its specific binding pattern to the HPV genome [126-129]. For HPV16 and HPV18, this most frequently involves repression of the early viral promoter, limiting expression of E6 and E7 [121, 130, 131]. In the 3' segment of the LCR, E2 has been shown to physically obstruct binding of transactivating proteins such as SP1 and the TATA-binding protein (TBP), which are required for transcription of E6 and E7 from the early viral promoter. Here, higher levels of E2 force displacement of SP1 and TBP from their respective LCR binding sites, instead favouring the E2-mediated recruitment of E1 and viral genome replication [132-137]. In addition, E2 inhibits early viral transcription by blocking access of a number of transcription factors such as NF1, AP-1, Oct, TEF1 and YY1 to the enhancer region of the LCR, which lies upstream of the early viral promoter (
Figure 1-5) [138-142]. Furthermore, E2 facilitates recruitment of cellular factors such as BRD4, EP400 and SMCX to epigenetically alter viral chromatin, which silences the early viral promoter region [143-145].

E8^E2 is another principal regulator of viral genome replication and transcription, which exclusively inhibits both early viral transcription and viral genome amplification throughout the viral life cycle [146, 147]. These spliced variants are derived from fusion transcripts containing a short E8 mRNA fragment, expressed from the E1 ORF, which is spliced together with the C-terminal of the E2 ORF [147-149]. As E8^E2 contains the C-terminal DNA binding domain and lacks the N-terminal transactivation domain of E2, this protein isoform

competitively binds to and efficiently blocks E2 binding sites within the viral genome. This inhibits E2-mediated recruitment and tethering of the viral DNA helicase, E1, to the viral origin of replication, thus repressing HPV DNA replication. Further, via its E8 segment, E8^E2 recruits host cellular corepressor complexes, NCoR/SMRT, to assist in its inhibition of viral genome replication and early promoter activity, resulting in even greater transcriptional and replicative repression than E2 [149-153].

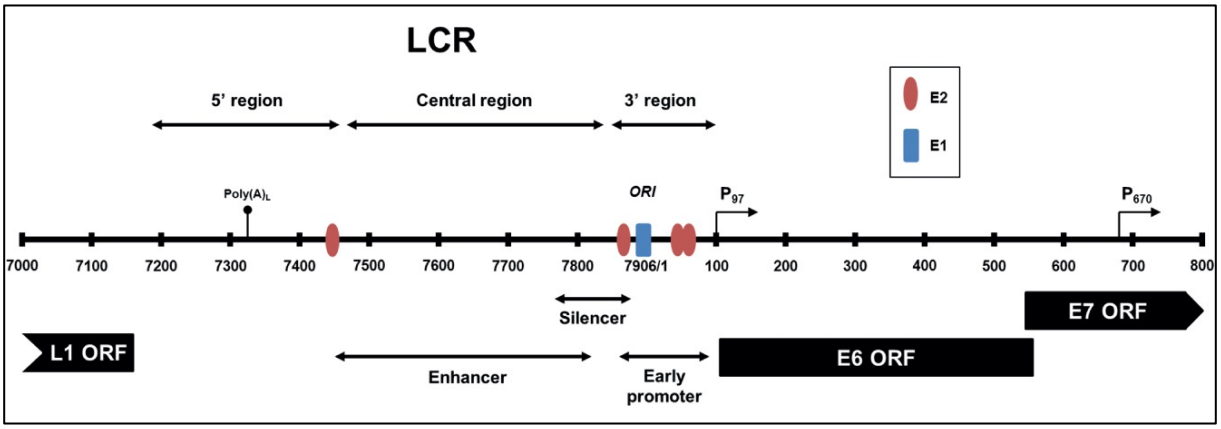


Figure 1-5: E2 and E1 binding sites within the HPV16 long control region (LCR)

Taken from Groves and Coleman [154]. The red ovals demonstrate positions of all four E2 binding sites within the LCR and the blue rectangle shows the E1 binding site in the 3' region and at the origin of replication (*ORI*). Here, E2 binds with E2 binding sites via its C-terminal DNA-binding domain, and to E1 via its N-terminal transactivation domain to facilitate binding of E1 to the *ORI* for viral genome replication. E2 binding within the 3' region also controls transcription of the early viral promoter (P_{97}), normally limiting early promoter activity and expression of E6/E7 [123-125]. The enhancer region of the LCR, which contains binding sites of transcription factors such as NF1, AP-1, Oct, TEF1 and YY1 is flanked by E2 binding sites, contributing to E2-mediated regulation of viral gene expression [138-142]. Finally, E8^E2 contains the C-terminal of E2 but not the N-terminal, allowing E8^E2 proteins to bind the same E2 binding sites within the LCR, where E8^E2 functionally blocks both the *ORI* and early viral promoter [149-153].

Epigenetic modulation of the viral genome is also important for viral genome maintenance and episome establishment, involving regulation of epigenetic processes such as histone acetylation, histone methylation, and CpG DNA methylation [155]. HPV-associated regulations of these epigenetic processes by E6, E7, E2, E8^E2 have been shown to regulate viral promoter usage and alternative splicing in a differentiation-dependent manner [95, 148, 156].

In the Parish lab, the chromatin regulator CCCTC-binding factor (CTCF), which alters chromosomal architecture and transcription via establishment of chromatin loops, has also been shown to be an important regulator of the HPV18 early promoter [157, 158]. Here, binding of CTCF to a conserved CTCF-binding site within the E2 ORF, was shown to physically interact with the transcriptional repressor protein, YY1, consequentially restricting expression of E6 and E7 by mediating formation of a transcriptionally repressive chromatin loop between the LCR and E2 ORF [159, 160]. As the HPV18-containing cells undergo terminal differentiation, YY1 production decreases, which alleviates this transcriptional block, upregulating E6 and E7 production to assist viral genome amplification [159]. In addition, alteration of CTCF has also been shown to modulate alternative splicing and late viral gene expression profiles in HPV18-HFK models, suggesting that this protein plays additional regulatory roles of the late viral promoter [98].

Another noteworthy regulatory factor is the epidermal growth factor, EGF, which has been shown to alternatively modulate E6 and E7 expression via activation of the ERK1-ERK2 signalling pathway [161]. This represses splicing between the HPV16 splice donor at nucleotide 226 (SD226) and the splice acceptor at nucleotide 409 (SA409), favouring

expression of full-length E6-encoding mRNA transcripts. Reduction of EGF however, increases SD226^SA409 splicing, instead favouring transcription of E6*I transcripts, which express E7 (see Figure 1-4, transcript species A, B, E, and I) [162-164].

1.3 The HPV life cycle in anogenital tissues

1.3.1 HPV infection and establishment in basal keratinocytes

The HPV life cycle is intrinsically linked to the differentiation status of the host keratinocyte. Successful infection relies on complex interplay between viral and cellular proteins, where HPV have evolved to take advantage of and manipulate the epithelial growth process of stratified squamous epithelia. Initial infection occurs in mitotically active basal cells, which line the bottom of the epidermis and are anchored to the basal lamina. In the cervix, HPV likely access the basal lamina via micro-wounds, which form during sexual or other forms of physical contact [165, 166].

The cervix contains a discrete population of specialised, cuboidal, non-stratifying, stem cells, known as cervical reserve cells, which are highly sensitive to both HPV infection and HPV-driven neoplasia and where the majority of cervical cancers arise [167, 168]. These are located within the cervical squamocolumnar junction (SCJ) of the metaplastic transformation zone, which lies between the squamous epithelium of the ectocervix and the columnar or glandular tissues of the endocervix [169]. While HR-HPVs are able to infect each of these cervical epithelia, HPV-driven carcinogenesis is primarily associated with cells, such as cervical reserve

cells that lack the ability to stratify. This means that unlike in stratifying epithelia such as in the ectocervix, HPV are unable to complete their productive viral life cycle, something that is thought to deregulate expression of E6 and E7 oncogenes, leading to neoplasia and cancer [84, 170, 171].

The icosahedral HPV capsid is made up of two proteins: the major capsid protein (L1) and the minor capsid protein (L2). At the basal lamina, L1 capsid proteins interact with and bind epithelial cell surface and extracellular matrix (ECM) macromolecules called heparin sulphate proteoglycans, as well as potential other host proteins such as laminins (Figure 1-6) [172-174]. This binding triggers a conformational change of both L1 and L2 proteins, mediated by cyclophilin B, leading to exposure of the L2 N-terminus [175]. This allows for interaction with secondary receptor complexes on the plasma membrane, which are required for internalisation [166, 173, 175, 176]. These include but are not limited to: α-integrins [177, 178], growth factor receptors such as EGFR [178], laminins [173], and the tetraspanin CD151 [179]. However, increasing evidence supports tetraspanin and tetraspanin enriched microdomains (TEMs) as particularly important platforms for viral uptake [179-181]. After internalisation, HPV16 is transferred to the endosome where acidification triggers viral uncoating of the capsid in a pH-dependent manner [182]. The remaining viral genome and L2 complex, now separated from L1, break free from the endosome, and are transported to the nucleus via the *trans*-golgi network [183, 184].

1.3.2 Initial replication and maintenance of the HPV genome in lower epithelial layers

In the nucleus, the circular extrachromosomal viral genome, or episome, is quickly replicated and then maintained at low copy numbers, normally between 50 to 200 copies per cell [176, 183, 185, 186]. In the first instance, viral replication is carried out independently of cellular replication. The early viral proteins, E1 and E2, are crucial for this initial replication phase and have been shown to be the first detectable viral mRNA species following HPV31 infection [187]. To initiate viral genome replication, E2 crucially binds and recruits the viral replication helicase, E1, along with other host DNA replication machinery to the viral origin of replication (see section 1.2.3, Figure 1-5), which contains three E2 binding sites and an A/T-rich region that contains several E1 bindings sites [188-191]. While E1 is able to bind to the origin of replication independently, its sequence specificity is greatly improved when complexed with E2 [192, 193]. E1 then melts and unwinds viral DNA at the origin of replication, allowing for access of host replication machinery and viral genome replication [194].

At these lower epithelial layers, viral transcription is initiated from the early promoter and ends at the early polyadenylation (pA^E) site, favouring expression of E6 and E7 over E1, E2, E1^E4 and E5, which are only weakly transcribed from the early promoter [195-198]. However, both viral replication and transcription are heavily restricted in undifferentiated cells, where viral episomes, mRNA, and proteins are maintained at low levels [176, 183, 185, 186, 199]. This is thought to be a crucial immune avoidance strategy, avoiding unwanted stimulation of local immune cells, which populate below the basement membrane. As discussed in section 1.2.3, E2 is a principal regulator of the early viral promoter in undifferentiated keratinocytes that limits expression of E6 and E7. In addition, viral replication and transcription are

negatively regulated by E8^E2 proteins, which competitively bind and transcriptionally block E2 binding sites at the origin of replication and early viral promoter. This is thought to be of great importance for the maintenance phase of persistent infection, where viral genome copy numbers are stably maintained at low levels in proliferating basal cells [200].

During the maintenance phase, in an alternate method of DNA replication, viral episomes are segregated by E2 and replicated in concordance with host DNA during cellular proliferation. Replicated HPV genomes are then divided equally between daughter cells [201]. This second DNA replication process is regulated by E2, which via its C-terminal and N-terminal domains, tethers itself and the viral genome to mitotically active chromatin, via attachment to chromatin-binding proteins [201-204]. Chromatin-binding targets include BRD4, MKlp2, ChlR1, and TopBP1 [185, 205, 206]. The cellular chromatin adaptor protein BRD4, has been identified as a major E2 target that plays multiple roles throughout the viral life cycle, including chromosomal E2 tethering [145, 207], initiation and regulation of viral transcription [208-210], and facilitation of viral genome amplification [211, 212]. During the maintenance phase of the viral life cycle, E2 has crucially been shown to bind with BRD4 at highly transcriptionally active host fragile sites, something that likely assists successful viral genome replication. However, these fragile sites are highly susceptible to DNA damage, which may induce integration of the viral genome into host chromatin [212].

1.3.3 Conservation of cell proliferation capability in differentiating keratinocytes

In normal epithelia, basal cells divide either symmetrically, to generate other basal cells, or asymmetrically, to generate keratinocytes that migrate towards the epithelial surface [213].

As daughter keratinocytes migrate upwards, the cells undergo terminal differentiation and exit the cell cycle, losing their replicative capabilities. At the surface, dead keratinocytes are sloughed off as part of the epithelial-renewal process [214]. However, given that vegetative viral genome amplification is dependent on cellular DNA replication machinery, HR-HPV have developed methods to recommence cell cycle re-entry in cells that would have normally exited the cell cycle. This is primarily mediated by the viral oncoproteins E6 and E7, which work together to re-initiate the cell cycle and host DNA synthesis in differentiating cells, effectively uncoupling cellular proliferation and differentiation programmes [16, 17].

E6 and E7 mRNA are transcribed from the early viral promoter and are increasingly expressed from lower to intermediate epithelial layers (Figure 1-6) [195, 196]. E6 and E7 oncoproteins, target a wide variety of cellular pathways to promote cell growth and proliferation, suppress anti-proliferative responses and alter the immune response. In normal dividing cells, two important tumour suppressor protein families play central roles in the regulation of cell growth and proliferation. These are the tumour suppressor protein p53, a DNA damage transcription factor that mediates cell cycle arrest and apoptosis, and retinoblastoma (Rb) proteins, which regulate cell cycle progression and differentiation, which are primary targets of E6 and E7 respectively. These mechanisms and the roles that they play in HPV-driven transformation are discussed in section 1.4.4.

Following mitosis, HPV E6 and E7 target p53, Rb and other cellular factors, further discussed in section 1.4.4, to potently induce unregulated cell cycle re-entry in parabasal and intermediate epithelial layers. This maintains production of the cellular replication machinery

that is required for viral genome amplification, resulting in extended expression of proliferation markers like PCNA, p16, MCM7, cyclin A, and cyclin E into upper layers and delayed expression of differentiation markers like involucrin, keratins (K)1/K10 and K4/K13 [88, 215-219]. In addition, E6- and E7-induced proliferation promotes expansion of basal and parabasal layers, resulting in increased thickness of tissues at sites of productive infection [199, 216]. Importantly, these pro-proliferative, anti-apoptotic, and DNA damage inducing oncoprotein activities are potent promoters of cellular transformation.

1.3.4 Vegetative viral genome amplification in G2 arrested, differentiating keratinocytes

As HPV-containing keratinocytes move to mid/upper-intermediate layers of the epithelium and following completion of cellular DNA replication in the S-phase, vegetative viral genome amplification is primarily carried out in the G2 phase, specifically in differentiating cells undergoing G2 arrest. Here, viral episomes are amplified to high copy numbers, ready for progeny virion formation [89, 220]. Differentiation-dependent activation of the viral late promoter, along with continued use of the pA^E site, upregulates transcription of viral genes that play varying roles in viral DNA amplification including E1, E2, E1^E4 and E5 (Figure 1-6) [197, 198]. While these viral genes are expressed from both early and late viral promoters, unlike the early viral promoter, E2 does not regulate late viral promoter usage, resulting in increased expression following differentiation [128].

HPV-mediated regulation of cyclin B1 is thought to be critical for G2 arrest. Cyclin B1 is an E2F responsive, mitosis promoting factor that is highly expressed during the G2/M phase of the cell cycle. In normal, mitotically active basal cells, cyclin B1 associates with CDK1 and is

translocated to the nucleus. Here, cyclin B1/CDK1 complexes play essential roles in cell cycle progression, pushing cells from the G2 phase into mitosis [221]. In HPV-containing anogenital epithelia however, studies have shown that cells undergoing HPV genome amplification express markers of G2 arrest, characterised by high levels of cytoplasmic cyclin B1 [89, 220]. Given that cyclin B1/CDK1 complexes are required in the nucleus to promote mitosis, accumulation of these proteins in the cytoplasm maintains keratinocytes in a G2-like state.

In normal keratinocytes, at the G2/M checkpoint, G2 arrest is initiated in response to DNA damage during cellular DNA replication. G2 arrest and DNA repair mechanisms are principally mediated by ataxia telangiectasia-mutated (ATM) and ATM and Rad3-related (ATR) DNA-damage response (DDR) pathways via regulation of cyclin B1/CDK1 and activation of DNA damage repair factors [222]. The ATM and ATR family members consist of serine/threonine protein kinases of the phosphatidylinositol 3-kinase-related kinase (PIKK) family, which work alongside DNA-dependent protein kinases (DNA-PK) to mediate repair of damaged DNA. Crucially, ATM and DNA-PK proteins respond to double strand DNA breaks and ATR proteins respond to single strand breaks at hindered replication forks. Both ATM and ATR pathways then promote phosphorylation of downstream effector proteins such as CHK2, H2AX, RAD51, BRCA1 and SMC1, as well as Fanconi Anaemia proteins [223-225].

To induce G2 arrest, HPV promote the retention of cyclin B1 and CDK1 in the cytoplasm. The exact mechanism by which HPV stimulates G2 arrest is not completely understood. However, given that G2/M arrest is ordinarily initiated by ATM and ATR, it is likely that HPV-induced activation of these pathways are a contributing factor [226]. It has been shown that E7 alone

is able to stimulate cytoplasmic accumulation of cyclin B1 and CDK1 in differentiating cells [89, 220, 227-229]. In addition, E1^E4 accumulates in the cytoplasm at extremely high levels in mid-epithelial layers, coinciding with viral genome amplification and is considered to play an important supporting role in this process by promoting a more replication-competent environment, which is discussed further in Chapter 3, section 3.5.4 [230-238].

During G2 arrest, HR-HPV E1, E6, and E7 have been shown to activate and utilise the ATM and ATR DNA damage response (DDR) pathways to assist productive viral genome amplification [227, 239-241]. Here, E1, E6 and E7 preferentially recruit and alter activities of downstream DDR factors including RAD51, BRCA1, H2AX, CHK2 and SMC1 away from host DNA to HPV replication sites [242-244]. These repair factors, in concert with E1, are then thought to replicate viral DNA in a 'recombination-dependent replicative' manner, greatly amplifying viral genome copy numbers [239-241]. This has shown to be crucial for both productive viral genome amplification and HPV life cycle success overall [227]. In addition, levels of DDR factors have been shown to increase during each stage of cervical cancer progression, suggesting that activation of the DDR is also important for transformation [245].

1.3.5 Viral capsid production and virion assembly

In the upper epithelial layers, and final stages of the viral life cycle, transcription of E6 and E7 is reduced following increased production of viral E2 proteins, which block transcription from the early viral promoter (Figure 1-6) [246]. Subsequent suppression of E6 and E7 causes keratinocytes to finally exit the cell cycle [90, 198, 247]. Supported by high levels of E2, a switch of transcriptional termination from the pA^E to the late polyadenylation (pA^L) site

initiates transcription of the late viral capsid proteins, L1 and L2 [94, 95]. In the nucleus, L1 and L2 proteins encapsulate viral genomes, forming infectious virions, which are sloughed off from the epithelial surface along with host cells. It is hypothesised that E1^E4 proteins assist these final processes. Here, E1^E4 exhibits an N-terminal leucine-rich motif that allows for direct interaction and modification of host keratin networks, an interaction that is considered to assist viral capsid formation and release from superficial epithelium [233, 236, 248-251]

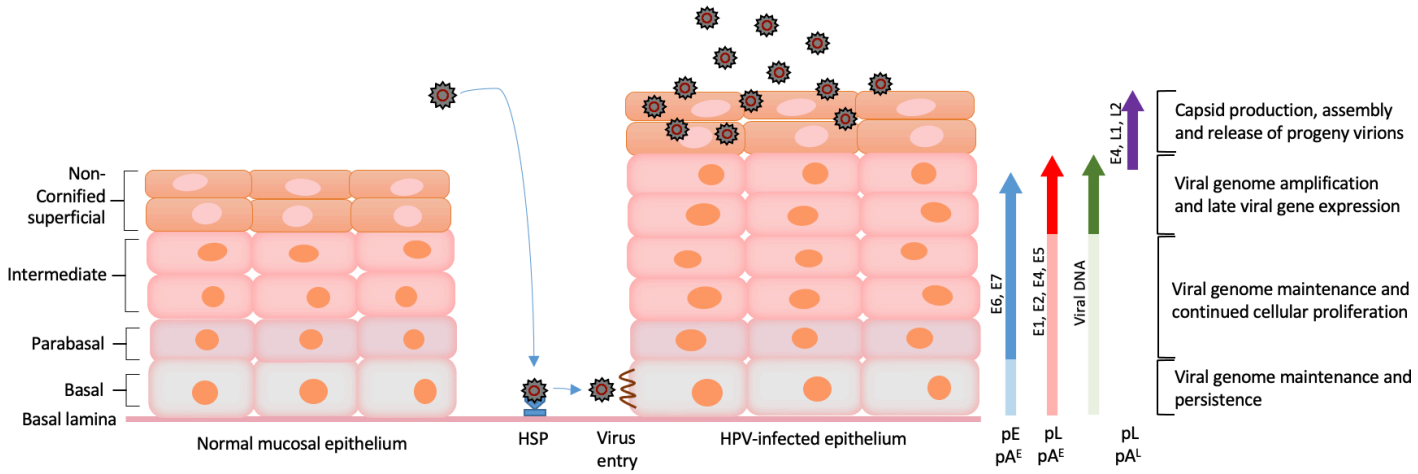


Figure 1-6: The HR-HPV life cycle

In stratified-squamous mucosal epithelium, HPV virions attach to heparin sulphate proteoglycan (HSP) on the basal lamina. HPV interact with secondary cell surface receptor complexes and are internalised into replication competent basal keratinocytes, where viral genomes are uncoated and translocated to the nucleus. Here, HPV genomes undergo an initial DNA amplification phase, establishing at 50 – 200 copies per cell. Transcription of early viral genes, initiated from the early viral promoter (pE) and terminating at the early polyadenylation (pA^E) site, assist initial episomal replication and maintenance, where E1, E2, E1^E4 (E4) and E5 are transcribed at low levels. In the lower epithelial layers, expression of E6 and E7 oncogenes are upregulated to promote cell-cycle re-entry and delay cellular differentiation. This also promotes expansion of basal and parabasal layers. In upper intermediate layers, differentiation-dependent activation of the late viral promoter (pL), along with continued use of the pA^E site, upregulates expression of E1, E2, E1^E4 and E5, to assist vegetative viral genome amplification in cells undergoing G2 arrest. A switch from the early to late polyadenylation (pA^L) site usage in the upper epithelial layers initiates expression of L1 and L2, which make up the viral capsid, along with continued expression of E1^E4. Virion assembly transpires in the upmost superficial layers, where new progeny virions, assisted by E1^E4, are released along with shedding epithelial cells.

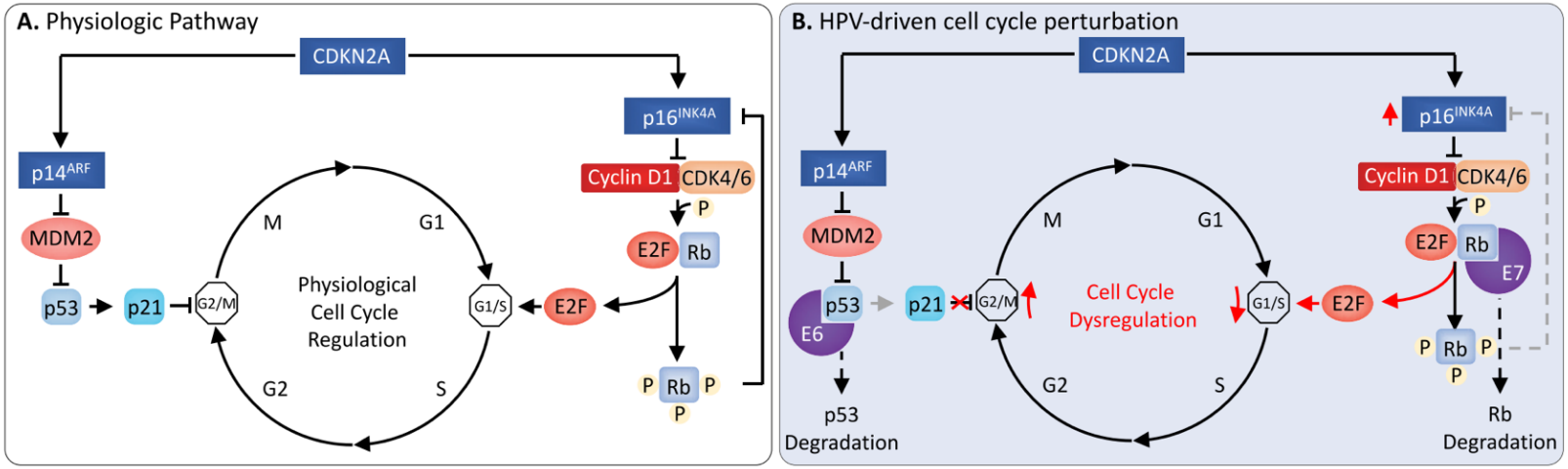


Figure 1-7: HPV cell cycle deregulation pathways.

Taken from Faraji, Zaidi [252]. In the normal physiological pathway of non-infected keratinocytes, the cyclin-dependent kinase Inhibitor 2A (CDKN2A) encodes two notable tumour suppressor proteins $p14^{ARF}$ and $p16^{INK4A}$. $p14^{ARF}$ is an inhibitor of a p53 regulator, otherwise known as MDM2, which maintains p53 at low concentrations. p53 are key DNA damage proteins, which act as stress induced transcription factors in response to replication errors and DNA damage, following DNA replication, during the G2/M transition phase. When activated, p53 proteins trigger p21, a cyclin-dependent kinase inhibitor (also known as CDKN1A) that inhibits several cyclin-CDK complexes to interrupt the cell cycle [253]. During HR-HPV-infection, p53 is degraded by E6 viral oncoproteins resulting in inhibition of p53, pushing the cell cycle into the mitotic phase [254, 255]. The other CDKN2A-encoded tumour suppressor protein, $p16^{INK4A}$, or p16 inhibits cyclin D1/CDK4 and cyclin D1/CDK6 complexes, which catalyse phosphorylation of Rb to release E2F. E2F transcription factors promote transcription of S-phase genes that stimulate cell cycle re-entry [256]. These are normally regulated by Rb tumour suppressor proteins. However, during HPV-infection Rb proteins are inactivated by E7 viral oncoproteins, relieving repression of E2F from Rb to stimulate cell cycle re-entry. This results in overexpression of $p16^{INK4A}$ via loss of negative feedback control [217].

1.4 HPV-associated head and neck cancers: characteristics and molecular mechanisms

Unlike in the cervix, the life cycle and natural history of HPV-mediated progression in the head and neck is not well characterised. As such, current understanding of HPV biology at this site is largely based on findings in anogenital models, lesions, and cancers. Interestingly, HPV-HNSCC display a number of distinct characteristics that diverge from both HPV-associated anogenital cancers and HPV-negative HNSCC. This section will further discuss current understanding of HPV-infection and cancer progression in the head and neck and will highlight important distinctions between these other comparable cancers.

1.4.1 Processes of HPV-driven transformation in the cervix and oropharynx

Progression of cervical cancers is well characterised due to a surplus of available cancer and pre-cancerous lesions at varying stages of disease progression. Pre-malignant cervical lesions are subsequently classed into one of three stages of cervical intraepithelial neoplasia (CIN), ranging from early dysplasia, CIN1, to pre-malignant lesions, CIN3. HPV infected tissues that are supporting HPV life cycles, still undergoing cellular differentiation, and producing low levels of E6 and E7 are generally described as CIN1, or low-grade squamous intraepithelial lesions (LSIL) [257, 258]. These are normally cleared by the immune system in under two years [24]. CIN3, or high-grade squamous intraepithelial lesions (HSIL), usually represent persistent infections in cells that are no longer able to differentiate or support viral life cycles. These lesions mostly comprise basaloid-like cells that extend from lower to upmost epithelial layers and have high expression of E6 and E7 [88, 215, 258, 259]. Around 40% of these pre-cancerous

lesions have been reported to develop into cancer [260]. In contrast, productive HPV-infections and lesions have rarely been identified in oropharyngeal regions, where HPV-OPSCC are almost always discovered at advanced disease stages [261, 262]. This has substantially limited study of HPV biology and early carcinogenic events at this body site [263].

Cancer progression of HPV-infected cells can develop in numerous ways. Events of central importance include deregulated and continuous expression of the viral E6 and E7 oncoproteins, host genomic instability, and viral persistence. Crucially, persistent infections have an increased chance to acquire epithelial cell abnormalities, which can lead to the development of cancer. Both viral persistence and tumourigenesis are principally aided by E6 and E7, which are able to promote every cancer hallmark within an infected cell (see section 1.4.4) [264]. This includes deregulation of cellular proliferation, induction of DNA damage, and inhibition of both anti-tumour and antiviral responses, where E6 and E7 further promote viral persistence via targeted impairment of cellular antiviral immune responses (see section 1.4.5). As such, an increase in E6 and E7 activities, along with a growing acquisition of genetic and epigenetic changes, drive malignant transformation to invasive carcinoma.

Deregulated expression of E6 and E7 can occur by multiple events including integration of viral DNA into host chromosomes (section 1.4.6) [265], or by epigenetic modification of the viral LCR via DNA methylation or chromatin modification, as reviewed by Groves and Coleman [154]. Another important factor is the anatomical site of infection. HPV-driven cancers consistently arise within cervical and oropharyngeal transformation zones such as the cervical squamocolumnar junction and the tonsil crypts. As further discussed in sections 1.3.1 and

1.4.2, these specialised, undifferentiating cells, with unique expression profiles are thought to be unable to support the normal replication cycle of HPV, resulting in deregulated expression of the early viral oncogenes [167, 266-269].

1.4.2 HPV infection and progression locality in the head and neck

HPV-HNSCC most commonly arise in the oropharynx, specifically within tonsil and base of tongue subsites [31, 32, 35]. A recent meta-analysis study conducted by Haeggblom, Ramqvist [35], revealed that 67% of assessed HPV-OPSCC originated in the tonsils, along with 25% in the base of tongue. Within the oropharynx, networks of invaginated tissue consisting of specialised reticulated epithelia known as crypts cover a significant area of the oropharyngeal surface, almost exclusively in the palatine tonsils and base of tongue (see Figure 1-8) [34, 56]. HPV-OPSCC have been shown to recurrently originate in the crypts. Interestingly, this is in direct contrast with HPV-negative OPSCC, which typically develop in normal stratified surface squamous epithelia [34, 35, 56]. The tonsil crypts must therefore represent a site that is more susceptible to HPV infection and/or HPV-mediated carcinogenic progression.

In the absence of micro-wounds that allow for HPV-entry to basal lamina during sexual intercourse, it is not clear what mode of entry HPV exploit in the oropharynx. However, it has been theorised that HPV infection may occur through reticulations in the epithelial barrier of the crypts (Figure 1-8) [34, 56]. These are characterised by a discontinuous basement membrane that allow immune cells from the lymphoidal tissues to access and process external antigens [270]. It is thought that HPV may take advantage of these reticulations to access and infect basal cells [31, 34, 271-273]. It is not completely understood why HPV-positive tonsil

cancers are more likely to develop in crypt epithelium, but viral access by reticulations, as well as alternative differentiation pathways and keratin profiles could be contributing factors.

Unlike the completely stratified surface tonsil epithelium, tonsil crypt epithelium does not undergo complete differentiation [266, 267]. Since the virus life cycle is dependent on the differentiation program of the epithelium, it is possible that like in the cervical reserve cells in the SCJ (section 1.3.1), a non-differentiating environment may disrupt HPVs normal replication cycle, leading to deregulated E6/E7 oncoprotein expression [269]. In addition, while K5/K14 are expressed in stratified surface tonsil epithelia, specifically in basal keratinocytes, tonsil crypt epithelia additionally express keratins characteristic of simple epithelia, K7, K8, K18 and K19 [274, 275]. Intriguingly, these keratins are also biomarkers of cervical squamocolumnar cells of the transformation zone, the primary site of HPV infection and cancer progression in the cervix (section 1.3.1) [274, 276, 277]. These keratins have also been indicated to play supportive roles in the HPV life cycle, as further discussed in chapter 3.5.7, and may represent cells that are sensitive to HPV-driven carcinogenesis [275, 278, 279]. Finally, unlike surface tonsil epithelia, basal keratinocytes of the tonsil crypts carry certain immune privileges that could further support viral persistence and carcinogenesis by limiting adaptive immune responses against infected cells and tumours (see section 1.4.5) [77, 280, 281].

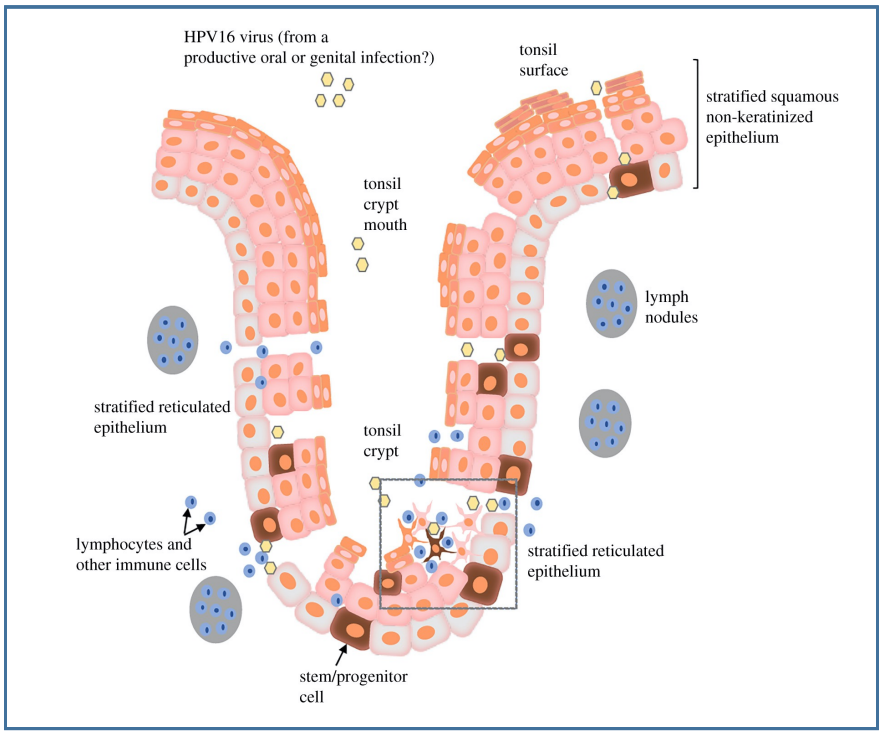


Figure 1-8: Schematic of tonsil epithelia

Comparing stratified squamous surface epithelia and specialised, often incompletely differentiated, reticulated crypt epithelia. Lymphocytes infiltrate reticulations to access oral antigens within the crypt. It is thought that HPV may access the basal lamina via these reticulations. This illustration was initially created by myself and adapted by Dr Roberts for publication [269].

1.4.3 HPV16 but not HPV18 drive the vast majority of HPV-OPSCC

HPV16 and HPV18 are the main carcinogenic drivers of HPV-associated cancers, which have been estimated to make up around 70 - 75% of all global HPV-associated cancer cases [5, 28, 282-284]. HPV16 and HPV18 have shown alternate preferentiality towards SCC and adenocarcinomas respectively. In a large survey entailing 1467 HPV-positive pre-neoplasia's of the cervix and 214 HPV-containing cervical cancers, HPV16 was reported in 62% of all cervical SCC cases and 35% adenocarcinomas. In contrast, HPV18 accounted for 66% of

adenocarcinomas and just 12% of SCCs. Put together, HPV16 makes up 57% of global cervical cancers and HPV18 accounts for 16% [282]. Interestingly, HPV16 and HPV18 were not nearly as proportionately dominant in cytologically normal cervical smears, constituting 30% and 10% of cases respectively [285]. Altogether this shows that HPV16 and HPV18 have much higher carcinogenic potential than other HR-HPV types in the cervix.

While HPV16 accounts for just over a third of oral/oropharyngeal HR-HPV infections [43], numerous studies have identified HPV16 as the most predominant HPV type in HPV-OPSCC, constituting upwards of 90% of cases [40, 286, 287]. Intriguingly, other HR-HPV types such as HPV18, which is commonly found in cervical cancer, are rarely detected in cancers at this body site (<2%) [33, 38-41, 282]. However, it is still unclear what advantages HPV16 has over other HR-HPV types to support more persistent infection or molecular progression in oropharyngeal subsites or how these differ to infection or pathogenesis in the cervix. Given that HPV-OPSCC frequently arise in lymphoepithelial tonsil subsites, it has been stated that HPV16 could elicit more pervasive anti-immune effects that may support longer lasting infections at this body site [1, 269]. There is currently limited understanding of HPV-mediated immune regulation in the oropharynx, as well as differences in immune regulation between HPV16 and other HR-HPV types like HPV18.

1.4.4 The HPV16 and HPV18 E6 and E7 oncoproteins and their role in carcinogenesis

The expression of HR-HPV E6 and E7 oncoproteins are required for initiation and maintenance of HPV-associated tumours. As further described in section 1.2.3, E6 and E7 are translated

from polycistronic mRNA transcripts that have been transcribed from the early viral promoter, which is regulated by host cellular factors such as AP-1 and SP1, as well as HPV E2 and E8^E2 proteins. E6 and E7 are highly multifunctional proteins that interact with a wide variety of cellular factors, importantly modulating all six of the major cancer hallmarks to promote viral persistence, cell survival and proliferation. This section will briefly overview some of the key physiological and carcinogenic activities of E6 and E7 and highlights some intriguing functional differences of the oncoproteins between HPV16 and HPV18 that may contribute to their distinct biological and pathological traits (see section 1.4.3).

HR-HPV E6 proteins are approximately 150 amino acids in length, and are comprised of two zinc finger, C-terminal and N-terminal binding domains, along with an alpha-helix connecting the two [288, 289]. The most notable interaction of E6 is the tumour suppressor protein p53, a key pro-apoptotic DNA damage regulator that acts as an essential safeguard of genomic integrity throughout the cellular replication cycle. During HPV infection, the viral oncoprotein E6 coordinates the degradation of p53 via ubiquitin-dependent proteolysis. This involves formation of a trimeric complex comprising E6, p53 and the cellular ubiquitination enzyme, E6-AP [254, 255]. Loss of p53 impairs cell cycle arrest and apoptotic functions, supporting unregulated cell cycle progression. However, this also instigates chromosomal instability, which over time substantially increases potential for carcinogenic mutations [290]. The ability of HR-HPV E6 to bind to and promote ubiquitination of p53 tumour suppressor proteins characterises one key difference between high-risk oncogenic HPVs and low risk HPVs [291]. Interestingly, HPV16 E6 has been shown to be able to target and inhibit p53 with greater efficiency than HPV18 E6 [292, 293].

A second mechanism of HPV-mediated inactivation of p53 involves the association of E6 with the acetyltransferases p300 and CREB binding protein (CBP), where HPV16 E6 was shown to inhibit the co-activating activities of p300 by binding to its C/H1 domain, the C/H3 domain and the C-terminus. Further, this interaction was shown to inhibit p300/CBP-mediated acetylation of p53, leading to impaired stability and activity of the tumour suppressor protein [290, 294-296]. In addition, E6-mediated impairment of p300/CBP also inhibited activation of nuclear factor kappa B (NF-kB), a central pro-inflammatory transcription factor that also plays important roles in cell survival and proliferation [295, 297].

Another important target of E6, as well as E7, is the multifunctional transcription factor and oncogene, c-myc. Here, HR-HPV E6 and E7 have been shown to both induce and interact with c-myc to deregulate cellular proliferation, apoptosis, immortalisation, and transformation [298-301]. To promote cellular longevity and immortalisation, HR-HPV E6 proteins, in cooperation with c-myc and E6AP, have also been shown to bind and induce transcription from the human telomerase reverse transcriptase (hTERT) promoter, driving functional activation of the telomerase complex [301-305]. Crucially, telomerase is required for synthesis of telomeric DNA sequences which lengthen the 3’ end of telomeres, suspending the process of senescence.

HR-HPV, but not low-risk HPV, E6 proteins also contain C-terminal PDZ binding motifs (PBM), which allow for interaction and E6-mediated proteasomal degradation of cellular proteins containing PDZ domains [306-310]. These PDZ proteins are tumour suppressors that play

numerous roles in cellular proliferation, polarity, and signalling [311, 312]. E6-mediated regulation of these proteins has been shown to be important for viral genome maintenance, the viral life cycle and carcinogenesis, relating to numerous mechanisms that have been reviewed by Ganti, Broniarczyk [311]. This notably includes a number of tumour suppressor proteins and members of the membrane-associated guanylate kinase (MAGUK) family, which play central roles in cellular polarity control pathways, including DLG1, Scribble (hScrib), and PAR3, along with tight junction proteins such as MAGI-I. These are all targeted by E6 for proteasomal degradation, leading to impaired cellular polarity and junctional stability [306, 307, 313-317]. Importantly, mutational inactivation of PBMs has shown to impair the transformative ability of E6 in both *in vitro* and *in vivo* models, highlighting this as an important feature of HR-HPV-induced carcinogenesis [318, 319].

Intriguingly, although all high-risk E6 proteins contain a canonical PBM, variable sequence diversity is observed between each HR-HPV type, resulting in distinct E6-PDZ affinity profiles. The biological relevance of this was first evidenced by Thomas, Massimi [316], who assessed and compared HPV16 and HPV18 E6 interactions with DLG1 and hScrib, showing that a single Leucine/Valine substitution in the E6 PBM greatly altered substrate specificity. Here, HPV16 E6 presented comparably higher binding specificity for hScrib but lower specificity for DLG1. A more recent study by Vincentelli, Luck [320], measured binding affinities of HPV16 and HPV18 E6 with a large panel of PDZ motifs, and showed distinct affinity patterns for cellular target proteins. It is likely that these differences in E6-PDZ binding profiles contribute to distinct biological and pathological traits of HPV16 and HPV18 in the tonsils. However, further study is required to better characterise these differences.

Interestingly, previous studies have suggested that HPV16 and HPV18 may differentially target specific interferon regulatory factors and pattern recognition receptors, which are important elements of the antiviral immune response [321, 322]. Here, HPV16 E6 was shown to bind and inhibit IRF3 [322], and only weak IRF3 regulatory functions were observed with HPV18 E6 [323]. In addition, HPV16 E6 and E7 have been shown to inhibit TLR9 expression in primary human keratinocytes, resulting in the functional loss of downstream TLR9 signal transduction [324]. However, Hasan, Bates [324] showed that HPV18 was not able to suppress TLR9 expression to the same effect as HPV16. This altogether suggests that HPV16 could elicit more stringent anti-immune activities than HPV18, possibly permitting better immune evasion and more persistent infections.

Like E6, HR-HPV E7 play numerous roles in malignant transformation. This largely involves induction and maintenance of uncontrolled proliferation but also includes deregulation of pathways relating to the immune response, chromosomal stability, and cell death signalling. At around 100 amino acids, HPV E7 are smaller than E6 and are made up of three conserved regions (CR)1/2/3 [325]. Importantly, HR-HPVs contain a conserved LXCXE motif within the CR2, or N-terminus, which have a unique affinity to strongly bind members of the Retinoblastoma protein (pRb) family of tumour suppressors, including Rb-p105, p107, and p130, which are all central regulators of cell cycle re-entry [326-328]. In contrast, this affinity is only weakly evoked by E7 of low risk-HPV types [329, 330].

Amongst their many functions, pRb importantly inhibit E2F transcription factor family members. E2F induce expression of certain S-phase genes that are required to push the cell cycle through the G1 phase and past the G1 restriction point, where the cell cycle becomes committed to replication without assistance from additional extracellular stimulants [256]. In differentiating cells, E7 promote this G1 to S-phase transition by binding Rb proteins via a conserved LXCXE motif located within the N-terminus of E7. This disrupts Rb-E2F binding, resulting in release of E2F proteins and subsequent activation of E2F-stimulated genes, leading to S-phase re-entry and cellular DNA synthesis (Figure 1-7) [218, 331, 332]. In addition, E7-mediated overexpression of the cyclin-dependent kinase inhibitor $p16^{INK4a}$ (p16), and E7-mediated release of p16 from pRb/E2F complexes, strongly inhibits cellular senescence and promotes cell survival [333].

In uninfected differentiating keratinocytes, cell cycle arrest occurs via differentiation-dependent upregulation of the cyclin-dependent kinase inhibitors, p21 and p27 [334]. These inhibit cyclin-dependent kinase 2 (CDK2), a central cell cycle progression catalyst that stimulates G1 to S-phase entry via association with both cyclin E and cyclin A [334, 335]. To suppress growth arrest, E7 targets and dysregulates several cellular proteins that induce cell cycle arrest following differentiation, including p21. In addition, E7 also stimulates the CDK2 activators cyclin A, cyclin E, and Cdc25a phosphatase to promote cell cycle progression [331, 336-339].

Protein-protein interaction studies carried out by White, Sowa [340], have highlighted some intriguing differential targets of between HPV16 and HPV18 E7. This notably includes a unique

interaction of HPV16 E7 with ZER1, a subunit of cullin 2 (CUL2), which is a core component of the cullin-RING E3 ubiquitin ligases. This E7-ZER1 interaction was shown to direct binding of E7 to CUL2 in order to stimulate its ubiquitination of the pRb tumour suppressor proteins. This is thought to further contribute to aberrant pRb degradation, further destabilising RB1-controlled cellular proliferation mechanisms [341]. Importantly, this E7-ZER1 interactions has recently been reported to contribute to the carcinogenic activities of HPV16 E7 by further promoting the growth of primary keratinocytes (Nouel and White, 2022, awaiting peer-review) [342]. If true, this could provide further evidence as to how HPV16 drives so many more cancers than HPV18 globally.

1.4.5 HR-HPV-mediated immune evasion and persistence

Contributing to their primary purpose as the first line of bodily defence, basal keratinocytes are widely regarded as important immune effector cells and nonprofessional antigen-presenting cells. During viral infection of a keratinocyte, innate cellular immune responses are activated following recognition of viral DNA/RNA by PRRs including TLR3, 7, 8 or 9, RIG-I, or cGAS. Sensitisation of these PRRs stimulates downstream activation of TNF-receptor-associated factor (TRAF) proteins, which potentiate nuclear translocation and activation of the interferon regulatory factors, (IRF)3 and IRF7, as well as the essential immune-stimulating transcription factor, NF-kB. These cooperatively induce transcription of type-I interferons (IFNs), most notably IFN-α and IFN-β, and other pro-inflammatory cytokines. Following activation, type I IFNs stimulate expression of IFN-stimulated genes (ISGs), which promote inflammation and activate downstream immune alarm functions and antiviral responses including by activation of JAK-STAT signalling pathways [343-350]. In addition, IFN-α and IFN-

β bind and activate IFN-α/β receptor (IFNAR) complexes to further augment type I IFN activation and IFN-stimulated gene expression in a positive feedback loop [351].

Activation of these immune pathways and production of ISGs also stimulate adaptive immune response mechanisms via production of immune stimulatory molecules such as cytokines, chemokines, toll-like receptors (TLRs), and MHC class I and II complexes. As such, keratinocytes are important mediators of anti-viral immunity that activate, recruit, and regulate a multitude of lymphoid- and myeloid-derived immune cells [352, 353]. This is of particular importance in the tonsils, which are highly immunocompetent lymphoid organs that mount initial immune responses against ingested or inhaled pathogens. Immune system priming is largely carried out in the tonsil crypts, which are highly saturated with B- and T-lymphocytes and other immune sentinel cells such as dendritic cells and macrophages [354, 355].

In order to persist in this hostile environment, HPV must therefore be able to effectively circumnavigate intracellular host-immune responses, as well as wider immune detection. An important but passive immune evasion strategy is in controlling release of immunogenic viral antigens to avoid detection by pattern recognition receptors (PRRs) and pathogen-associated molecular patterns or (PAMPs) in immunostimulatory basal cells, which activate immune signalling cascades and anti-viral responses [356]. As such, production of viral capsid proteins is exclusively carried out in superficial epithelial layers, where extracellular immune surveillance is lowest [357, 358].

HR-HPVs can also elicit more aggressive anti-immune strategies via activities of E6, E7 and E5, which mediate expression or functions of numerous immune stimulatory elements at various stages of immune signalling cascades. This results in suppression of central immune signalling pathways, often targeting and inhibiting activation of type I IFNs and NF-kB. This downregulates transcription of ISGs and NF-kB-inducible genes, which comprise essential components of the inflammatory response, as well as JAK-STAT and immune cell signalling pathways, significantly hampering antiviral responses [359-362]. The HPV oncoproteins have been shown to dysregulate these cellular immune responses by multiple mechanisms, which are briefly overviewed in Table 1-1.

In addition to the cellular immune response, HR-HPV are also able to modulate recruitment and activation of innate and adaptive immune cells, including CD4+ and CD8+ T-lymphocytes, B-lymphocytes, natural killer cells (NKT), Langerhans cells, dendritic cells, and macrophages [343, 363]. This is largely driven by E6-, E7-, and E5-mediated inhibition of immune signalling elements, such as interleukins and cytokines, or via repression of cell-surface immune cell associating elements, including major histocompatibility complexes (MHC), CD1D, PD-L1, and tumour necrosis factor receptors (TNFRs) [364-368]. This altogether limits recognition of infected keratinocytes by adaptive immune cells.

Manipulation of intracellular immune responses by HPV will likely support successful viral infection and persistence in the tonsils. In addition, tonsillar keratinocytes are immune privileged cells, which may favour more persistent HPV infections. Here, crypt epithelia secrete high amounts of PD-L1, a PD-1 antagonist that binds PD-1 receptors on local immune

cells in order to evade immune cell responses [74, 75]. In addition, HPV have been shown to overstimulate PD-L1 expression in OPSCC, which could further contribute to immune avoidance at this biological subsite [76, 77].

HR-HPV immune avoidance strategy	Mechanism	Host immune consequences
Epigenetic alteration of PAMP-sensing PRRs and cytokines	*HR-HPV E6 and E7 oncoproteins alter host epigenetic chromatin organization via recruitment of histone modifying and histone methylating enzymes to mediate host gene expression.*	
	E7-mediated recruitment of the methyltransferase, DNMT1 inhibits expression of the chemokine, *CXCL14* via hypermethylation of its upstream promoter [369].	Suppresses CXCL14-mediated recruitment of T-cells, DCs and NKT cells [370].
	HPV16 E7 recruits the histone modifying enzymes HDAC1 and JARID1B to alter a promoter region upstream of the PRR, TLR9 [371].	Inhibits viral sensing by TLR9, leading to downstream suppression of IRFs, type I IFNs, and NF-kB.
Mediation of PRR-inducible signal transduction pathways	*Virus sensing by PRRs such as TLR3, 7, 8 or 9, RIG-I, or cGAS, induces signal transduction pathways that activate innate immune responses. These notably activate type I IFNs and NF-kB, which induce transcription of innate and adaptive immune genes [343].*	
	E7 impairs the cytosolic cGAS-STING pathway by binding and inhibiting STING via its LXCXE motif [372].	Suppresses downstream activation of IRF3 and NF-kB [373].
	HR-HPV upregulate transcription of the ubiquitin ligase, *UCHL1*, an inhibitor of TRAF3 and TRAF6, which are required for activation of IRF3 and IRF7 [374, 375].	Suppresses downstream expression of type I IFNs, prevents phosphorylation of p65, and inhibits NF-kB [376].
	HR-HPV E6 complexes with and promotes activity of the ubiquitin ligase, E6-AP, initiating degradation of IL-1β [377].	IL-1β is a master regulator of the inflammatory response and canonical activator of NF-kB [378].
	HPV16 E7 directly binds and inhibits the interferon regulatory factors (IRF)1 and IRF9 [379, 380].	Inhibits transcription of IFN-β and production of ISGF3, which induces transcription of ISGs.
Suppression of pro-inflammatory transcriptional mediators	HR-HPV E6 and E7 further inhibit NF-kB via mediation of regulatory elements, including PCAF and IRFD1 [381, 382].	Downregulates transcription of NF-kB-inducible genes, hampering the innate immune response [362].
	HR-HPV inhibit IFNAR signalling pathways via inactivation of STAT1, STAT2 and IRF9, which induce transcription of ISGs [321, 383].	Suppresses downstream activation of type I IFNs and ISGs, impeding innate and adaptive responses [362].
Mediation of antigen presentation pathways	*Keratinocytes express cell surface MHC class I complexes, which display antigen peptide fragments to cytotoxic T-cells. Required for activation of the T-cell response [384]. Transcription of MHC components are largely mediated by IFNs and NF-kB [385, 386].*	
	HR-HPV inhibit transcription of MHC-I elements via dysregulation of IFN-α, IFN-β, and NF-kB [387], and by E7-mediated epigenetic modification [364].	Limits recognition of infected keratinocytes by T-cells, protecting against T-cell destruction and further T-cell-mediated immune stimulation.
	HR-HPV E5 and E7 inhibit trafficking of MHC-I to the cell surface [388-390].	

Table 1-1: Overview of notable HR-HPV immune avoidance strategies and mechanisms

1.4.6 The oncogenic consequences of HPV genome integration

Insertion of HR-HPV genomes into host chromatin, or integration, has been widely linked to selective growth advantages and promotion of genomic instability, usually due to dysregulated expression of E6 and E7 oncogenes [391, 392]. This dysregulation is thought to occur by at least one of a number of described mechanisms, which have been overviewed in Figure 1-9. Notably, when linearized HPV genomes, broken at E1 or E2 regions undergo recombination to free ends of human DNA at fragile sites, this results in loss of the viral transcriptional regulators E2 and E8^E2, which promotes upregulated expression of E6 and E7 oncogenes [393, 394].

During persistent HPV infection, viral integration is a critical oncogenic event in the development of a substantial proportion of anogenital and oropharyngeal cancers. Viral integration is an unintentional event that marks the end of the HPV life cycle. After which, production of viral episomes is no longer possible due to breakage of the viral genome during its insertion into host DNA thus ending productive viral infection in the host.

Chromosomal integration of HPV genomes are extremely common events in HPV cervical cancers, accounting for >80% of cases [395]. In HPV-positive OPSCCs however, integration is a far rarer event comprising around 39 - 43% of cases, with remaining HPV-OPSCC cases presenting mostly episomal or mixed episomal and integrated status (cells containing both HPV episomes and integrated HPV DNA) [396, 397]. While HPV genome integration in cervical cancers is generally associated with worse clinical outcome than tumours maintaining viral

episomes, the situation in OPSCC is less clear, with conflicting reports indicating either worse or unchanged clinical outcomes [277, 397, 398].

Intriguingly, studies have demonstrated un-enhanced expression of E6 and E7 in a number of integrated HPV-HNSCC biopsies [396, 398-400]. This was also shown in primary HTK immortalised with HPV16 genomes [401]. This altogether suggests that constitutive but not necessarily high levels of viral oncoproteins are required for HPV-driven carcinogenesis at this body site, which may suggest alternative cancer-driving events. It has been proposed instead that HPV integration may drive oncogenesis via alterations of the host genome or chromatin regulatory elements at/or bordering the site of incision, which are known as integration breakpoints [399, 400], an event that can result in inhibition of tumours suppressor genes or enhancement of oncogenes. This seems likely given that HPV-integrated OPSCC frequently present distinct DNA methylation arrangements and gene expression profiles from episomal HPV-OPSCC [400]. In HPV-OPSCC, HPV are commonly found to integrate at 'genomic hotspots', which consist of transcriptionally active regions, as well as regions containing fragile sites [402-404]. Interestingly, genomic analysis of HPV-HNSCC have revealed a number of key cancer-related genes that were recurrently deregulated by HPV integration. This included *RAD51B, ETS2, PDL1, TP63,* and *NR4A2* [399, 400, 405].

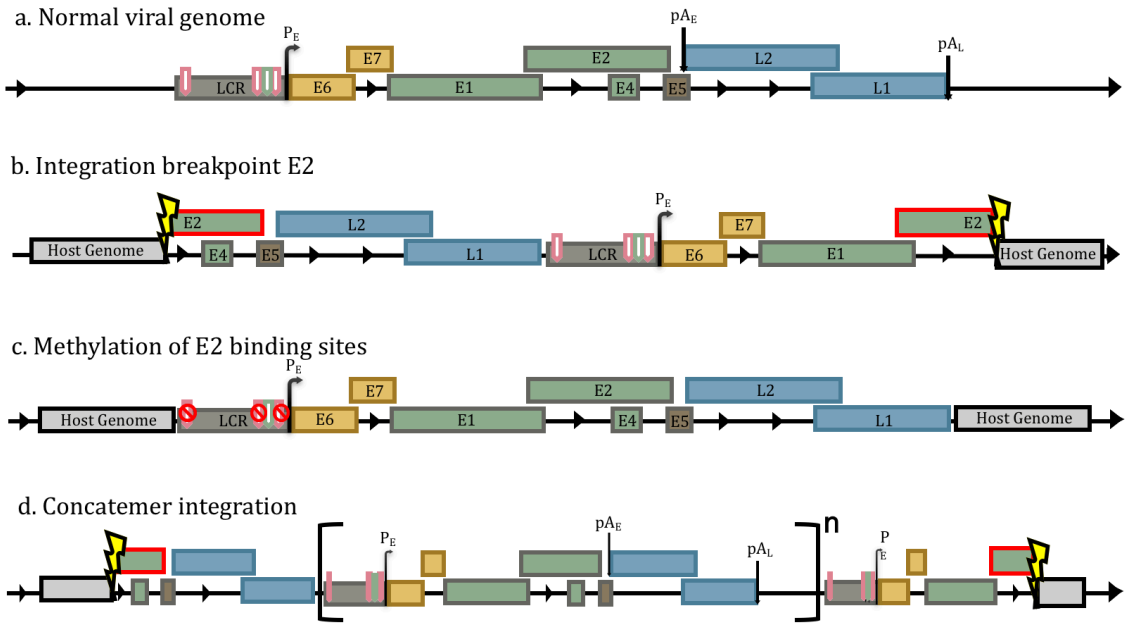


Figure 1-9: Models of oncogenic integration events that dysregulate E6 and E7 expression

a. shows a normal linearized HPV genome. P_E (early promoter), pA_E and pA_L (early and late polyadenylation sites). The pink blocks located in the LCR represent E2 binding sites and the green block is an E1 binding site. **b.** Breakage of the E1 or E2 ORF during integration disrupts expression of E2, a transcriptional repressor of the pE, facilitating unregulated transcriptional activation of E6 and E7 expression [393, 394]. **c.** E2 binding sites of the URR may become methylated during integration, inhibiting E2 protein binding of the pE and resulting in upregulated expression of E6 and E7 [406]. **d.** The HPV genome can either integrate into the host genome as one single transcript or as concatemers. These are integrated head-to-tail repeats of the whole HPV genome, which may serve as 'transcriptional super-enhancers' that promote oncogenesis by expressing a number of E6/E7 repeats [399]. In addition, but not shown above, epigenetic events occurring near the site of integration can stimulate E6 and E7 expression without directly affecting E2 binding to URR DNA sites [407]. Chromatin regulatory elements near to the site of integration may also upregulate expression of E6 and E7 [408]. This figure has been adapted from McBride and Warburton [409].

1.4.7 The distinct molecular characteristics of HPV-HNSCC

Distinct molecular and epigenetic differences have been described between HPV-HNSCC and HPV-negative HNSCC tumours, signified by alterations in transcription profiles and mutational spectrums. It is important to note that HPV-positive HNSCC do not normally present nearly as many genomic mutations as HPV-negative HNSCC. This is largely due to the fact that HR-HPV E6 and E7 oncoproteins elicit direct oncogenic activities on host cells, largely via deregulation of tumour suppressor proteins such as p53 and Rb. In HPV-negative HNSCC however, deregulation of tumour suppressing functions is more reliant on a multitude of genetic mutations [410]. As such, alternate activity of pathways involving cell cycle regulation, DNA-damage repair, and mitogenic signalling are observed between HPV-positive and HPV-negative tumours, which have been recently reviewed by Faraji, Zaidi [252] and Pan, Issaeva [411]. These differences in mutational status appear to be one likely explanation as to why HPV-HNSCCs are generally more responsive to radiotherapy and have more favourable clinical outcomes than HPV-negative tumours (section 1.1.5) [412, 413].

Genome wide analysis using data derived from Cancer Genome Atlas [410] has built up an excellent database of somatic mutations and molecular markers of HNSCC, revealing distinct molecular discrepancies between HPV-positive and -negative HNSCCs. Here, genome wide analysis of 279 HNSCCs took place, in which 36 were HPV positive. Importantly, 56% of these HPV-positive cancers were shown to have activating mutations of *PIK3CA*, encoding the catalytic subunit phosphoinositide 3-kinase (PI3K), which is involved in the AKT/PI3K/AKT/mTOR pathway. This is an essential pathway for protein synthesis, cell growth, proliferation, and survival, in which overactivation can be a significant driving factor of

carcinogenesis [414]. *PIK3CA* is the most common and sometimes only mutated oncogene observed in HPV-OPSCCs, indicating an important role in OPSCC development [415, 416]. Other notable, albeit less frequent HPV-OPSCC mutations include *ZNF750, KMT2D, EP300, FGFR3, TRAF3, PTEN, CYLD,* and *DDX3X*, many of which only rarely mutate in HPV-negative HNSCC [410, 417].

Many HPV-specific mutations exhibit distinct cytosine to thymidine mutations at TpC sites, which are thought to be driven via dysregulation of the cytidine deaminase, APOBEC family members - something that has been observed in both HPV-HNSCC and cervical cancers [395, 410, 418]. During an innate immune response, APOBEC elicit potent antiviral effects by mutating viral genomes, impairing viral replication. However, overexpression has now also been shown to mutate host DNA in HPV-associated cancers, signified by APOBEC-specific mutational signatures such as the mutational activation of *PIK3CA* [395, 410, 418, 419].

Global expression studies of HPV-HNSCC revealed distinct molecular signatures between HPV-HNSCC and cervical cancers, indicating alternative natural histories in carcinogenesis between the two body sites [420]. While some mutational similarities are noted between HPV-HNSCC and cervical cancers, including mutations of *PIK3CA*, *EP300*, *PTEN*, *FBXW7*, and *TGFBR2*, other mutations have been identified that were more exclusive to HPV-HNSCC, including *FGFR3*, *CASZ1*, *TRAF3, CYLD*, *ZNF750*, and *RIPK4* [417]. Mutational inactivation of *TRAF3* or *CYLD* could be of particular importance in progression of a subset of HPV-OPSCCs. These are negative regulators of NF-kB transcription factors, which positively regulate proliferation and cell-mediated immune responses, and negatively regulate apoptosis. Disruptive mutations of

TRAF3 or *CYLD* occurs in up to 36% of HPV-OPSCC cases and is much less frequent in cervical cancers <5% [410, 417, 421]. Intriguingly, this occurs in both integrated and non-integrated HPV-OPSCC and may play important roles in survival.

1.5 Project hypothesis and aims

As discussed throughout this introduction, research of HPV biology and carcinogenesis in the oropharynx has been greatly limited by a shortage of available HPV-positive pre-malignant biopsies and tissues. As such, much of our current understanding of HPV infections and the natural history of HPV-driven carcinogenesis in the oropharynx primarily stems from study of cervical lesions, cancers, and primary anogenital infection models. This encompasses fundamental areas such as the HPV life cycle, viral transcriptome organisation, viral integration events, regulation of host immune responses, and early carcinogenic events. Yet, a number of important distinctions have been described between HPV-OPSCC and HPV-cervical carcinomas (CC) that may signify different biological activities of HPV and transformative events between the two anatomical sites. These include distinct mutational profiles (section 1.4.7), differences in detectable HPV infections and disease presentation (section 1.4.3), and an alternate predisposition for integration of the viral genome (section 1.4.6). In addition, contrasting virus-type distributions are observed between HPV-OPSCC and cervical cancers, which show HPV16 as a much more dominant driver of oropharyngeal carcinogenesis in comparison to other HR-HPV types such as HPV18 (section 1.4.3).

Overall, these observations could reflect phenotypic or molecular distinctions between tonsillar and cervical subsites that may constitute an alternate permissibility for HPV infection

or susceptibility for HPV-driven carcinogenesis, which in specific tonsil subsites, favours HPV16 over other common HR-HPV types such as HPV18. These may also emphasise alternative biological or carcinogenic behaviours between HPV16 and HPV18 at this body site, as is hypothesised below. These site-specific differences in carcinogenic potential between HPV16 and HPV18 in the tonsils are intriguing and present a unique method for investigation of HPV16 persistence and carcinogenesis, forming the key aims of this project.

Hypothesis: HPV16 and HPV18 differentially manipulate host tonsil keratinocytes leading to differences in viral persistence and disease progression.

Key Aims:

1. Assist in the establishment of a complete set of isogenic HPV16- and HPV18-immortalised primary human tonsil keratinocyte (HTK) models (Chapter 3).
2. Characterise and compare HPV16 and HPV18 life cycle organisation in the primary HTK models (Chapter 3).
3. Characterise and compare HPV16 and HPV18 transcriptome organisation in the primary HTK models (Chapter 4).
4. Characterise and compare HPV16- and HPV18-induced transcriptional reprogramming of the primary HTK models (Chapter 5).

Chapter 2: Materials and methods

2.1 Cell culture

Cell culture was conducted in sterile conditions using a HEPA-filtered laminar cabinet (Thermo Fisher Scientific, UK). Cells were grown in appropriate growth media, components of which were filter sterilized using 0.22µm hydrophilic polyvinylidene fluoride (PVDF) membrane filters (Starlab™). Cells were maintained in incubators at 37°C with 5% CO_2. For general cell maintenance, all media was warmed at 37°C prior to use. Cells were harvested using trypsin-containing solution and pelleted by gentle centrifugation at 200 x g for 5 minutes. To determine cell concentrations, cells were counted using Kova fast-read® cell counting chambers (Biosigma). Human keratinocytes were cultured on Falcon® 100mm cell culture dishes (Corning) and all other cells including NIH-3T3 J2 fibroblasts on 100mm cell culture dishes (Corning). Routine testing of cell lines for mycoplasma infection was carried out according to manufacturer's instructions using the MycoAlert mycoplasma detection kit (Lonza, UK).

2.1.1 Culture and irradiation of NIH-3T3 J2 fibroblasts

NIH-3T3 J2 mouse fibroblasts (J2-fibroblasts) were cultured in Dulbecco's modified Eagle's medium (DMEM) with HEPES modification (Sigma-Aldrich, Gillingham, UK), supplemented with adult bovine serum (10% vol/vol) (Gibco, Loughborough, UK) and 4mM L-glutamine. J2-fibroblasts were grown up to 90% confluency before passage. To passage, cells were washed

with PBS, incubated for 5 minutes in 1ml TrypLE™ (Life Technologies, UK) until detached and harvested in J2 media to inactivate TrypLE™. Cells were centrifuged and resuspended in J2 media. Cells were typically diluted 1:20 between each passage and reseeded onto Corning 100mm tissue culture dishes. J2 fibroblasts were cultured up to passage 20, at which point low passage J2 fibroblasts would be thawed from cryogenic storage.

Harvested J2-fibroblasts were re-suspended in complete E-media at a density of $2x10^6$ cells per ml. The cells were irradiated with 30 Grays of either a caesium-137 source or with X-ray radiation using a Cell Rad Faxitron and were stored at 4°C for up to one week.

2.1.2 Primary human keratinocytes

2.1.2.1 Keratinocyte cell culture

Primary human tonsil keratinocytes (HTKs) donor #2 and #3 (HTK#2 and HTK#3) and human foreskin keratinocytes (HFKs) donor #13 (HFK#13) were isolated from patients via tonsillectomy and from neonate foreskin respectively by Dr Sally Roberts (Wilson, Ryan [422]; REC approval number 06/Q1702/45).

Primary keratinocytes were initially cultured in serum-free keratinocyte media (SFM, Thermo Fisher Scientific, UK). SFM was changed every two days until HTKs reached 80% confluency. Cells were washed twice with 10ml sterile PBS and trypsinized in 2ml TrypLE™ until detached. To inactivate the trypsin, cells were resuspended in pre-warmed E-media supplemented with 1µg/ml epidermal growth factor (EGF, BD Biosciences), referred to as complete E-media. This was made up in 2L batches and consisted of 1200ml DMEM without calcium (Sigma-Aldrich

D1152), 640ml Ham's F12 nutrient media (Thermo Fisher Scientific, UK), 100ml Hyclone™ foetal calf serum (FCS), 40ml penicillin and streptomycin (Sigma-Aldrich), 2ml hydrocortisone (5mg/ml, Sigma-Aldrich), 2ml cholera toxin A (10mg/ml, Corning), and 20ml 20x cocktail. The cocktail was made up of 20ml insulin (5mg/ml, 0.1M HCL, Sigma-Aldrich), 20ml transferrin (5mg/ml, Sigma-Aldrich), 20ml tri-iodo–lL-thyronine ($2x10^{-8}$M, Sigma-Aldrich), 20ml adenine (0.18M, Sigma-Aldrich) and 120ml PBS. Cells were counted, pelleted, and resuspended in complete E-media. HTKs were re-seeded onto fresh plates containing $2x10^6$ irradiated J2 fibroblasts and 10ml complete E-media, typically at $2x10^5$ cells per plate. Excess cells were frozen at -80°C, either suspended in freezing media for long term storage or as cell pellets for later experimentation (see 2.1.3).

2.1.2.2 Generation and maintenance of HPV-positive keratinocytes

HPV-positive (HPV+) keratinocyte lines were previously generated by members of the Roberts group as described by Wilson, Ryan [422]. In brief, low passage primary human keratinocytes were cultured up to around 50% confluency on 60mm tissue culture dishes in SFM. HTKs were co-transfected with 1µg recircularized HPV16 or HPV18 genomes and 1µg neomycin resistant plasmid (pSV_2Neo) in SFM containing the transfection reagent FuGENE™ (Promega Corporation). The FuGENE/HPV DNA mixture was incubated at room temperature and added to the primary cells. The following day, transfected cells were passaged onto cell culture plates containing irradiated J2 fibroblasts and complete E-media. G418 selection took around 5 days, after which HPV+ keratinocytes were expanded in complete E-media on a feeder layer of irradiated J2-fibroblasts.

To passage HPV+ keratinocytes, irradiated J2-fibroblasts were first plated onto cell culture dishes at $2x10^6$ cells per dish, 4 to 24 hours prior to seeding the keratinocytes. HPV+ keratinocytes were grown up to around 90% confluency before further passage. Irradiated J2-fibroblasts were removed by washing with media followed by a further 3 washes with PBS. Keratinocytes were incubated in 2ml TrypLE™ for 10 minutes until detached. An equal volume of complete E-media was added to inactivate trypsin, cells were counted, and $2x10^5$ cells were added to cell culture dishes containing 10ml complete E-media and evenly spread around the dish. E-media was changed every two days and the irradiated J2 fibroblasts were changed every 5-7 days.

2.1.3 Storage of cells

2.1.3.1 Cryogenic storage of cells

Keratinocytes were harvested for liquid nitrogen storage at each passage and J2-fibroblasts were cryogenically frozen at early passage. Harvested cells were resuspended gently in freezing medium at a density of $2x10^6$ cells per cryotube and slowly cooled at -80°C in Mr. Frosty™ freezing containers (Thermo Fisher Scientific, UK) containing isopropanol. The vials were then placed on dry ice and transferred to liquid nitrogen. Keratinocyte freezing media consisted of sterile complete E-media, supplemented with 10% (vol/vol) Hyclone™ FCS and 20% (vol/vol) glycerol. J2-fibroblast freezing media comprised J2 media supplemented with 10% (vol/vol) DMSO (Sigma-Aldrich) and 10% (vol/vol) adult bovine serum (Thermo Fischer Scientific, UK).

2.1.3.2 Storage of cell pellets

During each passage, keratinocytes were pelleted for DNA, RNA, or protein studies. For pellets, typically $5x10^6$ to 10^7 keratinocytes were centrifuged. Cell pellets were resuspended in 10ml sterile PBS, transferred to 15ml conical centrifuge tubes and re-centrifuged. Residual pellets were resuspended in 1ml PBS and transferred to 1.5ml microfuge tubes and cells pelleted by centrifugation. The supernatant was removed, and cell pellets were stored at -80°C.

2.1.4 Organotypic raft culture

2.1.4.1 Steel grid preparation

Meshed stainless steel grids were cut into small, approximately 4 x 4cm, rectangular grids. Edges were equally bent on two opposite ends to provide raised support for the rafts. A spirit level was used to establish a flat horizontal plane. Grids were treated in a fume hood with concentrated sulphuric acid for one hour and rinsed overnight with running tap water. Washed grids were autoclaved for sterilisation.

2.1.4.2 Collagen plug preparation

Prior to use, all reagents were chilled on ice and pipettes cooled at -20°C. Low passage J2-fibroblasts were harvested, pelleted, and cooled on ice ($2x10^6$ per collagen plug). Cell pellets were resuspended in 0.3ml filtered reconstitution buffer (RCB; 260mM $NaHCO_3$, 200mM HEPES, 0.05M NaOH) and 0.3ml filtered DMEM (10x) per collagen plug. Type I rat-tail collagen (Corning; product number 354249) was used at a final concentration of 4mg/ml, diluted in

0.02M acetic acid to a final volume of 2.4ml per plug. This was combined and mixed with the J2 fibroblast suspension. To bring the suspension to the correct pH, filtered 1M NaOH was added drop by drop until the mix turned an orange-red colour. Finally, 3ml of the collagen mix was added to 3cm tissue culture dishes and placed in an incubator at 37°C with 5% CO_2 for 30 minutes to set, after which 2ml complete E-media was added to the surface of each. Collagen plugs were maintained for up to 3-days and complete E-media was replaced each day until used.

2.1.4.3 Raft culture

Primary human keratinocytes were placed on top of the plugs and allowed to expand. To seed, media was removed from each plug and replaced with $2x10^6$ keratinocytes suspended in 2ml complete E-media. Complete E-media was replaced each day until the media had turned yellow over a single 24-hour period, or up to 4 days. Once ready, rafts were carefully lifted, using sterile spatulas, onto steel grids centrally positioned in 10cm tissue culture dishes. E-media (without EGF) was transferred to each dish so as to form a liquid-air interface (see Figure 2-1). Rafts were incubated at 37°C with 5% CO_2 for 13 days. E-media was changed every other day.

2.1.4.4 Raft fixation, paraffin-embedding, and haematoxylin and eosin (H&E) staining

Six hours prior to fixation, raft culture medium was taken from raft-containing plates, supplemented with 20µM 5-Bromo-2'-deoxyuridine (BrdU, BD Biosciences) and replaced. BrdU is a thymidine analogue that can substitute the nucleoside during DNA replication,

allowing for visualization of proliferating cells. After 6 hours, media was removed, and the rafts were flooded with 10% vol/vol formaldehyde in DMEM. Paraffin-embedding and sectioning was carried out by Propath Ltd., (Hereford, UK), along with H&E staining. Four-micron unstained sections were prepared from the paraffin blocks on poly-L-lysine coated slides by the University of Birmingham Human Biomaterials Resource Centre (HBRC).

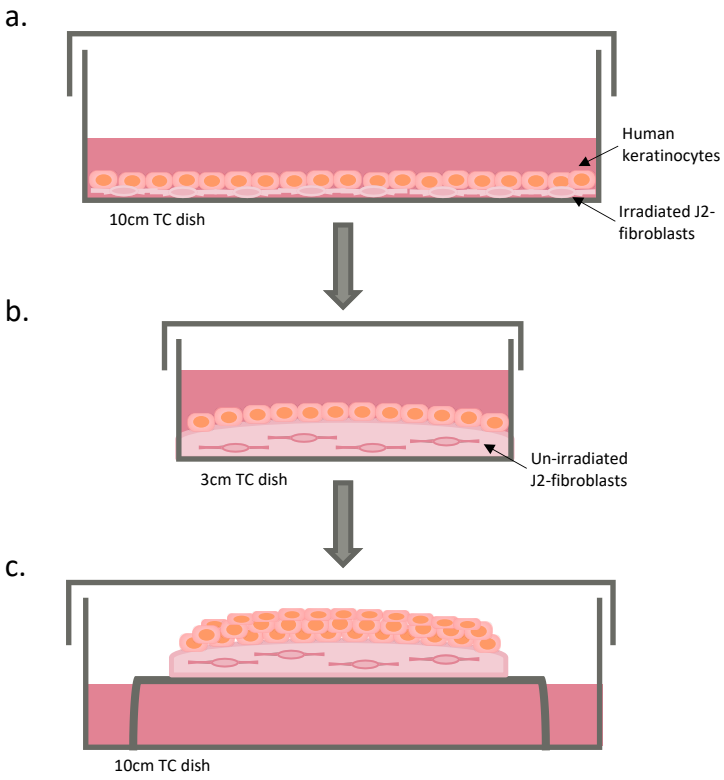


Figure 2-1: Organotypic raft culture diagram

1. Primary human keratinocytes were cultured in monolayer on irradiated J2-fibroblasts.
2. Collagen plugs embedded with J2-fibroblasts were prepared in small 3cm tissue culture dishes. Keratinocytes were seeded onto collagen plug surfaces and expanded in E-media supplemented with EGF.
3. After expansion, plugs were lifted onto steel grids in a 10cm tissue culture dish. E-media without EGF was added underneath rafts to form a liquid-air interface and stimulate stratification of basal keratinocytes.

2.1.4.5 Raft immunofluorescence staining and imaging

Immunofluorescence staining was carried out on paraffin-embedded raft sections prepared by HBRC. Agitated low temperature epitope retrieval (ALTER, Watson, Rollason [423]) was performed to expose antigens in the formalin-fixed tissue. To deparaffinize and rehydrate, the sections were immersed in Histo-Clear (National Diagnostics, UK) for 10 minutes and then into 100% industrial methylated spirits (IMS) for 5 minutes. Excess IMS was removed, and slides were washed several times in water. For ALTER, slides were placed in a large beaker with a magnetic stirrer containing 1 litre of EDTA/TWEEN buffer (1mM EDTA (pH 8.0), 0.1% vol/vol Tween®20), pre-heated to 65°C. Slides were agitated in the solution overnight at 65°C. Slides were rinsed in water and blocked in heat inactivated goat serum (HINGS) blocking buffer (20% vol/vol HINGS, 0.1% wt/vol bovine serum albumin (BSA) in PBS) for 2 hours at room temperature. All antibody incubations were carried out in humidity chambers. Primary antibodies (see Table 2-3) were diluted in HINGS blocking buffer and were added to the surface of the section. These were incubated overnight at 4°C and washed 2-3 times in PBS. Secondary antibodies; Alexa Fluor 488 goat anti-mouse and Alexa Fluor 594 goat anti-rabbit (Invitrogen; Table 2-4) were diluted 1:1000 in HINGS blocking buffer and were added to each section. These were incubated for 1 hour at 37°C and washed 4 times in PBS. On the final wash, 4',6-diamidino-2-phenylindole (DAPI; Sigma-Aldrich) was added to the PBS as a nuclear stain at around 0.1µg/ml and slides were washed for 10 minutes. Finally, a few drops of Fluoroshield™ mounting media (Sigma-Aldrich) was added to each section and cover slips were placed on top.

Immunofluorescence and H&E staining were visualized using a Nikon E600 epifluorescence microscope. Images of the stained raft sections were acquired using a Leica DC200 camera and processed using ImageJ.

2.2 DNA analysis

2.2.1 Isolation and purification of HPV DNA from keratinocytes

DNA was extracted via phenol-chloroform extraction. Primary human keratinocytes were washed in PBS and resuspended in DNA lysis buffer (400mM NaCl, 10mM Tris-HCl, 10mM EDTA, pH 7.4) and RNase A, which was added to a final concentration of 400ng/ml. The cell suspension was mixed and incubated at 37°C for 30 minutes. Proteinase K was added to a final concentration of 50µg/ml followed by 0.2% (vol/vol) SDS. The cell suspension was incubated overnight at 37°C.

After overnight incubation, DNA was sheared by passing through a blunt 19-gauge needle 8 times. Six millilitres of phenol/chloroform/isoamyl alcohol (25:24:1) was added, mixed, and centrifuged at 16000 x g for 5 minutes to separate. The upper aqueous phase was collected, added to a fresh 15ml conical tube, and the extraction was repeated. This was repeated twice more using 6ml chloroform/isoamyl alcohol (24:1). DNA was precipitated from the resulting aqueous phase by addition of 2 volumes of ethanol along with one-tenth volume of 0.3M sodium acetate (pH 5.2) and incubated overnight at -20°C. The DNA/ethanol solution was centrifuged at 16000 x g for 30 minutes at 4°C to isolate the DNA precipitate. This was washed and re-centrifuged twice with ice-cold 70% vol/vol ethanol. The final pellet was air dried for a

few minutes and resuspended in 100 - 200µl of Tris-EDTA (TE) buffer by incubating the sample at 60°C for 10 minutes. DNA concentration was measured using the NanoDrop Spectrophotometer and DNA was stored at -20°C.

2.2.2 Preparation of HPV genomes for copy number controls and Southern blot probes

To prepare HPV DNA probes, HPV genome containing plasmids were digested using restriction enzymes to extricate HPV genomes from pUC-HPV16 and pGEMII-18 vectors. Twenty units of restriction enzymes and 5µl of appropriate buffer were added to 10µg of pUC-HPV16 and pGEMII-18 plasmids containing whole HPV16 and HPV18 genomes. Total volume was made up to 50µl with DNase free H_2O. *Bam*H I was used to cut the HPV16 genome from the pUC-HPV16 vector and *Eco*RI was used for HPV18. After addition of the restriction enzymes, DNA solutions were vortex mixed, briefly centrifuged, and incubated overnight in a water bath at 37°C. Two hundred nanograms of digested HPV DNA was run on a 1% wt/vol DNA agarose gel alongside 200ng of undigested HPV DNA. The 1% agarose gel consisted of 2g DNA agarose (Bioline), 40ml TBE buffer (89mM Tris (pH 7.6), 89mM boric acid, 2mM EDTA), 160ml H_2O, and ethidium bromide (0.2µg/ml). Gels were run in 0.5% wt/vol TBE buffer. Complete digestion of plasmids was monitored: a band at 7.9kb for the HPV genomes and a second at 2-3kb, for remaining pUC and pGEMII vectors. The gels were placed on a UV illuminator and the 7.9kb band was cut out using a sterile scalpel. A QIA quick gel extraction kit (Qiagen) was used to extract the HPV genomes according to manufacturer's instructions. DNA concentration was determined using the NanoDrop spectrophotometer (see 2.3.3) and DNA was stored at -20°C.

For Southern blotting, HPV16 and HPV18 copy number standards were prepared equivalent to 10, 50, 100, or 150 copies per cell. Preparations were carried out using carrier salmon sperm DNA (10µg/ml). The size of each plasmid was used to determine molecular weight in Daltons (g/mol): Molecular weight was equal to the number plasmid base pairs multiplied by 660g/mol, where 660g/mol is the average mass of a single base pair. The following formula was used to determine the weight of one plasmid:

$$\textit{One copy of viral plasmid (g) = molecular weight (g/mol) / } 6.02\times10^{23} \textit{ (molecules/mole)}$$

This was multiplied by 50 to determine the number of grams for 50 copies of plasmid. Given that 5µg of each DNA sample was being assessed in a Southern blot, to prepare an equivalent 50 copies of plasmid per cell, the following formula was used to determine an approximate number of cells in 5µg (or 5×10^{6}pg) of sample, where one cell contains 6.6×10^{-12}g (or 6.6pg) of DNA:

$$\frac{5\times10^{6}\text{pg}}{6.6\text{pg}} = 7.6\times10^{5} \text{ cells}$$

Finally, to determine how many grams of purified HPV16- or HPV18-plasmid to load to attain an equivalent of 50 copies per cell, the following formula was used:

$$\textit{50 copies of HPV genome per cell (g) = } (7.6\times10^{5}) \textit{ x (50 copies of plasmid (g)).}$$

2.2.3 Nucleic acid spectrometry

Concentration and purity of nucleic acid samples were measured using the NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific). This first involved cleaning of the upper and lower measurement pedestals using nuclease free water and lint-free lab wipes, also carried out between each measurement. The NanoDrop was blanked using 1µl of nuclease free water or TE buffer, and 1µl of DNA or RNA was added to the lower measurement pedestal. Nucleic acid concentrations were measured in ng/µl, and 260/280 ratios were used to measure purity. A 260/280 ratio of ~1.8 for DNA and ~2.0 for RNA was indicative of good sample purity.

2.2.4 Southern blotting

2.2.4.1 Digestion of total DNA isolated from HPV+ keratinocytes

Five µg of total DNA from HPV16- or HPV18-containing keratinocyte samples were digested with restriction enzymes. DNA was incubated overnight at 37°C with 10 units of restriction enzymes and 2µl of appropriate buffers. Additionally, *Dpn*I was added to all digests to remove any residual input DNA via specific cleavage of methylated sites, not present in HPV genomes. Total volumes were made up to 20µl with DNAase free H_2O. *Bam*H I was used to linearize HPV16 by cutting the genome at a single site and *Hind*III was used as a non-cutter. For HPV18 genomes, *Eco*RI was used as a single cutter and *Bgl*II as a non-cutter. Following overnight incubation, enzymes were heat inactivated by incubating samples at 65°C for 10 minutes and samples were stored at -20°C.

2.2.4.2 Agarose gel electrophoresis

Digested total DNA samples were run on 0.8% wt/vol agarose gels, consisting of 4g molecular grade DNA agarose dissolved in 500ml 1 x TBE (89mM Tris-HCl, 89mM Boric acid, 2mM EDTA) and 5µl of 10mg/ml ethidium bromide. Once set, gels were placed in Fisherbrand™ horizontal electrophoresis tanks (Thermo Fisher Scientific) and flooded with 2.5L TBE containing 20µl ethidium bromide. Each digested DNA sample (see 2.2.2) was mixed with 10µl DNA loading buffer (0.3% wt/vol bromophenol blue, 0.1% wt/vol xylene cyanol FF, 30% vol/vol glycerol, in water), and was loaded alongside an HPV copy number control (50 copies of HPV genome per cell, see 2.2.2) and a 1kb DNA ladder (Quick-load® Purple 1kb DNA ladder; New England Biolabs). Gels were run overnight at 50V until the methylated blue had migrated at least two thirds of the gel. To assess equal loading of DNA samples, ultraviolet (UV) light was used to detect the intercalating agent, ethidium bromide. UV images of each gel were obtained using a GeneFlash transluminometer (Syngene, UK).

2.2.4.3 Transfer protocol

After electrophoresis, DNA was transferred from the agarose gel to Gene Screen plus™ nylon membranes (PerkinElmer, UK) via capillary transfer. Depurination of the HPV DNA containing agarose gels involved two washes in 0.25M HCl for 20 minutes. This was followed by a neutralisation and denaturation phase, where agarose gels were washed once in 0.4M NaOH for 30 minutes.

For the transfer, four sheets of Whatman™ 3MM filter paper (23 x 33cm) were soaked in 0.4M NaOH and draped over a glass plate into a tray filled with 2L 0.4M NaOH, setup so that each

opposite end was submerged in solution. The agarose gel was placed face down on top of the Whatman paper stack, followed by the nylon membrane (22.5 x 20cm), pre-soaked in 0.4M NaOH, and four more NaOH soaked Whatman squares (21 x 23.5cm). Air bubbles were carefully removed after each addition to the stack. To complete, dry paper towels were stacked on top, evenly covering all of the filter paper without any gaps. Another glass plate was placed on top along with a centred weight to compress the paper towels. Transfer stacks were left overnight at room temperature. NaOH diffused from the reservoir up towards the dry paper towels, drawing DNA from the agarose gel to the nylon membrane via this capillary action.

After capillary transfer, the membrane was soaked in 0.5x saline-sodium citrate (SSC, 300mM NaCl, 30mM sodium citrate, pH 7.0) and DNA was covalently bonded to the membrane via crosslinking using a UV crosslinker (Stratalinker, Stratagene, UK). The nylon membrane was finally wrapped in cling film and stored at -20°C.

2.2.4.4 Preparation of radiolabelled HPV DNA probe

HPV16 and HPV18 genomes were linearized and isolated from pUC-HPV16 and pGEMII-18 vectors via restriction enzyme treatment as described (2.2.4.1). To form the probe, 50ng of the HPV DNA was added to RNase free H_2O to make up a total volume of 13µl. DNA was denatured by boiling at 95°C for 5 minutes, followed by 2 minutes incubation in an ice slurry mix. In a specialised lab for radioactive working and behind a Perspex safety screen, 5µl of oligonucleotide labelling buffer (1mg/ml random deoxynucleotide primers, 1M HEPES, pH 6.6) was added to each probe along with 2.5µl molecular-grade BSA (10mg/ml), 50µCi α-^{32}P-dCTP

(EasyTides, Perkin Elmer), and 1.5µl Klenow DNA polymerase I (Thermo Fisher). This tube was mixed and briefly centrifuged before incubating in a lead pot at room temperature for 2-3 hours. Finally, the volume was made up to 50µl with TE and the probe was purified using an Illustra Probe Quant G-50 microcolumn (Amersham Biosciences). This involved centrifugation at 780 x g for 2 minutes to separate radiolabelled DNA from unincorporated nucleotides. Radiolabelled HPV DNA was stored in a lead pot at -20°C until use.

2.2.4.5 Probe hybridization to DNA, membrane washing and autoradiography

Two times hybridization buffer was composed of 5x SSC (750mM NaCl, 1.5M sodium citrate, pH 7.0), 10x Denhardt's solution (0.2% wt/vol Ficoll 400, 0.2% wt/vol polyvinylpyrrolidone, 0.2% wt/vol BSA fraction V, and 20% wt/vol dextran sulphate; all Sigma Aldrich). This was mixed 1:1 with deionised formamide along with SDS, which was added to a final concentration of 0.1% wt/vol. To complete the pre-hybridization buffer, 200µl of 10mg/ml salmon sperm was denatured via boiling for 10 minutes at 100°C, cooled for 2 minutes in an ice slurry and added to 10ml of 1x hybridization buffer.

Nylon membranes containing transferred DNA were thawed and laid over a square of gauze that had been soaked in distilled water. These were rolled up together and positioned into a glass hybridization canister (Hybaid, UK). Pre-hybridization buffer was added to the canister, and this was incubated for an hour in a hybridization oven (Hybaid, UK), which gently rotated the canister at 42°C for 1 hour, blocking non-specific DNA binding sites on the membrane.

During this incubation period, the hybridization buffer was prepared, containing the radiolabelled DNA probe. Two hundred microlitres of 10mg/µl salmon sperm DNA was added to the radiolabelled probe (2.3.2.4) and was boiled at 100°C for 10 minutes. This was cooled on an ice slurry for 2 minutes before adding to 10ml 1x hybridization buffer. After 1 hour of incubation, the pre-hybridization buffer in the canister was replaced with the hybridization buffer containing the radiolabelled probe and was incubated overnight rotating at 42°C.

Stringency washes were carried out to remove non-specifically bound DNA to the hybridized membranes. Washes were carried out in a large glass dish, where the membranes were gently agitated on a plate shaker at room temperature in varying SSC-SDS solutions. Prior to the first wash, membranes were placed in the dish containing 100ml 2x SSC, 0.1% SDS and wiped with a soaked sponge to remove unbound probe. For the first wash, the 2x SSC, 0.1% SDS solution was replaced with an additional 400ml of the same solution and agitated for 15 minutes. For each wash, solutions were replaced with 500ml fresh SSC-SDS solution. Table 2-1 shows the complete set of washes that were carried out.

For the final wash, 0.1x SSC, 1% SDS solution was pre-warmed to 55°C and added to the membranes, which were rocked for 30 minutes at 55°C. For autoradiography, membranes were enfolded in cling film in the dark; positioned into a cassette with autoradiograph film (GE Healthcare) and an intensifying screen. X-ray film was exposed for varying times at -80°C to slow β-particles. Films were developed using a Compact ×4 automatic processor (Xograph Healthcare, UK). For copy number controls, radio-probed membranes were exposed on phosphor screens and radioactive signals were imaged on a FLA-9000 scanner (FujiFilm, USA).

Table 2-1: Order of SSC-SDS solutions and conditions for washes of DNA probed nylon membranes

Solution	No. of washes	Temp	Time (min)	20X SSC	20% SDS	ddH_2O
2X SSC, 0.1% SDS	2	RT	15	100ml	5ml	895ml
0.5X SSC, 0.1% SDS	2	RT	15	25ml	5ml	970ml
0.1X SSC, 0.1% SDS	2	RT	15	5ml	5ml	990ml
0.1X SSC, 1% SDS	1	55°C	30	2.5ml	25ml	470ml

2.3 RNA analyses

2.3.1 RNA extraction

RNA extractions were carried out according to manufacturer's instructions using the RNeasy Mini kit (Qiagen, UK). In brief, cell pellets frozen at -80°C (2.1.3.2) were thawed on ice and resuspended in an appropriate volume of lysis buffer (buffer RLT), supplemented with 10µl β-mercaptoethanol. To disrupt and homogenize, cells were passed through a 19-gauge needle eight times. One volume of 70% vol/vol ethanol was added to the lysate, which was centrifuged at 8000 x g in a RNeasy Mini spin column containing a silica gel membrane. The spin column, now containing bound RNA, was re-centrifuged with a stringent wash buffer (buffer RW1) to remove biomolecules such as proteins and fatty acids, followed by two spins in a mild wash buffer (buffer RPE) to remove salt traces. Finally, spin columns were centrifuged

with RNAse free water or TE buffer to elute the RNA. RNA concentration was determined using the NanoDrop spectrophotometer and RNA was stored at -80°C.

2.3.2 DNase treatment of RNA

RNA-extracts were DNase treated using the DNase I, RNase-free kit (ThermoFisher, UK). This involved incubating 1µg of RNA extracts with 1µl (1U) DNase I (cat. No EN0521), 1µl 10x reaction buffer with $MgCl_2$, and nuclease free water made up to a final volume of 10µl for 30 minutes at 37°C. These volumes could be adjusted accordingly to DNase treat larger quantities of RNA. To inactivate DNase, 1µl of chelating agent EDTA (50mM) was added, and RNA was incubated at 65°C for 10 minutes. DNase treated RNA concentration was read on the NanoDrop and cDNA synthesis was carried out immediately.

2.3.3 cDNA synthesis

Reverse transcription of RNA was performed using the Tetro cDNA Synthesis Kit (Bioline, UK). A primer premix was prepared which consisted of up to 5µg RNA, 1µl Random Hexamer primers, 1µl 10mM dNTP mix, 4µl 5x RT buffer, 1µl RiboSafe RNase Inhibitor, 1µl Tetro Reverse Transcriptase (200 u/µL). The premix was made up to a final volume of 20µl with nuclease-free water. Using a PTC-200 PCR machine (Biorad), samples were incubated for 30 minutes at 45°C, and then at 85°C for 5 minutes to terminate the reaction. After cooling at 4°C, cDNA was stored at -20°C.

2.3.4 qPCR

2.3.4.1 Oligonucleotide primer design

Previously published oligonucleotide primer sequences were obtained from either the Harvard Primer Bank (https://pga.mgh.harvard.edu/primerbank/), or the NCBI Genbank database (https://pga.mgh.harvard.edu/primerbank/). HPV16- and HPV18-E8^E2 primers were designed using Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Basic local alignment search tool (BLAST) searches (http://www.ncbi.nlm.nih.gov/BLAST/) were carried out to validate that all primer sequences were specific to target sequences only. Primers were synthesized by Sigma-Aldrich and melting temperatures (T_m) used were according to the manufacturer's instructions. Primer sequences are summarised in Table 2-2.

Table 2-2: qPCR primers

Target protein	Forward primer sequence (5'–3')	Reverse primer sequence (5'–3')	Amplicon length (bp)	Melting temperature (T_m) °C
HPV16 E8^E2	GGCAATACTGAAGTGGAAACTCAGC	GTTTCTTCGGTGCCCAAGGC	119	60
HPV18 E8^E2	GGCTGTTCTGAAGTGGAAGCAA	GGTGAGGGGGTGTGCTGTAG	96	60
CD40	ACATACAACCAAACTTCTCCCCG	GCAAAAAGTGCTGACCCAATCA	119	59
CCL2	CAGCCAGATGCAATCAATGCC	TGGAATCCTGAACCCACTTCT	190	63
IL18R1	CCTTGACCCTTTGGGTGCTTA	CTCATGTGCAAGTGAACACGA	125	60
IL1B	ATGATGGCTTATTACAGTGGCAA	GTCGGAGATTCGTAGCTGGA	132	57
IL15	TTTCAGTGCAGGGCTTCCTAA	GGGTGAACATCACTTTCCGTAT	129	60
β-actin	CTCTTCCAGCCTTCCTTCCT	GGATGTCCACGTCACACTTC	88	59

2.3.4.2 Relative quantification of gene expression

Quantitative polymerase chain reaction (qPCR) was performed to detect and determine relative quantities of a target nucleic acid sequence or gene present in a sample. Each qPCR reaction was made up to a total volume of 20µl and consisted of 10µl SensiFAST™ SYBR® (Bioline, UK), 0.125µl forward and reverse primers (10mM), 7.75µl nuclease-free water, and 2µl of cDNA sample. Input samples were typically used at 50ng per reaction. Reactions from each sample were added in triplicate to a 96 well-plate. No template controls and standard curves were included on every plate to monitor contamination and assess primer efficiency. Acceptable standard curve ranges were between 80 and 120%.

After preparation, plates were either run immediately or were stored at 4°C for up to 8 hours. Reactions were carried out on an MX3000P qPCR System (Agilent, UK). Cycling conditions consisted an initial single denaturing step at 95°C for 2 minutes, followed by 40 cycles of: denaturation for 10 seconds at 95°C, annealing at a primer specific temperature (Table 2-2) for 30 seconds, and an extension step for 75°C for 15 seconds. MxPro software (Agilent) was used to analyse qPCR data, including generation of amplification plots, standard curves, and dissociation curves. Amplification plots showed the cycle number plotted against a fluorescent value. This was used to determine the threshold cycle (C_t), a horizontal line of intersection set in the linear phase, defined as the number of cycles required for the fluorescent signal to be detected above background signal. This was used to assign C_t values to each reaction. The delta delta ($\Delta\Delta$)C_t method was used to calculate relative expression fold change between each sample and an endogenous control. Where ΔC_t = Ct (gene of interest) – Ct (endogenous

control), ΔΔCt = (endogenous control ΔCt) – (gene of interest ΔCt), fold change in expression = $2^{-\Delta\Delta Ct}$.

2.3.5 Illumina RNA-sequencing and Nanopore-sequencing

RNA was extracted from untransfected and HPV16- and HPV18-containing HTK#2 and HTK#3 cell cultures, comprising six samples in total at passages 4 – 5 (between 30 - 40 population doublings). Extracted RNA was DNase treated (section 2.2.1), then sample preparation and RNA-sequencing were conducted by Genomics Birmingham. Here, poly(A)+ RNA was isolated with oligo-dT primers using the Oligotex mRNA mini kit (Qiagen). RNA quality was verified using the TapeStation (Agilent), where RNA integrity numbers (RIN), used to measure RNA quality, demonstrated high-quality, intact mRNA, showing at between 9.6 and 10.0 RIN (scaled between 1 – 10) for all samples (data not shown).

Nanopore sequencing (Nanopore) was carried out using direct Oxford Nanopore sequencing kits (Oxford Nanopore technologies, UK [SQK-RNA002]), by Genomics Birmingham and according to manufactures instructions (see Schwenzer, Abdel Mouti [424]) and sequenced at around 1 – 2 million reads per sample. Initial bioinformatics analysis was conducted by Dr Boris Noyvert as described in Chapter 4, section 4.2.

For short-read Illumina RNA-sequencing (RNA-seq), mRNA libraries were prepared using NEBnext® Ultra™ RNA library prep kit (New England Biolabs, US). Libraries were run on a NextSeq 550 System (Illumina), using a high-output flow cell, providing read-depths of around 70 – 80 million reads. Bioinformatic analyses including transcript mapping, quantification of

gene expression, differential analysis of gene expression, and gene set enrichment analysis were carried out by Dr Boris Noyvert, as described in Chapter 5 (section 5.6). Further differential gene expression analyses in the form of volcano plots were carried out using GraphPad prism and protein-protein interaction analysis was performed using StringDB (https://string-db.org/).

2.4 Protein analysis

2.4.1 SDS-PAGE and western blotting

2.4.1.1 Protein lysate extraction

Cell pellets stored at -80°C (2.1.3.2) were thawed on ice and resuspended in urea lysis buffer (ULB; 8M Urea, 150mM NaCl, 25mM Tris [pH7.5]). Before adding to cell pellets, 1ml of urea lysis buffer was supplemented with 10µl, 150mM β-Mercaptoethanol and 10µl of protease inhibitors. Two hundred microlitres ULB was added to each cell pellet and mixed. Protein lysates were sonicated for 10 seconds at 30% amplitude, centrifuged at 15,000 x g for 15 minutes at 4°C, and the supernatant was transferred to freshly labelled 1.5ml micro tubes. Protein lysates were stored at -20°C.

2.4.1.2 Bradford assays

Bradford assays were performed to determine protein concentration of each sample. Two-fold serial dilutions of BSA in ddH_2O were used to construct a reference curve of known protein concentrations, ranging from 2mg/ml to 0.125mg/ml. Ten microlitres of BSA standards (undiluted) and 10µl protein lysates (diluted 1:5 in ddH_2O) were added to individual cuvettes. Standards and samples were resuspended in 1ml Bradford reagent (Biorad), also diluted 1:5

in ddH_2O. Absorbance was measured using an Eppendorf Biophotometer (Eppendorf) at 495nm, after the spectrophotometer was blanked with a sample containing 10µl ULB and 1ml diluted Bradford reagent. Absorbance readings of each standard were plotted onto a graph on Microsoft Excel and protein lysate concentration was determined by comparison to the standard curve.

2.4.1.3 Gel electrophoresis

Protein lysates were run on 10% wt/vol polyacrylamide Mini-PROTEAN® precast gels (Bio-rad, UK). To set up electrophoresis tanks, gel cassettes were placed into an electrode assembly and fitted into the tank (Mini-PROTEIN Tetra Cell; Bio-Rad), which was flooded with 1x SDS-running buffer (192mM glycine, 25mM Tris (pH 7.5), and 0.1% vol/vol SDS). The comb was carefully removed from the gel cassette and wells were gently rinsed out with running buffer. Prior to loading, 20µg of protein lysate was mixed at a 1:1 ratio with loading buffer (5% vol/vol β-mercaptoethanol in Laemmli buffer) and made up to a final volume of 20µl with ddH_2O. Samples were boiled at 95°C for 10 minutes, briefly centrifuged and added to the gel alongside a lane containing 6µl of PageRuler™ Plus prestained protein ladder (Thermo Scientific, UK). Once set up, proteins were separated by electrophoresis for 1 – 1.5 hours at 100 volts.

2.4.1.4 Semi-dry protein transfer

Semi-dry transfer was performed using a Trans-Blot Turbo Transfer System (Bio-Rad, UK). Eight pieces of Whatman™ chromatography paper (Sigma-Aldrich) and one piece of Polyvinylidene difluoride (PVDF) membrane (brand) were cut to a proportionate size to the gel. Prior to setup, Whatman paper was soaked in semi-dry transfer buffer (48mM Tris, 39mM

glycine, 0.037% vol/vol SDS, 20% vol/vol methanol) and the PVDF membrane was soaked in methanol. The transfer sandwich composed; four pieces of pre-soaked Whatman paper, PVDF membrane, the gel and four more pieces of Whatman paper. Bubbles were removed after addition of each layer to the stack. The transfer sandwich was positioned into a cassette and run on the Trans-Blot Turbo at a constant 25 volts and up to 1 amp for 30 minutes. Once complete, PVDF membranes containing protein were quickly transferred into Tris buffered saline with 0.1% vol/vol Tween 20 (TBST; 50 mM Tris-HCl and 150 mM NaCl, pH 7.6).

2.4.1.5 Membrane blocking, antibody staining and imaging

Ponceau S (0.1% wt/vol Ponceau S, 5% vol/vol acetic acid) was used to assess protein transfer and was subsequently washed off with PBS. The PVDF membrane was then blocked in either 5% wt/vol milk or 3% wt/vol BSA in TBST for 1 hour at room temperature. All washing and blocking were carried out on a rocker. After blocking, primary antibodies, along with housekeeping gene mouse anti-β-actin monoclonal antibodies, were added to the membrane at specified dilutions in blocking buffer and incubated overnight at 4°C (see Table 2-3 for list of antibodies). Membranes were washed 3 times for 5 minutes in TBST to remove non-specifically bound proteins. Secondary conjugated horseradish peroxidase (HRP) antibodies, anti-mouse or anti-rabbit (Table 2-4), were added to the membrane at a 1:10,000 dilution in blocking buffer and rocked at room temperature for 1 hour. The membrane was washed for a final 3 x 5 minutes in TBST and treated with an enhanced chemiluminescent (ECL) HRP substrate (Supersignal^{TM} West Pico PLUS; ThermoFisher, UK). Protein bands were visualized using a Fusion FX digital chemiluminescence detection system (Vilber Lourmat, Collégien, France).

Table 2-3: Primary antibodies

	Primary antibodies	Animal	Size (kDa)	Supplier/No.	Western dilution	IF dilution
Cytokeratin and other differentiation markers	K1/K10	Mouse	-	8.60	-	1:50
	NCL-CK10	Mouse	56.5	Novasastra	-	1:50
	K7	Mouse	51.4	Dako - M701829-2	-	1:50
	K8	Mouse	40/68	Sigma – M20	-	1:100
	K13	Mouse	-	Roberts	-	1:100
	K14	Mouse	-	Roberts	-	1:100
	Involucrin	Mouse	120	Sigma-Aldrich	-	1:100
HPV	E4	Rabbit	66	R424	-	1:1000
	E4	Mouse	66	1D11 (N-term)	-	1:2
	HPV18 L1	Mouse	-	Nova Costra – HPV18 NCL	-	1:100
	HPV16 L1	Mouse	-	Roberts	-	1:100
	HPV16 E7	Mouse	14	Santa Crus – sc-6981	1:200	-
	HPV18 E7	Mouse	14	Abcam – NCL	1:1000	-
Cell cycle and proliferation	P16 (N-20) SC-467	Rabbit poly	16	Santa Cruz	-	1:50
	MCM7	Mouse	80	Santa Cruz – sc-9966	1:200	1:100
	Cyclin B (H433)	Rabbit	48	Santa Cruz	-	1:500
	P53 (D01) sc-126	Mouse	53	Santa Cruz	-	1:100
	RB p105 (RB1)	Mouse	105	BD bio (554136)	-	1:250
	RB p130 (RB2)	Mouse	130	BD bio (610267)	-	1:250
Immune response	CD40	Rabbit	30	Abcam - ab228818	1:10,000	1:500
DNA damage	γH2AX	Mouse	14	Upstate	-	1:250
Housekeeping genes	Beta actin	Mouse	42	Sigma	1:50,000	-

Table 2-4: Secondary antibodies

Secondary antibodies	Animal	Supplier	Western dilution	IF dilution
Alexa-Fluor® 488 Goat Anti-Mouse (Green)	Goat	Invitrogen	-	1:500
Alexa-Fluor 594® Goat Anti-Rabbit (Red)	Goat	Invitrogen	-	1:500
Goat anti-Mouse IgG, HRP conjugated	Goat	ThermoFisher	1:10,000	-
Goat anti-Rabbit IgG, HRP conjugated	Goat	ThermoFisher	1:10,000	-

Chapter 3: High-risk HPV life cycle organisation in primary tonsil keratinocyte models

3.1 Introduction

Even though the global incidence rate of HPV-positive oropharyngeal squamous cell carcinomas (OPSCC) is growing [31, 32], the natural history of HPV-OPSCCs still remains largely uncharacterised. A major factor to this shortage of knowledge has been due to an absence of pre-malignant lesions in the oropharynx. In the cervix, the study of disease progression has been driven by discovery of neoplastic lesions of increasing dysplasia, which have been of enormous benefit to understanding the natural history of cervical carcinogenesis.

One alternative route to biopsies would be establishment of HPV replication models that accurately recapitulate the molecular phenotype of oropharyngeal subsites sensitive to HPV-driven carcinogenesis. Establishment of primary human keratinocyte models - primary keratinocytes of the foreskin and cervix transfected with HR-HPVs - have provided a foundation for investigation of early viral infection after it was found that the HPV genomes were maintained in an episomal form in extended culture [425]. HPV was shown to immortalise but not transform the primary cells and importantly the HPV containing cells were able to undergo complete productive life cycles in organotypic raft culture [426, 427]. Given that the tonsil crypts are the primary site of HPV-positive tumours in the oropharynx (see

introduction, section 1.4.2), several research groups (including the Roberts group) have used primary tonsil keratinocytes as a recipient cell for high-risk HPV genomes and used these models to gather information regarding HPV replication, virion production and viral genome integration in this cell type [401, 428].

To better understand HPV16 infection and identify key virus-host events in viral pathogenesis in the oropharynx, the Roberts group established an HPV life cycle model in primary human tonsil keratinocytes (HTKs). HTKs isolated from three donors were transfected with recircularized genomes of HPV16 or HPV18 and grown in extended cell culture to establish isogenic cell lines (HPV-HTK#1, #2, and #3). Because HPV18 is rarely found in HPV-OPSCC (see introduction, section 1.4.3), the comparison of isogenic cell lines may enable identification of causes behind the differences of pathogenicity at this anatomical site [31, 32, 282]. Thus, we hypothesize that HPV16 life cycle organization in the tonsils is distinct to that of HPV18. If the molecular organization is different, then differences in virus-host interactions might have differential effects upon disease progression at this site. Since cell lines established from cervical premalignant disease show molecular progression upon extended cell culture [394, 429], the isogenic HPV16- and HPV18-HTK cell lines were also grown in extended cell culture to establish longitudinal models. Characterisation of donors #1 and #2 was initiated by former members of the group - Peter Rae and Elizabeth Marsh; this included the analysis of the structure of the viral genomes and analysis of virus and cellular events upon differentiation of the cells in three-dimensional (3D) organotypic cell culture.

The overall objective of this study was to complete the characterisation of the three donor cell lines, including the isogenic cell lines established from donor #3 to provide a fully characterised HPV-HTK longitudinal model system. To this end, HPV16- and HPV18-HTK#3 (HPV-HTK#3) cells were also grown in extended cell culture and Southern blotting was carried out to assess the physical status of HPV DNA. In addition, 3D organotypic rafts were developed and analysed to compare the spatial organization of the HPV life cycles between HPV16 and HPV18. In order to better conceptualise the viral life cycle in the tonsils, comparisons were made to an anogenital site – the foreskin – using HPV16 or HPV18 –positive cell lines established in primary foreskin keratinocytes [159, 430, 431]. The HFK cell lines were generated by previous members of the Roberts group.

Together these studies aimed to: (1) characterize the HPV life cycle and compare to foreskin, and (2) compare the virus life cycle organization of two oncogenic HPV types (16 and 18) with very different frequencies in OPSCC in primary tonsil keratinocytes.

3.2 Generation of longitudinal model of HPV-HTK#3

HPV16- and HPV18-HTK cell lines in donor #3 were established by Lucia Minoni (visiting Ph.D. student from International Agency for Research on Cancer, Lyon, France) and Sally Roberts, but they had not undergone extended cell culture or Southern analysis to assess the physical status of the viral genomes. Alongside this experiment, another HPV16-positive cell line was established in HFK (donor #12). The cells were revived from liquid nitrogen storage and grown on a feeder layer of irradiated 3T3 J2-fibroblasts with continuous passaging of approximately 170 - 200 population doublings (PD) over one- to two-year periods. At the end of the study

period, HPV16-HTK#3 and HPV18-HTK#3 had each undergone approximately 168 and 194 PD respectively. The HPV16-positive foreskin cell line was grown to around 170 PD.

Total DNA was extracted from each passage (P) and Southern blotting was carried out to identify whether HPV genomes were established as extrachromosomal episomes and whether they were maintained in an episomal form during extended cell culture or whether HPV DNA integrated into the host genome. Prior to Southern analysis, DNA was digested with restriction enzymes that either digested the HPV genome at a single site as to linearize the viral genome or that did not have a recognition site in the viral genomes. DNA from HPV16-HTK#3 was digested with *Bam*HI, which cuts the viral genome at a single position, or *Hind*III, which does not cut the HPV16 genome. HPV18 was digested with either *Eco*RI (single cutter) or *Bgl*II (non-cutter). *Dpn*I was added to all digests to remove any residual input DNA. If viral genomes were replicating as episomes, they would migrate through a DNA agarose gel as supercoiled DNA and open/relaxed circles following incubation with an enzyme that does not cut the viral genome, or as a linear structure of approximately 7900bps upon restriction with an enzyme that has a single site within the viral DNA. For example, Southern blot analysis of early passage HPV18-HTK#3 DNA showed the presence of supercoiled and open circle forms of the viral DNA after *Hind*III treatment, and linear forms following *Eco*RI treatment (Figure 3-1a).

3.2.1 HPV18-HTK#3 longitudinal model

Southern blotting of HPV18-HTK#3 cells revealed that while HPV18 established as episomes, a marked decrease of episome copy numbers was observed with increasing cell passage (Figure 3-1). Additional bands were visible in the later passages, particularly P11 onwards (>84

PD), that likely indicate virus-host junction fragments (identified by red arrowheads in Figure 3-1b and c). In *Eco*R1 digestions, linearized genomes were detected up to P22 but there was no detection of episomal forms in these high passage cells. These data together suggest that in HTK#3 donor cells, HPV18 established as episomes, which were maintained in early passages but were not persistent in longer term culture. The detection of possible virus-host DNA fragments in addition to a linear genome suggests integration of the HPV18 genome. It is possible that cells containing integrated HPV DNA attained selective growth advantage over cells containing HPV episomes, resulting in the gradual decrease of detectable HPV episomes. An event similarly noted by Pett, Alazawi [394] who observed progressive increases in colony-forming efficiency after progressive integration of HPV16 genomes in W12 cells grown in monolayer.

3.2.2 HPV16-HTK#3 longitudinal model

In HPV16-HTK#3 cells, HPV16 established and underwent stable persistent replication as episomes up to P11 (~78 PD) (Figure 3-2a). However, at P11 the cells ceased proliferating and entered a crisis-like phase, in which cell division stalled, cells appeared non-uniform in shape and size, and colonies lost their distinctive cobblestone appearance (Figure 3-2b). After approximately 2-3 months in this quiescent-like state cell growth reinitiated and HTK colonies reverted to typical cobblestone-like appearances. Southern blotting detected possible virus-host junction fragments. This crisis-like phase was a consistent event in HPV16-HTK#3 cells, revived from liquid nitrogen at different times, but not observed in the growth of HPV16-HTK#2 or HPV18-HTK#3 cells.

3.2.3 HPV16 HFK#12 longitudinal model

In the human foreskin cell line HPV16-HFK#12, HPV16 similarly established as episomes, which persisted up to P19 (125 PD) and then copy number decreased and was not detectable after P23 (151 PD) (Figure 3-3). The emergence of two additional bands of approximately 6000 and 11000bps were observed between P12 and P25, most likely detecting viral DNA that had integrated into host cellular DNA - virus-host junction fragments.

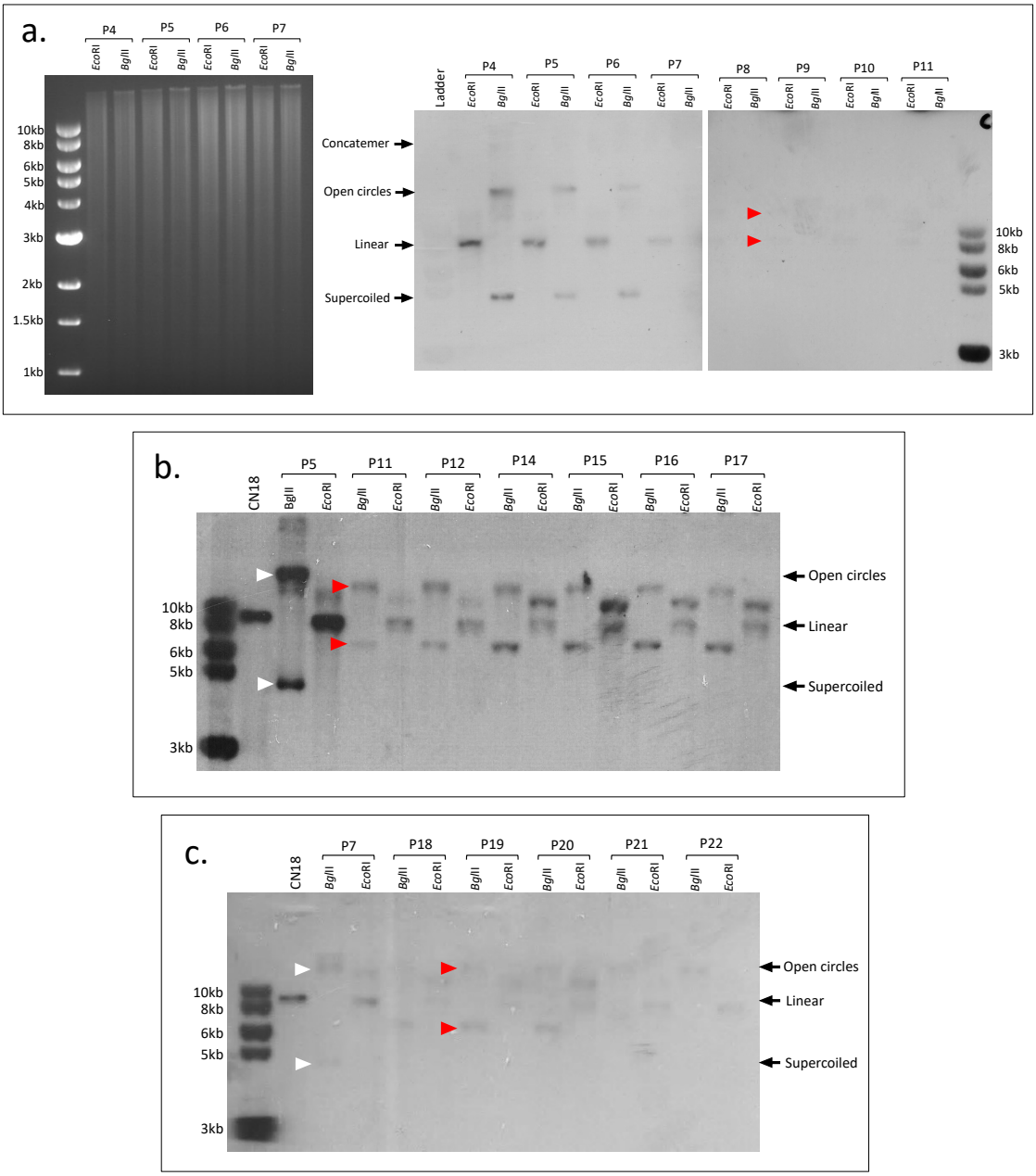


Figure 3-1: Southern blotting of HPV18-HTK#3 cells

a. Left hand image shows example DNA-agarose gel prior to capillary transfer, demonstrating equal DNA loading. Southern blots (middle and right) demonstrate genomic HPV DNA between P4 and P11. **b.** Southern blots for HPV18-HTK#3 between P11 and P17. **c.** Southern blots for high passage HPV18-HTK#3 cells between P18 and P22. White arrows indicate episomal bands. Red arrows indicate potential virus-host junctions. CN18 = HPV18 copy number, equivalent to 50 whole genome copies per cell. Episomal HPV18 DNA treated with *Bgl*II migrated as circular supercoiled genomes and open/relaxed circles. Bands at ~8.0 kb indicate linear HPV DNA after *Eco*RI restriction enzyme digestion.

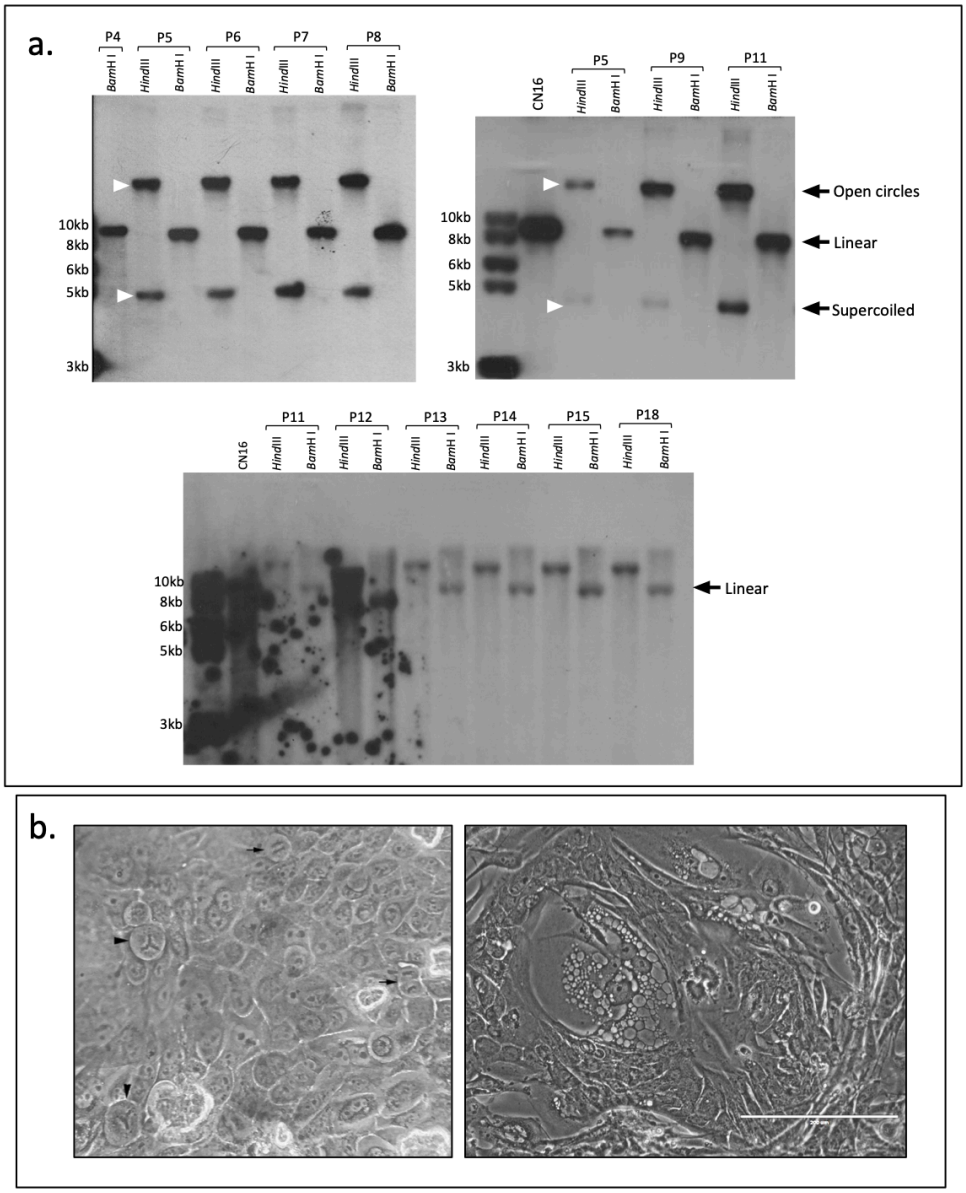


Figure 3-2: Southern blotting and morphology of HPV16-HTK#3 longitudinal model

a. HPV16-HTK#3 Southern blots. P = passage number, CN16 = HPV16 copy number, equivalent to 50 whole genome copies per cell. HPV16 DNA treated with *Hind*III display circular supercoiled whole HPV genomes (4kb) and HPV whole genome open circles. Bands at approximately 8.0kb show linearized HPV DNA after *Bam*H I restriction enzyme digestion. White arrows indicate episomal bands. **b.** left side image shows morphology of HPV16-HTK#3 at P5. Cells present typical cobblestone-like appearance along with high numbers of mitotic cells, marked with arrows, many of which present abnormal mitotic tendencies such as tripolar spindle mitoses. The right-side image shows atypical HPV16-HTK#3 colony during the crisis-like phase at P11. Scale bars indicate 200 microns.

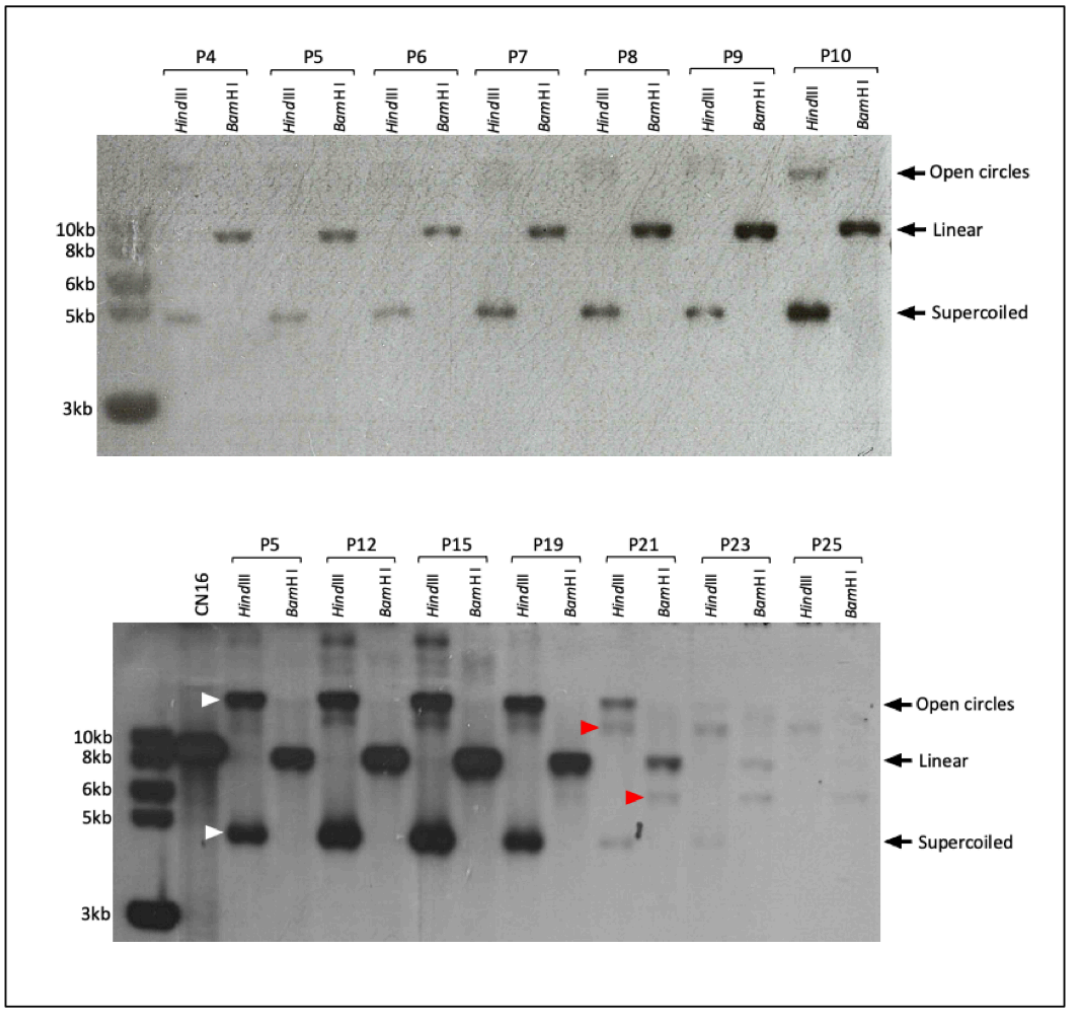


Figure 3-3: Southern blotting of HPV16-HFK#12 longitudinal model

P = passage number, CN16 = HPV16 copy number, equivalent to 50 whole genome copies per cell. HPV16 DNA treated with *Hind*III display circular supercoiled whole HPV genomes (4kb) and HPV whole genome open circles. Bands at approximately 8.0kb show linearized HPV DNA after *Bam*H I restriction enzyme digestion. White arrows indicate episomal bands. Red arrows indicate potential virus-host junctions.

3.2.4 Comparisons to previously established HPV-HTK models

Figure 3-4 illustrates a comparative summary of HPV16 and HPV18 physical status in all three isogenic HTK donor lines and one of the HPV16-HFK donors, #12 over extended cell culture. Southern analysis of the HPV-HTK cell lines revealed that two out of three HPV16-HTK donor lines (#2 and #3) presented persistent replication of episomes upon extended culture. Integration events were not evident in HPV16-HTK#2 even at high PD (~150 PD) but did occur in HPV16-HTK#3 cells at around 78 PD.

In HTK#1 primary cells, HPV16 DNA did not appear to establish as detectable episomes, but as the primary cells showed extended life span, HPV genomes most likely integrated very early on in growth of the cells (data not shown). When the same donor cells (HTK#1) were transfected with HPV18, episomes were initially established but copy number decreased over long-term passage, an event that occurred in all three HPV18-transfected HTK donor models between P7–P12 (40 – 70 PD). Concordant with loss of episomes, emergence of additional bands of different molecular weight, likely virus-host junction fragments, were observed in all three HPV18-HTK donor lines suggesting integration of HPV DNA into the host genome. While HPV16 genomes in HTK#3 cell lines did appear to integrate, the process of integration was substantially different to the HPV18-HTK lines, including its HPV18-HTK#3 counterpart: where episome copy numbers decreased gradually in HPV18-HTK#3 lines (undetectable ~50 PD), episomes in HPV16-HTK#3 cells were stably maintained up to P11 (~80 PD), but then undetectable following the crisis-like phase, where cell growth stalled for one to two months.

Altogether, these findings show that both HPV16 and HPV18 extended growth of the primary HTK models. In addition, variations in the physical status of viral DNA over extended cell culture may be indicating HPV16- and HPV18-specific differences in viral genome maintenance that could reflect differences in the biology of these alternative virus types in the tonsils.

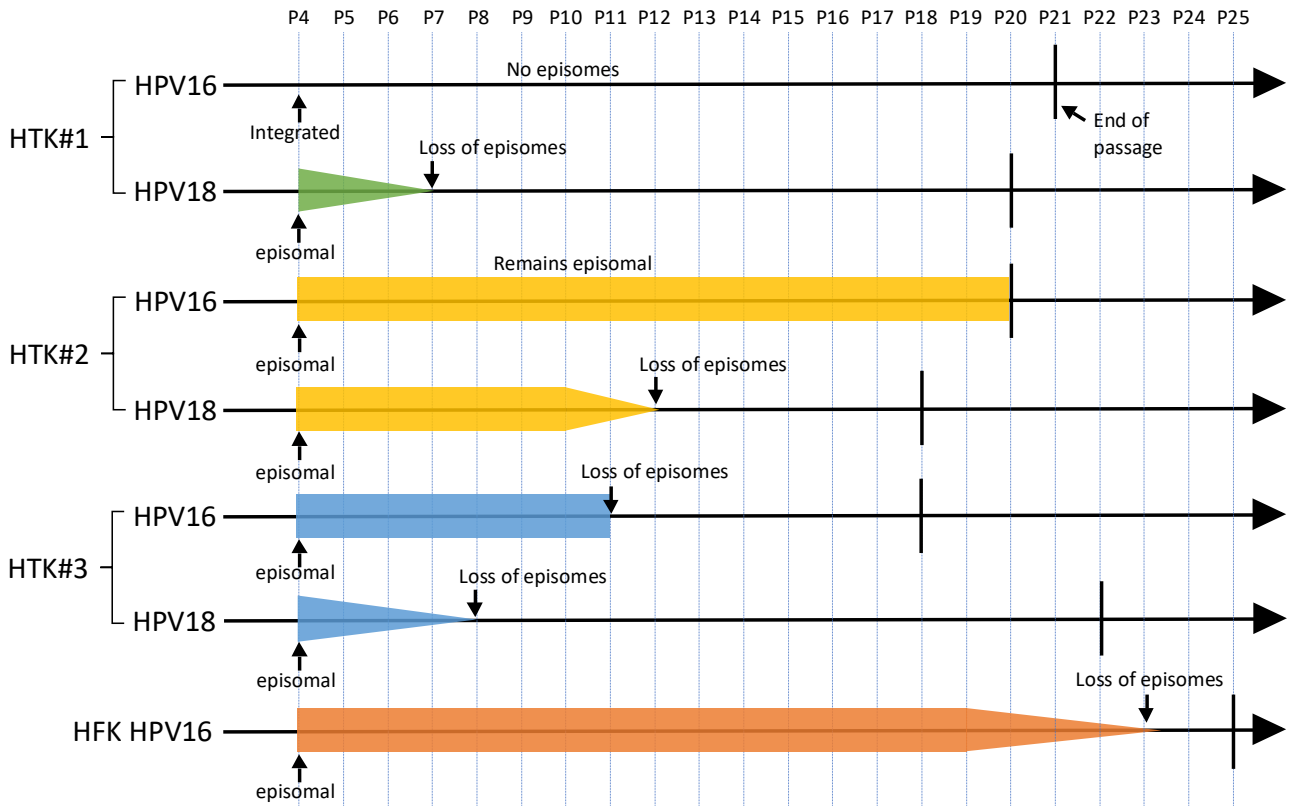


Figure 3-4: Overview of HPV genome status in isogenic longitudinal keratinocyte models.

Diagram illustrates the physical status of HPV16 and HPV18 genomes in each HTK and HFK donor lines during longitudinal passage. Here, cell cultures that exhibited viral episomes are shown as coloured boxes. Thinning boxes show gradual loss of episomes during cell culture, likely indicating integration of the viral genome into host DNA. HPV-HTK cell lines were passaged up to passage (P)18 and P25, where end of passage is shown with a horizontal line.

3.3 Temporal and spatial organization of HPV16 and 18 life cycle in human tonsil keratinocyte organotypic raft culture

Once it was recognised that HPV16 and HPV18 could stably replicate as episomes in the HTK longitudinal models, further analyses were carried out to characterize and compare HPV16 and HPV18 life cycle organisation in the tonsils. Given that the HPV life cycle is dependent on the epithelial differentiation program [432], organotypic 3D raft culture was established from early cell passages of HPV-HTK lines to assess the spatial organisation of the HPV life cycle in the tonsils. HTK#3 and HTK#2 lines were both selected for organotypic raft culture due to their ability to maintain HPV episomes in extended culture as in natural infection, or which were at least able to maintain episomes up to >50 PD (Figure 3-4). Organotypic raft culture and H&E staining of HPV-HTK#2, untransfected HTK#3 and HPV18-HFK#12 lines had previously been carried out by the Roberts group. However, raft culture of HPV16-HTK#3 and HPV18-HTK#3 was still outstanding.

HPV16-HTK#3 and HPV18-HTK#3 cells from early cell culture in which the viral genomes were episomal (between 29 – 38 PD) were seeded and expanded on collagen plugs containing mouse 3T3 J2-fibroblasts - dermal equivalents. Once confluent, the plugs were placed on top of raised steel grids and fed from underneath with media, creating an air-liquid interface. The fibroblasts and media (devoid of epidermal growth factor) provided nutrients and growth factors, which allowed for stratification and differentiation of basal-like HTKs over a two-week period. Rafts were then formalin-fixed, paraffin-embedded, and sectioned.

3.3.1 Morphology of HPV-HTK organotypic raft cultures

In the first instance, H&E staining was carried out to examine the histology of each raft. HTK raft cultures were non-keratinized, recapitulating the physiological structure of mucosal epithelia. In comparison, HPV18-HFK raft culture showed development of a cornified layer, where keratinocytes had differentiated into highly keratinized corneocytes in the most superficial layers of the epidermis, the stratum corneum (Figure 3-5). In addition, HTK raft cultures had a characteristic basket-weave-like organisation, typical of mucosal epithelia, as demonstrated in published examples of cervical and tonsil biopsies [270, 423, 433]. As such, presence of lace-like structure in HTK raft culture further support these epithelial models as mucosal.

HPV16-HTK#3 and HPV18-HTK#3 raft cultures were harvested at 13 days and H&E staining was carried out to examine the histology of each raft in comparison to untransfected-HTK#3 and HPV16/18-HTK#2 raft culture models. Epithelial thickening was observed in HPV18-HTK#3 and both HPV16- and HPV18-HTK#2 raft culture, when compared to untransfected HTK#3 culture. Interestingly, HPV18-HTK rafts were consistently thicker in both donors compared to HPV16-containing counterparts, suggesting that HPV18 was having a greater proliferative effect than HPV16 (Figure 3-6). When compared with HPV-HTK#2 cells, HPV-HTK#3 cells presented with more limited stratification, even though rafts were cultured in the same conditions. In addition, the basal layers of the HPV-HTK#3 rafts appeared much less organized compared to the HPV-HTK#2 rafts and untransfected HTK#3 counterparts. In these rafts, fibroblast cells appeared to have migrated through the plug and accumulated underneath the basal cells, and in some areas, such as in Figure 3-6, basal cells and fibroblast cells appeared

embedded, and it was difficult to distinguish basal keratinocytes from fibroblasts. To ensure that this was not down to technical error, HPV-HTK#2 raft cultures were established on three separate occasions, using multiple collagen plugs each time. Low level stratification was consistently observed.

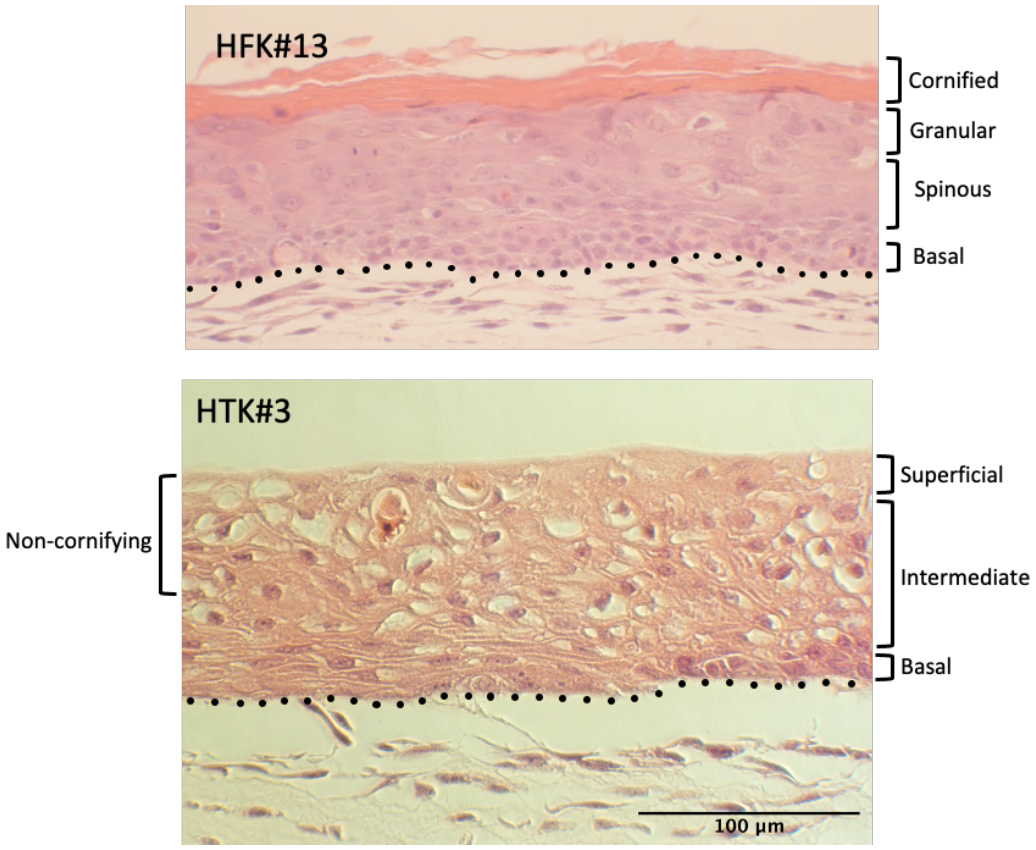


Figure 3-5: H&E staining of human keratinocyte raft culture

Top image shows primary human foreskin keratinocyte (HFK) #12 raft culture. Presents a thick cornified layer of keratinized corneum, composed of layers of protective terminally differentiated cornified cells, below which lies the granular layer. Bottom image shows primary human tonsil keratinocytes (HTK) raft culture, grown from donor #3 basal cells. Consists of mucosal stratified squamous non-keratinized epithelium. As mitotically active basal keratinocytes move away from the basement membrane, cells begin to terminally differentiate and then flatten as they move up intermediate epithelium towards superficial epithelial layers. HTK but not HFK raft cultures show a characteristic basket-weave organisation in the intermediate and superficial epithelial layers, typical of mucosal epithelia. Dotted lines show the basal lamina, which separates basal cells from the fibroblast-containing collagen plug. Scale bars indicate 100 microns.

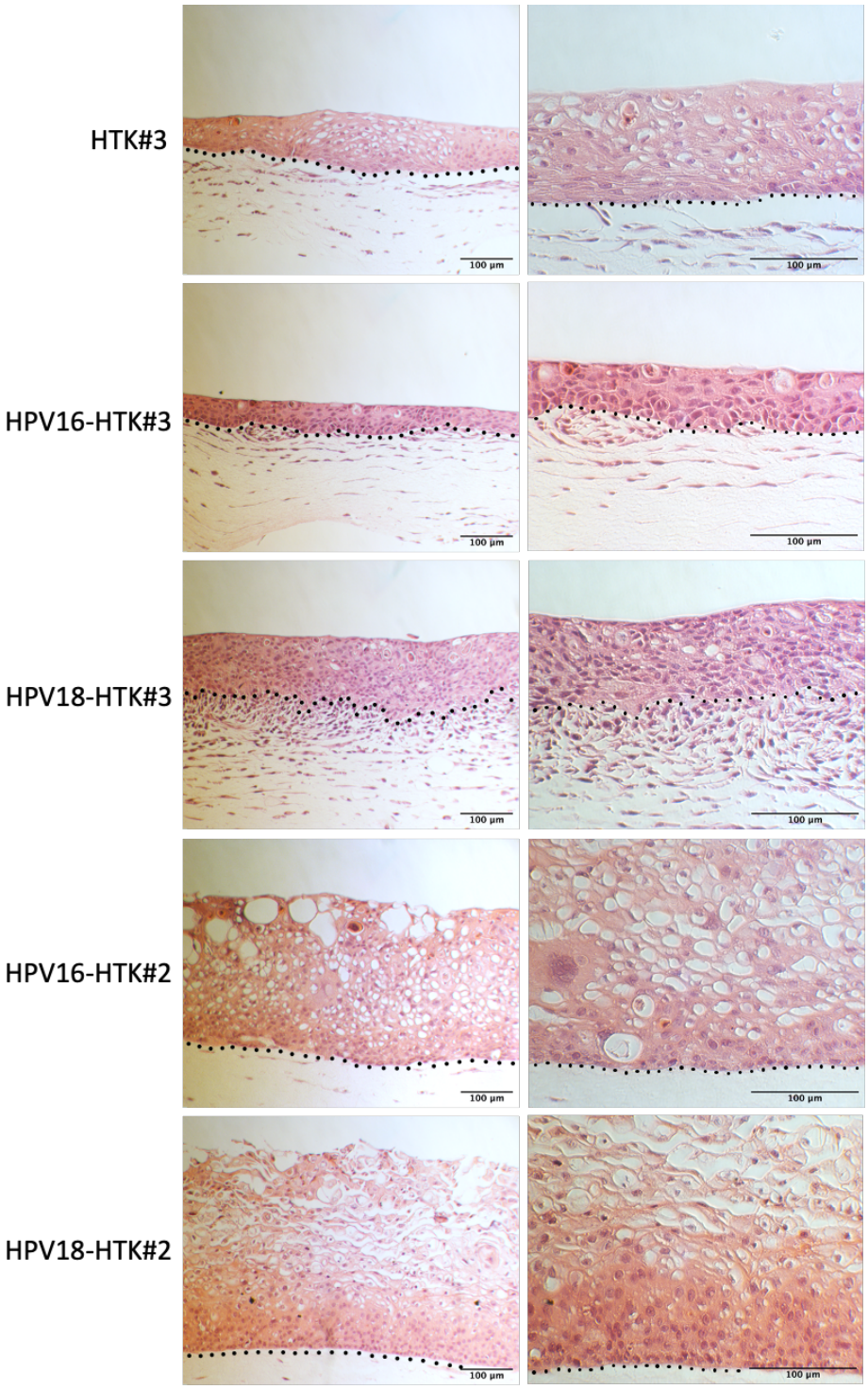


Figure 3-6: H&E staining of human tonsil keratinocyte raft culture

Comparison of H&E-stained raft culture from HTK#2 and HTK#3 lines. Basement membrane is indicated by black dotted line. Images in right hand column show higher power magnification. Scale bars indicate 100 microns.

3.3.2 Immunofluorescence microscopy of early HPV life cycle markers

To identify whether the HTK support complete HPV16 and HPV18 life cycles, immunofluorescence (IF) staining of HPV-HTK#3 raft sections was performed to assess expression of early and late viral proteins required for completion of the viral life cycle. IF staining was also carried out on untransfected HTK#3, HPV16-HTK#2, HPV18-HTK#2, and HPV18-HFK rafts, which had previously been cultured by the Roberts group. Examination of the early viral oncoproteins E6 and E7 was carried out using surrogate markers due to a lack of robust antibodies that could detect E6 and E7 via immunostaining.

Given that the E6 oncoprotein directly suppresses tumour suppressor p53 protein expression via ubiquitin-dependent proteolysis [292, 293] (see introduction section 1.3.3 for more detail), p53 levels were used to assess E6 activity. In untransfected HTK#3 rafts, p53 was primarily expressed in basal and occasional cells of the lower immediate layers - parabasal cells. Expression of p53 was markedly reduced in both HPV-HTK#2 as well as HPV18-HFK raft culture (Figure 3-7) indicating that HPV16 and HPV18 are effectively downregulating p53 protein expression via expression of viral E6 oncoprotein.

In order to promote cell cycle re-entry in differentiated keratinocytes, HPV16 and HPV18 E7 oncoproteins interact with and promote proteasomal degradation of members of the retinoblastoma protein (Rb) family of tumour suppressors, Rb-p105, p107 and Rb-p130. These ‘pocket proteins’ are key negative regulators of cell cycle progression that bind to and dysregulate E2F transcription factors, resulting in transcriptional repression of genes required for G1 to S phase transition as well as activation of genes involved in DNA synthesis (see

introductory section 1.3.3 for more information). Rb staining was therefore used to examine E7 activity.

While the three pocket proteins all regulate E2F transcriptional activity, each protein has variable affinity to different E2F transcription family members [434]. Rb-p105 is mainly expressed in basal and proliferating epithelial cells and has high affinity to the activating transcription factors E2F1, 2 and 3 [435, 436]. In the untransfected HTK#3 raft culture, Rb-p105 was expressed in the nuclei of cells in basal and occasional cells of the lower intermediate regions (Figure 3-8a). In contrast, Rb-p105 was mostly undetected in basal layers of all HPV-containing raft cultures but was present in intermediate epithelial layers, and more so in HPV16 than HPV18-containing raft culture. Rb-p130, which has higher affinity for inhibitory E2F4 - 5 proteins and is predominantly expressed in differentiated epithelium [436, 437], was detected in abundance in the nuclei of suprabasal cells of untransfected HTK rafts. In contrast, Rb-p130 was not detected in all HPV16 and HPV18-containing HTK#2 rafts, and undetectable in HPV18-HFK rafts, apart from upmost superficial HPV16-HTK epithelial layers (Figure 3-8b).

Together, these data suggest that both HPV16 and HPV18 E7 have a strong negative effect on basal Rb-p105 expression in stratified tonsil and foreskin basal keratinocytes but that does not extend to suprabasal epithelial layers. In contrast, Rb-p130 protein levels were significantly diminished in HPV-containing raft culture indicating a profound inhibitory effect from HPV16 and HPV18 E7.

3.3.3 IF microscopy of late HPV life cycle markers

In the later stages of the HR-HPV life cycle, E1^E4 proteins are abundantly expressed in the cytoplasm of suprabasal cells following upregulation of the late viral promoter, resulting in increased transcription of genes involved in viral genome amplification consisting of E1, E2, E1^E4 and E5 [235, 422, 432, 438]. As such, E1^E4 now widely serves as a reliable biomarker of late viral gene expression [235, 439]. In addition, because there is a strong correlation between E1^E4 expression and viral genome amplification [440], E1^E4 also acts as a marker of productive viral infection. However, given that E1^E4 is not directly involved in viral genome amplification [237], not all E1^E4 positive cells are likely to be undergoing DNA amplification. High expression of E1^E4 was detected in the cytoplasm of cells located in intermediate layers of HPV16-HTK#2 and HPV18-HTK#2 raft culture, and in the upper suprabasal (upper spinous and granulosum) layers of the HPV18-HFK model. As expected E1^E4 was undetected in the untransfected HTK#3 model (see Figure 3-7 - Figure 3-12). Surprisingly, E1^E4 was undetected in both HPV16- and HPV18-containing HTK#3 raft culture even though these cells contain the episomal forms of the viral genomes, suggesting that in this donor the oncogenic viruses were not able to complete a productive infection (see all HPV-HTK#3 immunostaining in Figure 3-7 – Figure 3-12).

For both viruses, E1^E4 expression was absent from basal and lower intermediate layers as expected for an HPV protein that is primarily expressed from transcripts activated from the late promoter. However, the spatial suprabasal expression of E1^E4 was slightly different between HPV16- and HPV18-HTK#2 raft culture models. In HPV18-HTK#2 raft culture, E1^E4 expression was limited/confined to cells of the intermediate epithelium and largely absent in

the upper regions of the epithelium, whereas in HPV16-positive rafts, E1^E4 was initially detected higher up in intermediate epithelium and generally extended upwards to the superficial layers (Figure 3-7 – Figure 3-12).

The major capsid proteins L1 and L2 are also transcribed from the viral late promoter [89, 441]. Here, L1 was detected using HPV16 and HPV18 specific antibodies and as expected was not detected in untransfected HTK raft culture (Figure 3-9). As with E1^E4, L1 was also undetected in HPV16-HTK#3 and HPV18-HTK#3 raft culture, further indicating no expression of late viral transcripts or a defect in their translation in this cell line. However, L1 was largely detected in the nuclei of HPV16-HTK#2 and HPV18-HTK#2 raft culture, primarily colocalising with E1^E4 in intermediate epithelial layers. L1 was also detected in the cytoplasm of cells in upper layers of HPV16-HTK#2 raft culture, which could indicate virion formation. Larger clusters of L1 positive nuclei were frequently detected in HPV16-HTK#2 rafts than in HPV18-HTK#2, where L1 positive cells were less common.

Some distinctions in L1 locality were observed between HPV-HTK and HPV-HFK raft culture. In HPV18-HTK raft culture, nuclear L1 was detected in abundance in granular and cornified epithelial layers and did not always colocalise with E1^E4. Here, L1 was commonly detected in more superficial, cornified cells, where E1^E4 production had largely reduced. In contrast, in HPV-HTK models, nuclear L1 was detected lower down in intermediate layers, and almost always in cells that were strongly producing E1^E4

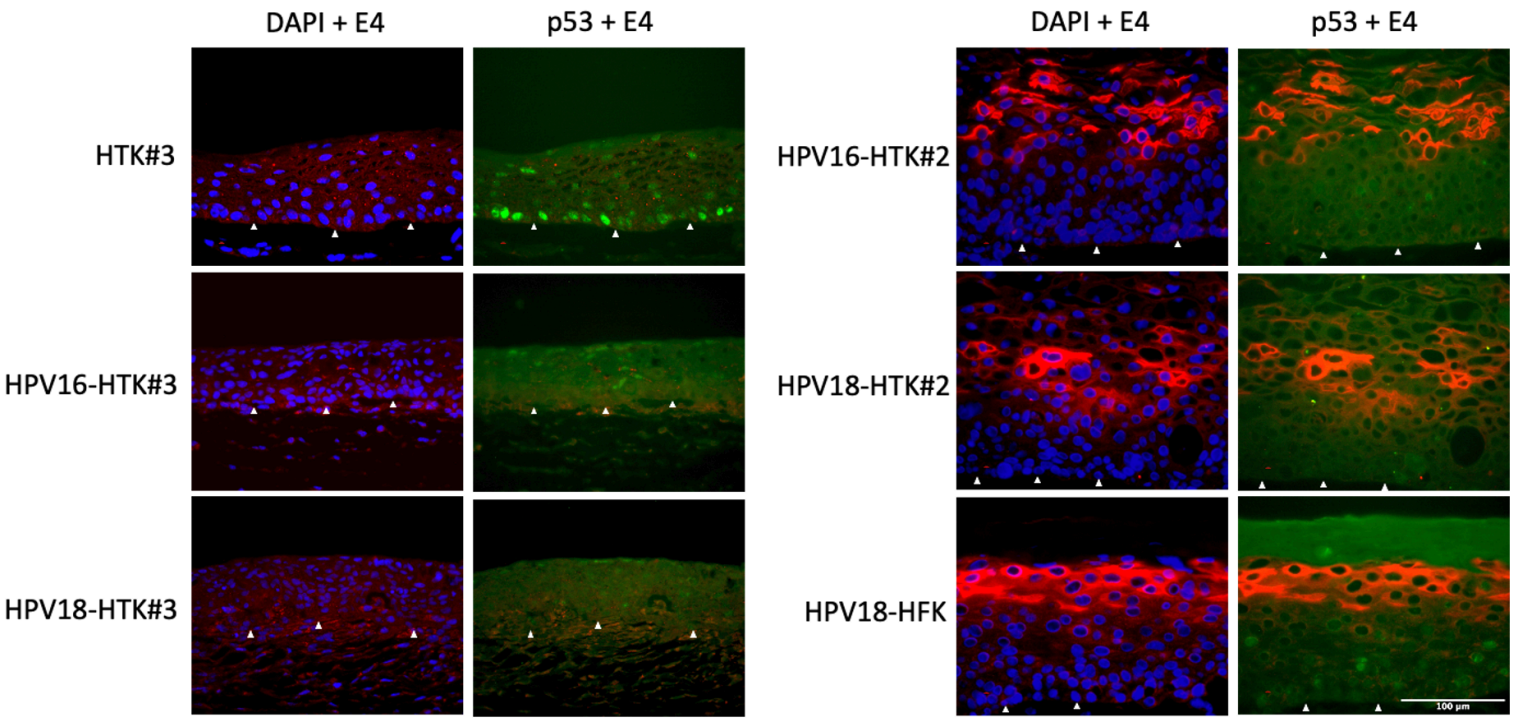


Figure 3-7: Immunofluorescence staining for p53 and E1^E4 in HTK raft culture

Shows HTK and HFK organotypic raft culture, either untransfected or containing HPV16 or HPV18. Raft culture was formalin fixed, paraffin embedded and immunostained with antibodies specific for p53 (green), E1^E4 (red), and DAPI (blue). White arrows indicate basal membrane. Scale bar in bottom right image indicates 100 micrometres.

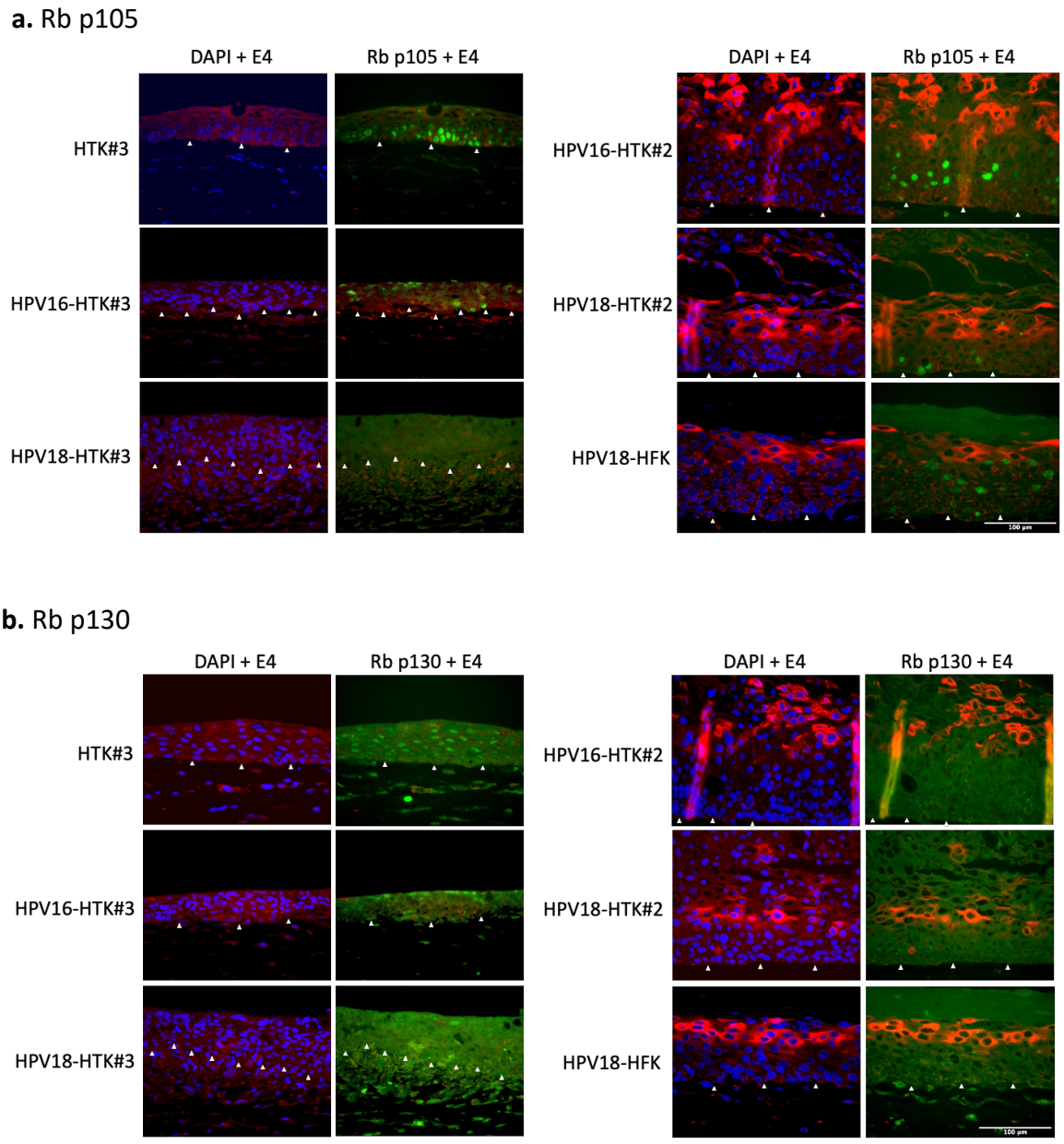


Figure 3-8: Immunofluorescence staining for Rb and E1^E4 in HTK raft culture

Shows HTK and HFK organotypic raft culture, immunostained with antibodies specific for **a.** Rb-p105 (green), E1^E4 (red), and DAPI (blue) and **b.** Rb-p130 (green), E1^E4 (red), and DAPI (blue). White arrows indicate basal membrane. Scale bar indicates 100 micrometres.

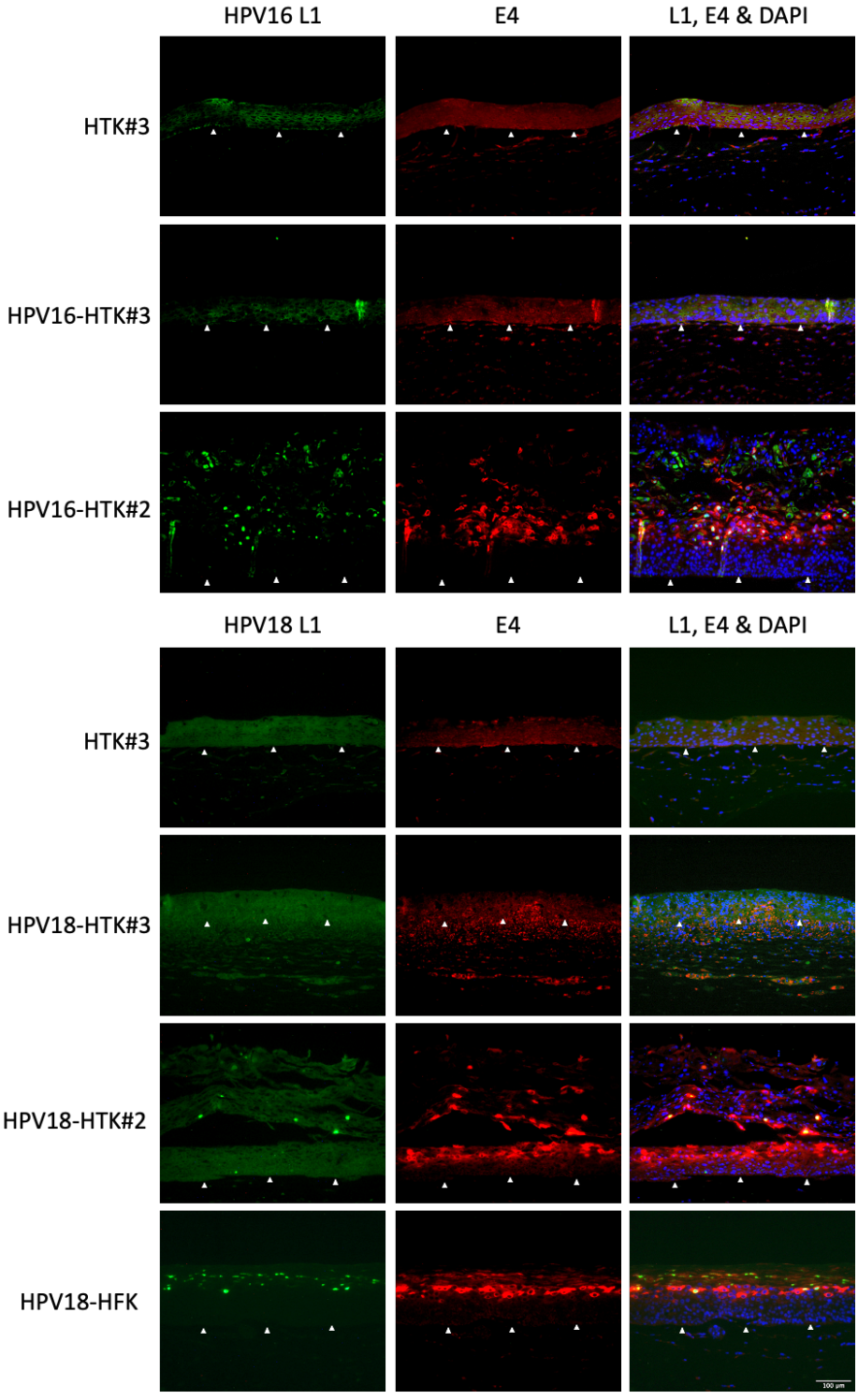


Figure 3-9: Immunofluorescence staining for HPV L1 and E1^E4 in HTK raft culture

HTK and HFK organotypic raft cultures were immunostained with antibodies specific for HPV16 L1/ E1^E4, or HPV18 L1/E1^E4. L1 (green), E1^E4 (red) and DAPI (blue). White arrows indicate basal membrane. Scale bars in the bottom right indicate 100 microns.

3.3.4 HPV16 and HPV18 E7 regulate cell cycle proteins in HTK raft culture

Additional immunofluorescence staining of HPV-HTK raft culture was performed to further evaluate effects of HPV16 and HPV18 on host cell cycle proteins known to be dysregulated by viral early proteins. To better understand and compare spatial HPV16 and HPV18 life cycle events in the tonsils, expression of the cell cycle markers MCM7, p16 and cyclin B1 were assessed in the organotypic HPV-HTK raft models.

The minichromosome maintenance protein complex (MCM) is a DNA helicase composed of six proteins, encoded by the E2F-responsive genes MCM2 – MCM7. The MCM complex controls cell cycle progression via mediation of DNA replication initiation and elongation [442]. Downregulation of Rb pocket protein activity by HPV E7 is known to induce MCM transcription, including that of MCM7, via activation of E2F, thus making MCM7 a suitable marker of E7 activity [88, 215, 216]. Following detection of Rb-p105 in intermediate epithelial layers of HPV-containing raft culture, immunostaining for MCM7 was carried out to better understand whether HPV16 and HPV18 E7 were functionally inactivating the pocket proteins irrespective of Rb-p105 presence.

MCM7 was detected in the nuclei of cells in basal/parabasal layers of untransfected HTK#3 raft culture but not in the post-mitotic differentiating cells of intermediate and superficial epithelium (Figure 3-10). In HPV16-HTK#2 and HPV18-HTK#2 raft culture, MCM7 was mainly detected in basal and parabasal keratinocytes as well as in intermediate epithelial layers. In contrast, MCM7 was abundantly expressed throughout HPV16- and HPV18-HTK#3 raft culture from basal through to upmost superficial layers suggesting uninhibited E7 expression

throughout the raft culture. In HPV18-HFK raft culture, MCM7 was detected in basal, spinous, and granular epithelial layers.

Next, protein expression of the cyclin dependant kinase inhibitor 2A (*CDKN2A*) gene was examined, which encodes the tumour suppressor protein, $p16^{INK4a}$ (p16). Inactivation of Rb-p105 via HR-HPV E7 drives upregulated expression of p16, which is usually regulated by Rb-p105 through a negative feedback loop [217]. As a result, p16 is universally overexpressed in HPV cancers; however, its function as a tumour suppressor protein is lost due to HPV oncoprotein activity. Conversely in HPV-negative HNSCC, p16 expression is often reduced [217, 443], making this protein an important biomarker of HPV oncogenic activity that is widely used in both diagnosis and grading of HPV-associated dysplasia in cervical and oropharyngeal squamous epithelium.

P16 expression was mainly localised to basal cells of untransfected HTK#3 raft culture with some expression also detected in superficial layers (Figure 3-11). In HPV-HTK#2 raft culture, p16 was detected mainly in basal layers in both HPV16- and HPV18-containing lines; however, expression was also extended to lower intermediate layers in the HPV16-HTK#2 models, further demonstrating prolonged E7 expression into cells of intermediate layers. This was similarly witnessed in HPV18-HFK. Strikingly, p16 expression was excluded from epithelial layers expressing E1^E4. In HPV-HTK#3 raft culture, p16 was overexpressed throughout the entire thickness of the epithelium, further indicating viral E7 expression all throughout the epithelia.

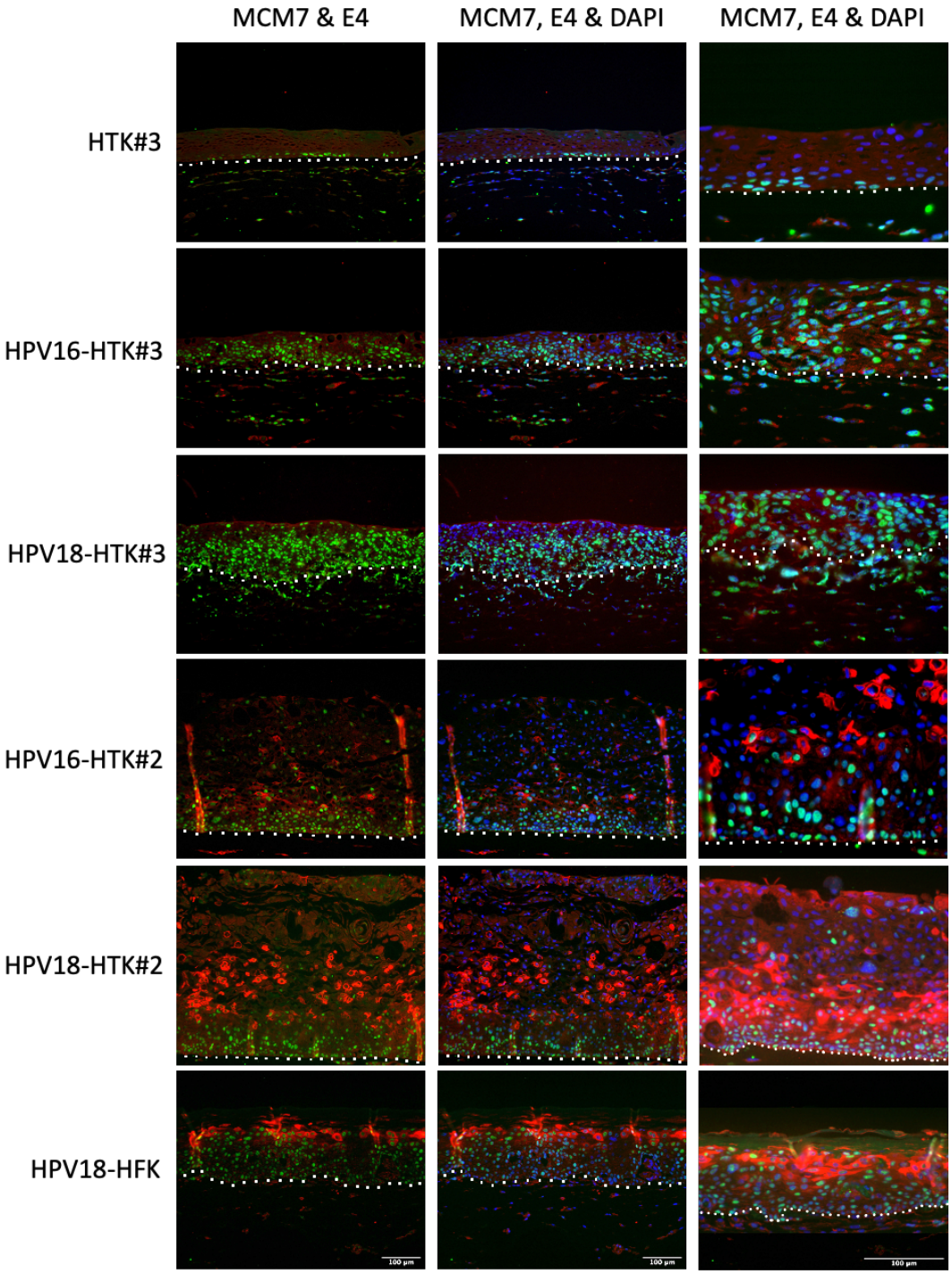


Figure 3-10: Immunofluorescence staining for MCM7 and E1^E4 in HTK raft culture

HTK and HFK organotypic raft cultures were immunostained with antibodies specific for cell cycle/surrogate E7 marker, MCM7 (green), E1^E4 (red) and DAPI (blue). The white dotted lines indicate basal membrane, but this was difficult to determine in the HPV-HTK#3 raft cultures. Scale bars in the bottom row of each column indicate 100 microns.

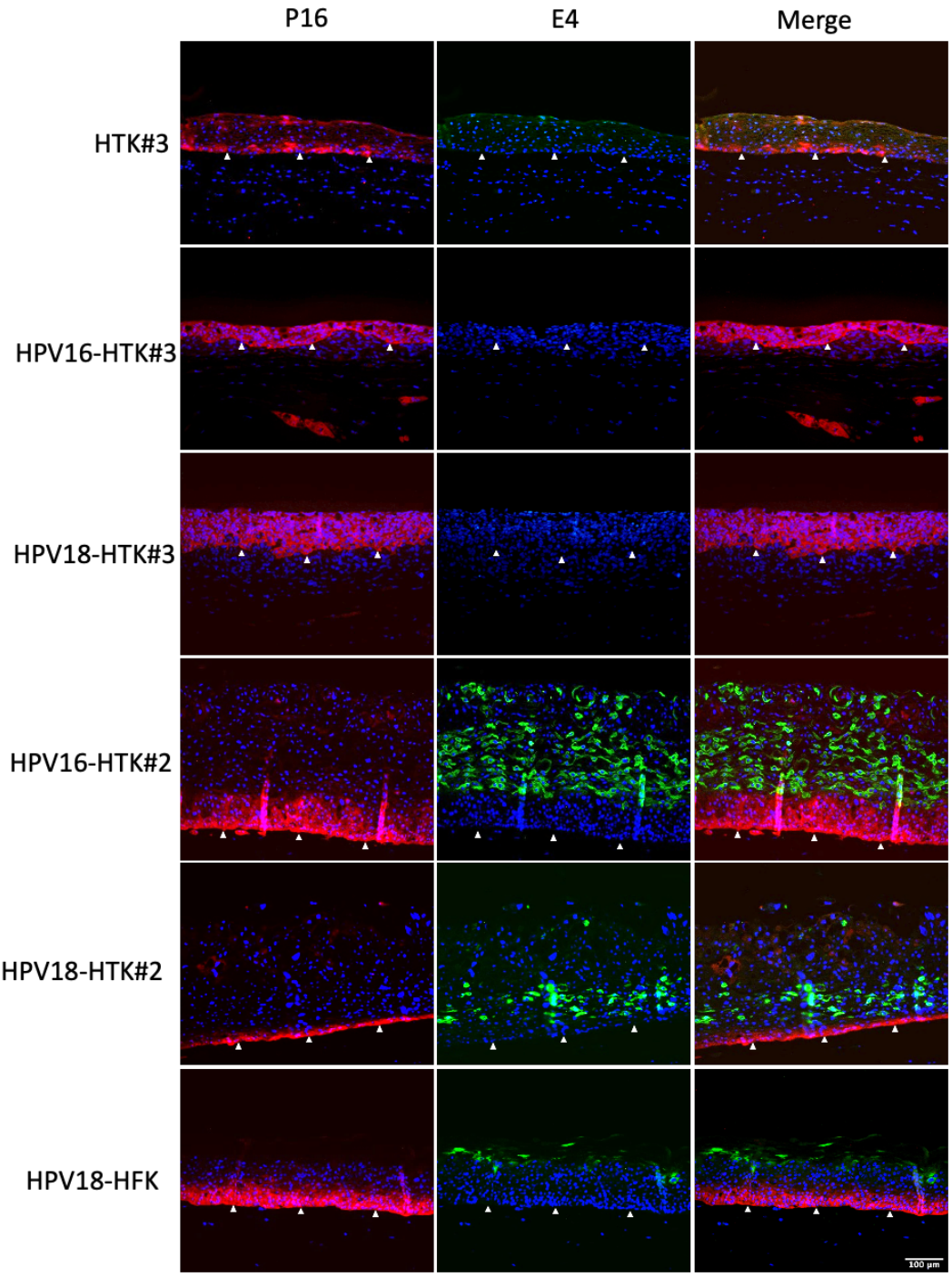


Figure 3-11: Immunofluorescence staining for p16 and E1^E4 in HTK raft culture

HTK and HFK organotypic raft cultures were immunostained with antibodies specific for p16 (red), E1^E4 (green) and DAPI (blue). White arrows indicate basal membrane. Scale bar indicates 100 microns (bottom right).

3.3.5 HPV16 and 18 activate G2-arrest and viral genome amplification in different areas of HTK and HFK raft culture

Previous studies have revealed that productive HPV genome replication is carried out in G2 arrested cells following completion of cellular DNA replication in S phase cells [89, 220]. HPV have been shown to induce G2 arrest by promoting translocation of cyclin B1 into the cytoplasm and away from the nucleus, where it is required for cell cycle progression from the G2-phase to mitosis (see introduction section 1.3.4) [89, 220]. As such, cytoplasmic cyclin B1 was used as a marker of cells in G2 arrest.

In untransfected HTK#3 rafts, cyclin B1 was detected in the cytoplasm and the nuclei of occasional basal cells (Figure 3-12). Comparably significant increases of cyclin B1 positive cells were seen in HPV16/18-HTK and HPV-HFK raft culture, where expression was primarily concentrated in the cytoplasm. However, nuclear accumulation of cyclin-B1 was observed in occasional HPV16-HTK cells, indicating progression from the G2-phase to mitosis and perhaps implying less HPV16-mediated induction of G2-arrest than in the HPV18-containing models. Distribution of cyclin B1-positive cells varied between the two tissue types. In foreskin keratinocytes, cyclin B1-positive cells were distributed throughout the epithelia, often in granular epithelial layers overlapping with E1^E4 positive regions and often colocalising with E1^E4 positive cells, as described previously by Knight, Pugh [237]. Conversely, in HPV-HTK#2 raft culture, areas of cytoplasmic cyclin-B1 - and E1^E4 – positive cells were mutually exclusive, where cyclin B1 expression was limited to lower intermediate and basal layers, and E1^E4 positive cells confined in upper intermediate layers. In both HPV16-HTK#3 and HPV18-

HTK#3 raft culture, cyclin B1 was detected in the cytoplasm of cells located in basal, intermediate, and superficial epithelial layers.

In situ hybridization studies to detect genomic HPV18 DNA in HPV18-HTK#2 and HPV18 HFK raft culture were previously conducted and imaged by the Roberts group to assess viral genome amplification (Figure 3-13). This study showed important distinctions in HPV DNA amplification locality between the two donor sites that were of great relevance following observed differential spatial expression of cyclin B1. Similarly, to cyclin B1, viral genome amplification was shown to take place in lower intermediate epithelial layers in HPV18-HTK#2 raft culture, beginning in parabasal cells. In contrast, amplification was shown to take place in upper spinous and granular epithelial layers in HPV18-HFK raft culture.

3.3.6 Evaluation of DNA damage markers in HPV-HTK raft culture

HR-HPV-mediated stimulation of DNA damage response (DDR) pathways has been shown to be an essential mechanism for differentiation-dependent viral genome amplification [227, 239-241]. Given that viral genome amplification was shown to be carried out at different regions between HPV-HTK and HPV-HFK raft culture, further IF-based investigations were carried out to assess γH2AX, which is phosphorylated and activated in response to HPV-induced DDR pathway activation [227, 432, 444, 445]. In HTK#3 raft sections, γH2AX was infrequently detected in the nuclei of basal or parabasal keratinocytes (Figure 3-14). In HPV16- and HPV18-HTK#2 raft culture, γH2AX was detected in occasional cells in lower intermediate epithelial layers, prior to E1^E4 production, and generally at low level. In upper intermediate and suprabasal, following upregulated production of E1^E4, γH2AX was frequently strongly

detected in the nuclei of E1^E4-positive cells. This was particularly abundant in HPV16-HTK#2 models. In HPV18-HFK raft culture, γH2AX was primarily detected in E1^E4-positive cells at granular epithelial layers.

3.3.7 Overview of HPV life cycle organisation in tonsil and foreskin raft culture

Overall, these studies indicated differential spatial organisation of important HPV life cycle events between HPV-HTK#2 and HPV-HFK raft culture models, which have been summarised and described in Figure 3-15. This figure shows viral life cycle comparisons between HPV16-HTK, HPV18-HTK and HPV18-HFK raft culture models, illustrating expression patterns of E1^E4, L1, viral DNA amplification (HPV18-HTK and HPV18-HFK only), as well as markers of E7 activity, based on MCM7 and p16 immunostaining (Figure 3-10 –Figure 3-12).

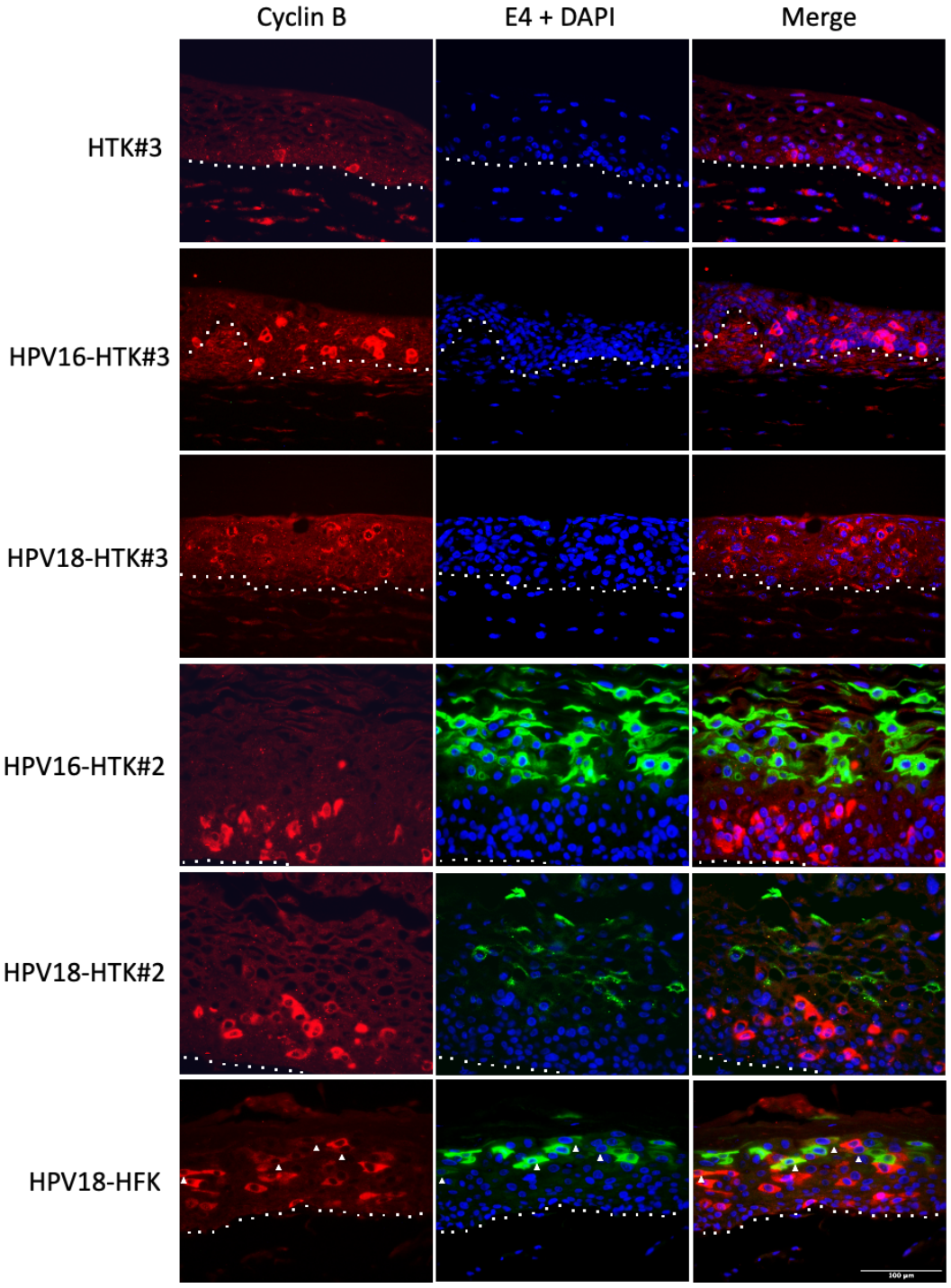


Figure 3-12: Immunofluorescence staining for cyclin B1 and E1^E4 in HTK raft culture

HTK and HFK organotypic raft cultures were immunostained with antibodies specific for cyclin B1 (red), E1^E4 (green) and DAPI (blue). White arrows indicate cyclin B1 and E1^E4 colocalization. White dots indicate basal membrane. Scale bar in the bottom right image indicates 100 microns.

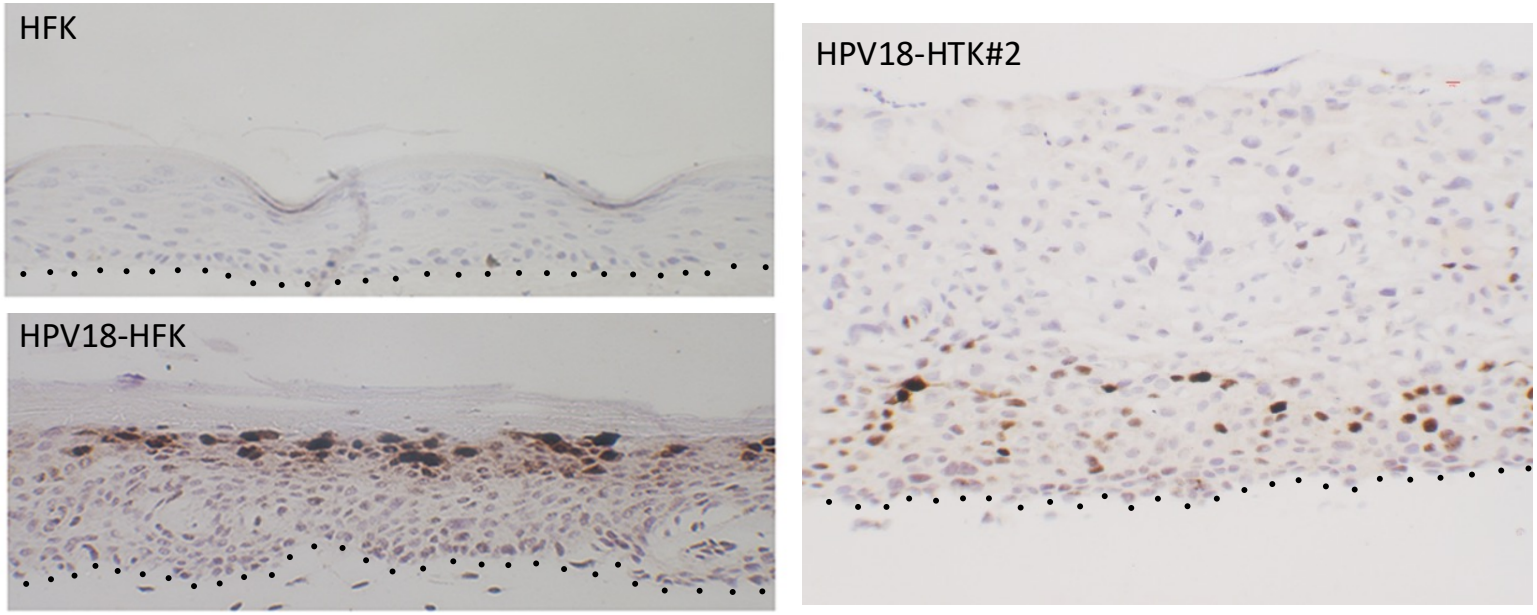


Figure 3-13: In situ hybridization of HPV18-HTK and HPV18-HFK raft cultures

Previously carried out by the Roberts group to assess viral genome amplification in untransfected-HFK, HPV18-HFK and HPV18-HTK#2 raft culture. Here, raft sections were probed for HPV18 genomic DNA, which are observed as brown-positive nuclei. Black dots indicate the basement membrane.

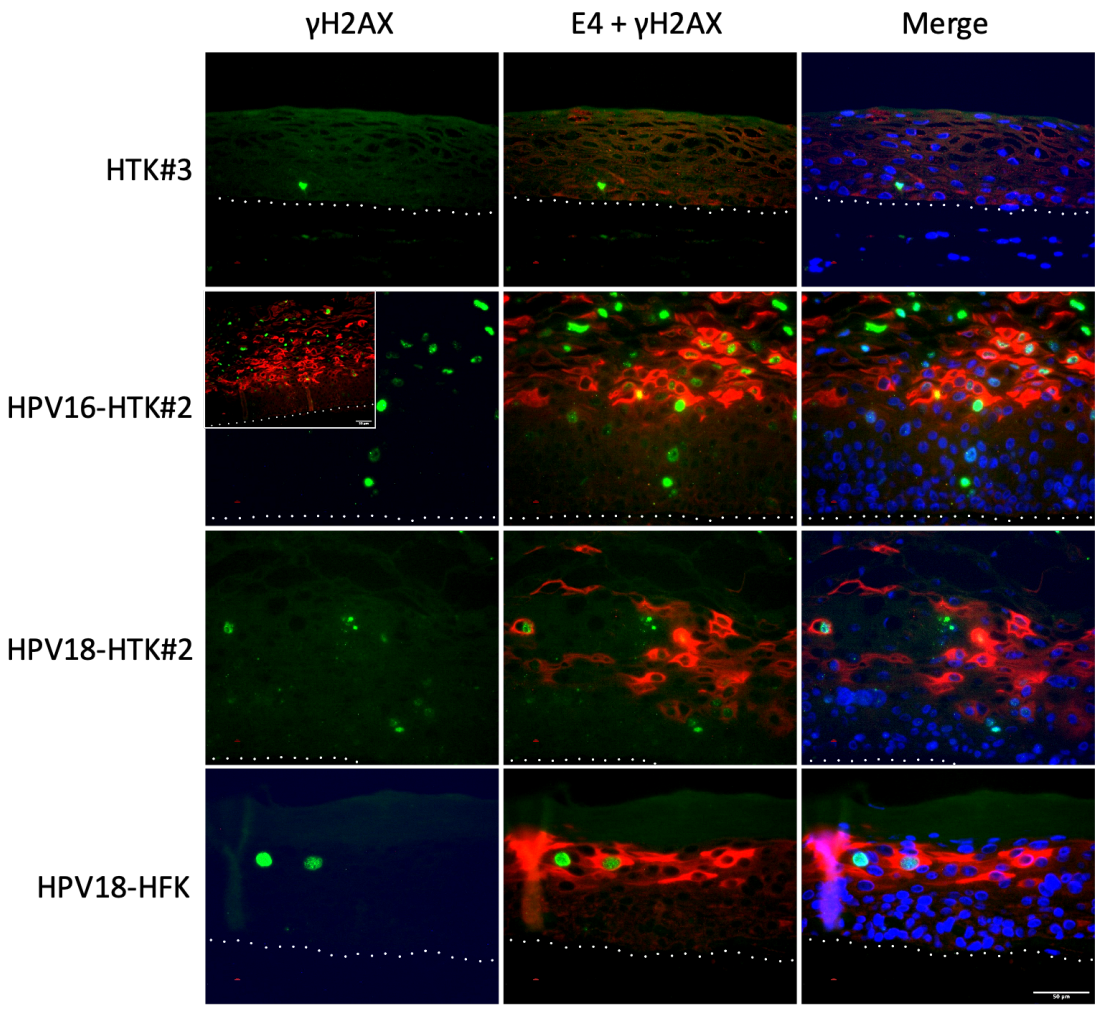


Figure 3-14: Immunofluorescence staining for γH2AX and E1^E4 in HTK raft culture

HTK#3, HPV16/18-HTK#2 and HPV18-HFK organotypic raft culture was immunostained with antibodies specific for γH2AX (green), E1^E4 (red) and DAPI (blue). The smaller box shows E1^E4 and γH2AX immunostaining of HPV16-HTK#2 raft culture at lower magnification. White dots indicate basal membrane. Scale bars indicate 50 microns.

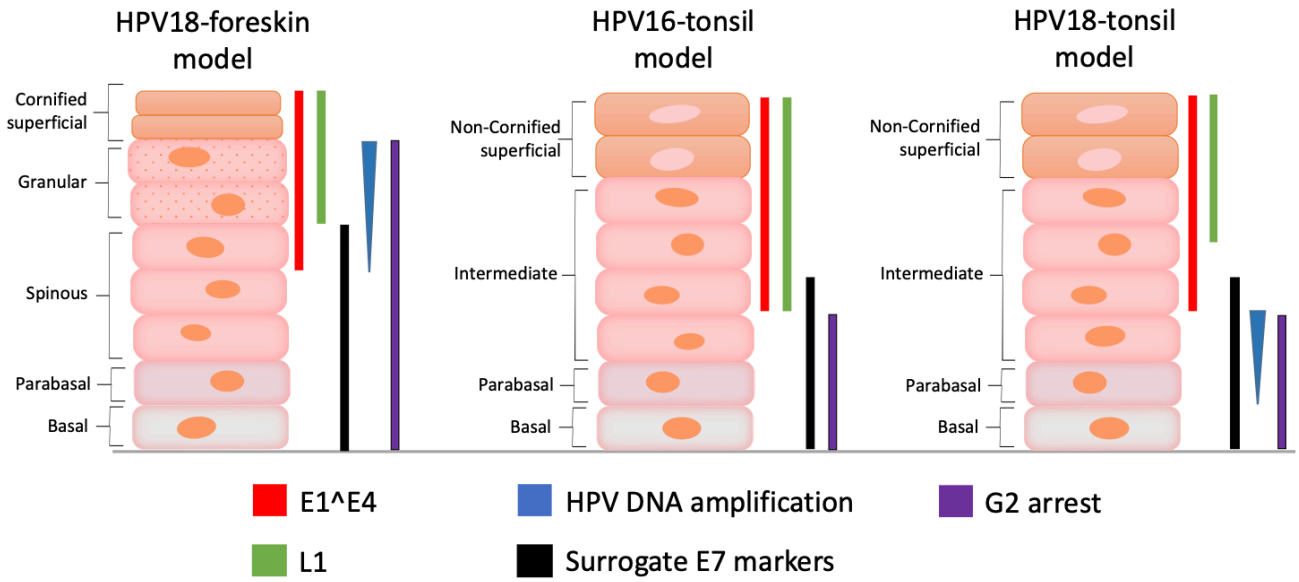


Figure 3-15: Proposed comparative spatial organisation of the HPV16 and HPV18 life cycles in HTK and HFK raft culture

E7 surrogate markers consisted of MCM7 and p16. Cells undergoing G2 arrest were signified by detection of cytoplasmic cyclin B1. Production of late viral proteins E1^E4 and L1 were detected at different stages of epithelial development, occurring much lower down and closer to basal layers in HPV-HTK raft culture. This was also true for viral genome amplification, as signified by increased HPV DNA accumulation via in situ hybridization studies. Importantly, regional separation of E1^E4 and cyclin B1 was uniquely observed in HPV-HTK raft culture, suggesting that E1^E4 was not being abundantly produced during G2 arrest. Further IF-studies of differentiation markers were next carried out to better assess these body site-specific spatial variations of HPV life cycle events (section 3.4). Very few differences in HPV life cycle organisation were observed between HPV16 and HPV18 in HTK raft culture. However, slight variations in E1^E4 and L1 locality was observed between the two viruses, where these viral proteins were frequently detected in higher superficial epithelial layers in HPV16-HTK#2 raft culture than HPV18-HTK#2.

3.4 Keratin and differentiation profiles of HPV-HTK organotypic raft culture

During the HPV life cycle, the switch from usage of the early viral promoter to the late promoter is dependent on cellular differentiation. However, differentiation is delayed by viral oncoproteins E6 and E7, which promote continued cell cycle progression in cells that would have normally terminally differentiated, driving production of host machinery required for viral genome amplification. Following observed spatial differences in expression of late viral genes, E1^E4 and L1 between HPV-HTK#2 and HPV-HFK raft cultures (section 3.3), expression of differentiation markers was performed by IF staining to characterise and compare differentiation between the HPV-HTK and HPV-HFK organotypic raft models. Differentiation was also compared between HPV-HTK#2 and HPV-HTK#3 raft culture, following a lack of late viral protein expression in the HPV-HTK#3 lines, which are normally produced in differentiated keratinocytes. Further, keratin markers specific to stratified tonsil epithelia and tonsil crypt epithelia were used to characterise the physiology of the HTK organotypic raft culture.

3.4.1 IF microscopy-based identification of mitotically active cells in HPV-HTK raft culture

Keratin 14 (K14) is a type I keratin that forms a heterodimer with type II keratin, K5 to form the cytoskeleton of basal cells in stratified epithelia. K14 and K5 are expressed exclusively in mitotically active basal keratinocytes and their expression is downregulated during cellular differentiation [446, 447]. Immunofluorescence staining of HPV-HTK and HPV-HFK raft sections was carried out using anti-K14 antibodies to identify mitotically active keratinocytes that had not yet differentiated (Figure 3-16). In untransfected HTK#3 raft culture, K14 was predictably detected in basal epithelial layers. In HPV16-HTK#2 and HPV18-HTK#2 raft culture,

K14 was detected in basal and lower epithelial layers, indicating HPV-induced delayed differentiation. K14 expression was more extensive in HPV16-HTK#3 and HPV18-HTK#3 raft culture and was detected from basal to the upper intermediate layers of HPV18-HTK#3 models as well as into upmost superficial epithelial layers of the HPV16-HTK#3 raft culture.

While mitotically active cells were confined to basal cells in untransfected HTK#3 raft culture as signified by K14, K14 positive cells were observed in intermediate layers of both HPV-HTK#2 and HPV-HTK#3 isogenic models, further demonstrating extended HPV E6 and E7-driven cellular proliferation of basal keratinocytes and delayed cellular differentiation. Further, H&E-stained HPV-HTK#3 raft culture appeared to show greater numbers of mitotic cells that had expanded to superficial cells (section 3.3.1). To verify this, immunofluorescence staining was carried out on multiple HTK and HFK raft sections with anti-phospho-histone H3 (Ser10) antibodies towards detection of cells undergoing mitosis (Figure 3-17a). In agreement, these data showed presence of mitotic (P-H3 positive) cells that extended into much higher superficial epithelial layers in HPV-HTK#3 raft culture when compared with the HPV-HTK#2 or HPV18-HFK models. The overall percentage of cells stained positive for phospho-histone H3 (P-H3) was calculated by dividing P-H3 positive nuclei by DAPI-positive nuclei, the data taken from five random fields of view, over multiple raft sections (Figure 3-17b). The data supported significantly higher mitotic cell counts in HPV-HTK#3 raft culture than in their HPV-HTK#2 and HPV-HFK counterparts, where p-values were no higher than 0.0052 in any comparison.

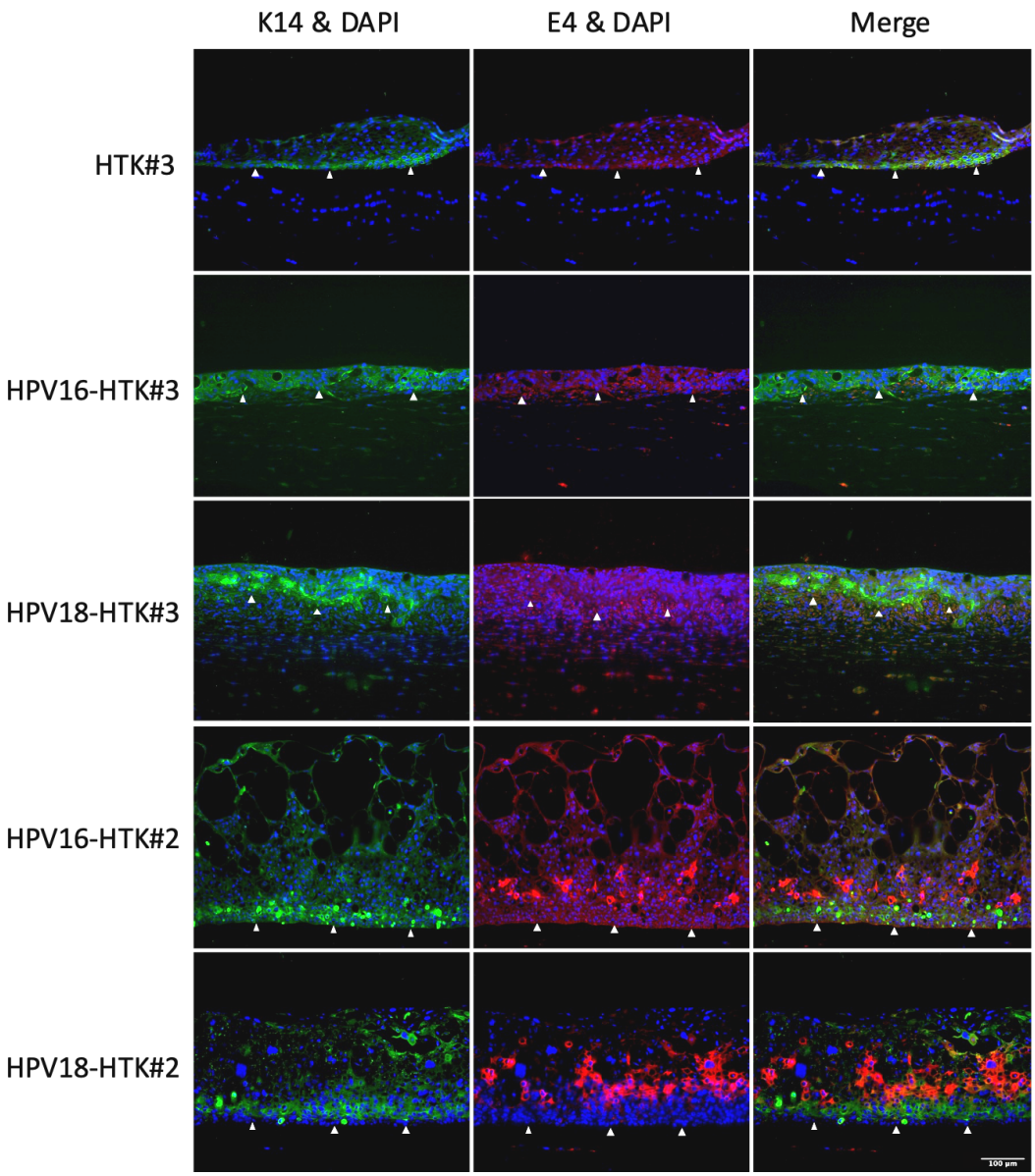
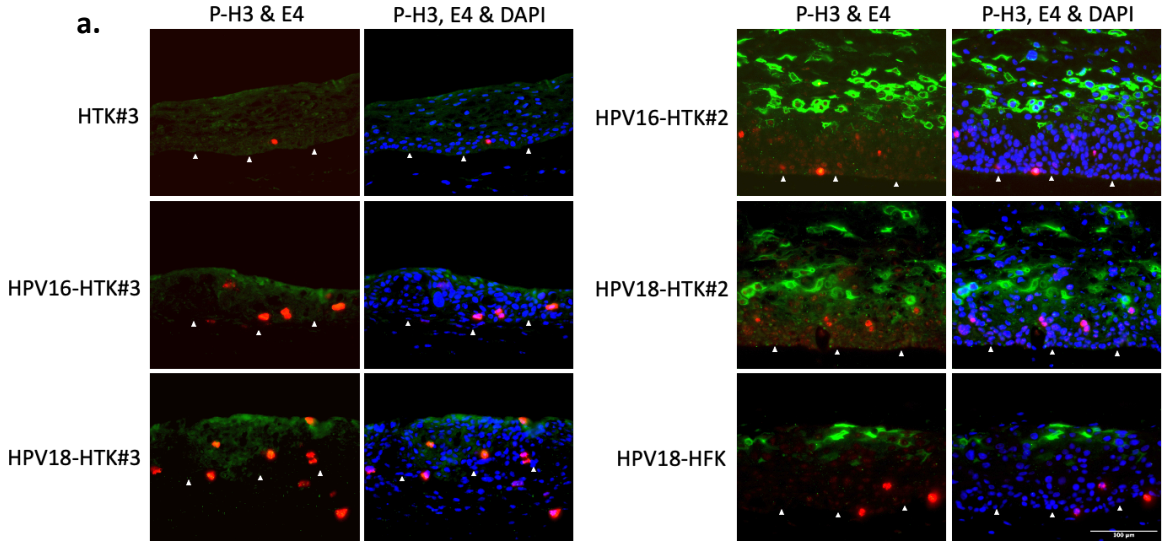


Figure 3-16: Immunofluorescence staining for K14 and E1^E4 in HTK raft culture

Raft cultures were immunostained with antibodies specific for keratin 14 (green), which is expressed in mitotically active basal cells, as well as E1^E4 (red) and DAPI (blue). White arrows indicate basal membrane. Scale bar in the bottom right image indicates 100 microns.



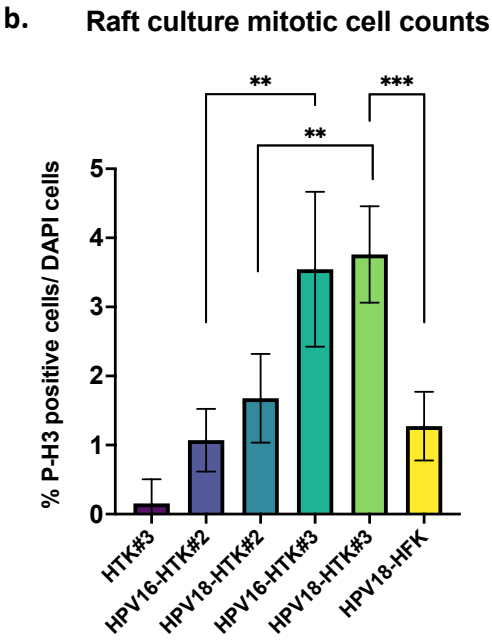


Figure 3-17: Mitotic cell counts of organotypic raft culture

a. Immunofluorescence staining of HTK and HFK organotypic raft culture, either untransfected or containing HPV16 or HPV18. Raft culture was formalin fixed, paraffin embedded and immunostained with antibodies specific for phospho-histone H3 (P-H3) (red), E1^E4 (green), and DAPI (blue). White arrows indicate basal membrane. Scale bar in bottom right image indicates 100 micrometres. **b.** Percentage of mitotic cells was calculated by number of P-H3 positive nuclei /total DAPI positive nuclei. Standard deviation was calculated from the mean and unpaired parametric T-tests were carried out to assess significance ($p < 0.05$).

3.4.2 IF microscopy of differentiation markers

During differentiation, expression of new keratin pairs is upregulated and expression of basal keratins such as K5/K14 is downregulated. In oral mucosal epithelia, such as stratified tonsil epithelia, differentiated cells produce K4/K13 and in cornified epithelium K1/K10 [274, 448]. As such, IF staining using anti-K13 and anti-K1/K10 antibodies was carried out to identify terminally differentiated keratinocytes in the HPV- raft cultures. As expected, K13 was detected in abundance in intermediate and superficial epithelial layers in HTK#3 raft culture and was not observed in basal keratinocytes (Figure 3-18a). Onset of K13 positive cells was delayed in all HPV-containing HTK raft culture when compared against the HPV-negative model but was even further delayed in HPV16- and HPV18-HTK#3 raft culture, where K13 was mainly detected in upper intermediate/superficial epithelial layers. Onset of K13 was shown to coincide with E1^E4 in both HPV16-HTK#2 and HPV18-HTK#2 raft cultures. However, pockets of dampened K13 positivity were observed in areas where E1^E4 had strongly accumulated. K13 was also more concentrated around the cell periphery in HPV-HTK#2 cells when compared with untransfected-HTK#3 and HPV-HTK#3 models.

As K13 is not expressed in cornifying epithelium, K13 was expectantly absent from HPV18-HFK raft culture. However, K1/K10 was detected in suprabasal HPV18-HFK epithelium; starting in lower spinous epithelium, and abundant production was observed in cornified epithelial layers (Figure 3-18b). Conversely, K1/K10 was not detected in donor HTK#3 or HPV-HTK#2 raft culture. In HPV-HTK#2 raft culture, K13 positive cells were first detected in intermediate E1^E4 positive regions. In HPV-HFK raft culture however, K10 was initially detected in lower spinous epithelium, which have been shown on numerous occasions not to exhibit E1^E4, which was

shown to be consistently detected in more upper spinous and granular epithelial layers (see Figure 3-7 –Figure 3-12). In this case, E1^E4-specific antibodies were unfortunately not added to these raft sections.

Involucrin is a well-established early marker of terminal differentiation in squamous epithelium. As with K13 and K1/K10, involucrin was detected in suprabasal epithelium of all assessed raft cultures but not in basal keratinocytes (Figure 3-19). Involucrin was only detected in upper superficial epithelial layers of both HPV-HTK#3 models, further demonstrating more limited differentiation in this HTK donor line.

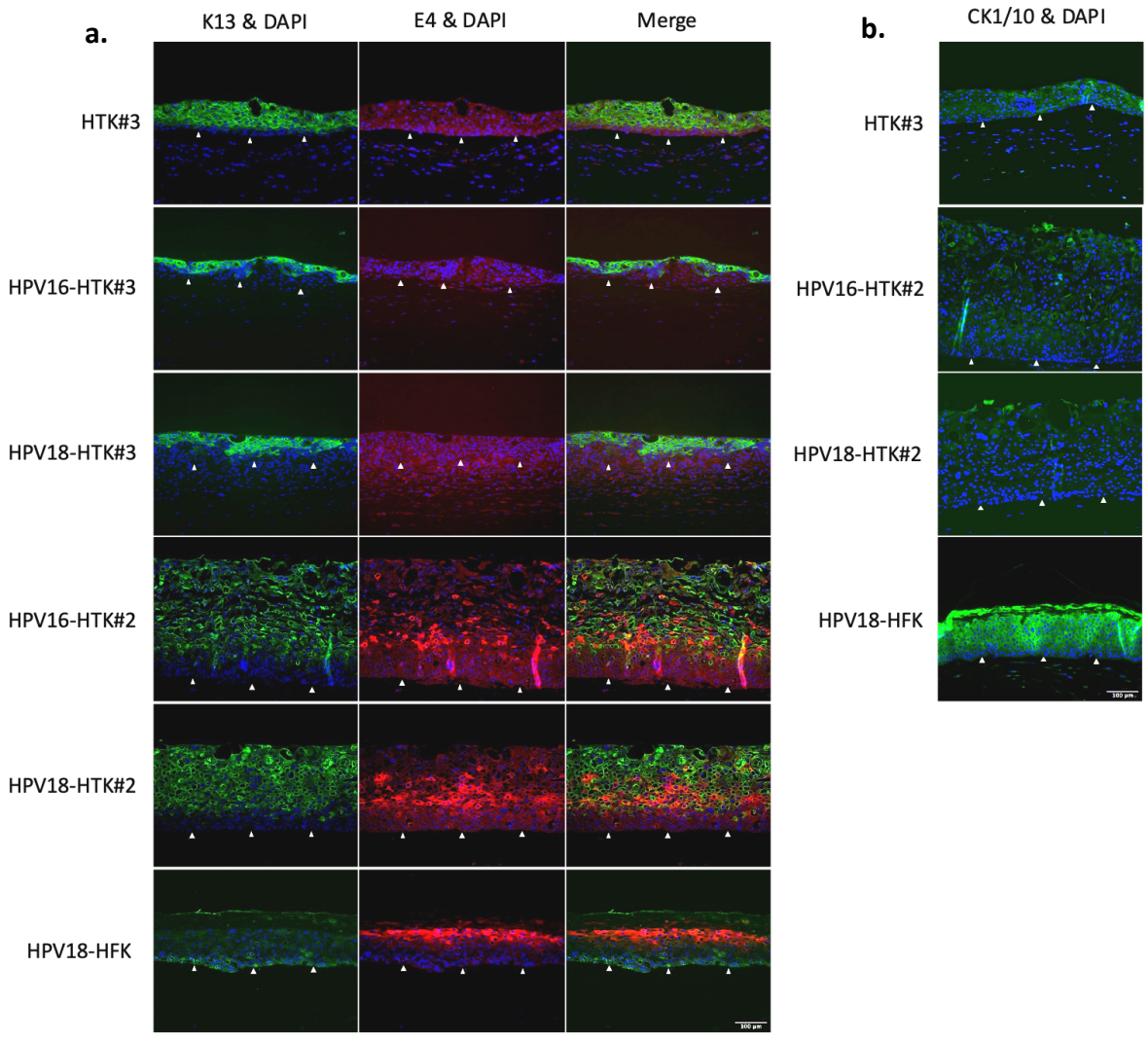


Figure 3-18: Immunofluorescence staining for K13/K10 and E1^E4 in HTK raft culture

a. Raft culture was immunostained with antibodies specific for K13 (green), a marker of differentiated cells in oral mucosal epithelium, as well as E1^E4 (red) and DAPI (blue). **b.** Raft culture immunostained with anti-K1/K10 antibodies (green), a differentiation marker in cutaneous epithelium, and DAPI (blue), but not E1^E4. White arrows indicate basal membrane. Scale bar in the bottom right image indicates 100 microns.

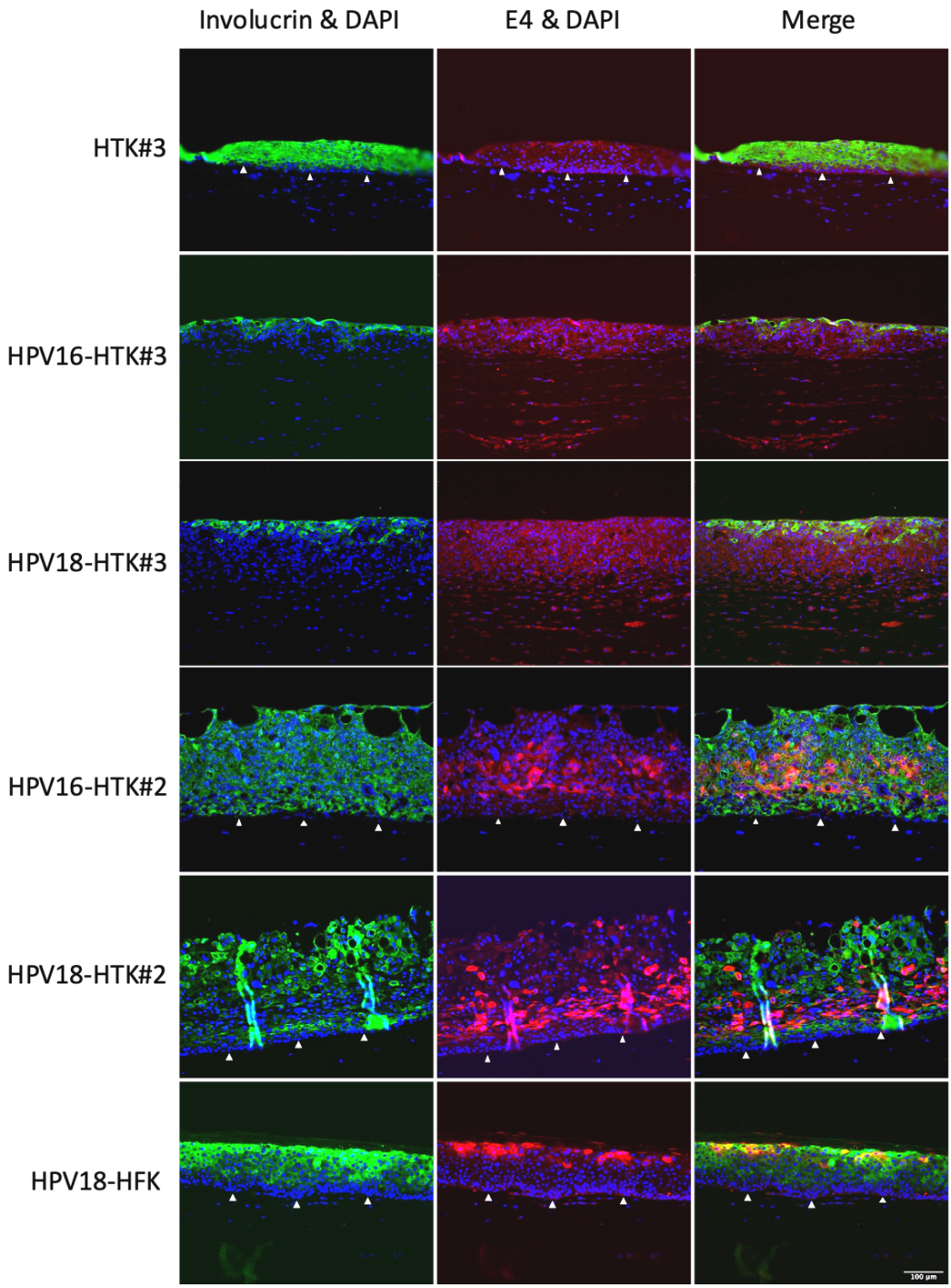


Figure 3-19: Immunofluorescence staining for involucrin and E1^E4 in HTK raft culture

Raft culture was immunostained with antibodies specific for involucrin (green), an early marker of terminal differentiation, as well as E1^E4 (red) and DAPI (blue). White arrows indicate basal membrane. However, the basement membrane was difficult to determine in both HPV-HTK#3 raft cultures. Scale bar in the bottom right image indicates 100 microns.

3.4.3 Organotypic tonsil models express biomarkers specific to primary site of HPV-positive tumours in the oropharynx

As discussed in the introduction (section 1.4.2), HPV-positive OPSCCs most frequently originate in tonsillar and base of tongue subsites, commonly in the tonsil crypts [34, 35, 56]. Unlike in normal tonsillar epithelia, the tonsil crypts produce keratins that are characteristic of simple epithelia, such as K7, K8, K18 and K19 [274, 275]. These keratins are also highly expressed in cervical squamocolumnar cells of the transformation zone, which are the primary site of HPV-driven cervical carcinogenesis [274, 276]. As such, expression of these keratins may represent anatomical subsites that are sensitive to HPV-driven disease progression.

Immunofluorescence staining of K8 and K7 was performed to identify whether HTK#2 or HTK#3 organotypic raft culture expressed any of these crypt specific biomarkers. Immunofluorescence staining for K8 in HTK#3, HPV16 and 18 HTK#2 and HPV18-HFK raft culture was initially performed by Peter Ray [269]. These studies showed limited K8 expression in the cytoplasm of occasional basal keratinocytes of the differentiated donor cells (Figure 3-20a). K8 was however detected in abundance in upper intermediate and superficial epithelial layers of HPV16 and HPV18-HTK#2 raft culture, as well as in basal and parabasal cells. In both HPV-HTK#3 raft culture, which showed more limited stratification, K8 was strongly expressed throughout basal and suprabasal epithelia. In contrast, K8 was undetected in the HPV18-HFK model.

K7 was detected mainly in upmost superficial layers of untransfected HTK#3 raft culture (Figure 3-20b). As in K8, K7 overexpression was observed in upper intermediate and superficial

layers of both HPV-HTK#2 raft cultures, largely following E1^E4 expression. K7 was also strongly expressed all throughout HPV16 and HPV18-HTK#3 raft culture but was only weakly detected in basal and superficial HPV-HFK epithelial layers. As with K13, pockets of dampened K7 and K8 positivity were observed in areas or cells that were strongly positive for E1^E4.

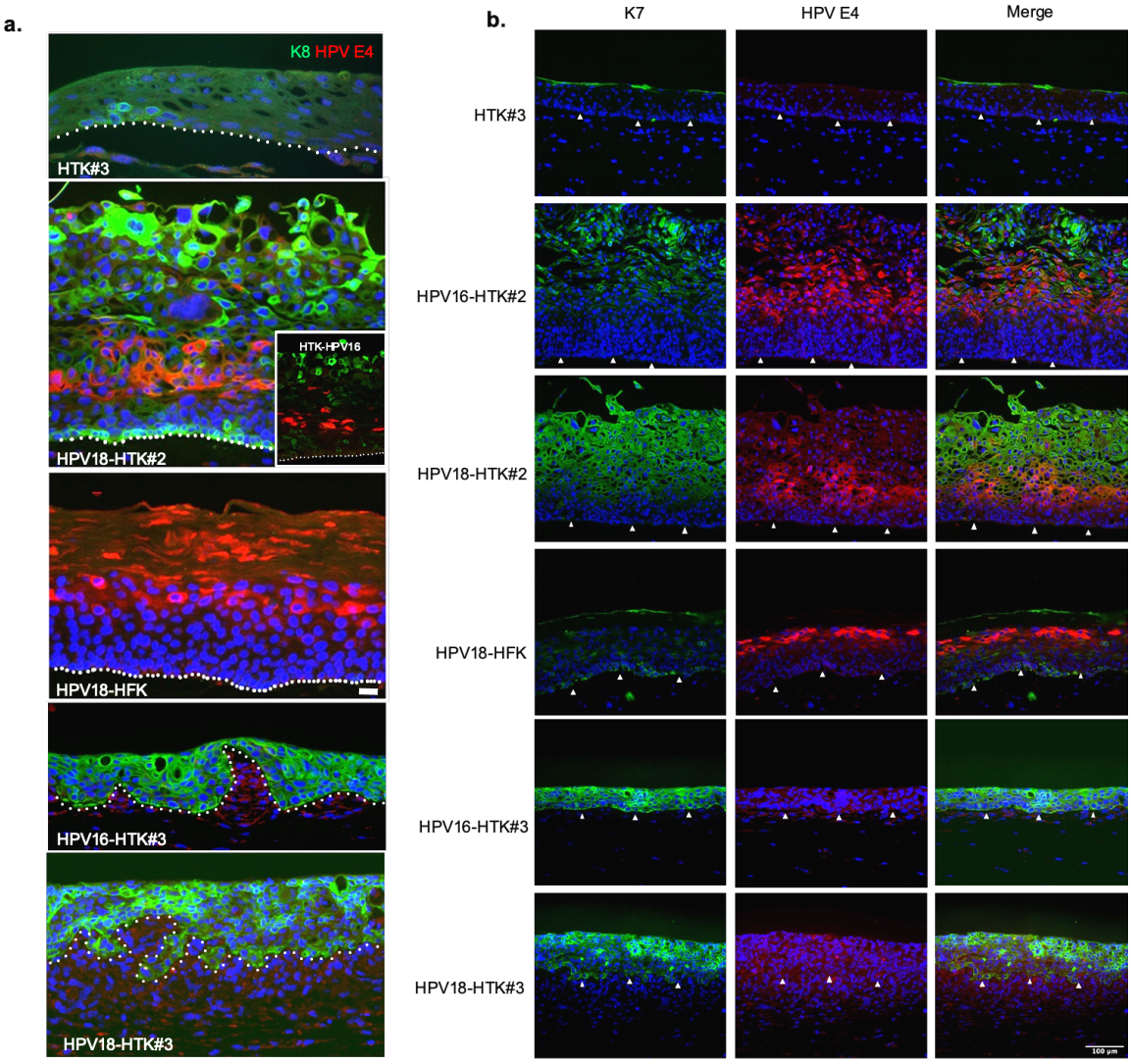


Figure 3-20: Immunofluorescence staining for K8/K7 and E1^E4 in HTK raft culture

a. Raft culture was immunostained with antibodies specific for K8 (green), E1^E4 (green) and DAPI (blue). **b.** Show's raft culture immunostained with K7 (green), E4 (red) and DAPI (blue). White dotted lines and arrows indicate basal membrane. Scale bar in the bottom right image indicates 100 microns. Here, K8 immunostaining of HTK#3, HPV18-HTK#2, HPV16-HTK#2 and HPV18-HFK raft culture was initially conducted by Peter Ray. All additional K8 and K7 immunostaining was carried out thereafter, and these data were collectively published in Roberts, Evans [269].

3.5 Discussion

To gain a better understanding of the HPV infectious cycle in the oropharynx, isogenic HPV16- and HPV18-genome maintaining tonsil keratinocyte models were developed to evaluate and compare viral life cycle organisation. By transfecting the same primary HTK donor cells with different virus types, we have established powerful model systems that allow for effective comparison of HPV16, which causes >90% HPV-OPSCC cases and HPV18, which is rarely found in OPSCC. This methodology effectively eliminates donor variability, allowing for more focused comparison between HPV16 and HPV18, to better understand how HPV16 dominates in cancers at this body site.

3.5.1 HPV genome maintenance in HPV-HTK models

Similar to previously established HPV-transfected primary tonsil keratinocyte lines [449, 450], our models have shown that both HPV16 and HPV18 were able to induce extended growth of primary HTKs in all three donors. Overall, HPV18 presented a greater tendency to integrate in primary HTK cells than HPV16, whereas HTK#2 cells maintained viral genomes as episomes. Here, Southern blotting revealed that the evolving physical status of HPV18 genomes was consistent between all three donors, where HPV18 DNA established as episomes, then integrated at between 40 – 70 population doublings.

As the donor sample size is small, it is not feasible to conclude whether HPV16 genomes have greater potential to maintain as episomes in HTK as HPV18. However, given the differences between HPV16 and HPV18 integration tendencies in the same donors, it would support the

hypothesis of differential episome maintenance of HPV16 and HPV18 genomes at this body site. Crucially, variations in episomal maintenance may be indicative of differences in biology between these two viruses in tonsil keratinocytes and may reveal HPV18 genomes as unable to persist long-term as episomes in tonsil keratinocytes. It is possible that the HPV-tonsil models could be recapitulating HPV16 and HPV18 genome maintenance and integration tendencies in the oropharynx. However, repeat transfections of the same donors and additional donors is required to validate the consistency of these behaviours.

While HPV16 DNA is integrated in around 60 – 80% of cervical cancers, HPV18 DNA has been shown to integrate in almost 100% of cases [391, 395, 402, 451, 452]. Integration of the HPV18 genome therefore appears to be a key requirement for HPV18-driven tumourigenesis, whereas HPV16 may have a greater potential to drive carcinogenesis as extrachromosomal genomes than HPV18 in the cervix. In HPV-positive OPSCCs, which mainly present with HPV16 DNA, integration is a less frequent event, occurring in only 39 - 43% of cancers, with remaining HPV-OPSCC cases, genomes are present mostly as episomal, as well as presenting with a mixed episomal and integrated status [396, 397]. However, given the rarity of HPV18 in OPSCC cancers (< 2.0%), it is difficult to speculate whether HPV18 has the same predisposition to integrate in the oropharynx as in the cervix.

Given that HPV18 is so rarely detected in OPSCCs, and that integration is such an important step in HPV18-driven cervical carcinogenesis, it was somewhat surprising to find that HPV18 was able to both induce extended growth of our primary HTK and also integrate its DNA so consistently within each donor. Interestingly, studies of integrated HPV-HNSCC biopsies

proposed that constitutive but not necessarily high levels of HPV16 E6/E7 can be sufficient for HPV-driven carcinogenesis at this body site [396, 398-400]. As such, it is speculated that integration at, or near to, specific genomic regions that provide infected cells with a selective advantage could be of particular importance in the oropharynx (see introductory section 1.4.6). It is therefore possible that HPV16 might favour specific integration events or integration at specific genomic hotspots that are more effective at promoting cancer progression than HPV18 in the tonsils.

Differences in HPV16 and HPV18 integration distributions have been evidenced in several studies of cervical cancers [402, 453-455]. Notably, HPV18 has been shown to integrate in proximity of c-*myc*, in 30% of assessed HPV18 cervical cancers, an event that has since been shown to drive overexpression of the proto-oncogene to provide a selective advantage [453, 454]. In contrast, no similar integration events around *c-myc* were observed in HPV16 cervical cancers, or in HPV-driven HNSCCs, potentially indicating this as a less important event in HPV16-mediated HNSCC carcinogenesis [410, 453]. It would be interesting to conduct further direct comparison of HPV16 and HPV18 integration events in cervical cancers using data from the cancer genome atlas and see whether HPV16-specific integration events are consistent with those observed in HNSCCs.

How HPV16 and HPV18 differentially regulate molecular targets and pathways may also play a role in how frequently these virus-types undergo integration. Interestingly, host-transcriptomic analyses of our HPV-HTK donor lines, as discussed in Chapter 5, suggest that HPV16 and HPV18 differentially regulate numerous pathways relating to the cellular immune

response. Here, HPV16 was shown to inhibit IFN-type I signalling and inflammation to a significantly greater extent than HPV18 (see sections 5.6 - 5.10). Crucially, stimulation of inflammatory and type I IFN responses have been shown to accelerate spontaneous HPV episomal loss, leading to an emergence of HPV-integrated clones [456-458]. Less stringent regulation of these pathways by HPV18 may therefore increase probability of viral genome integration.

A possible mechanism for this could involve the apolipoprotein B mRNA-editing catalytic polypeptide 3 (APOBEC3) proteins. These antiviral cytosine deaminases are theorised to increase the chance of viral integration in OPSCC by inducing double strand DNA breaks of the host genome and inducing hypermutation of HPV E2 [459-461]. Crucially, APOBEC3 proteins are interferon-inducible antiviral factors, so greater antiviral activities in HPV18-infected cells could potentially assist APOBEC3-associated integration [459]. To explore what role this is playing in the HTK lines, it would be interesting to first examine levels of type I IFN-inducible genes and APOBEC family members in our HPV-HTK lines, to identify possible changes in expression as the physical status of the viral genome changes from episomal to integrated. Possible knockdown studies of *APOBECA* and *APOBECB* could also be carried out to see whether these APOBEC proteins impede HPV genome integration over longitudinal passage.

3.5.2 HPV life cycle organisation in organotypic HPV-HTK raft culture models

Organotypic raft culture of the HPV-HTK models showed that both HPV16- and HPV18-HTK stratified into mucosal-like epithelia and that the HTK#2 models were able to support both HPV16 and HPV18 life cycles (see Figure 3-6 to Figure 3-9), something that has also been

previously shown in episomal HPV18-HTK donor #1 raft culture (Roberts, Rae and Marsh, unpublished data). Intriguing differences in viral life cycle organisation were observed between HPV-HTK#2 and HPV-HFK raft culture. Indeed, late viral life cycle events such as viral genome amplification and production of differentiation-dependent viral proteins, E1^E4 and L1, occurred at different stages of cellular differentiation and stratification in these two models.

Here, accumulation of these late viral proteins in both HPV16 and HPV18-HTK#2 raft culture initiated in lower epithelial layers and closer to the basal layer than in HPV-HFK rafts, coinciding with a switch in keratin profiles from undifferentiated (K13) (Figure 3-18) to differentiated keratins (K14) (Figure 3-16). This indicates a less delayed activation of the late viral promoter after cellular differentiation, as was observed in the HPV18-HFK models and has previously been reported in other HR-HPV-foreskin models and cervical LSIL biopsies [235, 236, 248, 432, 462]. These findings might reflect spatial differences in the switching of early to late viral late promoters, which are alternatively regulated by a range of transcription factors and transcriptional repressors, whose activities alter during cellular differentiation. This could be further attributed to discrepancies in differentiation programmes and keratin profiles between tonsil, foreskin, and cervical epithelia [129, 463-466].

These studies have revealed important distinctions in HPV biology between the mucosal tonsillar and cornifying foreskin epithelial models, and are further discussed in sections 3.5.3 through 3.5.7. However, when comparing HPV16 and HPV18 directly, few obvious differences were observed in life cycle organisation or in the regulation of the cellular differentiation

programmes in HTK. This suggests that discrepancies in HPV16/18 life cycle organisation are not likely playing a significant role in differences in disease potential between the two virus-types in the tonsils. However, some variable growth was observed between HPV16- and HPV18-HTK raft cultures. Here, HPV18 raft cultures frequently demonstrated greater epithelial thickening than donor equivalent HPV16 raft cultures, perhaps indicating that HPV18 promoted cellular proliferation to a greater extent than HPV16.

Indeed, RNA-seq profiles of HPV16- and HPV18-HTK donor lines showed that HPV18, compared to HPV16, supported greater activity of numerous pro-proliferative pathways, including KRAS, MTORC1, STAT3, STAT5, and TNF-α/NF-kB signalling pathways (see Chapter 5, section 5.6.1). This may have contributed to a greater proliferative capacity of HPV18-HTK, and the aforementioned differences in epithelial thickening as seen in Figure 3-6. Further analysis of the RNA-seq data is required to pinpoint specific differential targets of HPV18 and HPV16 that are modulating the activities of these pathways. However, it would first be important to perform growth assays to better evaluate and measure differences in growth capacity between the two virus-types in HTK.

3.5.3 HPV-HTK#3 raft culture did not support HPV life cycles and presented HSIL-like phenotypes

In contrast to HPV-HTK#2, donor #3 cells harbouring HPV16 and HPV18 showed reduced epithelial thickening, less obvious morphological differentiation, and did not support the infectious cycle, as signified by a consistent lack of E1^E4 and L1 expression. This was surprising given that Southern analysis showed that viral episomes were established and

maintained in this donor. Future *in situ* hybridization studies in these cell lines for HPV DNA are required to identify whether viral DNA amplification is taking place. However, a lack of late viral proteins indicates that this donor does not support the HPV16 and 18 life cycle.

Deregulation of viral oncoprotein expression along with identification of how far basal-like cells extend into suprabasal epithelia are important distinctions in the grading of HPV+ neoplasms [257, 258]. In high-grade squamous intraepithelial lesions (HSIL), E7-driven overexpression of MCM7 and p16 often extends to the full thickness of epithelia, delaying epithelial differentiation almost indefinitely [88, 215, 258, 259]. As such, HSIL do not support the HPV life cycle due to a lack of activation of the late viral promoter and in such cases HPV E1^E4 proteins are often isolated to just a few cells of upmost superficial layers of the epithelium and are sometimes completely absent [88]. Both HPV16-HTK#3 and HPV18-HTK#3 organotypic raft models exhibited all of these signature morphological characteristics indicative of higher-grade dysplasia, and this was a consistent observation between multiple rafting of HPV-HTK#3 cells.

These HSIL-like phenotypes are also seen *in vitro* raft culture of normal immortalized human keratinocytes (NIKS), which maintained episomal HPV16 genomes [258]. Like the HPV-HTK#3 raft culture, these studies demonstrated comparable increased expression of E7 into the upmost epithelial layers, abortive HPV16 life cycles, signified by a scarcity of E1^E4, and reduced responsiveness to cell-cell contact inhibition. Given that HPV-cancers usually take decades to develop, such early potential high-grade neoplastic changes may be met with some

scepticism. However, studies from the HPV vaccine trials have shown that HSIL can develop shortly after HPV infection [467-469].

Overall, the data herein has highlighted significant differences in the differentiation potential of HPV genome-containing cells between two tonsil donors (see Figure 3-6, and Figure 3-16 – 3-19). Notably, the epithelium of HPV-HTK#3 rafts also presented with poorly defined basal layers, where it was difficult to define the basal keratinocytes and distinguish them from fibroblasts in the collagen plug (Figure 3-6 – 3-12). There are two likely scenarios in this situation: (1) an influx of fibroblasts from the collagen plug enter the epithelium; or (2) keratinocytes invade the underlying dermal equivalent. Future IF-based experiments should be carried out on HPV-HTK#3 raft culture to better discern keratinocytes from fibroblasts by immunostaining for markers of the two cell types such as the E-cadherin (keratinocyte marker) and vimentin (fibroblast marker).

It is possible that embedding of basal cells and fibroblasts could have been due to increased motility and invasiveness of the HTK#3 cells following transfection with HPV16 and HPV18. Intriguingly, additional analysis of the RNA-sequencing data, as described in Chapter 5 (section 5.2), demonstrated significant differences in epithelial-mesenchymal transition (EMT) signatures between the two donors following transfection with HPV16 and HPV18. Here, gene set enrichment analysis (GSEA) was carried out to investigate biological pathways that were significantly, functionally altered between HPV-HTK#2 and HPV-HTK#3 cells, based on cumulative differences in expression of a defined gene set (see section 5.6). This analysis denoted significantly altered transcriptional activity of genes defining EMT in HPV-HTK#3 lines

than in HPV-HTK#2 lines, where the q-value (or adjusted p-value, see section 5.2) was 0.0026 (supplementary Figure S-1a). This entailed a number of important markers/regulators of differentiation, motility, and invasion, such as genes encoding growth factors, cell adhesion molecules, integrins and cytokines such as TGF-beta, CADM1, elastin, Gremlin1, ITGB3, ITGA2, IL8, and IL6 (supplementary Figure S-1b).

Although both donors were transfected with identical HPV16 and HPV18 recircularized genomes, contained episomal and not integrated HPV genomes, and were cultured in the same conditions, these organotypic raft culture studies potentially highlight HTK#3 cells as a donor more susceptible to HPV-driven high-grade dysplasia. Importantly, HTK#3 cells underwent abortive HPV16 and HPV18 infections and demonstrated more deregulated expression of pivotal EMT markers, which could, potentially, impair their ability to differentiate and may promote enhanced motility and invasiveness of the cells. However, given the small number of donors available in this study it is difficult to speculate the normality of this behaviour in HPV-containing tonsil tissues as a whole, again highlighting the need for additional donors and repeat experimentation.

If these HPV-HTK#3 lines are representative of higher-grade dysplasia, theses cell lines may represent a useful model to study differences between LSIL- and HSIL-like phenotypes following HPV infection in HTK. Additional investigation of cell growth and cell-cell contact inhibition, as well as cell motility and invasion, by the way of migration and invasion assays could be carried out to better characterise and confirm these HSIL-like phenotypes. It would also be useful to compare viral gene expression patterns between the two donor lines, to

better comprehend differences in viral transcriptome organisations that could be contributing to a more HSIL-like phenotype. Finally, in addition to the EMT pathway, more in-depth analysis of the RNA-seq data should also be carried out to investigate other alternatively regulated biological pathways and protein-protein interaction networks to pinpoint specific donor characteristics that could be contributing to these differing phenotypes.

3.5.4 Differential spatial expression of E1^E4, cyclin B1 and viral genome amplification between HPV-HTK and HPV-HFK raft culture models

In basal keratinocytes, HPV genomes are maintained as extrachromosomal episomes at low copy numbers. Here, viral DNA replication occurs in concordance with cellular DNA [201]. However, as keratinocytes differentiate, activation of the late viral promoter and vegetative viral genome amplification occurs. Following completion of cellular DNA replication in the S-phase, viral genome amplification instead takes place at the DNA damage checkpoint G2/M phase boundary, during HPV-induced G2 arrest [89, 220]. Here, HR-HPV E6, E7 and E1 utilise DNA damage response (DDR) ATM and ATR pathways to recruit DNA damage factors necessary for viral genome amplification (see introduction, section 1.3.4) [227, 239-241].

HR-HPV E1^E4, which is transcribed from spliced E2 and E1 open reading frames, has also been suggested to play an important supporting role in viral genome amplification [232, 233, 235, 236]. While these viral transcripts are located in early regions of HPV genomes, E1^E4 are only expressed at low levels from the early viral promoter and are more abundantly expressed in later stages of the viral life cycle upon differentiation-dependent activation of the late viral promoter [91, 97, 131, 470]. In anogenital tissues, E1^E4 overexpression occurs in conjunction

with HR-HPV genome amplification in upper-mid epithelial layers following cellular differentiation and subsequent activation of the late viral promoter, an event that has been consistently demonstrated in multiple high-risk HPV types, including HPV16 and HPV18 [89, 199, 230, 231, 233, 430].

Importantly, HR-HPV E1^E4 proteins exhibit an N-terminal leucine-rich motif that allows for direct interaction and modification of host keratin networks. An interaction that has been suggested to assist G2 arrest functions, viral genome amplification, capsid formation, and virion release from superficial epithelium [233, 236, 248-251]. Overexpression of HPV16 and HPV18 E1^E4 proteins has been shown to contribute to G2 arrest in cervical cancer lines and primary foreskin infection models [230, 231, 234, 237]. Here, upregulated production of E1^E4 briefly overlaps with E6 and E7 in upper-mid epithelial layers, where it's theorised that E1^E4-assisted G2 arrest inhibits E7-driven cell cycle re-entry, which cooperatively sustains cells in a replication competent G2-like state that is supportive of viral genome amplification [88, 232, 233, 471]. In this manner, E1^E4-induced cell cycle arrest functions have been demonstrated to enhance amplification efficiency in both HPV16 and HPV18 life cycles. However, this is not an essential requirement for viral genome amplification of either virus type, which continues in a delayed and less efficient fashion without E1^E4, mediated primarily by E7 [229, 233, 237].

In the HPV16 and HPV18-containing HTK#2 raft culture, almost complete regional separation was witnessed between cells that had accumulated E1^E4 and cells undergoing G2-arrest, denoted by cytoplasmic accumulation of the nuclear pro-mitotic mediator, cyclin B1 (Figure 3-12). A finding that was further confirmed in HPV18-HTK#1 raft culture, developed by the

Roberts group (data not shown). This was in direct contrast to the HPV-HFK raft culture models, as well as similar HPV18-HFK raft culture models of different donors that have been previously documented by the Roberts group [237], which showed regional colocalization of E1^E4 and cytoplasmic cyclin B1. Contrary to previous reports, exclusion of E1^E4 from cytoplasmic cyclin B1-positive regions suggests that E1^E4 was not contributing to G2-arrest functions in the tonsils. Importantly, these distinctions suggest an alternate mechanism of HR-HPV-induced G2 arrest altogether in the tonsils. It has been shown that E7 alone is able to stimulate cytoplasmic accumulation of cyclin B1 and CDK1 in differentiating cells, so it is possible that E7 could have been driving G2 arrest in HTK independently of E1^E4 [220, 229].

In situ hybridization studies conducted by the Roberts group highlighted similar spacial differences in viral genome amplification between HPV18-HTK and HPV18-HFK raft culture, showing that viral DNA accumulated much earlier in epithelial stratification of HTK than in HFK – in parabasal and lower intermediate epithelial layers prior to E1^E4 accumulation. Put together, this may indicate that viral genome amplification was being carried out in cells undergoing E1^E4-indapendent G2 arrest. However, it will be important to validate these theories in raft sections that have been concertedly probed and immunostained for viral DNA, E1^E4 and cyclin B1. Given that regional separation of cyclin B1 and E1^E4 was also observed in HPV16-HTK raft culture (Figure 3-12), it seems likely that viral genome amplification occurred prior to E1^E4 accumulation in these models too, although further *in situ* hybridization studies are required to verify this.

Altogether, these studies reveal substantial distinctions in viral life cycle organisation between HTK and HFK models and highlight important questions as to how HPV16 and HPV18 are directing G2 arrest and viral genome amplification in the tonsils. Previous studies stipulate vegetative viral genome amplification as being a differentiation-dependent life cycle event, occurring primarily in cells undergoing G2 arrest and following activation of the late viral promoter [89, 199, 230, 231, 233, 430]. Yet, in HPV-HTK raft culture, G2 arrest and viral genome amplification appeared to be carried out preceding increased production of E1^E4, which is a reliable marker of late viral promoter activation [89, 472]. This might also suggest that amplification was occurring prior to cellular differentiation and activation of the viral late promoter in HTK. More direct comparison of viral life cycle, cell cycle and cellular differentiation markers used in these studies should allow for more definitive characterisation of the HPV life cycle organisation in the tonsils. In addition, it will also be important to examine spatial organisation of other viral proteins that are required for viral genome amplification, namely E1 and E2. Like E1^E4, these viral proteins are upregulated in a differentiation-dependent manner and are essential for viral genome amplification in G2 arrested cells [197, 198]. As such, it will be important to verify the point at which production of these proteins increases during HTK stratification.

3.5.5 DNA damage responses were primarily located in concordance with E1^E4

Upon epithelial differentiation, HR-HPV utilize the ATM and ATR DDR pathways for vegetative viral genome amplification during G2-arrest [227]. This event is signified by differentiation-dependent activation of central DDR response elements, including phosphorylated forms of ATM, NBS1, CHK2, and H2AX [227, 432, 444, 445]. Here, γH2AX, which is activated in both

ATM and ATR pathways, has been shown to associate directly with HPV genomes and accumulate within viral replication centres [227, 242].

While G2 arrest and viral genome amplification was observed in lower-epithelial regions of HPV16/18-HTK#2 raft culture, γH2AX was surprisingly detected primarily in upper epithelial layers of HPV16/18-HTK#2 raft cultures in concordance with E1^E4 accumulation, similar to HPV18-HFK and previously published anogenital models (Figure 3-14) [473, 474]. This could indicate that viral genome amplification may not taking place in DDR-activated cells in the HTK models. However, the findings presented here are still somewhat preliminary and will require further investigation. Additional examination of γH2AX and other important DDR factors such as phosphorylated CHK1, CHK2, ATM, ATR, BRCA1 and Rad51 should be carried out and directly compared with markers of cell cycle arrest, viral genome amplification and differentiation. ATM and CHK2 could be assessed by IF of HPV-HTK raft culture, as demonstrated by Moody and Laimins [227].

3.5.6 HPV E1^E4-mediated keratin reorganisation in HTK raft culture

In the epithelium, keratins mainly exist as either soluble monomers or insoluble keratin filaments, which can be solubilised, reorganised, or collapsed to support cellular functions such as migration, mitosis, and apoptosis [475-477]. This process is largely regulated via phosphorylation or association with solubility factors, which stimulate ubiquitination and proteasomal degradation [478, 479]. Following differentiation, E1^E4 has been shown to impede solubility factor functions and hyper-phosphorylate keratins, generally leading to reorganisation of keratin to the cell periphery, as well as depletion of keratins via

ubiquitination [236, 248, 251, 472]. E1^E4-mediated reorganisation of the keratin network is highly regulated throughout epithelial differentiation and is thought to assist keratin-cross linking, virus release and transmission [480]. Aspects of this re-organisation were also observed in the HPV-HTK raft culture models, where positivity of K7, K8 and K13 dampened specifically in cells and pockets of cells that were undergoing productive viral infection, as demonstrated by strong E1^E4-positivity (Figure 3-18 and Figure 3-20). In addition, K13 had noticeably accumulated around the cellular periphery of HPV-HTK#2 cells when compared with untransfected-HTK raft culture, where K13 was more evenly distributed in the cytoplasm.

3.5.7 HPV-mediated epidermal reprogramming in primary HTK models

One ongoing difficulty when modelling HPV-associated diseases *in vitro* is that HPV-driven neoplasia is often limited to specific sites of specialized epithelia that are sensitive to transformation. In the cervix for example, HPV-driven transformation ordinarily takes place in the cervical transformation zone [171]. In the tonsils, HPV-driven carcinogenesis consistently arise in the tonsil crypts [274]. Current organotypic raft culture models however most closely recapitulate normal stratified epithelia such as the ectocervix or tonsil surface, which are less sensitive to HPV-driven transformation.

In the tonsils, normal surface epithelia and crypt epithelia undergo distinct differentiation programmes, evidenced by alternative keratin profiles. Two distinctive biomarkers of the crypts are the type II keratin pairs K7/K19 and K8/K18. These simple keratins are normally highly expressed in differentiated cells of crypt epithelium but are detected at low levels in normal surface epithelium [274, 275]. Further, these keratins are also highly expressed in

cervical squamocolumnar cells of the transformation zone and are commonly overexpressed in both HPV-OPSCC and cervical cancers [276, 279, 481]. While HPV infections of the foreskin are so common [482], HPV-related malignancies at this body site are extremely rare. It is therefore possible that alternative keratin programmes between non-cornifying mucosal, K7/K19- and K8/K18-expressing epithelia and cornifying epithelia like the foreskin, which does not express these keratins, could be a significant distinction in this contrasting susceptibility to HPV-driven transformation.

Intriguingly, immunostaining of our HTK organotypic rafts revealed a transformation in keratin expression profiles following transfection with either HPV16 or HPV18, demonstrating increased production of K7 and K8 in upper intermediate and superficial layers, largely coinciding with productive viral life cycle stages (Figure 3-20) [269]. Importantly, no similar induction of these keratins was observed in HPV18-HFK, suggesting that this was an HTK-specific event. Together, these data indicate that HPV16 and HPV18 can transform tonsil keratin profiles to imitate a more crypt-like phenotype which may be beneficial for the virus. As such, this could also highlight the HTK models as a suitable model of HPV-infection in the crypts.

K7/K19 and K8/K18 are likely important elements in HPV-mediated carcinogenesis for several reasons: (1) they are specifically produced in epithelia predisposed to HPV-associated transformation [483, 484], (2) they are commonly overexpressed in HPV-positive cancers [279, 481], and (3) they directly associate with E7 and E1^E4 to assist viral life cycle functions [248, 275, 278, 279]. As such, HPV-induced epidermal reprogramming of surface tonsil epithelia to

express K7/K8 and partnering keratins could be key processes for viral life cycle success and potentially HPV-driven disease progression in the tonsils. Future K7 and K8 knockout experiments would be a useful method of evaluating HPV life cycle dependency on altered expression of these keratinocytes in the tonsils.

3.6 Conclusions

Altogether, the studies detailed herein have highlighted clear differences in the organization of the HPV productive cycle between tonsillar and anogenital infection *in vitro* and show that E1^E4 is not required for induction of G2-arrest and possibly viral genome amplification in HTK. As well, HPV16/18 were demonstrated to alter the keratin profiles of the tonsils to promote expression of simple keratins, K7 and K8, possibly to promote a productive viral life cycle. However, when comparing HPV16-HTK and HPV18-HTK raft culture, very few differences were observed in life cycle organisation or in regulation of the cellular differentiation programmes. As such, viral and host transcriptome analysis was next performed to evaluate differences in virus and host transcription profiles between the isogenic virus-type HTK models (chapters 4 and 5).

3.7 Future objectives:

These findings raise a number of questions into the biology of HR-HPV in the tonsils. How is HPV inducing G2 arrest and supporting high production of viral genomic DNA in the tonsils prior to E1^E4 accumulation? Is viral genome amplification being carried out in undifferentiated epithelial layers? What role, if any, is the DNA damage response playing in viral genome amplification?

It will next be important to further characterise and compare spatial and temporal onset of HPV life cycle events that were shown to be differentially organised between HPV-HTK and HPV-HFK models. Given the regional separation of E1^E4 and cytoplasmic cyclin B1, further *in situ hybridization* studies should be carried out in combination with IF to ascertain whether productive viral genome amplification is being carried out prior to both cellular differentiation and activation of the late viral promoter in HPV-HTK raft cultures. In addition, IF-and Western blot-based investigations of E1 and E2, which respectively carry out and regulate viral genome replication, may further our understanding of how viral genome amplification is taking place. This could be examined using organotypic raft and methylcellulose differentiated HTK. Knockdown of E1^E4 in the HPV-HTK models may also allow for a better understanding of the role of this protein in the viral life cycle overall.

Most importantly, development of additional episomal-maintaining HPV-HTK donor models will be critical to assure overall validity of these findings. It will be crucial to develop additional raft culture models that can also support the viral life cycle, as with the HPV16/18-HTK#2 and HPV18-HTK#1 lines. This will allow for more definitive confirmation of reported differences in episomal maintenance between HPV16 and HPV18 in HTK, differences in viral life cycle organisation and epidermal reprogramming between tonsil and anogenital raft culture systems. In addition to HPV18-HFK models, establishment of HPV16-HFK and HPV16- and HPV18-human cervical keratinocyte (HCK) models could also allow for more complete comparison between these alternative body sites.

Chapter 4: High-risk HPV transcriptome in primary tonsil keratinocyte virus life cycle models

4.1 Introduction

In Chapter 3, investigations of our HTK raft culture models assessed a potential divergence in HPV16 and HPV18 biology that might contribute to the contrasting differences in disease burden between the two virus-types in the oropharynx (described in introduction, section 1.4.3) [31, 32]. Here, HTK#2 was shown to support both HPV16 and HPV18 life cycles, however little variation in viral life cycle organisation was observed between the two virus-types overall. As such, it seems unlikely that discrepancies in HPV life cycle organisation are greatly contributing to differences in disease potential in the tonsils.

We therefore propose that differences underlying viral persistence and carcinogenesis in the oropharynx could be manifested at the molecular level. To date, little study into the molecular biology of HR-HPV has taken place in pre-cancerous tonsil tissues or primary tonsil infection models, greatly limiting our understanding of early transformation events at this distinct anatomical site. As such, work detailed here aimed to both investigate and compare viral transcriptome organisation of HPV16 and HPV18 in the tonsils using our primary HPV episome-containing HTK lines. We hypothesised that differences in HPV16 and HPV18 transcription profiles in primary HTK would correspond directly with the differences in disease burden at this body site.

Prior establishment of detailed HPV transcription maps has been crucial to our understanding of HPV gene expression and its immensely complex transcriptional and post-transcriptional regulation (see introduction, section 1.2). Deciphering viral gene expression patterns is important to understand how the virus manipulates the host throughout its life cycle. However, until very recently, HPV transcriptome mapping has not allowed for accurate quantification of mature viral mRNA transcripts [98, 485].

The Parish and Roberts' groups have been amongst the first to pilot Nanopore as a more comprehensive method of HPV transcriptome profiling [98]. This third generation (long-read) sequencing technique allows for direct characterisation of unamplified, full-length poly(A)$^+$ selected mRNA by detecting changes in electrical currents as transcripts are passed through microscopic protein channels (nanopores) [486, 487]. Unlike conventional short-read sequencing techniques, this allows for characterisation of native, unfragmented HPV transcripts and importantly supports quantification of different viral mRNA variants. However, in the absence of RNA enrichment and fragmentation steps, Nanopore lacks the same accuracy and read-depth of short-read sequencing techniques, such as Illumina-based RNA-sequencing (RNA-seq), Given these differences we also utilised Illumina RNA-seq alongside the Nanopore to globally map host HTK transcriptomes (Chapter 5; detailed in section 5.2) [486-488].

4.2 HPV-HTK mRNA preparation and Nanopore-sequencing

To investigate HPV16/18 transcriptome organisation in the primary HTK, early cell culture passages, of less than 40 population doublings (PD), were selected which contained viral

genomes as episomes, as determined from Southern blot data (section 3.2, Figure 3-4). Total RNA was extracted from all six untransfected and HPV16- and HPV18-containing HTK#2 and HTK#3 cell cultures; HTK cultures were grown in monolayer on a feeder layer of irradiated 3T3 J2-fibroblasts, which were supplemented with EGF to conserve cells in a more undifferentiated- and basal-like state (methods section 2.1.2) [489, 490]. To confirm this, analysis of the Illumina RNA-seq data was carried out to compare expression of basal cell-specific keratins (*KRT5* and *KRT14*), and differentiation-dependent keratins (*KRT4* and *KRT13*) (supplementary Table S-1). Here, expression of the basal-specific keratins was consistently and significantly higher than the differentiation-dependent keratins, showing an average log_2 fold increase of 2.8, confirming that cells were in a more undifferentiated-like state opposed to differentiated.

RNA was extracted from untransfected and HPV16- and HPV18-containing HTK#2 and HTK#3 cell cultures (methods section 2.3). Sample preparation, quality control, and sequencing steps are described in methods section 2.3.5. Poly A+ selected RNA (mRNA) was Nanopore sequenced at around 1 to 2 million reads. Sequence alignment, performed by Dr Boris Noyvert, was carried out using minimap2, where mRNA fragments were aligned to HPV16 (K02718) and HPV18 (AY262282.1) genomes, as well as to the human genome (GRCh37) to identify and characterise any virus-host fusion transcripts.

Final read depths, along with numbers of HPV and virus-host fusion transcripts are overviewed in Table 4-1, following normalisation to the sequencing depth of each HPV-HTK line, presented as reads per million (RPM). Expression of episome-derived transcripts occurred at near equivalent levels between HPV16- and HPV18-containing cell lines of the

same donor (Table 4-1). However, expression of virus-host fusion transcripts varied between cell lines. HPV18-HTK#3 and HPV16-HTK#2 presented a greater number of viral reads (650.7 and 509.5 RPM) than HPV16-HTK#3 and HPV18-HTK#2 (182.2 and 363.4 RPM) overall. This was largely due to differences in production of virus-host fusion transcripts, where the former two HTK lines contained much greater numbers of virus-host fusion transcripts, which respectively made up 61.4% and 34.2% of total viral transcripts.

In addition, a number of HPV mRNA with early termination sites were identified, where transcripts were shown to terminate upstream of the pA^E signal (Figure 4-5). Early terminating transcripts were present in all four HPV-HTK lines but were particularly abundant in HPV18-HTK#3 cells (80.6 RPM). Lower numbers were observed in HPV16-HTK#2 (8.3 RPM), HPV18-HTK#2 (5.5 RPM), and HPV16-HTK#3 lines (3.8 RPM) (Table 4-1). Unlike episome-derived transcripts, which ended at the pA^E, the viral mRNA segment of virus-host fusion transcripts commonly ended around nucleotide 880 in HPV16-HTK cells, or around nucleotides 929 or 3607 in HPV18-HTK cells, likely making use of polyadenylation sites located in the human mRNA region (data not shown). Equivalent endpoints were observed in the early terminating transcripts, possibly indicating these as additional integrant-derived RNA. However, early termination could also be caused by RNA degradation or low fidelity polymerase, which could result in unprocessed polyadenylated RNA.

The HPV integration events have been explored further in section 4.10. However, given that these studies aimed to assess HPV16/18 transcriptome organisation in the tonsils, as in a normal viral infection, all other transcriptomic analyses exclusively involved transcripts derived from episomal HPV DNA.

Table 4-1: Overview of Nanopore viral reads

Cell line	Total number of detected transcripts, millions	Total HPV transcripts (RPM)	HPV episome-derived transcripts (RPM)	HPV episome-derived transcripts: % of total HPV transcripts	Virus-host fusion transcripts (RPM)	Virus-host fusion transcripts: % of total HPV transcripts	Early terminating transcripts (RPM)	Early terminating transcripts: % of total HPV transcripts
HTK#2	1.55	-	-	-	-	-	-	-
HPV16-HTK#2	0.97	509.5	326.9	64.2%	170.2	34.2%	8.3	1.6%
HPV18-HTK#2	1.08	363.4	341.3	93.9%	16.6	4.6%	5.5	1.5%
HTK#3	1.43	-	-	-	-	-	-	-
HPV16-HTK#3	1.59	182.2	162.0	88.9%	16.4	9.0%	3.8	2.1%
HPV18-HTK#3	2.10	650.7	170.8	26.2%	399.3	61.4%	80.6	12.4%

4.3 HPV16 and HPV18 transcription organisation in primary tonsil keratinocytes

During the HPV life cycle, viral gene expression is stringently and adjustably regulated in concert with the epithelial differentiation programme, ensuring that specific viral proteins are present as required for each stage of epithelial development. This is largely carried out via differentiation-dependent regulation of viral promoter regions, polyadenylation, and splice sites (see introductory section 1.2). Analysis of the Nanopore data was conducted to identify potential discrepancies in activities of these transcriptional elements between HPV16- and HPV18-HTK lines.

4.3.1 HPV16 and HPV18 promoter usage in primary tonsil keratinocytes

To compare HPV16 and HPV18 promoter usage, transcriptional start sites (TSS) of all episome-derived HPV transcripts were analysed. HPV16/18 TSS clustered into five distinct promoter regions, which have been described in Table 4-2. The exact location of each promoter region within the HPV16 and HPV18 genome was based on transcriptional start site clustering in the HTK lines, as demonstrated in Figure 4-1, in addition to definitions made previously in the literature [90, 91, 96-98, 491].

In the first instance, violin plots were prepared to visualise and appreciate TSS distributions within each viral genome (Figure 4-1). Expectedly, given that cell cultures were grown in an undifferentiated-like state, the highest concentration of TSS were detected in the pE of both HPV16 and HPV18 genomes, in both HTK donors. The most distinct differences in TSS distributions centred around the pE8 promoter, where a much greater number of TSS were observed at this promoter region in HPV18 than in HPV16, possibly denoting differential E8^E2

expression between the isogenic HTK lines. In contrast to the other three HPV16/18-HTK lines, HPV18-HTK#2 presented higher usage of the p3000 promoter, indicating donor- and virus type-specific differences in its regulation.

4.3.2 Quantitative comparison of HPV16 and HPV18 promoter usage in primary tonsil keratinocytes

Numerical differences in viral promoter usage are summarised in Figure 4-2. These data confirmed the pE as the most consistently active promoter in all four HPV-HTK lines. Proportionally higher usage of the pE was shown in the HPV16-HTK donor lines than in donor-equivalent HPV18-HTK counterparts, where pE-derived transcripts accounted for 66.7% of HPV16-HTK#2 transcripts and 39.7% of HPV18-HTK#2 transcripts (Figure 4-2a). Here, normalised expression of HPV16- and HPV18-HTK#2 pE-derived transcripts was respectively 218.7 RPM and 134.7 RPM (Figure 4-2b). pE activity was also proportionately higher in HPV16-HTK#3 cells than in HPV18-HTK#3 cells, where 75.5% (132.4 RPM) and 61.4% (97.3 RPM) of transcripts were shown to start in pE regions of HPV16- and HPV18-HTK#3 respectively.

Little variability in viral late promoter (pL) or p500/p520 usage was observed between donor equivalent HPV16- and HPV18-containing HTK. However, distinctions in pE8 activity were further exemplified. Here, HPV18 transcripts beginning at the pE8 respectively accounted for 13.3% (45.2 RPM) and 10.2% (16.2 RPM) of total HTK#2 and HTK#3 transcripts (Figure 4-2a and b). In contrast, very few pE8-initiating TSS were detected in HPV16-HTK#2 (0.6%; 2.1 RPM) or HPV16-HTK#3 (1.1%; 1.9 RPM). This suggests much higher pE8 activity in HPV18-HTK than HPV16-HTK, likely indicating higher transcription of the viral transcriptional repressor, E8^E2.

Donor-specific differences in HPV promoter usage were evident between both HTK#2 and HTK#3 transfection models. In comparison to the isogenic HPV-HTK#3 lines, HPV-HTK#2 lines presented consistently higher activity of each viral promoter (excluding HPV16 pE8), suggesting higher transcriptional activity in HTK#2 cells than HTK#3 cells (Figure 4-2b). Given that viral gene expression is restricted at low-level in infected basal cells and increases during cellular stratification and differentiation, it is possible that these finding could reflect disparities in differentiation capabilities and EMT signatures observed between the two donors following HPV transfection (Chapter 3; Figure 3-2, Figure 3-10, Figure 3-17 - Figure 3-20, and supplementary Figure S-1 [492-496]).

4.3.3 HPV16 and HPV18 polyadenylation in primary tonsil keratinocytes

Usage of early (pA^E) and late (pA^L) polyadenylation signals are strictly regulated in a differentiation-dependent manner, providing an important distinction between production of early and late viral transcripts [103]. Here, early viral transcripts are commonly polyadenylated at the pA^E signal located at nucleotide 4215 in the HPV16 genome and 4235 in the HPV18 genome. Late viral transcripts are instead polyadenylated at pA^L sites located at nucleotides 7321 (HPV16) and 7278 (HPV18) (see introduction, section **Error! Reference source not found.**; Figure 1-3).

To assess HPV16 and HPV18 polyadenylation activity in the HTK lines, additional violin plots were constructed to map the 3' ends of each episome-derived mRNA transcript. Transcript termination was observed exclusively at the early polyadenylation (pA^E) signal in all four HPV-

HTK lines (Figure 4-3a), as expected in undifferentiated keratinocytes. Termination sites varied slightly at the pA^E between HPV16-HTK#2 and HPV16-HTK#3. In HTK#2 cells, HPV16 transcripts ended at nucleotide 4212 only (Figure 4-3b). In HTK#3 cells however, HPV16 transcripts ended at nucleotides 4217, 4225 and 4232. In contrast, termination site locality appeared consistent between HPV18-HTK#2 and HPV18-HTK#3, where transcripts most frequently terminated at nucleotide 4273 of the HPV18 genome. This suggests differential regulation of the HPV16 pA^E between HTK#2 and HTK#3. However, this does not present any obvious biological significance.

In summary, initial assessment of the Nanopore data revealed high activity of the pE and pA^E in each assessed HPV-HTK line, as is expected in HPV-transfected undifferentiated keratinocytes. When comparing HPV16 and HPV18 promoter usage, these studies demonstrated variable pE and pE8 promoter usage between the two virus-types, indicating differing promoter regulation and gene expression profiles between HPV16- and HPV18-HTK.

Table 4-2: HPV16 and HPV18 promoter positions

HPV16 viral promoters	HPV16 promoter region: nucleotide position on genome	HPV18 viral promoters	HPV18 promoter region: nucleotide position on genome
Major early viral promoter (pE)	1 - 280	Major early viral promoter (pE)	1 - 350
P500	390 - 579	P520	351 - 680
Major late viral promoter (pL)	580 - 890	Major late viral promoter (pL)	681 - 920
E8 promoter (pE8)	1100 - 1400	E8 promoter (pE8)	1100 - 1400
p3500	2800 - 3900	p3000	2800 - 4050

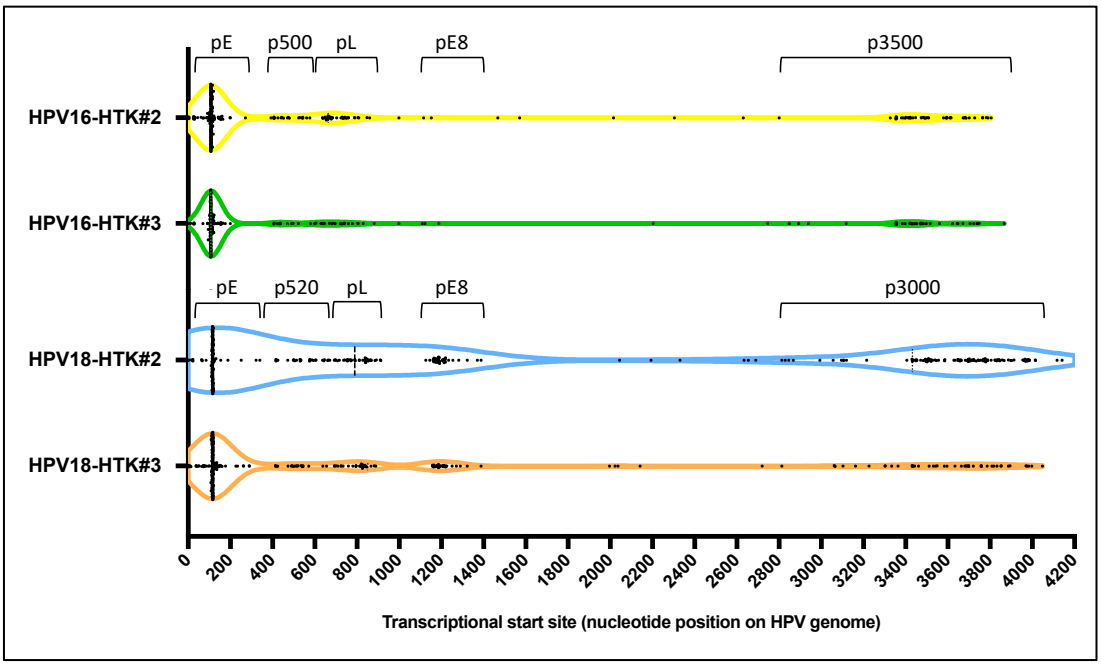


Figure 4-1: HPV16 and HPV18-HTK transcriptional start site maps.

Violin plots were used to map the 5' ends, or transcriptional start sites (TSS) of HPV16 and HPV18 mRNA transcripts. These are similar to box plots but allow for better visualisation of the full distribution of numeric data using density curves, or kernel density estimates, which are a non-parametric method of estimating probability density of a variable. It is important to note that the maximal height of a density curve does not specify a direct number of TSS but instead indicates a higher probability of clustering. The frequency distribution is instead shown by the number of points at a specified location. The median value is indicated with black dashed lines, and the 1st and 3rd quartiles are signified with

small, dotted lines. Black dots indicate individual transcriptional start sites. Promoter regions are labelled above their corresponding nucleotide regions.

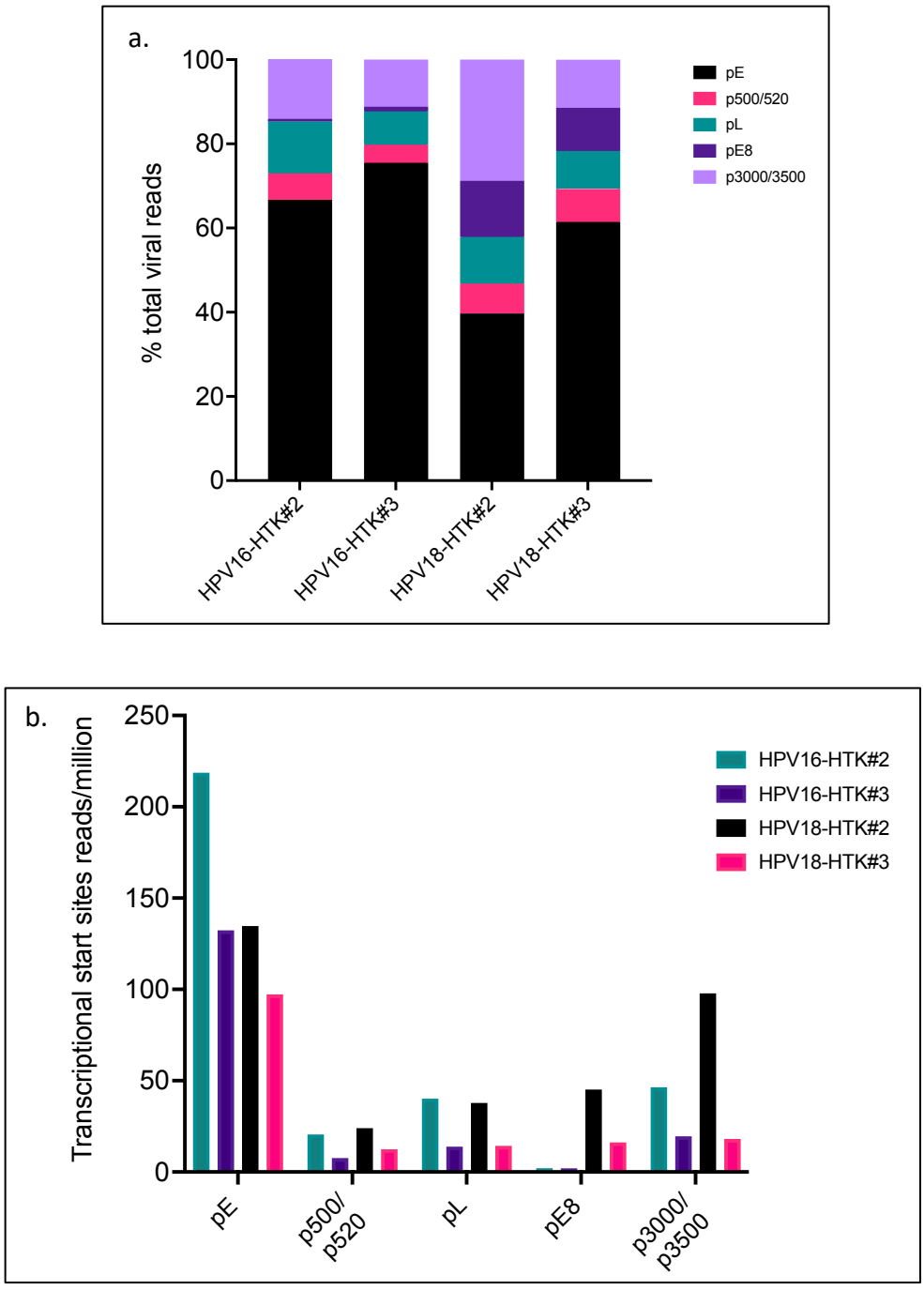


Figure 4-2: Comparison of HPV16 and HPV18 promoter usage in HPV-HTK lines

a. Stacked bar graph demonstrates proportional promoter usage as a percentage of the total number of episome-derived HPV transcripts in each sample. **b**. Shows the number of detected transcripts starting at each promoter region in reads/million (RPM) of total Nanopore reads of each sample, including both viral and human mRNAs.

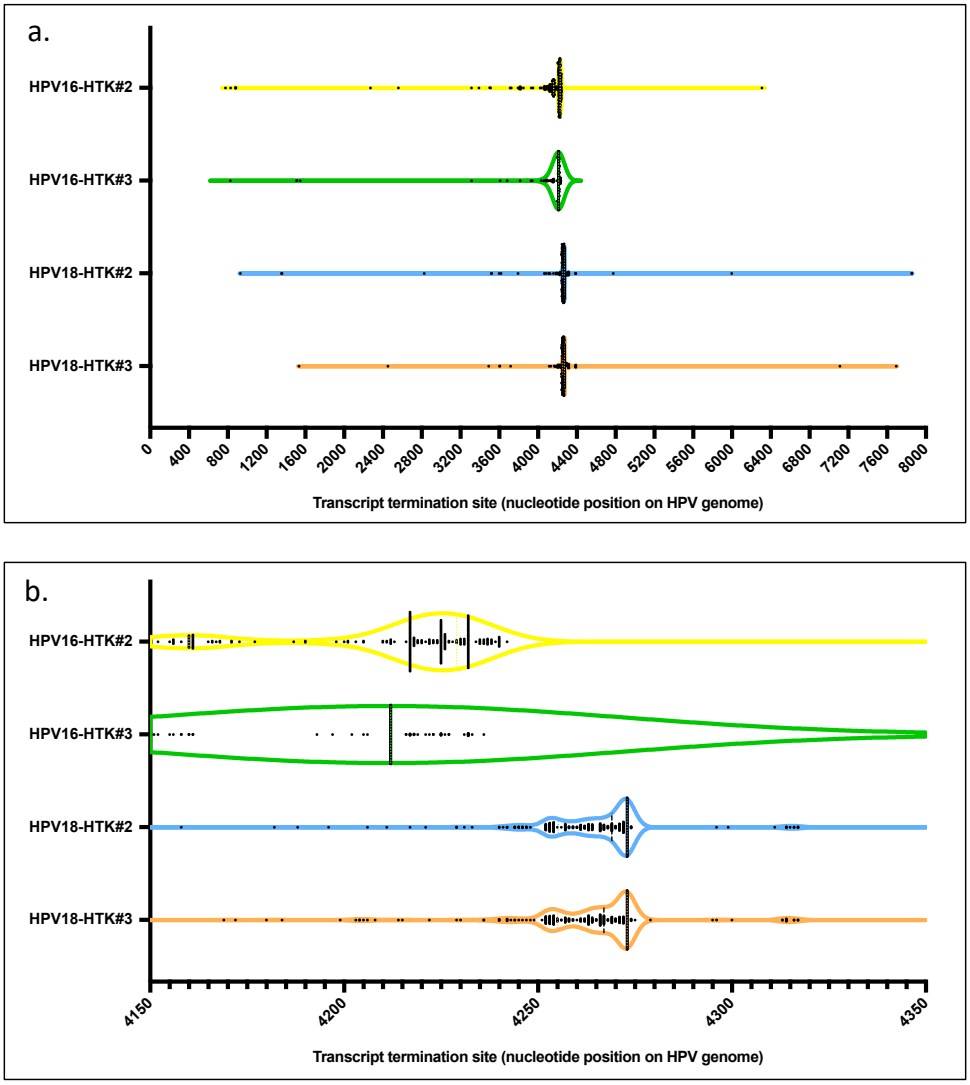


Figure 4-3: Transcriptional termination sites of HPV16 and HPV18-HTK lines

Violin plots were used to map the 3' ends, or transcriptional termination sites of MinION sequenced HPV16, and HPV18 mRNA transcripts. **a**. Shows all detected HPV16 and HPV18-HTK termination sites. **b**. Shows a magnified region of all transcript termination sites detected at the early polyadenylation (pA^E) region.

4.4 HPV16 and HPV18 alternative splicing in primary tonsil keratinocytes

In addition to TSS and termination sites, HPV mRNA species are further defined by alternative splicing. This is another highly regulated process that permits a single polycistronic pre-mRNA HPV transcript to encode multiple functionally distinct protein isoforms (reviewed in introductory chapter, sections 1.2.2 - 1.2.3). Short-read (Illumina) RNA-seq was performed on HPV16- and HPV18-HTK lines as described in section 5.2. Sequenced mRNAs were aligned to HPV16 (K02718) and HPV18 (AY262282.1) genomes to examine alternative splicing within the HPV16- and HPV18-HTK lines. Samples were sequenced at ~70 - 80 million reads per sample (Table 5-1), allowing for more accurate assessment of viral splice events than Nanopore, where sequencing depths reached ~1 - 2 million reads per sample (Table 4-1). However, unlike Nanopore, this method does not allow us to effectively visualise full-length transcripts. Sashimi plots were constructed using the Integrated Genomics Viewer (IGV) (http://software.broadinstitute.org/software/igv/), which was used to visualise HPV read coverage and splice junctions (Figure 4-4). However, it is important to note that any virus-host fusion transcripts were not excluded from these analyses, likely inflating total viral reads and splice activities in HPV18-HTK#3 and HPV16-HTK#2 models.

Alternative splicing occurs via the differential usage of splice donor (SD) sites and splice acceptor (SA) sites. SD and SA sites are depicted in Figure 4-4. HPV16 and HPV18 genomes contain several distinct SD and SA sites, leading to differential transcript diversity between the two virus types [497]. In the E6/E7 open reading frames (ORF), unlike HPV18, HPV16 genomes contain two additional SA sites (SA526 and SA742), which generate distinct spliced E6 (E6*)

variants such as HPV16 E6*II and E6*X [120, 498] (Figure 4-4a). In contrast, HPV18 contain two unique and rarely used SA sites (SA3440 and SA3506) and splice donor sites (SD3165 and SD3284) in the E2 ORF (Figure 4-4b).

In the HPV16-HTK lines, four distinct E6*-encoding splice events were identified, E6*I (SD233^SA416), E6*II (SD233^SA526), E6*III (SD233^SA3357), and E6*X (SD226^SA742). In comparison, just two HPV18 E6* spliced variants were observed, E6*I (SD233^SA416) and E6*II (SD233^SA3434). Expectedly, the highest number of splice junctions in both virus-type lines typically occurred between SD226^SA409 (HPV16) and SD233^SA416 (HPV18). This splice event takes place in the E6 ORF and is required for expression of E6*I mRNA, which primarily encodes E7 oncoproteins, and to a lesser extent E6*I proteins [114]. These were shown to make up between 85-93% of all HPV16 E6* transcripts and almost all (98-99%) of HPV18 E6* transcripts.

In the HPV16-HTK lines, splicing from SD226 to the SA526 and SA742 sites are respectively required for production of lesser defined E6 splice proteins, E6*II, and E6*X [120, 498]. HPV16 E6*II transcripts accounted for ~5.4% of E6* transcripts and very few E6*X transcripts were observed (<0.1%). Transcripts spliced between SD233^SA3357 encode E6*III proteins and made up ~9% of HPV16 E6* transcripts (Figure 4-4a). While the function of many of these E6* isoforms (E6*II, E6*III, and E6*X) is not currently well understood, high expression of E6*II and E6*III isoforms has been linked to more severe oropharyngeal and cervical pathogenesis [499-501]. The unique production of these spliced variants by HPV16 may therefore contribute to

varying biological activities between the two virus types. As such, further research is required to elucidate the structures and functions of these viral isoforms.

Further differences in E8^E2 transcription were indicated between HPV16-HTK and HPV18-HTK. HPV16 and HPV18 E8^E2 transcripts were identified via use of the SD1302^SA3357 and SD1357^SA3434 splice junctions, a splice event that is used almost exclusively for production of E8^E2 transcripts [502]. Here, just 20 HPV16-HTK#2 E8^E2-specific splice events compared with 190 HPV18-HTK#2, and 4 HPV16-HTK#3 E8^E2-specific splice events compared with 68 HPV18-HTK#3. When normalised to the total number of reads, this showed a $\log_2$ fold change of 3.3 (HTK#2) and 4 (HTK#3).

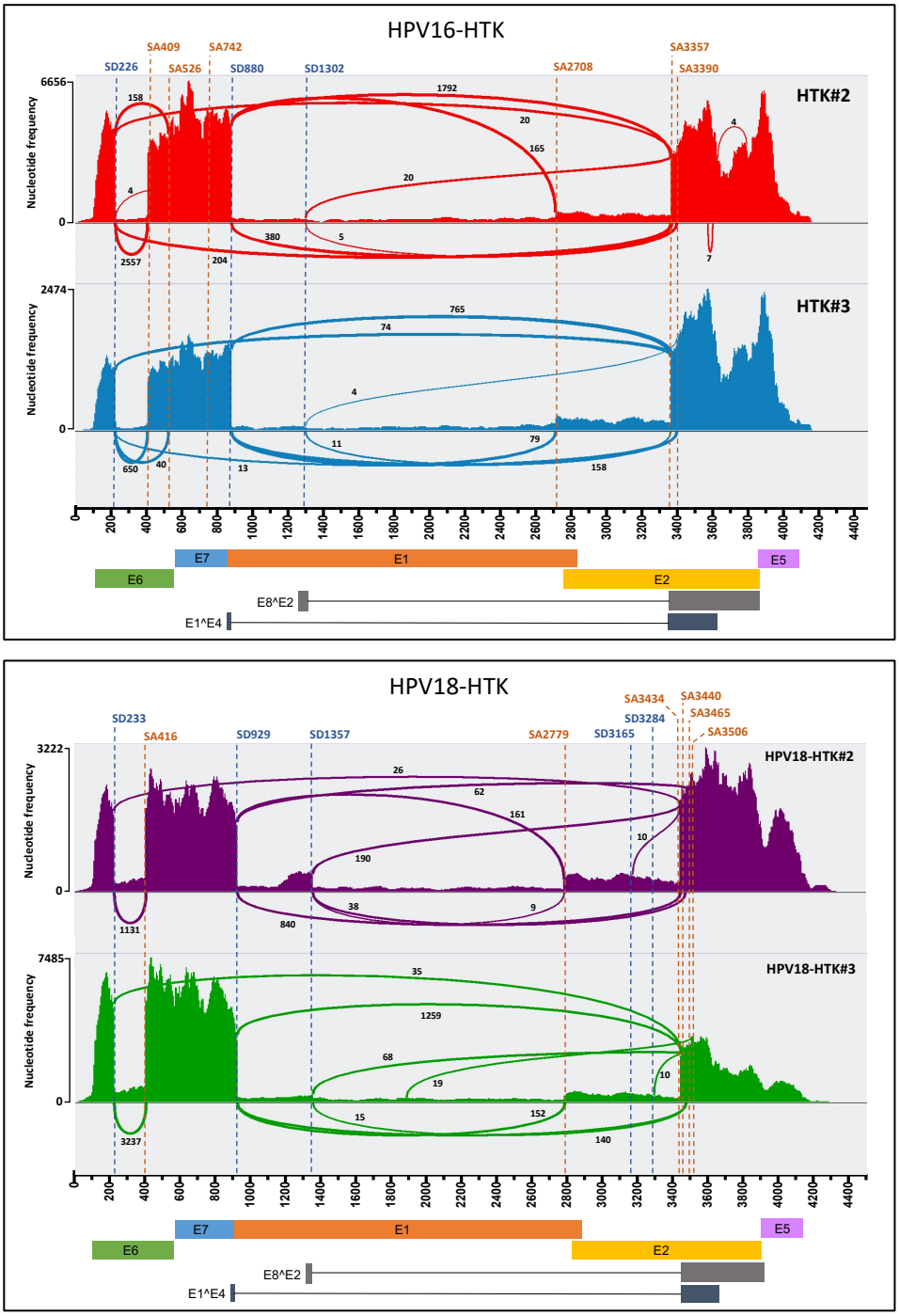


Figure 4-4: Overview of HPV16 and HPV18 splicing in human tonsil keratinocytes

Sashimi plots were generated from Illumina-based RNA-seq data to visualise HPV16 and HPV18 splicing in HTK#2 and HTK#3 donors. Splice donor (SD) sites are marked with blue lines and splice acceptor sites (SA) with orange lines. Splice junctions are presented as arches that connect SD and SA sites, where numbers indicate splice junction reads. HPV16 and HPV18 open reading frames are shown below each plot along with a scale indicating nucleotide position.

4.5 Transcriptome mapping of primary HPV16- and HPV18-containing tonsil keratinocytes

Comprehensive analysis of the Nanopore data was carried out to examine and compare HPV16 and HPV18 transcription profiles in the HTK lines. Here, TSS, splice junctions and termination sites were used to determine the species of each episome-derived mRNA (Figure 4-5). The relative abundance of each mRNA species is presented as RPM, following normalisation to the sequencing depth of each HPV-HTK line. A total of 20 and 22 distinct HPV16 and HPV18 mRNA species were identified.

4.5.1 Identification of novel HPV16 and HPV18 mRNA species in primary HTK

Seven novel mRNA species were identified in the HPV16-HTK lines (Figure 4-5a). This included transcript 20, which was expressed at 12.1 and 5.4 RPM in HPV16-HTK#2 and HPV16-HTK#3 cells respectively. These unspliced mRNA were shown to initiate from the p3500 promoter and were predicted to encode E2C and E5, exhibiting similar organisation to HPV18 transcript 23 (Figure 4-5b), a transcript that has been previously characterised by Toots, Mannik [96]. Very little difference in expression of these transcripts were observed HPV16- and HPV18-containing donor lines (log_2 fold change <0.2), suggesting similar levels of E2C between the two virus type cell lines.

The other six novel transcripts (9, 11, 12, 14, 16, 19) shared comparable transcriptional arrangements to previously mapped HPV16 mRNA species but were shown to use the SA3390 site as opposed to SA3357 [90, 109]. This SA3390 consensus was previously described by

Milligan, Veerapraditsin [90] for an mRNA species equivalent to transcript 10 (Figure 4-5a). Splicing at SA3390 is expected to trim the first 11 amino acids from the 5' end of the E2 DNA-binding domain (E2C), which is conserved within E4 and E8 open reading frames [90]. As such, use of this splice acceptor would likely alter the E2C reading frame. However, the existence or function of any resulting protein isoforms has yet to be validated. Alternatively, this splice event could potentially inhibit production of E1^E4, E8^E2, or E2C, instead favouring production of E5.

A potentially comparable HPV18 splice acceptor, SA3465, which similarly trims the E2C reading frame was identified in our HPV18-HTK lines (Figure 4-5b; transcripts 12 – 17). These transcripts have been well characterised in HPV18-HFK models and have been suggested to encode E5 [91, 98]. However, it is possible that these mRNA species could also encode E4, E8, and E2C spliced variants. As with HPV16 SA3390-spliced transcripts, further verification is required.

Of the HPV16 SA3390-spliced transcripts, transcript 10, was the most highly expressed and was detected at 14.4 and 7.6 RPM in HTK#2 and HTK#3 cells respectively (Figure 4-5a). This was predicted to encode E6*I, E7, E5, in addition to a possible E1^E4 splice variant, termed E1^E4s. The novel transcripts 9, 11, 14, and 19 were only expressed at very low level, usually ≤1 RPM, suggesting little biological relevance. However, transcripts 12 and 16, which potentially encode E5, and E1^E4s/E5 were expressed at 3.1 and 8.3 RPM in HPV16-HTK#2 cells but ≤1 RPM in HPV16-HTK#3 cells.

In the HPV18-HTK lines, the majority of identified mRNA species have previously been defined in HPV18-HFK or in HPV18-containing human osteosarcoma-derived (U2OS) cells [96, 98]. Just one novel mRNA species (transcript 14) was identified in HPV18-HTK#2 cells only. This started at the pE and was shown to splice between SD233 and SA3465 and was predicted to encode E5. However, this transcript was only expressed at low level, at 3.7 RPM and was otherwise undetected in HPV18-HTK#3 cells.

4.5.2 Differential expression of E6- and E7-encoding transcripts between HPV16- and HPV18-HTK

Expectedly, the most abundantly expressed variant, in both HPV16 and HPV18-containing HTK lines, was transcript 3 (Figure 4-5). This mRNA species is predicted to encode for E6*I, E7, E1^E4 and E5 in both virus types and has also been shown to be the most abundantly expressed viral mRNA species in anogenital models, cancers and HPV-OPSCCs [90, 98, 114, 417, 503]. The normalised expression of transcript 3 was higher in the HPV16-HTK lines than in donor equivalent HPV18-HTK lines, which could suggest higher translation of E6*I/E7. This accounted for 133.1 and 74.8 RPM in HPV16-HTK#2 and HPV16-HTK#3, and 95.9 and 60.6 RPM in HPV18-HTK#2 and HPV18-HTK#3 (Figure 4-5).

To better compare relative HPV expression profiles between the isogenic HPV16-HTK and HPV18-HTK donor lines, the aggregated coding potential of each distinct ORF was determined by calculating the total number of viral reads which mapped to a characteristic viral promoter and splice junction (Table 4-3). This was normalised to RPM, then $\log_2$ fold change ratios were used to measure differences in expression between donor-equivalent HPV16- and HPV18-

containing HTK lines. Finally, the proportional abundance of a viral gene was calculated by dividing the aggregated coding potential by the total number of episome-derived HPV transcripts.

Alternative splicing between SD226^SA409 (HPV16) and SD233^SA416 (HPV18) is a major splicing event that disrupts production of full-length E6-encoding mRNAs, instead producing E6*I mRNA isoforms. These isoforms are primary precursors of E7 oncoproteins, which are efficiently translated from the E7 AUG. In addition, translation from the E6 AUG, which is less efficient than the E7 AUG, produces E6*I proteins, which can act as functional E6 inhibitors [112, 114, 162, 504]. As such, splicing between SD226^SA409 and SD233^SA416 was used to define transcripts predicted to encode E6*I/E7, consisting of HPV16 transcripts 1, 3, 7, and 10 and HPV18 transcripts 1, 3, 11, 13, and 19 (Figure 4-5). Higher normalised expression of these E6*I/E7-encoding transcripts was observed in the HPV16-HTK lines than donor equivalent HPV18-HTK lines, showing a $\log_2$ fold change of between 0.32 and 0.52 (Table 4-3). Higher proportional expression of these E6*I isoforms was also observed in HPV16-HTK lines, representing 50.5% and 55.3% of all episome-derived HPV transcripts, 13 – 20% higher than in respective HPV18-HTK lines (Table 4-3).

HPV16 E7 are also likely translated from E6*II mRNA isoforms, which use the SA226^SD256 junction (4, 8, and 11), in addition to transcripts beginning at the p500 promoter, which lies just upstream of the E7 ORF (transcripts 13 and 14) (Figure 4-5a) [501]. Likewise, HPV18 E7-encoding mRNAs also include transcripts starting at the p520 promoter (transcripts 5, 15, and 20) (Figure 4-5b). While HPV16 and HPV18 unspliced polycistronic E6/E7 mRNAs (transcripts

2 and 9 [HPV16] and 4 and 18 [HPV18]) also contain the E7 ORF, translation of E7 is considered to be very inefficient due to a potent upstream translational block imposed by the E6 AUG (see introduction, section 1.2.2). Consequently, unspliced-E6/E7 transcripts primarily produce full-length E6 [112, 114, 161].

Excluding unspliced-E6/E7 transcripts, the aggregated coding potential of all E7-encoding transcripts was respectively 0.52 and 0.32 log2 fold higher in HPV16-HTK#2 and HPV16-HTK#3 lines than in donor-equivalent HPV18-HTK lines (Table 4-3). This made up between 60-64% of all HPV16 episome-derived reads and 41-43% of HPV18-HTK reads, altogether indicating higher production of E7 in HPV16-HTK than in HPV18-HTK. In contrast, expression of unspliced E6-encoding transcripts was collectively higher in the HPV18-HTK lines (transcripts 4, 12, 18, and 25) than in donor equivalent HPV16-HTK lines (transcripts 2, 9, and 22), showing log_2 fold expressional increases of 0.84 (HTK#2) and 2.03 (HTK#3) (Table 4-3). Put together, these findings could suggest higher splicing of the E6 ORF in HPV16-HTK lines than in HPV18-HTK lines, potentially inducing higher expression of E7 and lower expression of E6.

The HPV16 E6*III and HPV18 E6*II transcripts (5 and 2) exhibit comparable splicing arrangements, where splice junctions were observed between SD226^SA3357 (HPV16) and SD233^SA3434 (HPV18), which in both cases spliced the first part of E6 ORF to the E2C ORF (Figure 4-5). However, sizeable differences in expression of these transcripts were observed between the isogenic HTK lines, which showed comparatively higher expression in HPV16-HTK lines than donor-equivalent HPV18-HTK lines (log_2 fold change = 2.33 and 1.52) (Table 4-3). These made up around 9.5% of total episome-derived HPV16 transcripts and between 1.6 –

3.1% of HPV18 transcripts. However, given that the exact structure and function of these spliced isoforms is not currently well characterised, it is difficult to speculate what functional differences this could provide between the two virus types overall.

4.5.3 HPV18 support higher expression of E8^E2 than HPV16 in primary HTK

Previous analysis of TSS and alternative splicing highlighted potential differential expression of transcripts encoding E8^E2 between HPV16-HTK and HPV18-HTK lines (Figure 4-2 and Figure 4-4). Transcription of E8^E2 uniquely originates at the pE8, and transcripts are characteristically spliced from SD1302^SA3357 (HPV16) or SD1357^SA3434 (HPV18). E8^E2 proteins are highly potent regulators of viral genome replication and transcription from the pE and are derived from a short E8 fragment, contained within the E1 ORF, which are spliced to the E2C ORF (see introductory section 1.2.3) [146, 147]. E8^E2-encoding transcripts accounted for ≤0.8% of total episome-derived HPV16-HTK reads (Figure 4-5;Table 4-3). In contrast, E8^E2-encoding transcripts made up between 5.3 – 8.5% of viral HPV18-HTK reads. These distinct deviations in expression suggest that HPV16 and HPV18 differentially regulate viral transcription and DNA replication in primary HTK.

4.5.4 Expression of E2-encoding mRNAs in HPV16- and HPV18-containing HTK

Overall, expression of transcripts containing the E2 ORF was collectively higher in the HPV16-HTK lines (transcripts 1, 7, 8, and 17) than in donor equivalent HPV18-HTK lines (transcripts 1, and 18 - 21) (Figure 4-5), showing a $\log_2$ fold difference of between 2.37 (HTK#2) and 0.67 (HTK#3) (Table 4-3). This may suggest somewhat higher production of E2 in the HPV16-HTK

lines, however it is important to note that almost all of these HPV16 and HPV18 transcripts initiated from the early viral promoter. Crucially, studies by Alloul and Sherman [505] and Zheng, Cui [506] have shown that E2 is poorly translated from early viral transcripts due to the presence of upstream AUG and strong Kozak sequences, which impede scanning ribosomes from accessing start codon sequences at the E2 ORF. Instead, these studies show that E2 is most efficiently translated from transcripts originating from the late viral promoter [506], however little to no expression of any such E2-encoding transcripts was witnessed in any of our HPV-HTK lines (Figure 4-5), perhaps indicating little in the way of E2 translation. As such, it is difficult to speculate on differences in E2 protein levels between the two virus-types in HTK based on these data. This needs to be determined by western blotting.

4.5.5 HPV-HTK transcriptome mapping conclusions

When comparing HPV16 and HPV18 gene expression profiles in HTK, the key findings of these transcriptional analyses are as follows: (1) HPV16-HTK supported higher expression of E7-encoding mRNAs but lower expression of E6-encoding mRNAs than donor-equivalent HPV18-HTK lines, and (2) HPV18-HTK demonstrated much greater expression of E8^E2-encoding transcripts than HPV16-HTK. Finally, no conclusive differences in E1, E1^E4, or E5 expression was observed between the two virus-type lines. However, it was also difficult to compare translational tendencies of these viral genes due to variable and diverse AUG positioning, which is an important determinant of translation efficiency (see introduction, section 1.2.2).

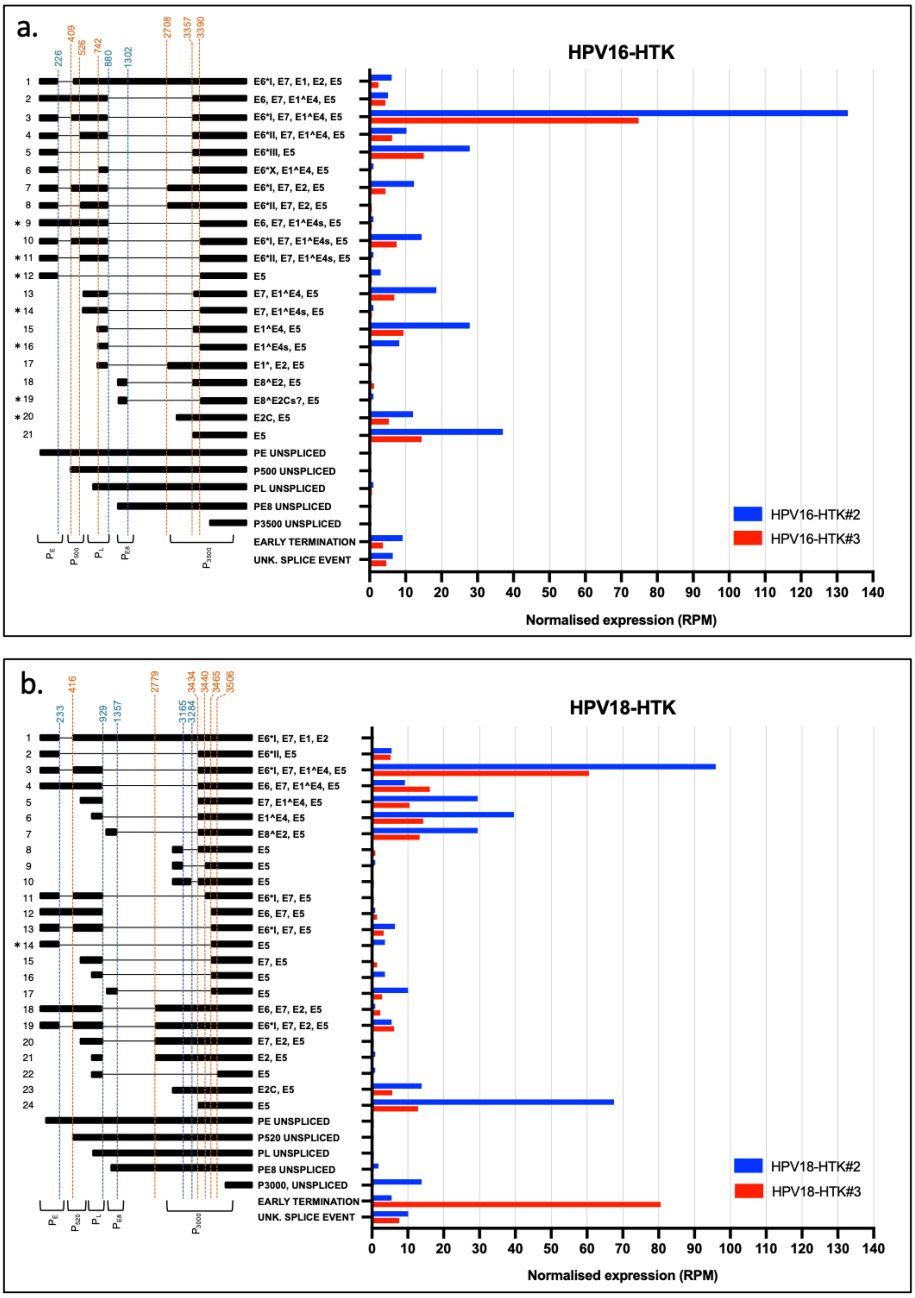


Figure 4-5: HPV16-HTK and HPV18-HTK transcript maps

Distinct viral mRNA species are labelled numerically. Coding regions are demonstrated by black boxes. Splice donor (SD) sites are marked with blue dashed lines and splice acceptor sites (SA) with orange dashed lines. Transcripts end at the early polyadenylation site. The coding potential of each mRNA species is labelled on the right and transcript abundance is detailed in reads/million (RPM). Approximate promoter regions are depicted at the bottom of each splice map. Previously unidentified novel transcripts are signified with asterisks. **a**. Transcript maps of HPV16-HTK#2 (blue) and HPV16-HTK#3 (red). **b**. Transcript maps of HPV18-HTK#2 (blue) and HPV18-HTK#3 (red). Transcript maps were constructed in collaboration with Dr Roberts.

Table 4-3: Relative aggregated expression of HPV16 and HPV18 mRNA isoforms in primary HTK lines

Viral genes	HPV16-HTK#2 reads/mill	HPV16-HTK#2 % Viral transcripts	HPV18-HTK#2 reads/mill	HPV18-HTK#2 % Viral transcripts	Log_2 fold change	HPV16-HTK#3 reads/mill	HPV16-HTK#3 % Viral transcripts	HPV18-HTK#3 reads/mill	HPV18-HTK#3 % Viral transcripts	Log_2 fold change
E6	6.2	1.9%	11.1	3.2%	0.84	5.0	3.1%	20.51	12.3%	2.03
E7	197.0	69.9%	137.4	39.2%	-0.52	103.8	64.2%	83.0	49.6%	-0.32
E6*I	166.1	50.5%	107.9	30.8%	-0.62	89.3	55.3%	70.6	42.2%	-0.34
E6*II	11.43	2.8%	5.5	1.6%	-1.04	6.9	4.3%	5.3	3.1%	-0.40
E6*III	27.85	8.5%	N/A	N/A	-	15.1	9.3%	N/A	N/A	-
E2	38.2	11.6%	7.4	2.1%	-2.37	14.5	9.0%	9.1	5.4%	-0.67
E8^E2	1	0.30%	29.5	8.4%	4.88	1.3	0.8%	13.4	8.0%	3.41
Total number of episome-derived viral transcripts	329.1	100%	350.9	100%	N/A	165.8	100%	251.4	100%	1.00

4.6 Investigation of E6 and E6* splicing variations between HPV16-HTK and HPV18-HTK models

Mapping of HPV16 and HPV18 transcriptomes in the primary HTK lines revealed virus-type specific variations in expression of E7- and E6-encoding mRNAs (Table 4-3). As explained in section 4.5.2, alternative generation of E7 over E6 expressing transcripts is largely contingent on splicing of the E6 ORF [114, 504]. In comparison to the HPV18-HTK lines, HPV16-HTK showed higher expression of transcripts predicted to encode E6*I/E7 and lower expression of polycistronic unspliced-E6/E7 transcripts, which primarily encode E6 (Table 4-3) [114].

Ratios of spliced and unspliced E6 mRNAs (E6*/E6) were calculated to better comprehend differences in splicing of the E6 ORF between HPV16- and HPV18-HTK. Recent studies by Carper, Troutman [507] showed in HPV-HNSCCs that a ratio between spliced E6*I and unspliced E6 (E6*I/E6) mRNAs accurately reflected both E7 protein abundance and gene expression of the host-cellular E7 targets, *E2F1* and *RB*. As such, using the aggregated variant expression data shown in Table 4-3, E6*I/E6 ratios were used to predict and compare expected E7 abundance between HPV16- and HPV18-HTK (Figure 4-6). Here, E6*I/E6 ratios were respectively 2.7- and 5.2-fold higher in HPV16-HTK#2 and HPV16-HTK#3 lines than in donor-equivalent HPV18-HTK lines, advocating comparatively higher levels of E7 proteins (Figure 4-6).

Taking this a step further, E7/E6 ratios compared the aggregated expression of all transcripts predicted to encode E7 against unspliced E6-encoding transcripts (Table 4-3). Each assessed HPV16/18-HTK line showed much higher expression of E7-encoding mRNAs than E6-encoding

mRNAs, indicating a higher comparative production of E7 than E6 (Figure 4-6). However, ratios varied substantially between cell lines, ranging between 4.1 and 31.9. When comparing HPV16 and HPV18, E7/E6 ratios were shown to be respectively 2.6- and 5.1-fold higher in HPV16-HTK#2 and HPV16-HTK#3 lines in comparison to donor-equivalent HPV18-HTK lines, further indicating greater production of HPV16 E7 than HPV18 E7 in HTK overall.

Additional E6*/E6 ratios were calculated to measure relative expression of alternatively spliced transcripts predicted to encode E6*II and E6*III against unspliced E6-encoding transcripts. HPV16 E6*II-encoding mRNA isoforms are spliced between SD226 to a unique splice acceptor located at nucleotide 526 (Figure 4-5a). E6*II/E6 ratios of 1.8 and 1.4 demonstrated relatively similar levels between the two mRNA variants in HPV16-HTK#2 and HPV16-HTK#3 respectively (Figure 4-6). Ratios between HPV16 E6*III-encoding mRNAs, which use the SD226^SA3357 splice junction, and unspliced E6 (E6*III/E6) demonstrated higher proportional expression of E6*III-encoding transcripts than unspliced E6-encoding transcripts, ranging between 3 and 4.5 (Figure 4-6). In contrast, ratios between HPV18 E6*II, which uses a comparable SD233^SA3434 junction, and unspliced E6 mRNAs (E6*II/E6) demonstrated lower expression of spliced E6*II transcripts than unspliced E6 transcripts, where ratios ≤0.5.

Together, these data imply that the E6 ORF was subjected to greater levels of splicing in HPV16-HTK than in HPV18-HTK, driving much higher proportional expression of spliced E6* mRNA variants overall. These findings further exemplify E6*I/E7 transcripts as the most abundantly expressed E6/E7-derived mRNA isoform in both HPV16/18-HTK but suggest greater comparative production of E7 in HPV16-HTK than in HPV18-HTK.

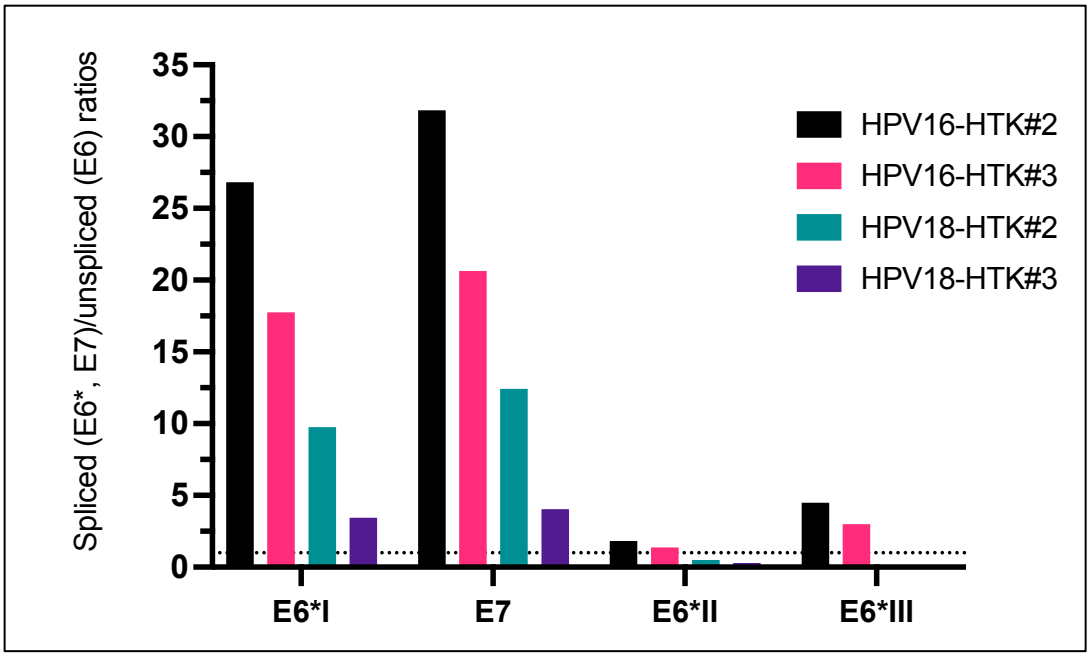


Figure 4-6: HPV16-HTK and HPV18-HTK spliced and unspliced E6 ratio comparisons

HPV16 and HPV18 E6*/E6 ratios were calculated by dividing total number of E6*, or E7 splice encoding transcripts by the total number of unspliced E6 transcripts, carried out for E6*I-, E7, E6*II-, and E6*III-encoding transcripts. The dotted line shows y = 1, which would indicate a E6*/E6 ratio of 1:1.

4.7 HPV16-HTK demonstrate higher E7 protein production and activity than HPV18-HTK

To see whether differential spliced/unspliced E6 ratios corresponded with differences in E7 protein levels, Western blots were performed using three independent replicate samples of early passage HPV-HTK (<50 population doublings [PD]), comprising three sequentially passaged cell lysates of each HPV16/18-HTK cell line. Immunoblotting was carried out using separate HPV16 and HPV18 E7 antibodies and densitometry was performed using ImageJ to quantify band intensity. β-actin antibodies were also added at a consistent dilution to normalise band intensity and estimate relative E7 abundance between samples and membranes. Significantly higher levels of E7 were indicated in HPV16-HTK cells when compared with respective HPV18-HTK donor counterparts (Figure 4-7). In addition, differential E7 levels were consistent with E6*I/E6 ratios shown in Figure 4-6. However, given that two separate antibodies were required, it is possible that HPV16 and HPV18 E7 could have variable affinity to each respective antibody.

To further validate virus-type specific differences in E7 production, additional immunoblots were prepared to examine and compare production of the E7 surrogate marker, MCM7 (Figure 4-8a) [88, 215, 216]. Significant differences in MCM7 protein levels were demonstrated between HPV16-HTK and HPV18-HTK lines, following a similar trend to E7, and altogether supporting higher production and activity of E7 in HPV16-HTK than in HPV18-HTK. When compared with untransfected-HTK cells, both HPV16-HTK and HPV18-HTK demonstrated increased production of MCM7. However, more significant increases were demonstrated in the HPV16-HTK lines overall.

4.7.1 HPV-HTK E6 Western blot findings

Western blots were also attempted for HPV16 and HPV18 E6 but with little success in our laboratory when using the commercial antibodies HPV16 E6-2E-3F8 (Euromedex, France), HPV16 E6-6F4 (Euromedex, France). Therefore, immunoblotting for the E6 target p53 was carried out to compare downstream E6 protein activity (Figure 4-8b). In HTK#2 cells, p53 was significantly reduced in HPV16-HTK#2 when compared against both untransfected- and HPV18-containing HTK cells, suggesting higher activity of E6 in HPV16-HTK#2 than in HPV18-HTK#2. However, no significant differences in p53 levels were observed between HPV16- or HPV18-HTK#3 cell lines. These data contrast with the Nanopore findings, which indicated higher expression of unspliced E6-encoding transcripts in HPV18-HTK than in HPV16-HTK (Table 4-3), highlighting a need for further experimental validation.

4.7.2 RNA-Seq analysis may support higher E7 activity in HPV16-HTK than HPV18-HTK

Illumina-based RNA-Seq was carried out to assess HPV-induced transcriptional reprogramming of the isogenic HPV-HTK lines (Chapter 5). Here, gene set enrichment analysis (GSEA) was performed to identify biological pathways that were significantly, functionally altered between HPV16-HTK and HPV18-HTK, based on cumulative differences in gene expression (see section 5.6.2, Figure 5-6). To account for the scale of this analysis, multiple testing corrections were used to adjust p-values for the false discovery rate, increasing statistical reliability.

When comparing HPV16-HTK and HPV18-HTK, GSEA of characteristic E6- and E7-target pathways showed that adjusted p-values (q-values) did not reach significance. However, significant differences in expression of the E7-target pathways, 'E2F targets' and 'G2/M checkpoints' was observed between the two virus-type models when testing (non-adjusted) p-values. The E2F targets gene set comprises genes that encode cell cycle modulated E2F transcription factors. Overall expression of this pathway was significantly higher in the HPV16-HTK lines than in the HPV18-HTK lines, where $p = 0.0136$ (data not shown). Further, the G2/M checkpoints pathway, which entails genes involved in cell cycle progression from the G2-phase to mitosis, was also shown to be significantly higher in HPV16-HTK than HPV18-HTK ($p = 0.0335$).

While this statistical methodology does not provide the same level of confidence as the q-value, a significant p-value could still suggest biologically significant differences in functional activities of a pathway between the isogenic HPV16/18-containing HTK lines. Put together, these findings further indicate increased E7 oncoprotein activity in HPV16-HTK than in HPV18-HTK. No significance differences in p53 target pathways were observed between the two virus-type lines ($p = 0.69$), suggesting little difference in E6-associated regulation of p53 between HPV16- and HPV18-HTK lines.

In conclusion, these data suggest higher expression and translation of E7 in HPV16-HTK than in HPV18-HTK, resulting in comparatively increased activity of E7 in HPV16-HTK, as indicated by a higher abundance of MCM7, along with potential differential regulation of E7-target pathways.

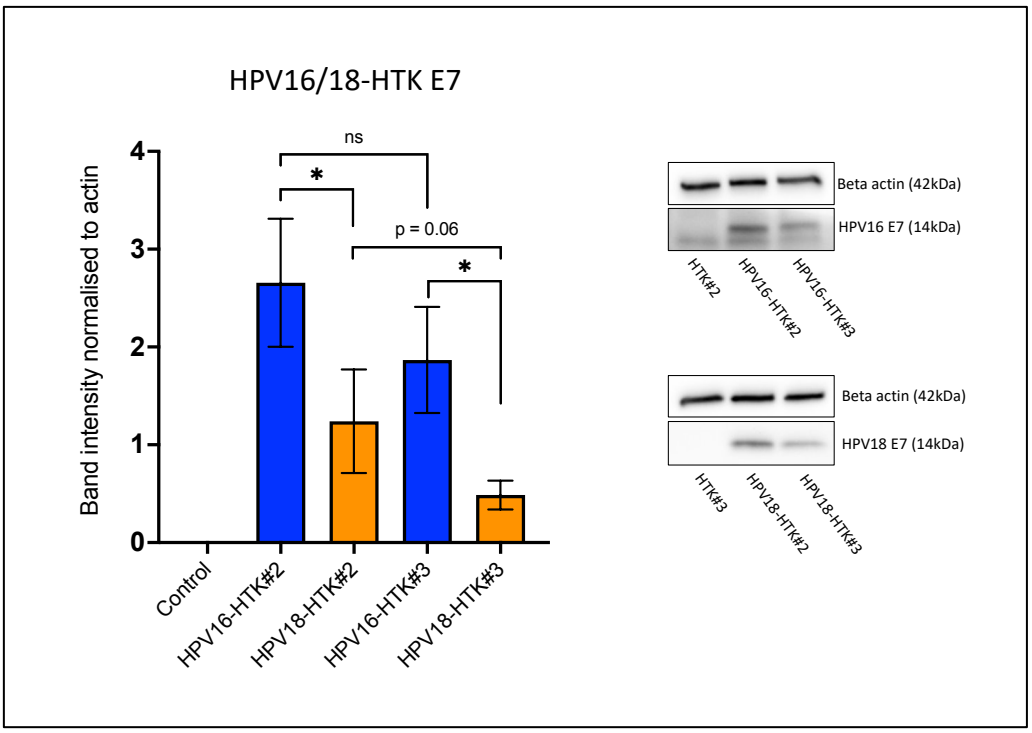


Figure 4-7: Analysis and comparison of HPV16 and HPV18 E7 protein abundance in HTK

Densitometry was performed to calculate the intensity of each band following normalisation against β-actin. Representative western blots are shown on the right. A complete set of E7 western blots can be found in supplementary Figure S-2. Error bars demonstrate standard deviation, unpaired parametric T-tests were carried out to assess significance (* = $p < 0.05$).

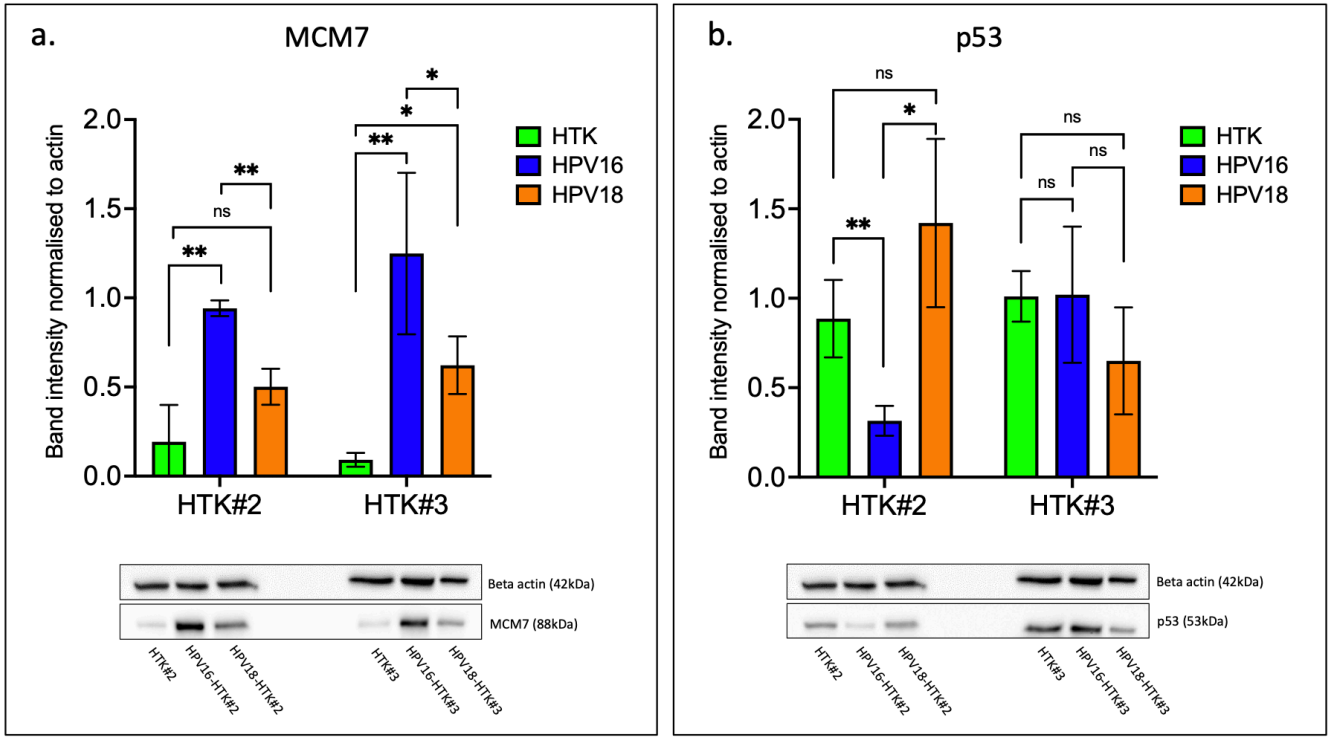


Figure 4-8: Comparison of HPV16 and HPV18 oncoprotein targets MCM7 and p53 production in HTK

a. Western blotting of MCM7 protein expression. Densitometry was used to quantify and compare band intensity between HPV16-HTK and HPV18-HTK, following normalisation against actin. **B**. Similar western blotting and densitometry analyses were performed on p53, a surrogate marker of E6 activity. Representative western blots are shown below each respective graph. Graphs show standard deviation. Unpaired parametric T-test were carried out to assess significance. ns = non-significant, * = <0.05. A complete set of western blots is shown in supplementary Figure S-2.

4.8 Comparison of E6 and E7 protein levels between early and late passage HPV-HTK models

To investigate potential differences HPV16- and HPV18-driven transformation in the tonsils, additional Western blotting studies were performed to further investigate and compare differences in E6 and E7 activity over long-term culture of the primary HPV-HTK lines. Total protein lysates were extracted from low passage cells, ranging between P4 – P7 (<40 PD) and high passage cells, ranging between P18 – P20 (>120 PD).

Previous Southern blot analyses demonstrated primarily episomal HPV16/18 genome organisation in low passage (LP) HPV-HTK#2 and HPV-HTK#3 cells below P8 (<40 PD) (section 3.2.4, Figure 3-4). However, in high passage (HP) cells >120 PD, HPV16-HTK#3, HPV18-HTK#2 and HPV18-HTK#3 presented loss of episomal DNA and likely presence of virus-host junction fragments, suggesting integration of the HPV genome. In contrast, HPV16-HTK#2 was the only HPV-HTK model to remain episomal over extended passage.

In agreement with the findings presented in Figure 4-7, blotting for E7 indicated significantly higher abundance of the oncoprotein in LP HPV16-HTK than in LP HPV18-HTK lines (Figure 4-9a). In both HTK donor lines, HPV16 and HPV18 E7 protein levels were not significantly different between LP and HP. However, non-significant increases of E7 were seen from low- to high-passage HPV18-HTK#2 cells. Similar trends were observed when immunoblotting for MCM7 (Figure 4-9b). Notably, significant increases in MCM7 levels were also demonstrated between low- and high-passage HPV18-HTK#2 cells, suggesting increased activity of E7. Whilst levels of MCM7 were slightly greater in HP HPV18-HTK#3 cells than the LP cells, the differences

did not reach significance. Levels of MCM7 in untransfected donor cells was lower than those containing the viral genomes reflecting the increased cell proliferation in the presence of the virus.

The surrogate marker p53 was used to compare E6 activity between early and late passage HPV-HTK cells. p53 protein levels decreased in HP cells compared to LP for both viruses and in both models, indicating an increase in E6 activity over long-term culture (Figure 4-10). Statistical significance was reached for HPV18-HTK#2, HPV16-HTK#3 and HPV18-HTK#3, and non-significant decreases were shown in HPV16-HTK#2 cells. No significant change in p53 levels were observed between HPV16- and HPV18-containing HTK lines. As expected, p53 levels were highest in the donor cells than those carrying the HPV genomes.

In summary, these data demonstrate little significant change in E7 production and activity between LP and HP HPV16/18-HTK lines, other than in HPV18-HTK#2 cells. However, consistent reductions in p53 suggest increased E6 activity over long-term culture irrespective of the physical status of the HPV genome. It's not clear why changes in E6 activity were not consistent with E7. However, it's possible that the high passage cells may be promoting higher expression of unspliced E6-encoding transcripts. Increased E6 and E7 oncogene expression is a common product of HPV integration into the host genome, accentuated by increases in transcription and mRNA stability [508]. This was likely the case for HPV18-HTK#2 cells, which indicated higher E6 and E7 activities.

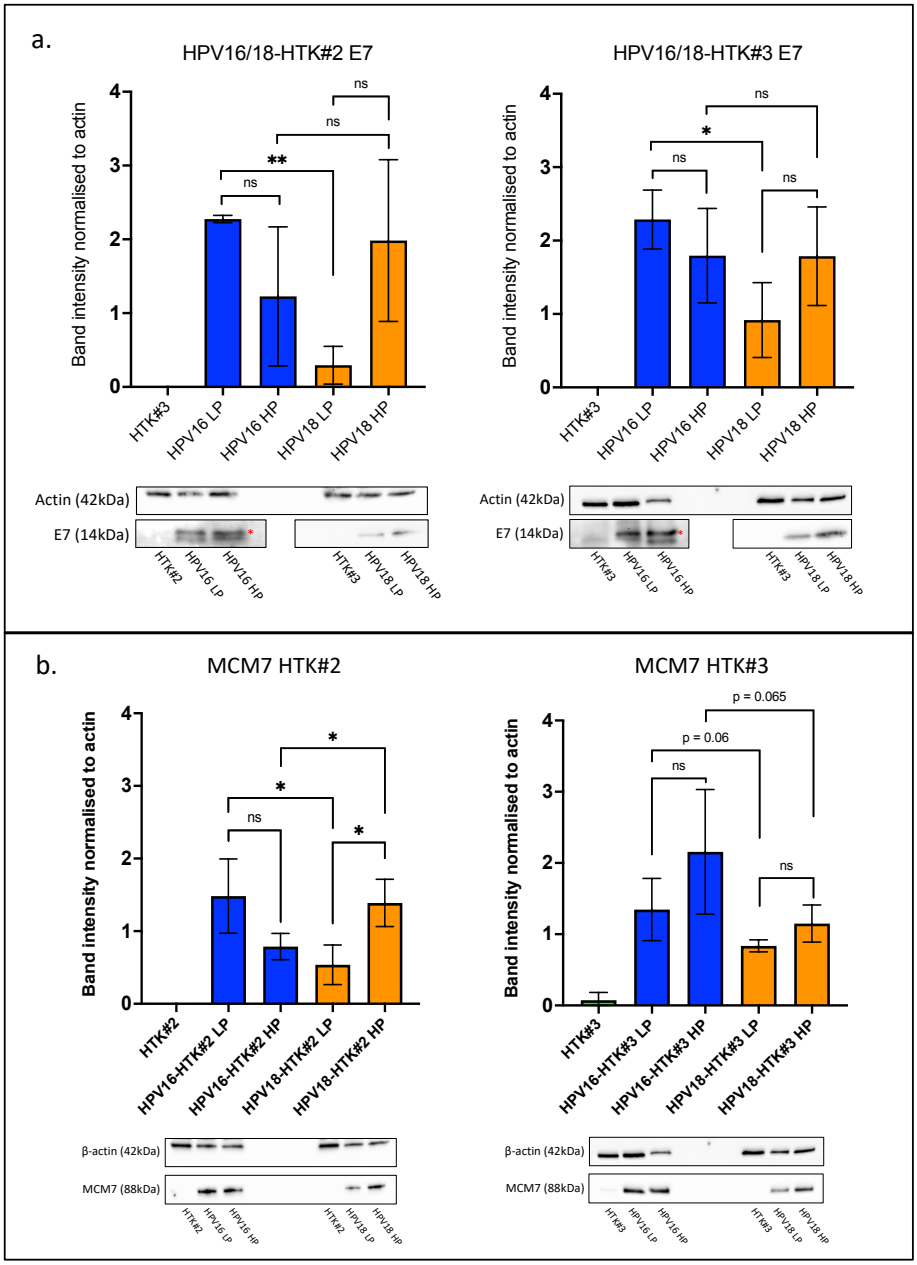


Figure 4-9: Comparison of E7 production in early and late passage HPV-HTK

Western blotting was performed to compare differences in E7 expression over long-term culture. **a**. Densitometry based analysis of HPV16 and HPV18 E7. **B**. Densitometry based analysis of E7 surrogate marker, MCM7. Not shown in graphs, all comparisons between untransfected-HTK models and HPV-HTK models are significant. Representative western blots are shown below each graph (red asterisks demonstrate expected E7 bands, positioned at 14kDa). Graphs show standard deviation. Unpaired parametric T-test were carried out to assess significance. ns = non-significant, * = <0.05. LP = low passage and HP = high passage. The complete set of E7 Western blots can be seen in supplementary Figure S-3.

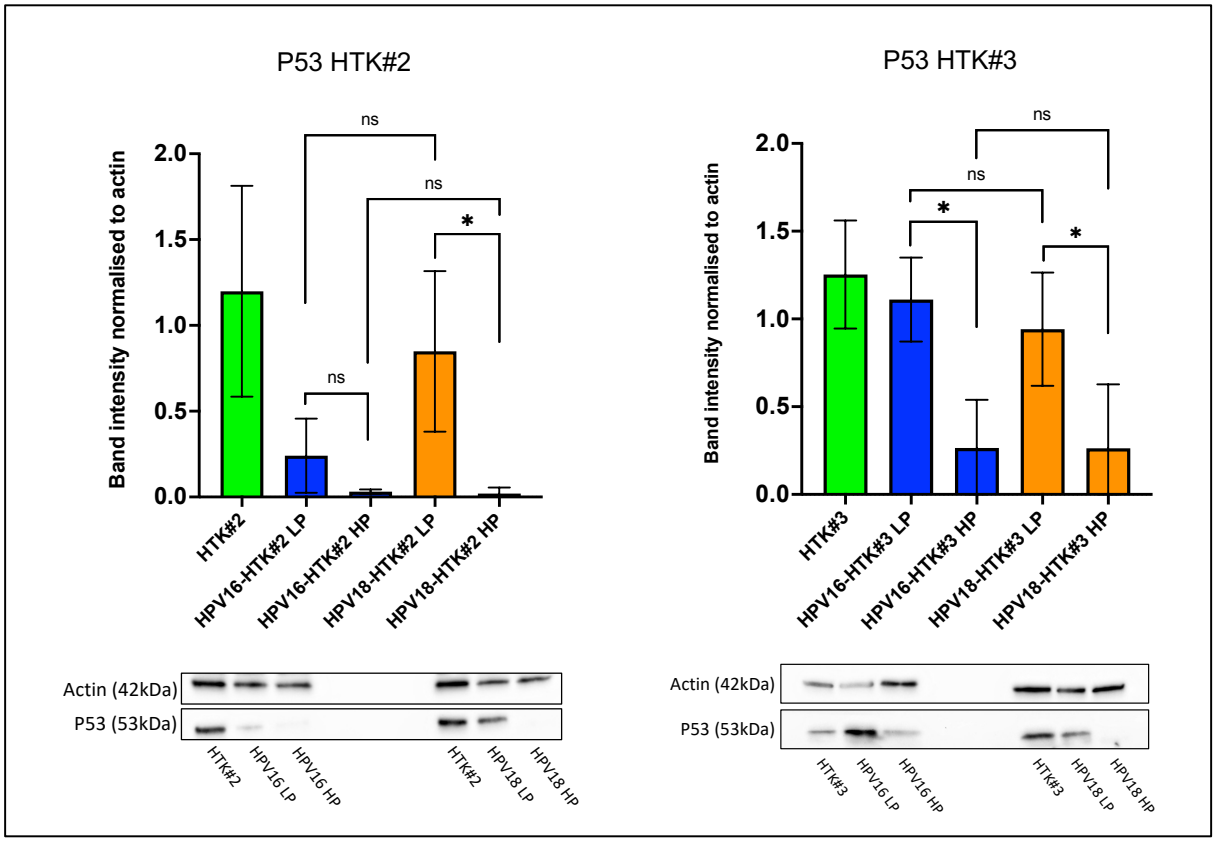


Figure 4-10: Comparison of p53 production in early vs late passage HPV-HTK

Western blotting of the surrogate E6 marker, p53 was performed to compare E6 production between early and late passage HPV-HTK models. Representative western blots are shown below each respective graph. Graphs show standard deviation. Unpaired parametric T-test were carried out to assess significance. ns = non-significant, * = <0.05. LP = low passage and HP = high passage. The complete set of western blots can be seen in supplementary Figure S-3.

4.9 HPV16 and HPV18 differentially express E8^E2 in primary HTK lines

Distinct differences in pE8 promoter usage and E8^E2 transcript abundance were observed between HPV16-HTK and HPV18-HTK lines, which suggested comparatively higher expression of E8^E2 in HPV18-HTK lines than in HPV16-HTK lines (Figure 4-1 - Figure 4-5). Additional qPCR studies, using three sequentially passaged independent replicates of each HTK line, were conducted to validate differences in E8^E2 expression. HPV16- and HPV18-E8^E2 primers designed for this analysis contained nucleotide sequences of the E8 ORF in the forward primer and of the E2 ORF in the reverse primer (see materials and methods, Table 2-2). Based on Nanopore-HPV sequences (Figure 4-5), only spliced E8^E2 and full-length mRNA transcripts contained both of these forward and reverse primer regions. However, full-length amplicons would likely be too long for amplification (>2kb). DNA agarose gel electrophoresis of the qPCR products confirmed single size-specific products (Figure 4-11a). In addition, qPCR dissociation curves showed no evidence of non-specific amplification (Figure 4-11b).

Given that alternative HPV16- and HPV18-E8^E2 primer sets were used for these studies, the data was presented as delta Ct, following normalisation against β-actin. Based on the Ct values, the data was shown to follow the trend that E8^E2 expression is higher in HPV18-HTK than HPV16-HTK (Figure 4-11a). However, clearer differences in expression were observed between HTK#2 models than between HTK#3 models, which did not reach significance.

4.9.1 HPV genome copy number differences between HPV16 and HPV18 isogenic HTK models

Given that E8^E2 are potent regulators of viral genome replication (introduction, section 1.2.3), additional Southern blotting was performed to see whether differences in E8^E2 expression between HPV16- and HPV18-HTK lines correlated to differences in episome copy numbers. Blots hybridized to radiolabelled HPV16 or HPV18 DNA were exposed on Fujifilm and densitometry analysis was carried out using ImageJ (Figure 4-12). Copy number controls, equivalent to 10, 50, 100 and 150 copies of HPV16 and HPV18 genomes per cell were used to estimate HPV copy numbers via generation of linear trendlines. The coefficient of determination, (R^2), demonstrated good trendline correlation, where R^2 >0.93 (data not shown).

At P5 (~30-40 PD), HPV16 episomes accounted for around 95 (HTK#2) and 58 (HTK#3) copies per cell (Figure 4-12b). This was 2.9- and 13.1-fold higher than donor-equivalent HPV18-containing cells, which supported just 33 (HTK#2) and 4 (HTK#3) copies per cell. These findings demonstrated differences in episome copy numbers between HPV16-HTK and HPV18-HTK lines and correlated with the differences in E8^E2 expression, which could suggest greater E8^E2-mediated control of HPV DNA replication in HPV18-HTK.

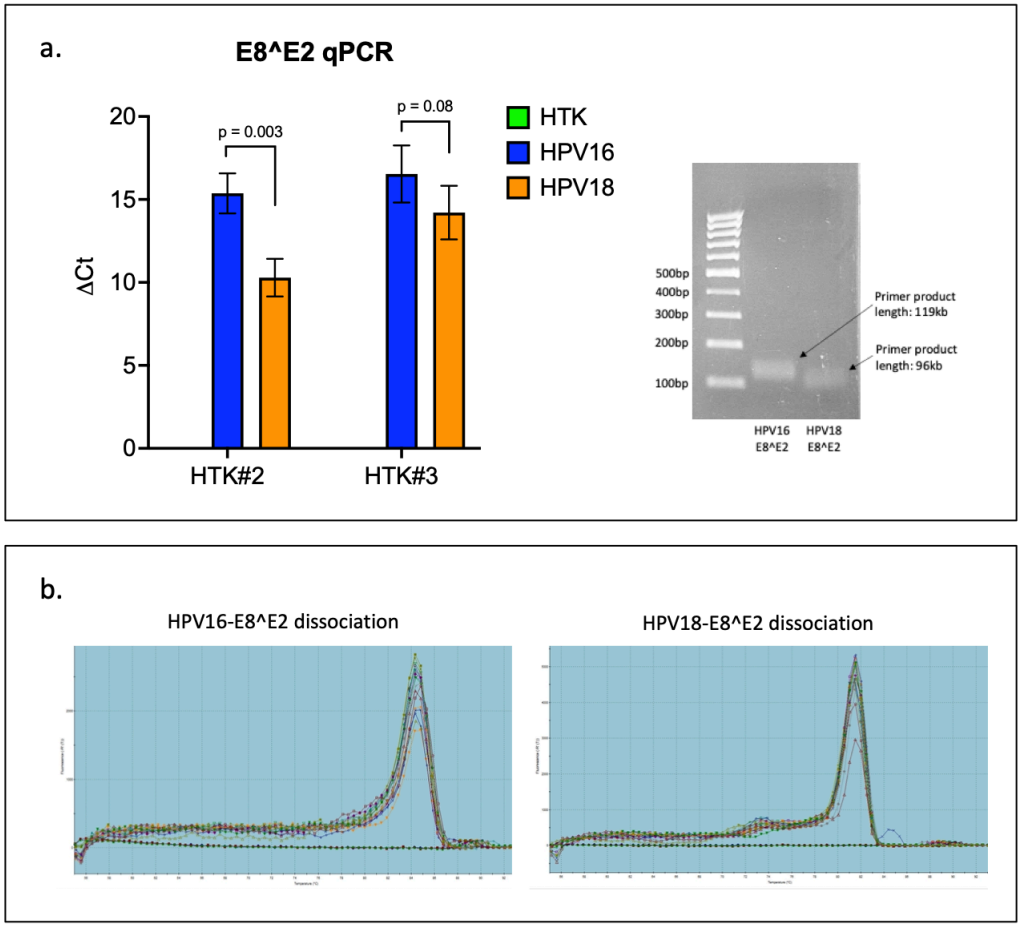


Figure 4-11: Differential gene expression of E8^E2 between HPV16-HTK and HPV18-HTK

a. (left) qPCR analysis shows delta Ct, where E8^E2 Ct values had been normalised against β-actin. E8^E2 was undetected in the HTK controls (green). One-tailed parametric paired T-tests assessed significance, where $p < 0.05$ (*). Error bars show standard deviation. DNA agarose gel electrophoresis was used to confirm E8^E2-specific amplification, using qPCR cDNA product (right). Band sizes correspond with expected HPV16- and HPV18-E8^E2 primer lengths of 119kb (HPV16) and 96 kb (HPV18). **b**. Dissociation curves from the HPV16- and HPV18-E8^E2 qPCR runs show the melting temperature of the qPCR products. Here, single peaks indicate amplification of a distinct mRNA species.

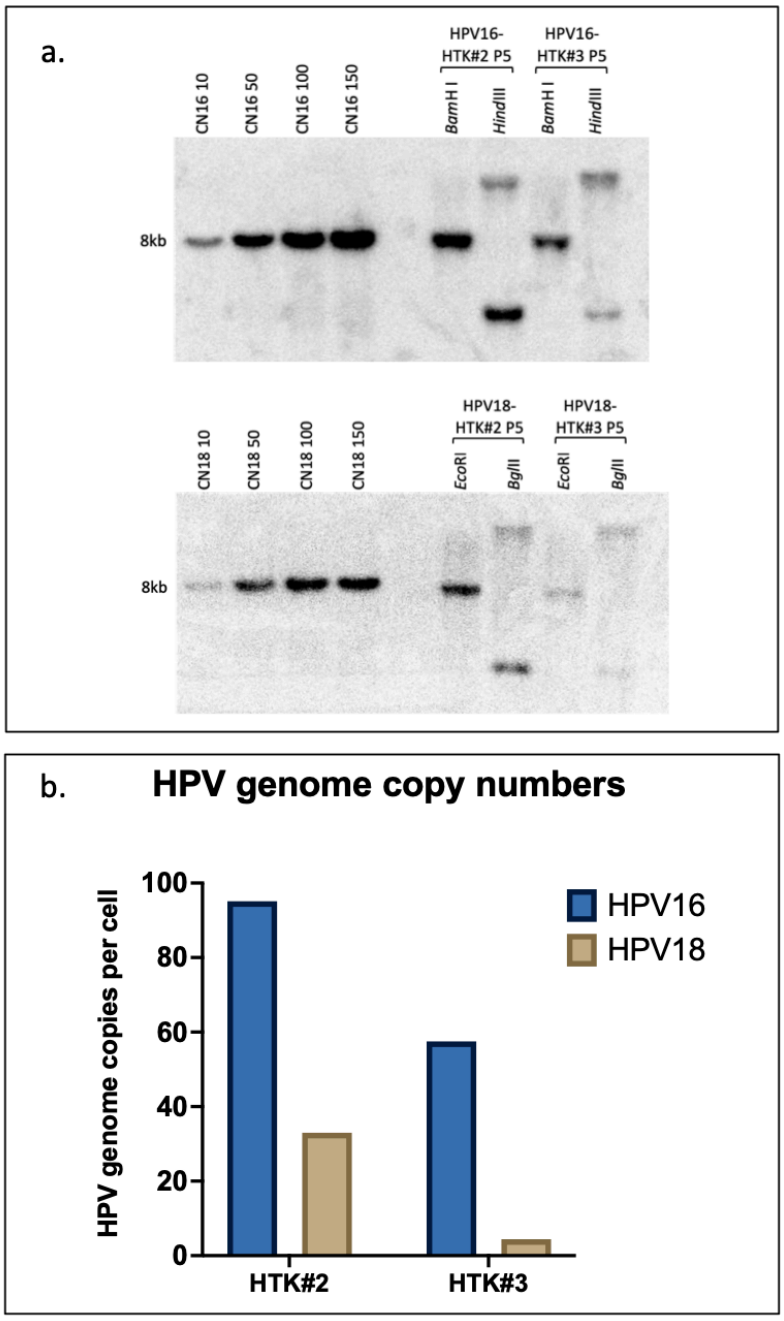


Figure 4-12: Quantification of HPV16 and HPV18 copy numbers

Linearised HPV genome copy numbers were compared against copy number controls (CN16 and CN18), equivalent to 10, 50 100 and 150 whole genome copies per cell. **a**. HPV16 and HPV18 DNA, extracted from HTK#2 and HTK#3 isogenic cell lines was linearised using restriction enzymes *Bam*H I (HPV16) and *Eco*RI (HPV18). Digestion with non-cutter restriction enzymes *Hind*III (HPV16) and *Bgl*II (HPV18) was carried out to identify circular whole HPV genomes **b**. densitometry analyses were performed to estimate HPV genome copies per cell.

4.10 Investigation of virus-host fusion transcripts in primary tonsil keratinocytes

Integration of HR-HPV DNA into the host genome has been estimated to occur in around 39 – 43% of HPV-OPSCCs [396, 397]. In the primary HPV-HTK lines, Southern blotting denoted HPV DNA integration in HPV16-HTK#3, HPV18-HTK#3 and HPV18-HTK#2 in extended cell culture (section 3.2.4, Figure 3-4). In early cell culture however, all four HPV16/18-HTK lines established and maintained viral episomes up to at least <50 PD, so were considered suitable for Nanopore-based transcription profiling. However, alignment of the Nanopore-sequencing reads to HPV and human reference genomes revealed the presence of virus-host fusion transcripts, suggesting possible early integration events (Table 4-1). The total number episome-derived transcripts and virus-host fusion transcripts identified by Nanopore, as described in Table 4-1, is illustrated in Figure 4-13a. HPV18-HTK#3 cells contained the highest number of virus-host fusion transcripts, at 399.3 RPM, comprising 61.4% of total (integrated and non-integrated) viral reads. This compared with 16.6 RPM (4.6%) of HPV18-HTK#2 reads, 16.4 RPM (9%) of HPV16-HTK#3 reads, and 170.2 (34.2%) of HPV16-HTK#2 reads. Further analysis of the virus-host fusion transcript data was carried out to identify possible integration events or tendencies in HTK that may contribute to HPV-driven progression.

Analysis of human mRNA sequences in virus-host fusion transcripts was carried out to identify possible HPV integration breakpoint locations in the host genome. Seventy-four unique integration sites were identified in total, consisting of 11 HPV16-HTK#2, 9 HPV16-HTK#3, 18 HPV18-HTK#2, and 38 HPV18-HTK#3, which were mapped onto representative chromosomes (Figure 4-13b). Chromosomal locations varied substantially, where integration breakpoints were identified at many different sites within each chromosome, excluding chromosome 9.

Some clustering was seen around the centromere of chromosome 5, where a collection of unique integration breakpoints was detected in HPV16-HTK#3 and HPV18-HTK#2 lines, as well as a further three unique breakpoints in HPV18-HTK#3. Potential clustering was also seen in chromosomes 3, 7, 8 and 17. The highest number of unique integration events were seen in chromosome 1, which displayed a total of 9 integration breakpoints over the four HPV-HTK models.

Inspection of each virus-host fusion transcript was carried out to further examine the prospective locality of each fusion point in the host genome. Table 4-4 shows an overview of findings, including the nearest human gene to the prospective site of integration, and whether the fusion transcript was located upstream, downstream, or within the coding sequence of a gene. In addition, the table demonstrates the number of fusion transcript copies at a specific location, and whether the virus-host fusion transcript integrates into a fragile site in the host genome. A number of integration events were identified at specific loci of cancer-related genes. Notably, both HPV18-HTK#3 and HPV18-HTK#2 virus-host fusion transcripts were shown to contain human sequences corresponding to the 3q28 locus, and directly upstream of the central tumour suppressor *TP63* (Table 4-4). In addition, 23 HPV16-HTK#2 fusion transcript copies demonstrated human sequences synonymous to intron 1 of the *TRAF3* ORF, which is centrally involved in the transactivation of NF-kB. Other integration events were observed within or upstream of genes that play roles in a number of cancer related pathways, including *TP63, WWOX, FBXW8, RTKN, FGF14, NFYC, CDH13*, and *ROCK1* (Table 4-4).

Multiple copies of a fusion transcript were observed at a number of specific loci, particularly in HPV18-HTK#3 and HPV16-HTK#2 cells, suggesting that HPV DNA had integrated at a transcriptionally active region of the host genome (Figure 4-13b). Notably, in HPV18-HTK#3, the vast majority of virus-host fusion transcripts were detected at three distinct transcriptionally active loci, contained within the *FBXW8* ORF, and proximal to *RNU6-34P* and *RNF180* genes, which altogether accounted for 55.5% of all viral HPV18-HTK#3 transcripts (Table 4-4). Similarly, in HPV16-HTK#2 cells, virus-host fusion transcripts expressed upstream of *AC073901.1*, a processed pseudogene, made up 26.3% of total viral reads.

Examination of viral promoter usage and splice junctions was carried out to predetermine potential viral gene expression of virus-host fusion transcripts (Table 4-4). Here, the vast majority of identified HPV16- and HPV18-fusion transcripts (>90%) started at the pE, producing polycistronic E6/E7 mRNAs. These were primarily spliced between SD226^SA409 (HPV16) and SD233^SA416 (HPV18), which made up 74.3% and 80.4% of total virus-host fusion transcripts respectively, resulting in production of E7-encoding E6*I mRNA isoforms (Figure 4-5) [114]. Unspliced E6-encoding transcripts respectively accounted for 2.7% and 11.2% of total HPV16- and HPV18-HTK virus-host fusion transcripts. In addition, 3.7% of HPV16-HTK used SD226^SA526 splice junctions, so were predicted to encode E6*II and E7, and 8% of HPV16-HTK fusion transcripts contained SD226^SA3357 splice junctions, which contains the E6*III ORF (Figure 4-5). In contrast, virus-host fusion transcripts containing E1, E2, E8^E2, or E1^E4 ORF were rarely observed. As such, high levels of virus-host fusion transcripts, as found in HPV18-HTK#3 and HPV16-HTK#2 cells may have elevated expression of E7, and to a lesser extent, E6.

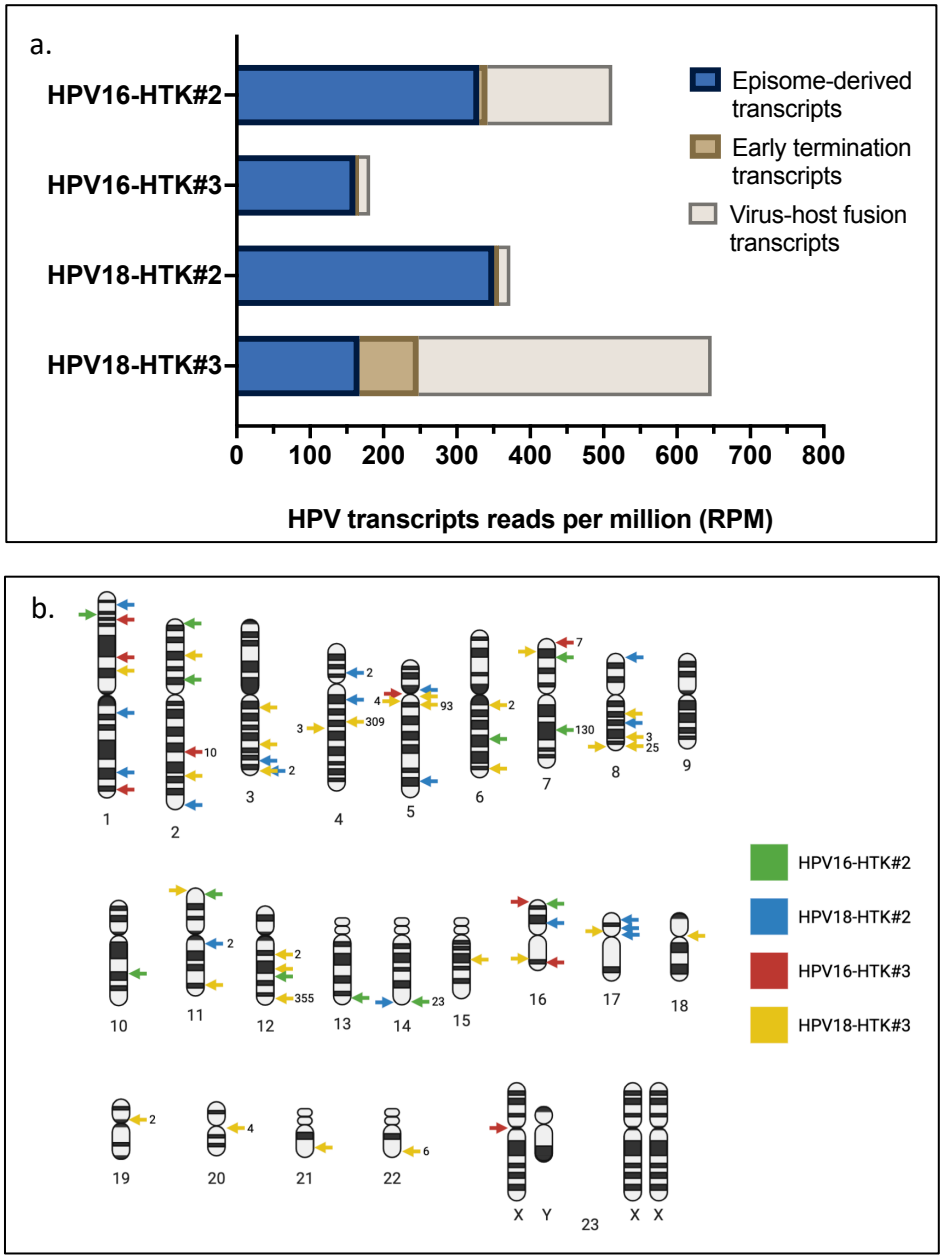


Figure 4-13: HPV16 and HPV18 integration tendencies in primary HTK

a. Stacked bar graph compares total number of episome-derived mRNA transcripts (blue), as identified by Nanopore, against the total number of virus-host fusion transcripts (cream), and the number of viral transcripts which ended before the pAE signal, which are also possible integrated transcripts. Equivalent numbers of episome-derived transcripts are seen between HPV16-HTK and HPV18-HTK of the same donor. **b.** Approximate integration breakpoint locations of virus-host fusion transcripts were mapped onto representative chromatin, marked at specified loci with coloured arrows. Each colour is representative of one of the four HPV-HTK models. The number of virus-host fusion transcript copies is indicated next to each relevant breakpoint position. Figure created with BioRender.com.

Table 4-4: HPV16- and HPV18-HTK virus-host fusion gene coordinates in the human genome

Summarises Nanopore-detected virus-host fusion transcripts. The colour and position of each integration breakpoint matches those depicted in Figure 4-13. Approximate integration breakpoint locations are shown in relation to the human gene coordinates (gene distance). A gene distance of '0' indicates HPV genome incision within the coding sequence of a host gene, (-) indicates a position upstream of a host gene, and (+) indicates a downstream position. HPV16/18 TSS and splice sites were examined to characterise the coding potential of each fusion transcript.

HPV-HTK model	Closest human gene	Human gene coordinates	Gene distance	No. of transcripts	Common fragile site	HPV open reading frame
HPV18-HTK#2	*NECAP2*	**Chr1**: 16440721-16460078	0	1	FRA1A	-
HPV16-HTK#2	*GMEB1*	**Chr1**: 28668732-28719353	0	1	-	E5
HPV16-HTK#3	*FTLP18*	**Chr1**: 36630335-36630857	(-)26937	1	-	-
HPV16-HTK#3	*NFYC*	**Chr1**: 40722094-40724064	0	1	-	E6*I, E7
HPV18-HTK#3	*CCDC18-AS1*	**Chr1**: 93262186-93346025	0	1	FRA1D	E6*I, E7
HPV18-HTK#2	*ARNT*	**Chr1**: 150809713-150876708	0	1	FRA1F	E5
HPV18-HTK#2	*NSL1*	**Chr1**: 212726153-212791782	0	1	-	E5
HPV16-HTK#2	*MCOLN2*	**Chr1**: 84925583-84997113	0	1	FRA1D	E7, E2C?, E5
HPV16-HTK#2	*LINC00487*	**Chr2**: 6728177-6770311	0	1	-	E6, E7
HPV18-HTK#3	*CCDC85A*	**Chr2**: 56183990-56386172	(+)3560	1	-	E6*I, E7
HPV16-HTK#2	*RTKN*	**Chr2**: 74425836-74442422	0	1	-	E6*I, E7
HPV16-HTK#3	*RAPGEF4*	**Chr2**: 172735274-173052893	0	10	FRA2G	**7** (E6*I, E7, E1^E4), **1** (E6*III), **1** (E6*I, E7, E1)
HPV18-HTK#3	*NUP35*	**Chr2**: 183117513-183161680	0	1	FRA2D	E6*I, E7
HPV18-HTK#2	*AC114814.2*	**Chr2**: 235147016-235147500	(+)2524	1	-	E6*I, E7
HPV18-HTK#3	*CBLB*	**Chr3**: 105655461-105869552	0	1	-	E6*I, E7
HPV18-HTK#3	*LINC02006*	**Chr3**: 153384934-153980186	0	1	-	-
HPV18-HTK#2	*SENP2*	**Chr3**: 185582496-185633551	(+)2524	1	FRA3C	-
HPV18-HTK#3	*TP63*	**Chr3**: 189631389-189897276	(-)46920	1	-	E6*I, E7
HPV18-HTK#2	*TP63*	**Chr3**: 189631389-189897276	(-)16514	2	-	E6*I, E7
HPV18-HTK#2	*PDS5A*	**Chr4**: 39822863-39977956	0	2	-	E6*I, E7
HPV18-HTK#2	*AC239584.1*	**Chr4**: 74552584-74589481	(+)18906	1	-	E5
HPV18-HTK#3	*RNU6-34P*	**Chr4**: 96152297-96152403	(+)20818	309	-	**295** (E6*I,E7), **23** (E6, E7), **7** (E7)
HPV18-HTK#3	*AC096661.1*	**Chr4**: 96212279-96213100	(-)15953	3	-	**2** (E6*I, E7), **1** (E6, E7)
HPV18-HTK#2	*MFSD4BP1*	**Chr5**: 52579555-52581106	(+)12786	1	-	E6*I, E7
HPV16-HTK#3	*MFSD4BP1*	**Chr5**: 52579555-52581106	(+)45263	1	-	E6*I, E7
HPV18-HTK#3	*PART1*	**Chr5**: 60487713-60548813	0	1	-	E6*I, E7
HPV18-HTK#3	*AC122707.1*	**Chr5**: 63957893-63981043	(+)5142	4	-	**3** (E5), **1** (E7, E1^E4, E5)
HPV18-HTK#3	*RNF180*	**Chr5**: 64165843-64372869	(-)25095	93	-	**65** (E6*I, E7), **12** (E6, E7), **4** (E7)
HPV18-HTK#2	*LSM1P2*	**Chr5**: 163887506-163887827	(+)81247	1	-	-
HPV18-HTK#3	*AL391417.1*	**Chr6**: 86432244-86432669	(-)85559	2	-	E6, E7
HPV16-HTK#2	*HS3ST5*	**Chr6**: 114055586-114343045	0	1	FRA6F	E7
HPV18-HTK#3	*SYNE1*	**Chr6**: 152121684-152637801	0	1	-	E6*I, E7
HPV16-HTK#3	*FBXL18*	**Chr7**: 5431335-5513809	0	7	FRA7B	**3** (E6*I, E7), **1** (E5), **1** (E6*II, E7)
HPV16-HTK#2	*CRPPA*	**Chr7**: 16087525-16421538	0	1	-	E6*I, E7
HPV16-HTK#2	*AC073901.1*	**Chr7**: 115503367-115504081	(-)71138	130	-	**107** (E6*I, E7), **6** (E6*II, E7), **9** (E7), **3** (E6, E7), **4** (E6*III)
HPV18-HTK#2	*CSMD1*	**Chr8**: 2935353-4994972	0	1	-	E6, E7
HPV18-HTK#3	*CNBD1*	**Chr8**: 86866442-87615219	0	1	-	E6*I, E7
HPV18-HTK#2	*LINC00534*	**Chr8**: 90221341-90687863	0	1	-	E6*I, E7
HPV18-HTK#3	*SAMD12*	**Chr8**: 118189455-118622112	0	3	FRA8C	**3** (E6*I, E7)
HPV18-HTK#3	*EFR3A*	**Chr8**: 131904093-132013642	0	25	-	**23** (E6*I, E7), **1** (E7, E5)
HPV18-HTK#3	*AC104257.1*	**Chr8**: 131308545-131317632	0	1	-	E6*I, E7
HPV16-HTK#2	*HPSE2*	**Chr10**: 98457077-99235862	0	1	-	E8^E2, E5
HPV18-HTK#3	*POLR2L*	**Chr11**: 837356-842529	0	1	-	E6*I, E7
HPV16-HTK#2	*ST5*	**Chr11**: 8693351-8910951	0	1	-	E5
HPV18-HTK#2	*C11orf80*	**Chr11**: 66744451-66843328	0	2	FRA11H	E6, E7
HPV18-HTK#3	*AMBRA1*	**Chr11**: 46396414-46594125	0	1	-	E6, E7, E1^E4, E5
HPV18-HTK#3	*HMGA2-AS1*	**Chr12**: 65851340-65882167	0	2	-	**2** (E6*I, E7, E1^E4)
HPV18-HTK#3	*OSBPL8*	**Chr12**: 76351797-76559809	0	1	-	E6*I, E7
HPV16-HTK#2	*RNU6-1329P*	**Chr12**: 92952639-92952745	(+)9892	1	-	E6*I, E7, E1^E4
HPV18-HTK#3	*FBXW8*	**Chr12**: 116910950-117031148	0	355	FRA12C (rare)	**262** (E6*I, E7), **58** (E6, E7), **6** (E7)
HPV16-HTK#2	*FGF14*	**Chr13**: 101710804-102402457	0	1	-	E6*I, E7
HPV16-HTK#2	*TRAF3*	**Chr14**: 102777476-102911500	0	23	-	**11** (E6*I, E7), **9** (E6*III), **1** (E7), **1** (E6, E7)
HPV18-HTK#2	*IGHVIII-82*	**Chr14**: 106879563-106879812	(+)245918	1	-	E6*I, E7
HPV18-HTK#3	*MYO5A*	**Chr15**: 52307283-52529050	0	1	-	E6*I, E7
HPV16-HTK#3	*MAPK8IP3*	**Chr16**: 1706183-1770317	0	1	-	E6*I, E7, E1^E4, E5
HPV16-HTK#2	*NPRL3*	**Chr16**: 84271-138860	0	1	-	E6*III, E5
HPV18-HTK#2	*AC142381.4*	**Chr16**: 32008438-32084785	0	1	-	E6*I, E7
HPV18-HTK#3	*CDH13*	**Chr16**: 82626965-83800640	0	1	-	-
HPV16-HTK#3	*WWOX*	**Chr16**: 78099430-79212667	(-)1822	1	FRA16D	E6*I, E7
HPV18-HTK#2	*NTN1*	**Chr17**: 9021510-9244000	0	1	-	-
HPV18-HTK#2	*SNORD3B-2*	**Chr17**: 19063346-19064136	(+)3394	1	-	E6*I, E7
HPV18-HTK#3	*CCDC144NL-AS1*	**Chr17**: 20868433-21002276	0	1	-	E6, E7
HPV18-HTK#2	*B9D1*	**Chr17**: 19334308-19378193	0	1	-	E6*I, E7
HPV18-HTK#3	*ROCK1*	**Chr18**: 20946906-21111813	(-)531849	1	-	E6*I, E7
HPV18-HTK#3	*KCNN1*	**Chr19**: 17951293-18000080	0	2	FRA19B (rare)	**1** (E6*I, E7), **1** (E6*I, E7, E5)
HPV18-HTK#3	*EFCAB8*	**Chr20**: 32858923-32961609	0	4	-	**2** (E6*I, E7), **1** (E6, E7)
HPV18-HTK#3	*DOP1B*	**Chr21**: 36156782-36294274	0	1	-	-
HPV18-HTK#3	*CELSR1*	**Chr22**: 46360834-46537170	0	6	-	**3** (E6*I, E7), **1** (E6, E7), **2** (E5)
HPV16-HTK#3	*ZC4H2*	**Chr23** (XY): 10059-10404	0	1	-	E6*I, E7

4.11 Discussion

To seek out differing virus-associated transformation mechanisms between HPV16 and HPV18 in the tonsils, global transcriptomic analysis of our primary HPV-HTK lines was performed to characterise and compare HPV16 and HPV18 transcriptional tendencies and gene expression profiles. In allowing for accurate characterisation and quantification of full-length, native mRNA isoforms, Nanopore was performed on these undifferentiated, episome-maintaining HPV-HTK cells to assess HPV16 and HPV18 promoter usage, polyadenylation, and splicing, and to construct comprehensive viral transcription maps [509, 510].

4.11.1 HPV16- and HPV18-HTK presented varying early viral promoter activities

HPV gene expression profiles alter in a differentiation-dependent manner in response to specific phenotypic or functional changes of the host-cellular environment, ensuring strategic production of specific viral proteins at different stages of the viral life cycle (see introductory section 1.2). Centrally regulated transcriptional elements include the two major pE and pL promoters, as well as pA^E and pA^L signals [87-89]. In undifferentiated, episome maintaining cells, high pA^E activity and a strong pA^E signal promote expression of early viral transcripts required for translation of early proteins such as E6 and E7. High pE activity coupled with almost exclusive transcriptional termination at the pA^E signal was evident in our primary HPV16- and HPV18-HTK lines, signifying the undifferentiated state of the cells (Figure 4-2 and Figure 4-3).

Interestingly, the total normalised expression of all HPV16 and HPV18 episome-derived transcripts remained at near equivalent levels between the two virus-type lines of the same donor. Yet, the pE was shown to be 14 – 27% more active in the HPV16-HTK donor lines than in respective HPV18-containing counterparts, indicating differing pE regulation and denoting higher expression of pE-derived E6/E7 polycistronic transcripts (Figure 4-2).

HR-HPV pE activity is largely regulated by viral E2 and E8^E2 proteins, which can bind one of four palindromic E2 binding sites located within the long control region (LCR) of the viral genome (reviewed in introductory section 1.2.3) [123-125]. In the HTK lines, quantitative analysis of the Nanopore data demonstrated little expression of transcripts that efficiently translate E2 in HPV16- and HPV18-HTK. However, distinct expressional differences of E8^E2, which is a strong negative regulator of pE-transcription, was observed between the two virus-type lines (Table 4-3). Here, relatively high expression of HPV18 E8^E2 contrasted with a near absence of HPV16 E8^E2 expression in HTK, making this a likely candidate for observed differences in pE activities, as is further discussed in section 4.11.3.

HR-HPV pE activities are also regulated by numerous cellular transcription factors including AP-1, SP1, YY1, NF1, Oct-1, and CDP, which have been shown to bind specific transcriptional promoter regions of the viral LCR, upstream of the TATA box and transcription initiation sites [129, 138, 493]. Intriguingly, differential gene expression analysis of the RNA-seq data, comparing HPV16- and HPV18-HTK host transcriptomes (Chapter 5), revealed significant differences in expression of two functionally divergent AP-1 components, *FOS* and *FOSL1* (see Figure 5-8 and supplementary Table S-2). In comparison to HPV18-HTK, HPV16-HTK showed

significantly higher expression of *FOS*, which encodes the oncoprotein c-Fos ($p = 0.015$, $\log_2$ fold change = 0.6) and significantly lower expression of *FOSL1*, which encodes Fra-1 ($p = 0.036$, $\log_2$ fold change = -0.67).

AP-1 are heterodimers formed from a Fos family member (c-Fos, FosB, Fra-1 or Fra-2) and a Jun family member (c-Jun, JunB or JunD), which promote or enhance expression of genes involved in cell proliferation and survival, differentiation, apoptosis, and immunity [511]. Both Fos and Jun family members have varying transcriptional efficacies, where c-Fos- and FosB-containing AP-1 complexes are described as "strong transcriptional activators", and Fra-1- and Fra-2-containing AP-1 complexes are described as "weak transcriptional activators" [512, 513].

AP-1 are constitutively active in HPV-driven HNSCC and cervical cancers and have been reported to be essential for HR-HPV pE-driven E6 and E7 transcription [514-519]. In addition, HPV16 E7 has been shown to directly upregulate expression of *FOS* [520]. Importantly, habitual increases in *FOS* (c-Fos) expression, observed throughout tumourigenesis of HPV-containing oral and cervical cancers, inversely correlate with decreasing expression of *FOSL1* (Fra-1) [521-523]. Further, an orchestrated switch in AP-1 complex composition from a c-Fos/C-Jun heterodimer to Fra-1/c-Jun heterodimers was shown to significantly downregulate E7 transcription in cervical cancer cell lines [524]. As such, opposing trends of *FOS* and *FOSL1* in the HPV16- and HPV18-HTK lines could reflect differences in pE activity, and may be an important distinguishing factor of HPV16-driven OPSCC that requires further exploration.

Another possible factor of these observed differences in pE activity could relate to differing accessibility of viral chromatin around the pE. In 1977, Studies by Favre, Breitburd [525] revealed that HPV DNA was packaged in nucleosomes. The LCR of HPV16 and HPV18 genomes have since been shown to contain at least two nucleosomes, which respectively overlap the viral promoter and the origin of replication/SP1 binding site within the pE [526]. As such, the precise arrangement, and epigenetic regulations of these HPV nucleosomes are likely important elements of transcriptional pE regulation. Importantly, studies by Groves, Knight [408] show that high HPV16 oncogene expression is associated with more open chromatin of the viral LCR. As such, it is feasible that higher pE activity in the HPV16-HTK lines could be associated with comparable derepression of the HPV16 URR, which might better facilitate transcription factor binding and transcription initiation.

Research carried out by the Parish lab have shown that epigenetic modification of HPV18 genomes can alter pE activity and viral genome replication in a differentiation-dependent manner. Here, the host chromatin regulator CCCTC-binding factor (CTCF) and the transcriptional repressor protein Yin Yang 1 (YY1), respectively bind to conserved binding sites within the E2 ORF and LCR, before physically interacting with one another to form a transcriptionally repressive chromatin loop [98, 159]. It is possible that HPV18 DNA could be subject to a greater degree of CTCF/YY1-mediated regulation than HPV16 in HTK, leading to decreased pE activity and viral genome replication, as observed in our HPV16/18-HTK lines (see Figure 4-2 and Figure 4-12). However, differences in pE accessibility could also relate to virus-type modulated variations in histone acetylation, histone methylation, or even differential methylation of CpG islands in HPV DNA. In the first instance, Formaldehyde-

assisted isolation of regulatory elements (FAIRE) could be carried out to directly compare chromatin accessibility around the early promoter of HPV16 and HPV18 genomes in our HPV-HTK lines to assess this possibility.

4.11.2 HPV16 demonstrated higher production of E7 than HPV18 in primary HTK

Characterisation and quantification of episome-derived HPV mRNAs was next carried out to examine relative expression of HPV16 and HPV18 transcripts in our HTK lines. Here, HPV16 and HPV18 transcription profiles were similar to those previously investigated in undifferentiated, low-grade cervical lesions [109, 485, 491, 527], and undifferentiated, anogenital HPV-immortalised anogenital keratinocyte models [90, 91, 528]. Like these previous models and biopsies, transcription profiles of our HPV16- and HPV18-HTK lines appeared more LSIL-like in organisation, presenting a diverse array of viral mRNA isoforms with the potential to encode early life cycle proteins such as E1, E2, E5, E6, E7, E1^E4, and E8^E2 (Figure 4-5, Table 4-3). In contrast, HPV-containing cervical cancers predominately express E6 and E7 oncogenes [109, 485]. Here, increasing expression of E6 and E7, particularly expression of E7-encoding E6*I transcripts, has been evidenced in HPV-associated cervical lesions with increasing histological grades, ranging from CIN1 to squamous cell carcinomas [109, 485].

This analysis demonstrated comparatively higher expression of early polycistronic transcripts spliced within the E6 ORF in HPV16-HTK than in HPV18-HTK, particularly at SD226^SA409 (HPV16) and SD233^SA416 (HPV18) splice junctions (Figure 4-5, Table 4-3). This major splice event disrupts transcription of full-length E6-encoding mRNAs to instead produce E6*I

isoforms, which have been described as principal E7-producing transcripts [114], and are the most abundant HPV mRNA isoforms in cervical cancers and tissues [491, 503, 529]. In addition, recent mass genomic studies have shown E6*I to be the most dominantly expressed E6 splice variant in HPV-OPSCC, where unspliced E6 are only detected at low level [417, 503].

Comparatively higher expression of E6*I isoforms and other prospective E7-encoding mRNA variants in the HPV16-HTK lines correlated with significantly higher production of E7 oncoproteins (Figure 4-9). Further, gene set enrichment analysis of host-HTK transcriptomes indicated comparably higher transcriptional activity of principal E7-induced pathways, including E2F targets and G2/M checkpoints. This suggests that higher HPV16-E7 levels were functionally inducing these cell cycle pathways to a significantly greater extent than HPV18-E7 in HTK and further indicate increased E7 oncoprotein activity in HPV16-HTK than in HPV18-HTK. Crucially, further deregulation of these anti-tumour responses could further promote uncontrolled cellular proliferation, immortalisation, and chromosomal instability, altogether aiding HPV16-driven carcinogenesis in the tonsils.

Given that HPV16 demonstrated higher pE activity, higher expression of unspliced E6-encoding transcripts was also expected. Yet, in contrast to E6*I levels, higher expression of unspliced E6-encoding transcripts was instead observed in the HPV18-HTK lines (Figure 4-5, Table 4-3). However, these expressional trends did not reflect differences in E6 activity, where western blotting for the E6-target, p53, indicated significantly higher E6 activity in HPV16-HTK#2 cells than in HPV18-HTK#2 cells and no significant change between HPV16/18-HTK#3 cells (Figure 4-10). However, it is possible that such discrepancies could reflect differences in

HPV16- and HPV18-E6 affinities for p53, where HPV16 E6 has been shown to more efficiently bind and promote degradation of p53 than HPV18 E6 [292, 293]. As such, p53 levels may not accurately reflect differences in E6 translation between HPV16- and HPV18-HTK lines. It will therefore be important to validate comparable E6 levels by qPCR and western blotting, to clarify whether differences in E6 mRNA and protein levels correspond with the Nanopore findings.

Considering the Nanopore findings, contrasting trends of unspliced- and spliced-E6 mRNAs suggests that HPV16 E6 ORF were subjected to greater levels of splicing than HPV18 E6 in HTK. This was exemplified by E6*I/E6 ratios, which were 2.7- and 5.2-fold higher in HPV16-HTK#2 and HPV16-HTK#3 lines than in donor-equivalent HPV18-HTK lines (Figure 4-6).

So why did more splicing of the E6 ORF occur in HPV16-HTK? Interestingly, differential gene expression analysis of the RNA-seq data demonstrated significantly lower expression of *EGF* in HPV16-HTK lines when compared with HPV18-HTK lines, showing $\log_2$ fold expressional decreases of -0.7 ($p = 0.0425$) (see Figure 5-12 and supplementary Table S-2). Studies conducted by Rosenberger, De-Castro Arce [161] showed that efficient SD226^SA409 splicing of the HPV16 E6 ORF corresponds to cellular EGF levels. Here, differential expression of EGF was suggested to alter activities of cellular splice factors, hnRNPA1, hnRNPA2, BRM and Sam68, which regulate E6 ORF splicing (see introductory section 1.2.2). Depletion of EGF promoted E6 splicing, resulted in increased expression of E6*I isoforms, coupled with increased translation of E7 and degradation of pRb. Conversely, a higher presence of EGF favoured higher expression of unspliced E6-encoding mRNAs and greater inhibition of p53

[161]. It is therefore possible that differences in *EGF* expression could have contributed to these variations in E6 splicing. As such, further validations of *EGF* expression and protein production should be carried out to confirm differential expression.

Altogether, these differences in pE activity, splicing of the E6 ORF and consequential variations of E7 levels between HPV16- and HPV18-HTK may be an important mechanistic distinction between the two virus-types in the tonsils. HPV-driven carcinogenesis is largely driven by cooperative E6 and E7 activities. However, a number of reports have advocated E7 as the more intrinsic driver and upholder of carcinogenesis in both HPV-HNSCCs and cervical cancers, where E6 was reported to play more of a secondary role in carcinogenesis [530-532]. Importantly, differences in E7 levels and activities suggest that HPV16 and HPV18 may differentially manipulate the host cellular environment during these early viral life cycle stages, which may help to explain contrasting oncogenic behaviours in the tonsils [31, 32, 282].

It is plausible that the variation in HPV16 and HPV18 E7 levels also contributed to differences in immune regulation in the tonsil keratinocyte lines, as discussed in Chapter 5. Both HPV16 and HPV18 E7 play critical roles in the regulation of cell-mediated immune response pathways, such as the type I IFN and NF-kB signalling pathways, which were shown to be strongly inhibited in HPV16-HTK in comparison to HPV18-HTK and controls (see Chapter 5, Figure 5-8). Such E7-mediated mechanisms include direct inhibition of IFN-type I-inducing interferon regulatory factors (IRF)1 and IRF9 [379, 380], inhibition of NF-kB regulatory elements like PCAF and IRFD1 [381, 382], and inhibition of the cytosolic cGAS-STING pathway, which induces downstream activation of IRF3 and NF-kB [372]. Importantly, downregulation of these pivotal

pathways correlated with higher levels of E7, possibly suggesting that HPV16 promotes higher expression of E7 in the tonsils to establish a more anti-immune environment, something that may support more persistent infections and aid long-term virus-driven tumourigenic alterations.

4.11.3 HPV16 and HPV18 demonstrated distinct differences in E8^E2 expression in primary HTK

HR-HPV E8^E2 are primary inhibitors of viral genome replication and transcription that are transcribed from a unique pE8 promoter, located within the E1 ORF [146, 147]. E8^E2 transcripts contain a short E8 mRNA fragment, expressed from the E1 ORF, which are spliced to the C-terminal of the E2 ORF [147-149]. The spliced removal of E2 N-terminal sequences renders E8^E2 unable to bind and recruit E1 helicases, as well as N-terminal-associating transcriptional mediators. However, E8^E2 retains the ability to bind with E2 C-terminal-specific nucleotide sequences, such as those contained within the origin of replication and early viral promoter [152, 533, 534]. Further, the E8 portion of E8^E2, which is highly conserved between HR-HPVs, strongly recruits host cellular corepressor complexes, NCoR/SMRT, which potently repress viral genome replication and gene expression [149-153]. E8^E2 therefore acts as an antagonist of E2 that blocks E2-mediated induction of viral genome replication and further inhibits transcription from the early viral promoter with greater efficiency [153, 535].

Distinct differences in E8^E2 expression were observed between our HPV16-HTK and HPV18-HTK lines, where the relative normalised expression of E8^E2-encoding transcripts was 3.4-

and 4.9- log_2 fold higher in the HPV18-HTK than in donor-equivalent HPV16-HTK lines (Figure 4-5, Table 4-3). This importantly reflected differences in pE8 promoter activity, where high pE8 activity in the HPV18 lines contrasted with a near absence of usage in the HPV16 lines, suggesting that the HPV16 and HPV18 pE8 promoters are being differentially regulated in the tonsils. However, little is currently known about how the pE8 and E8^E2 expression is regulated.

In HPV16-containing normal keratinocytes and immortalised keratinocytes, and in HPV18-containing U2OS (osteosarcoma) cell lines, disrupting mutations of E8^E2 (ΔE8^E2) have been shown to drastically increase HPV16 and HPV18 genome copy numbers by around 5- to 30-fold in comparison to cells containing wild-type E8^E2 [146-148, 153]. We therefore theorised that differential E8^E2 expression between the HPV16- and HPV18-HTK lines underlined differences in viral episome abundance. Expectedly, Southern blotting denoted a greater number of HPV episomes per cell in the HPV16-HTK lines, which was shown to be 2.9- and 13.1-fold higher than donor-equivalent HPV18-HTK#2 and HPV18-HTK#3 cells respectively (Figure 4-12).

These same E8^E2 mutant studies also showed 2.5- to 6-fold increases in E6 and E7 expression between ΔE8^E2- and wild-type E8^E2-containing cells, denoting greater pE activity in absence of E8^E2 [146-148, 153]. As discussed earlier (4.11.1), pE activity was found to be around 1.4 – 1.6-fold higher in HPV16-HTK lines than in donor-equivalent HPV18-HTK lines, contributing to proportionately higher E7 expression, along with significantly increased E7 protein production and activity (4.11.2). It is therefore likely that differing levels of E7 were

also associated with expressional differences of E8^E2. Put together, these data indicate greater E8^E2-mediated regulatory control of viral DNA replication and early viral transcription in HPV18-HTK than HPV16-HTK.

Another principal pE regulator to consider is E2. HPV16-HTK were shown to express greater levels of mRNAs containing the E2 ORF than HPV18-HTK, however almost all of these E2-encoding mRNA species started at the pE (Table 4-3). Importantly, such pE-derived transcripts have been shown to be inefficient translators of E2 in comparison to mRNAs transcribed from the pL, perhaps indicating little in the way of overall E2 translation [505, 506]. However, it is still possible that these observed differences in E2 expression could be biologically relevant, so it will be important to confirm and compare levels of E2 by western blotting.

As discussed in introductory section 1.2.3, higher levels of E2 force displacement of cellular transcription factors and pE-transactivators such as SP1 and TBP from the viral LCR. This inhibits transcription from the viral pE, instead favouring the E2-mediated recruitment of E1 and viral genome replication [132-137]. Yet, in our HTK lines, we observed a comparable increase not just in viral genome copy numbers but also of early viral transcription in the HPV16-HTK lines. Put together, this seems to suggest that differential expression of E8^E2 was likely playing a bigger role in the regulation of these processes than E2. However, it is still possible that higher expression of E2 in the HPV16-HTK lines could have contributed to differences in episome copy numbers via viral genome replication and segregation/tethering of the viral genome to host chromatin, as discussed in section 1.3.4.

HPV genomes are maintained as extrachromosomal episomes in low copy numbers, usually between 10 – 200 copies per cell [536]. However, in differentiated suprabasal epithelial layers, viral genome amplification produces thousands of viral episomes per cell [537]. Maintaining low viral copy numbers and low levels of viral transcripts in undifferentiated cells functions as a potential immune avoidance strategy by limiting production of large quantities of the viral DNA, RNA, and proteins. It's possible that these differences in E8^E2 expression and E7 production could reflect varying immune avoidance strategies of HPV16 and HPV18 in the tonsils. For HPV18, comparatively higher E8^E2 expression may promote a less immune-stimulatory environment via greater limitation of viral episomes and early viral transcripts. In comparison, reduced E8^E2-associated regulation by HPV16 could instead favour higher E7 activity to directly dysregulate central elements of the host immune response, as has been later evidenced in Chapter 5. Given that HPV16 makes up the vast majority of HPV-OPSCC, we theorise that this more direct method of immune regulation could better promote viral persistence in the tonsils, perhaps supporting long-term tumourigenesis [31, 32, 282].

To better understand the consequences of differential E8^E2 expressions, it would be interesting to investigate the effects of mutational inactivation of E8^E2 HPV-HTK lines. Mutational inactivation of HPV18 E8^E2 could provide insight into how this viral protein alters HPV18 transcriptome organisation and DNA replication in the tonsils, and how this compares with that of HPV16-HTK, which supports very little E8^E2 expression. It will also be interesting to examine E8^E2 levels in HPV16- and HPV18-HTK following cellular differentiation. Are E8^E2 levels maintained at higher levels in HPV18-HTK than HPV16-HTK during late viral life cycle stages?

4.11.4 HPV16 and HPV18 integration events in primary host-HTK genomes

Current understanding of HPV integration in the oropharynx is principally based on late-stage tumour biopsies. As such, analysis of virus-host fusion transcripts, as identified by Nanopore, was carried out to assess and compare HPV16 and HPV18 integration events in primary HTK lines, which at the point of sequencing, at 30 – 40 PD, had demonstrated primarily episomal HPV16/18 genome organisation (section 3.2.4, Figure 3-4). Surprisingly, multiple virus-host fusion transcripts were detected within each of the HPV16/18-HTK lines, even within HPV16-HTK#2 cells, which were shown to maintain viral episomes in long term cell culture (>150 PD) and support the complete virus life cycle in organotypic raft culture (Chapter 3, Figure 3-4). These findings suggest that viral integration could be a frequent event in HTK *in vitro* that occurs shortly after viral transfection.

Analysis of human mRNA sequences of virus-host fusion transcripts was performed to pinpoint possible integration breakpoint locations in host chromatin. However, it should be noted that this may not always accurately reflect the exact site of viral DNA incision, as integration into intronic DNA sequences could result in splicing between viral and host exonic RNA. It is also important to note that mRNA analysis only allows for identification of virus-host fusion transcripts which had integrated into transcriptionally active areas of the host genome.

A number of potential integration events were identified at previously characterised integration hotspots, or at specific loci either proximal to, or within genes which are frequently mutated in HPV-associated cancers (Table 4-4). Notable integration events, occurring at

previously documented integration hotspots, included the tumour suppressors *TP63* (HPV18-HTK#2 and #3) [399, 538, 539], *WWOX* (HPV16-HTK#3) [540-542], and *TRAF3* (HPV16-HTK#2). Importantly, mutational inactivation of *TRAF3* occurs in approximately 25% of integrated and non-integrated HPV-HNSCC and is thought to result in constitutive activation of NF-kB, promoting cell survival and proliferation [410, 417, 421]. Other identified cancer-related genes included *FBXW8, RTKN, FGF14, NFYC, CDH13*, and *ROCK1* (Table 4-4). However, examination of the Illumina-based RNA-Seq data showed no obvious changes in expression of any of these genes, including *TP63*, *WWOX,* and *TRAF3* (data not shown). As such, it seems unlikely that these particular integration events were directly promoting transformation of the cells.

While these integration events did not appear to directly dysregulate host gene expression, virus-host fusion transcripts were shown to primarily encode E6* transcripts, along with some unspliced E6 transcripts, likely elevating total E7 and E6 levels within the HPV-HTK lines. This was particularly true for HPV18-HTK#3 and HPV16-HTK#2, which demonstrated high expression of virus-host fusion transcripts, where HPV DNA appeared to integrate into a number of highly transcriptionally active areas of host HTK#2 and HTK#3 genomes. It is therefore possible that such integration events could drive overexpression of E7 and E6 during early infection of the tonsils, in both integrated and episome maintaining cells, which may promote tumourigenesis via aberrant cell cycle progression and increased genomic instability.

4.12 Conclusions

In conclusion**,** we identify Nanopore analysis has allowed for accurate characterisation and quantification of full-length, native HPV transcripts in our undifferentiated HPV16/18-HTK

models. Comparison of HPV16 and HPV18 transcriptome organisation in the primary HTK lines revealed divergent transcriptional tendencies during the early stages of the viral life cycle that underlined differences in HPV16/18 transcription profiles. In comparison to HPV18-HTK lines, HPV16-HTK presented comparatively higher activity of the major early promoter and greater splicing of the E6 ORF, resulting in comparably higher expression, translation, and activity of E7.

Several underlying mechanisms were identified that could have contributed to these differences in pE activity and splicing, including alternative expression of the transcriptional regulators, *EGF*, *FOS*, and *FOSL1*, as well as distinct expressional differences of the viral transcription and DNA replication regulator, E8^E2. Importantly, differences in E7 levels and activities suggest that HPV16 and HPV18 may differentially manipulate the host cellular environment during these early viral life cycle stages and could help to explain contrasting oncogenic behaviours in the tonsils. As such, RNA-seq was carried out to characterise and compare HPV16- and HPV18-induced changes of host-HTK transcriptomes (Chapter 5).

4.13 Future studies

Additional HPV16- and HPV18-transfected donor models are required to support the findings of this chapter. Given that HPV16/18-HTK#3 cells did not support the viral life cycle, establishment of additional HPV-HTK lines should allow for examination of HPV16 and HPV18 transcription profiles following differentiation. Here, Nanopore could be carried out on HPV16- and HPV18-HTK models following organotypic raft culture or methylcellulose-mediated differentiation, as was performed on HPV-HFK models in the Parish lab [98]. This

will allow for a better understanding of how HPV16 and HPV18 are manipulating the host-HTK throughout the viral life cycle and may provide important additional insights into differences in viral transcriptome organisation that may contribute to HPV16-induced transformation. In particular, it will be interesting to examine comparative E7 and E8^E2 levels to see whether these are maintained at differing levels, as in the undifferentiated cells. In addition, as with E6 and E7, it will also be important to further examine and validate expression and translation of other early viral genes, such as E1, E2, E1^E4 and E5.

These studies have provided preliminary evidence that HPV16 and HPV18 differentially regulate expression of transcription factors, *EGF*, *FOS,* and *FOSL1,* which are crucially required for activation of the pE and splicing of the E6 ORF. Initial qPCR- and western blot-based validations of these genes are required to confirm the RNA-Seq findings. It will then be important to further explore how these transcriptional mediators are being differentially regulated by HPV16 and HPV18 in the tonsils.

Chapter 5: High-risk HPV-mediated transcriptional reprogramming of primary tonsil keratinocytes

5.1 Introduction

Transcriptional reprogramming of host keratinocytes is a key evolutionary function of high risk (HR) HPV, which creates a suitable environment supportive of the virus life cycle. In particular, the viral transcription regulators, E6 and E7 have been identified as key modulators of host gene expression that alter multiple cellular pathways, including host-cellular proliferation, differentiation, and metabolism programmes [543-545]. Further, E6 and E7-mediated regulation of the host transcriptome plays an important role in viral persistence by adjusting a range of antiviral pathways such as apoptosis and immune regulation. Incidentally, increased viral persistence along with perturbed anti-tumour mechanisms and dysregulated proliferation are thought to significantly contribute to disease progression [264, 546, 547].

In Chapter 4, Nanopore-based investigations of HPV16 and HPV18 transcription profiles of our primary HTK donor-lines demonstrated comparatively higher expression, translation, and activity of E7 in HPV16-HTK lines than in donor-equivalent HPV18-HTK lines. This was partly evidenced by significantly increased transcriptional activities of two central E7-target pathways involved in cell cycle progression, ‘E2F targets’ and ‘G2/M checkpoints’, indicating that HPV16 had altered these transcriptional programmes to a greater extent than HPV18. It

is therefore possible that HPV16 and HPV18 could have been differentially regulating other aspects of the host-HTK transcriptome that are significant for carcinogenesis.

Global gene expression studies in the oropharynx have so far focused on tumour biopsies and cancer cell lines, often comparing HPV-positive and HPV-negative tumours to identify HPV-specific mutations or comparable gene expression changes. Together these studies have highlighted clear differences in molecular pathogenesis between HPV-positive and HPV-negative HNSCC, revealing alternative gene expression profiles, mutations, and epigenetic modifications (see introductory section 1.4.7) [400, 410, 417, 548-552]. Further, global expression studies have revealed some specific molecular differences between HPV-HNSCC and cervical cancers, which indicate potential differences in the natural history of progression between these two body sites (introductory chapter 1.4.7) [420].

Given that the vast majority of HPV-OPSCC are first detected at advanced disease stages, examination of host genome expression profiles within pre-malignant, HPV-infected oropharyngeal tissues has proved difficult due to a lack of HPV-positive pre-lesion biopsies [57, 553]. As a result, a significant gap in our understanding of the natural history of HPV-mediated disease in the oropharynx remains.

Study of primary HPV-oropharyngeal models is an important next step towards a greater understanding the HPV infectious cycle and natural history in the oropharynx. As such, RNA-sequencing (RNA-Seq) of our HPV-HTK lines was carried out to characterise and compare HPV16- and HPV18-mediated reorganisation of the primary tonsil keratinocyte transcriptome,

the primary site of HPV-induced oncogenic progression in the oropharynx [31, 269]. Firstly, global gene expression changes were assessed by comparing transcriptomes of the HTK donor cells with transcriptomes of HPV16- and HPV18-HTK cells. Next, host transcription profiles were compared between HPV16- and HPV18-HTK cell lines to evaluate differential HPV-type specific gene expression changes. We hypothesise that HPV16 and HPV18 differentially induce transcriptional reprogramming in the tonsils, leading to differences in viral persistence and disease progression.

5.2 Comprehensive analysis of HPV16 and HPV18-altered host HTK transcription profiles

Global gene expression of the two donor primary keratinocytes HTK#2 and HTK#3, and the donor cells transfected with HPV16 or HPV18 was examined by RNA-sequencing (RNA-Seq) using a NextSeq 550 (Illumina) to sequence polyA+ selected mRNA. This is a next generation, short-read sequencing technique that allows for accurate quantification of mRNA transcripts. RNA-Seq was performed on HTK isolated from tonsil tissue and the HPV containing cells, all grown in monolayer tissue culture and on a feeder layer of murine fibroblasts (see methods section 2.1.2.1). To better assess HPV-driven transcriptome reorganisation as in a typical episome maintaining infection, Southern blotting was performed to ensure that each HPV-HTK cell line contained episomal HPV DNA at the point of sequencing, as discussed in Chapter 3, section 3.2 (Figure 3-1 – Figure 3-4). Therefore, isogenic HPV-HTK#1 donor cells were excluded due to very early loss of HPV16 episomes during the establishment of the cell lines and likely integration of HPV16 DNA.

RNA was extracted from untransfected and HPV16- and HPV18-containing HTK#2 and HTK#3 cell cultures. Sample preparation, quality control, and sequencing steps are described in methods section 2.3.5. RNA sequence alignment and mapping was carried out by Dr Boris Noyvert using the RNA-Seq alignment tool, STAR [554], which was also used to calculate read counts. mRNA sequences were aligned to the hg19 genome and DESeq2 was used to normalise mRNA read counts and conduct differential gene expression analysis. The read depth and proportion of fragments aligned to the HPV and host genomes in each sample is detailed in Table 5-1. Regularised log transformation was calculated to appropriately adjust for any non-linear read count distributions and fold change was presented on a $\log_2$ scale.

5.3 Principal component analysis

Prior to differential expression analysis, principal component analysis (PCA) was performed by Dr Boris Noyvert to visualise and compare global gene expression variation between each HTK donor and HPV-HTK line (Figure 5-1). The first principal component (PC1) and second principal component (PC2) represent the first and second most variable linear combination of variables. Here, PC1 accounted for 40% of variation in overall gene expression and PC2 accounted for 30% of variation. In total, PC1 and PC2 represented 70% of the overall assessed variability of gene expression between the six HTK lines.

Altogether, these data showed distinct variation in global gene expression between HTK#2 and HTK#3 donor cells indicating significant differences in gene expression profiles between the two donor cell populations. In addition, no similar patterns of variation were observed between the two assessed untransfected-HTK donor cells when compared against their HPV16

and HPV18-containing counterparts. While HTK#2 and HPV18-HTK#2 clustered closely together, indicating similar global gene expression profiles, HTK#2 and HPV16-HTK#2 cells displayed a much wider variance in PC2. These data suggest that HPV16 induces a much greater global gene expression change than HPV18 in this HTK donor line. In contrast, little variation was observed between the isogenic HPV16-HTK#3 and HPV18-HTK#3. However, neither HPV-HTK#3 lines closely correlated with untransfected-HTK#3, indicating that both HPV16 and HPV18 had greatly altered the HTK#3 transcriptome in a similar manner. Together, these data indicate that HPV-mediated transcriptional reorganisation was variable between each donor and that the two donors had distinct differences prior to HPV establishment.

Table 5-1: RNA-seq read mapping statistics.

This table overviews the total number of reads mapped during RNA-sequencing of each HTK and HPV-HTK models. In addition, these data show the total number of mRNA reads that mapped to human, HPV16 and HPV18 genomes. The vast majority of reads (96.9-97.7%) that mapped to the human genome were mapped to exons.

Cell line	Total number of fragments, millions	Proportion of total mapped fragments to hg19: % of total	Proportion of fragments mapped to exons: % of total mapped	Proportion of fragments mapped to K02718.1 (HPV16): % of total mapped	Proportion of fragments mapped to AY262282.1 (HPV18): % of total mapped
HTK#2	78.2	85.8	97.5	0	0
HPV16-HTK#2	76.7	90.3	97.6	0.047	0
HPV18-HTK#2	72	88.1	97.4	0	0.0299
HTK#3	75	87.3	97.6	0	0
HPV16-HTK#3	69.5	87.3	97.7	0.019	0
HPV18-HTK#3	74.9	87.9	96.9	0	0.0489

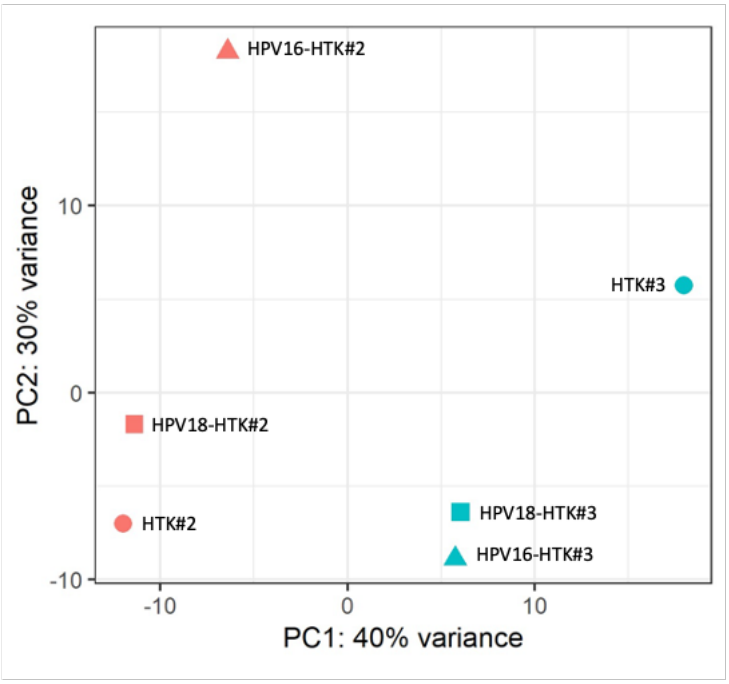


Figure 5-1: Principle component analysis.

Demonstrates overall variance of global gene expression profiles between HTK, HPV16-HTK and HPV18-HTK models. Produced by Boris Noyvert.

5.4 Overview of differential gene expression in HPV16-HTK and HPV18-HTK

In order to effectively examine and compare HPV16- and HPV18-induced transcriptional reprogramming in HTK cells, evaluation of the RNA-Seq data was carried out via three main comparisons: (1) comparing HPV16-HTK cells against untransfected-HTK cells, (2) comparing HPV18-HTK against untransfected HTK cells, and (3) comparing HPV16-HTK against HPV18-HTK cells. For each aforementioned comparison, volcano plots were prepared to visualise differences in expression of all genes quantified by RNA-Seq (Figure 5-2 and Figure 5-3). Volcano plots are a form of scatter plot that demonstrate statistical significance (p value) versus magnitude of change (regularised $\log_2$ fold change) for each assessed gene.

HPV16-HTK and untransfected-HTK comparisons showed a total of 1583 significantly differentially expressed genes ($p < 0.05$). Of which, 983 were significantly upregulated in HPV16-HTK (62.1%) and 600 were significantly downregulated (37.9%) (Figure 5-2a). When comparing HPV18-HTK and untransfected-HTK lines, a total of 1471 genes were shown to be significantly differentially expressed, consisting of 835 upregulated genes (56.8%) and 636 downregulated genes (43.2%) (Figure 5-2b). Together, these data broadly demonstrate that both HPV16 and HPV18 extensively altered host HTK gene expression profiles, and that both virus-types upregulated expression of more genes than they downregulated.

The similarities and differences in differential gene expression following HPV16 and HPV18 establishment in the two HTK donors are summarised in a Venn diagram shown in Figure 5-3a. Overall, 909 genes were shown to be modulated by both HPV16 and HPV18 when compared against untransfected-HTK cells. A total of 674 unique gene expression changes were observed

between HPV16-HTK and control-HTK cells, accounting for 42% of all genes that were significantly altered by HPV16. A further 562 unique gene expression changes were observed between HPV18-HTK and control-HTK cells (39%). Altogether, this data shows that while HPV16 and HPV18 significantly altered expression of many of the same genes in both HTK donors, these two oncogenic viruses also uniquely regulated a large number of different genes.

To better evaluate the HPV16 and HPV18 specific targets, a volcano plot was prepared to visualise these genes (Figure 5-3b). In total, expression of 536 genes was shown to be significantly altered between HPV16-HTK and HPV18-HTK. Of this, expression of 303 genes were significantly higher in HPV16-HTK when compared with HPV18-HTK, and expression of 234 genes were significantly lower in HPV16-HTK than HPV18-HTK.

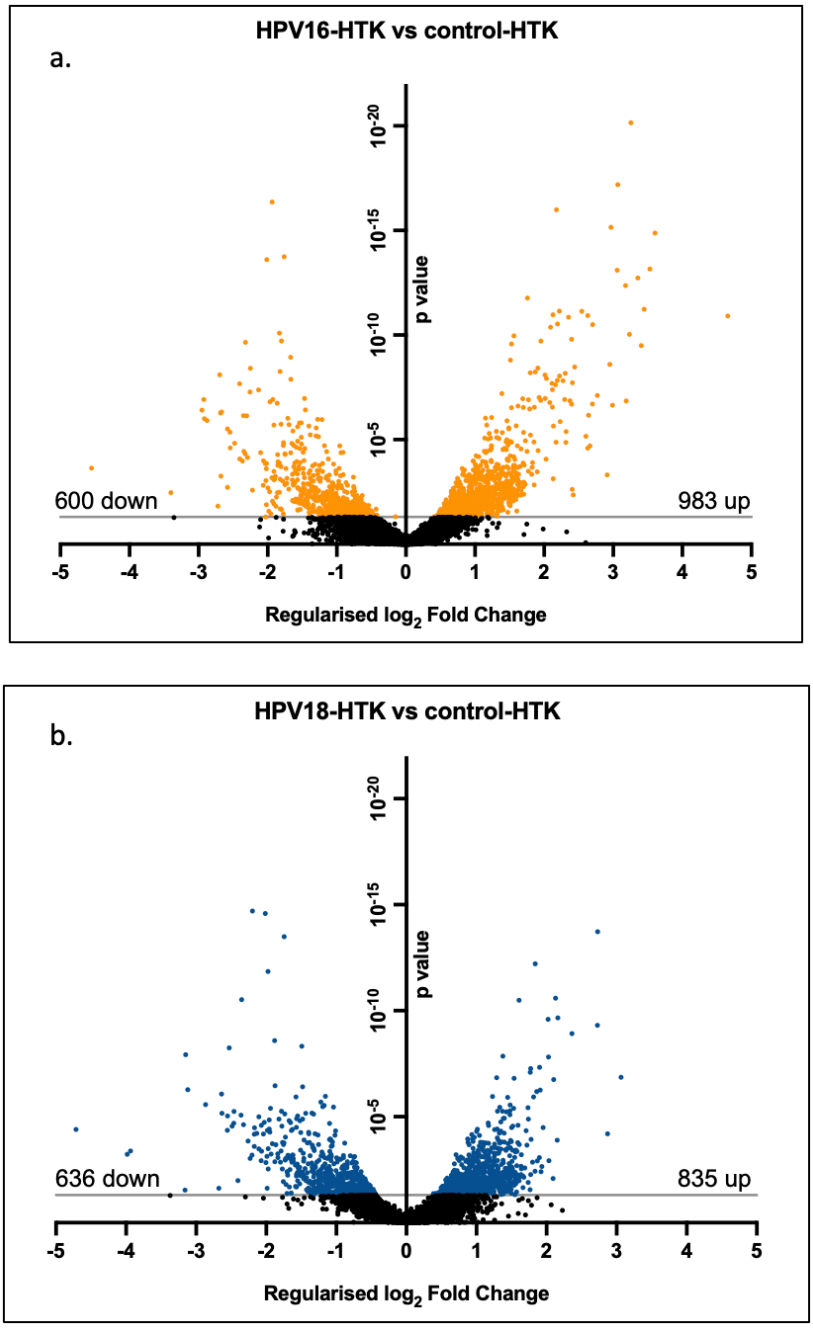
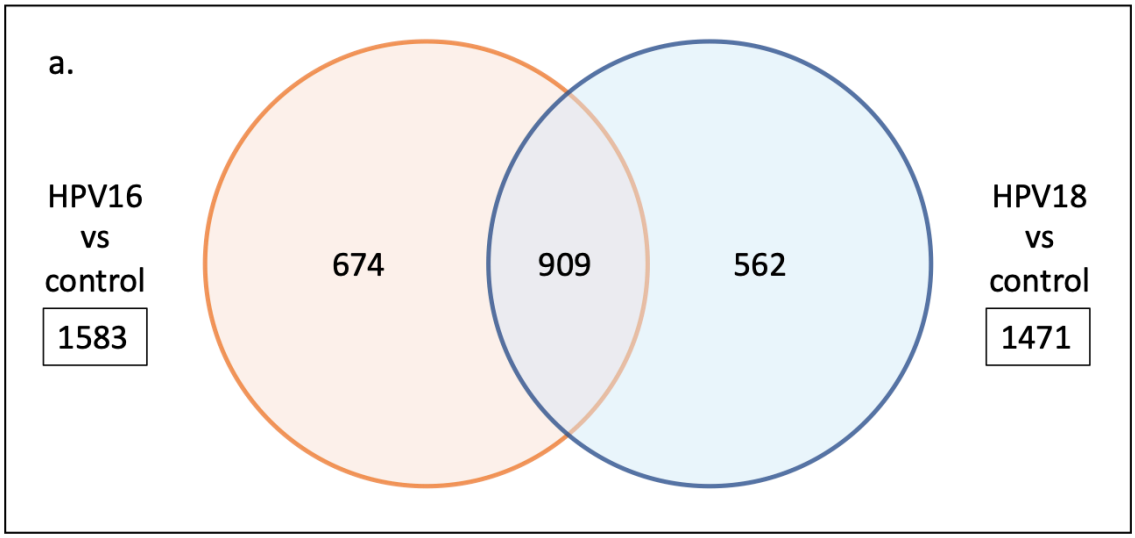


Figure 5-2: Comprehensive differential gene expression analysis of HPV- and untransfected-HTK

Volcano plots compare p-values against regularised log_2 fold change to fully visualise overall differences in gene expression between HPV-HTK cells and controls. **a**. compares HPV16-HTK and control-HTK cells. Here, HPV16 was shown to significantly downregulate expression of 600 genes and upregulate 983. **b**. compares HPV18-HTK and control-HTK cells. HPV18 comparatively downregulated 636 genes and upregulated 835 genes. The y-axis shows log^{10} of the p-value, where the grey lines mark significance ($p < 0.05$). All significantly differentially expressed genes are coloured orange (HPV16) or blue (HPV18).



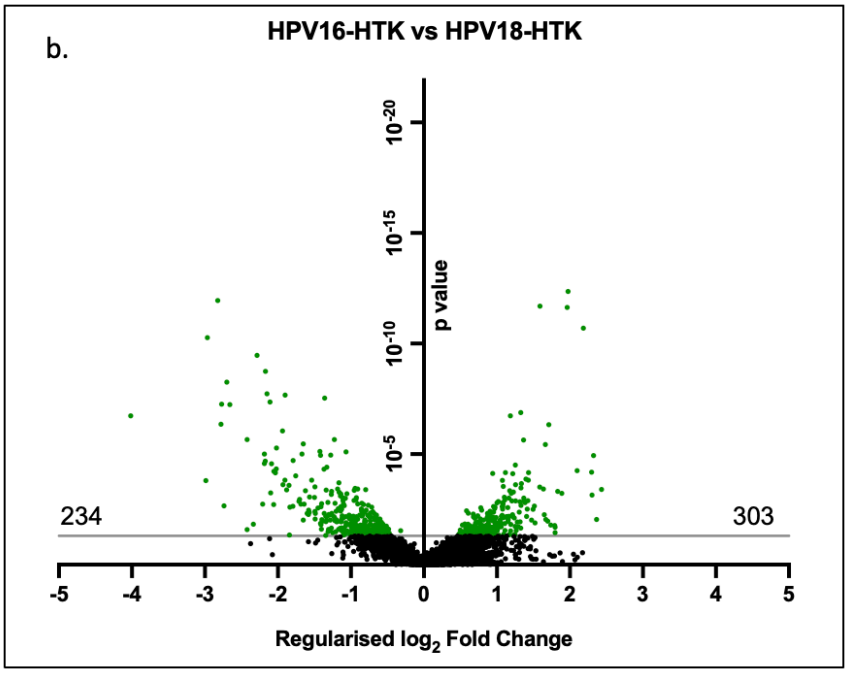


Figure 5-3: Comprehensive differential gene expression analysis of HPV16- and HPV18-HTK lines

a. Venn diagram broadly compares overlap of genes that were significantly differentially expressed between HPV-HTK lines and controls, showing that HPV16 and HPV18 both altered expression of 909 same gene targets but exclusively altered the expression of 674 and 562 genes respectively. **b**. Volcano plot compares differences in gene expression between HPV16-HTK and HPV18-HTK lines. Expression of 234 genes was shown to be significantly lower in HPV16-HTK than HPV18-HTK lines and expression of 303 genes were shown to be significantly higher in HPV16-HTK than HPV18-HTK. The y-axis shows $\log^{10}$ of the p-value, where the grey line mark significance ($p < 0.05$). All significantly differentially expressed genes are coloured green.

5.5 Heat maps and hierarchical cluster analysis of top differentially expressed genes in HPV16-HTK and HPV18-HTK lines

Heat maps were generated by Boris Noyvert, using Pheatmap R package to visualise expression of the top 50 most significantly differentially expressed genes between HPV16-HTK and untransfected-HTK models, as well as between HPV18-HTK and HTK models (Figure 5-4a and b). Here, unsupervised hierarchical cluster analysis grouped genes and HTK models into clusters according to their expression, allowing for easier comparison of gene set clusters between different defined groups.

When comparing HPV16-HTK and untransfected-HTK models, clustering was observed into two distinct groups (Figure 5-4a). One where genes were significantly upregulated in HPV16-HTK when compared with untransfected-HTK, and a second group where genes were significantly downregulated. Of the most significant genes, 39 were upregulated and 11 were downregulated, indicating that HPV16 was having an overall up-regulatory effect on these genes. The regularised $\log_2$ fold expression ranged from -2.7 to 4.7. In contrast, 22 of the 50 genes were comparably upregulated in HPV18-HTK cells in comparison to the control and 28 genes were downregulated (Figure 5-4b). Regularised $\log_2$ fold expression ranged from -5.2 to 6.5.

Dendrograms at the top of each heatmap compared overall gene expression similarity between HPV16-HTK, HPV18-HTK and control models. In both HPV vs control heat maps, gene expression was shown correlate more closely between HPV16-HTK and HPV18-HTK than to

control-HTK models. A total of 18 genes overlapped between the two heat maps (Figure 5-4a and b). These consisted of *KLK12, UCHL1*, *RHCG*, *PLCE1*, *ALDH1L1*, *CACNB4*, *CDKN2A*, *RP11-94L15.2*, *PEG10*, *COCH*, *LAMA1*, *CAMK4*, *PRTFDC1*, *NEFH*, *SLC35G2*, *LHX6*, *PRSS27*, *MEST*; the direction of regulation being the same for both virus types. This showed that HPV16 and HPV18 were strongly regulating the expression of a group of similar host genes. This included upregulated expression of the tumour suppressor gene, *CDKN2A* (p16), along with a number of *CDKN2A*-associated genes, such as *NES* and *UCHL1* in HPV16-HTK models and *NOTCH3*, *MAF*, and *UCHL1* in the HPV18-HTK models (Figure 5-4a and b). Greatly altered expression of *CDKN2A* was expected as HR-HPV E7-mediated inactivation of Rb leads to overexpression of *CDKN2A*, which is commonly used as a biomarker for identification of HPV positive OPSCC [555, 556].

Heat maps comparing the top 50 most significantly altered genes between HPV16- and HPV18-HTK models were generated (Figure 5-4c). Of the 50 genes, expression of 33 was significantly higher in HPV16-HTK lines when compared with HPV18-HTK, and 17 were significantly lower. Regularised $\log_2$ fold expression ranged from -4.0 to 2.3. Dendrograms showed that gene expression correlated more closely between HPV18-HTK and control-HTK than between HPV16-HTK and control cells, indicating that HPV16 was altering overall gene expression to a greater extent than HPV18 in this gene set.

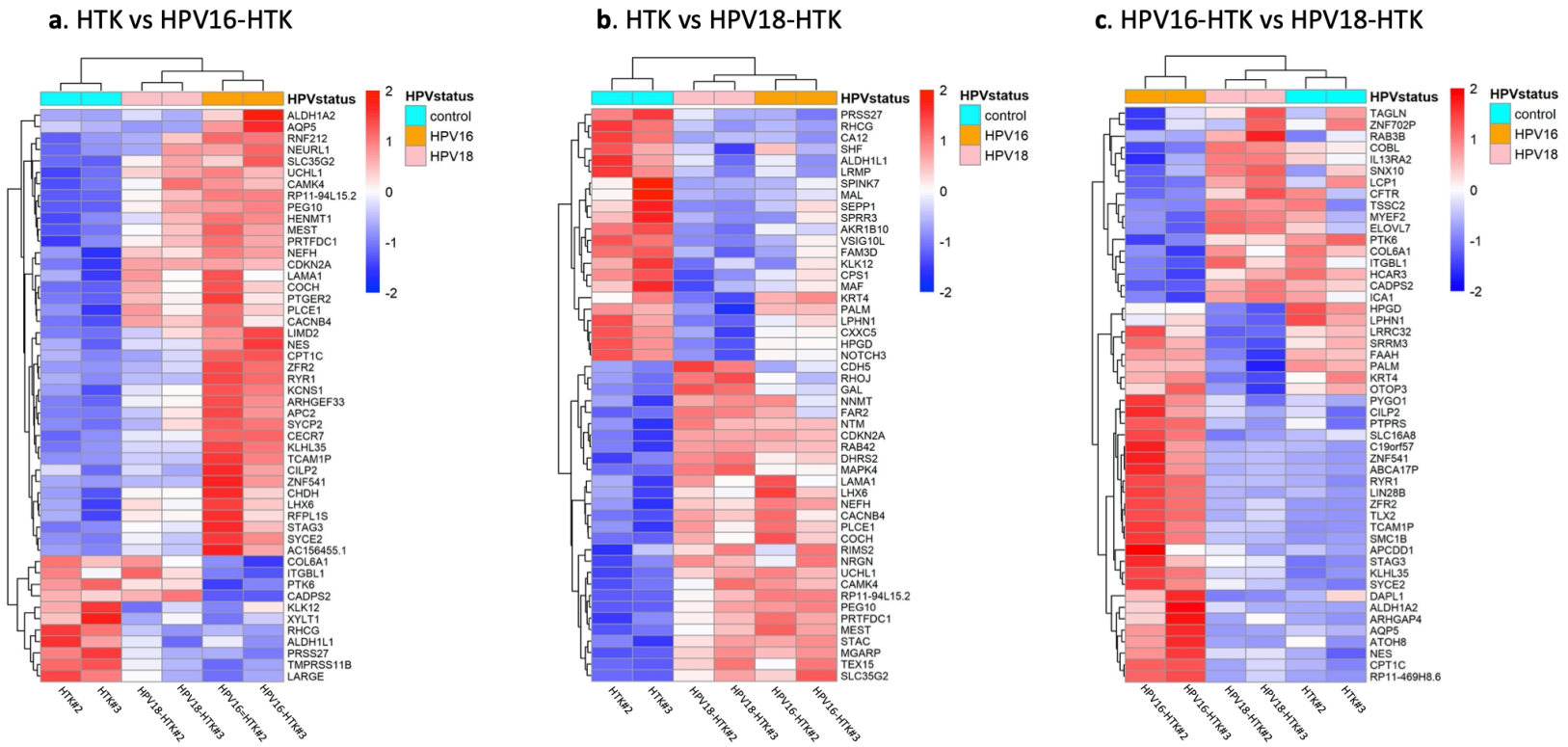


Figure 5-4: Heat maps and hierarchical clustering analysis

Demonstrates gene expression changes for the top 50 most significantly altered genes for the following model comparisons: **a**. HPV16-HTK vs control-HTK, **b**. HPV18-HTK vs control-HTK, and **c**. HPV16-HTK vs HPV18-HTK. Dendrograms, located at the top and left-hand side of each heat map indicate similarity between hierarchical clusters and HTK models. The colour indicates the z-score of log-transformed normalised expression values, as shown on the column on the right of each heat map. Here, deeper red or blue colouration is representative of larger differences of expression from the gene mean, where red indicates upregulated expression, and blue indicates downregulated expression. Heat maps were prepared by Boris Noyvert.

5.6 Identification of alternatively regulated biological pathways by gene set enrichment analysis (GSEA)

Comprehensive differential gene expression analyses showed that HPV16 and HPV18 significantly altered expression of around 1,500 host genes (section 5.2). To better understand the biological consequences this had on the host keratinocyte, GSEA was carried out to identify the top alternatively regulated biological pathways between HPV-HTK and control-HTK, as well as between HPV16-HTK and HPV18-HTK models.

This analysis was performed by Dr Boris Noyvert using GAGE Bioconductor software (https://bioconductor.org/). GSEA was used to determine the statistical significance and concordance of overall gene expression changes in pre-defined gene sets between the HTK and HPV-HTK lines. The gene sets used for these analyses were obtained from the Molecular Signatures Database (MSigDB; https://www.gsea-msigdb.org/gsea/msigdb/). MSigDB Hallmark pathways consist of refined gene sets that accurately portray select biological processes, allowing for recognition of altered biological pathways in which the differentially expressed genes participate [557]. For these analyses, multiple testing corrections were first carried out to adjust p-values, producing q-values, which better estimate and adjust for positive false discovery rates.

5.6.1 GSEA of HPV16-HTK and HPV18-HTK in comparison to untransfected-HTK

In the first instance, GSEA was used to identify alternatively regulated biological pathways between HPV16-HTK and untransfected-HTK controls. In total, 16 gene sets were shown to be

significantly differentially regulated between the two model systems, where $q < 0.05$ (Figure 5-5a). These generally clustered into two main groups: (1) cell cycle progression pathways and (2) immune signalling pathways. When comparing HPV18-HTK and control-HTK, 16 gene sets were shown to be significantly altered (Figure 5-5b). These largely fell into one of two groups: (1) cell cycle progression and (2) cell metabolism related pathways. In total, nine of the same gene sets were similarly alternatively regulated in both HPV16-HTK and HPV18-HTK lines when compared with control-HTK. These mainly consisted of cell cycle progression and DNA damage response pathways, including E2F targets, G2M checkpoint, MYC targets, KRAS signalling, mitotic spindle, and p53 pathways, all of which were up/downregulated in the same direction. Other comparable gene sets included spermatogenesis, EMT and IFN-α signalling.

Notably, a number of immune pathways were significantly differentially regulated in HTK lines following HPV16 genome transfection. These included interferon gamma (IFN-γ), interferon alpha (IFN-α), inflammatory response and complement signalling pathways, all of which were downregulated in HPV16-HTK (Figure 5-5a). All four of these pathways play important roles in viral detection and clearance in tonsil keratinocytes [354, 558, 559]. In contrast, only IFN-α signalling was shown to be significantly downregulated in HPV18-HTK when compared with control-HTK (Figure 5-5b). Further, GSEA indicated an overall increased expression of another immune response gene set, TNF-α signalling via NF-kB, in HPV18-HTK lines when compared with untransfected-HTK, although this was just short of significance, where $q = 0.0596$ (data not shown). Altogether, these data reveal that HPV16 was a stronger negative regulator of immune signalling pathways in the host tonsil keratinocyte in comparison to HPV18.

HPV18-HTK vs control-HTK comparisons showed significant alternative expression of several metabolism related pathways, including mTORC1 signalling, haem metabolism, xenobiotic metabolism and adipogenesis (Figure 5-5b). Other than mTORC1 signalling, which was showed to be significantly upregulated in HPV18-HTK lines, overall expression of all other metabolism-related pathways was significantly reduced. No similar differences were observed when comparing HPV16-HTK and control-HTK lines, potentially highlighting metabolism dysregulation as a HPV18-specific function in HTK.

5.6.2 GSEA of HPV16-HTK in comparison to HPV18-HTK

GSEA of MSigDB gene sets between HPV16-HTK and HPV18-HTK models identified 19 pathways that were differentially expressed between the two virus-type models (Figure 5-6). The most notable difference were differences in immune signalling/response pathways. Eight immune pathways were alternatively regulated, including the top five most significantly differentially expressed pathways - IFN-γ, inflammatory response, allograft rejection, interferon alpha, and TNF-α signalling via NF-kB. In addition, significant differential expression of complement, IL2-STAT5 signalling, and IL6-JAK-STAT3 signalling pathways was also observed. Expression of the eight immune pathways was significantly lower in HPV16-HTK than HPV18-HTK.

While the IFN-α response was previously shown to be downregulated by both HPV16 and HPV18 when compared with untransfected-HTK (Figure 5-5), direct GSEA comparisons demonstrated significantly lower expression of IFN-α response in HPV16-HTK than HPV18-HTK

(Figure 5-6). This suggests that HPV16 is having an even greater inhibitory effect on IFN-α signalling than HPV18. Altogether, these data strongly indicate that HPV16 was downregulating host immune responses to a much greater extent than HPV18. This could have important consequences for establishment and persistence of HPV16 infection in tonsil keratinocytes.

In addition, these data demonstrated significant differential regulations of five cellular metabolism pathways, including mTORC1 signalling, fatty acid metabolism, xenobiotic metabolism, oxidative phosphorylation, cholesterol homeostasis, and adipogenesis. Other differentially regulated, and cancer-associated pathways included KRAS signalling, epithelial mesenchymal transition (EMT), hypoxia, and spermatogenesis pathways.

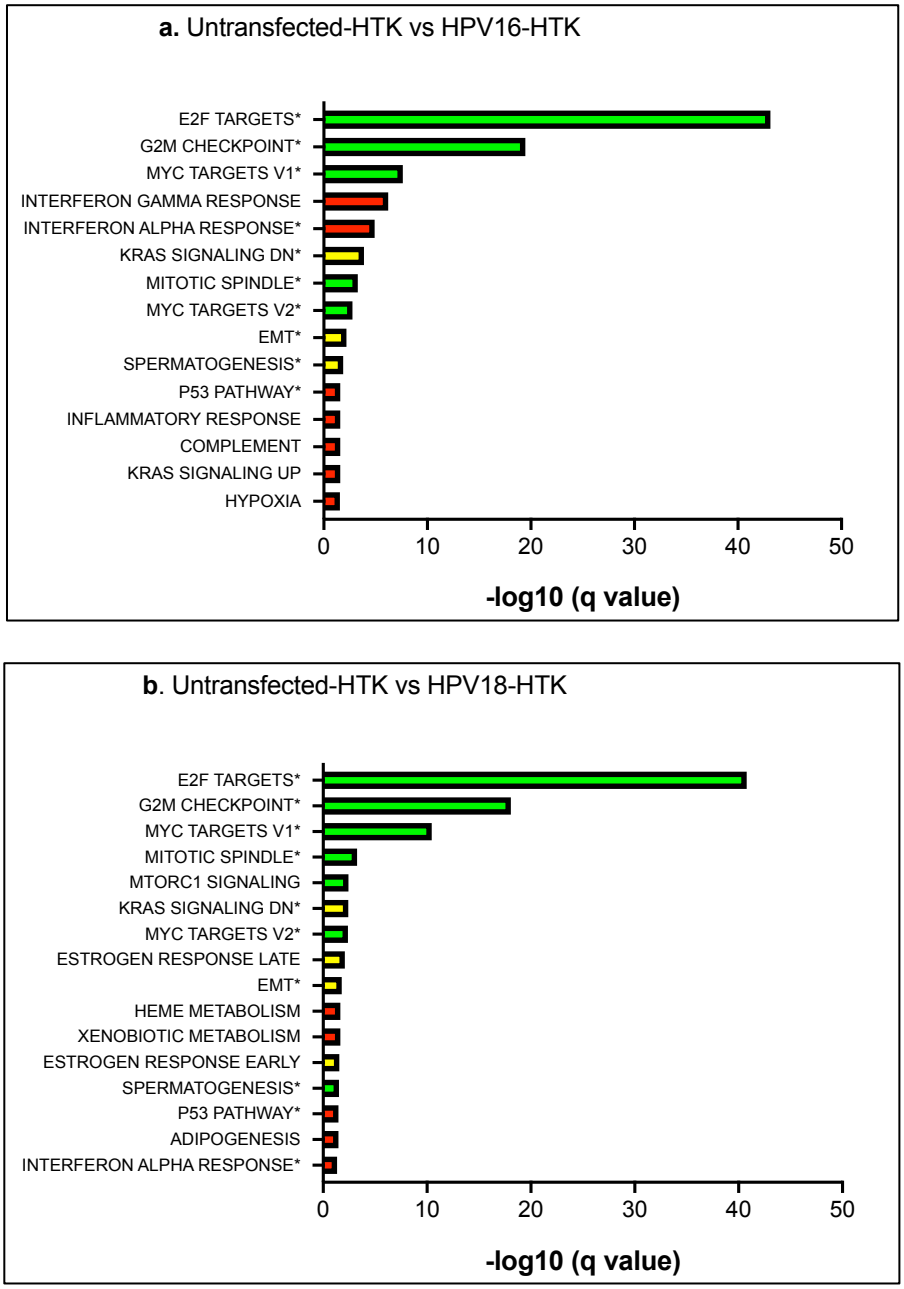


Figure 5-5: GSEA comparing HPV16/18-HTK against untransfected-HTK controls

Demonstrates all significantly differentially regulated MSigDB Molecular Hallmark gene sets between **a.** HPV16-HTK and untransfected-HTK models, and **b.** HPV18-HTK and untransfected-HTK models, where $q < 0.05$ (equivalent to -log10 >1.3). Pathways are ordered in descending significance, where E2F targets was identified as the most significantly alternatively expressed gene set in both analyses. Relative differences in overall expression of each pathway are signified by colour. Green bars indicate gene sets that were upregulated in HPV-HTK in comparison to control-HTK, red bars indicate comparatively downregulated expression, and yellow bars indicate both up/downregulation. Common pathways are labelled with an asterisk.

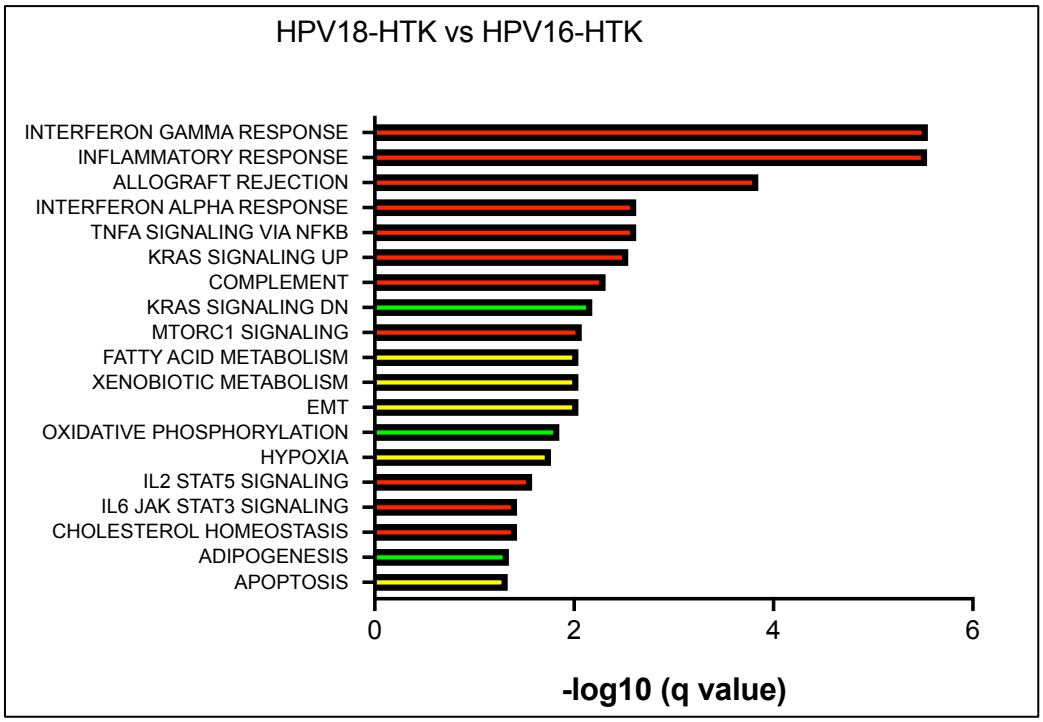


Figure 5-6: GSEA comparing HPV18-HTK against HPV16-HTK.

Shows all significantly differentially expressed MSigDB Molecular Hallmark gene sets between the two virus type models ($q < 0.05$, or $-\log10 > 1.3$). The interferon gamma response was the most significantly differentially expressed gene set. Colours indicate comparable differences in overall direction of expression of each gene set between the two HPV-type lines. Red bars indicate gene sets in which overall expression is significantly lower in HPV16-HTK than in HPV18-HTK lines; green bars indicate comparatively higher expression in HPV16-HTK than HPV18-HTK and yellow bars indicate up and down expression.

5.7 Differential gene expression analysis: comparison of HPV16-HTK and HPV18-HTK immune signalling gene sets

Following GSEA-based identification of alternatively expressed pathways between HPV16-HTK and HPV18-HTK, volcano plots were prepared to examine specific gene expression changes within significantly altered MSigDB Hallmark gene sets of interest (Figure 5-7 and Figure 5-8). Given the clear distinctions in immune regulation between HPV16 and HPV18, ongoing analyses are primarily focused on differences in immune regulation between the two virus-type lines. All eight differentially regulated immune pathways play important roles in the cell-mediated antiviral response and are described below. In the volcano plots, gene expression of each immune pathway was shown to be negatively skewed for the majority of genes, further signifying lower expression in HPV16-HTK lines than in HPV18-HTK.

It is important to note that these volcano plots only demonstrate relative gene expression differences between HPV16-HTK and HPV18-HTK lines and do not indicate how gene expression is altered in comparison to HPV-negative controls. The complete list of significantly differentially expressed chemokines, cytokines, immune signalling receptor genes, and other notable immune regulatory genes are depicted in supplementary Table S-2.

5.7.1 IFN-γ response

The type-II interferon signalling pathway, IFN-γ signalling pathway was the most significantly differentially expressed gene set (q value = 2.82×10^{-6}). This entails genes that are commonly upregulated following IFN-γ stimulation [557]. Out of 177 genes within this gene set, 11 were

significantly differentially expressed between the two virus type lines, 10 of which had comparatively reduced expression in HPV16-HTK (Figure 5-7a). These included positive regulators of lymphocyte activation and proliferation *CCL2*, *CD40*, *CD74*, *IL15,* and *IL7*, along with cytokine regulators, *NLRC5* and *CD74*, which are also known to positively regulate MHC class I (MHC-I) and MHC class II (MHC-II) antigen presentation pathways respectively [560, 561]. Only *CCL5* was more highly expressed in HPV18-HTK than HPV16-HTK. This is a potent chemotactic cytokine of T-cells, eosinophils, and basophils.

5.7.2 Inflammatory response

This pathway comprises genes that are upregulated following activation of the inflammatory response (Figure 5-7b). Genes that were downregulated by HPV16 included the IL-1 pro-inflammatory family members *IL1A*, *IL1B*, and *IL18R1*. Other potent activators of the inflammatory response include TNF receptors (TNFR) such as *TNFRSF9* and *CD40*, which were also transcriptionally repressed in HPV16-HTK models.

5.7.3 Allograft rejection

This gene set overviews genes of the alloimmune response that are activated in response to nonself antigens, many of which are involved in MHC-mediated antigen presentation and activation of T-cells [557]. In the case of HPV-infection, induction of the allograft rejection pathway would indicate viral antigen detection and presentation towards T-cell activation. On the other hand, HPV-mediated suppression of this pathway would likely indicate immunosuppression. HPV16 has a net negative effect on overall gene expression of the

allograft rejection pathway when compared with HPV18-HTK (Figure 5-7c). Significant reductions in expression were seen in nine genes; comprising five cytokines (*CCL2*, *IL7*, *IL11*, *IL1B*, and *IL15*), as well as four genes involved in cell surface immune cell signalling (*CD40*, *CD1D*, *HLA-DRA*, and *CD74*). Expression of *TLR6*, which plays a key role in NF-kB activation, was also reduced in HPV16-HTK models but just outside significance (p = 0.058).

5.7.4 IFN-α response

IFN-α is a central transcriptional activator that regulates a multitude of downstream immune response pathways. Significantly altered IFN-α response genes included *IL7, CD74, IL15, IFI27, PROCR,* and *LAMP3* (Figure 5-7d). IFI27 and LAMP3 are pro-apoptotic proteins. Expression of the *ISG20*, which exhibits direct antiviral functions [562], was also reduced in HPV16-HTK (p = 0.059).

5.7.5 TNFα signalling via NF-kB

This gene set comprises genes that are regulated by the transcription factor NF-kB in response to activation by TNFα via the canonical NF-kB activation pathway. NF-kB positively regulate numerous immune- and cell growth/survival-related pathways. Expression of nine, mostly immune-related genes were significantly reduced in HPV16-HTK (Figure 5-8a). These included cytokines (*CCL2, IL1A*, and *IL1B*) and cytokine receptors (*IL7R* and *IL15RA*), as well as *TNFRSF9* and NF-kB-mediators, *CXCL1* and *TRAF1*, though these did not quite reach significance.

This analysis also highlighted a number of differentially expressed genes which play variable regulatory roles in cell growth, development, and metabolism. These included *ID2*, *FOS, FOSL1, BTG2, EGR3, BTG2*, and *SLC2A3*. Expression of *ID2* and *BTG2*, which are negative regulators of differentiation and proliferation, was higher in HPV16-HTK models. These data also revealed contrasting expression of the FOS gene family members, *FOSL1* and *FOS*. FOS family and JUN family members dimerize to form the transcription factor AP-1, which regulates expression of genes involved in cellular proliferation, survival, and apoptosis. Here, *FOS* expression was higher in HPV16-HTK and *FOSL1* expression was lower. However, *FOS* is thought to be a much more transcriptionally active alternative to *FOSL1* [512, 513]. This could potentially indicate more active AP-1 activity in the HPV16 lines.

5.7.6 Complement system

This system contains innate immune genes that are components of the complement system. Genes whose expression was significantly lower in HPV16-HTK than HPV18-HTK included *SERPING1*, *LTF*, and *DGKG*. SERPING1 is a protease inhibitor that primarily inhibits the complement system (Figure 5-8b). LTF on the other hand is a multifunctional first line defence mediator that stimulates the complement system. Genes with higher expression in HPV16-HTK cells included *KLK1, CD36,* and *CCL5*. While KLK1 is thought to negatively regulate type I interferon production [563], CCL5 and CD36 both play pro-immune stimulatory roles.

5.7.7 IL2-STAT5 signalling

This gene set comprises genes that are upregulated by STAT5 following IL2 stimulation. This is a major proinflammatory signalling pathway that regulates T-cell proliferation, survival, and developmental processes. *IL2* itself was not detected in the RNA-seq data. However, expression of STAT5-inducable genes *COL6A1, TNFRSF9, SLC39A8, SLC2A3* and *IL18R1* were all significantly lower in HPV16-HTK cells (Figure 5-8c).

5.7.8 IL6-JAK-STAT3 signalling

This pathway contains genes that are upregulated by IL6 via the JAK-STAT3 signalling cascade (Figure 5-8d). While no significant differences in *IL6* expression was observed between HPV16-HTK and HPV18-HTK lines, *IL6* was significantly upregulated in both HPV-HTK lines when compared against controls, but to a greater extent in HPV18-HTK (log_2 fold change = 1.95) than in HPV16-HTK (log_2 fold change = 1.26) (supplementary Table S-2). In agreement, expression of IL6-JAK-STAT3-stimulated genes were generally lower in HPV16-HTK than HPV18-HTK (Figure 5-8d). Significantly altered genes included *IL18R1, IL1B, IL15RA, PLA2G2A,* and *CD36*.

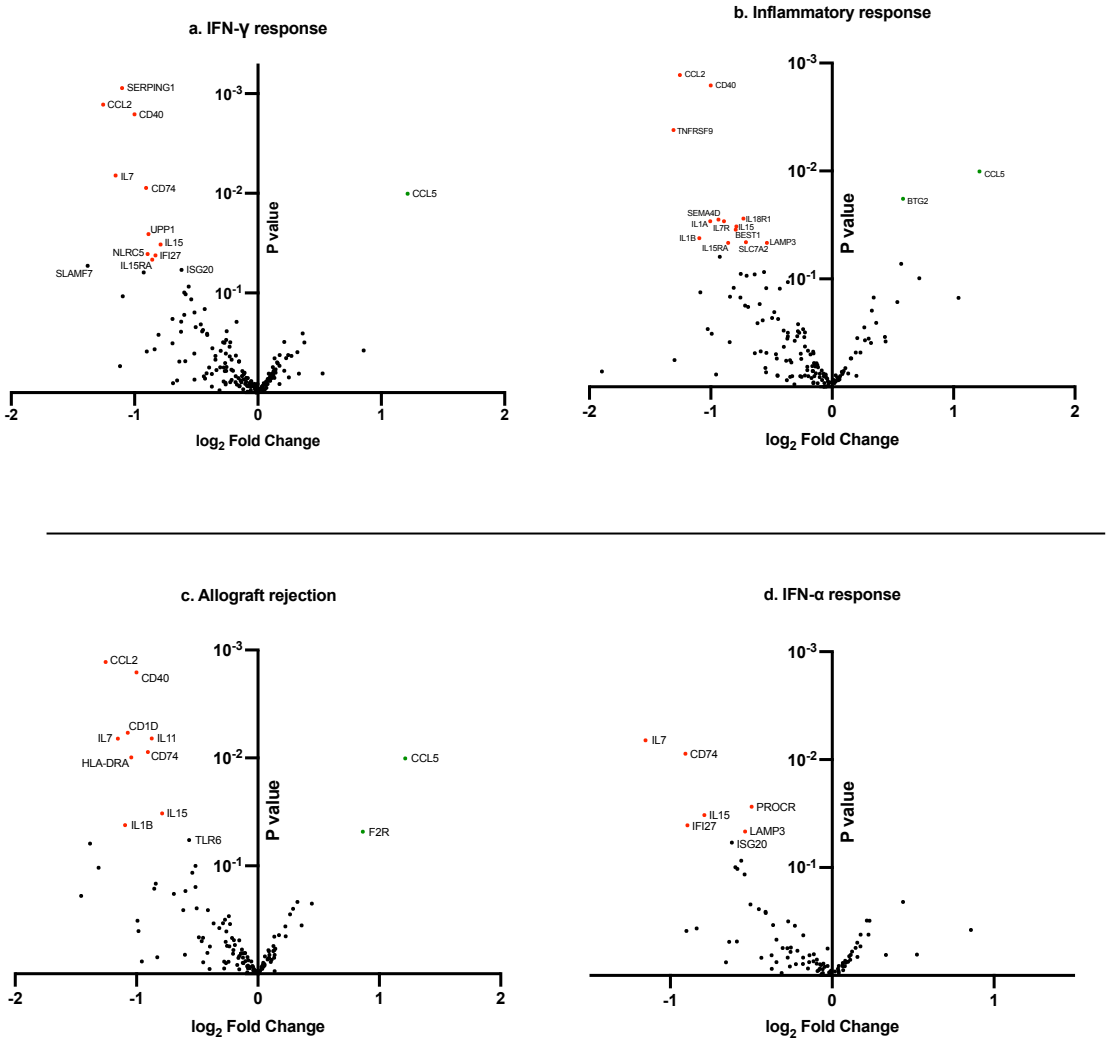


Figure 5-7: HPV16-HTK vs HPV18-HTK, differential gene expression analysis of altered immune pathways (1)

Volcano plots depicting gene expression changes within the top four immune pathways, shown by GSEA to be significantly differentially expressed between HPV16-HTK and HPV18-HTK. All significantly altered genes are coloured red or green ($p < 0.05$). Green points indicate genes where expression was significantly higher in HPV16-HTK than HPV18-HTK, or vice versa. Red shows the opposite.

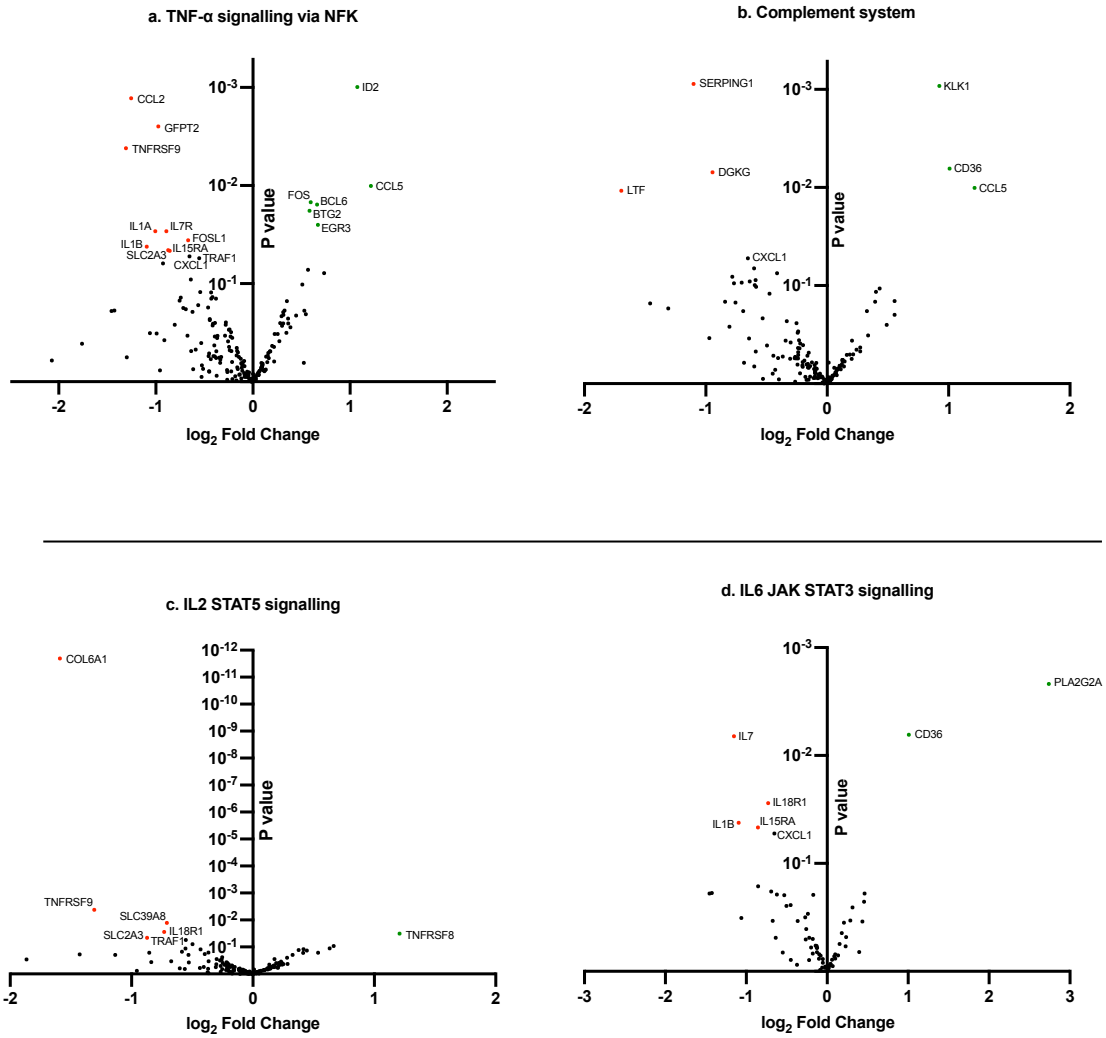


Figure 5-8: HPV16-HTK vs HPV18-HTK, differential gene expression analysis of altered immune pathways (2)

Volcano plots depicting gene expression changes within immune pathways, shown by GSEA to be significantly differentially expressed between HPV16-HTK and HPV18-HTK. Volcano plots compare p-values against regularised $\log_2$ fold change to fully conceptualise differences in gene expression. All significantly altered genes are coloured red or green and are labelled with gene names ($p < 0.05$). Green points indicate genes where expression was significantly higher in HPV16-HTK than HPV18-HTK, or vice versa. Red shows the opposite.

5.8 Differential gene expression analysis, comparison of HPV16-HTK and untransfected-HTK immune signalling gene sets

HPV16 was also shown to significantly reduce gene expression profiles in four immune pathways when compared with untransfected-HTK lines. Additional volcano plots were prepared to examine specific gene expression changes within these significantly altered immune pathways (Figure 5-9).

5.8.1 IFN-γ and IFN-α responses

In the IFN-γ response gene set, 16 genes were shown to be significantly downregulated by HPV16 (Figure 5-9a), including expression of potent immune stimulators *CCL2, CD40, IL7,* and *IFI27*. Suppression of genes involved in MHC class I antigen presentation was also observed. This included MHC class I components, *B2M, HLA-B*, and *HLA-G*, and the MHC class I antigen loading mediator, *TAP-1*. *MX1* and *MX2*, which have strong intracellular antiviral functions were also downregulated, as well as the critical IFN-α and IFN-β-transcriptional inducer, *IRF7*. *IL6* and *CCL5* expression was upregulated in HPV16-HTK. As demonstrated in Figure 5-9c, expression of the IFN-γ receptor component, *IFNGR2*, which is critical for IFN-γ activation, was significantly downregulated in HPV16-HTK cells.

Transcription of many of these significantly altered genes are also known to be induced by IFN-α and in the IFN-α response pathway in Figure 5-9b. While not included in these IFN gene sets, expression of *IFNB1*, which encodes the other main type I interferon, IFN-β, was also

significantly reduced in HPV16-HTK in comparison to control-HTK ($p = 0.029$, $\log_2$ fold change = -0.95).

5.8.2 Inflammatory response

HPV16 replication was shown to significantly downregulate expression of 16 inflammatory response-related genes and induce expression of eight others (Figure 5-9c). Downregulated genes notably included central PAMP recognition receptors (PRRs), *TLR3* and *NOD2*. Other notable transcriptionally repressed proinflammatory proteins included *CCL2, CD40, RASGRP1, IRF7,* and *IL18R1*. However, expression of proinflammatory genes *CCL5, IL6, CX3CL1*, and *IL7R*, were upregulated in HPV16-positive cells.

5.8.3 Complement system

Finally, HPV16 was also shown to dysregulate expression of a large number of genes within the MSigDB complement system gene set (Figure 5-9d). While many of these genes play important roles in various areas of the immune response, only a few significantly altered genes were identified that play direct roles in the complement cascade itself. As such, it was not clear from these analyses to what extent HPV16 was regulating the complement system.

5.8.4 Concluding remarks

HPV16-mediated dysregulation of immune response pathways was shown to comparatively downregulate expression of a large number of genes that play critical and varying roles in virus sensing, host immune alarm functions, and antiviral response mechanisms.

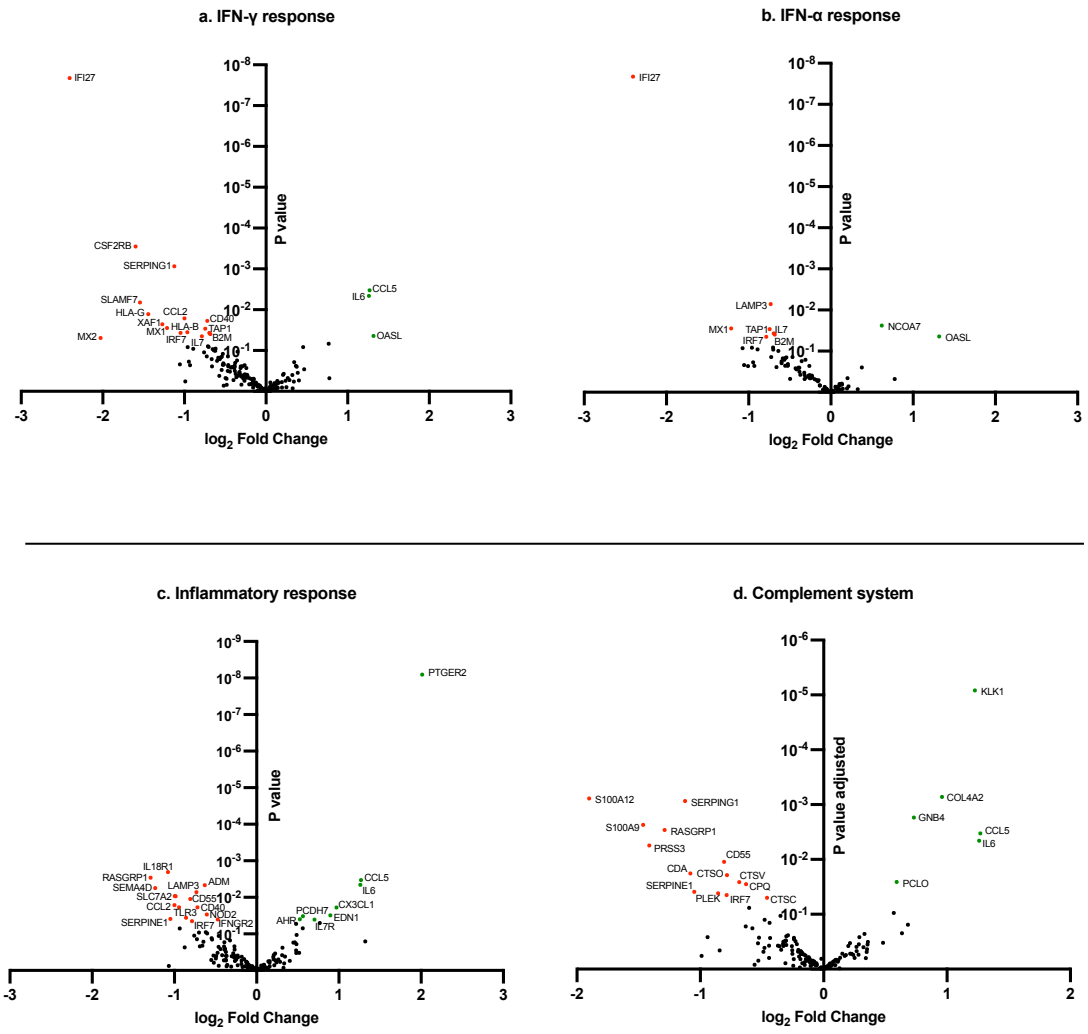


Figure 5-9: Control-HTK vs HPV16-HTK, differential gene expression analysis of altered immune pathways

Volcano plots depicting gene expression changes within all four immune signalling pathways that were shown to be significantly differentially expressed between untransfected-HTK and HPV16-HTK cells by GSEA. All significantly altered genes are coloured red or green and are labelled with gene names, where $p < 0.05$. Green points indicate genes in which expression significantly increased in HPV16-HTK when compared with control-HTK cells. Red points indicate significant decreases in expression in HPV16-HTK when compared with control-HTK.

5.9 Functional interaction analysis

Functional network analysis of the RNA-Seq data was carried out using the STRING database tool to evaluate potential post-translational physical interactions and functional associations between genes that were significantly differentially expressed between HPV16-HTK, HPV18-HTK, and untransfected-HTK lines (https://string-db.org/) [564]. These analyses also allowed for visualisation of wider protein networks that were collectively up and/or downregulated following HPV-mediated transcriptome reprofiling, as well as identification of central proteins that could be playing key regulatory roles to support viral infection and persistence.

In the first instance, interaction networks were used to evaluate genes that were differentially expressed between HPV-HTK and control-HTK. Genes were only included in this analysis if they were significantly differentially expressed ($p \leq 0.05$) and presented a $\log_2$ fold expression change > 1. Interaction networks were split into two analyses: (1) Genes that were significantly downregulated in HPV-HTK cells when compared with control-HTK, and (2) genes that were significantly upregulated in HPV-HTK cells. Interaction networks excluded any genes that did not present any interactions.

5.9.1 Functional interaction analysis, comparison of HPV16-HTK and untransfected-HTK

Functional interaction networks were first used to examine potential post-translational interactions between genes that were significantly differentially expressed between control-HTK and HPV16-HTK lines (Figure 5-10). Here, proteins have been linked depending on physical or functional associations, corroborated from multiple primary databases [564]. The

thickness of a line that connects two proteins indicates the strength of the data supporting this interaction. This is dependent on several selected factors including experimentally defined protein-protein interactions, gene/protein co-expression/co-occurrence, as well as genes which are co-expressed from a close neighbourhood within the genome [564].

In the first instance, interaction networks assessed genes that were significantly downregulated in HPV16-HTK lines (Figure 5-10a). This analysis highlighted a large network comprising 45 genes, which are known to be implicated in numerous immune response pathways including type I interferon responses, cytokine-mediated signalling, and the inflammatory response. Central genes that connected different immune pathways included *CCL2, CD40, B2M, TLR3*, and *IL2RG*. Altogether, 12 genes were mapped that play roles in lymphocyte activation, including *CD40, TNFSF4, CD1D, B2M, LEPR, LEP, IL11, IL7, IFNB1, IL18R1, CLEC7A,* and *RORA.* In addition, nine genes were mapped that are involved in the inflammatory response, including *CCL2, ACKR2, CD40, TNFSF4, TLR3, LY96, S100A9, S100A8*, and *S100A12*. A large cluster of highly interconnected immune genes was revealed, consisting of *HLA-B, HLA-G, IFI6, IFI27, IRF7, MX1, MX2, OAS2,* and *XAF*. Along with *B2M* and *IFNB1* (encodes IFN-β), these genes are all heavily involved in antiviral immune responses, including type I interferon signalling, mediation of type I IFN transcription, mediation of interferon and TNF-α induced apoptosis, and MHC-I antigen presentation.

A second large cluster of interconnected downregulated genes were observed comprising components of the cornified envelope. These play roles in cornification, keratinization, keratinocyte differentiation, epidermal development. Many of these genes encode proteins

that make up the cell envelope for example involucrin (*IVL*), *TCHH, EVPL, PPL, SPRR2E, SPRR2G, SPRR1A, SPRR1B, RPTN,* and *CNFN*. In addition, around 34 genes were included that play various roles in cellular metabolism pathways including, lipid metabolism, fatty acid metabolism, xenobiotic metabolism, and organic hydroxy compound metabolism. This included a large cluster of genes which centred around *CYP2E1*, which encodes the protein, cytochrome P450, which catalyses xenobiotic metabolism reactions and assists synthesis of lipids and cholesterols. Other central genes included *CFTR*, *EGF*, and *SERPINE1*.

Interaction networks were next used to examine potential downstream interactions between genes that were significantly upregulated in HPV16-HTK lines when compared with controls (Figure 5-10b). These were highly clustered and largely consisted of genes complicated in cell growth and cell cycle progression. Some major cell cycle pathways included cell cycle checkpoint and checkpoint progression, mitosis, and DNA synthesis and replication pathways. A large number of interactions were also observed between genes involved in p53 mediated pathways such as DNA repair pathways. In addition, this analysis also highlighted gene networks involved in chromatin remodelling, meiosis, and senescence.

5.9.2 Functional interaction analysis, comparison of HPV18-HTK and untransfected-HTK

Further interaction networks were prepared to evaluate genes that were significantly differentially expressed between HPV18-HTK and control-HTK (Figure 5-11). In the first instance, interaction maps looked at all genes downregulated in HPV18-HTK models (Figure 5-11a). As with HPV16 vs HTK interactions, many downregulated functional networks centred

around *CYP2E1* and *EGF*. Around *CYP2E1*, a high proportion of interconnecting genes were involved in a variety of cell metabolism processes including xenobiotic metabolism, fatty acid metabolism, hormone metabolism, primary alcohol metabolism, and retinoid metabolism. Together these data suggest that HPV18 is strongly modulating wider cellular metabolism networks. A large number of genes were also identified which play roles in cell development and survival. These included cornification, keratinization, differentiation, epidermal development, cell adhesion, and cell death related pathways. In contrast to HPV16-HTK vs control comparisons, few immune-related proteins were observed.

Next, interaction networks were prepared to examine genes that were significantly upregulated in HPV18-HTK models (Figure 5-11b). As in HPV16-HTK vs control comparisons, HPV18-HTK were shown to stimulate higher expression of similar genes involved in cell cycle progression and cell growth, DNA replication and DNA repair pathways. Further similarities were seen between upregulation of pathways related to meiosis, senescence, and chromatin remodelling. A number of significantly upregulated immune-related genes were observed, associated with MHC-II signalling (*HLA-DRB1, HLA-DRA, CD74* and *HLA-DPA1*), toll-like receptor signalling (*TLR4* and *TLR6*), in addition to important interleukin/chemokine signalling components (*IL6, CXCL* and *CXCL2)*.

5.9.3 Functional interaction analysis, comparison of HPV16-HTK and HPV18-HTK lines

Final interaction maps evaluated potential post-translational interactions between genes that were differentially expressed between HPV16-HTK and HPV18-HTK models. First, interaction

networks looked at genes in which expression was significantly lower in HPV16-HTK cells than in HPV18-HTK (Figure 5-12a). This analysis highlighted a large network comprising immune signalling genes that are known to play central roles in both innate and adaptive immune responses, such as cytokine-mediated immune signalling, lymphocyte recruitment and activation, and the inflammatory response. Central genes included *CCL2, CD40, IL1B, IL1A, IL15,* and *HLA-DRA*. Genes involved in lymphocyte activation included *CCL2, CD40, IL1B, TLR4, IL15, IL7R, CD1D, CD5, CD74*, and *PDCD1LG2*. These are all additionally implicated in cytokine-mediated signalling along with *IL1A, HLA-DRA, IL15RA, TNFRSF9, IL11,* and *IL13RA2*. Inflammatory response related genes included *CCL2, IL1A, IL1B,* and *TLR4.* A second large cluster appeared to centre around the epidermal growth factor, *EGF*. Many of these genes play roles in cell development.

Interaction maps examining genes in which expression was significantly higher in HPV16-HTK than in HPV18-HTK, were largely centred around *SOX2, FOS, BCL6, PPARGC1A, FZD7, ALDH3A2* and *AKR1C3* (Figure 5-12b). A large gene cluster centred around *ALDH3A2* and *AKR1C3*, consisted of genes associated with a wide variety of metabolism related functions. The largest network, mainly centred around *FOS, BCL6* and *SOX2* are genes generally involved in regulation of cell proliferation. However, many of these genes play opposing roles in both positively and negatively regulating proliferation.

5.9.4 Concluding remarks

Altogether, these functional and physical interaction networks highlighted a collection of central differentially expressed genes that could be contributing to wider HPV16-mediated immune dysregulation in the primary HTK lines. This notably included, *CD40*, *CCL2*, and the PRR, *TLR3*. In addition, these analyses indicated downregulated networks of essential immune regulators, including IL-1 family members (*IL1B, IL1A*, and *IL18R1*), as well as genes involved in MHC-I and MHC-II antigen presentation, (*HLA-B, HLA-G, TAP1, B2M, HLA-DR1,* CD1D).

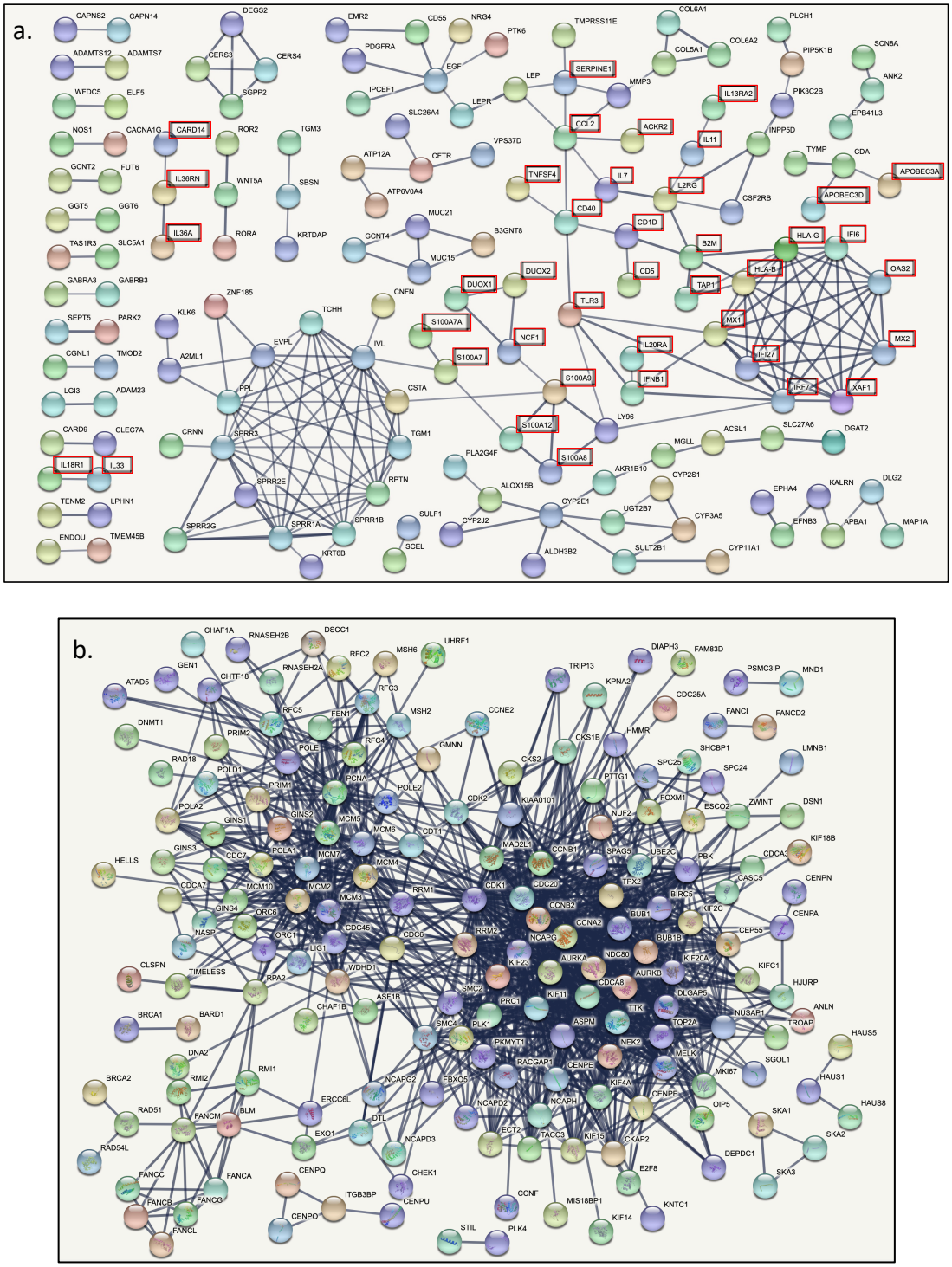


Figure 5-10: Functional interaction network analysis, HPV16-HTK vs untransfected-HTK lines

Interaction analysis was carried out to evaluate potential post-translational interactions between genes that were shown to be **a**. significantly downregulated in HPV16-HTK lines when compared with untransfected-HTK, or **b.** significantly upregulated in HPV16-HTK lines when compared with untransfected-HTK**.** Red boxes indicate any genes involved in the antiviral immune response.

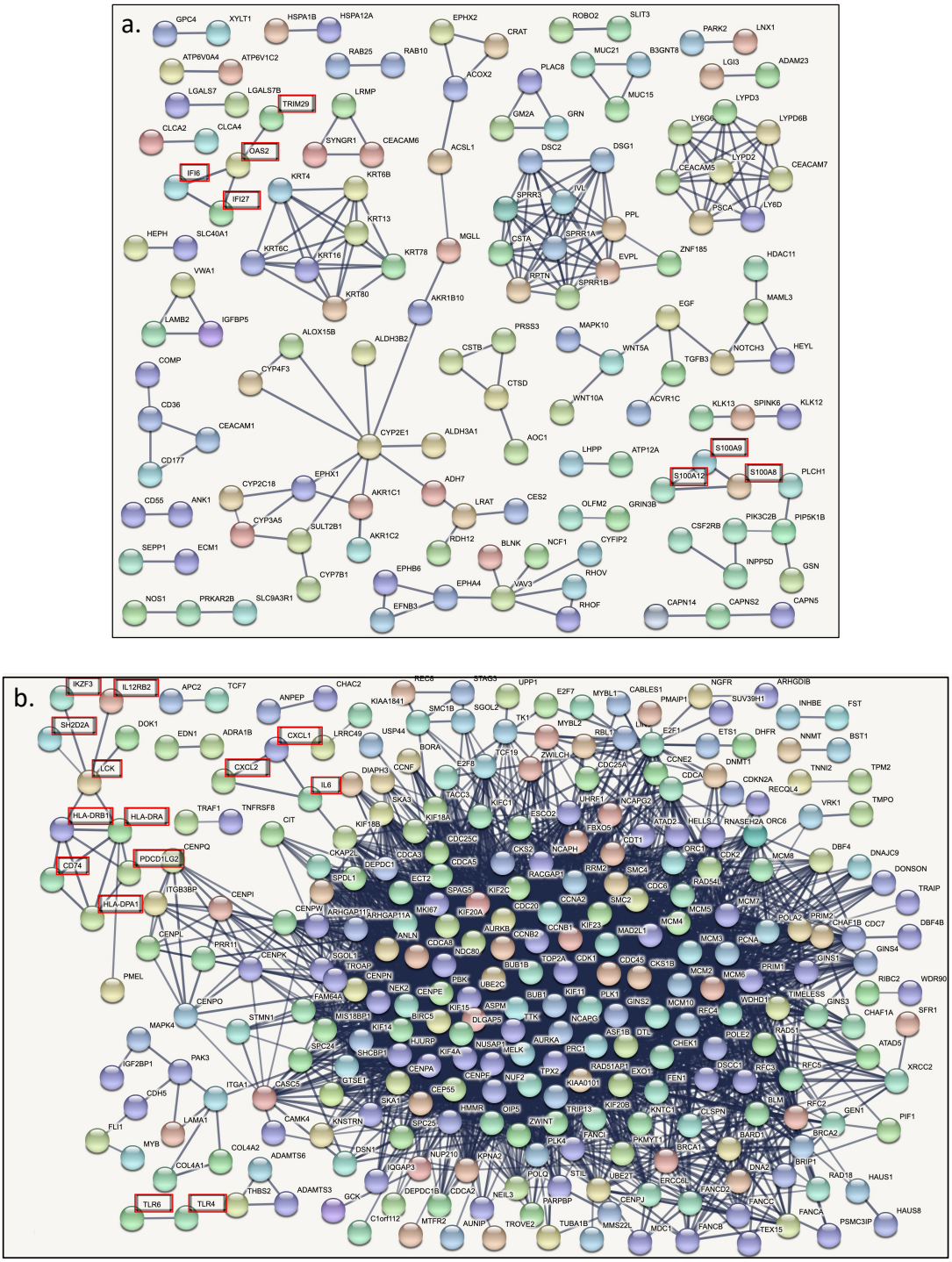


Figure 5-11: Functional interaction network analysis, HPV18-HTK vs untransfected-HTK lines

Interaction analysis was carried out to evaluate potential post-translational interactions between genes that were shown to be **a**. significantly downregulated in HPV18-HTK lines when compared with untransfected-HTK, or **b.** significantly upregulated in HPV18-HTK lines when compared with untransfected-HTK**.** Red boxes indicate any genes involved in the antiviral immune response.

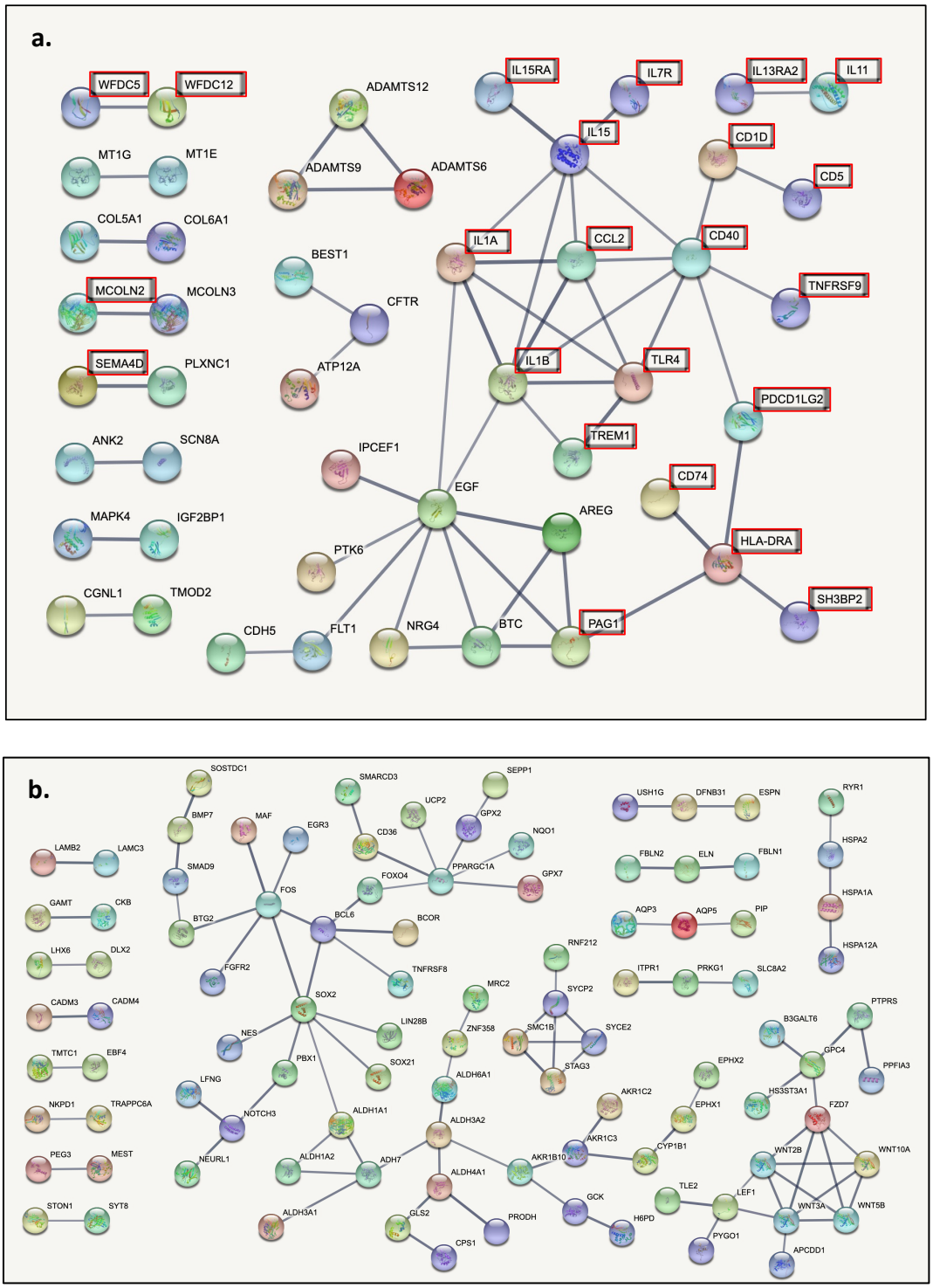


Figure 5-12: Functional interaction network analysis, HPV16-HTK vs HPV18-HTK

Network analysis was carried out to evaluate potential post-translational interactions between significantly differentially expressed genes that were shown to be **a**. more highly expressed in HPV16-HTK lines than in HPV18-HTK lines, or **b.** more highly expressed in HPV18-HTK lines than in HPV16-HTK lines. Red boxes indicate any genes involved in the antiviral immune response.

5.10 Validation of differential gene expression of immune signalling genes between HPV16-HTK and HPV18-HTK lines

To validate some of the key changes identified by RNA-Seq, qPCR was carried out using three independent replicates of each HTK and HPV-HTK model. RNA was extracted from three sequentially passaged cell populations, which were DNase I treated prior to cDNA synthesis (see methods section 2.3).

5.10.1 qPCR-based validation of differentially expressed IL-1 family members

In the first instance, qPCR analysis was carried out to validate differences in expression of selected IL-1 family members that play central roles in inflammation. IL-1β is the primary modulator of inflammation that plays crucial roles in a number of innate and adaptive immune response mechanisms [565]. The RNA-Seq data indicated lower expression of *IL1B* in HPV16-HTK than in HPV18-HTK (Figure 5-13a). However, greater reductions were observed in HPV16-HTK#2 than HPV16-HTK#3. These findings were consistent with those determined by qPCR, where *IL1B* expression was significantly lower in HPV16-HTK lines than donor equivalent HPV18-HTK and untransfected-HTK lines (with a slight exception for HPV16-HTK#3) (Figure 5-13b). However, unlike in the RNA-seq data, a significant reduction in *IL1B* was also observed in HPV18-HTK#2 cells in comparison to HTK#2 controls.

IL18 expression remained at high and relatively unchanged levels between HTK and HPV-HTK models (supplementary Table S-2). However, according to the RNA-seq data, expression of the IL18 receptor component, *IL18R1*, which is required for IL18-mediated signalling, was

lower in HPV16-HTK models than in HPV18-HTK and controls (Figure 5-13c) [566]. In agreement, qPCR demonstrated significant reductions in *IL18R1* levels in HPV16-HTK#2, as opposed to HPV18-HTK#2 (Figure 5-13d). Non-significant reductions in expression were observed when comparing HPV16-HTK#3 with HPV18-HTK#3. However, comparisons between HPV16/18-HTK lines and controls showed greater and more significant reductions in *IL18R1* expression by HPV16 overall. As such, these data followed a similar trend to the RNA-Seq data.

5.10.2 qPCR-based validation of differentially expressed cytokines

Next, qPCR was performed to validate altered expression of cytokines known to play central roles in keratinocyte-immune cell signalling. CCL2 is a potent chemoattractant of lymphocytes and myeloid cells [567]. In agreement with RNA-Seq findings [567]. In agreement with RNA-seq findings (Figure 5-14a), qPCR detected very few *CCL2* transcripts in both HPV16-HTK donor lines (Figure 5-14b). However, *CCL2* remained at relatively high levels in HPV18-HTK and controls, altogether validating the RNA-Seq data.

IL15 is another potent stimulator of lymphocyte activation and proliferation [568]. The RNA-Seq data indicated lower *IL15* expression in both HPV16-HTK donors when compared with equivalent HPV18-HTK lines (Figure 5-14c). While qPCR also showed significantly lower *IL15* expression in HPV16-HTK#2 than HPV18-HTK#2, this analysis did not reveal any significant change in expression between any HTK#3 model (Figure 5-14d).

5.10.3 Validation of differential *CD40* expression

CD40 is a transmembrane protein and member of the tumour necrosis factor receptor (TNFR) family, which is highly expressed on the cell surface of keratinocytes, notably on oral and cervical basal keratinocytes, and plays an essential role in cellular, humoral, and adaptive immune responses [569-571]. The RNA-Seq data demonstrated lower expression of *CD40* in HPV16-HTK than in HPV18-HTK and controls (Figure 5-15a). These findings were also supported and validated by qPCR, which demonstrated comparable distinctions in *CD40* expression between HTK lines (Figure 5-15b).

Given the importance of CD40 in the keratinocyte-antiviral response, western blotting was carried out to assess and compare CD40 protein abundance in the HPV-HTK lines (Figure 5-15c and Figure 5-15d). As in the genomic studies, these data demonstrated a significant downregulation of CD40 in HPV16-HTK lines in comparison to untransfected- and HPV18-containing-HTK lines, altogether confirming marked reductions in CD40 levels in the presence of HPV16 but not HPV18. Since CD40 is commonly overexpressed in high grade HPV-cervical lesions and cervical cancers [571], CD40 protein abundance was also assessed in high passage HPV16-HTK and HPV18-HTK lines (>120 PD). This showed that CD40 levels remained very low in the presence of HPV16 and high in HPV18-HTK cells. However, it should be noted that CD40 levels were highly variable between HPV18-HTK independent replicates.

Finally, immunofluorescence staining of HTK, HPV16-HTK, HPV18-HTK, and HPV18-HFK organotypic raft cultures was performed to determine CD40 subcellular localisation and expression in differentiated epithelia (Figure 5-16). CD40 expression was typically cytoplasmic

and was frequently detected at the cell periphery. In untransfected rafts (HTK#3), CD40 was detected primarily in basal keratinocytes (also in occasional superficial cells). In the presence of HPV16 very few CD40-positive cells were detected, and this was consistent in both donors (HTK#2 and HFK#3). In contrast, CD40 positive cells were detected in the HPV18 rafts; in HPV18-HTK#2 these were largely confined to the basal layers whereas in HPV18-HTK#3 rafts cells throughout the thickness of the raft were positive. The distribution of CD40 in the HPV18 HTK#3 rafts likely reflects the more undifferentiated and dysplastic nature of these rafts (see chapter 3, Figure 3-12 - Figure 3-15 and Figure 3-19 - Figure 3-20). Since CD40 is primarily expressed in undifferentiated basal keratinocytes, high suprabasal CD40 expression was somewhat expected. In contrast, even though HPV16-HTK#3 raft culture showed a similar inability to differentiate, CD40 was mostly undetected within basal and suprabasal cells. Overall, these IF analyses demonstrated clear differences in CD40 expression profiles between HPV16-HTK and HPV18-HTK raft culture.

5.10.4 Concluding remarks

These qPCR validations were shown to generally correspond with trends observed in the RNA-Seq data and support findings that HPV16 downregulates expression of *IL1B, IL18R1, CCL2,* and *CD40* to a greater extent than HPV18 in HTK. Further, HPV16-mediated inhibition of *CD40* at a transcriptional level appeared to impact downstream translation and cell surface expression of this important innate and adaptive immune mediator.

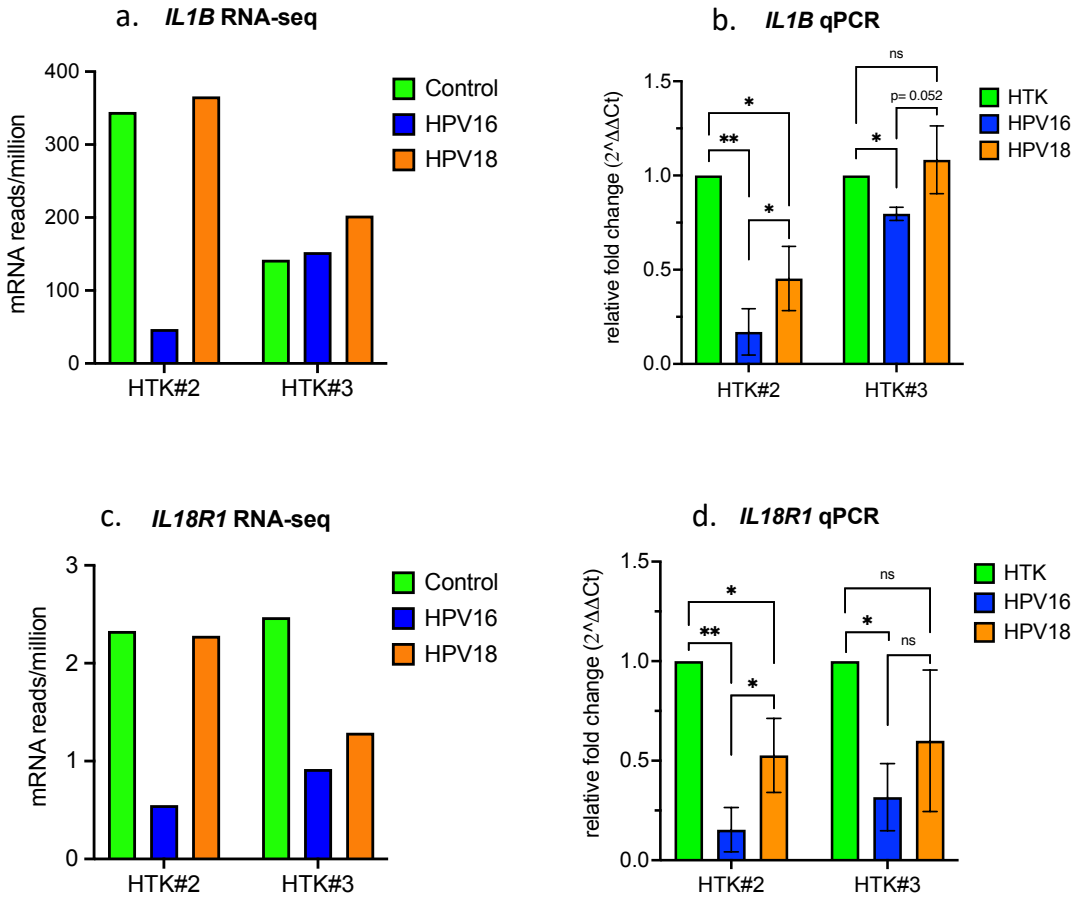


Figure 5-13: Differential gene expression validation of *IL1B* and *IL18R1*

a and **c**, shows the total number of *IL1B* and *IL18R1* transcripts, as detected by RNA-Seq, calculated in reads/million **b** and **d,** Further qPCR analysis was performed to validate differences in *IL1B* and *IL18R1* gene expression between HTK lines. qPCR analysis involved the $2^{-\Delta\Delta Ct}$ method, which was used to calculate relative fold difference relative to untransfected-HTK control (HTK) lines, where beta actin was used as a housekeeping gene. One-tailed paired or unpaired parametric T-tests were carried out to assess significance: non-significant (NS), $p < 0.05$ (*), $p < 0.01$ (**). Error bars show standard deviation.

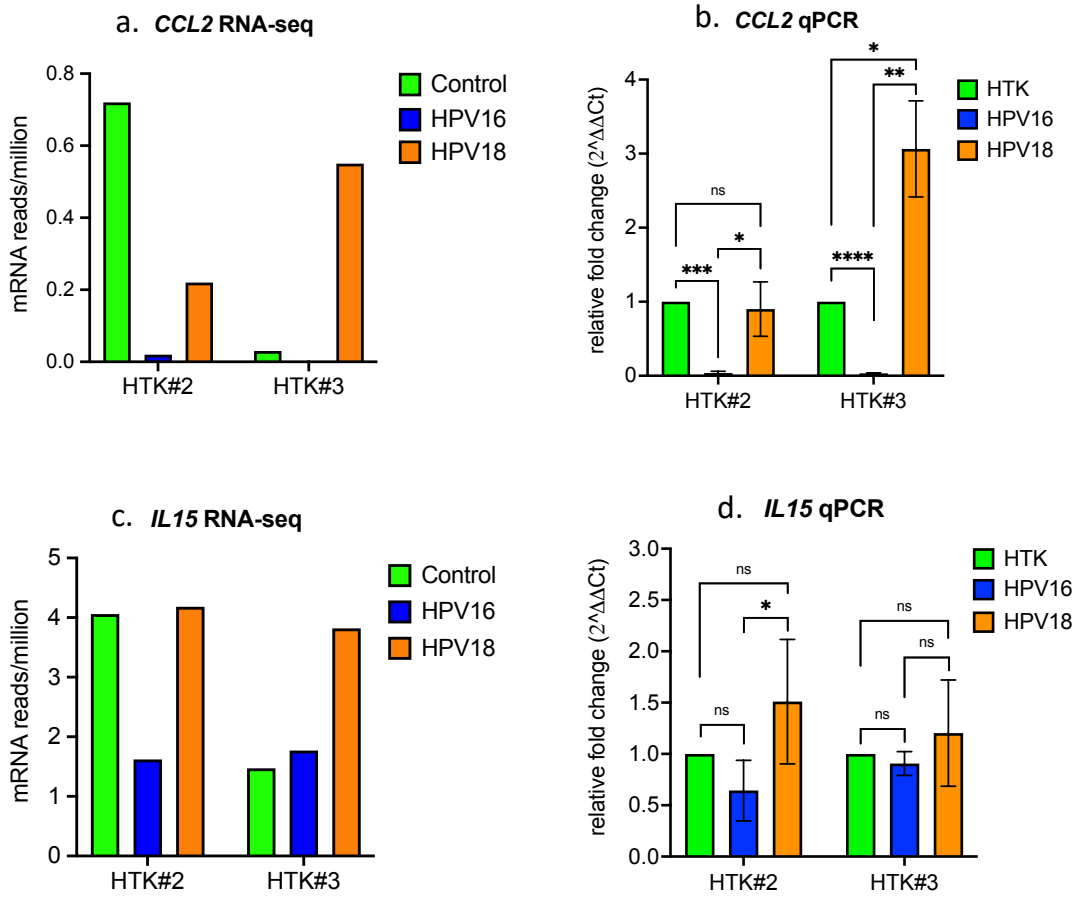


Figure 5-14: Differential gene expression validation of *CCL2* and *IL15*

a and **c**, shows the total number of *IL1B* and *IL18R1* transcripts, as detected by RNA-Seq, calculated in reads/million **b** and **d,** Further qPCR analysis was performed to validate differences in *IL1B* and *IL18R1* gene expression between HTK lines. qPCR analysis involved the $2^{-\Delta\Delta Ct}$ method, which was used to calculate relative fold difference relative to untransfected-HTK control (HTK) lines, where beta actin was used as a housekeeping gene. One-tailed paired or unpaired parametric T-tests were carried out to assess significance: non-significant (NS), $p < 0.05$ (*), $p < 0.01$ (**). Error bars show standard deviation.

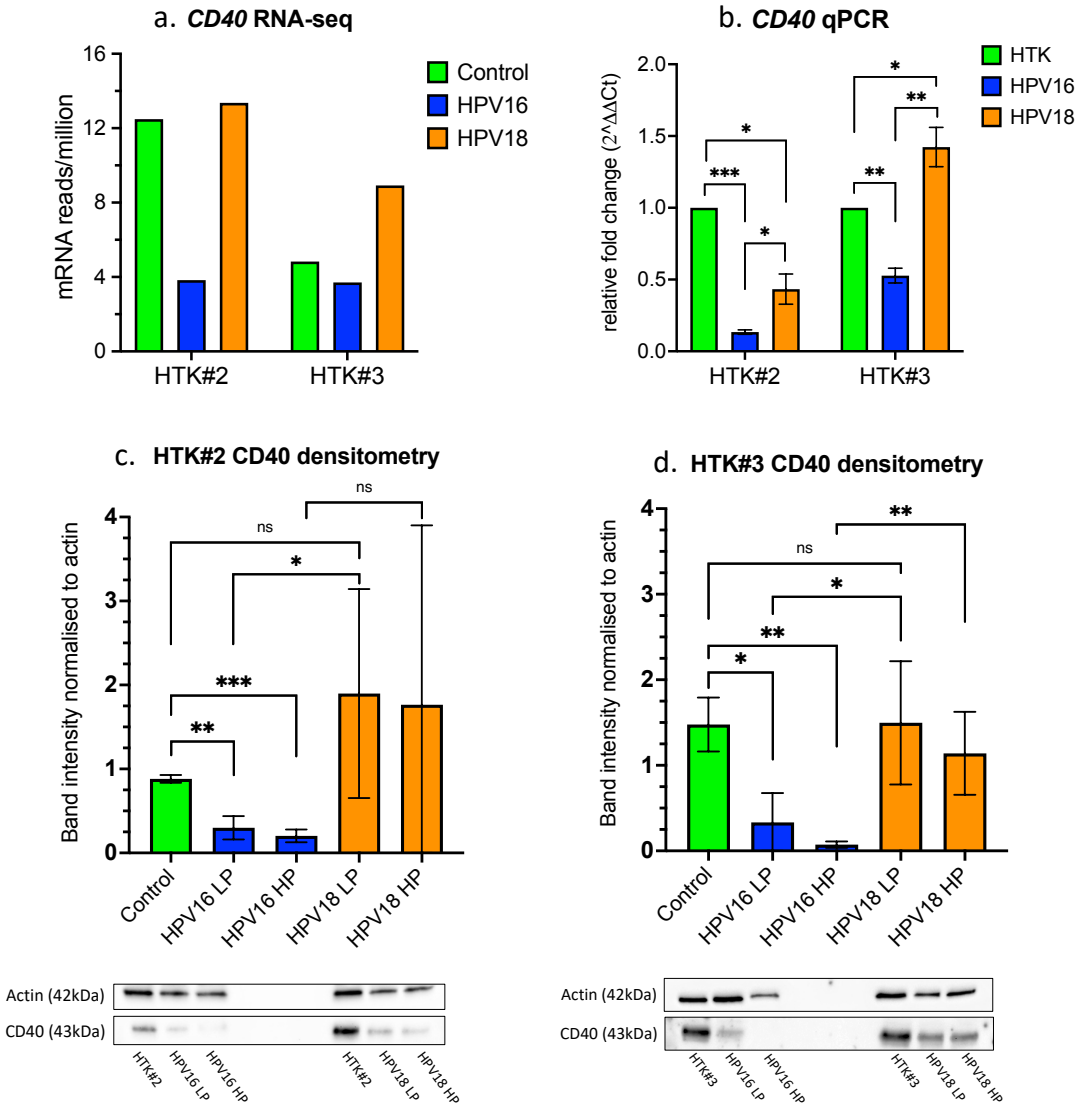


Figure 5-15: Differential expression validation of *CD40*

a. Shows total number of detected *CD40* transcripts detected by RNA-Seq in reads/million **b**. qPCR validation of comparable *CD40* expression using the $2^{-\Delta\Delta Ct}$ method, as previously described. Western blot analysis demonstrates comparable CD40 protein levels in isogenic low passage (LP) and high passage (HP) **c.** HPV-HTK#2 and **d.** HPV-HTK#3 lines. Densitometry was used to quantify and compare band intensity following normalisation against beta actin. Representative western blots are shown below each respective graph (see supplementary Figure S-3). One-tailed parametric T-tests assessed significance: Non-significant (NS), $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***). Error bars show standard deviation.

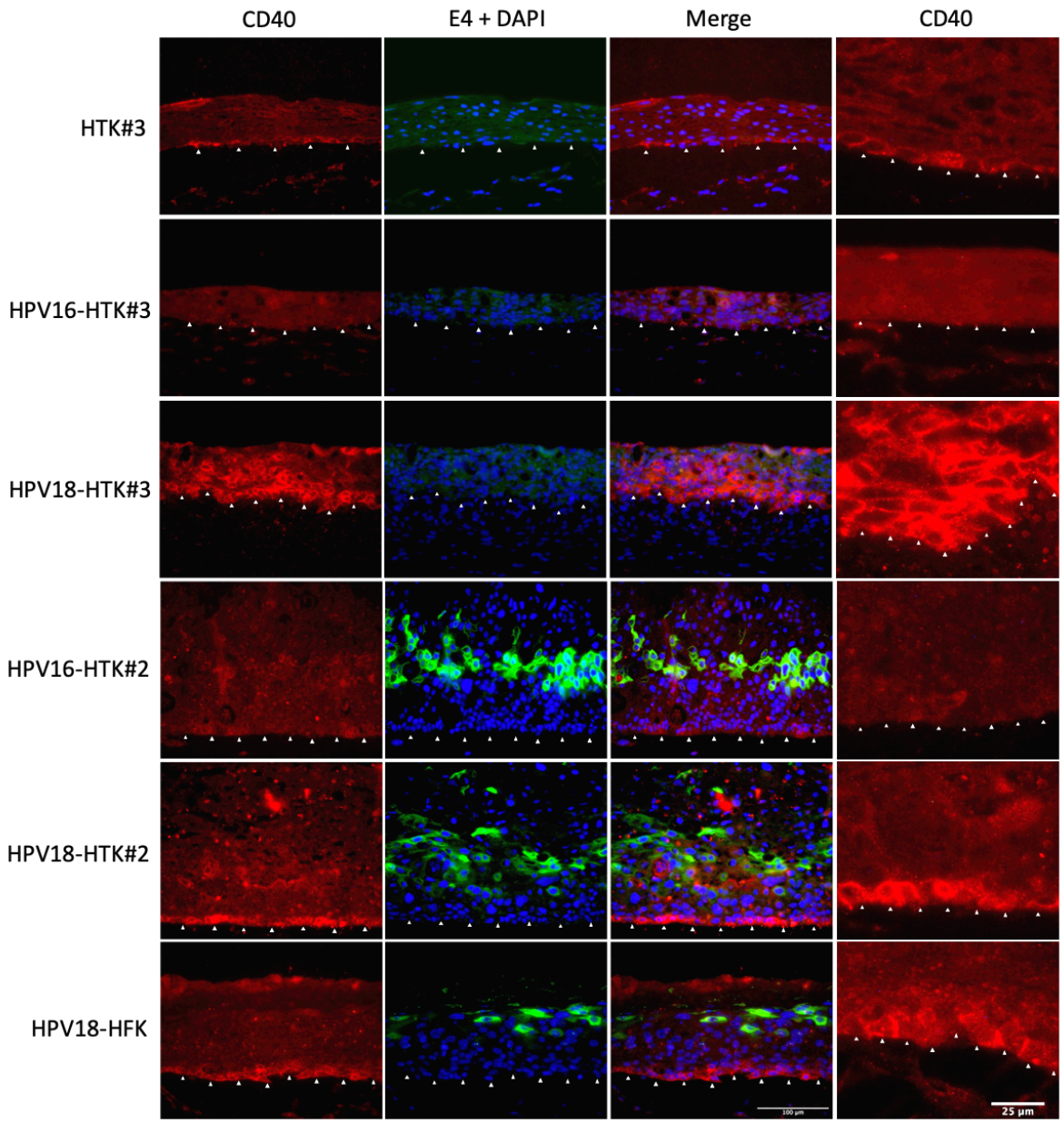


Figure 5-16: Immunofluorescence staining for CD40 and E1^E4 in HTK raft culture

HPV16/18-HTK, untransfected-HTK, and HPV18-HFK raft cultures were formalin fixed, paraffin embedded and immunostained with antibodies specific for CD40 (red), E1^E4 (green) and DAPI (blue). White arrows indicate basal membrane. Scale bars in the bottom row specify magnification in microns. The right-hand side column illustrates computer-enlarged images to better visualise CD40 localisation within basal keratinocytes.

5.11 Discussion

Towards a better understanding of HPV16-driven disease progression in the oropharynx, RNA-Seq was performed to characterise and compare host mRNA transcription profiles in primary untransfected-HTK and donor-equivalent HPV16- and HPV18-HTK lines. We hypothesised that HPV16, which is present in the vast majority of HPV-OPSCC, induces differential transcriptional reprogramming in the tonsils to HPV18, in a manner that further augments viral persistence and disease progression at this body site.

5.11.1 Comprehensive analysis and comparison of HPV16-HTK and HPV18-HTK transcriptome reorganisation

Principal component analysis indicated that HPV16- and HPV18-induced transcriptional reprogramming of HTK was variable between the two donor model systems. However, comprehensive analysis of the differential gene expression data revealed substantial HPV-type-specific variations of host transcriptomes. While HPV16 and HPV18 did alter transcription of many of the same host genes, a high proportion (~40%) of differentially expressed genes were shown to be exclusively regulated by each HPV type (Figure 5-1 - Figure 5-3). This suggests that HPV16 and HPV18 differentially regulated the transcription of a high proportion of alternative genes.

Gene set enrichment analysis (GSEA) was carried out to identify the biological pathways HPV16 and HPV18 were differentially targeting and regulating in HTK (Figure 5-6). In comparison to untransfected-HTK, HPV16 and HPV18 both, as expected, enriched expression

of E7-inducible pathways centred around cell cycle progression, such as cell proliferation, cell growth, cell division, mitosis, and DNA replication (see introductory section 1.3.3) [17, 362, 572-574]. In addition, expression of the 'p53 targets' pathway was shown to significantly decrease in both HPV16/18-HTK lines, indicating HPV16- and HPV18-E6-mediated inhibition of p53 activity in the HTK lines, likely driving inhibition of cell cycle arrest and apoptosis, along with increased chromosomal instability [290].

Interestingly, GSEA revealed that in comparison to controls and HPV16-HTK, HPV18-HTK was uniquely altering a large number of metabolism related pathways, including mTORC1 signalling, fatty acid metabolism, cholesterol homeostasis, heme metabolism, and adipogenesis (Figure 5-6). mTORC1 is a master transcriptional regulator of cell metabolism and growth that regulates many of the above pathways, such as promoting a switch in glucose metabolism from oxidative phosphorylation to glycolysis [575]. In addition, mTORC1 plays complementary roles in HPV-associated carcinogenesis via the PI3K/AKT/MTOR signalling cascade [576]. mTORC1 appeared to be more active in HPV18-HTK than in HPV16-HTK and controls, and relatively unchanged between HPV16-HTK and controls. This was somewhat surprising as mTORC1 overactivity is commonly observed in HPV-OPSCCs, which primarily contain HPV16 and rarely HPV18 [419]. mTORC1 signalling activities are enhanced by HR-HPV E6 and are theorised to assist viral replication by altering cell metabolism processes in a manner that provides nutrient support for protein synthesis [577-579].

Intriguingly, analysis of the RNA-Seq data showed differential expression of the DEP domain-containing mTOR-interacting protein (*DEPTOR*), a primary inhibitor of mTORC1 and mTORC2

activities that contains a PDZ binding domain, highlighting this protein as a target of E6 [580, 581]. Here, DEPTOR expression was significantly lower in HPV18-HTK lines than HPV16-HTK lines ($p = 0.0004$, regularised log2 fold change -0.8). It is possible that this finding could reflect differences in E6 expression between HPV16- and HPV18-HTK lines, as indicated in Chapter 4 (section 4.5.2, Table 4-3), or possibly differences in PDZ affinities between HPV16 and HPV18 E6, as discussed in introductory section 1.4.4. These differences in mTORC1 and DEPTOR are intriguing and may provide an important distinction between HPV16 and HPV18 in HTK that requires further investigation.

5.11.2 HPV16 represses multiple immune pathways in HTK in comparison to HPV18 and HPV-negative controls

Studies of cervical lesions, cancers, and primary anogenital infection models show that HR-HPVs have evolved a multitude of mechanisms to evade or inhibit host antiviral defences at varying stages of the innate and adaptive immune response, which has been overviewed in introductory section 1.4.5. These responses are primarily regulated by E6, E7 and E5 oncoproteins (see introductory Table 1-1). However, little is currently known about how HR-HPVs mediate host immunity during infection of oropharyngeal subsites like the tonsils, which are primary sites of HPV-driven carcinogenesis [31, 32, 35]. Given that the tonsils and tonsil crypts produce and maintain high levels of lymphocytes and other immune surveillance cells for immune system priming (introductory section 1.4.5), HPV-mediated immune regulation is likely critical for viral persistence at this body site.

Comparative analysis of transcription profiles of isogenic HPV16-HTK, HPV18-HTK and donor cells demonstrate that HPV16 uniquely impaired transcription of a diverse range of genes and wider gene networks in HTKs that are essential for innate and adaptive cell-mediated antiviral responses (Figure 5-5 - Figure 5-12). Importantly, these altered immune regulatory genes, networks and pathways all have highly interconnected and inter-regulatory roles, showing that HPV16 was able to broadly dampen host HTK mRNA profiles spanning across the transcriptional immune network relative to both untransfected and HPV18-transfected HTKs.

5.11.3 HPV16 and HPV18 alternatively regulate activities of central transcriptional immune mediators

In our HTK lines, HPV16 was found to suppress activities of a number of central transcriptional immune mediators, and to a significantly greater extent than HPV18 (Figure 5-6). These included IFN-α, IFN-γ, NF-kB, STAT3, and STAT5, which are known to be manipulated via E6- and E7-mediated regulation of virus-sensing PRRs, PRR-inducible signal transduction pathways, and other up/downstream signalling pathways, which are overviewed in introductory section 1.4.5, Table 1-1 [343, 363].

Differential regulation of the transcriptional immune mediators, IFN-α and NF-kB was an obvious difference between HPV16-HTK and HPV18-HTK lines. These are vastly interlinked and cooperative transcriptional inducers of intraepithelial immunity that play essential mediating roles for the majority of differentially regulated immune genes and pathways identified in these studies. Notably, this includes the inflammatory response and the IL6-JAK-STAT3 and

IL2-STAT5 signalling pathways, in addition to a number of downstream immune effector genes that could contribute to HPV16-mediated impairment of the antiviral immune response in the tonsils. These differentially expressed genes included, but was not limited to *CD40*, *CCL2*, *IL18R1*, *IL15, TLR3*, *IFN1B*, as well as a number of IL-1 and MHC-I and MHC-II family members, which are each important elements of the antiviral response, as is further discussed in section 5.11.4. Efficient regulation of these transcriptional mediators could potentially alter the host transcriptome in a manner that promotes a more anti-inflammatory, anti-immune environment that could hamper intracellular antiviral mechanisms and immune cell activation. If these observations hold true *in vivo* HPV infections in the tonsils, this could provide a significant survival advantage for HPV16 over HPV18 at this highly immunological anatomical site.

Given the central regulatory roles of E6 and E7 in the cell-mediated immune response, it seems likely that variable expression or activities of these oncoproteins could have contributed significantly to differing immune regulation observed between our HPV16- and HPV18-HTK lines. Crucially, HPV16 and HPV18 E6/E7 oncoproteins exhibit distinct structural and functional differences, as discussed in introductory section 1.4.4, that likely contribute to these viruses characteristic biological and pathological traits. This notably includes differing capabilities of HPV16 and HPV18 E6/E7 to bind and regulate TLR9 and IRF3, which are both critical upstream elements of the type I IFN signalling cascade [582]. Here, studies by Hasan, Bates [324] have shown that HPV16 E6/E7 inhibit TLR9 expression with much greater efficiency than HPV18 E6/E9 in primary human keratinocytes, resulting in functional loss of downstream TLR9 signal transduction. In addition, while HPV16 E6 has been shown to directly

bind and potently inhibit transcriptional activities of IRF3 [322], only weak IRF3 regulatory functions have been observed by HPV18 [323].

Findings presented in Chapter 4 also determined significantly higher transcription, translation, and activity of E7 in HPV16-HTK lines in comparison to donor-equivalent HPV18-HTK lines (see section 4.7). We therefore theorise that these differences in E7 levels, perpetuated by higher activity of the HPV16 early viral promoter and comparably higher splicing of E6/E7-encoding mRNAs (sections 4.3.1, 4.3.2, and 4.6), could have contributed to greater HPV16 E7-mediated deregulation of the cellular immune response.

5.11.3.1 HPV16 downregulates expression of IFN-α-inducible genes in primary HTK lines

The type I-IFN cytokines, IFN-α and IFN-β are the most important cellular antiviral response mediators that induce downstream transcription of a diverse range of immune effector genes, termed IFN-stimulated genes (ISGs) [583]. Studies have emphasised an important role for the type-I IFN response in the spontaneous clearance of HPV, where loss of viral episomes during serial passaging of W12 cells was witnessed alongside an unprovoked spike in type-I IFN-inducible genes, or separately following IFN-β treatments [456, 458]. As such, strong HPV-mediated inhibition of type-I IFN activity are likely to have significant biological consequences *in vivo*.

While both HPV16 and HPV18 were both shown to inhibit the IFN-α signalling pathway in HTK (Figure 5-5), expression of these IFN-α-inducible genes were suppressed to a significantly greater extent in HPV16-HTK than HPV18-HTK (Figure 5-6), suggesting stronger HPV16-

mediated impairment of the pathway. Modulation of the IFN-α response is a well-known immune avoidance strategy of HR-HPVs, for which a number of viral oncoprotein-mediated mechanisms have been described [361, 362]. Reported upstream targets of E6/E7 include PRRs such as TLR9 and STING, and transcriptional regulators including NF-kB, and IRFs, notably IRF3 and IRF7, which are master regulators of type-I IFN transcription (see introductory section 1.4.5, Table 1-1). It is important to note that expression of *IFNA1* itself, which encodes IFN-α was not detected in any HTK or HPV-HTK cell line. However, expression of *IFNB1*, which encodes IFN-β, was significantly reduced in HPV16-HTK but not HPV18-HTK when compared with controls ($p = 0.029$, $\log_2$ fold change = -0.95) (supplementary Table S-2).

HPV16/18 E6/E7 upregulate expression of *UCHL1*, a deubiquitinating enzyme that strongly inhibits activation of TRAF3 and TRAF6, which are required for activation of IRF3 and IRF7 [374, 375]. In the RNA-Seq data, in comparison to untransfected-HTK, *UCHL1* was shown to be the first and second most significantly upregulated gene in HPV18-HTK and HPV16-HTK respectively, likely contributing to impeded IFN-α signalling in both virus-type lines (Figure 5-4). However, this finding does not clarify observed differences in IFN-α-regulation between HPV16- and HPV18-HTK. As previously discussed, differential targeting of TLR9 and IRF3 could likely have contributed to lower IFN-α activities in the HPV16-HTK lines. Other contributing factors likely included lower activity of NF-kB, which partners with specific IRFs to enhance IFN and ISG production [584], and lower expression of toll-like receptors (TLRs), notably *TLR3* and *TLR4*, which trigger downstream activation of IFN-, IRF-, and NF-kB [585].

5.11.3.2 HPV16 downregulates expression of NF-kB-inducible genes in primary HTK

NF-kB is another critical transcriptional mediator of the cellular immune response that is normally activated by one of two signalling pathways, the canonical or non-canonical pathway, which are triggered by numerous inflammatory stimuli, resulting in upstream activation of virus-sensing PRRs, TNF-receptors (TNFRs), and cytokine receptors, notably including IL-1 family receptors (IL-1R) [586, 587]. Activated NF-kB elicits potent pro-inflammatory and immune stimulatory effects by inducing transcription of a wide range of immune genes, including cytokines and cytokine receptors, as well as genes that promote type-I IFN and JAK-STAT signalling pathways, and MHC antigen presentation [588]. In addition, NF-kB promotes expression of genes that support cellular proliferation and development and inhibit apoptosis and necrosis [589, 590].

HPV16 oncoprotein-mediated suppression of NF-kB activation has been reported to be highly advantageous for immune avoidance during early HPV infection of primary keratinocytes of the human cervical transformation zone [591]. Importantly, HPV16- and HPV18-E7 has been shown to both directly and indirectly impair canonical TNF-α-mediated NF-kB translocation and activation, notably via association with essential NF-kB coactivators such as IkB and the NF-kB p65 subunit [295, 382, 387, 592, 593]. It is therefore feasible that differences in NF-kB activity between HPV16- and HPV18-HTK (Figure 5-6), could reflect differences in E7 expression, as demonstrated in Chapter 4 (see Figure 4-7 and Figure 4-8). However, activation of NF-kB is also induced following activation of TLRs, CD40, and IL-1 receptor complexes upon interaction with IL-1α and IL-1β, which were all transcriptionally repressed by HPV16 [586, 587, 594].

Little is currently known about how HR-HPV interacts with NF-kB during an oropharyngeal infection. Interestingly, NF-kB activity has been reported to be significantly lower in HPV-OPSCCs than in HPV-negative OPSCCs, which may indicate some level of HPV-mediated suppression [518]. Yet, constitutive NF-kB activation has also been identified in a subset of HPV-OPSCC and is thought to support carcinogenesis and chemoresistance [356, 595, 596]. The findings presented here suggest that HPV16, but not HPV18, hampers activation of NF-kB in transfected HTK, perhaps indicating more inhibitory activities of NF-kB during early infection of the tonsils, which may play an important role in early viral persistence.

5.11.3.3 HPV16 downregulates expression of IFN-γ-inducible genes in primary HTK lines

The type-II interferon, IFN-γ is primarily secreted by lymphocytes to stimulate adaptive antiviral responses in keratinocytes [597, 598]. It is a critical pathway for the T helper cell response, enhancing antigen presentation by increasing expression of MHC-I and MHC-II genes in virus-infected epithelial cells [599, 600].

In our HTK lines, HPV16 was shown to uniquely suppress IFN-γ signalling, leading to impaired expression of IFN-γ-inducible immune effector genes including *CCL2, CD40, IL7, IRF7, MX1, MX2*, as well as genes involved in MHC-I and -II antigen presentation (Figure 5-6 and Figure 5-7). This was somewhat surprising, as unlike type-I IFNs, IFN-γ is not thought to be produced by keratinocytes. However, it is possible that some IFN-γ-stimulation could have occurred from J2-fibroblasts or bovine serum [601, 602]. *IFNG* was not detected in HTK by RNA-Seq, but these cells had been well cleansed of fibroblasts and growth media prior to sequencing. To

clarify these findings, it would be interesting to treat these cells with IFN-γ and further examine differences in expression of some of these IFN-γ-inducible genes.

5.11.3.4 Concluding remarks and future study of IFN-α, IFN-γ and NF-kB activity in HPV-HTK

Impairment of IFN and NF-kB activities are likely critically important mechanisms for immune evasion in the tonsils that could reflect a significant distinction between HPV16 and HPV18 persistence at this body site. As such, further validation needs to be carried out to directly compare IFN-α and NF-kB activities between HPV16-HTK, HPV18-HTK and the control donor cells. ELISA-based methods, as described by Rodero, Decalf [603], could be used to sufficiently compare IFN-α activity. In addition, staining organotypic HPV-HTK raft sections to compare NF-kB cellular location should allow for identification of activated NF-kB. Here, detection of NF-kB in the nucleus would support successful nuclear translocation, strongly indicating NF-kB activation [604]. It would be interesting to reassess expression of some specific IFN-γ-induced genes by qPCR following IFN-γ treatment, to account for a lack of IFN-γ-producing immune cells.

5.11.4 HPV16 transcriptionally represses expression of important immune signalling elements

Differential gene expression and interaction network analysis of our HPV-HTK lines highlighted a selection of IFN-α, IFN-γ and NF-kB-inducible genes that had been significantly downregulated by HPV16 and play fundamental roles in the innate and adaptive immune response, as discussed below.

5.11.4.1 Cell surface expression of CD40 was suppressed at a transcriptional level in HPV16-HTK lines

Efficient crosstalk between keratinocytes and lymphocytes are essential for viral clearance. CD40 is a tumour necrosis factor receptor (TNFR) family member that is highly expressed on the surface of oral and cervical basal keratinocytes and plays a broad and active role in the innate and adaptive immune response [569-571, 605-607]. In the tonsils, ligation of CD40 to CD40-ligands (CD40L), presented on the surface of T-lymphocytes, crucially stimulates both T-cell and keratinocyte antiviral responses [608]. In keratinocytes, this ligation event stimulates downstream activation of NF-kB, MAPK, JAK/STAT, and PI3K signalling pathways, resulting in upregulated expression of cytokines, chemokines, and MHC components, along with potent induction of the inflammatory response [605, 607, 609-611].

Increased CD40-CD40L ligation reportedly enhances immune targeting and destruction of HPV-containing cervical cancer cells via activation of CD8+ T-cells, highlighting CD40 as an important chaperone of antiviral immunity [571, 612]. However, both HPV16- and HPV18-E6 have been shown to significantly limit the CD40-CD40L response in cervical and foreskin keratinocytes [613], possibly via E6-mediated inhibition of the AT-hook transcription factor (AKNA), which promotes transcriptional activities at *CD40* promoter regions [614]. However, neither study reported dissimilarities in *CD40* levels or activities between HPV16- and HPV18-containing cell lines.

In our HPV-HTK lines, genomic and protein-based studies revealed novel differences in *CD40* expression and protein production, which was uniquely suppressed by HPV16 (Figure 5-15 and Figure 5-16). Consequential reductions in CD40 cell surface expression were also uniquely observed in the basal layers of HPV16-HTK raft culture models, likely hampering prospective CD40-CD40L binding and downstream activation of NF-kB, MAPK, JAK/STAT, and PI3K signalling pathways. However, further studies will be required to confirm differences in functional CD40 activities between HPV16- and HPV18-HTK following stimulation with CD40-L. *CD40* transcription has been shown to be highly induced by IFN-γ in both normal and HPV-transfected keratinocytes [570, 607, 615], as well as by IL-1β, IL-1α, NF-kB, TGF-β and TNF-α [616-620]. It is therefore likely that comparatively lower *CD40* levels in HPV16-HTK were associated with reduced IFN-γ or NF-kB activities, and potentially lower expression of IL-1β and IL-1α.

5.11.4.2 HPV16 uniquely downregulated a network of genes involved in MHC antigen presentation

MHC class I (MHC-I) and class II (MHC-II) complexes display antigen peptide fragments and interact with T-cell receptors located on the cell surface of CD8+ cytotoxic T-cells and CD4+ helper T-cells respectively. These are critical mechanisms for T-cell activation that prompt immediate cell destruction by cytotoxic T-cells, as well as helper T-cell-mediated activation/recruitment of macrophages, B-cells, and cytotoxic T-cells [384]. MHC complexes are functionally expressed on the surface of antigen presenting cells and keratinocytes, and their expression is centrally mediated by NF-kB, and by type-I and type-II IFNs [385, 386, 621, 622].

To evade lymphocyte recognition, HPV16- and HPV18-E7 reportedly suppress MHC-I transcription, including via E7-mediated suppression of NF-kB, which has been shown to both impede MHC-I surface expression and NKT detection [364, 623-625]. This involves transcriptional repression of *HLA-A*, *HLA-B*, *HLA-C,* as well as *TAP1* which is an essential chaperone for processing and loading of antigens onto MHC-I complexes [626]. Our RNA-Seq analyses highlighted a similar network of differentially expressed genes involved in MHC-I antigen presentation, which were all significantly downregulated by HPV16 but not HPV18 in HTK (Figure 5-7 and Figure 5-8). These included comparable MHC-I components and associators, *HLA-B*, *HLA-G*, *B2M*, and *TAP1*.

In addition, alternative expression of genes involved in the MHC-II antigen presentation pathway were also observed, showing higher expression of the MHC-II components *HLA-DRA*, *HLA-DRB1,* and *HLA-DPA1* in HPV18-HTK than in HPV16-HTK (supplementary Table S-2). This was also true for *CD74,* which is required for maturation of MHC-II molecules. Altogether, these findings potentially indicate that higher activity of NF-kB and IFN signalling pathways in HPV18-HTK supported more productive MHC-I and -II production than in HPV16-HTK. Importantly, lower MHC-I expression is a hallmark of HPV-OPSCCs [627], potentially highlighting this differential regulation as an important distinction between the two virus-types in the tonsils. However, further validation and comparison of MHC protein production and surface expression are still required.

5.11.4.3 HPV16 uniquely inhibited expression of CCL2 in HTK

Numerous cytokines and chemokines were shown to be expressed at significantly lower levels in HPV16-HTK lines than in HPV18-HTK lines, including *IL1A*, *IL1B*, *IL15,* and *IL7* as well as associated receptors, *IL18R1, IL13RA2, IL31RA, IL7R* and *IL15RA* (Figure 5-7, Figure 5-8, and supplementary Table S-2). This also included *CCL2*, which is a highly potent chemoattractant and regulator of lymphocytes and myeloid cells, and a critical element of the adaptive immune response [567, 628]. Crucially, *CCL2* transcription has been shown to be frequently eradicated in both cervical cancers and E6- and E7-containing primary cervical cells, highlighting this cytokine as an important oncoprotein target [629-631].

Expression of *CCL2* was almost completely and uniquely lost in our HPV16-HTK lines. No direct mechanism of HR-HPV-mediated inhibition of *CCL2* has yet been described. However, transcription of *CCL2* is induced by numerous pro-inflammatory stimuli including NF-kB, IL-1, IL-4, IL-6, TNFα, TGFβ, IFN-γ, VEGF, and MTORC1, many of which are regulated by HR-HPV [567, 632, 633]. It is therefore possible that alternate activities/expression of NF-kB, MTORC1, IFN-γ, or IL-1, observed in the HPV16/18-HTK lines could have contributed to these reduced *CCL2* levels.

5.11.5 HPV16 uniquely downregulated expression of inflammatory response genes in HTK

During inflammation, keratinocytes stimulate production and secretion of pro-inflammatory cytokines and chemokines to promote activation and recruitment of immune cells from surrounding tissues and the blood stream. However, in anogenital models and cancers, HR-HPVs have been shown to promote more anti-inflammatory environments via modulation of

pro-inflammatory PRRs and transactivators including NF-kB, type-I IFNs and STATs [634, 635]. In our HPV-HTK lines, the inflammatory response pathway was shown to be significantly repressed in HPV16-HTK in comparison to HPV18-HTK and untransfected cells (Figure 5-5 and Figure 5-6), something that was likely driven by reduced activity of the NF-kB, IL6-JAK-STAT3, and IL2-STAT5 pro-inflammatory pathways.

The IL-1 cytokine family are critical pro-inflammatory mediators that are constitutively expressed in basal keratinocytes and are amongst the most prolific and active keratinocyte-derived cytokines in the skin that also potently induce the canonical NF-kB signalling cascade [636-641]. In our HTK lines, HPV16 was shown to uniquely alter expression of *IL1A, IL1B, IL33, IL36A*, and the cytokine receptor, *IL18R1* (supplementary Table S-2), in a manner that could impair inflammation and NF-kB activation. Crucially, these IL-1 proteins have all been shown to play important pro-inflammatory roles in the epidermis and are known to be transcriptionally modified by HR-HPVs [642]. Given their collective importance in epithelial immunity, proposed HPV16-mediated suppression of all these IL-1 genes has the potential to substantially weaken inflammatory effects and antiviral immune responses, contributing to viral persistence and thus highlighting this gene set as an important area for future study.

5.11.6 HPV16 and HPV18 appear to differentially favour immune avoidance and cell growth/survival pathways

In addition to their roles in inflammation, the TNFα/NF-kB, IL6- JAK-STAT3, and IL2-STAT5 signalling pathways are all important promoters of cellular proliferation, differentiation, and

survival. Furthermore, these pathways are commonly found to be overactive in primary HPV-infected anogenital cell lines and cervical cancers and are also thought to directly contribute to cervical carcinogenesis [643-646]. It was therefore somewhat surprising to find in our HTK lines that these pathways had been downregulated by HPV16, which is present in the vast majority of HPV-HNSCCs, in comparison to HPV18, which is rarely detected in HNSCCs [31, 32, 40, 282, 286]. Moreover, HPV18 also showed comparably higher activity of several other pro-proliferative/cell growth pathways including MTORC1 and KRAS signalling (Figure 5-6), which are also deregulated in HPV-driven cancers, including HPV-HNSCCs [647, 648].

Altogether, these findings appear to suggest that HPV18 favours higher activity of such cell growth/survival pathways over immune avoidance during early infection of the tonsils. In contrast, HPV16 seems to instead prioritise immune avoidance strategies over such pro-proliferative and pro-survival activities at this anatomical site. Since the tonsils contain such high levels of immune cells, more effective immune avoidance strategies, like those implemented by HPV16, could be crucial for more productive and persistent infections at this anatomical site. However, longer lasting infections may also increase long-term risk of virus-induced carcinogenic alterations, potentiated by sustained activity of the HPV oncoproteins and ongoing genomic instability.

The findings presented in Chapter 4 and 5 altogether highlight differing immune avoidance strategies between HPV16 and HPV18 in the tonsils. Here, HPV16 appeared to promote a more aggressive and direct method of immune avoidance via more potent E6/E7-mediated alteration of the transcriptional immune network, perhaps aided by higher activity of the early

viral promoter, along with higher levels and activity of E7 (see sections 4.3 - 4.7). In contrast, HPV18 appeared to limit expression of early viral genes and viral genome copy numbers, perhaps instead favouring lower production of immunogenic viral antigens as an immune avoidance strategy, limiting activation of PRRs, and downstream NF-kB and IFN-I signalling. It is not clear whether similar differences are observed in other anogenital subsites. However, in the tonsils, it is possible that higher E7 activities and greater suppression of the immune response could provide a better survival advantage that could help to explain why HPV16 but not HPV18 contributes to so many more HNSCCs.

5.11.7 Comparison of HPV16-mediated immune regulations in HTK, HPV-HNSCC, and cervical tissues

In the HTK lines, HPV16 was shown to suppress expression or activities of a number of immune genes/proteins that are also commonly impaired in HPV-HNSCCs. This suggests that HPV16 deregulates these immune responses from early infection and could further exemplify the importance of HPV16-mediated impairment of the antiviral response in order to persist and promote cancer at this anatomical site.

Akin to our HTK lines, STAT3 expression has been demonstrated to be expressed at relatively low levels in HPV-OPSCCs, particularly in comparison to HPV-negative OPSCC and cervical cancers [419, 518]. Like our studies, these differences were thought to be attributed to lower activity of NF-kB, which has also reported to be significantly less active in HPV-OPSCCs than in HPV-negative OPSCCs, further indicating HPV-mediated suppression of the pro-inflammatory

factor [518]. However, constitutive NF-kB mutational activation has also been identified in a subset of HPV-OPSCCs and is thought to support carcinogenesis and chemoresistance in an alternate manner of cancer progression [356, 595, 596]. Additional examples of immune genes that were inhibited in both our HPV16-HTK lines and HPV-OPSCCs include *IL1B* [649, 650], *CCL2* [629-631], and MHC-I. Importantly, reduced expression of MHC-I has been described as a hallmark of HPV-OPSCCs [627].

HPV-mediated transcriptional reprogramming has been well studied in low- and high-grade anogenital lesions, cancers and primary anogenital models, which have provided insight into the natural history of HPV and HPV-associated disease progression in the cervix [572, 651-655]. Comparison of our transcriptomic data with other global gene expression studies in primary cervical models have highlighted many similarities in how HPV16 reorganises host immune transcription profiles at these distinct anatomical sites. Importantly, HPV16 has been shown to broadly downregulate expression of both IFN- and NF-kB-inducible genes, such as *IL1A, IL1B, SERPINE1, and S100A7,* at both sites [572, 651, 652, 656] (Table S-2). This suggests that HPV16 downregulates immune responses in a largely similar manner during early infection of tonsil and cervical epithelia. However, since HPV18 are commonly found in cervical cancers but are rarely detected in HPV-driven tonsil cancers, it seems likely that these HPV16-specific immune regulatory functions could be particularly important to persist in the tonsils. It would be interesting to compare host mRNA transcription profiles in primary HPV16- and HPV18-transfected cervical lines, to understand whether these differences in immune regulation are specific to the tonsils or are also present in the cervix.

5.12 Conclusions

Effective immune evasion strategies are likely of particular importance in the tonsils and tonsil crypts, which are highly immunocompetent lymphoid organs that maintain high concentrations of lymphocytes and other immune sentinel cells [354, 355]. Nevertheless, HPV16-driven cancers have been shown to consistently arise at this anatomical subsite [34, 35, 56]. These studies suggest that from early infection, HPV16 regulates host HTK mRNA levels, in an alternative manner to HPV18, to create a more baseline anti-inflammatory and immune-repressive environment. Differential regulation of immune responses between these two HR-HPV types may be a key distinction for viral success, persistence, and subsequent oncogenesis at this body site that could underlie differences in disease progression between HPV16 and HPV18 in the tonsils.

5.13 Future study

Altogether, these data have highlighted a large number of exciting directions of research. However, the findings presented in this chapter have been largely explorative and require much more confirmational study. Importantly, increasing the total number of biological repeats by production of additional isogenic HPV-HTK donor lines is essential to confirm the overall validity of these findings. In addition, global transcriptomic analysis of differentiated HPV-HTK model organotypic raft culture could also provide a broader picture of the transcriptional landscape more similar to *in vivo* infection. Other potential and important project directions have been discussed in section 6.6.

qPCR was performed to validate transcription of a select number of genes that appeared to be differentially expressed between HPV16-HTK and HPV18-HTK. While these trends were largely consistent with the RNA-Seq data, given the limited number of available biological repeats, it will still be important to carry out additional mRNA and protein-based validations of other distinguished unvalidated differentially expressed genes. This includes notable cytokines and chemokines such as *CCL2, IL7, TLR3*, *IFNB1,* and IL-1 family member genes, as well as genes involved in MHC-I and MHC-II antigen presentation. It would also be interesting to confirm differences in STAT3 and STAT5 phosphorylation between these isogenic HPV-HTK lines. Importantly, further experimental validation will be required to directly compare activity of central transcriptional regulators such as NF-kB and IFN-α, as previously described in section 5.11.3.4.

Given that these studies have highlighted the importance of immune regulation by HPV16 in the tonsils, this could highlight specific immunotherapies as potentially useful strategies for the treatment of HPV-OPSCCs. Immunotherapies such as pembrolizumab and nivolumab, which target the PD-1/PD-L1 checkpoint axis, have already shown great promise in treating recurrent or metastatic HNSCCs (see section 1.1.5). However, the findings presented in this thesis highlight several other possible targets worth investigating, including HPV16-E7, CD40, CCL2, and TNF-α/NF-kB. Potentially relevant clinical trials that are currently underway include: HPV16-E7 (NCT03978689, NCT04287868, NCT03258008, NCT02865135, and NCT02858310), and CD40 (NCT02225002). Interestingly, in primary HPV16- and HPV18-immortalised HFK lines, HPV16-HFK was reported to be more sensitive to TNF-α treatment than HPV18-HFK, where HPV18-HFKs demonstrated much greater resistance to this inhibitory cytokine. This

further highlights the TNF-α/NF-kB pathway as a potentially important immunotherapy target in HPV-OPSCCs, which are primarily driven by HPV16 [657].

Finally, these RNA-Seq data also highlighted intriguing differences between HPV16-HTK and HPV18-HTK lines in the regulation of numerous other pathways. Many of these revolved around cellular metabolism. Further differences were observed in gene sets defining apoptosis, EMT, and KRAS signalling. Differential transcriptional regulation of each one of these pathways may provide important novel insights into alternative HPV16 and HPV18 carcinogenic potential.

Chapter 6: Final discussion

6.1 Establishment of isogenic primary HPV-HTK models

HPV16 accounts for the vast majority of HPV-OPSCC and is also the most commonly identified HR-HPV type in cervical cancers. However, while HPV18 is the next most common virus type in cervical cancers, comprising nearly a fifth of global cases, this virus is rarely detected in OPSCC [31, 32, 40, 282, 286]. These disparities exemplify differences in the natural history of disease driven by these two viruses and highlight a key question in the field; how is HPV16 causing so many more cancers in the oropharynx than other common HR-HPV types? To begin to characterise the viral infectious cycle and HR-HPV-driven functional changes in the oropharynx, the Roberts group established cell infection models from primary human tonsil keratinocytes (HTK) - the primary site of HPV-driven carcinogenesis in the oropharynx, transfected with either HPV16 or HPV18 genomes [34, 35, 56].

We hypothesised that HPV16 and HPV18 differentially manipulate host tonsil keratinocytes, leading to differences in both viral persistence and pathogenesis. The studies presented in this thesis aimed to characterise and compare HPV16 and HPV18 life cycles, viral transcriptome organisation, and virus-induced alteration of host transcriptomes in the primary HTK models.

Longitudinal passage and Southern blotting of the three isogenic HPV16- and 18-HTK donor models, revealed intriguing differences in HPV16 and HPV18 genome maintenance. These findings indicate a propensity for integration of HPV18 genomes following viral episome

establishment and maintenance, whereupon viral DNA integrated between 40 to 70 population doublings. In contrast HPV16 presented much more variable integration tendencies between each of the three donors, varying from no integration, very early, or late integration during cell passage. These variations in establishment/maintenance replication cycles could underline differences in biology between the two viruses at this anatomical site and might suggest HPV18 genomes as unable to persist long-term as episomes in tonsil keratinocytes. However, this finding is currently limited by a low number of biological repeats and will require further exploration.

6.2 Characterisation and comparison of HPV16 and HPV18 life cycles in primary HTK models

Organotypic raft culture was developed from HPV16- and HPV18-HTK cells containing viral episomes and stratified into mucosal-like epithelia. HTK#2 raft culture was shown to support both HPV16 and HPV18 life cycles, as signified by production of both early and late viral proteins. However, HTK#3 did not; instead presenting morphological characteristics reminiscent of high-grade dysplasia, as signified by less obvious morphological differentiation, deregulated E7 oncoprotein expression that extended the full thickness of the raft and failed expression of viral proteins associated with the productive stages of the virus life cycle (section 3.5.3). This transpired upon repeated rafting of both HPV16- and HPV18-containing HTK and may indicate HTK#3 cells as being more susceptible to HPV-driven high-grade dysplasia. An observation that may be further evidenced by significant differential expression of EMT signatures between the two donors.

When comparing HTK#2 raft cultures, very little variation was observed between HPV16 and HPV18 life cycles overall. However, comparison with an HPV18-positive primary foreskin keratinocyte model (HPV18-HFK) raft cultures revealed intriguing and novel differences in viral life cycle organisation between these anatomical sites. In comparison to HPV-HFK raft culture and previously published cervical models, key differentiation-dependent life cycle events, including viral genome amplification and increased production late phase viral proteins (E1^E4 and L1) occurred much earlier in epithelial stratification of HPV-HTK raft cultures [199, 235, 237, 248, 438]. Crucially, in HPV16/18-HTK#2 raft cultures, almost complete regional separation of E1^E4 and cytoplasmic cyclin B1, a marker of cells undergoing G2 arrest, suggests that E1^E4 was not contributing to G2 arrest functions in HTK, as has been previously defined in numerous anogenital-derived models [230, 231, 233, 234, 237]. E1^E4-induced G2 arrest greatly enhances viral genome amplification efficiency by creating a replication competent environment supportive of viral DNA replication. However, these studies may be indicating alternative, E1^E4-independent mechanisms of virus-induced G2 arrest in the tonsils, likely via E7 [220, 229].

HPV16 and HPV18 were shown to alter keratin expression profiles of HTK raft cultures, as signified by upregulated expression of the type II keratins, K7 and K8. These keratins are distinctive biomarkers of specialised tonsil crypt epithelium, which is the primary site of HPV-driven carcinogenesis in the tonsils [274, 277]. Importantly, no similar induction of these keratins occurred in HPV18-HFK, suggesting that this epidermal reprogramming event was unique to tonsil keratinocytes. Together, these data indicate that HR-HPV may alter the keratin profile of the tonsils to imitate a more crypt-like phenotype which may be beneficial

for the virus and could potentially highlight these HTK models as applicable models of HPV-infection in the head and neck which primarily arises in the tonsil crypts.

6.3 Characterisation and comparison of HPV16 and HPV18 transcriptome organisation in primary HTK models

Long-read Nanopore sequencing was used to both characterise and quantify full-length viral mRNA isoforms in undifferentiated basal-like HPV16/18-HTK models. This permitted examination of viral promoter usage, polyadenylation signalling, alternative splicing, as well as production of transcription maps, altogether supporting effective characterisation and comparison of HPV16 and HPV18 transcriptome organisation in the tonsils during early stages of the viral life cycle.

Quantification of viral transcripts between HPV16 and HPV18 indicate underlying virus-type-specific differences in transcriptome maintenance within the tonsils. Notably, the HPV16-HTK lines demonstrated higher usage of the early viral promoter (pE) than donor-equivalent HPV18-HTK lines, along with an increased propensity for splicing of the E6 open reading frame (ORF), together driving higher expression of E7-encoding transcripts. Consequentially, HPV16-HTK presented significantly more enhanced production of E7 oncoproteins than HPV18-HTK, in addition to increased activity of E7-stimulated cell cycle pathways, as shown by significantly higher expression of E2F- and G2/M-stimulated genes.

In comparison to HPV16, HPV18 supported higher expression of E8^E2, which is a principal and potent repressor of early viral transcription and DNA replication that physically and

mechanistically inhibits both the pE and viral origin of replication [146, 147]. Contrasting distinctions of pE8 promoter usage and abundance of E8^E2 transcripts between HPV16-HTK and HPV18-HTK may have greatly contributed to differential usage of the pE between these isogenic models. Further, viral genome copy numbers were shown to be 3- and 13-fold lower in HPV18-HTK lines than in corresponding HPV16-HTK donor cells, perhaps linked to increased activity of E8^E2.

Altogether, these data suggest that HPV18 was imposing greater restrictions on early viral transcription and viral genome replication than HPV16 in the primary HTK lines. Crucially, these disparities in viral genome copy numbers, early viral promoter usage, and expression of E7-encoding transcripts were consistent with differences in E8^E2 expression. Therefore, differential expression of E8^E2 could be an important distinction in viral transcriptome organisation between these two viruses in the tonsils, a hypothesis that requires additional validation and study, perhaps by inactivating mutations of E8^E2 in HPV18-HTK. Crucially, higher, and less regulated usage of the pE, along with greater production of E7 in HPV16-HTK could potentially contribute to increased persistence and/or disease progression of HPV16 in the tonsils.

Nanopore also allowed for identification and characterisation of virus-host fusion transcripts originating from transcriptionally active areas of the host genome. These ‘fusion’ transcripts were detected within just 30 – 40 population doublings of each assessed HPV16/18-HTK cell line and at multiple genomic locations, even within lines stably maintaining viral episomes. These findings indicate viral integration as a common and early event during transfection of

tonsil keratinocytes. Intriguingly, high numbers of fusion transcripts, detected in HPV18-HTK#3 and HPV16-HTK#2 cells, correlated to a small number of specific genomic locations, indicating that HPV16/18 DNA had integrated at very transcriptionally active regions of the host genome. These events appeared to greatly enhanced overall abundance of fusion transcripts. Further, the vast majority (75 – 80%) of virus-host fusion transcripts aligned to E6*I isoforms, which efficiently translate E7 [161], perhaps upregulating production of these transformative proteins in cell lines that presented high numbers of fusion transcripts. Altogether, this analysis has highlighted several potentially disease contributing events, which occurred early on following HPV16/18 transfection of the HTK models.

6.4 HPV16 and HPV18 differentially reorganise the host transcriptome in primary HTK

To better understand how HPV16 and HPV18 were differentially manipulating the host cellular environment, short-read RNA-sequencing (RNA-seq) was performed to characterise and compare HPV16- and HPV18-induced changes of the host HTK transcriptomes. Differential gene expression (DGE) analyses, including gene set enrichment analysis (GSEA) were carried out to identify significantly differentially expressed gene sets, which broadly signify altered function or activity of a select biological pathway. These analyses highlighted intriguing differences between HPV16 and HPV18 in HTK, which were shown to differentially regulate numerous pathways relating to the cellular immune response, metabolism, and KRAS signalling.

The most striking finding of these analyses centred around the cellular immune response. Here, HPV16 was shown to significantly repress downstream transcriptional activities of multiple antiviral pathways in comparison to both HPV18-HTK and untransfected-HTK controls. This notably included pathways relating to key transcriptional immune mediators, IFN-α, IFN-γ and NF-kB, which centrally regulate downstream expression of immune effector genes required for innate and adaptive antiviral responses, which are highly regulated by HPV16/18 E6 and E7 oncoproteins [343]. Additional DGE and interaction network analysis highlighted a number of IFN- and NF-kB-inducible genes, significantly downregulated by HPV16, which play critical roles in the keratinocyte-mediated antiviral response. This notably included *CD40*, *CCL2*, *IL1B, IFN1B, IL15*, *IL18R1,* and *TLR3,* as well as multiple genes involved in MHC class I and II antigen presentation. Differential expression of a select number of these genes was validated by qPCR. Moreover, HPV16-mediated inhibition of *CD40* was shown to hamper translation and cell surface expression of this essential immune-stimulating element.

DGE analyses suggested significantly reduced activities of several essential pro-inflammatory pathways in HPV16-HTK models when compared with HPV18-HTK models. In addition to IFN-α, IFN-γ and NF-kB, this also included IL6-JAK-STAT3 and IL2-STAT5 signalling pathways. Importantly, comparable downregulation of these pathways corresponded with significant reductions of inflammatory response-related genes in HPV16-HTK, which notably included crucial pro-inflammatory cytokines and IL-1 family members such as *IL1A, IL1B, IL33, IL36A*, and *IL18R1*.

Collectively, these findings suggest that HPV16 elicit more stringent anti-immune effects than HPV18 in HTK by extensively dampening the host transcriptional immune network. Effective immune avoidance strategies are critical for productive and persistent infections that allow for prolonged periods of potential transmission. This is particularly true for the tonsils, which are highly active and responsive lymphatic tissues. However, augmented HPV16 persistence may also increase long-term risk of virus-induced carcinogenic alterations. We therefore hypothesise that observed differential regulation of the transcriptional immune network may underlie perceived differences in disease progression between HPV16 and HPV18 at this anatomical site.

6.5 Conclusions

Put together, the findings presented in this thesis reveal distinct differences in viral transcriptome organisation between HPV16 and HPV18 in the tonsils that could potentially underlie diverging strategies of immune evasion. Higher expression of E8^E2, coinciding with reduced pE usage and lower copies of viral episomes, could suggest that HPV18 attempts to limit HTK immune responses by restricting production of immunogenic viral antigens. In contrast, HPV16 was shown to regulate the antiviral response in a much more direct manner via potent mediation of the transcriptional immune network, possibly reinforced by higher expression, translation, and activity of E7. Given that HPV16 accounts for >90% of HPV-OPSCC and that HPV18 is rarely detected at this body site [31], it is possible that higher E7 activities and more aggressive mediation of the immune response could provide an essential survival advantage at this body site.

6.6 Final thoughts

These explorative investigations of our primary HPV-HTK cell lines contribute to current understanding of the viral life cycle, viral transcriptome organisation and HR-HPV-driven functional changes during early infection of the tonsils. Moreover, comparison of HPV16 and HPV18 within the same donors overcomes limitations of donor variability, allowing for novel insights as to how HPV16 may be better persisting or promoting carcinogenesis at this distinct anatomical subsite.

Altogether, these studies present many exciting avenues for further exploration. However, some findings presented in this thesis are currently limited by a low number of available HTK donor models. Establishment of additional HPV16/18-HTK cell lines is currently underway in the Parish lab, which should allow for further validation of significant novel findings that have been outlined in this summary. In particular, contrasting pE8 activities and consequential disparities in E8^E2 expression between HPV16-HTK and HPV18-HTK could have greatly contributed to observed differences in virus and host-HTK transcriptome organisation, likely via influencing differential expression of the E7 oncogenes. Along with further initial validations of this described mechanism (described in section 4.11.3), investigating how the pE8 is being differentially regulated between HPV16- and HPV18-HTK will be crucial to understand this pivotal mechanism. Studies carried out by Straub, Fertey [502] suggest that HPV16 E1 and E2 play a role in pE8 regulation, in addition to host-cellular transcription factors of the erythroblast transformation specific (ETS) family. In the first instance, further investigation of differential levels of these viral proteins, as well as how HPV16 and HPV18 are regulating activity of ETS family members in HTK should be carried out.

Greater HPV16-mediated immune regulation could be a crucially important feature of this virus-type in the tonsils that requires further exploration. Being devoid of immune cell surveillance *in vitro*, it is likely that the HPV-HTK cells were receiving little external immune stimulation. Yet, *in vivo*, the tonsils maintain high levels of infiltrating lymphocytes [280]. Treatment of the HPV-HTK lines with immunostimulants such as poly(I:C), IFN-γ, or TNF-α, as previously demonstrated by Termini, Boccardo [658] and Karim, Meyers [361], could assess more practicable immunosuppressive effects of HPV16 in comparison to HPV18. Ideally, development of HTK raft culture/lymphocyte co-culture systems would provide a more *in vivo*-like model to effectively compare immune modulatory effectiveness of these two virus-types in the presence of responsive immune cells. Similar HNSCC co-culture systems have been successfully established by Dijkstra, Cattaneo [659] and Baruah, Bullenkamp [660].

Finally, a significant limitation of these primary HTK models is that they are derived from normal surface tonsil keratinocytes, as opposed to keratinocytes of specialised tonsillar crypt epithelia, where the vast majority of HPV-driven tonsil carcinomas originate [34, 35, 56]. Incredibly, investigation of our HPV-HTK raft culture models showed that HPV16/18 had reorganised HTK keratin expression profiles, promoting production of K7 and K8, which are characteristic of the tonsil crypts. This could highlight our HTK models as more suitable models of HPV-infection in the crypts than first thought. However, it is likely that these models still do not exhibit certain molecular characteristics of the crypts. Studies by Kang, Kannan [661] have described a method to isolate and identify tonsil crypt progenitor cells from surface progenitors, which were used to establish *in vitro* cell culture and organotypic raft culture

models. This could present a powerful method for investigation of HPV16-induced transformation events in the crypts. Here, transfection of both tonsil crypt and surface progenitors of the same donor with wild-type HPV16 (and HPV18) episomes could be carried out to assess differences in HPV life cycle, viral transcriptome organisations and virus-induced host transcriptome reorganisation, as was carried out in this project.

List of references

1. Schiffman, M., et al., *The carcinogenicity of human papillomavirus types reflects viral evolution.* Virology, 2005. **337**(1): p. 76-84.
2. Bernard, H.U., et al., *Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments.* Virology, 2010. **401**(1): p. 70-9.
3. zur Hausen, H., *Papillomaviruses in the causation of human cancers - a brief historical account.* Virology, 2009. **384**(2): p. 260-5.
4. Hughes, G. and N. Field, *The epidemiology of sexually transmitted infections in the UK: impact of behavior, services and interventions.* Future Microbiol, 2015. **10**(1): p. 35-51.
5. de Martel, C., et al., *Worldwide burden of cancer attributable to HPV by site, country and HPV type.* Int J Cancer, 2017. **141**(4): p. 664-670.
6. Garcia-Vallve, S., A. Alonso, and I.G. Bravo, *Papillomaviruses: different genes have different histories.* Trends Microbiol, 2005. **13**(11): p. 514-21.
7. Gottschling, M., et al., *Multiple evolutionary mechanisms drive papillomavirus diversification.* Mol Biol Evol, 2007. **24**(5): p. 1242-58.
8. Rollison, D.E., et al., *An Emerging Issue in Oncogenic Virology: the Role of Beta Human Papillomavirus Types in the Development of Cutaneous Squamous Cell Carcinoma.* J Virol, 2019. **93**(7).
9. Tommasino, M., *The biology of beta human papillomaviruses.* Virus Res, 2017. **231**: p. 128-138.
10. Forslund, O., *Genetic diversity of cutaneous human papillomaviruses.* J Gen Virol, 2007. **88**(Pt 10): p. 2662-2669.
11. Ekstrom, J., et al., *High throughput sequencing reveals diversity of Human Papillomaviruses in cutaneous lesions.* Int J Cancer, 2011. **129**(11): p. 2643-50.
12. Goon, P., et al., *Recurrent respiratory papillomatosis: an overview of current thinking and treatment.* Eur Arch Otorhinolaryngol, 2008. **265**(2): p. 147-51.
13. Gerein, V., et al., *Incidence, age at onset, and potential reasons of malignant transformation in recurrent respiratory papillomatosis patients: 20 years experience.* Otolaryngol Head Neck Surg, 2005. **132**(3): p. 392-4.
14. Lacey, C.J., C.M. Lowndes, and K.V. Shah, *Chapter 4: Burden and management of non-cancerous HPV-related conditions: HPV-6/11 disease.* Vaccine, 2006. **24 Suppl 3**: p. S3/35-41.
15. Mauz, P.S., et al., *HPV vaccination as preventive approach for recurrent respiratory papillomatosis - a 22-year retrospective clinical analysis.* BMC Infect Dis, 2018. **18**(1): p. 343.
16. Sherman, L., et al., *Inhibition of serum- and calcium-induced differentiation of human keratinocytes by HPV16 E6 oncoprotein: role of p53 inactivation.* Virology, 1997. **237**(2): p. 296-306.
17. Hebner, C.M. and L.A. Laimins, *Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity.* Rev Med Virol, 2006. **16**(2): p. 83-97.
18. Oh, S.T., M.S. Longworth, and L.A. Laimins, *Roles of the E6 and E7 proteins in the life cycle of low-risk human papillomavirus type 11.* J Virol, 2004. **78**(5): p. 2620-6.
19. Burk, R.D., A. Harari, and Z. Chen, *Human papillomavirus genome variants.* Virology, 2013. **445**(1-2): p. 232-43.
20. Chen, Z., et al., *Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67.* PLoS One, 2011. **6**(5): p. e20183.
21. Egawa, N. and J. Doorbar, *The low-risk papillomaviruses.* Virus Res, 2017. **231**: p. 119-127.
22. Chesson, H.W., et al., *The estimated lifetime probability of acquiring human papillomavirus in the United States.* Sex Transm Dis, 2014. **41**(11): p. 660-4.
23. Giuliano, A.R., et al., *Incidence and clearance of genital human papillomavirus infection in men (HIM): a cohort study.* Lancet, 2011. **377**(9769): p. 932-40.
24. D'Souza, G., et al., *Six-month natural history of oral versus cervical human papillomavirus infection.* Int J Cancer, 2007. **121**(1): p. 143-50.

25. Holowaty, P., et al., *Natural history of dysplasia of the uterine cervix.* J Natl Cancer Inst, 1999. **91**(3): p. 252-8.
26. Castellsague, X., *Natural history and epidemiology of HPV infection and cervical cancer.* Gynecol Oncol, 2008. **110**(3 Suppl 2): p. S4-7.
27. Mork, J., et al., *Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck.* N Engl J Med, 2001. **344**(15): p. 1125-31.
28. de Martel, C., et al., *Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis.* Lancet Glob Health, 2020. **8**(2): p. e180-e190.
29. Global Burden of Disease Cancer, C., et al., *Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study.* JAMA Oncol, 2019. **5**(12): p. 1749-1768.
30. Johnson, D.E., et al., *Head and neck squamous cell carcinoma.* Nat Rev Dis Primers, 2020. **6**(1): p. 92.
31. Schache, A.G., et al., *HPV-Related Oropharynx Cancer in the United Kingdom: An Evolution in the Understanding of Disease Etiology.* Cancer Res, 2016. **76**(22): p. 6598-6606.
32. Chaturvedi, A.K., et al., *Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers.* J Clin Oncol, 2013. **31**(36): p. 4550-9.
33. de Sanjose, S., et al., *Burden of Human Papillomavirus (HPV)-Related Cancers Attributable to HPVs 6/11/16/18/31/33/45/52 and 58.* JNCI Cancer Spectr, 2018. **2**(4): p. pky045.
34. Kim, S.H., et al., *HPV integration begins in the tonsillar crypt and leads to the alteration of p16, EGFR and c-myc during tumor formation.* Int J Cancer, 2007. **120**(7): p. 1418-25.
35. Haeggblom, L., et al., *Time to change perspectives on HPV in oropharyngeal cancer. A systematic review of HPV prevalence per oropharyngeal sub-site the last 3 years.* Papillomavirus Res, 2017. **4**: p. 1-11.
36. Ellington, T.D., et al., *Trends in Incidence of Cancers of the Oral Cavity and Pharynx - United States 2007-2016.* MMWR Morb Mortal Wkly Rep, 2020. **69**(15): p. 433-438.
37. Du, M., et al., *Incidence Trends of Lip, Oral Cavity, and Pharyngeal Cancers: Global Burden of Disease 1990-2017.* J Dent Res, 2020. **99**(2): p. 143-151.
38. Tinhofer, I., et al., *Contribution of human papilloma virus to the incidence of squamous cell carcinoma of the head and neck in a European population with high smoking prevalence.* Eur J Cancer, 2015. **51**(4): p. 514-521.
39. Chaturvedi, A.K., et al., *Human papillomavirus and rising oropharyngeal cancer incidence in the United States.* J Clin Oncol, 2011. **29**(32): p. 4294-301.
40. Mehanna, H., et al., *Prevalence of human papillomavirus in oropharyngeal and nonoropharyngeal head and neck cancer--systematic review and meta-analysis of trends by time and region.* Head Neck, 2013. **35**(5): p. 747-55.
41. Ndiaye, C., et al., *HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis.* Lancet Oncol, 2014. **15**(12): p. 1319-31.
42. Van Dyne, E.A., et al., *Trends in Human Papillomavirus-Associated Cancers - United States, 1999-2015.* MMWR Morb Mortal Wkly Rep, 2018. **67**(33): p. 918-924.
43. Kreimer, A.R., et al., *Oral human papillomavirus in healthy individuals: a systematic review of the literature.* Sex Transm Dis, 2010. **37**(6): p. 386-91.
44. Marur, S., et al., *HPV-associated head and neck cancer: a virus-related cancer epidemic.* Lancet Oncol, 2010. **11**(8): p. 781-9.
45. Pytynia, K.B., K.R. Dahlstrom, and E.M. Sturgis, *Epidemiology of HPV-associated oropharyngeal cancer.* Oral Oncol, 2014. **50**(5): p. 380-6.
46. Anantharaman, D., et al., *Geographic heterogeneity in the prevalence of human papillomavirus in head and neck cancer.* Int J Cancer, 2017. **140**(9): p. 1968-1975.
47. Ragin, C., et al., *Prevalence of HPV Infection in Racial-Ethnic Subgroups of Head and Neck Cancer Patients.* Carcinogenesis, 2017. **38**(2): p. 218-229.
48. D'Souza, G., et al., *Differences in the Prevalence of Human Papillomavirus (HPV) in Head and Neck Squamous Cell Cancers by Sex, Race, Anatomic Tumor Site, and HPV Detection Method.* JAMA Oncol, 2017. **3**(2): p. 169-177.
49. Castellsague, X., et al., *HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients.* J Natl Cancer Inst, 2016. **108**(6): p. djv403.

50. Li, H., et al., *Sex differences in patients with high risk HPV-associated and HPV negative oropharyngeal and oral cavity squamous cell carcinomas.* Cancers Head Neck, 2018. **3**: p. 4.
51. Milagre, I., et al., *Gender Differences in Global but Not Targeted Demethylation in iPSC Reprogramming.* Cell Rep, 2017. **18**(5): p. 1079-1089.
52. Au Yeung, C.L., et al., *HPV-16 E6 upregulation of DNMT1 through repression of tumor suppressor p53.* Oncol Rep, 2010. **24**(6): p. 1599-604.
53. Saunders, C.L., et al., *Associations Between Sexual Orientation and Overall and Site-Specific Diagnosis of Cancer: Evidence From Two National Patient Surveys in England.* J Clin Oncol, 2017. **35**(32): p. 3654-3661.
54. Hashim, D., et al., *Hormone factors play a favorable role in female head and neck cancer risk.* Cancer Med, 2017. **6**(8): p. 1998-2007.
55. Sabatini, M.E. and S. Chiocca, *Human papillomavirus as a driver of head and neck cancers.* Br J Cancer, 2020. **122**(3): p. 306-314.
56. Syrjanen, S., *HPV infections and tonsillar carcinoma.* J Clin Pathol, 2004. **57**(5): p. 449-55.
57. McIlwain, W.R., et al., *Initial symptoms in patients with HPV-positive and HPV-negative oropharyngeal cancer.* JAMA Otolaryngol Head Neck Surg, 2014. **140**(5): p. 441-7.
58. Mahal, B.A., et al., *Incidence and Demographic Burden of HPV-Associated Oropharyngeal Head and Neck Cancers in the United States.* Cancer Epidemiol Biomarkers Prev, 2019. **28**(10): p. 1660-1667.
59. Cooper, T., et al., *Association of keratinization with 5-year disease-specific survival in oropharyngeal squamous cell carcinoma.* JAMA Otolaryngol Head Neck Surg, 2015. **141**(3): p. 250-6.
60. Liu, C., et al., *The molecular mechanisms of increased radiosensitivity of HPV-positive oropharyngeal squamous cell carcinoma (OPSCC): an extensive review.* J Otolaryngol Head Neck Surg, 2018. **47**(1): p. 59.
61. Lewis, J.S., Jr., *Morphologic diversity in human papillomavirus-related oropharyngeal squamous cell carcinoma: Catch Me If You Can!* Mod Pathol, 2017. **30**(s1): p. S44-S53.
62. Mehanna, H., et al., *Oropharyngeal cancer: United Kingdom National Multidisciplinary Guidelines.* J Laryngol Otol, 2016. **130**(S2): p. S90-S96.
63. Ang, K.K., et al., *Human papillomavirus and survival of patients with oropharyngeal cancer.* N Engl J Med, 2010. **363**(1): p. 24-35.
64. 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.
65. Chaturvedi, A.K., et al., *Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States.* J Clin Oncol, 2008. **26**(4): p. 612-9.
66. Chaturvedi, A.K., *Epidemiology and clinical aspects of HPV in head and neck cancers.* Head Neck Pathol, 2012. **6 Suppl 1**: p. S16-24.
67. Beaty, B.T., et al., *PIK3CA Mutation in HPV-Associated OPSCC Patients Receiving Deintensified Chemoradiation.* J Natl Cancer Inst, 2020. **112**(8): p. 855-858.
68. Psyrri, A., et al., *Future directions in research, treatment and prevention of HPV-related squamous cell carcinoma of the head and neck.* Head Neck Pathol, 2012. **6 Suppl 1**: p. S121-8.
69. Mehanna, H., et al., *Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial.* Lancet, 2019. **393**(10166): p. 51-60.
70. Gillison, M.L., et al., *Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial.* Lancet, 2019. **393**(10166): p. 40-50.
71. Smalley Rumfield, C., et al., *Immunomodulation to enhance the efficacy of an HPV therapeutic vaccine.* J Immunother Cancer, 2020. **8**(1).
72. Frazer, I.H. and J. Chandra, *Immunotherapy for HPV associated cancer.* Papillomavirus Res, 2019. **8**: p. 100176.
73. Lechner, M., et al., *HPV-associated oropharyngeal cancer: epidemiology, molecular biology and clinical management.* Nat Rev Clin Oncol, 2022. **19**(5): p. 306-327.
74. Kalender, W.A., R. Hebel, and J. Ebersberger, *Reduction of CT artifacts caused by metallic implants.* Radiology, 1987. **164**(2): p. 576-7.

75. Morse, M.L., et al., *Metabolism of lactose by Staphylococcus aureus and its genetic basis.* J Bacteriol, 1968. **95**(6): p. 2270-4.
76. Kwon, M.J., et al., *Clinical implication of programmed cell death-1 ligand-1 expression in tonsillar squamous cell carcinoma in association with intratumoral heterogeneity, human papillomavirus, and epithelial-to-mesenchymal transition.* Hum Pathol, 2018. **80**: p. 28-39.
77. Lyford-Pike, S., et al., *Evidence for a role of the PD-1:PD-L1 pathway in immune resistance of HPV-associated head and neck squamous cell carcinoma.* Cancer Res, 2013. **73**(6): p. 1733-41.
78. Yang, Y., *Cancer immunotherapy: harnessing the immune system to battle cancer.* J Clin Invest, 2015. **125**(9): p. 3335-7.
79. Huang, P.W. and J.W. Chang, *Immune checkpoint inhibitors win the 2018 Nobel Prize.* Biomed J, 2019. **42**(5): p. 299-306.
80. Almangush, A., I. Leivo, and A.A. Makitie, *Biomarkers for Immunotherapy of Oral Squamous Cell Carcinoma: Current Status and Challenges.* Front Oncol, 2021. **11**: p. 616629.
81. Burtness, B., et al., *Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study.* Lancet, 2019. **394**(10212): p. 1915-1928.
82. Falcaro, M., et al., *The effects of the national HPV vaccination programme in England, UK, on cervical cancer and grade 3 cervical intraepithelial neoplasia incidence: a register-based observational study.* Lancet, 2021. **398**(10316): p. 2084-2092.
83. Chaturvedi, A.K., et al., *Effect of Prophylactic Human Papillomavirus (HPV) Vaccination on Oral HPV Infections Among Young Adults in the United States.* J Clin Oncol, 2018. **36**(3): p. 262-267.
84. Egawa, N., et al., *Human Papillomaviruses; Epithelial Tropisms, and the Development of Neoplasia.* Viruses, 2015. **7**(7): p. 3863-90.
85. Graham, S.V. and A.A.A. Faizo, *Control of human papillomavirus gene expression by alternative splicing.* Virus Res, 2017. **231**: p. 83-95.
86. Van Doorslaer, K., et al., *The Papillomavirus Episteme: a major update to the papillomavirus sequence database.* Nucleic Acids Res, 2017. **45**(D1): p. D499-D506.
87. Jia, R., et al., *Control of the papillomavirus early-to-late switch by differentially expressed SRp20.* J Virol, 2009. **83**(1): p. 167-80.
88. Middleton, K., et al., *Organization of human papillomavirus productive cycle during neoplastic progression provides a basis for selection of diagnostic markers.* J Virol, 2003. **77**(19): p. 10186-201.
89. Wang, H.K., et al., *Robust production and passaging of infectious HPV in squamous epithelium of primary human keratinocytes.* Genes Dev, 2009. **23**(2): p. 181-94.
90. Milligan, S.G., et al., *Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells.* Virology, 2007. **360**(1): p. 172-81.
91. Wang, X., et al., *Construction of a full transcription map of human papillomavirus type 18 during productive viral infection.* J Virol, 2011. **85**(16): p. 8080-92.
92. Coupe, V.M., et al., *Transcriptional analysis of human papillomavirus type 16 in histological sections of cervical dysplasia by in situ hybridisation.* J Clin Pathol, 2012. **65**(2): p. 164-70.
93. Xue, Y., et al., *HPV16 E2 is an immediate early marker of viral infection, preceding E7 expression in precursor structures of cervical carcinoma.* Cancer Res, 2010. **70**(13): p. 5316-25.
94. Ozbun, M.A. and C. Meyers, *Human papillomavirus type 31b E1 and E2 transcript expression correlates with vegetative viral genome amplification.* Virology, 1998. **248**(2): p. 218-30.
95. Johansson, C., et al., *HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation.* EMBO J, 2012. **31**(14): p. 3212-27.
96. Toots, M., et al., *The transcription map of human papillomavirus type 18 during genome replication in U2OS cells.* PLoS One, 2014. **9**(12): p. e116151.
97. Doorbar, J., et al., *Detection of novel splicing patterns in a HPV16-containing keratinocyte cell line.* Virology, 1990. **178**(1): p. 254-62.
98. Ferguson, J., et al., *The chromatin insulator CTCF regulates HPV18 transcript splicing and differentiation-dependent late gene expression.* PLoS Pathog, 2021. **17**(11): p. e1010032.
99. Glahder, J.A., et al., *A promoter within the E6 ORF of human papillomavirus type 16 contributes to the expression of the E7 oncoprotein from a monocistronic mRNA.* J Gen Virol, 2003. **84**(Pt 12): p. 3429-3441.

100. Wahl, M.C., C.L. Will, and R. Luhrmann, *The spliceosome: design principles of a dynamic RNP machine.* Cell, 2009. **136**(4): p. 701-18.
101. Matera, A.G. and Z. Wang, *A day in the life of the spliceosome.* Nat Rev Mol Cell Biol, 2014. **15**(2): p. 108-21.
102. Cartegni, L., S.L. Chew, and A.R. Krainer, *Listening to silence and understanding nonsense: exonic mutations that affect splicing.* Nat Rev Genet, 2002. **3**(4): p. 285-98.
103. Johansson, C. and S. Schwartz, *Regulation of human papillomavirus gene expression by splicing and polyadenylation.* Nat Rev Microbiol, 2013. **11**(4): p. 239-51.
104. Wang, Z. and C.B. Burge, *Splicing regulation: from a parts list of regulatory elements to an integrated splicing code.* RNA, 2008. **14**(5): p. 802-13.
105. Howard, J.M. and J.R. Sanford, *The RNAissance family: SR proteins as multifaceted regulators of gene expression.* Wiley Interdiscip Rev RNA, 2015. **6**(1): p. 93-110.
106. Busch, A. and K.J. Hertel, *Evolution of SR protein and hnRNP splicing regulatory factors.* Wiley Interdiscip Rev RNA, 2012. **3**(1): p. 1-12.
107. Chaudhury, A., P. Chander, and P.H. Howe, *Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles.* RNA, 2010. **16**(8): p. 1449-62.
108. Wang, Y., et al., *Mechanism of alternative splicing and its regulation.* Biomed Rep, 2015. **3**(2): p. 152-158.
109. Chen, J., et al., *Mapping of HPV transcripts in four human cervical lesions using RNAseq suggests quantitative rearrangements during carcinogenic progression.* Virology, 2014. **462-463**: p. 14-24.
110. Ajiro, M., et al., *Serine/Arginine-Rich Splicing Factor 3 and Heterogeneous Nuclear Ribonucleoprotein A1 Regulate Alternative RNA Splicing and Gene Expression of Human Papillomavirus 18 through Two Functionally Distinguishable cis Elements.* J Virol, 2016. **90**(20): p. 9138-52.
111. Kozak, M., *An analysis of vertebrate mRNA sequences: intimations of translational control.* J Cell Biol, 1991. **115**(4): p. 887-903.
112. Zheng, Z.M., et al., *Splicing of a cap-proximal human Papillomavirus 16 E6E7 intron promotes E7 expression, but can be restrained by distance of the intron from its RNA 5' cap.* J Mol Biol, 2004. **337**(5): p. 1091-108.
113. Remm, M., A. Remm, and M. Ustav, *Human papillomavirus type 18 E1 protein is translated from polycistronic mRNA by a discontinuous scanning mechanism.* J Virol, 1999. **73**(4): p. 3062-70.
114. Tang, S., et al., *The E7 oncoprotein is translated from spliced E6*I transcripts in high-risk human papillomavirus type 16- or type 18-positive cervical cancer cell lines via translation reinitiation.* J Virol, 2006. **80**(9): p. 4249-63.
115. Stacey, S.N., et al., *Translation of the human papillomavirus type 16 E7 oncoprotein from bicistronic mRNA is independent of splicing events within the E6 open reading frame.* J Virol, 1995. **69**(11): p. 7023-31.
116. Graham, S.V., *Keratinocyte Differentiation-Dependent Human Papillomavirus Gene Regulation.* Viruses, 2017. **9**(9).
117. Mole, S., et al., *RNA splicing factors regulated by HPV16 during cervical tumour progression.* J Pathol, 2009. **219**(3): p. 383-91.
118. Li, X., et al., *Eight nucleotide substitutions inhibit splicing to HPV-16 3'-splice site SA3358 and reduce the efficiency by which HPV-16 increases the life span of primary human keratinocytes.* PLoS One, 2013. **8**(9): p. e72776.
119. Somberg, M. and S. Schwartz, *Multiple ASF/SF2 sites in the human papillomavirus type 16 (HPV-16) E4-coding region promote splicing to the most commonly used 3'-splice site on the HPV-16 genome.* J Virol, 2010. **84**(16): p. 8219-30.
120. McFarlane, M., et al., *Human Papillomavirus 16 Oncoprotein Expression Is Controlled by the Cellular Splicing Factor SRSF2 (SC35).* J Virol, 2015. **89**(10): p. 5276-87.
121. McBride, A.A., *The papillomavirus E2 proteins.* Virology, 2013. **445**(1-2): p. 57-79.
122. Gauthier, J.M., J. Dillner, and M. Yaniv, *Structural analysis of the human papillomavirus type 16-E2 transactivator with antipeptide antibodies reveals a high mobility region linking the transactivation and the DNA-binding domains.* Nucleic Acids Res, 1991. **19**(25): p. 7073-9.
123. Ustav, M. and A. Stenlund, *Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames.* EMBO J, 1991. **10**(2): p. 449-57.

124. Phelps, W.C. and P.M. Howley, *Transcriptional trans-activation by the human papillomavirus type 16 E2 gene product.* J Virol, 1987. **61**(5): p. 1630-8.
125. Cripe, T.P., et al., *Transcriptional regulation of the human papillomavirus-16 E6-E7 promoter by a keratinocyte-dependent enhancer, and by viral E2 trans-activator and repressor gene products: implications for cervical carcinogenesis.* EMBO J, 1987. **6**(12): p. 3745-53.
126. Kim, K., et al., *Methylation patterns of papillomavirus DNA, its influence on E2 function, and implications in viral infection.* J Virol, 2003. **77**(23): p. 12450-9.
127. McPhillips, M.G., K. Ozato, and A.A. McBride, *Interaction of bovine papillomavirus E2 protein with Brd4 stabilizes its association with chromatin.* J Virol, 2005. **79**(14): p. 8920-32.
128. Steger, G. and S. Corbach, *Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein.* J Virol, 1997. **71**(1): p. 50-8.
129. Carson, A. and S.A. Khan, *Characterization of transcription factor binding to human papillomavirus type 16 DNA during cellular differentiation.* J Virol, 2006. **80**(9): p. 4356-62.
130. Soeda, E., et al., *Repression of HPV16 early region transcription by the E2 protein.* Virology, 2006. **351**(1): p. 29-41.
131. Romanczuk, H., F. Thierry, and P.M. Howley, *Mutational analysis of cis elements involved in E2 modulation of human papillomavirus type 16 P97 and type 18 P105 promoters.* J Virol, 1990. **64**(6): p. 2849-59.
132. Dong, G., T.R. Broker, and L.T. Chow, *Human papillomavirus type 11 E2 proteins repress the homologous E6 promoter by interfering with the binding of host transcription factors to adjacent elements.* J Virol, 1994. **68**(2): p. 1115-27.
133. Hou, S.Y., et al., *Alleviation of human papillomavirus E2-mediated transcriptional repression via formation of a TATA binding protein (or TFIID)-TFIIB-RNA polymerase II-TFIIF preinitiation complex.* Mol Cell Biol, 2000. **20**(1): p. 113-25.
134. Tan, S.H., et al., *The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID.* J Virol, 1994. **68**(10): p. 6411-20.
135. Stubenrauch, F., I.M. Leigh, and H. Pfister, *E2 represses the late gene promoter of human papillomavirus type 8 at high concentrations by interfering with cellular factors.* J Virol, 1996. **70**(1): p. 119-26.
136. Bouvard, V., et al., *Characterization of the human papillomavirus E2 protein: evidence of trans-activation and trans-repression in cervical keratinocytes.* EMBO J, 1994. **13**(22): p. 5451-9.
137. Fujii, T., et al., *High and low levels of cottontail rabbit papillomavirus E2 protein generate opposite effects on gene expression.* J Biol Chem, 2001. **276**(2): p. 867-74.
138. Chong, T., et al., *The enhancer of human papillomavirus type 16: binding sites for the ubiquitous transcription factors oct-1, NFA, TEF-2, NF1, and AP-1 participate in epithelial cell-specific transcription.* J Virol, 1991. **65**(11): p. 5933-43.
139. Silva, R.C.O., et al., *Functional evaluation of human papillomavirus type 31 long control region variants.* Genomics, 2020. **112**(6): p. 5066-5071.
140. Bauknecht, T., et al., *Identification of a negative regulatory domain in the human papillomavirus type 18 promoter: interaction with the transcriptional repressor YY1.* EMBO J, 1992. **11**(12): p. 4607-17.
141. O'Connor, M. and H.U. Bernard, *Oct-1 activates the epithelial-specific enhancer of human papillomavirus type 16 via a synergistic interaction with NFI at a conserved composite regulatory element.* Virology, 1995. **207**(1): p. 77-88.
142. Chong, T., W.K. Chan, and H.U. Bernard, *Transcriptional activation of human papillomavirus 16 by nuclear factor I, AP1, steroid receptors and a possibly novel transcription factor, PVF: a model for the composition of genital papillomavirus enhancers.* Nucleic Acids Res, 1990. **18**(3): p. 465-70.
143. Schweiger, M.R., et al., *Brd4-independent transcriptional repression function of the papillomavirus e2 proteins.* J Virol, 2007. **81**(18): p. 9612-22.
144. Smith, J.A., et al., *Genome-wide siRNA screen identifies SMCX, EP400, and Brd4 as E2-dependent regulators of human papillomavirus oncogene expression.* Proc Natl Acad Sci U S A, 2010. **107**(8): p. 3752-7.
145. Wu, S.Y., et al., *Brd4 links chromatin targeting to HPV transcriptional silencing.* Genes Dev, 2006. **20**(17): p. 2383-96.

146. Straub, E., et al., *The viral E8^E2C repressor limits productive replication of human papillomavirus 16.* J Virol, 2014. **88**(2): p. 937-47.
147. Kurg, R., et al., *Human papillomavirus E2 protein with single activation domain initiates HPV18 genome replication, but is not sufficient for long-term maintenance of virus genome.* Virology, 2010. **408**(2): p. 159-66.
148. Lace, M.J., et al., *The E8--E2 gene product of human papillomavirus type 16 represses early transcription and replication but is dispensable for viral plasmid persistence in keratinocytes.* J Virol, 2008. **82**(21): p. 10841-53.
149. Fertey, J., et al., *Growth inhibition of HeLa cells is a conserved feature of high-risk human papillomavirus E8^E2C proteins and can also be achieved by an artificial repressor protein.* J Virol, 2011. **85**(6): p. 2918-26.
150. Stubenrauch, F., T. Zobel, and T. Iftner, *The E8 domain confers a novel long-distance transcriptional repression activity on the E8E2C protein of high-risk human papillomavirus type 31.* J Virol, 2001. **75**(9): p. 4139-49.
151. Stubenrauch, F., et al., *Expression of E8^E2 Is Required for Wart Formation by Mouse Papillomavirus 1 in Vivo.* J Virol, 2021.
152. Powell, M.L., et al., *NCoR1 mediates papillomavirus E8;E2C transcriptional repression.* J Virol, 2010. **84**(9): p. 4451-60.
153. Dreer, M., et al., *Interaction of NCOR/SMRT Repressor Complexes with Papillomavirus E8^E2C Proteins Inhibits Viral Replication.* PLoS Pathog, 2016. **12**(4): p. e1005556.
154. Groves, I.J. and N. Coleman, *Pathogenesis of human papillomavirus-associated mucosal disease.* J Pathol, 2015. **235**(4): p. 527-38.
155. Burley, M., S. Roberts, and J.L. Parish, *Epigenetic regulation of human papillomavirus transcription in the productive virus life cycle.* Semin Immunopathol, 2020. **42**(2): p. 159-171.
156. Bodaghi, S., R. Jia, and Z.M. Zheng, *Human papillomavirus type 16 E2 and E6 are RNA-binding proteins and inhibit in vitro splicing of pre-mRNAs with suboptimal splice sites.* Virology, 2009. **386**(1): p. 32-43.
157. Kim, T.H., et al., *Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome.* Cell, 2007. **128**(6): p. 1231-45.
158. Braccioli, L. and E. de Wit, *CTCF: a Swiss-army knife for genome organization and transcription regulation.* Essays Biochem, 2019. **63**(1): p. 157-165.
159. Pentland, I., et al., *Disruption of CTCF-YY1-dependent looping of the human papillomavirus genome activates differentiation-induced viral oncogene transcription.* PLoS Biol, 2018. **16**(10): p. e2005752.
160. Paris, C., et al., *CCCTC-binding factor recruitment to the early region of the human papillomavirus 18 genome regulates viral oncogene expression.* J Virol, 2015. **89**(9): p. 4770-85.
161. Rosenberger, S., et al., *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, 2010. **107**(15): p. 7006-11.
162. Pim, D., P. Massimi, and L. Banks, *Alternatively spliced HPV-18 E6* protein inhibits E6 mediated degradation of p53 and suppresses transformed cell growth.* Oncogene, 1997. **15**(3): p. 257-64.
163. Filippova, M., et al., *The large and small isoforms of human papillomavirus type 16 E6 bind to and differentially affect procaspase 8 stability and activity.* J Virol, 2007. **81**(8): p. 4116-29.
164. Muthusami, S., et al., *A review on the role of epidermal growth factor signaling in the development, progression and treatment of cervical cancer.* Int J Biol Macromol, 2022. **194**: p. 179-187.
165. Schiller, J.T., P.M. Day, and R.C. Kines, *Current understanding of the mechanism of HPV infection.* Gynecol Oncol, 2010. **118**(1 Suppl): p. S12-7.
166. Kines, R.C., et al., *The initial steps leading to papillomavirus infection occur on the basement membrane prior to cell surface binding.* Proc Natl Acad Sci U S A, 2009. **106**(48): p. 20458-63.
167. Herfs, M., et al., *A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer.* Proc Natl Acad Sci U S A, 2012. **109**(26): p. 10516-21.
168. Martens, J.E., et al., *Distribution pattern and marker profile show two subpopulations of reserve cells in the endocervical canal.* Int J Gynecol Pathol, 2009. **28**(4): p. 381-8.
169. Reich, O., et al., *Defining the Cervical Transformation Zone and Squamocolumnar Junction: Can We Reach a Common Colposcopic and Histologic Definition?* Int J Gynecol Pathol, 2017. **36**(6): p. 517-522.
170. Doorbar, J., et al., *Human papillomavirus molecular biology and disease association.* Rev Med Virol, 2015. **25 Suppl 1**: p. 2-23.

171. Doorbar, J. and H. Griffin, *Refining our understanding of cervical neoplasia and its cellular origins.* Papillomavirus Res, 2019. **7**: p. 176-179.
172. Giroglou, T., et al., *Human papillomavirus infection requires cell surface heparan sulfate.* J Virol, 2001. **75**(3): p. 1565-70.
173. Culp, T.D., et al., *Keratinocyte-secreted laminin 5 can function as a transient receptor for human papillomaviruses by binding virions and transferring them to adjacent cells.* J Virol, 2006. **80**(18): p. 8940-50.
174. Combita, A.L., et al., *Gene transfer using human papillomavirus pseudovirions varies according to virus genotype and requires cell surface heparan sulfate.* FEMS Microbiol Lett, 2001. **204**(1): p. 183-8.
175. Bienkowska-Haba, M., H.D. Patel, and M. Sapp, *Target cell cyclophilins facilitate human papillomavirus type 16 infection.* PLoS Pathog, 2009. **5**(7): p. e1000524.
176. Richards, R.M., et al., *Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection.* Proc Natl Acad Sci U S A, 2006. **103**(5): p. 1522-7.
177. Evander, M., et al., *Identification of the alpha6 integrin as a candidate receptor for papillomaviruses.* J Virol, 1997. **71**(3): p. 2449-56.
178. Surviladze, Z., A. Dziduszko, and M.A. Ozbun, *Essential roles for soluble virion-associated heparan sulfonated proteoglycans and growth factors in human papillomavirus infections.* PLoS Pathog, 2012. **8**(2): p. e1002519.
179. Scheffer, K.D., et al., *Tetraspanin CD151 mediates papillomavirus type 16 endocytosis.* J Virol, 2013. **87**(6): p. 3435-46.
180. Fast, L.A., et al., *Inhibition of Tetraspanin Functions Impairs Human Papillomavirus and Cytomegalovirus Infections.* Int J Mol Sci, 2018. **19**(10).
181. Spoden, G., et al., *Clathrin- and caveolin-independent entry of human papillomavirus type 16--involvement of tetraspanin-enriched microdomains (TEMs).* PLoS One, 2008. **3**(10): p. e3313.
182. DiGiuseppe, S., et al., *Cruising the cellular highways: How human papillomavirus travels from the surface to the nucleus.* Virus Res, 2017. **231**: p. 1-9.
183. Day, P.M., et al., *Identification of a role for the trans-Golgi network in human papillomavirus 16 pseudovirus infection.* J Virol, 2013. **87**(7): p. 3862-70.
184. Schelhaas, M., et al., *Entry of human papillomavirus type 16 by actin-dependent, clathrin- and lipid raft-independent endocytosis.* PLoS Pathog, 2012. **8**(4): p. e1002657.
185. Parish, J.L., et al., *ChlR1 is required for loading papillomavirus E2 onto mitotic chromosomes and viral genome maintenance.* Mol Cell, 2006. **24**(6): p. 867-76.
186. Pyeon, D., et al., *Establishment of human papillomavirus infection requires cell cycle progression.* PLoS Pathog, 2009. **5**(2): p. e1000318.
187. Ozbun, M.A., *Human papillomavirus type 31b infection of human keratinocytes and the onset of early transcription.* J Virol, 2002. **76**(22): p. 11291-300.
188. Sanders, C.M. and A. Stenlund, *Recruitment and loading of the E1 initiator protein: an ATP-dependent process catalysed by a transcription factor.* EMBO J, 1998. **17**(23): p. 7044-55.
189. Sanders, C.M. and A. Stenlund, *Transcription factor-dependent loading of the E1 initiator reveals modular assembly of the papillomavirus origin melting complex.* J Biol Chem, 2000. **275**(5): p. 3522-34.
190. Ustav, E., et al., *The bovine papillomavirus origin of replication requires a binding site for the E2 transcriptional activator.* Proc Natl Acad Sci U S A, 1993. **90**(3): p. 898-902.
191. Mohr, I.J., et al., *Targeting the E1 replication protein to the papillomavirus origin of replication by complex formation with the E2 transactivator.* Science, 1990. **250**(4988): p. 1694-9.
192. Sedman, J. and A. Stenlund, *Co-operative interaction between the initiator E1 and the transcriptional activator E2 is required for replicator specific DNA replication of bovine papillomavirus in vivo and in vitro.* EMBO J, 1995. **14**(24): p. 6218-28.
193. Sedman, T., J. Sedman, and A. Stenlund, *Binding of the E1 and E2 proteins to the origin of replication of bovine papillomavirus.* J Virol, 1997. **71**(4): p. 2887-96.
194. Yang, L., et al., *The E1 protein of bovine papilloma virus 1 is an ATP-dependent DNA helicase.* Proc Natl Acad Sci U S A, 1993. **90**(11): p. 5086-90.
195. Beyer-Finkler, E., et al., *Cell differentiation-related gene expression of human papillomavirus 33.* Med Microbiol Immunol, 1990. **179**(4): p. 185-92.

196. Stoler, M.H., et al., *Differentiation-linked human papillomavirus types 6 and 11 transcription in genital condylomata revealed by in situ hybridization with message-specific RNA probes.* Virology, 1989. **172**(1): p. 331-40.
197. Bodily, J. and L.A. Laimins, *Persistence of human papillomavirus infection: keys to malignant progression.* Trends Microbiol, 2011. **19**(1): p. 33-9.
198. del Mar Pena, L.M. and L.A. Laimins, *Differentiation-dependent chromatin rearrangement coincides with activation of human papillomavirus type 31 late gene expression.* J Virol, 2001. **75**(20): p. 10005-13.
199. Peh, W.L., et al., *Life cycle heterogeneity in animal models of human papillomavirus-associated disease.* J Virol, 2002. **76**(20): p. 10401-16.
200. Hoffmann, R., et al., *Different modes of human papillomavirus DNA replication during maintenance.* J Virol, 2006. **80**(9): p. 4431-9.
201. McBride, A.A., *Replication and partitioning of papillomavirus genomes.* Adv Virus Res, 2008. **72**: p. 155-205.
202. McBride, A.A., J.G. Oliveira, and M.G. McPhillips, *Partitioning viral genomes in mitosis: same idea, different targets.* Cell Cycle, 2006. **5**(14): p. 1499-502.
203. Dao, L.D., et al., *Dynamic localization of the human papillomavirus type 11 origin binding protein E2 through mitosis while in association with the spindle apparatus.* J Virol, 2006. **80**(10): p. 4792-800.
204. Van Tine, B.A., et al., *Human papillomavirus (HPV) origin-binding protein associates with mitotic spindles to enable viral DNA partitioning.* Proc Natl Acad Sci U S A, 2004. **101**(12): p. 4030-5.
205. Yu, T., Y.C. Peng, and E.J. Androphy, *Mitotic kinesin-like protein 2 binds and colocalizes with papillomavirus E2 during mitosis.* J Virol, 2007. **81**(4): p. 1736-45.
206. Donaldson, M.M., W. Boner, and I.M. Morgan, *TopBP1 regulates human papillomavirus type 16 E2 interaction with chromatin.* J Virol, 2007. **81**(8): p. 4338-42.
207. You, J., et al., *Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes.* Cell, 2004. **117**(3): p. 349-60.
208. Helfer, C.M., J. Yan, and J. You, *The cellular bromodomain protein Brd4 has multiple functions in E2-mediated papillomavirus transcription activation.* Viruses, 2014. **6**(8): p. 3228-49.
209. Jang, M.K., D. Kwon, and A.A. McBride, *Papillomavirus E2 proteins and the host BRD4 protein associate with transcriptionally active cellular chromatin.* J Virol, 2009. **83**(6): p. 2592-600.
210. McKinney, C.C., et al., *Brd4 Activates Early Viral Transcription upon Human Papillomavirus 18 Infection of Primary Keratinocytes.* mBio, 2016. **7**(6).
211. Sakakibara, N., et al., *Brd4 is displaced from HPV replication factories as they expand and amplify viral DNA.* PLoS Pathog, 2013. **9**(11): p. e1003777.
212. Jang, M.K., K. Shen, and A.A. McBride, *Papillomavirus genomes associate with BRD4 to replicate at fragile sites in the host genome.* PLoS Pathog, 2014. **10**(5): p. e1004117.
213. Gibbs, S. and M. Ponec, *Intrinsic regulation of differentiation markers in human epidermis, hard palate and buccal mucosa.* Arch Oral Biol, 2000. **45**(2): p. 149-58.
214. Baek, J.H., et al., *Common genes responsible for differentiation and senescence of human mucosal and epidermal keratinocytes.* Int J Mol Med, 2003. **12**(3): p. 319-25.
215. Brake, T., et al., *Comparative analysis of cervical cancer in women and in a human papillomavirus-transgenic mouse model: identification of minichromosome maintenance protein 7 as an informative biomarker for human cervical cancer.* Cancer Res, 2003. **63**(23): p. 8173-80.
216. Peh, W.L. and J. Doorbar, *Detection of papillomavirus proteins and DNA in paraffin-embedded tissue sections.* Methods Mol Med, 2005. **119**: p. 49-59.
217. Lu, D.W., S.K. El-Mofty, and H.L. Wang, *Expression of p16, Rb, and p53 proteins in squamous cell carcinomas of the anorectal region harboring human papillomavirus DNA.* Mod Pathol, 2003. **16**(7): p. 692-9.
218. Cheng, S., et al., *Differentiation-dependent up-regulation of the human papillomavirus E7 gene reactivates cellular DNA replication in suprabasal differentiated keratinocytes.* Genes Dev, 1995. **9**(19): p. 2335-49.
219. Demeter, L.M., et al., *Induction of proliferating cell nuclear antigen in differentiated keratinocytes of human papillomavirus-infected lesions.* Hum Pathol, 1994. **25**(4): p. 343-8.
220. Banerjee, N.S., et al., *Human papillomavirus (HPV) E7 induces prolonged G2 following S phase reentry in differentiated human keratinocytes.* J Biol Chem, 2011. **286**(17): p. 15473-82.

221. Porter, L.A. and D.J. Donoghue, *Cyclin B1 and CDK1: nuclear localization and upstream regulators.* Prog Cell Cycle Res, 2003. **5**: p. 335-47.
222. Taylor, W.R., et al., *Mechanisms of G2 arrest in response to overexpression of p53.* Mol Biol Cell, 1999. **10**(11): p. 3607-22.
223. Spriggs, C.C. and L.A. Laimins, *Human Papillomavirus and the DNA Damage Response: Exploiting Host Repair Pathways for Viral Replication.* Viruses, 2017. **9**(8).
224. Eyfjord, J.E. and S.K. Bodvarsdottir, *Genomic instability and cancer: networks involved in response to DNA damage.* Mutat Res, 2005. **592**(1-2): p. 18-28.
225. Sulli, G., R. Di Micco, and F. d'Adda di Fagagna, *Crosstalk between chromatin state and DNA damage response in cellular senescence and cancer.* Nat Rev Cancer, 2012. **12**(10): p. 709-20.
226. Kousholt, A.N., T. Menzel, and C.S. Sorensen, *Pathways for genome integrity in G2 phase of the cell cycle.* Biomolecules, 2012. **2**(4): p. 579-607.
227. Moody, C.A. and L.A. Laimins, *Human papillomaviruses activate the ATM DNA damage pathway for viral genome amplification upon differentiation.* PLoS Pathog, 2009. **5**(10): p. e1000605.
228. Moody, C., *Mechanisms by which HPV Induces a Replication Competent Environment in Differentiating Keratinocytes.* Viruses, 2017. **9**(9).
229. Chow, L.T., et al., *A highly efficient system to produce infectious human papillomavirus: Elucidation of natural virus-host interactions.* Cell Cycle, 2009. **8**(9): p. 1319-23.
230. Davy, C.E., et al., *Human papillomavirus type 16 E1 E4-induced G2 arrest is associated with cytoplasmic retention of active Cdk1/cyclin B1 complexes.* J Virol, 2005. **79**(7): p. 3998-4011.
231. Davy, C.E., et al., *Identification of a G(2) arrest domain in the E1 wedge E4 protein of human papillomavirus type 16.* J Virol, 2002. **76**(19): p. 9806-18.
232. Davy, C. and J. Doorbar, *G2/M cell cycle arrest in the life cycle of viruses.* Virology, 2007. **368**(2): p. 219-26.
233. Egawa, N., et al., *HPV16 and 18 genome amplification show different E4-dependence, with 16E4 enhancing E1 nuclear accumulation and replicative efficiency via its cell cycle arrest and kinase activation functions.* PLoS Pathog, 2017. **13**(3): p. e1006282.
234. Nakahara, T., et al., *Modulation of the cell division cycle by human papillomavirus type 18 E4.* J Virol, 2002. **76**(21): p. 10914-20.
235. Nakahara, T., et al., *Human papillomavirus type 16 E1circumflexE4 contributes to multiple facets of the papillomavirus life cycle.* J Virol, 2005. **79**(20): p. 13150-65.
236. McIntosh, P.B., et al., *E1--E4-mediated keratin phosphorylation and ubiquitylation: a mechanism for keratin depletion in HPV16-infected epithelium.* J Cell Sci, 2010. **123**(Pt 16): p. 2810-22.
237. Knight, G.L., et al., *A cyclin-binding motif in human papillomavirus type 18 (HPV18) E1^E4 is necessary for association with CDK-cyclin complexes and G2/M cell cycle arrest of keratinocytes, but is not required for differentiation-dependent viral genome amplification or L1 capsid protein expression.* Virology, 2011. **412**(1): p. 196-210.
238. Knight, G.L., A.S. Turnell, and S. Roberts, *Role for Wee1 in inhibition of G2-to-M transition through the cooperation of distinct human papillomavirus type 1 E4 proteins.* J Virol, 2006. **80**(15): p. 7416-26.
239. Mehta, K. and L. Laimins, *Human Papillomaviruses Preferentially Recruit DNA Repair Factors to Viral Genomes for Rapid Repair and Amplification.* mBio, 2018. **9**(1).
240. Pickering, M.T. and T.F. Kowalik, *Rb inactivation leads to E2F1-mediated DNA double-strand break accumulation.* Oncogene, 2006. **25**(5): p. 746-55.
241. Duensing, S. and K. Munger, *The human papillomavirus type 16 E6 and E7 oncoproteins independently induce numerical and structural chromosome instability.* Cancer Res, 2002. **62**(23): p. 7075-82.
242. Gillespie, K.A., et al., *Human papillomaviruses recruit cellular DNA repair and homologous recombination factors to viral replication centers.* J Virol, 2012. **86**(17): p. 9520-6.
243. Anacker, D.C., et al., *Productive replication of human papillomavirus 31 requires DNA repair factor Nbs1.* J Virol, 2014. **88**(15): p. 8528-44.
244. Fradet-Turcotte, A., et al., *Nuclear accumulation of the papillomavirus E1 helicase blocks S-phase progression and triggers an ATM-dependent DNA damage response.* J Virol, 2011. **85**(17): p. 8996-9012.
245. Spriggs, C.C., et al., *Expression of HPV-induced DNA Damage Repair Factors Correlates With CIN Progression.* Int J Gynecol Pathol, 2019. **38**(1): p. 1-10.

246. Thierry, F., *Transcriptional regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma.* Virology, 2009. **384**(2): p. 375-9.
247. Grassmann, K., et al., *Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA.* J Virol, 1996. **70**(4): p. 2339-49.
248. Wang, Q., et al., *Functional analysis of the human papillomavirus type 16 E1=E4 protein provides a mechanism for in vivo and in vitro keratin filament reorganization.* J Virol, 2004. **78**(2): p. 821-33.
249. Doorbar, J., et al., *Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network.* Nature, 1991. **352**(6338): p. 824-7.
250. Roberts, S., et al., *Mutational analysis of human papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells.* J Virol, 1994. **68**(10): p. 6432-45.
251. Roberts, S., et al., *Mutational analysis of the human papillomavirus type 16 E1--E4 protein shows that the C terminus is dispensable for keratin cytoskeleton association but is involved in inducing disruption of the keratin filaments.* J Virol, 1997. **71**(5): p. 3554-62.
252. Faraji, F., et al., *Molecular mechanisms of human papillomavirus-related carcinogenesis in head and neck cancer.* Microbes Infect, 2017. **19**(9-10): p. 464-475.
253. Macleod, K.F., et al., *p53-dependent and independent expression of p21 during cell growth, differentiation, and DNA damage.* Genes Dev, 1995. **9**(8): p. 935-44.
254. Martinez-Zapien, D., et al., *Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53.* Nature, 2016. **529**(7587): p. 541-5.
255. Scheffner, M., et al., *The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53.* Cell, 1993. **75**(3): p. 495-505.
256. Dyson, N., *The regulation of E2F by pRB-family proteins.* Genes Dev, 1998. **12**(15): p. 2245-62.
257. Jenkins, D., *Histopathology and cytopathology of cervical cancer.* Dis Markers, 2007. **23**(4): p. 199-212.
258. Isaacson Wechsler, E., et al., *Reconstruction of human papillomavirus type 16-mediated early-stage neoplasia implicates E6/E7 deregulation and the loss of contact inhibition in neoplastic progression.* J Virol, 2012. **86**(11): p. 6358-64.
259. Lee, D., et al., *Using Organotypic Epithelial Tissue Culture to Study the Human Papillomavirus Life Cycle.* Curr Protoc Microbiol, 2016. **41**: p. 14B 8 1-14B 8 19.
260. Martin, C.M. and J.J. O'Leary, *Histology of cervical intraepithelial neoplasia and the role of biomarkers.* Best Pract Res Clin Obstet Gynaecol, 2011. **25**(5): p. 605-15.
261. Huang, S.H., et al., *Atypical clinical behavior of p16-confirmed HPV-related oropharyngeal squamous cell carcinoma treated with radical radiotherapy.* Int J Radiat Oncol Biol Phys, 2012. **82**(1): p. 276-83.
262. Goldenberg, D., et al., *Cystic lymph node metastasis in patients with head and neck cancer: An HPV-associated phenomenon.* Head Neck, 2008. **30**(7): p. 898-903.
263. Leemans, C.R., P.J.F. Snijders, and R.H. Brakenhoff, *The molecular landscape of head and neck cancer.* Nat Rev Cancer, 2018. **18**(5): p. 269-282.
264. Pal, A. and R. Kundu, *Human Papillomavirus E6 and E7: The Cervical Cancer Hallmarks and Targets for Therapy.* Front Microbiol, 2019. **10**: p. 3116.
265. Pett, M. and N. Coleman, *Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis?* J Pathol, 2007. **212**(4): p. 356-67.
266. Mischke, D., et al., *Keratins as molecular markers of epithelial differentiation: differential expression in crypt epithelium of human palatine tonsils.* Ann Otol Rhinol Laryngol, 1991. **100**(5 Pt 1): p. 372-7.
267. Perry, M.E., M.M. Jones, and Y. Mustafa, *Structure of the crypt epithelium in human palatine tonsils.* Acta Otolaryngol Suppl, 1988. **454**: p. 53-9.
268. Mirkovic, J., et al., *Carcinogenic HPV infection in the cervical squamo-columnar junction.* J Pathol, 2015. **236**(3): p. 265-71.
269. Roberts, S., et al., *Modelling human papillomavirus biology in oropharyngeal keratinocytes.* Philos Trans R Soc Lond B Biol Sci, 2019. **374**(1773): p. 20180289.
270. Perry, M.E., *The specialised structure of crypt epithelium in the human palatine tonsil and its functional significance.* J Anat, 1994. **185 (Pt 1)**: p. 111-27.
271. Westra, W.H., *The morphologic profile of HPV-related head and neck squamous carcinoma: implications for diagnosis, prognosis, and clinical management.* Head Neck Pathol, 2012. **6 Suppl 1**: p. S48-54.

272. Chernock, R.D., et al., *Extensive HPV-related carcinoma in situ of the upper aerodigestive tract with 'nonkeratinizing' histologic features.* Head Neck Pathol, 2014. **8**(3): p. 322-8.
273. Rieth, K.K.S., et al., *Prevalence of High-Risk Human Papillomavirus in Tonsil Tissue in Healthy Adults and Colocalization in Biofilm of Tonsillar Crypts.* JAMA Otolaryngol Head Neck Surg, 2018. **144**(3): p. 231-237.
274. Clark, M.A., et al., *Differential cytokeratin and glycoconjugate expression by the surface and crypt epithelia of human palatine tonsils.* Histochem Cell Biol, 2000. **114**(4): p. 311-21.
275. Woods, R.S.R., et al., *Cytokeratin 7 in Oropharyngeal Squamous Cell Carcinoma: A Junctional Biomarker for Human Papillomavirus-Related Tumors.* Cancer Epidemiol Biomarkers Prev, 2017. **26**(5): p. 702-710.
276. Herfs, M., et al., *Deciphering the Multifactorial Susceptibility of Mucosal Junction Cells to HPV Infection and Related Carcinogenesis.* Viruses, 2017. **9**(4).
277. Begum, S., et al., *Tissue distribution of human papillomavirus 16 DNA integration in patients with tonsillar carcinoma.* Clin Cancer Res, 2005. **11**(16): p. 5694-9.
278. Kanduc, D., *Translational regulation of human papillomavirus type 16 E7 mRNA by the peptide SEQIKA, shared by rabbit alpha(1)-globin and human cytokeratin 7.* J Virol, 2002. **76**(14): p. 7040-8.
279. Favia, G., et al., *Possible association between HPV16 E7 protein level and cytokeratin 19.* Int J Cancer, 2004. **111**(5): p. 795-7.
280. Ward, M.J., et al., *Tumour-infiltrating lymphocytes predict for outcome in HPV-positive oropharyngeal cancer.* Br J Cancer, 2014. **110**(2): p. 489-500.
281. Perry, M. and A. Whyte, *Immunology of the tonsils.* Immunol Today, 1998. **19**(9): p. 414-21.
282. Li, N., et al., *Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: Variation by geographical region, histological type and year of publication.* Int J Cancer, 2011. **128**(4): p. 927-35.
283. Clifford, G.M., et al., *Human papillomavirus types in invasive cervical cancer worldwide: a meta-analysis.* Br J Cancer, 2003. **88**(1): p. 63-73.
284. Arbyn, M., et al., *Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis.* Lancet Glob Health, 2020. **8**(2): p. e191-e203.
285. Bulk, S., et al., *Preferential risk of HPV16 for squamous cell carcinoma and of HPV18 for adenocarcinoma of the cervix compared to women with normal cytology in The Netherlands.* Br J Cancer, 2006. **94**(1): p. 171-5.
286. Gotz, C., et al., *Detection of HPV infection in head and neck cancers: Promise and pitfalls in the last ten years: A meta-analysis.* Mol Clin Oncol, 2019. **10**(1): p. 17-28.
287. Abogunrin, S., et al., *Prevalence of human papillomavirus in head and neck cancers in European populations: a meta-analysis.* BMC Cancer, 2014. **14**: p. 968.
288. Zanier, K., et al., *Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins.* Science, 2013. **339**(6120): p. 694-8.
289. Nomine, Y., et al., *Domain substructure of HPV E6 oncoprotein: biophysical characterization of the E6 C-terminal DNA-binding domain.* Biochemistry, 2003. **42**(17): p. 4909-17.
290. White, E.A., et al., *Comprehensive analysis of host cellular interactions with human papillomavirus E6 proteins identifies new E6 binding partners and reflects viral diversity.* J Virol, 2012. **86**(24): p. 13174-86.
291. Fu, L., et al., *Degradation of p53 by human Alphapapillomavirus E6 proteins shows a stronger correlation with phylogeny than oncogenicity.* PLoS One, 2010. **5**(9).
292. Scheffner, M., et al., *The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53.* Cell, 1990. **63**(6): p. 1129-36.
293. Werness, B.A., A.J. Levine, and P.M. Howley, *Association of human papillomavirus types 16 and 18 E6 proteins with p53.* Science, 1990. **248**(4951): p. 76-9.
294. Thomas, M.C. and C.M. Chiang, *E6 oncoprotein represses p53-dependent gene activation via inhibition of protein acetylation independently of inducing p53 degradation.* Mol Cell, 2005. **17**(2): p. 251-64.
295. Patel, D., et al., *The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300.* EMBO J, 1999. **18**(18): p. 5061-72.
296. Zimmermann, H., et al., *The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300.* J Virol, 1999. **73**(8): p. 6209-19.
297. Mukherjee, S.P., et al., *Analysis of the RelA:CBP/p300 interaction reveals its involvement in NF-kappaB-driven transcription.* PLoS Biol, 2013. **11**(9): p. e1001647.

298. Peter, M., et al., *MYC activation associated with the integration of HPV DNA at the MYC locus in genital tumors.* Oncogene, 2006. **25**(44): p. 5985-93.
299. Wang, Y.W., et al., *HPV-18 E7 conjugates to c-Myc and mediates its transcriptional activity.* Int J Biochem Cell Biol, 2007. **39**(2): p. 402-12.
300. Abba, M.C., et al., *The c-myc activation in cervical carcinomas and HPV 16 infections.* Mutat Res, 2004. **557**(2): p. 151-8.
301. Veldman, T., et al., *Human papillomavirus E6 and Myc proteins associate in vivo and bind to and cooperatively activate the telomerase reverse transcriptase promoter.* Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8211-6.
302. Gewin, L., et al., *Identification of a novel telomerase repressor that interacts with the human papillomavirus type-16 E6/E6-AP complex.* Genes Dev, 2004. **18**(18): p. 2269-82.
303. Gewin, L. and D.A. Galloway, *E box-dependent activation of telomerase by human papillomavirus type 16 E6 does not require induction of c-myc.* J Virol, 2001. **75**(15): p. 7198-201.
304. Veldman, T., et al., *Transcriptional activation of the telomerase hTERT gene by human papillomavirus type 16 E6 oncoprotein.* J Virol, 2001. **75**(9): p. 4467-72.
305. Liu, X., et al., *The E6AP ubiquitin ligase is required for transactivation of the hTERT promoter by the human papillomavirus E6 oncoprotein.* J Biol Chem, 2005. **280**(11): p. 10807-16.
306. Kiyono, T., et al., *Binding of high-risk human papillomavirus E6 oncoproteins to the human homologue of the Drosophila discs large tumor suppressor protein.* Proc Natl Acad Sci U S A, 1997. **94**(21): p. 11612-6.
307. Lee, S.S., R.S. Weiss, and R.T. Javier, *Binding of human virus oncoproteins to hDlg/SAP97, a mammalian homolog of the Drosophila discs large tumor suppressor protein.* Proc Natl Acad Sci U S A, 1997. **94**(13): p. 6670-5.
308. Kuhne, C., et al., *Differential regulation of human papillomavirus E6 by protein kinase A: conditional degradation of human discs large protein by oncogenic E6.* Oncogene, 2000. **19**(51): p. 5884-91.
309. Massimi, P., et al., *Phosphorylation of the discs large tumour suppressor protein controls its membrane localisation and enhances its susceptibility to HPV E6-induced degradation.* Oncogene, 2006. **25**(31): p. 4276-85.
310. Nicolaides, L., et al., *Stabilization of HPV16 E6 protein by PDZ proteins, and potential implications for genome maintenance.* Virology, 2011. **414**(2): p. 137-45.
311. Ganti, K., et al., *The Human Papillomavirus E6 PDZ Binding Motif: From Life Cycle to Malignancy.* Viruses, 2015. **7**(7): p. 3530-51.
312. James, C.D. and S. Roberts, *Viral Interactions with PDZ Domain-Containing Proteins-An Oncogenic Trait?* Pathogens, 2016. **5**(1).
313. Zhang, Y., et al., *Structures of a human papillomavirus (HPV) E6 polypeptide bound to MAGUK proteins: mechanisms of targeting tumor suppressors by a high-risk HPV oncoprotein.* J Virol, 2007. **81**(7): p. 3618-26.
314. Thomas, M., et al., *Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation.* Oncogene, 2002. **21**(33): p. 5088-96.
315. Glaunsinger, B.A., et al., *Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins.* Oncogene, 2000. **19**(46): p. 5270-80.
316. Thomas, M., et al., *The hScrib/Dlg apico-basal control complex is differentially targeted by HPV-16 and HPV-18 E6 proteins.* Oncogene, 2005. **24**(41): p. 6222-30.
317. Takizawa, S., et al., *Human scribble, a novel tumor suppressor identified as a target of high-risk HPV E6 for ubiquitin-mediated degradation, interacts with adenomatous polyposis coli.* Genes Cells, 2006. **11**(4): p. 453-64.
318. Nguyen, M.L., et al., *The PDZ ligand domain of the human papillomavirus type 16 E6 protein is required for E6's induction of epithelial hyperplasia in vivo.* J Virol, 2003. **77**(12): p. 6957-64.
319. Watson, R.A., et al., *Activity of the human papillomavirus E6 PDZ-binding motif correlates with an enhanced morphological transformation of immortalized human keratinocytes.* J Cell Sci, 2003. **116**(Pt 24): p. 4925-34.
320. Vincentelli, R., et al., *Quantifying domain-ligand affinities and specificities by high-throughput holdup assay.* Nat Methods, 2015. **12**(8): p. 787-93.

321. Li, S., et al., *The human papilloma virus (HPV)-18 E6 oncoprotein physically associates with Tyk2 and impairs Jak-STAT activation by interferon-alpha.* Oncogene, 1999. **18**(42): p. 5727-37.
322. Ronco, L.V., et al., *Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity.* Genes Dev, 1998. **12**(13): p. 2061-72.
323. Albertini, S., et al., *HPV18 Persistence Impairs Basal and DNA Ligand-Mediated IFN-beta and IFN-lambda1 Production through Transcriptional Repression of Multiple Downstream Effectors of Pattern Recognition Receptor Signaling.* J Immunol, 2018. **200**(6): p. 2076-2089.
324. Hasan, U.A., et al., *TLR9 expression and function is abolished by the cervical cancer-associated human papillomavirus type 16.* J Immunol, 2007. **178**(5): p. 3186-97.
325. Phelps, W.C., et al., *The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A.* Cell, 1988. **53**(4): p. 539-47.
326. Barrow-Laing, L., W. Chen, and A. Roman, *Low- and high-risk human papillomavirus E7 proteins regulate p130 differently.* Virology, 2010. **400**(2): p. 233-9.
327. Roman, A., *The human papillomavirus E7 protein shines a spotlight on the pRB family member, p130.* Cell Cycle, 2006. **5**(6): p. 567-8.
328. Dyson, N., et al., *Homologous sequences in adenovirus E1A and human papillomavirus E7 proteins mediate interaction with the same set of cellular proteins.* J Virol, 1992. **66**(12): p. 6893-902.
329. Klingelhutz, A.J. and A. Roman, *Cellular transformation by human papillomaviruses: lessons learned by comparing high- and low-risk viruses.* Virology, 2012. **424**(2): p. 77-98.
330. Felsani, A., A.M. Mileo, and M.G. Paggi, *Retinoblastoma family proteins as key targets of the small DNA virus oncoproteins.* Oncogene, 2006. **25**(38): p. 5277-85.
331. Jones, D.L., R.M. Alani, and K. Munger, *The human papillomavirus E7 oncoprotein can uncouple cellular differentiation and proliferation in human keratinocytes by abrogating p21Cip1-mediated inhibition of cdk2.* Genes Dev, 1997. **11**(16): p. 2101-11.
332. Demers, G.W., et al., *Abrogation of growth arrest signals by human papillomavirus type 16 E7 is mediated by sequences required for transformation.* J Virol, 1996. **70**(10): p. 6862-9.
333. McLaughlin-Drubin, M.E., D. Park, and K. Munger, *Tumor suppressor p16INK4A is necessary for survival of cervical carcinoma cell lines.* Proc Natl Acad Sci U S A, 2013. **110**(40): p. 16175-80.
334. Missero, C., et al., *Involvement of the cell-cycle inhibitor Cip1/WAF1 and the E1A-associated p300 protein in terminal differentiation.* Proc Natl Acad Sci U S A, 1995. **92**(12): p. 5451-5.
335. Deshpande, A., P. Sicinski, and P.W. Hinds, *Cyclins and cdks in development and cancer: a perspective.* Oncogene, 2005. **24**(17): p. 2909-15.
336. Funk, J.O., et al., *Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein.* Genes Dev, 1997. **11**(16): p. 2090-100.
337. Nguyen, C.L. and K. Munger, *Direct association of the HPV16 E7 oncoprotein with cyclin A/CDK2 and cyclin E/CDK2 complexes.* Virology, 2008. **380**(1): p. 21-5.
338. Katich, S.C., K. Zerfass-Thome, and I. Hoffmann, *Regulation of the Cdc25A gene by the human papillomavirus Type 16 E7 oncogene.* Oncogene, 2001. **20**(5): p. 543-50.
339. Nguyen, D.X., T.F. Westbrook, and D.J. McCance, *Human papillomavirus type 16 E7 maintains elevated levels of the cdc25A tyrosine phosphatase during deregulation of cell cycle arrest.* J Virol, 2002. **76**(2): p. 619-32.
340. White, E.A., et al., *Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses.* Proc Natl Acad Sci U S A, 2012. **109**(5): p. E260-7.
341. Huh, K., et al., *Human papillomavirus type 16 E7 oncoprotein associates with the cullin 2 ubiquitin ligase complex, which contributes to degradation of the retinoblastoma tumor suppressor.* J Virol, 2007. **81**(18): p. 9737-47.
342. Xu, J., et al., *CUL2 overexpression driven by CUL2/E2F1/miR-424 regulatory loop promotes HPV16 E7 induced cervical carcinogenesis.* Oncotarget, 2016. **7**(21): p. 31520-33.
343. Zhou, C., Z.K. Tuong, and I.H. Frazer, *Papillomavirus Immune Evasion Strategies Target the Infected Cell and the Local Immune System.* Front Oncol, 2019. **9**: p. 682.
344. Paz, S., et al., *Induction of IRF-3 and IRF-7 phosphorylation following activation of the RIG-I pathway.* Cell Mol Biol (Noisy-le-grand), 2006. **52**(1): p. 17-28.

345. Andersen, J.M., D. Al-Khairy, and R.R. Ingalls, *Innate immunity at the mucosal surface: role of toll-like receptor 3 and toll-like receptor 9 in cervical epithelial cell responses to microbial pathogens.* Biol Reprod, 2006. **74**(5): p. 824-31.
346. Hemmi, H., et al., *The roles of two IkappaB kinase-related kinases in lipopolysaccharide and double stranded RNA signaling and viral infection.* J Exp Med, 2004. **199**(12): p. 1641-50.
347. Sankar, S., et al., *IKK-i signals through IRF3 and NFkappaB to mediate the production of inflammatory cytokines.* Cell Signal, 2006. **18**(7): p. 982-93.
348. Chiang, H.S. and H.M. Liu, *The Molecular Basis of Viral Inhibition of IRF- and STAT-Dependent Immune Responses.* Front Immunol, 2018. **9**: p. 3086.
349. Platanias, L.C., *Mechanisms of type-I- and type-II-interferon-mediated signalling.* Nat Rev Immunol, 2005. **5**(5): p. 375-86.
350. Thompson, M.R., et al., *Pattern recognition receptors and the innate immune response to viral infection.* Viruses, 2011. **3**(6): p. 920-40.
351. Ivashkiv, L.B. and L.T. Donlin, *Regulation of type I interferon responses.* Nat Rev Immunol, 2014. **14**(1): p. 36-49.
352. Albanesi, C., et al., *The Interplay Between Keratinocytes and Immune Cells in the Pathogenesis of Psoriasis.* Front Immunol, 2018. **9**: p. 1549.
353. Pivarcsi, A., L. Kemeny, and A. Dobozy, *Innate immune functions of the keratinocytes. A review.* Acta Microbiol Immunol Hung, 2004. **51**(3): p. 303-10.
354. van Kempen, M.J., G.T. Rijkers, and P.B. Van Cauwenberge, *The immune response in adenoids and tonsils.* Int Arch Allergy Immunol, 2000. **122**(1): p. 8-19.
355. Chung, W.O. and B.A. Dale, *Innate immune response of oral and foreskin keratinocytes: utilization of different signaling pathways by various bacterial species.* Infect Immun, 2004. **72**(1): p. 352-8.
356. Park, J.H., et al., *Phosphorylation of IkappaBalpha at serine 32 by T-lymphokine-activated killer cell-originated protein kinase is essential for chemoresistance against doxorubicin in cervical cancer cells.* J Biol Chem, 2013. **288**(5): p. 3585-93.
357. Cason, J., et al., *Identification of immunogenic regions of the major coat protein of human papillomavirus type 16 that contain type-restricted epitopes.* J Gen Virol, 1989. **70 (Pt 11)**: p. 2973-87.
358. Doorbar, J., *The papillomavirus life cycle.* J Clin Virol, 2005. **32 Suppl 1**: p. S7-15.
359. Woodby, B., M. Scott, and J. Bodily, *The Interaction Between Human Papillomaviruses and the Stromal Microenvironment.* Prog Mol Biol Transl Sci, 2016. **144**: p. 169-238.
360. Frazer, I.H., *Interaction of human papillomaviruses with the host immune system: a well evolved relationship.* Virology, 2009. **384**(2): p. 410-4.
361. Karim, R., et al., *Human papillomavirus deregulates the response of a cellular network comprising of chemotactic and proinflammatory genes.* PLoS One, 2011. **6**(3): p. e17848.
362. Nees, M., et al., *Papillomavirus type 16 oncogenes downregulate expression of interferon-responsive genes and upregulate proliferation-associated and NF-kappaB-responsive genes in cervical keratinocytes.* J Virol, 2001. **75**(9): p. 4283-96.
363. Hong, S. and L.A. Laimins, *Manipulation of the innate immune response by human papillomaviruses.* Virus Res, 2017. **231**: p. 34-40.
364. Georgopoulos, N.T., J.L. Proffitt, and G.E. Blair, *Transcriptional regulation of the major histocompatibility complex (MHC) class I heavy chain, TAP1 and LMP2 genes by the human papillomavirus (HPV) type 6b, 16 and 18 E7 oncoproteins.* Oncogene, 2000. **19**(42): p. 4930-5.
365. Miura, S., et al., *CD1d, a sentinel molecule bridging innate and adaptive immunity, is downregulated by the human papillomavirus (HPV) E5 protein: a possible mechanism for immune evasion by HPV.* J Virol, 2010. **84**(22): p. 11614-23.
366. Kim, H.S., et al., *Association Between PD-L1 and HPV Status and the Prognostic Value of PD-L1 in Oropharyngeal Squamous Cell Carcinoma.* Cancer Res Treat, 2016. **48**(2): p. 527-36.
367. Filippova, M., et al., *The human papillomavirus 16 E6 protein binds to tumor necrosis factor (TNF) R1 and protects cells from TNF-induced apoptosis.* J Biol Chem, 2002. **277**(24): p. 21730-9.
368. Clerici, M., et al., *Cytokine production patterns in cervical intraepithelial neoplasia: association with human papillomavirus infection.* J Natl Cancer Inst, 1997. **89**(3): p. 245-50.
369. Burgers, W.A., et al., *Viral oncoproteins target the DNA methyltransferases.* Oncogene, 2007. **26**(11): p. 1650-5.

370. Cicchini, L., et al., *Suppression of Antitumor Immune Responses by Human Papillomavirus through Epigenetic Downregulation of CXCL14.* mBio, 2016. **7**(3).
371. Hasan, U.A., et al., *The human papillomavirus type 16 E7 oncoprotein induces a transcriptional repressor complex on the Toll-like receptor 9 promoter.* J Exp Med, 2013. **210**(7): p. 1369-87.
372. Lau, L., et al., *DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway.* Science, 2015. **350**(6260): p. 568-71.
373. Decout, A., et al., *The cGAS-STING pathway as a therapeutic target in inflammatory diseases.* Nat Rev Immunol, 2021. **21**(9): p. 548-569.
374. Honda, K., et al., *Role of a transductional-transcriptional processor complex involving MyD88 and IRF-7 in Toll-like receptor signaling.* Proc Natl Acad Sci U S A, 2004. **101**(43): p. 15416-21.
375. Oganesyan, G., et al., *Critical role of TRAF3 in the Toll-like receptor-dependent and -independent antiviral response.* Nature, 2006. **439**(7073): p. 208-11.
376. Karim, R., et al., *Human papillomavirus (HPV) upregulates the cellular deubiquitinase UCHL1 to suppress the keratinocyte's innate immune response.* PLoS Pathog, 2013. **9**(5): p. e1003384.
377. Niebler, M., et al., *Post-translational control of IL-1beta via the human papillomavirus type 16 E6 oncoprotein: a novel mechanism of innate immune escape mediated by the E3-ubiquitin ligase E6-AP and p53.* PLoS Pathog, 2013. **9**(8): p. e1003536.
378. Chen, L., et al., *Inflammatory responses and inflammation-associated diseases in organs.* Oncotarget, 2018. **9**(6): p. 7204-7218.
379. Park, J.S., et al., *Inactivation of interferon regulatory factor-1 tumor suppressor protein by HPV E7 oncoprotein. Implication for the E7-mediated immune evasion mechanism in cervical carcinogenesis.* J Biol Chem, 2000. **275**(10): p. 6764-9.
380. Antonsson, A., et al., *The human papillomavirus type 16 E7 protein binds human interferon regulatory factor-9 via a novel PEST domain required for transformation.* J Interferon Cytokine Res, 2006. **26**(7): p. 455-61.
381. Tummers, B., et al., *The interferon-related developmental regulator 1 is used by human papillomavirus to suppress NFkappaB activation.* Nat Commun, 2015. **6**: p. 6537.
382. Huang, S.M. and D.J. McCance, *Down regulation of the interleukin-8 promoter by human papillomavirus type 16 E6 and E7 through effects on CREB binding protein/p300 and P/CAF.* J Virol, 2002. **76**(17): p. 8710-21.
383. Chang, Y.E. and L.A. Laimins, *Microarray analysis identifies interferon-inducible genes and Stat-1 as major transcriptional targets of human papillomavirus type 31.* J Virol, 2000. **74**(9): p. 4174-82.
384. Neefjes, J., et al., *Towards a systems understanding of MHC class I and MHC class II antigen presentation.* Nat Rev Immunol, 2011. **11**(12): p. 823-36.
385. Theobald, V.A., et al., *"Neutral allografts"--lack of allogeneic stimulation by cultured human cells expressing MHC class I and class II antigens.* Transplantation, 1993. **55**(1): p. 128-33.
386. Black, A.P., et al., *Human keratinocyte induction of rapid effector function in antigen-specific memory CD4+ and CD8+ T cells.* Eur J Immunol, 2007. **37**(6): p. 1485-93.
387. Li, H., et al., *Repression of MHC class I transcription by HPV16E7 through interaction with a putative RXRbeta motif and NF-kappaB cytoplasmic sequestration.* Biochem Biophys Res Commun, 2009. **388**(2): p. 383-8.
388. Gruener, M., et al., *The E5 protein of the human papillomavirus type 16 down-regulates HLA-I surface expression in calnexin-expressing but not in calnexin-deficient cells.* Virol J, 2007. **4**: p. 116.
389. Ashrafi, G.H., et al., *E5 protein of human papillomavirus type 16 selectively downregulates surface HLA class I.* Int J Cancer, 2005. **113**(2): p. 276-83.
390. Campo, M.S., et al., *HPV-16 E5 down-regulates expression of surface HLA class I and reduces recognition by CD8 T cells.* Virology, 2010. **407**(1): p. 137-42.
391. Hu, Z., et al., *Genome-wide profiling of HPV integration in cervical cancer identifies clustered genomic hot spots and a potential microhomology-mediated integration mechanism.* Nat Genet, 2015. **47**(2): p. 158-63.
392. Jeon, S., B.L. Allen-Hoffmann, and P.F. Lambert, *Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells.* J Virol, 1995. **69**(5): p. 2989-97.
393. Rampias, T., et al., *E6 and e7 gene silencing and transformed phenotype of human papillomavirus 16-positive oropharyngeal cancer cells.* J Natl Cancer Inst, 2009. **101**(6): p. 412-23.

394. Pett, M.R., et al., *Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes.* Cancer Res, 2004. **64**(4): p. 1359-68.
395. Cancer Genome Atlas Research, N., et al., *Integrated genomic and molecular characterization of cervical cancer.* Nature, 2017. **543**(7645): p. 378-384.
396. Olthof, N.C., et al., *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, 2014. **9**(2): p. e88718.
397. Vojtechova, Z., et al., *Analysis of the integration of human papillomaviruses in head and neck tumours in relation to patients' prognosis.* Int J Cancer, 2016. **138**(2): p. 386-95.
398. Koneva, L.A., et al., *HPV Integration in HNSCC Correlates with Survival Outcomes, Immune Response Signatures, and Candidate Drivers.* Mol Cancer Res, 2018. **16**(1): p. 90-102.
399. Akagi, K., et al., *Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability.* Genome Res, 2014. **24**(2): p. 185-99.
400. Parfenov, M., et al., *Characterization of HPV and host genome interactions in primary head and neck cancers.* Proc Natl Acad Sci U S A, 2014. **111**(43): p. 15544-9.
401. Lace, M.J., et al., *Human papillomavirus type 16 (HPV-16) genomes integrated in head and neck cancers and in HPV-16-immortalized human keratinocyte clones express chimeric virus-cell mRNAs similar to those found in cervical cancers.* J Virol, 2011. **85**(4): p. 1645-54.
402. Bodelon, C., et al., *Genomic characterization of viral integration sites in HPV-related cancers.* Int J Cancer, 2016. **139**(9): p. 2001-11.
403. Thorland, E.C., et al., *Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites.* Cancer Res, 2000. **60**(21): p. 5916-21.
404. Christiansen, I.K., et al., *Transcriptionally active regions are the preferred targets for chromosomal HPV integration in cervical carcinogenesis.* PLoS One, 2015. **10**(3): p. e0119566.
405. Walline, H.M., et al., *Integration of high-risk human papillomavirus into cellular cancer-related genes in head and neck cancer cell lines.* Head Neck, 2017. **39**(5): p. 840-852.
406. Leung, T.W., et al., *HPV 16 E2 binding sites 1 and 2 become more methylated than E2 binding site 4 during cervical carcinogenesis.* J Med Virol, 2015. **87**(6): p. 1022-33.
407. Dooley, K.E., A. Warburton, and A.A. McBride, *Tandemly Integrated HPV16 Can Form a Brd4-Dependent Super-Enhancer-Like Element That Drives Transcription of Viral Oncogenes.* mBio, 2016. **7**(5).
408. Groves, I.J., et al., *HPV16 oncogene expression levels during early cervical carcinogenesis are determined by the balance of epigenetic chromatin modifications at the integrated virus genome.* Oncogene, 2016. **35**(36): p. 4773-86.
409. McBride, A.A. and A. Warburton, *The role of integration in oncogenic progression of HPV-associated cancers.* PLoS Pathog, 2017. **13**(4): p. e1006211.
410. Cancer Genome Atlas, N., *Comprehensive genomic characterization of head and neck squamous cell carcinomas.* Nature, 2015. **517**(7536): p. 576-82.
411. Pan, C., N. Issaeva, and W.G. Yarbrough, *HPV-driven oropharyngeal cancer: current knowledge of molecular biology and mechanisms of carcinogenesis.* Cancers Head Neck, 2018. **3**: p. 12.
412. Mirghani, H., et al., *Increased radiosensitivity of HPV-positive head and neck cancers: Molecular basis and therapeutic perspectives.* Cancer Treat Rev, 2015. **41**(10): p. 844-52.
413. Poeta, M.L., et al., *TP53 mutations and survival in squamous-cell carcinoma of the head and neck.* N Engl J Med, 2007. **357**(25): p. 2552-61.
414. Laplante, M. and D.M. Sabatini, *mTOR signaling in growth control and disease.* Cell, 2012. **149**(2): p. 274-93.
415. Lechner, M., et al., *Identification and functional validation of HPV-mediated hypermethylation in head and neck squamous cell carcinoma.* Genome Med, 2013. **5**(2): p. 15.
416. Lui, V.W., et al., *Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers.* Cancer Discov, 2013. **3**(7): p. 761-9.
417. Gillison, M.L., et al., *Human papillomavirus and the landscape of secondary genetic alterations in oral cancers.* Genome Res, 2019. **29**(1): p. 1-17.
418. Henderson, S., et al., *APOBEC-mediated cytosine deamination links PIK3CA helical domain mutations to human papillomavirus-driven tumor development.* Cell Rep, 2014. **7**(6): p. 1833-41.

419. Sewell, A., et al., *Reverse-phase protein array profiling of oropharyngeal cancer and significance of PIK3CA mutations in HPV-associated head and neck cancer.* Clin Cancer Res, 2014. **20**(9): p. 2300-11.
420. Leemans, C.R., B.J. Braakhuis, and R.H. Brakenhoff, *The molecular biology of head and neck cancer.* Nat Rev Cancer, 2011. **11**(1): p. 9-22.
421. Hajek, M., et al., *TRAF3/CYLD mutations identify a distinct subset of human papillomavirus-associated head and neck squamous cell carcinoma.* Cancer, 2017. **123**(10): p. 1778-1790.
422. Wilson, R., et al., *The full-length E1E4 protein of human papillomavirus type 18 modulates differentiation-dependent viral DNA amplification and late gene expression.* Virology, 2007. **362**(2): p. 453-60.
423. Watson, R.A., et al., *Changes in expression of the human homologue of the Drosophila discs large tumour suppressor protein in high-grade premalignant cervical neoplasias.* Carcinogenesis, 2002. **23**(11): p. 1791-6.
424. Schwenzer, H., et al., *LARP1 isoform expression in human cancer cell lines.* RNA Biol, 2021. **18**(2): p. 237-247.
425. Frattini, M.G., H.B. Lim, and L.A. Laimins, *In vitro synthesis of oncogenic human papillomaviruses requires episomal genomes for differentiation-dependent late expression.* Proc Natl Acad Sci U S A, 1996. **93**(7): p. 3062-7.
426. Dollard, S.C., et al., *Production of human papillomavirus and modulation of the infectious program in epithelial raft cultures. OFF.* Genes Dev, 1992. **6**(7): p. 1131-42.
427. Meyers, C., et al., *Biosynthesis of human papillomavirus from a continuous cell line upon epithelial differentiation.* Science, 1992. **257**(5072): p. 971-3.
428. Bienkowska-Haba, M., et al., *A new cell culture model to genetically dissect the complete human papillomavirus life cycle.* PLoS Pathog, 2018. **14**(3): p. e1006846.
429. Alazawi, W., et al., *Changes in cervical keratinocyte gene expression associated with integration of human papillomavirus 16.* Cancer Res, 2002. **62**(23): p. 6959-65.
430. Delury, C.P., et al., *The role of protein kinase A regulation of the E6 PDZ-binding domain during the differentiation-dependent life cycle of human papillomavirus type 18.* J Virol, 2013. **87**(17): p. 9463-72.
431. Marsh, E.K., et al., *Mitotic control of human papillomavirus genome-containing cells is regulated by the function of the PDZ-binding motif of the E6 oncoprotein.* Oncotarget, 2017. **8**(12): p. 19491-19506.
432. Ruesch, M.N., F. Stubenrauch, and L.A. Laimins, *Activation of papillomavirus late gene transcription and genome amplification upon differentiation in semisolid medium is coincident with expression of involucrin and transglutaminase but not keratin-10.* J Virol, 1998. **72**(6): p. 5016-24.
433. Anderson, D.J., J. Marathe, and J. Pudney, *The structure of the human vaginal stratum corneum and its role in immune defense.* Am J Reprod Immunol, 2014. **71**(6): p. 618-23.
434. Hurford, R.K., Jr., et al., *pRB and p107/p130 are required for the regulated expression of different sets of E2F responsive genes.* Genes Dev, 1997. **11**(11): p. 1447-63.
435. Dyson, N., et al., *The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product.* Science, 1989. **243**(4893): p. 934-7.
436. Macaluso, M., M. Montanari, and A. Giordano, *Rb family proteins as modulators of gene expression and new aspects regarding the interaction with chromatin remodeling enzymes.* Oncogene, 2006. **25**(38): p. 5263-7.
437. Zhang, B., W. Chen, and A. Roman, *The E7 proteins of low- and high-risk human papillomaviruses share the ability to target the pRB family member p130 for degradation.* Proc Natl Acad Sci U S A, 2006. **103**(2): p. 437-42.
438. Wilson, R., F. Fehrmann, and L.A. Laimins, *Role of the E1--E4 protein in the differentiation-dependent life cycle of human papillomavirus type 31.* J Virol, 2005. **79**(11): p. 6732-40.
439. Doorbar, J., *The E4 protein; structure, function and patterns of expression.* Virology, 2013. **445**(1-2): p. 80-98.
440. Doorbar, J., et al., *Analysis of HPV-1 E4 gene expression using epitope-defined antibodies.* EMBO J, 1988. **7**(3): p. 825-33.
441. Hummel, M., H.B. Lim, and L.A. Laimins, *Human papillomavirus type 31b late gene expression is regulated through protein kinase C-mediated changes in RNA processing.* J Virol, 1995. **69**(6): p. 3381-8.
442. Lei, M., *The MCM complex: its role in DNA replication and implications for cancer therapy.* Curr Cancer Drug Targets, 2005. **5**(5): p. 365-80.

443. Li, W., et al., *Human papillomavirus positivity predicts favourable outcome for squamous carcinoma of the tonsil.* Int J Cancer, 2003. **106**(4): p. 553-8.
444. Moody, C.A., et al., *Human papillomaviruses activate caspases upon epithelial differentiation to induce viral genome amplification.* Proc Natl Acad Sci U S A, 2007. **104**(49): p. 19541-6.
445. Reinson, T., et al., *Engagement of the ATR-dependent DNA damage response at the human papillomavirus 18 replication centers during the initial amplification.* J Virol, 2013. **87**(2): p. 951-64.
446. Moll, R., et al., *The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells.* Cell, 1982. **31**(1): p. 11-24.
447. Fuchs, E. and H. Green, *Changes in keratin gene expression during terminal differentiation of the keratinocyte.* Cell, 1980. **19**(4): p. 1033-42.
448. Albers, K.M., *Keratin biochemistry.* Clin Dermatol, 1996. **14**(4): p. 309-20.
449. Park, N.H., et al., *Immortalization of normal human oral keratinocytes with type 16 human papillomavirus.* Carcinogenesis, 1991. **12**(9): p. 1627-31.
450. Lace, M.J., et al., *Human papillomavirus (HPV) type 18 induces extended growth in primary human cervical, tonsillar, or foreskin keratinocytes more effectively than other high-risk mucosal HPVs.* J Virol, 2009. **83**(22): p. 11784-94.
451. Hafner, N., et al., *Integration of the HPV16 genome does not invariably result in high levels of viral oncogene transcripts.* Oncogene, 2008. **27**(11): p. 1610-7.
452. Vinokurova, S., et al., *Type-dependent integration frequency of human papillomavirus genomes in cervical lesions.* Cancer Res, 2008. **68**(1): p. 307-13.
453. Ferber, M.J., et al., *Preferential integration of human papillomavirus type 18 near the c-myc locus in cervical carcinoma.* Oncogene, 2003. **22**(46): p. 7233-42.
454. Zhang, R., et al., *Dysregulation of host cellular genes targeted by human papillomavirus (HPV) integration contributes to HPV-related cervical carcinogenesis.* Int J Cancer, 2016. **138**(5): p. 1163-74.
455. Ojesina, A.I., et al., *Landscape of genomic alterations in cervical carcinomas.* Nature, 2014. **506**(7488): p. 371-5.
456. Herdman, M.T., et al., *Interferon-beta treatment of cervical keratinocytes naturally infected with human papillomavirus 16 episomes promotes rapid reduction in episome numbers and emergence of latent integrants.* Carcinogenesis, 2006. **27**(11): p. 2341-53.
457. Lace, M.J., et al., *Interferon-beta treatment increases human papillomavirus early gene transcription and viral plasmid genome replication by activating interferon regulatory factor (IRF)-1.* Carcinogenesis, 2009. **30**(8): p. 1336-44.
458. Pett, M.R., et al., *Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response.* Proc Natl Acad Sci U S A, 2006. **103**(10): p. 3822-7.
459. Wang, Z., et al., *APOBEC3 deaminases induce hypermutation in human papillomavirus 16 DNA upon beta interferon stimulation.* J Virol, 2014. **88**(2): p. 1308-17.
460. Kondo, S., et al., *APOBEC3A associates with human papillomavirus genome integration in oropharyngeal cancers.* Oncogene, 2017. **36**(12): p. 1687-1697.
461. Zapatka, M., et al., *The landscape of viral associations in human cancers.* Nat Genet, 2020. **52**(3): p. 320-330.
462. Flores, E.R., et al., *The human papillomavirus type 16 E7 oncogene is required for the productive stage of the viral life cycle.* J Virol, 2000. **74**(14): p. 6622-31.
463. Sichero, L., J.S. Sobrinho, and L.L. Villa, *Identification of novel cellular transcription factors that regulate early promoters of human papillomavirus types 18 and 16.* J Infect Dis, 2012. **206**(6): p. 867-74.
464. Wooldridge, T.R. and L.A. Laimins, *Regulation of human papillomavirus type 31 gene expression during the differentiation-dependent life cycle through histone modifications and transcription factor binding.* Virology, 2008. **374**(2): p. 371-80.
465. Gunasekharan, V., G. Hache, and L. Laimins, *Differentiation-dependent changes in levels of C/EBPbeta repressors and activators regulate human papillomavirus type 31 late gene expression.* J Virol, 2012. **86**(9): p. 5393-8.
466. Songock, W.K., M.L. Scott, and J.M. Bodily, *Regulation of the human papillomavirus type 16 late promoter by transcriptional elongation.* Virology, 2017. **507**: p. 179-191.

467. Paavonen, J., et al., *Efficacy of a prophylactic adjuvanted bivalent L1 virus-like-particle vaccine against infection with human papillomavirus types 16 and 18 in young women: an interim analysis of a phase III double-blind, randomised controlled trial.* Lancet, 2007. **369**(9580): p. 2161-2170.
468. Paavonen, J., et al., *Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women.* Lancet, 2009. **374**(9686): p. 301-14.
469. Quint, W., et al., *One virus, one lesion--individual components of CIN lesions contain a specific HPV type.* J Pathol, 2012. **227**(1): p. 62-71.
470. Smotkin, D., H. Prokoph, and F.O. Wettstein, *Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms.* J Virol, 1989. **63**(3): p. 1441-7.
471. Griffin, H., et al., *Stratification of HPV-induced cervical pathology using the virally encoded molecular marker E4 in combination with p16 or MCM.* Mod Pathol, 2015. **28**(7): p. 977-93.
472. Doorbar, J., et al., *Characterization of events during the late stages of HPV16 infection in vivo using high-affinity synthetic Fabs to E4.* Virology, 1997. **238**(1): p. 40-52.
473. Banerjee, N.S., et al., *Vorinostat, a pan-HDAC inhibitor, abrogates productive HPV-18 DNA amplification.* Proc Natl Acad Sci U S A, 2018. **115**(47): p. E11138-E11147.
474. Banerjee, N.S., et al., *Targeting DNA Damage Response as a Strategy to Treat HPV Infections.* Int J Mol Sci, 2019. **20**(21).
475. Ridge, K.M., et al., *Keratin 8 phosphorylation by protein kinase C delta regulates shear stress-mediated disassembly of keratin intermediate filaments in alveolar epithelial cells.* J Biol Chem, 2005. **280**(34): p. 30400-5.
476. Fausther, M., L. Villeneuve, and M. Cadrin, *Heat shock protein 70 expression, keratin phosphorylation and Mallory body formation in hepatocytes from griseofulvin-intoxicated mice.* Comp Hepatol, 2004. **3**(1): p. 5.
477. Toivola, D.M., et al., *Type II keratins are phosphorylated on a unique motif during stress and mitosis in tissues and cultured cells.* Mol Biol Cell, 2002. **13**(6): p. 1857-70.
478. Izawa, I. and M. Inagaki, *Regulatory mechanisms and functions of intermediate filaments: a study using site- and phosphorylation state-specific antibodies.* Cancer Sci, 2006. **97**(3): p. 167-74.
479. Ku, N.O. and M.B. Omary, *Keratins turn over by ubiquitination in a phosphorylation-modulated fashion.* J Cell Biol, 2000. **149**(3): p. 547-52.
480. Khan, J., et al., *Role of calpain in the formation of human papillomavirus type 16 E1^E4 amyloid fibers and reorganization of the keratin network.* J Virol, 2011. **85**(19): p. 9984-97.
481. Woods, R.S.R., et al., *Cytokeratin 7 and 19 expression in oropharyngeal and oral squamous cell carcinoma.* Eur Arch Otorhinolaryngol, 2021.
482. Hernandez, B.Y., et al., *Circumcision and human papillomavirus infection in men: a site-specific comparison.* J Infect Dis, 2008. **197**(6): p. 787-94.
483. Mills, A.M., et al., *CK7 Immunohistochemistry as a Predictor of CIN1 Progression: A Retrospective Study of Patients From the Quadrivalent HPV Vaccine Trials.* Am J Surg Pathol, 2017. **41**(2): p. 143-152.
484. Paquette, C., A.M. Mills, and M.H. Stoler, *Predictive Value of Cytokeratin 7 Immunohistochemistry in Cervical Low-grade Squamous Intraepithelial Lesion as a Marker for Risk of Progression to a High-grade Lesion.* Am J Surg Pathol, 2016. **40**(2): p. 236-43.
485. Brant, A.C., et al., *Characterization of HPV integration, viral gene expression and E6E7 alternative transcripts by RNA-Seq: A descriptive study in invasive cervical cancer.* Genomics, 2019. **111**(6): p. 1853-1861.
486. Goodwin, S., J.D. McPherson, and W.R. McCombie, *Coming of age: ten years of next-generation sequencing technologies.* Nat Rev Genet, 2016. **17**(6): p. 333-51.
487. Soneson, C., et al., *A comprehensive examination of Nanopore native RNA sequencing for characterization of complex transcriptomes.* Nat Commun, 2019. **10**(1): p. 3359.
488. McNaughton, A.L., et al., *Illumina and Nanopore methods for whole genome sequencing of hepatitis B virus (HBV).* Sci Rep, 2019. **9**(1): p. 7081.
489. Hebert, T.L., et al., *Culture effects of epidermal growth factor (EGF) and basic fibroblast growth factor (bFGF) on cryopreserved human adipose-derived stromal/stem cell proliferation and adipogenesis.* J Tissue Eng Regen Med, 2009. **3**(7): p. 553-61.

490. Hennings, H., et al., *Calcium regulation of growth and differentiation of mouse epidermal cells in culture.* Cell, 1980. **19**(1): p. 245-54.
491. Schmitt, M., et al., *The HPV16 transcriptome in cervical lesions of different grades.* Mol Cell Probes, 2011. **25**(5-6): p. 260-5.
492. Kajitani, N., et al., *Productive Lifecycle of Human Papillomaviruses that Depends Upon Squamous Epithelial Differentiation.* Front Microbiol, 2012. **3**: p. 152.
493. Desaintes, C. and C. Demeret, *Control of papillomavirus DNA replication and transcription.* Semin Cancer Biol, 1996. **7**(6): p. 339-47.
494. Lace, M.J., et al., *Functional mapping of the human papillomavirus type 16 E1 cistron.* J Virol, 2008. **82**(21): p. 10724-34.
495. Hummel, M., J.B. Hudson, and L.A. Laimins, *Differentiation-induced and constitutive transcription of human papillomavirus type 31b in cell lines containing viral episomes.* J Virol, 1992. **66**(10): p. 6070-80.
496. Ozbun, M.A. and C. Meyers, *Characterization of late gene transcripts expressed during vegetative replication of human papillomavirus type 31b.* J Virol, 1997. **71**(7): p. 5161-72.
497. Olmedo-Nieva, L., et al., *The Role of E6 Spliced Isoforms (E6*) in Human Papillomavirus-Induced Carcinogenesis.* Viruses, 2018. **10**(1).
498. Ajiro, M. and Z.M. Zheng, *E6^E7, a novel splice isoform protein of human papillomavirus 16, stabilizes viral E6 and E7 oncoproteins via HSP90 and GRP78.* mBio, 2015. **6**(1): p. e02068-14.
499. Pastuszak-Lewandoska, D., et al., *HPV16 E6*II gene expression in intraepithelial cervical lesions as an indicator of neoplastic grade: a pilot study.* Med Oncol, 2014. **31**(3): p. 842.
500. Cerasuolo, A., et al., *Comparative analysis of HPV16 gene expression profiles in cervical and in oropharyngeal squamous cell carcinoma.* Oncotarget, 2017. **8**(21): p. 34070-34081.
501. Cricca, M., et al., *Molecular analysis of HPV 16 E6I/E6II spliced mRNAs and correlation with the viral physical state and the grade of the cervical lesion.* J Med Virol, 2009. **81**(7): p. 1276-82.
502. Straub, E., et al., *Characterization of the Human Papillomavirus 16 E8 Promoter.* J Virol, 2015. **89**(14): p. 7304-13.
503. Nulton, T.J., et al., *Analysis of The Cancer Genome Atlas sequencing data reveals novel properties of the human papillomavirus 16 genome in head and neck squamous cell carcinoma.* Oncotarget, 2017. **8**(11): p. 17684-17699.
504. Ajiro, M., et al., *Intron definition and a branch site adenosine at nt 385 control RNA splicing of HPV16 E6*I and E7 expression.* PLoS One, 2012. **7**(10): p. e46412.
505. Alloul, N. and L. Sherman, *The E2 protein of human papillomavirus type 16 is translated from a variety of differentially spliced polycistronic mRNAs.* J Gen Virol, 1999. **80 (Pt 1)**: p. 29-37.
506. Zheng, Y., et al., *Efficient production of HPV16 E2 protein from HPV16 late mRNAs spliced from SD880 to SA2709.* Virus Res, 2020. **285**: p. 198004.
507. Carper, M.B., et al., *An Immunocompetent Mouse Model of HPV16(+) Head and Neck Squamous Cell Carcinoma.* Cell Rep, 2019. **29**(6): p. 1660-1674 e7.
508. Jeon, S. and P.F. Lambert, *Integration of human papillomavirus type 16 DNA into the human genome leads to increased stability of E6 and E7 mRNAs: implications for cervical carcinogenesis.* Proc Natl Acad Sci U S A, 1995. **92**(5): p. 1654-8.
509. Brancaccio, R.N., et al., *MinION nanopore sequencing and assembly of a complete human papillomavirus genome.* J Virol Methods, 2021. **294**: p. 114180.
510. Yang, W., et al., *Accurate Detection of HPV Integration Sites in Cervical Cancer Samples Using the Nanopore MinION Sequencer Without Error Correction.* Front Genet, 2020. **11**: p. 660.
511. Bejjani, F., et al., *The AP-1 transcriptional complex: Local switch or remote command?* Biochim Biophys Acta Rev Cancer, 2019. **1872**(1): p. 11-23.
512. Hess, J., P. Angel, and M. Schorpp-Kistner, *AP-1 subunits: quarrel and harmony among siblings.* J Cell Sci, 2004. **117**(Pt 25): p. 5965-73.
513. Bakiri, L., et al., *Promoter specificity and biological activity of tethered AP-1 dimers.* Mol Cell Biol, 2002. **22**(13): p. 4952-64.
514. Kyo, S., et al., *Expression of AP1 during cellular differentiation determines human papillomavirus E6/E7 expression in stratified epithelial cells.* J Gen Virol, 1997. **78 (Pt 2)**: p. 401-11.
515. Thierry, F., et al., *Two AP1 sites binding JunB are essential for human papillomavirus type 18 transcription in keratinocytes.* J Virol, 1992. **66**(6): p. 3740-8.

516. Chan, W.K., et al., *Transcription of the transforming genes of the oncogenic human papillomavirus-16 is stimulated by tumor promotors through AP1 binding sites.* Nucleic Acids Res, 1990. **18**(4): p. 763-9.
517. Prusty, B.K. and B.C. Das, *Constitutive activation of transcription factor AP-1 in cervical cancer and suppression of human papillomavirus (HPV) transcription and AP-1 activity in HeLa cells by curcumin.* Int J Cancer, 2005. **113**(6): p. 951-60.
518. Gaykalova, D.A., et al., *NF-kappaB and stat3 transcription factor signatures differentiate HPV-positive and HPV-negative head and neck squamous cell carcinoma.* Int J Cancer, 2015. **137**(8): p. 1879-89.
519. Sepiashvili, L., et al., *Integrated omic analysis of oropharyngeal carcinomas reveals human papillomavirus (HPV)-dependent regulation of the activator protein 1 (AP-1) pathway.* Mol Cell Proteomics, 2014. **13**(12): p. 3572-84.
520. Morosov, A., W.C. Phelps, and P. Raychaudhuri, *Activation of the c-fos gene by the HPV16 oncoproteins depends upon the cAMP-response element at -60.* J Biol Chem, 1994. **269**(28): p. 18434-40.
521. Gupta, S., et al., *Selective participation of c-Jun with Fra-2/c-Fos promotes aggressive tumor phenotypes and poor prognosis in tongue cancer.* Sci Rep, 2015. **5**: p. 16811.
522. Mishra, A., et al., *Transactivation and expression patterns of Jun and Fos/AP-1 super-family proteins in human oral cancer.* Int J Cancer, 2010. **126**(4): p. 819-29.
523. de Wilde, J., et al., *Alterations in AP-1 and AP-1 regulatory genes during HPV-induced carcinogenesis.* Cell Oncol, 2008. **30**(1): p. 77-87.
524. Wang, M., et al., *HPV E7-mediated NCAPH ectopic expression regulates the carcinogenesis of cervical carcinoma via PI3K/AKT/SGK pathway.* Cell Death Dis, 2020. **11**(12): p. 1049.
525. Favre, M., et al., *Chromatin-like structures obtained after alkaline disruption of bovine and human papillomaviruses.* J Virol, 1977. **21**(3): p. 1205-9.
526. Stunkel, W. and H.U. Bernard, *The chromatin structure of the long control region of human papillomavirus type 16 represses viral oncoprotein expression.* J Virol, 1999. **73**(3): p. 1918-30.
527. McNicol, P., et al., *Expression of human papillomavirus type 16 E6-E7 open reading frame varies quantitatively in biopsy tissue from different grades of cervical intraepithelial neoplasia.* J Clin Microbiol, 1995. **33**(5): p. 1169-73.
528. Meyers, C., T.J. Mayer, and M.A. Ozbun, *Synthesis of infectious human papillomavirus type 18 in differentiating epithelium transfected with viral DNA.* J Virol, 1997. **71**(10): p. 7381-6.
529. Schmitt, M. and M. Pawlita, *The HPV transcriptome in HPV16 positive cell lines.* Mol Cell Probes, 2011. **25**(2-3): p. 108-13.
530. Strati, K. and P.F. Lambert, *Role of Rb-dependent and Rb-independent functions of papillomavirus E7 oncogene in head and neck cancer.* Cancer Res, 2007. **67**(24): p. 11585-93.
531. Jabbar, S.F., et al., *Persistence of high-grade cervical dysplasia and cervical cancer requires the continuous expression of the human papillomavirus type 16 E7 oncogene.* Cancer Res, 2009. **69**(10): p. 4407-14.
532. Jabbar, S.F., et al., *Cervical cancers require the continuous expression of the human papillomavirus type 16 E7 oncoprotein even in the presence of the viral E6 oncoprotein.* Cancer Res, 2012. **72**(16): p. 4008-16.
533. Ammermann, I., et al., *Inhibition of transcription and DNA replication by the papillomavirus E8-E2C protein is mediated by interaction with corepressor molecules.* J Virol, 2008. **82**(11): p. 5127-36.
534. Fertey, J., et al., *Interaction of the papillomavirus E8--E2C protein with the cellular CHD6 protein contributes to transcriptional repression.* J Virol, 2010. **84**(18): p. 9505-15.
535. Stubenrauch, F., et al., *The E8E2C protein, a negative regulator of viral transcription and replication, is required for extrachromosomal maintenance of human papillomavirus type 31 in keratinocytes.* J Virol, 2000. **74**(3): p. 1178-86.
536. Doorbar, J., et al., *The biology and life-cycle of human papillomaviruses.* Vaccine, 2012. **30 Suppl 5**: p. F55-70.
537. Stubenrauch, F. and L.A. Laimins, *Human papillomavirus life cycle: active and latent phases.* Semin Cancer Biol, 1999. **9**(6): p. 379-86.
538. Schmitz, M., et al., *Non-random integration of the HPV genome in cervical cancer.* PLoS One, 2012. **7**(6): p. e39632.
539. Walline, H.M., et al., *Genomic Integration of High-Risk HPV Alters Gene Expression in Oropharyngeal Squamous Cell Carcinoma.* Mol Cancer Res, 2016. **14**(10): p. 941-952.

540. Srivastava, S., et al., *Downregulated Expression of WWOX in Cervical Carcinoma: A Case-Control Study.* Int J Mol Cell Med, 2020. **9**(4): p. 273-288.
541. Gao, G. and D.I. Smith, *Human Papillomavirus and the Development of Different Cancers.* Cytogenet Genome Res, 2016. **150**(3-4): p. 185-193.
542. Qu, J., et al., *WWOX induces apoptosis and inhibits proliferation in cervical cancer and cell lines.* Int J Mol Med, 2013. **31**(5): p. 1139-47.
543. Chen, J., *Signaling pathways in HPV-associated cancers and therapeutic implications.* Rev Med Virol, 2015. **25 Suppl 1**: p. 24-53.
544. Smith, S.P., et al., *Identification of host transcriptional networks showing concentration-dependent regulation by HPV16 E6 and E7 proteins in basal cervical squamous epithelial cells.* Sci Rep, 2016. **6**: p. 29832.
545. Hyland, P.L., et al., *Evidence for alteration of EZH2, BMI1, and KDM6A and epigenetic reprogramming in human papillomavirus type 16 E6/E7-expressing keratinocytes.* J Virol, 2011. **85**(21): p. 10999-1006.
546. Yim, E.K. and J.S. Park, *The role of HPV E6 and E7 oncoproteins in HPV-associated cervical carcinogenesis.* Cancer Res Treat, 2005. **37**(6): p. 319-24.
547. Taberna, M., et al., *Human papillomavirus-related oropharyngeal cancer.* Ann Oncol, 2017. **28**(10): p. 2386-2398.
548. Chung, C.H., et al., *Genomic alterations in head and neck squamous cell carcinoma determined by cancer gene-targeted sequencing.* Ann Oncol, 2015. **26**(6): p. 1216-1223.
549. Lechner, M., et al., *Targeted next-generation sequencing of head and neck squamous cell carcinoma identifies novel genetic alterations in HPV+ and HPV- tumors.* Genome Med, 2013. **5**(5): p. 49.
550. Stransky, N., et al., *The mutational landscape of head and neck squamous cell carcinoma.* Science, 2011. **333**(6046): p. 1157-60.
551. Wilson, G.A., et al., *Integrated virus-host methylome analysis in head and neck squamous cell carcinoma.* Epigenetics, 2013. **8**(9): p. 953-61.
552. Walter, V., et al., *Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes.* PLoS One, 2013. **8**(2): p. e56823.
553. Palmer, E., et al., *Human papillomavirus infection is rare in nonmalignant tonsil tissue in the UK: implications for tonsil cancer precursor lesions.* Int J Cancer, 2014. **135**(10): p. 2437-43.
554. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
555. Lewis, J.S., Jr., et al., *p16 positive oropharyngeal squamous cell carcinoma:an entity with a favorable prognosis regardless of tumor HPV status.* Am J Surg Pathol, 2010. **34**(8): p. 1088-96.
556. Geissler, C., et al., *The role of p16 expression as a predictive marker in HPV-positive oral SCCHN--a retrospective single-center study.* Anticancer Res, 2013. **33**(3): p. 913-6.
557. Liberzon, A., et al., *The Molecular Signatures Database (MSigDB) hallmark gene set collection.* Cell Syst, 2015. **1**(6): p. 417-425.
558. Wagar, L.E., et al., *Modeling human adaptive immune responses with tonsil organoids.* Nat Med, 2021. **27**(1): p. 125-135.
559. Huang, Q., et al., *Simple hypertrophic tonsils have more active innate immune and inflammatory responses than hypertrophic tonsils with recurrent inflammation in children.* J Otolaryngol Head Neck Surg, 2020. **49**(1): p. 35.
560. Yoshihama, S., et al., *NLRC5/MHC class I transactivator is a target for immune evasion in cancer.* Proc Natl Acad Sci U S A, 2016. **113**(21): p. 5999-6004.
561. Gil-Yarom, N., et al., *CD74 is a novel transcription regulator.* Proc Natl Acad Sci U S A, 2017. **114**(3): p. 562-567.
562. Espert, L., et al., *ISG20, a new interferon-induced RNase specific for single-stranded RNA, defines an alternative antiviral pathway against RNA genomic viruses.* J Biol Chem, 2003. **278**(18): p. 16151-8.
563. Seliga, A., et al., *Kallikrein-Kinin System Suppresses Type I Interferon Responses: A Novel Pathway of Interferon Regulation.* Front Immunol, 2018. **9**: p. 156.
564. Szklarczyk, D., et al., *STRING v10: protein-protein interaction networks, integrated over the tree of life.* Nucleic Acids Res, 2015. **43**(Database issue): p. D447-52.
565. Kaneko, N., et al., *The role of interleukin-1 in general pathology.* Inflamm Regen, 2019. **39**: p. 12.
566. Dinarello, C.A., et al., *Interleukin-18 and IL-18 binding protein.* Front Immunol, 2013. **4**: p. 289.

567. Gschwandtner, M., R. Derler, and K.S. Midwood, *More Than Just Attractive: How CCL2 Influences Myeloid Cell Behavior Beyond Chemotaxis.* Front Immunol, 2019. **10**: p. 2759.
568. Bulfone-Paus, S., et al., *Interleukin-15 protects from lethal apoptosis in vivo.* Nat Med, 1997. **3**(10): p. 1124-8.
569. Villarroel Dorrego, M., et al., *Transfection and ligation of CD40 in human oral keratinocytes affect proliferation, adhesion and migration but not apoptosis in vitro.* Clin Exp Dermatol, 2006. **31**(2): p. 266-71.
570. Denfeld, R.W., et al., *CD40 is functionally expressed on human keratinocytes.* Eur J Immunol, 1996. **26**(10): p. 2329-34.
571. Hill, S.C., et al., *Activation of CD40 in cervical carcinoma cells facilitates CTL responses and augments chemotherapy-induced apoptosis.* J Immunol, 2005. **174**(1): p. 41-50.
572. Kang, S.D., et al., *Effect of Productive Human Papillomavirus 16 Infection on Global Gene Expression in Cervical Epithelium.* J Virol, 2018. **92**(20).
573. Rosty, C., et al., *Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma.* Oncogene, 2005. **24**(47): p. 7094-104.
574. Garner-Hamrick, P.A., et al., *Global effects of human papillomavirus type 18 E6/E7 in an organotypic keratinocyte culture system.* J Virol, 2004. **78**(17): p. 9041-50.
575. Saxton, R.A. and D.M. Sabatini, *mTOR Signaling in Growth, Metabolism, and Disease.* Cell, 2017. **168**(6): p. 960-976.
576. Zhang, L., et al., *The role of the PI3K/Akt/mTOR signalling pathway in human cancers induced by infection with human papillomaviruses.* Mol Cancer, 2015. **14**: p. 87.
577. Spangle, J.M. and K. Munger, *The human papillomavirus type 16 E6 oncoprotein activates mTORC1 signaling and increases protein synthesis.* J Virol, 2010. **84**(18): p. 9398-407.
578. Zhou, Y., et al., *Increased phosphorylation of p70 S6 kinase is associated with HPV16 infection in cervical cancer and esophageal cancer.* Br J Cancer, 2007. **97**(2): p. 218-22.
579. Spangle, J.M. and K. Munger, *The HPV16 E6 oncoprotein causes prolonged receptor protein tyrosine kinase signaling and enhances internalization of phosphorylated receptor species.* PLoS Pathog, 2013. **9**(3): p. e1003237.
580. Peterson, T.R., et al., *DEPTOR is an mTOR inhibitor frequently overexpressed in multiple myeloma cells and required for their survival.* Cell, 2009. **137**(5): p. 873-86.
581. Srinivas, K.P., et al., *DEPTOR promotes survival of cervical squamous cell carcinoma cells and its silencing induces apoptosis through downregulating PI3K/AKT and by up-regulating p38 MAP kinase.* Oncotarget, 2016. **7**(17): p. 24154-71.
582. Mesev, E.V., R.A. LeDesma, and A. Ploss, *Decoding type I and III interferon signalling during viral infection.* Nat Microbiol, 2019. **4**(6): p. 914-924.
583. McNab, F., et al., *Type I interferons in infectious disease.* Nat Rev Immunol, 2015. **15**(2): p. 87-103.
584. Taniguchi, T. and A. Takaoka, *A weak signal for strong responses: interferon-alpha/beta revisited.* Nat Rev Mol Cell Biol, 2001. **2**(5): p. 378-86.
585. Hoebe, K., et al., *Identification of Lps2 as a key transducer of MyD88-independent TIR signalling.* Nature, 2003. **424**(6950): p. 743-8.
586. Osborn, L., S. Kunkel, and G.J. Nabel, *Tumor necrosis factor alpha and interleukin 1 stimulate the human immunodeficiency virus enhancer by activation of the nuclear factor kappa B.* Proc Natl Acad Sci U S A, 1989. **86**(7): p. 2336-40.
587. Medzhitov, R., P. Preston-Hurlburt, and C.A. Janeway, Jr., *A human homologue of the Drosophila Toll protein signals activation of adaptive immunity.* Nature, 1997. **388**(6640): p. 394-7.
588. Liu, T., et al., *NF-kappaB signaling in inflammation.* Signal Transduct Target Ther, 2017. **2**.
589. Guttridge, D.C., et al., *NF-kappaB controls cell growth and differentiation through transcriptional regulation of cyclin D1.* Mol Cell Biol, 1999. **19**(8): p. 5785-99.
590. Chen, F., V. Castranova, and X. Shi, *New insights into the role of nuclear factor-kappaB in cell growth regulation.* Am J Pathol, 2001. **159**(2): p. 387-97.
591. Vandermark, E.R., et al., *Human papillomavirus type 16 E6 and E 7 proteins alter NF-kB in cultured cervical epithelial cells and inhibition of NF-kB promotes cell growth and immortalization.* Virology, 2012. **425**(1): p. 53-60.

592. Spitkovsky, D., et al., *The human papillomavirus oncoprotein E7 attenuates NF-kappa B activation by targeting the Ikappa B kinase complex.* J Biol Chem, 2002. **277**(28): p. 25576-82.
593. Richards, K.H., et al., *The human papillomavirus (HPV) E7 protein antagonises an Imiquimod-induced inflammatory pathway in primary human keratinocytes.* Sci Rep, 2015. **5**: p. 12922.
594. Senftleben, U., et al., *Activation by IKKalpha of a second, evolutionary conserved, NF-kappa B signaling pathway.* Science, 2001. **293**(5534): p. 1495-9.
595. Rui, D.A.C., et al., *The NFkappaB Signaling Pathway in Papillomavirus-induced Lesions: Friend or Foe?* Anticancer Res, 2016. **36**(5): p. 2073-83.
596. Mishra, A., et al., *Differential expression and activation of NF-kappaB family proteins during oral carcinogenesis: Role of high risk human papillomavirus infection.* Int J Cancer, 2006. **119**(12): p. 2840-50.
597. Pastore, S., et al., *Interferon-gamma promotes exaggerated cytokine production in keratinocytes cultured from patients with atopic dermatitis.* J Allergy Clin Immunol, 1998. **101**(4 Pt 1): p. 538-44.
598. Lee, A.J., et al., *Inflammatory monocytes require type I interferon receptor signaling to activate NK cells via IL-18 during a mucosal viral infection.* J Exp Med, 2017. **214**(4): p. 1153-1167.
599. Takei, Y., et al., *Central role for interferon-gamma receptor in the regulation of renal MHC expression.* J Am Soc Nephrol, 2000. **11**(2): p. 250-261.
600. Thelemann, C., et al., *Interferon-gamma induces expression of MHC class II on intestinal epithelial cells and protects mice from colitis.* PLoS One, 2014. **9**(1): p. e86844.
601. Talbot, N.C., et al., *Quantitative and semiquantitative immunoassay of growth factors and cytokines in the conditioned medium of STO and CF-1 mouse feeder cells.* In Vitro Cell Dev Biol Anim, 2012. **48**(1): p. 1-11.
602. Hynds, R.E., et al., *Cross-talk between human airway epithelial cells and 3T3-J2 feeder cells involves partial activation of human MET by murine HGF.* PLoS One, 2018. **13**(5): p. e0197129.
603. Rodero, M.P., et al., *Detection of interferon alpha protein reveals differential levels and cellular sources in disease.* J Exp Med, 2017. **214**(5): p. 1547-1555.
604. Noursadeghi, M., et al., *Quantitative imaging assay for NF-kappaB nuclear translocation in primary human macrophages.* J Immunol Methods, 2008. **329**(1-2): p. 194-200.
605. Elgueta, R., et al., *Molecular mechanism and function of CD40/CD40L engagement in the immune system.* Immunol Rev, 2009. **229**(1): p. 152-72.
606. Companjen, A.R., et al., *CD40 ligation-induced cytokine production in human skin explants is partly mediated via IL-1.* Int Immunol, 2002. **14**(6): p. 669-76.
607. Altenburg, A., et al., *CD40 ligand-CD40 interaction induces chemokines in cervical carcinoma cells in synergism with IFN-gamma.* J Immunol, 1999. **162**(7): p. 4140-7.
608. Jaiswal, A.I. and M. Croft, *CD40 ligand induction on T cell subsets by peptide-presenting B cells: implications for development of the primary T and B cell response.* J Immunol, 1997. **159**(5): p. 2282-91.
609. Moschonas, A., M. Ioannou, and A.G. Eliopoulos, *CD40 stimulates a "feed-forward" NF-kappaB-driven molecular pathway that regulates IFN-beta expression in carcinoma cells.* J Immunol, 2012. **188**(11): p. 5521-7.
610. Dainichi, T., et al., *Immune Control by TRAF6-Mediated Pathways of Epithelial Cells in the EIME (Epithelial Immune Microenvironment).* Front Immunol, 2019. **10**: p. 1107.
611. Li, H. and E.P. Nord, *CD40 ligation stimulates MCP-1 and IL-8 production, TRAF6 recruitment, and MAPK activation in proximal tubule cells.* Am J Physiol Renal Physiol, 2002. **282**(6): p. F1020-33.
612. King, A.E., et al., *Cd40 expression in uterine tissues: a key regulator of cytokine expression by fibroblasts.* J Clin Endocrinol Metab, 2001. **86**(1): p. 405-12.
613. Tummers, B., et al., *CD40-mediated amplification of local immunity by epithelial cells is impaired by HPV.* J Invest Dermatol, 2014. **134**(12): p. 2918-2927.
614. Manzo-Merino, J., et al., *The Human Papillomavirus (HPV) E6 Oncoprotein Regulates CD40 Expression via the AT-Hook Transcription Factor AKNA.* Cancers (Basel), 2018. **10**(12).
615. Peguet-Navarro, J., et al., *CD40 ligation of human keratinocytes inhibits their proliferation and induces their differentiation.* J Immunol, 1997. **158**(1): p. 144-52.
616. Starke, A., R.P. Wuthrich, and Y. Waeckerle-Men, *TGF-beta treatment modulates PD-L1 and CD40 expression in proximal renal tubular epithelial cells and enhances CD8 cytotoxic T-cell responses.* Nephron Exp Nephrol, 2007. **107**(1): p. e22-9.

617. van Kooten, C., et al., *Synergistic effect of interleukin-1 and CD40L on the activation of human renal tubular epithelial cells.* Kidney Int, 1999. **56**(1): p. 41-51.
618. Galy, A.H. and H. Spits, *CD40 is functionally expressed on human thymic epithelial cells.* J Immunol, 1992. **149**(3): p. 775-82.
619. Propst, S.M., et al., *Proinflammatory and Th2-derived cytokines modulate CD40-mediated expression of inflammatory mediators in airway epithelia: implications for the role of epithelial CD40 in airway inflammation.* J Immunol, 2000. **165**(4): p. 2214-21.
620. Propst, S.M., K. Estell, and L.M. Schwiebert, *CD40-mediated activation of NF-kappa B in airway epithelial cells.* J Biol Chem, 2002. **277**(40): p. 37054-63.
621. Morhenn, V.B. and B.J. Nickoloff, *Interleukin-2 stimulates resting human T lymphocytes' response to allogeneic, gamma interferon-treated keratinocytes.* J Invest Dermatol, 1987. **89**(5): p. 464-8.
622. Niederwieser, D., et al., *IFN-mediated induction of MHC antigen expression on human keratinocytes and its influence on in vitro alloimmune responses.* J Immunol, 1988. **140**(8): p. 2556-64.
623. Heller, C., et al., *Identification of key amino acid residues that determine the ability of high risk HPV16-E7 to dysregulate major histocompatibility complex class I expression.* J Biol Chem, 2011. **286**(13): p. 10983-97.
624. Li, W., et al., *Down-regulation of HLA class I antigen in human papillomavirus type 16 E7 expressing HaCaT cells: correlate with TAP-1 expression.* Int J Gynecol Cancer, 2010. **20**(2): p. 227-32.
625. Bottley, G., et al., *High-risk human papillomavirus E7 expression reduces cell-surface MHC class I molecules and increases susceptibility to natural killer cells.* Oncogene, 2008. **27**(12): p. 1794-9.
626. Abele, R. and R. Tampe, *The ABCs of immunology: structure and function of TAP, the transporter associated with antigen processing.* Physiology (Bethesda), 2004. **19**: p. 216-24.
627. Nasman, A., et al., *MHC class I expression in HPV positive and negative tonsillar squamous cell carcinoma in correlation to clinical outcome.* Int J Cancer, 2013. **132**(1): p. 72-81.
628. Deshmane, S.L., et al., *Monocyte chemoattractant protein-1 (MCP-1): an overview.* J Interferon Cytokine Res, 2009. **29**(6): p. 313-26.
629. Riethdorf, S., et al., *Expression of the MCP-1 gene and the HPV 16 E6/E7 oncogenes in squamous cell carcinomas of the cervix uteri and metastases.* Pathobiology, 1998. **66**(6): p. 260-7.
630. Zijlmans, H.J., et al., *The absence of CCL2 expression in cervical carcinoma is associated with increased survival and loss of heterozygosity at 17q11.2.* J Pathol, 2006. **208**(4): p. 507-17.
631. Kleine-Lowinski, K., et al., *Selective suppression of monocyte chemoattractant protein-1 expression by human papillomavirus E6 and E7 oncoproteins in human cervical epithelial and epidermal cells.* Int J Cancer, 2003. **107**(3): p. 407-15.
632. Dutta, P., et al., *Transcriptional Regulation of CCL2 by PARP1 Is a Driver for Invasiveness in Breast Cancer.* Cancers (Basel), 2020. **12**(5).
633. Nakatsumi, H., M. Matsumoto, and K.I. Nakayama, *Noncanonical Pathway for Regulation of CCL2 Expression by an mTORC1-FOXK1 Axis Promotes Recruitment of Tumor-Associated Macrophages.* Cell Rep, 2017. **21**(9): p. 2471-2486.
634. Amador-Molina, A., et al., *Role of innate immunity against human papillomavirus (HPV) infections and effect of adjuvants in promoting specific immune response.* Viruses, 2013. **5**(11): p. 2624-42.
635. Scott, M., M. Nakagawa, and A.B. Moscicki, *Cell-mediated immune response to human papillomavirus infection.* Clin Diagn Lab Immunol, 2001. **8**(2): p. 209-20.
636. Sims, J.E. and D.E. Smith, *The IL-1 family: regulators of immunity.* Nat Rev Immunol, 2010. **10**(2): p. 89-102.
637. Ishida, Y., et al., *Absence of IL-1 receptor antagonist impaired wound healing along with aberrant NF-kappaB activation and a reciprocal suppression of TGF-beta signal pathway.* J Immunol, 2006. **176**(9): p. 5598-606.
638. Swindell, W.R., et al., *RNA-Seq Analysis of IL-1B and IL-36 Responses in Epidermal Keratinocytes Identifies a Shared MyD88-Dependent Gene Signature.* Front Immunol, 2018. **9**: p. 80.
639. Sanmiguel, J.C., et al., *Interleukin-1 regulates keratinocyte expression of T cell targeting chemokines through interleukin-1 receptor associated kinase-1 (IRAK1) dependent and independent pathways.* Cell Signal, 2009. **21**(5): p. 685-94.
640. Chan, B.C.L., et al., *IL33: Roles in Allergic Inflammation and Therapeutic Perspectives.* Front Immunol, 2019. **10**: p. 364.

641. Companjen, A.R., et al., *Human keratinocytes are major producers of IL-18: predominant expression of the unprocessed form.* Eur Cytokine Netw, 2000. **11**(3): p. 383-90.
642. Sugawara, S., et al., *Neutrophil proteinase 3-mediated induction of bioactive IL-18 secretion by human oral epithelial cells.* J Immunol, 2001. **167**(11): p. 6568-75.
643. Chen, C.L., et al., *Stat3 activation in human endometrial and cervical cancers.* Br J Cancer, 2007. **96**(4): p. 591-9.
644. Morgan, E.L. and A. Macdonald, *JAK2 Inhibition Impairs Proliferation and Sensitises Cervical Cancer Cells to Cisplatin-Induced Cell Death.* Cancers (Basel), 2019. **11**(12).
645. Sobti, R.C., et al., *Deregulation of STAT-5 isoforms in the development of HPV-mediated cervical carcinogenesis.* J Recept Signal Transduct Res, 2010. **30**(3): p. 178-88.
646. Morgan, E.L. and A. Macdonald, *Autocrine STAT3 activation in HPV positive cervical cancer through a virus-driven Rac1-NFkappaB-IL-6 signalling axis.* PLoS Pathog, 2019. **15**(6): p. e1007835.
647. Sasaki, E., et al., *Frequent KRAS and HRAS mutations in squamous cell papillomas of the head and neck.* J Pathol Clin Res, 2020. **6**(2): p. 154-159.
648. Molinolo, A.A., et al., *mTOR as a molecular target in HPV-associated oral and cervical squamous carcinomas.* Clin Cancer Res, 2012. **18**(9): p. 2558-68.
649. Al-Sahaf, S., et al., *The IL-1/IL-1R axis induces greater fibroblast-derived chemokine release in human papillomavirus-negative compared to positive oropharyngeal cancer.* Int J Cancer, 2019. **144**(2): p. 334-344.
650. Partlova, S., et al., *Distinct patterns of intratumoral immune cell infiltrates in patients with HPV-associated compared to non-virally induced head and neck squamous cell carcinoma.* Oncoimmunology, 2015. **4**(1): p. e965570.
651. Wan, F., et al., *Gene expression changes during HPV-mediated carcinogenesis: a comparison between an in vitro cell model and cervical cancer.* Int J Cancer, 2008. **123**(1): p. 32-40.
652. Klymenko, T., et al., *RNA-Seq Analysis of Differentiated Keratinocytes Reveals a Massive Response to Late Events during Human Papillomavirus 16 Infection, Including Loss of Epithelial Barrier Function.* J Virol, 2017. **91**(24).
653. Wong, Y.F., et al., *Genome-wide gene expression profiling of cervical cancer in Hong Kong women by oligonucleotide microarray.* Int J Cancer, 2006. **118**(10): p. 2461-9.
654. Sopov, I., et al., *Detection of cancer-related gene expression profiles in severe cervical neoplasia.* Int J Cancer, 2004. **112**(1): p. 33-43.
655. Slebos, R.J., et al., *Gene expression differences associated with human papillomavirus status in head and neck squamous cell carcinoma.* Clin Cancer Res, 2006. **12**(3 Pt 1): p. 701-9.
656. Gius, D., et al., *Profiling microdissected epithelium and stroma to model genomic signatures for cervical carcinogenesis accommodating for covariates.* Cancer Res, 2007. **67**(15): p. 7113-23.
657. Villa, L.L., et al., *Differential effect of tumor necrosis factor on proliferation of primary human keratinocytes and cell lines containing human papillomavirus types 16 and 18.* Mol Carcinog, 1992. **6**(1): p. 5-9.
658. Termini, L., et al., *Characterization of global transcription profile of normal and HPV-immortalized keratinocytes and their response to TNF treatment.* BMC Med Genomics, 2008. **1**: p. 29.
659. Dijkstra, K.K., et al., *Generation of Tumor-Reactive T Cells by Co-culture of Peripheral Blood Lymphocytes and Tumor Organoids.* Cell, 2018. **174**(6): p. 1586-1598 e12.
660. Baruah, P., et al., *TLR9 Mediated Tumor-Stroma Interactions in Human Papilloma Virus (HPV)-Positive Head and Neck Squamous Cell Carcinoma Up-Regulate PD-L1 and PD-L2.* Front Immunol, 2019. **10**: p. 1644.
661. Kang, S.Y.C., et al., *Characterization of Epithelial Progenitors in Normal Human Palatine Tonsils and Their HPV16 E6/E7-Induced Perturbation.* Stem Cell Reports, 2015. **5**(6): p. 1210-1225.

Appendix: Supplementary tables and figures

Table S-1: Expression of genes present either in undifferentiated or differentiated mucosal epithelial cells

As demonstrated by Illumina-based RNA-seq. For each gene, the table depicts the p-value and regularised log_2 fold change of each model comparison, along with the number of reads per million as detected by RNA-seq.

Category of genes	Gene	Related protein	HTK#2 vs HTK#3		RNA-seq reads per million					
			p-value	Regularised log_2 fold change	HTK#2	HPV16-HTK#2	HPV18-HTK#2	HTK#3	HPV16-HTK#3	HPV18-HTK#3
Keratin markers of proliferating cells	***KRT5***	Keratin 5	0.000963	-2.5	381906.8	907484.6	643518.6	1679201.5	672139.2	540955
	KRT14	Keratin 14	0.00254	-2.7	171416.7	542518.8	302214.8	804317.7	321335.8	281586.8
Keratin markers of differentiation	***KRT4***	Keratin 4	0.715	-1.088	50618	93975	13373.4	137662.3	129093.9	8263.5
	KRT13	Keratin 13	0.0865	-3.5	36539.2	43076.8	21050.7	437983	58029.8	8272.6
Other markers of differentiation	***IVL***	Involucrin	0.00923	-2.7	6080.9	2835.9	6790	43033.8	7962.2	2725.4
	FLG	Filaggrin	0.00102	-7.5	1	6.3	25.8	160.3	19.2	9.1

Table S-2: Overview of notable genes involved in immune signalling and immune responses

Overviews notable genes involved in immune signalling and immune responses, shown in the RNA-seq data to be significantly differentially expressed between HPV16-HTK and HPV18-HTK models, or between HPV16-HTK and controls models. For each gene, the table depicts the p-value and regularised log_2 fold change of each model comparison, along with the number of reads per million as detected by RNA-seq.

	Gene	HPV16 vs HPV18		HPV16 vs control		HPV18 vs control		RNA-seq reads per million					
		p-value	Regularised log_2 fold change	p-value	Regularised log_2 fold change	p-value	Regularised log_2 fold change	HTK#2	HTK#3	HPV16-HTK#2	HPV16-HTK#3	HPV18-HTK#2	HPV18-HTK#3
Cytokines and chemokines	*CCL2*	0.00129	1.255	0.0163	-0.999	0.334	0.256	56.1	2.1	1.8	0	15.9	40.9
	IL15	0.0326	0.79	0.255	-0.385	0.318	0.405	317.5	110	124	123	301.1	286.1
	IL7	0.00663	1.154	0.0372	-0.695	0.478	0.458	31.1	3.1	0.9	2.1	22.9	31.8
	IFNB1	0.203	0.328	0.029	-0.946	0.303	-0.618	47.1	2.1	0.9	1.1	5	4.5
	CXCL1	0.0528	0.654	0.731	0.116	0.0226	0.77	18376.8	7488.3	12983	13906.8	28434	24076.7
	CCL5	0.0101	-1.214	0.00336	1.27	0.714	0.056	142.2	54.5	875.8	128.3	101.3	91.7
	IL6	0.182	0.692	0.00457	1.261	3.24e-05	1.952	94.2	16.8	191.3	227.7	408.4	435.9
	IL18	0.34	0.288	0.833	-0.058	0.457	0.23	5876.5	1740.1	3923.7	2391.8	5259	3239.5
Cytokine receptors	*IL13RA2*	1.16e-05	2.324	0.00299	-1.407	0.15	0.917	184.3	114.2	6.3	53.5	372.6	367.8
	IL31RA	0.0155	1.039	0.222	-0.522	0.23	0.487	71.1	216.9	17.1	181.8	173.9	263.4
	IL36RN	0.706	0.294	0.00194	-1.481	0.00645	-1.248	331.5	1095.8	203.9	75.9	201.7	138
	IL7R	0.0292	0.891	0.0407	0.703	2.53e-05	1.594	76.1	93.2	95.2	364.6	430.2	540.4
	IL15RA	0.0463	0.857	0.226	-0.463	0.434	0.395	518.8	135.1	150	177.5	448.1	379.6
	IL2RG	0.247	0.505	0.0189	-1.407	0.233	-0.902	287.5	632.8	27.8	239.5	225.5	123.5
TNF receptors	*CD40*	0.00161	1.001	0.0188	-0.72	0.42	0.28	976.6	362.5	293.7	257.7	961.8	668.4
	TNFRSF9	0.00417	1.307	0.772	0.158	0.00163	1.465	30	90.1	47.6	89.8	424.3	172.6
	TNFRSF25	0.0225	0.67	0.191	-0.378	0.33	0.291	396.6	602.4	362	291.9	583.2	752
Transcription factors	*IRF7*	0.552	0.255	0.0488	-0.787	0.158	-0.532	3332.4	849.6	712.3	808.3	996.5	999
	FOS	0.0148	-0.595	0.0408	0.495	0.695	-0.1	39942.7	22576	70996.6	34832.8	31977	22972.4

	FOSB	0.279	-0.386	0.0162	0.751	0.186	0.365	5616.1	2065.9	13308.2	4148.4	5830.3	4454.6
	FOSL1	0.0361	0.667	0.844	-0.075	0.0575	0.592	2213.6	2042.8	1840.6	2095.6	5764.7	2600.1
	MAF	0.0261	-0.539	0.0115	-0.681	2.05e-06	-1.221	1249	2570.8	791.4	1030.7	369.6	671.1
	BCL6	0.0156	-0.66	0.75	0.089	0.0358	-0.572	3342.4	2015.6	3614.7	2233.5	1404.9	1443.1
IL-1 family members	***IL18R1***	0.0276	0.731	0.00204	-1.078	0.373	-0.507	182.3	185.4	42.2	64.2	163.9	96.3
	IL11	0.0066	0.875	0.0285	-0.704	0.596	0.17	59.1	125.7	20.7	57.7	87.4	127.1
	IL1A	0.0292	1.005	0.164	-0.647	0.43	0.358	20825.7	15917.4	4644.2	16986.1	28897	23583.6
	IL1B	0.0419	1.094	0.0706	-0.936	0.821	0.158	26964.7	10669.9	3612	10605.2	26351.4	15172.9
	IL33	0.138	1.323	0.0155	-1.881	0.348	-0.558	4779.7	550	121.3	384.9	2185.9	408.7
	IL36A	0.152	1.269	0.00881	-1.93	0.234	-0.661	1239	260.9	15.3	191.4	358.7	251.6
MHC class I antigen presentation	***B2M***	0.116	0.541	0.0397	-0.682	0.627	-0.141	71176.3	21134.5	21055.1	17851	41050.3	28458.7
	HLA-B	0.245	0.623	0.0356	-0.965	0.348	-0.342	26023.2	5306.1	3609.3	5381.2	10775.3	6776.8
	HLA-G	0.559	0.396	0.0129	-1.443	0.0569	-1.046	472.8	311.1	29.6	177.5	154	105.3
	TAP1	0.257	0.412	0.0292	0.745	0.295	-0.333	8912.4	3756.7	2376.9	3040.7	4798.9	3606.4
	CD1D	0.00585	1.073	0.0171	-0.93	0.708	0.143	97.2	129.9	16.2	66.3	126.2	137.1
MHC class II antigen presentation	***CD74***	0.00883	0.907	0.893	-0.06	0.013	0.847	1098.8	648.5	629.7	996.5	2290.2	1787.3
	HLA-DRA	0.00988	1.043	0.736	0.143	0.00383	1.186	11	18.9	7.2	37.4	42.7	116.2
	HLA-DRB1	0.15	0.649	0.497	0.24	0.0353	0.889	15	15.7	11.7	39.6	58.6	50.9
	HLA-DPA1	0.399	0.324	0.162	0.464	0.0263	1.801	8	17.8	19.8	36.4	36.8	48.1
PRRs	***TLR3***	0.171	0.597	0.0187	-0.943	0.325	-0.346	2478	698.8	422.2	580.6	1021.4	871.9
	TLR4	0.00745	1.089	0.633	-0.278	0.0256	0.812	299.5	1	173.4	1.1	442.1	33.6
	NLRC5	0.0419	0.831	0.102	-0.623	0.692	0.208	572.9	116.3	150.9	119.7	382.5	292.4
	NOD2	0.423	0.224	0.0294	-0.607	0.168	-0.383	773.3	1236.2	547.1	491.8	677.6	642.1
Other immune regulators and effectors	***UCHL1***	0.33	-0.338	6.38e-18	3.066	1.89e-14	2.728	116.2	282.9	8945.2	4878.7	3577.8	6418.1
	TRAF1	0.055	0.553	0.215	0.344	0.00163	0.897	115.2	119.4	142.8	210.6	252.4	386
	TNIP3	0.000707	3.556	0.23	-0.538	0.0279	1.765	158.3	13.6	15.3	43.8	340.8	390.5
	CCDC3	0.0108	-1.387	0.0226	1.383	0.788	-0.004	30	368.8	629.7	448	99.4	151.7
	MX1	0.65	0.374	0.028	-1.213	0.0812	-0.839	54476.3	1548.4	3604	2534	6819.9	2957
	MX2	0.542	1.118	0.0493	-2.028	0.175	-0.91	8522.8	93.2	59.3	245.9	550.4	374.2
	EGF	0.0425	0.704	1.51e-05	-1.65	0.0183	-0.946	202.3	109	15.3	23.5	66.6	40
	IFI27	0.0406	0.894	2.12e-08	-2.406	0.000376	-1.511	35948.2	5215	1325.9	981.5	6442.3	1465.8
	SERPING1	0.000879	2.011	0.000862	-1.124	0.933	-0.023	228.4	81.7	48.5	24.6	143.1	129

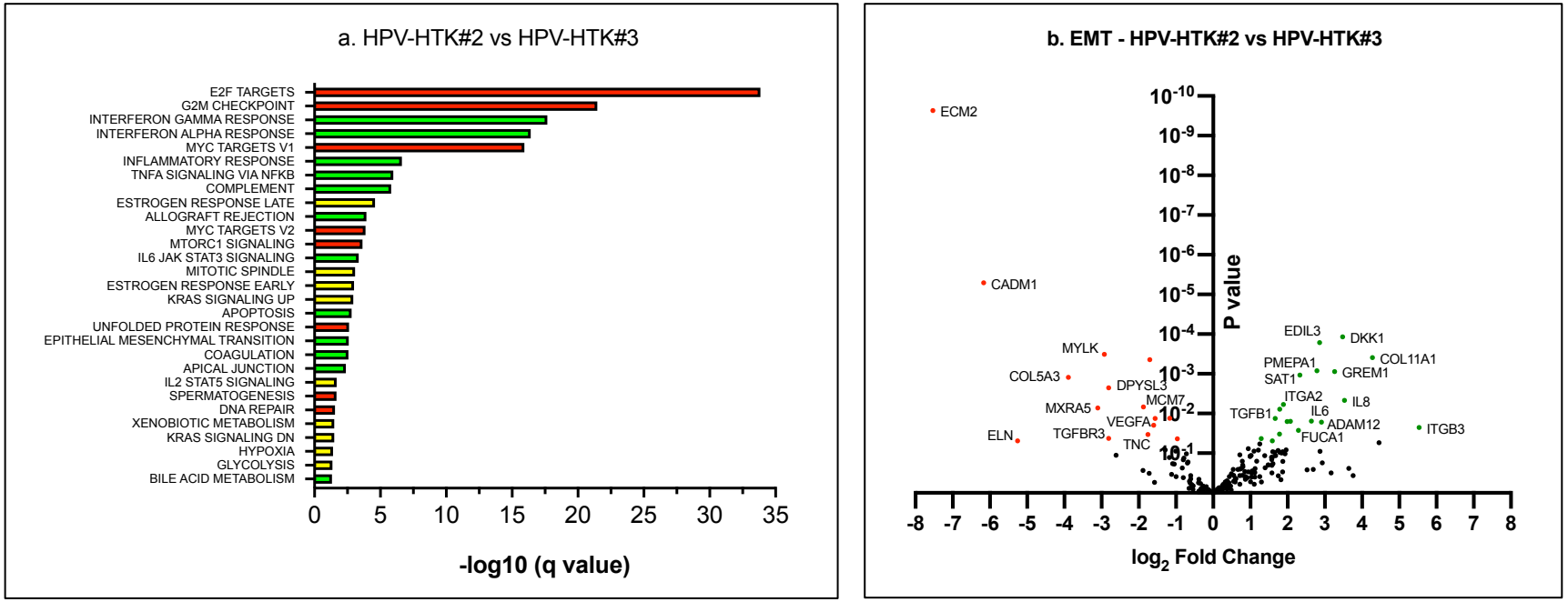


Figure S-1: Differential gene expression analysis comparisons between HTK#2 and HTK#3 donor cells following HPV16/18 transfection

a. GSEA comparing HPV-HTK#2 against HPV-HTK#3 lines. Shows all significantly differentially expressed MSigDB Molecular Hallmark gene sets between the two HPV-transfected donor lines ($q < 0.05$, or -log10 > 1.3). Red bars indicate gene sets in which overall expression is significantly lower in HPV-HTK#3 lines than in HPV-HTK#2 lines; green bars indicate comparably higher expression, and yellow bars indicate up and down expression. **b.** Volcano plots depicting gene expression changes of EMT signatures. Volcano plots compare p-values against regularised $\log_2$ fold change to fully conceptualise differences in gene expression. All significantly altered genes are coloured red or green and are labelled with gene names ($p < 0.05$). Green points indicate genes where expression was significantly higher in HPV-HTK#3 cells than HPV-HTK#2. Red shows the opposite.

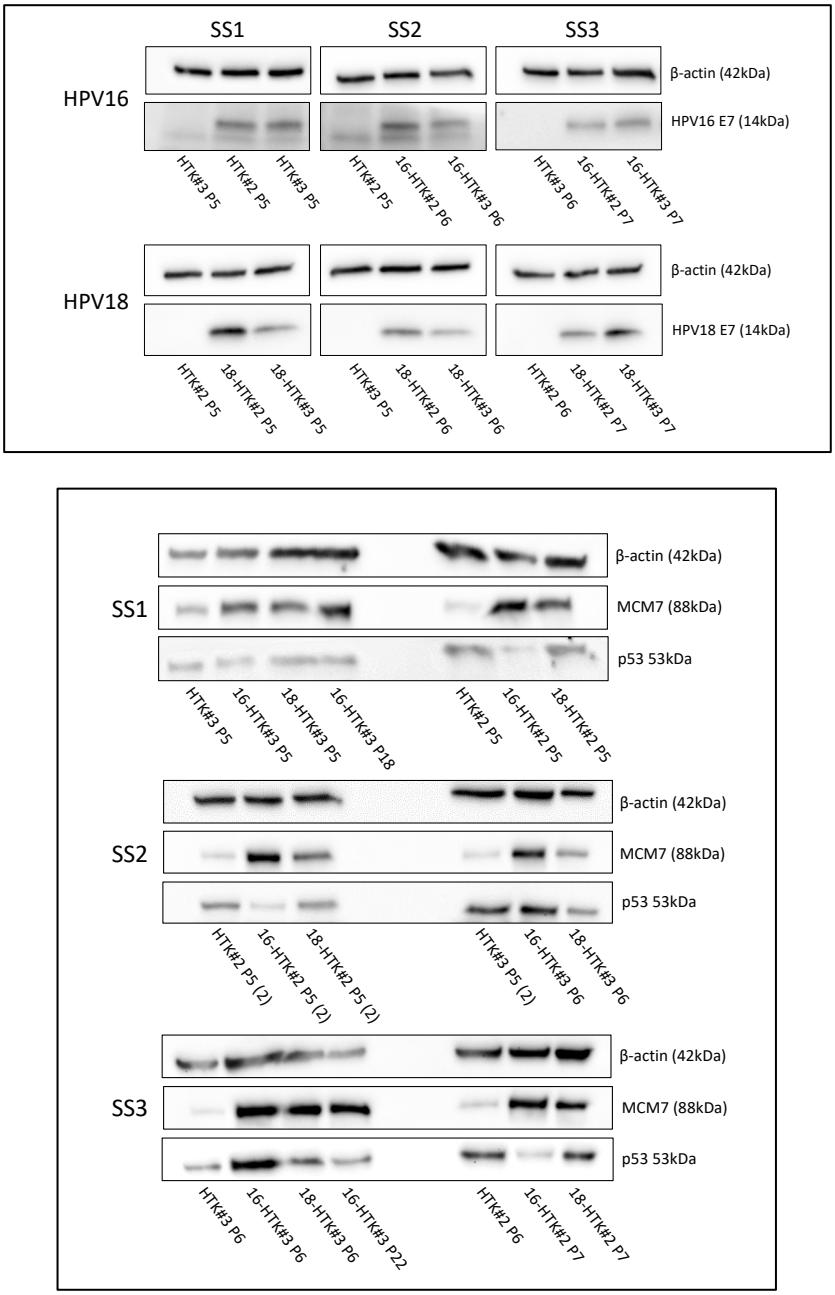


Figure S-2: Full western plot panel used for densitometry-based analyses comparing HPV16- and HPV18-HTK cell lines

a. Westerns blots examined levels of HPV16- and HPV18-E7 and **b.** Western blots examined levels of MCM7 and p53, as shown in Figure 4-7 and Figure 4-8 respectively. Passage number is shown as (P). The three sample sets (SS) comprise of three independent replicates of sequentially passaged cell lines. Primary and secondary antibodies used for these analyses are overviewed in Table 2-3.

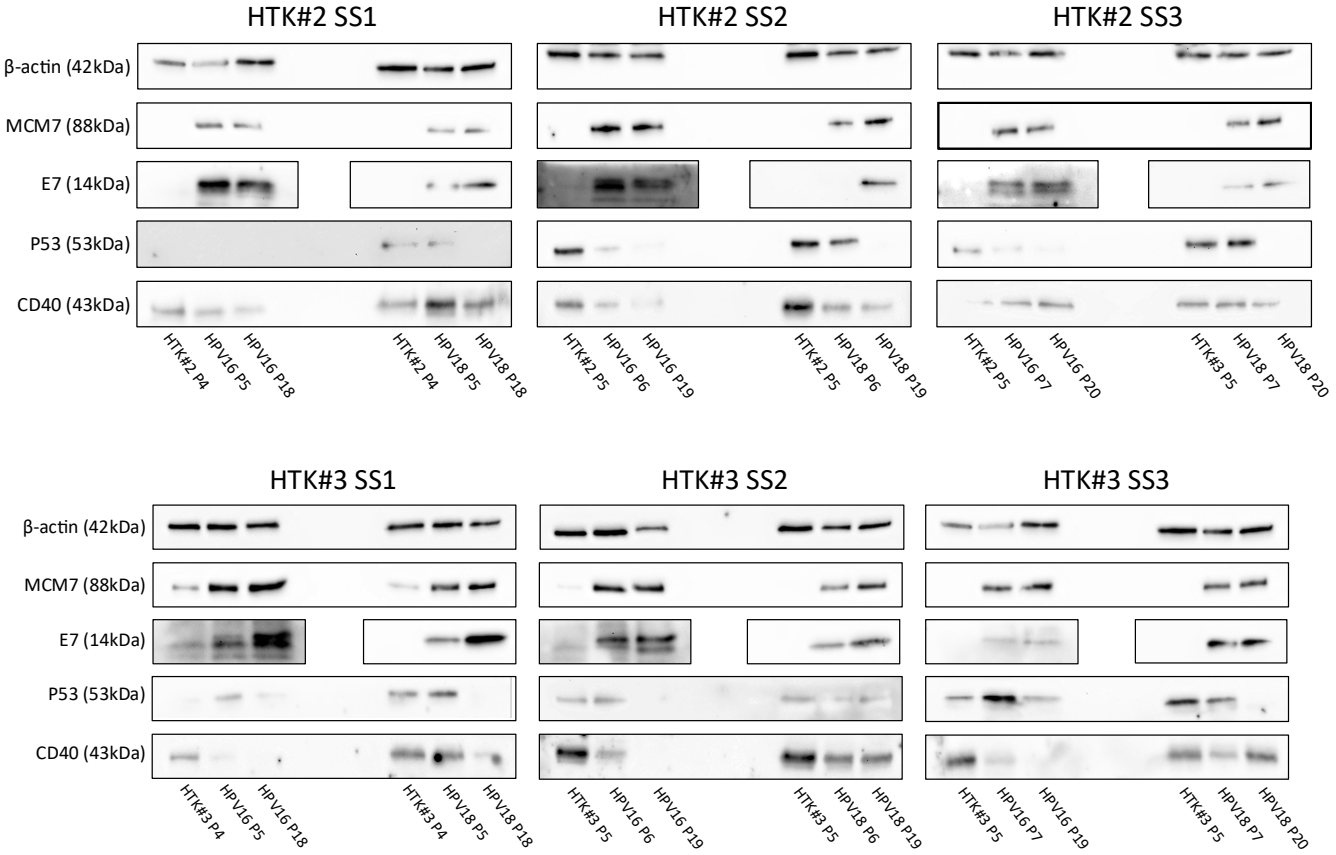


Figure S-3: Full western plot panel used for densitometry-based analyses comparing early vs late HPV16-HTK and HPV18-HTK cell lines

Western blots examined levels of β-actin, MCM7, HPV16-and HPV18-E7, p53, and CD40, as shown in Figure 4-9, Figure 4-10, and Figure 5-15. Primary and secondary antibodies used for these analyses are overviewed in Table 2-3. Passage number is shown as (P). The three sample sets (SS) consist of three independent replicates of sequentially passaged cell lines.