UNIVERSITY OF CAMBRIDGE

HUGHES HALL
UNIVERSITY OF CAMBRIDGE

Understanding and Targeting Clearance of Persistent HPV Infections

Ke Zheng

Department of Pathology

Hughes Hall

Supervisor: Prof. John Doorbar

Word count: 43,799

Date of submission: 30th September, 2022

Declaration:

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Abstract

Title: Understanding and Targeting Clearance of Persistent HPV Infections

Author: Ke Zheng

Given the great diversity of HPV types and tropism to epithelial sites (with associated diseases ranging from mild skin lesions to cancer), they share general genome organization and life strategy for long term infection after initial disease regression. Persistent infected cells form a viral reservoir, causing disease occurrence and are particularly problematic during carcinogenesis development for high-risk types; therefore, understanding HPV-mediated epithelial homeostasis is essential for understanding disease progression.

In this study, I applied the multiplex high-content confocal imaging system to reveal the spatial heterogenicity of keratinocytes with and without exogenous HPV E6 expression. Two *in vitro* model systems were developed to mimick the basal/suprabasal layers of the epithelium and evaluate HPV E6 functions in modulating cellular environment based on either the expression of a real-time cell cycle indicator (FUCCI) and early differentiation marker (KRT10); or the competition between differently fluorescently labelled infected versus uninfected cells.

Here, it is presented that HPV16E6 expressing keratinocytes exhibited increased proliferation, delayed contact inhibition and commitment to differentiation compared with normal keratinocytes. When the two populations are cultured together, HPV16E6 expressing cells have growth advantage over the adjacent uninfected neighbours in colonising the basal layer of epithelium.

Functional-deficient mutants of HPV16E6 revealed PDZ-domain containing proteins contributes to the competition advantage independent of the p53-regulated cell cycle progression. Further RNAi on keratinocytes of established HPV16E6 PDZ targets suggested MAGI and hPAR3 are two important modulators in maintaining basal cell life span and facilitates the formation of viral reservoir.

Similar phenotypes are also observed when keratinocytes expressing the E6 protein from different genera (α-HPV11/16/27, β-HPV8, γ-HPV65 and μ-HPV1), suggesting E6 is a conserved basal epithelial modulator which co-evolved with different epithelial sites for the establishment of local persistent HPV infections.

These results provide new insights of the function diverted biology of HPV during disease development, which may fuel the understanding of basal epithelial homeostasis and the development of successful anti-viral therapies.

Acknowledgement

Here, I would like to express my greatest gratitude to all who have helped me during this four-year period in Cambridge.

I would first like to thank my supervisor, Prof John Doorbar, for providing me this precious opportunity to study with him. With my sincere gratitude to his continuous support of my PhD study and research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

My exceptional appreciation goes to Dr Nagayasu Egawa, for his enthusiasm for science and day-to-day supervision and patience. I could not successfully conduct my research without him. I would like to acknowledge my funding resource, Maruho Co., Ltd, Japan, for supporting me to conduct this great project. It was also fantastic to work in a warm and friendly environment, appreciates all lab members' help to my experiments.

Thanks to all my lovely friends, for many memorable days and nights. With a special mention to Shengjun Ren, for always supporting me in the past four years and the next four million years.

Lastly, and most importantly, to my dearest parents, I love you!

Abbreviations

HPV	Human Papillomavirus
EGF(R)	Epidermal Growth Factor (Receptor)
HES	Hairy and Enhancer of Split
NICD	Notch Intracellular Domain
DAPT	N-[N-(3,5-Difluorophenacetyl)-L-alanyl]-S-phenylglycine t-butyl ester
FUCCI	Fluorescent Ubiquitination-based Cell Cycle Indicator
NIKS	Normal Immortalized Keratinocyte
HA	(Influenza) Hemagglutinin
MAML	Mastermind-like Protein
RBPJ	Signal Binding Protein for Immunoglobulin Kappa J Region
PBM	PDZ-Binding Motif

Table of Contents

1. Introduction

1.1 Diversity of HPV

More than 200 types of Human papillomavirus (HPV) have been identified in the last several decades and grouped into different genera based on genetic differences. The L1 capsid gene/open reading frame (ORF) is highly conserved in the HPV genome and thus has historically been used in the classification of Papillomavirus (PV) types (Harari *et al.* 2014). A novel PV is identified when the L1 ORF sequence is at least 10% different compared to existing PV types (de Villiers *et al.* 2004). HPVs are grouped into five genera (α, β, γ, μ and η), with tropism ranging from cutaneous to mucosal epithelia. Epithelial site-specific immune regulations are thought to have driven such genetic diversification and tissue tropism (Bordignon *et al.* 2017; Gottschling *et al.* 2011).

The largest genus is the α-papillomaviridae (α-PV), which comprises both high- and low-risk HPV types. Whereas a large proportion of HPVs is primarily associated with cutaneous infections, all high-risk α-HPV are primarily associated with mucosal epithelial infections (Figure 1.1; (Ostrow *et al.* 1989). As previously mentioned, HPV16 was the first high-risk HPV to be isolated and has the highest incidence rate in cervical cancer samples. Low-risk α-PV6/11 are responsible for the majority of genital wart cases and thus cause a significant economic burden to the general population (Forman *et al.* 2012). The HPV11 sequence was also used as the probe for the discovery of HPV16 (Durst *et al.* 1983). Therefore, due to the sequence similarity, it is of interest to determine the evolutionary mechanisms contributing to differences in disease pathogenesis between HPV11 and HPV16. Some α-PV also infect cutaneous epithelia and cause common skin warts, most frequently, HPV27 (Tomson *et al.* 2011). However, there are some (rare) cases of high-risk α-HPV types being isolated from cutaneous cancers.

Skin (plantar) warts were first documented in ancient Chinese medical literature more than two millennia ago (Guo 1999). The first HPV type, HPV1, was described in the 1960s (Klug &

Finch 1965). HPV1 is classified as μ-PV, a very small group that only comprises of 2 other known HPV types, HPV 63 and 204 (Šterbenc *et al.* 2017). Analysis of common warts revealed a unique clinical profile for HPV1 infection, largely related to the plantar site of infection in children under 12 years old (Bruggink *et al.* 2012; EGAWA *et al.* 1993; Orth *et al.* 1977).

From clinical genetic typing studies, γ-HPV 4/65 are primarily associated with cutaneous lesions (Hagiwara *et al.* 2005). Although it is not common, an association between persistent γ-HPV4/65 infection in epidermodysplasia verruciformis (EV) patients has been observed (Kawai *et al.* 2009). For a comprehensive analysis, we picked HPV65 as a representative for γ-HPVs, due to the unique phenotype of HPV65-induced pigmentation in skin warts (Egawa *et al.* 1993).

β -PV types are not typically found in common warts; however, their high incidence rates in patients with EV are concerning. Combined, EV and β-PV infection increase the patient's risk of non-melanoma skin cancer, especially with sun exposure (Orth 2006). Amongst the β-PVs, HPV8 infection is the most closely related to EV patients (Claudy *et al.* 1982).

By contrast, the η-PV genus only consists of one HPV type (HPV41), the clinical significance of which is yet to be discovered. Within the papillomaviridae, η-HPV is more closely related to bat-infecting PVs than human-PVs concerning sequence and phylogeny (Bravo & Félez-Sánchez 2015). Thus, η-PV is thought to be the result of a spontaneous host-shift event from bats to humans, perhaps explaining the absence of human pathogenicity.

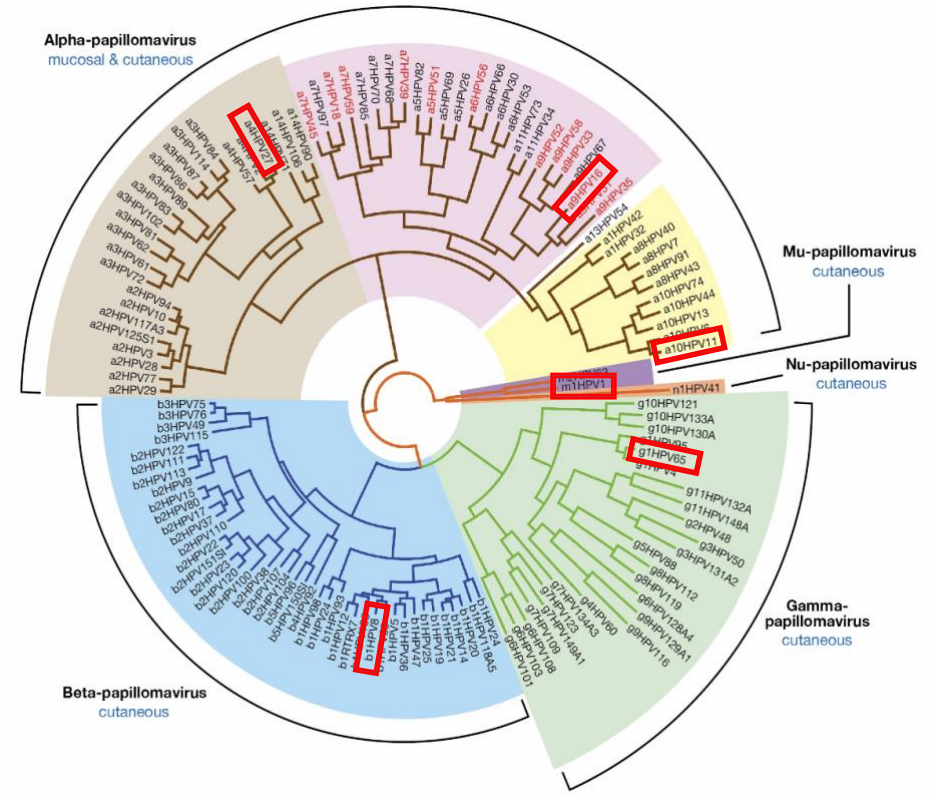


Figure 1.1: Evolutionary relationship between HPV genera. Phylogenetics is determined by E6/E7 sequence (modified from (Doorbar *et al.* 2012), with the HPV types tested in this report being highlighted in red boxes. For a comprehensive analysis, we picked clinically important representatives from each HPV group.

1.2 General HPV Life cycle

The successful HPV infection is thought to require virion particle access to actively proliferating cells at the basal epithelial layer. This is typically thought to be the result of a micro-wound or occurs at natural vulnerable sites such as hair follicles or the transformation zone of the cervix (Doorbar 2005). Thus, the wound healing response has been suggested to promote the initiation and propagation of epithelial cell infection (Doorbar 2006; Schiller *et al.* 2010). Though these early events are requisite for HPV infection, it is actually the persistence of the HPV genome in actively proliferating keratinocytes that increases the risk of neoplasia (Doorbar *et al.* 2012).

Upon binding the basal epithelial lamina, HPV particles have been shown to be endocytosed as a result of L1 and L2 viral capsid protein interaction with host membrane factors. Mock cellular infectivity assays using self-assembled viral-like particles with L1/L2 proteins from low- and high-risk HPV types indicate that HPV can infect a variety of epithelial cells (Müller *et al.* 1995). This led to the characterization of the host HPV receptor as a widely expressed and highly conserved cell surface ligand (Letian & Tianyu 2010). The putative host HPV receptor candidates include heparan sulphate proteoglycans (Combita *et al.* 2001), an evolutionarily conserved protein family (Sarrazin *et al.* 2011). This supports the theory that PV is an ancient virus that has co-evolved with different host organisms. Thus, elucidating the PV life cycle may reveal fundamental aspects of epithelial homeostasis, conserved throughout the animal kingdom.

The productive PV life cycle is dependent on the terminal differentiation scheme of the host epidermal tissue. Viral gene expression is co-regulated with keratinocyte differentiation programmes (Figure 1.2). In the basal layer, E2 is necessary for HPV genome replication and maintenance (Murakami *et al.* 2019). The viral oncoproteins E6 and E7 are also expressed in the basal layer, which is thought to delay differentiation in infected cells (Kranjec *et al.* 2017; White 2019). From RNAscope analysis, infected basal cells that are committed to differentiation have transient downregulation of viral E6 and E7 expression (unpublished material from Professor John Doorbar's group). In the upper epithelial layers, a reduction in E6 and E7 expression allows infected cells to respond to differentiation activating signals from adjacent cells, for example, via the Notch pathway (Kranjec *et al.* 2017; Macara *et al.* 2014a).

As groups, HPVs have evolved to infect and thrive at many different body sites by adapting to local epithelial conditions, which vary not only between cutaneous and mucosal epithelia but also amongst epithelial appendages such as hair follicles and sweat glands, with each of these sites providing distinctive epithelial niches and mechanisms of site-specific immune regulation (Bordignon *et al.* 2017; Gottschling *et al.* 2011).

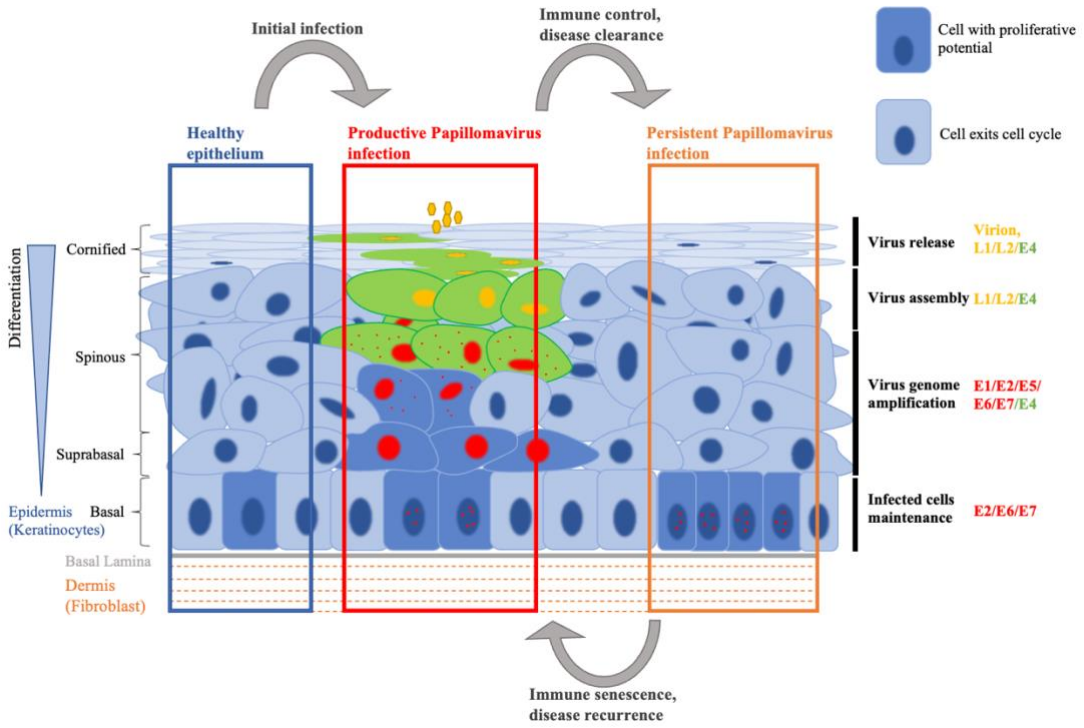


Figure 1.2 Human Papillomavirus Infection of Stratified Epithelium (Zheng *et al.* 2022)**.** In healthy epithelial tissue, a proportion of basal keratinocytes retain the potential to proliferate in order to maintain the multi-layered epithelial structure. Following HPV infection, viral gene expression can alter the normal differentiation program of the cell to facilitate the completion of the productive virus life cycle and is tightly regulated in the different epithelial layers. Viral gene expression may, in due course, trigger a successful immune response, leading to disease regression. It is thought that persistent infection with a limited expression of the viral genome can occur in the following regression. It has been suggested that changes in the efficiency of immune surveillance may allow the subsequent reactivation of disease.

1.3 Genome organization

The papillomavirus is a small double-stranded DNA virus of approximately 8kb in size. Genome arrangement is similar among HPV types (Figure 1.3), encoding four core proteins for viral capsid assembly (L1 and L2) and genome replication (E1 and E2). Other proteins (E4, E5, E6, E7 and E8) are not expressed by all PV types; therefore, they are considered ancillary proteins related to evolutionary outcomes. These proteins are important for epithelial tropisms, in

alteration of infected cells for progeny virus production and immune evasion (Egawa *et al.* 2015). In addition, fusion proteins E1^E4 and E8^E2 are encoded by spliced transcripts and modulate viral replication and transcription. These accessory proteins are important for the virus's productive life cycle and for maintaining persistent infection (Dreer *et al.* 2017; Wang *et al.* 2004).

Upon infection, the virus enters the cell through the endocytic pathway and is trafficked towards the nucleus. L1 is cleaved during this process (Day *et al.* 2013). During mitosis, the nuclear envelope breaks, and the virus uses the L2 protein to bind with the cellular condensed chromosomes (DiGiuseppe *et al.* 2016). The viral genome is equally segregated into daughter cells (Graham 2017b). The E1 and E2 proteins bind and unwind the replication origin and recruit cellular factors to synthesize viral genomes (McBride 2008). During virus persistence in the basal layer, it has been suggested that the E2 protein alone is sufficient for viral genome maintenance (Murakami *et al.* 2019). Expression of E4 is mainly restricted to the upper layer of epithelia, whose function is to remodel host keratin filament organization for HPV progeny release (Wang *et al.* 2004).

E5, E6 and E7 are classified as oncoproteins from high-risk HPV types. The E5 protein ORF is absent in β, μ and γ-PV, indicating that E5 is not necessary for the PV life cycle of these viruses but giving additional pathogenicity to the α-PVs. Among α-PV, the E5 protein is further categorized into four different groups (α, β, γ and delta) based on clinical manifestations (Bravo & Alonso 2004). One of the important functions of E5 is the downregulation of MHC I expression on the infected cell surface and facilitates α-PV immune surveillance by reducing CD8+ T cell recognition (Campo *et al.* 2010). The downregulation of MHC I expression is seen in both low-risk HPV6 and high-risk HPV16 (Venuti *et al.* 2011).

Initially, E6 and E7 were discovered as putative oncogenes due to their ability to immortalize keratinocytes (Hawley-Nelson *et al.* 1989a). E6 and E7 are expressed by bicistronic mRNA from the early promoter p97 and controlled by splicing in high-risk HPVs (Tang *et al.* 2006). The function of E6 and E7 have some redundancy. Both E6 and E7 are thought to be expressed

in the suprabasal layer of epithelia, driving differentiating cells to S-phase re-entry for viral genome amplification (Doorbar *et al.* 2015). E6 and E7 expressions are required for HPV-induced cancers (Pal & Kundu 2020). The expression pattern of E6 and E7 are considered to be one of the most important factors in distinguishing between high and low-risk HPV types (Smotkin *et al.* 1989). One of the well-established features of E7 is the interaction with protein retinoblastoma (pRb). The Rb protein was identified as the first tumour suppressor capable of inducing cell cycle arrest at the G1 phase by inhibiting the activity of E2F transcriptional factors (Weinberg 1995). Upon HPV infections, viral genome amplification in cytoplasm triggers host cellular DNA damage response (DDR) (Sakakibara *et al.* 2013). Apart from E1, E7 can also induce DDR (Moody & Laimins 2009), and the usual outcome of DDR is cell cycle arrest and senescence. Therefore, E7 promotes cell cycle entry into S-phase by targeting pRb protein for degradation via the proteasome pathway (Boyer *et al.* 1996) and is important for viral genome amplification in differentiated cells. The function of E7 in driving S-phase re-entry is important for the productive life cycle to ensure that viral genome amplification is happening in the differentiating keratinocyte. Previous work from our lab suggests that the E7 function in remodelling differentiated cell S-phase re-entry may play a limited role in driving cell proliferation in the basal layer (Kranjec *et al.* 2017). However, low-risk E7 binds with the Rb protein with much lower affinity, and whether the response to DDR is necessary during the low-risk HPV productive life cycle remains unclear (Moody 2017a).

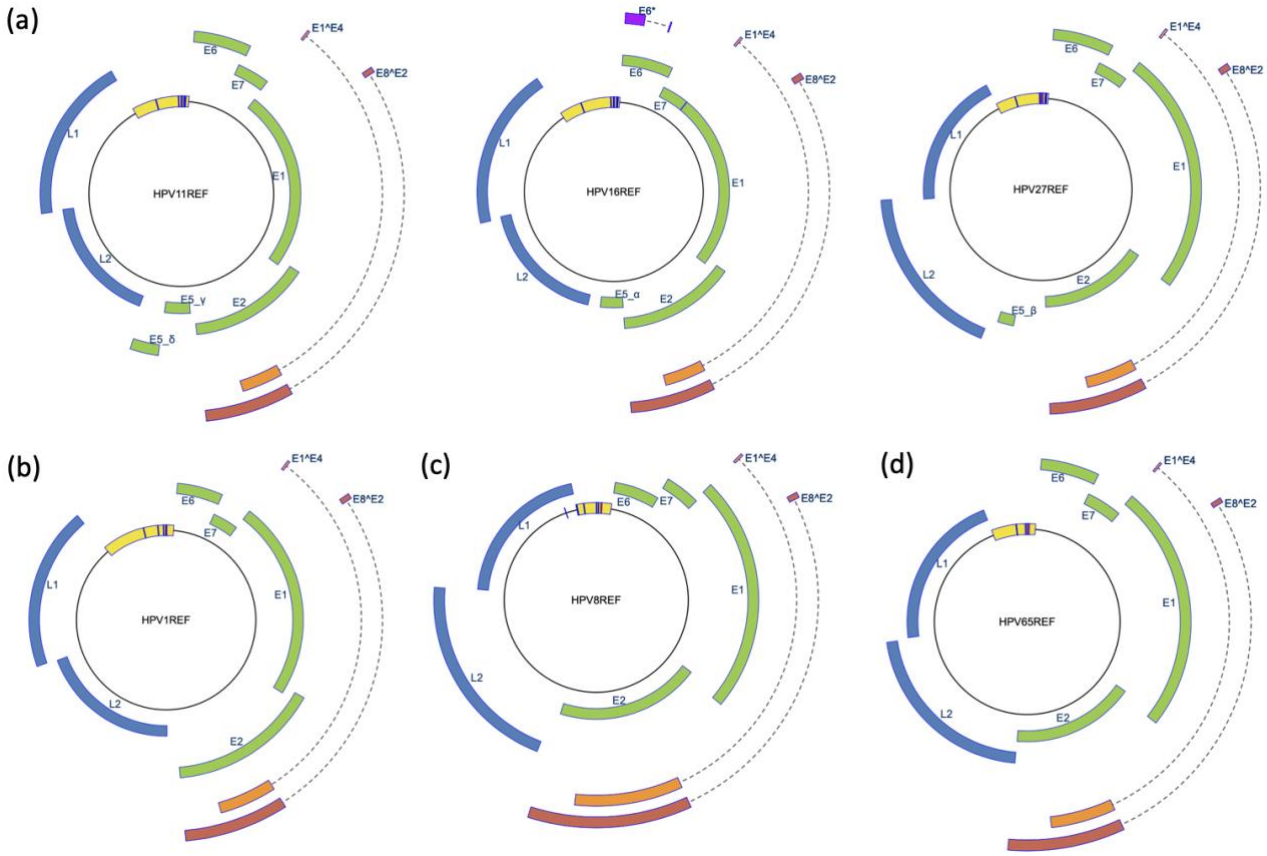


Figure 1.3: Similar genome arrangements of different HPV types (Papillomavirus Episteme PaVE). (a) α-PV, including low-risk mucosal HPV11, high-risk mucosal HPV16 and low-risk cutaneous HPV27. (b) μ-PV, low-risk cutaneous HPV1. (c) β-PV, low-risk HPV8. (d) γ-PV, low-risk cutaneous HPV65.

A mass-spectrometry based proteomic analysis has identified a greater diversity of host proteins interacting with HPV E6 compared to E7, and also proposed a distinctive interaction pattern between different HPV genera (White *et al.* 2012). The most well-studied role of high-risk E6 in driving carcinogenesis is the degradation of p53. Tumour suppressor p53 typically functions as a specific DNA binding protein regulating transcription of the genes involved in cell cycle arrest, apoptosis and senescence (Sherr 2004). The E6 protein, together with a ubiquitin-protein ligase UBE3A, also known as E6 associated protein (E6AP), targets p53 for proteasomal degradation (Martinez-Zapien *et al.* 2016). Although low-risk α-PV HPV11E6 has been found to be associated with E6AP and p53, downregulation of p53 by 11E6

remains controversial (Brimer *et al.* 2007; White *et al.* 2012). The major differences between low and high-risk HPV E6 proteins are the ability to degrade p53 and the additional PDZ-binding motif encoded by high-risk HPV E6s. Mutagenesis of E6AP revealed the interaction with HPV E6 is through the LXXLL sequence (Elston *et al.* 1998; Zanier *et al.* 2013b), which offers an anti-viral target for the high-risk HPV types. However, low-risk HPV11E6 was found to repress p53 transactivation and p53-dependent transcription in an E6AP-independent manner through inhibition of histone acetylation (Thomas & Chiang 2005). Another protein, p300, forms a complex with p53 for transcriptional activation of downstream target genes. Interaction between β-PV E6 and p300 was discovered by White *et al.*, and cutaneous HPV E6 can also inhibit p53-dependent transcriptional activation by promoting p300 degradation (Howie *et al.* 2011). Both α- and β-PV can alter the cell cycle by affecting p53-dependent gene activation, while the p53 degradation function is an evolved phenotype inherited from a most recent common ancestor of the high-risk species that do not always separate with carcinogenicity (Fu *et al.* 2010).

Induction of p53 by cellular stress triggers multiple cellular programs, including DNA damage repair, cell cycle arrest and apoptosis. In epithelial tissue, expression of p53 is mainly restricted to the basal layer (Botchkarev & Flores 2014), and it has been shown that sustained p53 induction but not pulse induction (signal with same amplitude but shorter period) can upregulate the expression of genes associated with cellular terminal fates (Purvis *et al.* 2012). Although it is not widely accepted, it has been proposed that viral genome maintenance may cause sustained cellular stress as a result of the recognition of viral DNA in the cytoplasm as a DNA damage event (Graham 2017c; Stanley 2012). Altogether, HPV requires strategies that antagonize the host cell response in driving the cell towards terminal differentiation, and HPV modulation of basal cell fate is proposed to be largely dependent on the E6 protein.

1.4 The burden of HPV-associated diseases

1.4.1 Cervical cancer

Cervical cancer is the fourth most common cancer to affect women worldwide, accounting for 6.6% of all female cancers ("WHO | Cervical cancer" 2018), with nearly all of these being due to persistent HPV infections (NHS 2008). The mortality rate of late-stage cervical cancers is high, around 50%, with the majority of fatalities being reported in low- and middle-income countries (Forman *et al.* 2012).

The correlation between HPV infection and cervical cancer was first hypothesized in the 1970s, with the HPV16 genome being discovered in cell lines established from cervical cancer biopsies (Durst *et al.* 1983; zur Hausen 1977). To date, more than 12 HPV types have been isolated from clinical samples and found to be associated with cervical cancer development, namely HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59 (Serrano *et al.* 2018). These carcinogenic types are classified as high-risk HPVs, with HPV16 and 18 being responsible for 70% of all cervical cancers.

1.4.2 Other diseases

Persistent HPV infection is also implicated in other cancers. Although these cancers are rare in the general population, over 50% of vaginal, vulvar, penile and oropharynx cancers are also associated with infection by high-risk HPV types (Chaturvedi 2010), and HPV now accounts for nearly 5% of all cancers worldwide ("HPV and Cancer - National Cancer Institute" 2019).

By contrast, low-risk HPV types are typically non-carcinogenic and are usually associated with benign hyper-proliferative lesions at cutaneous and mucosal infection sites in the general population. For example, HPV 6 and 11 cause 90% of genital warts as well as recurrent respiratory papillomatosis (RRP) in immunocompetent humans (Carifi *et al.* 2015). Also, cutaneous low-risk HPV infection and severe disease progression have been observed in

immunocompromised patients, such as those individuals suffering from Epidermodysplasia Verruciformis (EV) (Egawa & Doorbar 2017; Howley & Pfister 2015; Kawai *et al.* 2009).

1.4.3 Prevention of HPV-related diseases

Vaccination for HPV-mediated disease has been ongoing for over ten years. The vaccines have evolved over this period to cover 9 HPV types (Gardisil 9). The coverage of multiple HPV types means that nearly 90% of cervical cancers can now be prevented, while coverage of the low-risk HPV types (6 & 11) means the morbidity associated with benign diseases such as genital warts can be markedly decreased (Cheng *et al.* 2020).

A significant decline has been reported in the number of both low-and high-grade cervical disease in vaccinated cohorts (Pollock *et al.* 2014). Unfortunately, this success has not been emulated across the globe and worryingly, in some countries (e.g. Japan), the incidence and mortality of cervical cancer are rising (Simms *et al.* 2020).

Vaccine hesitancy, as seen with the MMR vaccine (vaccine against measles, mumps, and rubella), has had a significant impact on HPV immunization. While cases of Postural tachycardia syndrome (PoTS) have been dubiously associated with the vaccine, leading to a fall in vaccine uptake (Butts *et al.* 2017). Indeed, in a country like Japan where vaccine uptake was >70%, uptake has dropped to <1% due to these unfounded risks (Kunitoki *et al.* 2021). Low uptake rates are also seen in other High-Income countries such as the USA (<50%) (Orenstein *et al.* 2016). While education of the community is key to improving uptake, stemming the tide of false information is also a key aspect of improving vaccination rates.

Low and Middle-income countries also suffer from low vaccine uptake rates (Dorji *et al.* 2021; LaMontagne *et al.* 2017). However, here such low rates stem from either an inability to afford mass vaccination or a lack of infrastructure in order to provide for a school-based immunization programme. This means that the vaccine is not utilized in countries where they

are needed the most, and this problem is further compounded by the fact that such countries lack defined screening programmes, thus leading to a higher disease burden.

1.4.4 Treatments of HPV-related diseases

While the evidence of primary prevention of HPV with vaccination is clear, it is also clear that widespread uptake and complete coverage of disease-causing HPV types will take a few decades (Spayne & Hesketh 2021). This has necessitated the continuation of screening programmes for cervical cancer, which is now done by primary HPV nucleic acid testing. While screening programmes have been successful in reducing deaths (Jansen *et al.* 2020) due to cervical cancer, the excisional treatment utilized to treat high-grade cervical disease has significant consequences. In particular, as the depth of excision increases, the risk of cervical incompetence and, therefore, the risk of pre-term labour also increases (Castanon *et al.* 2014). This risk is further compounded by the fact that the majority of women that require excisional treatment are of childbearing age. Indeed less traumatic modes of treatment have been trialled, such as cryotherapy or ablation. However, these have not proved effective due to their inability to guarantee the depth of treatment and therefore complete ablation of a lesion (Basu *et al.* 2018).

Excisional treatment is also the mainstay of treatment for oropharyngeal and vulval tumours. In the case of vulval intra-epithelial neoplasia (VIN), treatment with laser ablation or immune modulators such as imiquimod or 5-FU has been shown to improve the clearance of such lesions; however, in a significant number of cases, excisional treatment is still required leading to unpleasant scarring and the requirement of a general anaesthetic (Kaushik *et al.* 2011). Benign diseases caused by HPV, including genital warts and respiratory papillomas, are treated either conservatively or medically with podophyllotoxin (for genital warts) or with excisional therapy (Goon *et al.* 2007; Patel *et al.* 2012; Silverberg *et al.* 2003). In the case of respiratory papillomas, the mainstay of treatment is excisional, depending on symptoms. This involves invasive endoscopy in order to visualize and remove the papillomas. Unfortunately,

excision in these cases may not be curative, with a significant risk of recurrence necessitating further procedures (Goon *et al.* 2007).

The incidence of new cases of genital warts per year in the UK is 0.16% (Woodhall *et al.* 2008), and the total cost to the healthcare system is £58.44 million/year (Coles *et al.* 2016). Whether caused by high- or low-risk HPV, the fact that these infections are common and persistent suggests that a cost-effective treatment modality would attract significant attention from the clinical community. Indeed, as our understanding of the molecular biology of HPV infections at different sites has improved, there is now scope to translate this knowledge into identifying effective antiviral molecules, be it new or repurposing of older established drugs.

1.5 Current efforts toward the development of anti-HPV agents

1.5.1 Viral Genome replication

HPV replication has been proposed to occur in at least two different phases; an amplification phase in the upper layers of the epithelium (where newly produced viral genomes are packaged into viral particles) and a maintenance phase in the basal layers of proliferating cells (where viral DNA is maintained at low copy number). The maintenance phase continues throughout the latency period, during which the virus is present at 50-100 copies per cell (Maglennon *et al.* 2011a). Viral proteins E1 and E2 are both required for viral DNA replication, while only E1 exhibits the DNA helicase and associated ATPase activity required for unwinding viral DNA during replication (Bergvall *et al.* 2013). E2 is a sequence-specific DNA binding protein that can simultaneously bind to sites in the origin of replication and to the E1 helicase (McBride 2013). The most successfully exploited antiviral targets have been viral enzymes, specifically the viral nucleotide kinases and polymerases (for example, acyclovir). During productive infection, the papillomaviruses replicate their viral DNA as episomes physically separate from the host cell genome (McBride 2017). DNA polymerase and most other associated replication factors are supplied by the host cell. HPV6 E1 helicase

inhibitors have been proposed by high throughput screening of the Boehringer Ingelheim compound collection, using an ATPase assay based on a novel scintillation proximity detection methodology, but a single non-conserved mutation in the NTP-binding site in HPV11 resulted in no inhibitory activity (Bergvall *et al.* 2013; Faucher *et al.* 2004). Despite these inhibitors being active in vitro, they were not effective in cell culture assays, thus limiting their potential as anti-HPV lead compounds.

Due to a lack of proteins that resemble traditional targets for medicinal chemistry in HPV, scientists have focused on the inhibition of the interaction between E1 and E2 to initiate viral genome amplification. This is also an attractive treatment for HPV proliferative lesions, such as condylomas. An inhibitor that reversibly bind to the transactivation domain of E2 and prevent interaction with E1 has been identified using a scintillation proximity assay (SPA) that measures the cooperative binding of recombinant HPV11 E1 and E2 to radiolabelled origin DNA (White *et al.* 2003). Unfortunately, these indandione inhibitors were inactive against high-risk HPV16 and 18. One of the compounds was capable of antagonizing E1-E2 interaction and inhibiting reporter gene expression in HPV11 transiently transfected cells but a significant reduction in compound efficacy has been observed in vivo and in vitro (White *et al.* 2003).

Another idea is to prevent the binding of the E1/E2 complex to the viral genome, the origin of replication (ori). Pyrrole–imidazole polyamides bind to the minor groove of DNA in an AT-rich region (He *et al.* 2014) to prevent the DNA bend that is induced by E2 and needed for stabilizing its association with its target site (Schaal *et al.* 2003). Bromodeoxyuridine (BrdU) incorporation demonstrates that DNA synthesis in the suprabasal layer of organotypic cultures is suppressed upon compound treatment with no evidence of cytotoxicity, correlating with a loss of HPV16 and HPV18 episomes (Edwards *et al.* 2011). Computational modelling has predicted some peptides that may potentially inhibit E1-E2 complex formation with higher affinity and efficacy (Kantang *et al.* 2016). Up to date, antiviral development targeting E1-E2 dependent viral genome amplification has been halted by the complexity of cellular machinery in vivo, but these issues might be solved in the future with new technology with more specific drug design and delivery methods. Recently, evidence has been published

suggesting E2 solely is sufficient for viral genome maintenance at basal epithelium (Murakami *et al.* 2019), suggesting targeting E2 protein interaction with viral genome might be a better solution to treat HPV-induced lesions as those latent basal cells will also respond to treatment. In addition to being necessary for replication, E2 also has essential roles in regulating the transcription of the viral genes (E6 and E7). Therefore, there is a potential carcinogenic risk that targeting E2 may select for those infected cells that have genome integrated (E2 expression was disrupted and E6/E7 overexpression).

1.5.2 Viral pathogenesis

The combined actions of the high-risk E6 and E7 oncoproteins are essential for the maintenance of the neoplastic phenotype and the evasion of apoptosis (Moody & Laimins 2010). Abrogation of either E6 or E7 function (or both) in neoplastic cells by targeting gene expression or protein-protein interactions should be an effective strategy. The most direct approach to targeting E6/E7 transcripts is by small interfering RNA (siRNA). Accumulation of cellular p53 protein level and successful inhibition of HPV-positive cancer cell growth has been observed both in vivo and in vitro (Fujii *et al.* 2006; Jonson *et al.* 2008; Niu *et al.* 2006; Yamato *et al.* 2008).

Research on small molecules against E7 protein function has not achieved promising progress so far. One study by Finzer et al. suggested a Histone deacetylase (HDAC) inhibitor, sodium butyrate, induces an intrinsic type of apoptosis in human papillomavirus (HPV)-positive cells by disrupting the mitochondrial transmembrane potential and restoration of pRb-mediated cell senescence, which is only observed in E7-expressing cells (Finzer *et al.* 2004). Notably, although HDAC inhibitors can efficiently revert both dysregulated cell proliferation and suppression of apoptosis signal in HPV-positive cancer cells, alteration of E6/E7 expression levels is not essential for the function (De La Cruz-Hernández *et al.* 2007).

Several studies have tested the inhibition of the E6-E6AP interaction through different molecules, which is essential for the E6-mediated degradation of host cellular proteins

through the LXXLL motif. The LXXLL containing α-helix binds in a hydrophobic pocket of E6 (Zanier *et al.* 2013a), therefore offering significant opportunities for antiviral drug targeting. Some early studies proposed minimal aqueous stable single α-helical peptides to disturb E6-E6AP interaction (Liu *et al.* 2004; Zanier *et al.* 2014), but they are comparatively low in binding affinity. A chimeric protein that contains the LXXLL motif and the PDZ motif displays a much higher binding affinity when compared to the LXXLL motif alone (Karlsson *et al.* 2015; Ramirez *et al.* 2015). Therefore, more recent studies have changed drug discovery direction towards small molecules to inhibit the E6-E6AP association. Despite the large overall surface area for the hydrophobic groove, flavonoid-derived compounds were identified as E6 inhibitory compounds (Cherry *et al.* 2013) and modification of charged groups on small molecules has increased inhibitory activity (Rietz *et al.* 2016). Quenched molecular dynamics simulations with HPV E6 have identified a pharmacophore comprised of three general hydrophobic centres, two hydrophilic centres, and one exclusion sphere (Baleja *et al.* 2006). Thus, a few other small molecules have been proposed to disrupt E6-E6AP interaction and rescue cellular p53 levels using a high throughput screening method (Baleja *et al.* 2006; Malecka *et al.* 2014).

Since HPV infections are constrained to local areas, intracellular antibody (intrabody) can be a good candidate to target viral proteins, which normally have problems with efficient delivery. Intrabodies against high-risk HPV E6 have shown an inhibitory effect with E6AP association *in vitro* and effectively prevent the degradation of p53 in HPV-positive cervical cancer cell lines (Griffin *et al.* 2006).

Proteasome inhibitors act by inhibiting the ubiquitin-proteasome-dependent protein degradation. Since HPV E6 targets p53 for degradation through the ubiquitin-proteasome system, proteasome inhibitors are good candidates to rescue wild-type p53 expression in cervical carcinoma cells. Two proteasome inhibitors, MG132 and PS341, have been tested in SiHa (HPV16 positive) and HeLa (HPV18 positive) cells and increased p53 protein level and transcriptional activity has been reported with increased sensitivity to radiotherapy for solid tumours (Hougardy *et al.* 2006; Lin *et al.* 2009).

Repurposing HIV protease inhibitors to treat HPV-induced cancers has been successful *in vitro* studies, in which the introduction of high-dose PIs into E6-transfected C33A cells (HPV-

negative cervical cancer) results in an increase in cellular p53 level (Hampson *et al.* 2006). Inhibition of the HIV protease results in the abrogation of polyprotein processing, which leads to the accumulation of non-infectious, morphologically immature HIV virions (Ghosh *et al.* 2016). The mechanism of HIV PIs in HPV E6 inhibition is different. Originally it was considered to be the selective inhibitory effect on proteasomes and prevent E6-E6AP mediated p53 degradation via the proteasome pathway. A chemical mapping study suggested that the main changes in E6 expressing cells treated with HIV PIs happen in the nucleus and might have other undiscovered functions (Kim *et al.* 2010). However, using HIV-PI for HPV-related diseases has been challenged since there is no clear benefit of antiretroviral treatment in decreasing HPV-related cervical disease in HIV-infected women from clinical studies (H. Adler 2010). This clinical insignificance might be due to the requirement of high-dose PI topically applied to HPV-related cancer for effective antiviral treatment.

1.5.3 Viral entry

Prevention of virus entry by microbicides or viricides is an alternative approach to targeting viral proteins after exposure. HPV binds initially to heparin sulphate proteoglycans (HSPGs) in the basal membrane (BM) via the L1 protein before attaching to and entering the keratinocyte (Kines *et al.* 2009). Carrageenan, a sulphated polysaccharide extracted from red seaweed, is a potent inhibitor of the attachment of the virus particle to HSPGs on the basal membrane and the cell surface. In an experimental cervicovaginal challenge model in the mouse, carrageenan was highly effective at preventing HPV transmission to cervical epithelial cells (Roberts *et al.* 2007).

Viral protein	Function in the viral life cycle	Proposed drug and mechanism
E1	DNA binding protein, helicase, ATPase, viral genome amplification	• Low-risk E1 helicase inhibitors (Faucher *et al.* 2004) • Inhibitors of tyrosyl-DNA phosphodiesterase one and PARP1 (Toots *et al.* 2017)

E2	Transcriptional regulation of early genes, initiation of DNA replication and partitioning of the viral genome	• Indandione inhibitors reversibly bind to the transactivation domain of E2, inhibiting E1/E2 interaction (White *et al.* 2003; Yoakim *et al.* 2003) • Pyrrole–imidazole polyamides bind to viral DNA and prevent E1-E2 initiation of replication, with potential modification by computational methods (Edwards *et al.* 2011; He *et al.* 2014; Kantang *et al.* 2016; Schaal *et al.* 2003)
E4	Accessory protein helping genome amplification, and virus synthesis, might have an additional role during virus disassembly and transmission	
E5	Only encoded by α-HPVs, transforming activity *in vitro,* role in viral immunosurveillance	

E6/E7	Oncogenes modify cell cycle status for viral genome amplification in the differentiated epithelium, role in viral persistence	• Small interfering RNA (siRNA) downregulate E6/E7 mRNA levels and induces cancer apoptosis (Fujii *et al.* 2006; Jonson *et al.* 2008; Niu *et al.* 2006; Yamato *et al.* 2008; Yoshinouchi *et al.* 2003) • Antisense deoxynucleotides (Yamayoshi *et al.* 2007) • Ribozymes (Zheng *et al.* 2004) • Intrabodies target E6 protein (Doorbar & Griffin 2007; Griffin *et al.* 2006) • Flavonoid-derived compounds disrupt E6-E6AP binding (Cherry *et al.* 2013) • Other small molecules inhibit E6 and cellular protein interaction (Baleja *et al.* 2006; Malecka *et al.* 2014) • Proteasome inhibitors prevent p53 degradation by E6(Hougardy *et al.* 2006; Lin *et al.* 2009) • HIV protease inhibitors prevent E6-mediated p53 degradation in HPV-positive cells (Batman *et al.* 2010; Hampson *et al.* 2006; Kim *et al.* 2010) • Histone deacetylase (HDAC) inhibitors revert E7 oncogenic functions (Finzer *et al.* 2004) • MEK/ERK inhibitors downregulate oncogenic E6/E7 transcription and protein expression (Luna *et al.* 2021)
L1/L2	Structural capsid proteins, viral entry	• Inhibitor prevents L1 and host membrane receptor HSPG interaction (Roberts *et al.* 2007) • Protamine sulfate inhibits attachment and cell entry of genital α-HPV types (Young *et al.* 2021)

Table 1.5: Summary of current HPV antiviral research (Zheng *et al.* 2022).

While the potential of antiviral HPV drugs has been realized for decades (summarized in Table 1.5), no specific anti-HPV drug has yet made it into routine clinical use (Fradet-Turcotte & Archambault 2007; Phelps & Alexander 1995). As well as there being few direct viral targets, any evaluation of treatment efficacy must take into account the possibility of spontaneous regression. HPV has developed sophisticated interactions with host cellular, and molecular mechanisms and requires a fully differentiated epithelium to complete its productive viral life cycle (Figure 1), which introduces an added layer of complexity for *in vitro* research studies. Reconstituting the productive HPV life cycle *in vitro* is a costly and time-consuming process, and as an alternative, there are only a few appropriate animal model systems – the best of which only approximate the important PV disease associations that occur in humans. As a result, no single drug treatment has yet been shown to have a high success rate in eliminating HPV infections.

1.6 Persistent HPV infections and E6 protein

Persistent infection is characterized as the virus is not cleared but remains in specific cells of infected individuals. The mechanisms by which persistent infections are maintained involve both modulation of virus and cellular gene expression and modification of the host immune response. Reactivation of persistent infection may be triggered by various stimuli, and host immunosuppression is one of the key factors for viral reactivation. Persistent high-risk HPV infection has been shown to constitute the most critical risk factor for the development of cervical precancers cervical intraepithelial neoplasia (CIN) grade 2 and 3 (Doorbar *et al.* 2021; Koshiol *et al.* 2008). Although the concept of a clear viral reservoir has been proposed for a long time, no measures to eradicate persistent viruses have been developed; and specifically for HPV, there is no antiviral drug available. Vaccination and immunotherapy can reduce the frequency of clinical recurrence and lighten disease symptoms, yet the virus continues to remain associated with the host. To achieve effective HPV clearance, it is important to identify the infected reservoir and specifically target these cells for terminal differentiation. Based on

previous research (Brimer *et al.* 2012; Kranjec *et al.* 2017) and unpublished data from our lab, we hypothesized that there is an evolutionarily conserved role for the E6 protein during HPV persistent infection at the basal layer of epithelia. Ultimately, by further understanding of the HPV reservoir population and epithelial homeostasis, therapeutic targets that can widely target the HPV persistence for effective clearance will be available.

1.7 Basal epithelial homeostasis

For epithelial homeostasis, attachment to the basal lamina, contact inhibition of proliferation, exiting basal layer (including cell division directionalities and apical extrusion) and commitment to differentiation are the main processes proposed for multi-layered epithelial development and maintenance.

1.7.1 Proliferation (contact inhibition) *in vitro*

The contact inhibition classically regulates proliferation in multicellular colonies, a process by which cultured cells stop dividing when they become confluent and cover the full 2D space (Aragona *et al.* 2013). In monolayer tissue culture, contact inhibition of proliferation has been demonstrated by a plateau in the growth curve as cells reach confluency (Eagle & Levine 1967). This behaviour is representative of proliferative arrest in most epithelial tissues, typically leading to terminal cell differentiation or death. Thus, contact inhibition is critical in determining tissue architecture. However, for immortalized cells, cell growth is not limited to such monolayer boundaries (Eagle & Levi 1967). Thus, loss of contact inhibition is considered a hallmark of cancer (Hanahan & Weinberg 2000, 2011).

Despite the fact that HPV types co-evolved over time to adapt to different niches, it has been discovered that similar pathways, but different proteins, are targeted by the viral E6s. The α-HPV E6 protein, including both low- and high-risk types, targets the p53 pathway regulating transcription of the genes involved in the cell cycle arrest and differentiation. While the high-

risk HPV16 E6 directly targets p53 for proteasomal degradation (Martinez-Zapien *et al.* 2016), low-risk HPV11 E6 also represses p53 transactivation and p53-dependent transcription by repressing p300-dependent acetylation (Thomas & Chiang 2005). In epithelial tissue, expression of p53 is mainly restricted to the basal layer (Botchkarev & Flores 2014), and sustained p53 induction, but not pulse induction, has been shown to upregulate the expression of genes associated with terminal cellular fates (Purvis *et al.* 2012).

With a heterogeneous population of basal epithelial cells, identifying individual molecular differences governing cell fate determination is essential. This is particularly important to the development of antiviral agents; by identifying the key viral contributor for infected cells to remain at the basal epithelium, antagonizing this viral function will lead us towards viral reservoir clearance (Zheng *et al.* 2022).

An early theory of apical extrusion of basal epithelial cells is dependent on cell division; daughter cells acquire different cellular fate at a fixed rate (Marques-Pereira & Leblond 1965), mutations in several pathways will alter the rate of mitosis and the possibility of cells acquiring terminal differentiation fate (Alcolea & Jones 2015; Lechler & Fuchs 2005; Matamoro-Vidal & Levayer 2019; Nagata & Igaki 2018; Williams *et al.* 2011). The orientation of epithelial cell division is an essential factor that defines cell morphology and tissue architecture. Mitosis can occur along three planes; division along the x- or y-axis of the epithelial layer will lengthen or widen the tissue respectively, whereas division along the z-axis will generate multiple sheets (Williams *et al.* 2011). Different mechanisms regulate these directionalities: division along the x/y-axis by planar cell polarity and the z-axis by apical-basal polarity (Macara *et al.* 2014b). However, there are limitations to this theory as reinsertion into the basal layer following apical budding has been observed/demonstrated (Packard *et al.* 2013).

1.7.2 Competition

Another mechanism of epithelial homeostasis is the extrusion of 'less fit' cells that ultimately undergo terminal differentiation and replenish the basal cell population (Ohsawa *et al.* 2018). This process has been observed in human colon epithelia (Eisenhoffer *et al.* 2012a), zebrafish epidermis (Eisenhoffer *et al.* 2012a) and Drosophila larvae epithelia (Lolo *et al.* 2012). The definition of basal cell fitness still needs investigation. In normal epithelia, faster-proliferating cells will be smaller in size than slow-cycling cells (Tsuboi *et al.* 2018). When the basal cell population is overcrowded, larger cells with greater mechanical deformation on the plasma membrane will adopt delamination as a cell fate. This process is governed by a membranous mechanosensitive ion channel (Eisenhoffer *et al.* 2012b). The cell is extruded towards the suprabasal layer by producing Sphingosine 1-Phosphate (S1P) to activate adjacent cell ROCK-dependent actin contraction (Figure 1.7.2) (Gudipaty & Rosenblatt 2017). Once started delamination from the basal lamina, the Notch pathway is activated, and the keratinocytes trigger the irreversible terminal differentiation steps.

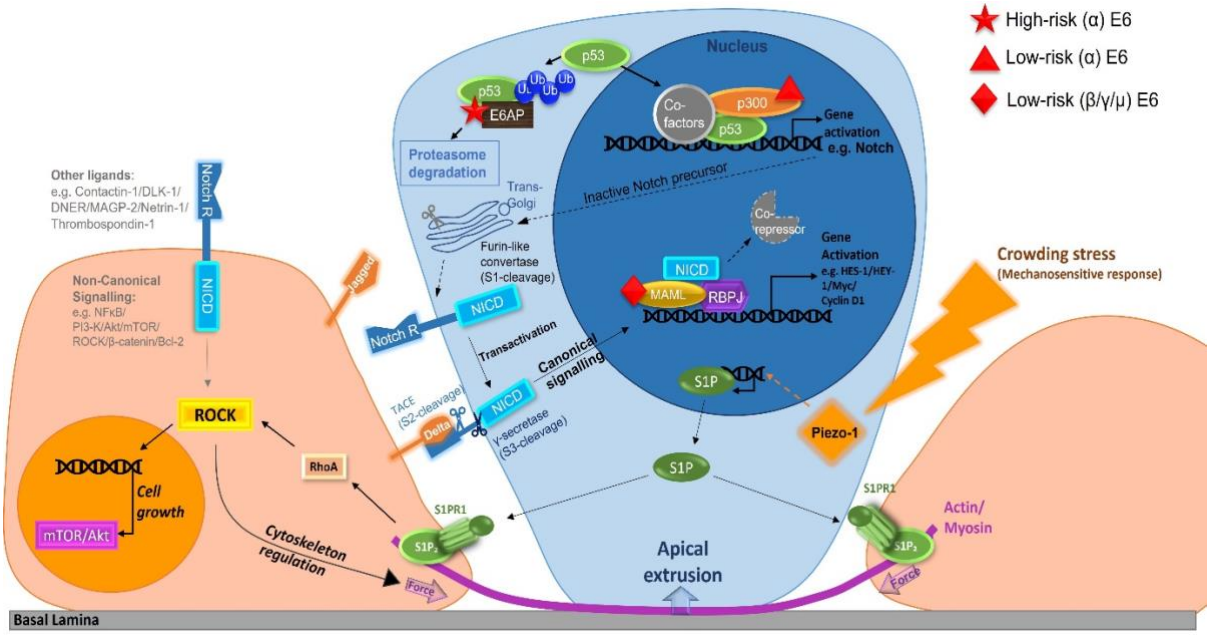


Figure 1.7.2: Cell Competition model of Apical Extrusion in Response to Basal Cell Overcrowding. Upregulation and translocation of the stress response protein Piezo-1 stimulate the release of the sphingosine-1-phosphate (S1P) molecule to adjacent cells. The binding of S1P to its receptor leads to induction of the ROCK pathway, with actin cytoskeleton

contraction, cell proliferation in the surrounding cells, and cell extrusion. The cell that delaminates from the basal layer has transiently higher expression of the Notch receptor on the cell surface, and becomes a Notch pathway signal-receiving cell. p53 is one of the transcriptional regulators of Notch precursor protein. High-risk E6 protein interacts with E6AP and targets p53 for proteasome degradation, downregulating total Notch receptor expression level. Low-risk E6 proteins can inhibit p300 and alter p53-dependent gene transcription. Cutaneous low-risk E6 protein bind with MAML to inhibit Notch-pathway activation. ROCK can also be activated by a non-canonical Notch signalling pathway.

There are a number of mechanisms underlying epithelial self-organization, from cell-cell signalling pathways, primarily the Notch pathway (Macara *et al.* 2014a). Functions of the Notch pathway in development can be largely divided into three categories: lateral inhibition, boundary formation and cell lineage decisions (Watt *et al.* 2008). Notch signalling in the epidermis is also considered to regulate cell adhesion to the extracellular matrix. Importantly, this defines the location of a cell and determines the signals received from the local microenvironment (Lowell *et al.* 2000). Reduced extracellular matrix adhesion stimulates cultured keratinocyte terminal differentiation (Watt 2002). Furthermore, cell extrusion from the basal layer is dependent on the activation of the small GTPase Rho effector kinase, ROCK, which leads to microtubule contraction (Slattum *et al.* 2009). Non-canonical Notch activation can activate ROCK independent of Notch transcriptional activity and induce keratinocyte differentiation (Yugawa *et al.* 2013). Therefore, Notch signalling is considered to be central in many epithelial homeostatic pathways.

The function of the Notch gene was first identified by genetic studies on the development of embryonic *Drosophila* epidermal cells (Poulson 1937). Human Notch1 and 2 genes are expressed in the suprabasal layer, whereas the Notch ligands, Delta & Jagged, are expressed in the basal layer (Uhlen *et al.* 2015). Delta-Notch signalling then drives early and late squamous epithelial differentiation of the spinous cell layer (Watt *et al.* 2008). Thus, inactivating mutations of Notch results in human malignancies, importantly squamous cell carcinoma (Wang *et al.* 2011).

The Notch receptor is synthesized as the inactive, ligand-inaccessible form in the endoplasmic reticulum, which is proteolytically cleaved and activated before insertion into the plasma membrane. In the trans-Golgi network, S1 cleavage by a Furin-like convertase generates a C-terminal fragment N(TM) and an N-terminal fragment N(EC). Following Notch receptor binding with Delta/Jagged ligands expressed by adjacent cells, Notch undergoes S2 cleavage by TNF-α converting enzyme (TACE), which yields a membrane-associated intermediate fragment, the Notch extracellular truncation (NEXT). Upon canonical activation, this fragment undergoes S3 cleavage by γ-secretase to release the Notch intracellular domain (NICD). NICD translocates into the cell nucleus and forms a complex with the RBP-J DNA binding protein. In the absence of NICD, RBP-J is bound to the transcriptional co-repressor protein Co-R. Thus, NICD-complex displaces Co-R on target gene promoters and recruits the transcriptional co-activator Mastermind-like (MAML1) protein (Kopan & Ilagan 2009a).

Notch mediates the transcriptional activation of target genes, including the HES/HEY family. In silico analysis of different tissues following Notch activation has shown that HES mRNAs are mostly upregulated in epithelial and endothelial tissues, whereas HEY mRNAs are more commonly observed in neural cells, depending on different ligand activations of Notch receptors (Katoh & Katoh 2007; Nandagopal *et al.* 2018). Within the HES family, Notch-activated HES-1 transcription is restricted to undifferentiated epithelial cells (Hirata *et al.* 2001).

The HES/HEY family encodes basic helix-loop-helix (bHLH) transcriptional repressors (Fischer & Gessler 2007). One important target of HES-1 is the p21 gene, the expression of which is a marker for Notch-induced growth suppression and keratinocyte early differentiation (Rangarajan *et al.* 2001a). Differential interaction with ligands can define cell fate during development. For example, interaction with the Notch-pathway Delta-like ligands (DLL1 and DLL4), directs a cell towards myogenesis and angiogenesis respectively (Nandagopal *et al.* 2018). The high-risk HPV16 E6 not only downregulates total Notch receptor expression level through a p53-dependent manner (Kranjec *et al.* 2017), but also changes the surface ligand

distribution favours DLL4 expression (Khelil *et al.* 2021). For epidermal homeostasis, early Notch activation at the basal layer causes the recipient cell to commit to differentiation. Induction of spinous gene expression is mediated by HES-1 expression, whereas basal gene repression occurs independently of HES-1 (Blanpain *et al.* 2006). However, the role of Notch in epithelial maintenance is less well established compared to early development. The differential role of ligands in Notch activation remains unclear, however, it is clear that Notch activation is associated with terminal epithelial differentiation.

Altogether, we hypothesize that HPV has evolved to target the Notch pathway under pressure to maintain a persistent infection of the basal cell layer. Notch1 transcriptional activity can be regulated directly by the degradation of p53 by high-risk HPV E6, which also reduces Notch1 expression on the surface of basal cells (Kranjec *et al.* 2017; Lefort *et al.* 2007). Therefore, HPV E6 can inhibit basal cell differentiation by targeting p53 degradation or transactivation. In addition, β-PV and μ-PV E6 proteins have been demonstrated to bind MAML1, the co-activator of Notch-induced transcription (Brimer *et al.* 2012). Altogether, this shows that a variety of HPV E6 proteins are able to either directly or indirectly downregulate Notch pathway activation and so dysregulate epithelial differentiation and homeostasis, contributing to persistent HPV infection.

1.7.3 Delamination

As the cells differentiate towards the apical epithelium, paracellular interactions change from desmosomes to adherent junctions and finally tight junctions (Figure 1.7.3). Reduced extracellular matrix adhesion (hemidesmosomes/integrin α6β4) also stimulates cultured keratinocyte terminal differentiation (Watt 2002). In response, membranous protein composition is tightly regulated (Marchiando *et al.* 2010).

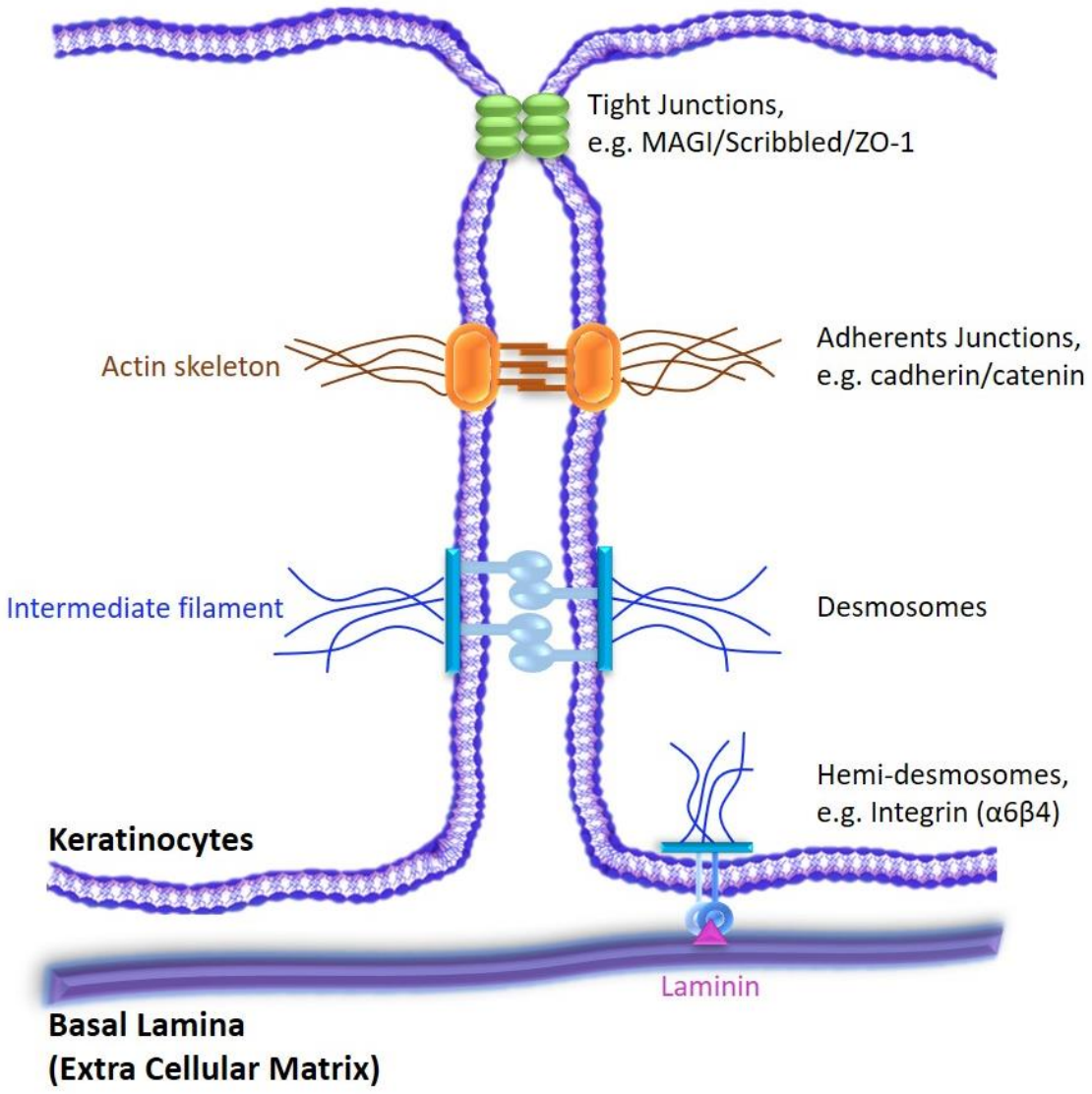


Figure 1.7.3: Cellular junctions that are predicted to be targeted by high-risk HPV proteins for pathogenesis. Types of cellular junctional proteins changes as the epithelium differentiates, as HPV life cycle is tightly co-regulated with different layers of the epithelium, the junctional proteins are important targets for HPV pathogenesis. Basal layer keratinocytes are attached to the basal lamina via Hemi-desmosomes. Adjacent basal keratinocytes are connected with desmosomes and intermediate filaments. While keratinocytes enter the suprabasal layer, actin skeleton bound adherent junctions (including the cadherin and catenin signalling pathways) becomes more abundant. Tight junctions form the most protective layer of the epithelium, consisting of MAGI, Scribbled and ZO-1 proteins, which are identified to be targted by high-risk HPV E6 proteins. Adapted from (Orr *et al.* 2019).

All high-risk α-HPV types are primarily associated with mucosal epithelial infections. The major difference distinguishing high-risk and low-risk HPV E6 protein is the additional PDZ-binding motif. Comprehensive analysis of the PDZ-binding motif revealed a positive relationship between the clinical significance of HPV E6 types and the number of associated PDZ-containing cellular proteins, also suggesting that the PDZ-binding motif is important for HPV pathogenesis (Thomas *et al.* 2016a). Apart from the degradation of p53, high-risk HPV E6 modify cell-to-cell junction arrangement by targeting proteins with PDZ-domains for proteasome degradation, namely zonula occludens-1 (ZO-1), Discs Large (Dlg) and Membrane-associated guanylate kinase, WW and PDZ domain-containing protein (MAGI) (Ganti *et al.* 2015). ZO-1 is a tight junction scaffolding protein essential for the development of unified apical surfaces and regulates epithelial proliferation in vitro (Balda *et al.* 2003). Complete knockout of ZO-1 does not cause spontaneous disease in mouse epithelium, possibly due to the redundancy of para-cellular junctional proteins. However, knockdown of ZO-1 and PDZ-deficient ZO-1 mutant expressing cells both have a thinner and more elongated shape to maintain cell volume without affecting the total basal-lateral area (Odenwald *et al.* 2018). Dlg protein is involved in a scaffolding protein complex with Scribble to regulate cellular basolateral polarity. Loss of Dlg results in activation of cell invasion and metastasis in several cancers (Tomiyama *et al.* 2015; Zhou *et al.* 2015). Dlg knockout mice display prominent defects in cell polarity and impaired progenitor cell maintenance and proliferation (Wang *et al.* 2014). The MAGI family of proteins has been shown to be one of the most strongly bound and most susceptible groups of proteins to be targeted by the high-risk E6 proteins (Thomas *et al.* 2002). The loss of tight junction integrity is a direct consequence of the loss of MAGI-1 in cervical cancer cell lines HeLa and CaSKi, which is reinstated upon E6 ablation and the re-emergence of MAGI-1 expression in these cells (Kranjec & Banks 2011). Up to date, all PDZ-associated protein and high-risk HPV E6 studies have been devoted to the oncogenic activity of E6 protein to disrupt cell contact and EMT. There is relatively little information about how the high-risk E6 PDZ-binding motif is involved in the modification of the para-cellular junction to remain at the basal lamina for persistent infection.

1.7.4 Differentiation

In normal epithelial homeostasis, keratinocytes are committed to differentiation as they enter the suprabasal layer of the epithelium. An early microarray analysis showed that downregulation of differentiation-related genes is primarily driven by high-risk E6 expression, with comparable results when E6/E7 full-length compared with E7 only or empty vector (Duffy *et al.* 2003). *In vitro* studies have also shown that exogenous expression of HPV E6 instead of E7 resulted in a delay in early differentiation marker Keratin 10 expression (Kranjec *et al.* 2017; Murakami *et al.* 2019). It is important to remember that E6 protein does not abrogate keratinocyte's ability to differentiate but only alters the cellular environment and favours the viral life cycle. Completion of the HPV life cycle requires a full differentiation scheme, and viral genome amplification for assembly happens at the upper layers of the epithelium (Moody 2017b). While for the aim to target viral reservoir clearance, understanding HPV E6 modulation in basal cell homeostasis and specifically targeting the infected cells for differentiation and extrusion will facilitate the development of 'true' anti-HPV therapies.

1.8 Aims

Based on the literature reviews, HPV E6 appears to play a major role in driving basal cell persistence following infection, and as a result of increased proliferation and reduced differentiation, acts to increase cellular fitness. Indeed, E6 expressing cells are predicted over time, to gradually displace adjacent uninfected keratinocytes, and to give rise to a clonal infection centre that can support the development of subsequent disease (Figure 1.8).

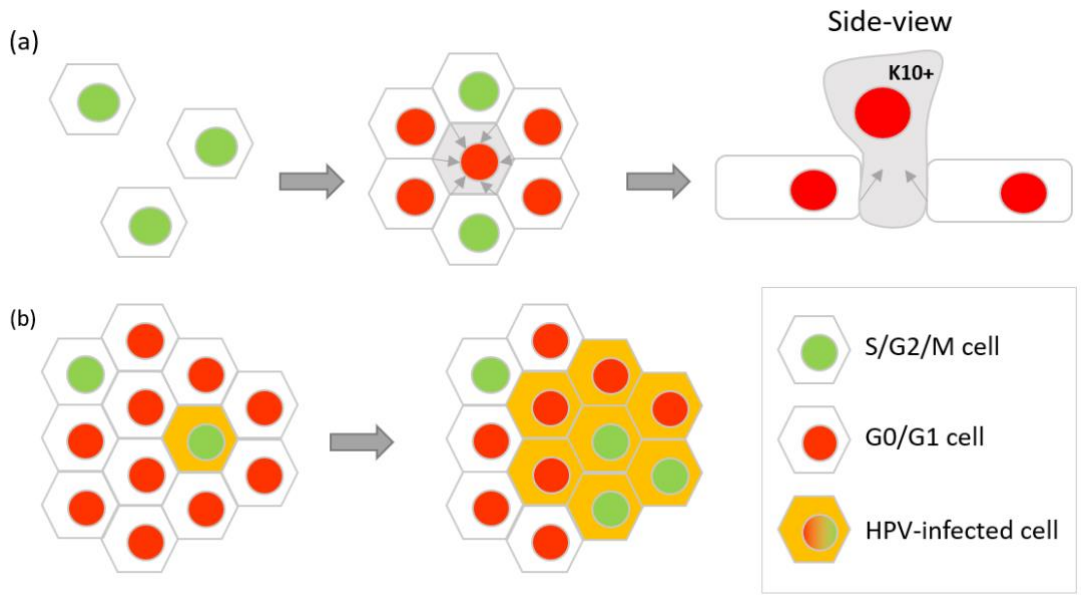


Figure 1.8: In vitro modelling of HPV E6 protein modulation in the basal epithelial organization. (a) Proliferating keratinocytes exit the cell cycle upon response to contact inhibition signal (at confluence), while a small proportion still retains the proliferation potential. Some keratinocytes will commit to differentiation (express K10) which delaminated from the basal layer and extruded towards the upper layers. (b) HPV-infected keratinocytes will gradually displace the uninfected basal cells by E6-mediated modulation of the cellular environment.

This thesis aims to:

i. **Establishment of *in vitro* three-dimensional modelling system of HPV E6 modulation in the basal epithelial organization.** With the use of a cell cycle indicator, differentiation markers and fluorescent cell competition assays, a reproducible and reliable *in vitro* modelling system suggested an evolutionarily conserved role of HPV E6 in modulating basal epithelial behaviour.

ii. **Identification of small molecule inhibitors against HPV16E6 functions during viral persistence.** Utilizing the model system for high-throughput screening to repurpose current drugs for anti-HPV effects.

iii. **Exploration of molecular pathways governing basal to suprabasal transitions, and key determinants for individual cell fate decision**. Identified proteins

essential for basal keratinocytes homeostasis targeted by HPV16E6 firstly at the populational level, then single-cell RNA sequencing further revealed individual cellular differences for future analysis.

2. Material and Methods

2.1 Molecular Cloning

2.1.1 Plasmid design

2.1.1.1 HPVE6

Sequences for HPV1E6, 8E6, 27E6 and 65E6 were obtained from PaVE (Papillomavirus Episteme) database. Genome plasmid for HPV1/HPV65 and a plasmid containing 8E6 were used for Polymerase chain reaction (PCR) with primers designed containing attB1/attB2 sites, KOZAK sequence and HA-tag (Figure 2.1). The sequence of 27E6 was synthesized by Invitrogen (ThermoFisher Scientific) in pMK-RQ plasmid backbone with Kanamycin resistance as a selection marker. HPV11E6 was cloned from pBR322_HPV11 gifted from Neil Christensen, and HPV11E6 was cloned form pTZH_HPV16 gifted from Margaret Stanley.

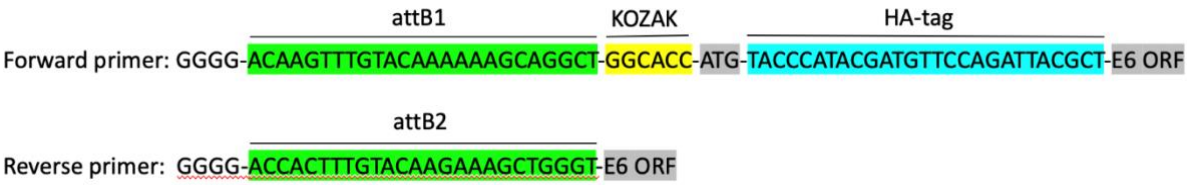


Figure 2.1.1.1: Primer design for PCR amplification of HPV E6 (5'-3').

2.1.1.2 shRNA

RNA knockdown plasmids (pSI_CLXPX_shRNA) were gifted from Dr. Tohru Kiyono.

2.1.1.3 Fluorescent expression vector (FUCCI/eGFP/mCherry)

Expression plasmid vector pBOB_EF1_FastFUCCI (Gift from Kevin Brindle & Duncan Jodrell), pQCXIP_eGFP and pQCXIP-mCherry were created by Nagayasu Egawa & Taylor Saunders-Wood

2.1.2 Polymerase chain reaction (PCR)

PCR was performed with KOD Hot Start DNA Polymerase (Sigma-Aldrich) using either genomic DNA plasmid or plasmid previously constructed containing the E6 sequence (Table 2.1.1/2.1.2). All PCR products were cleaned with Wizard® SV Gel and PCR Clean-Up System (Promega).

Reagent	Volume in each reaction
10 µM forward primer	5 µl
10 µM reverse primer	5 µl
2 mM dNTP	4 µl
25 mM Mg^{2+}	3 µl
10x KOD polymerase buffer	5 µl
KOD Hot Start DNA Polymerase	1 µl
DNA template	(10ng-100ng)
Nuclease Free Water	Make up to 50 µl
Total	50 µl

Table 2.1.2.1: Reaction mixture of HPV E6 PCR amplification.

Step	Temperature	Duration	Number of cycles
Denaturation	94 ℃	2 min	1
Denaturation	94 ℃	30 sec	30
Annealing	60 ℃	30 sec	30
Extension	68 ℃	30 sec	30
Extension	68 ℃	2 min	1
Cooling	4 ℃	∞	1

Table 2.1.2.2: PCR reaction condition set-up.

2.1.3 Plasmid construction (Gateway® cloning)

PCR amplicons of HPV E6 were recombined with donor vector pDONR221 to generate entry vector pENTR_HPVE6 using Gateway™ BP Clonase™ II Enzyme Mix (Invitrogen™). Then pENTR_HPVE6 was recombined with two different destination vectors using Gateway™ LR Clonase™ II Enzyme Mix (Invitrogen™) to obtain the final plasmid (pLXSN_HPVE6 and pQCXIN_HPVE6) (Plasmid backbone design shown in Appendix Figure 1). Both BP and LR reactions were carried out at 25℃ for one hour, by mixing 1μl of each plasmid and 0.5μl of clonase enzyme mix.

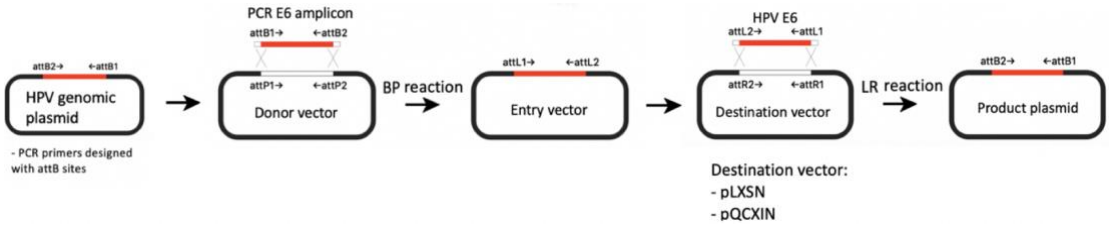


Figure 2.1.3: Schematic illustration of Gateway® cloning system for HPV E6 plasmid construction.

2.1.4 Agarose gel electrophoresis

Samples were stained with 6x DNA loading dye (ThermoFisher). Agarose gel (1%) were made by adding 1 gram of agarose in 0.5x TAE (same as running buffer), heated by microwave until fully dissolved. Gels were cast and run at a constant voltage of 100V using a Mupid-One Submarine gel electrophoresis system in 0.5x TAE buffer. For visualization, the gel was added to TAE buffer containing SybrSafe (10,000x dilution, Invitrogen). Gels were visualised using UV and DNA were recovered from the gel using Wizard® SV Gel and PCR Clean-Up System (Promega).

2.1.5 *E.coli* transformation

For entry and expression vector plasmids, total 2.5μl reaction mixture was incubated with 25μl of One Shot™ TOP10 Chemically Competent *E.coli* for 30min on ice. After heat-shock at

42℃ for 45 seconds, the reaction mixture was immediately transferred back on the ice for 2min. Transformed competent *E.coli* was pre-incubated in 200μl of SOC media (Invitrogen™) for 30min-1hr at 37℃, and then plated onto LB agar plates with antibiotics for overnight culture. Kanamycin (50μg/ml) and Ampicillin (100μg/ml) resistance were used for pENTR and pLXSN/pQCXIN selection respectively.

For donor and destination vector plasmids, total 2.5μl reaction mixture was incubated with 25μl of One Shot™ ccdB Survival 2 T1R Competent Cells, following the same protocol of bacterial transformation. Additional Chloramphenicol (50μg/ml) resistance were used for selection together with Kanamycin (50μg/ml) for pDONR / Ampicillin (100μg/ml) for pDEST.

2.1.6 Bacterial plasmid DNA extraction

A single colony was picked from the LB agar plate after overnight incubation at 37℃ and transferred into 2ml LB broth with same antibiotics. Extraction of plasmids from *E.coli* using QIAprep Spin Miniprep Kit (Qiagen). For expression vector to be used in cell line transfection, liquid culture was expanded to 250ml LB broth with antibiotics, then plasmids extracted using HiSpeed Plasmid Maxi Kit (Qiagen).

2.1.7 Sequencing confirmation of plasmids

Universal primers M13/pBABE5 were used for sequencing entry/expression vectors respectively (Appendix Figure 1).

2.2 Tissue culture

2.2.1 Media and supplements

Media	Components

Dulbecco's Modified Eagle Medium (DMEM)	Dulbecco's Modified Eagle Medium (DMEM, Sigma-Aldrich), with 10% foetal bovine serum (FBS, labtech.com, UK) and 1% Penicillin/Streptomycin (Gibco, UK)
F-Incomplete (FI)	DMEM:F-12 (Sigma-Aldrich) = 1:3, 0.66mM Ca2+, with 5% FBS (labtech.com, UK), 1% Penicillin/Streptomycin (Gibco, UK), 5μg/ml insulin (Sigma, UK), 8.4 ng/ml cholera toxin (Sigma, UK), 24 μg/ml adenine (Sigma, UK), 0.4 μg/ml hydrocortisone (Calbiochem)
F-Complete (FC)	FI media, 10ng/ml epidermal growth factor (R&D Systems, USA)
Freezing media	10% DMSO (Merck) in FBS (labtech.com, UK) or TC-Protector Cell Freezing Medium (Bio-Rad Laboratories Ltd.)

Table 2.2.1: Tissue culture media and supplements.

2.2.2 Subculture of cell lines

Culture vessel	Seeding density as 1x	Volume of media	Volume of Trypsin-EDTA
T175 flask	$3x10^6$	25ml	5ml
T75 flask	$1x10^6$	10ml	3ml
T25 flask	$3x10^5$	3ml	1ml
6-well plate	$1x10^5$	2ml	1ml
96-well plate	$6x10^3$	200μl	-
8-chamber slides	$2x10^4$	500μl	-

Table 2.2.2: Cell culture seeding density and media usage

2.2.2.1 HEK293TT

HEK293TT cells were maintained in DMEM (37℃, 5% CO_2) with media change every 48hr. For passaging, HEK293TT cells were firstly washed with 5ml sterile 1xPBSA twice to remove FBS in media, then detached by treated with 3ml of Gibco™ Trypsin-EDTA (0.25%) phenol red (ThermoFisher Scientific) for 5-15min at 37℃. Detached cells were neutralized with 5ml of DMEM, and then pelleted by centrifugation at 1500rpm for 5min at 4℃. Cells were resuspended in 6-8ml of culture medium.

Cell number was counted using Z Series Coulter Counter (Beckman Coulter), by adding 200μl of cell suspension into 9.8ml of Beckman Coulter™ ISOTON™ II Diluent (FisherScientific).

2.2.2.2 J2-3T3 Mouse Fibroblasts

J2-3T3 cells were maintained same as HEK293TT cells in DMEM up to 90% confluency. For making irradiated feeders, J2-3T3 cells were irradiated for 14 minutes and 14 seconds (220kV, 14mA; resulting dosage was 60Gy). Frozen stocks were stored at -80℃ and expected recovery rate is 50%.

2.2.2.3 EF-1F Human Foreskin Fibroblasts

EF-1F cells were maintained same as HEK293TT cells in DMEM up to 90% confluency. For preparing dermal equivalents, see section 2.3.2.

2.2.2.4 NIKS Normal Immortalised Keratinocytes

NIKS (Normal Immortalized Keratinocyte) were maintained at sub-confluence (70-80% confluency) with 0.5x seeding density of γ-irradiated J2-3T3 feeder cells (Flores *et al.* 1999), in F-complete media. Feeders were plated 24hr prior to keratinocytes for proper attachment.

For NIKS cells passage, feeder cells were removed by incubation with 5ml of PBSA + 0.5mM EDTA (diluted from UltraPure™ 0.5M EDTA, pH 8.0, ThermoFisher Scientific) for 5min at room temperature. After wash twice in 5ml of sterile PBSA, NIKS cells were then incubated with 3ml of Gibco™ Trypsin-EDTA (0.25%) phenol red (ThermoFisher Scientific) for 15-20min at 37℃. Detached cells were neutralized with 5ml of FC medium, and then pelleted by centrifugation at 1500rpm for 5min at 4℃. Cells were resuspended in 6-8ml of culture medium. After counting, NIKS were seeded depending on different experimental purposes.

2.2.3 Expression cell line establishment

2.2.3.1 Transfection of HEK293TT

HEK293TT cells were seeded at a density of $4x10^6$ cells per T75 flask 24hr prior to transfection. Expression vector (5μg), Gag-pol (3.3μg) and VSV-G (1.7μg) in 500μl of Gibco™ Opti-MEM™ I Reduced Serum Medium (FisherScientific) and 30μg of Polyethyleneimine (PEI) in 500μl of Gibco™ Opti-MEM™ I Reduced Serum Medium (FisherScientific) were mixed together and developed for 30min at room temperature. Then total 1ml of the reaction mixture was added into 293TT cells dropwise. Media was changed 24 hours after transfection.

2.2.3.2 Extraction of retrovirus

Retrovirus was collected in culture media 48 hours after changing media and filtered through 0.45μm filter to remove cell debris. Retroviruses are stocked in 1ml per vial at -80℃ for future uses.

2.2.3.3 Transduction of NIKS

For transduction, NIKS cells were plated at $2x10^5$ cells per well in 6-well plates, 24 hours in advance. Retrovirus stock (1ml) and FC media (1ml) was added into NIKS cells with polybrene (10μg/ml) and incubated for at least 2 hours before replaced with FC media. Cells were selected by 400μg/ml Gibco™ Geneticin™ Selective Antibiotic (FisherScientific) / 100-

200μg/ml Gibco™ Hydromycin B (Invitrogen) in FC media for 48 hours for successful transduction.

2.2.4 Cell treatment with small molecules

All small molecules were stocked at 10mM in DMSO, at -20℃.

2.2.4.1 1st screening

Prestwick Chemical Library® were supplied by Maruho Co.Ltd (Appendix File 1). Working concentration of the small molecules were 10μM (1:1000 dilution with FC media). Same amount of DMSO (0.01%) was used as vehicle control.

2.2.4.2 2nd confirmation screening

Confirmation of small molecule effect with HIT molecules only. IC50 were carried out for some of the high-priority candidates.

HIT PLATE-A	1	2	3	4	5	6	7	8	9	10	11	12
A		P01C04	P02C02	P02E07	P03A11	P03B10	P03D05	P03E07	P03F11	P05B02	P05H08	
B		P01C08	P02C05	P03A02	P03B02	P03C03	P03D06	P03E08	P03G02	P05B04	P05H09	
C		P01D09	P02C08	P03A04	P03B03	P03C07	P03D07	P03E09	P03H10	P05D05	P06G08	
D		P01F05	P02C10	P03A05	P03B04	P03C08	P03D08	P03F03	P04F04	P05E07	P06H06	
E		P01F10	P02D04	P03A06	P03B05	P03C09	P03D09	P03F04	P04G08	P05E10	P07A04	
F		P02B02	P02D07	P03A07	P03B06	P03C10	P03D10	P03F05	P04H07	P05E11	P07H06	
G		P02B03	P02D08	P03A08	P03B07	P03C11	P03E03	P03F06	P04H11	P05G08	P07H07	
H		P02B08	P02D09	P03A09	P03B08	P03D03	P03E06	P03F10	P05A09	P05H06	P07H09	
HIT PLATE-B	1	2	3	4	5	6	7	8	9	10	11	12
A		P07H11	P09E07	P10D09	P12C10	P12E04	P12G02	P13B07	P13E05	P13H07	P15A06	P15E03
B		P08B03	P09G10	P10D11	P12D05	P12E07	P12G03	P13C03	P13E06	P14A04	P15A11	P15E05
C		P08D03	P10A06	P10F06	P12D06	P12E09	P12G04	P13C05	P13E07	P14B04	P15B05	P15F05
D		P08E06	P10A09	P10G08	P12D07	P12F02	P12G06	P13C09	P13E08	P14D04	P15C06	P15F09
E		P08E09	P10B05	P11F04	P12D09	P12F04	P12G09	P13D07	P13E09	P14F07	P15D03	P15G03
F		P08G06	P10C05	P11G06	P12D10	P12F06	P12H03	P13D10	P13E10	P14G05	P15D04	P15H02
G		P08H08	P10C09	P11H06	P12E02	P12F07	P12H08	P13E03	P13F05	P14G08	P15D05	P15H05
H		P09C06	P10D05	P12C07	P12E03	P12F09	P13B03	P13E04	P13F10	P14H07	P15D10	

Table 2.2.4.2: HIT molecule layout for 2nd screening.

2.2.4.3 Follow-up analysis

Chemical name	CAS ID	Catalogue number

Tipranavir	174484-41-4	CAY21028
Lopinavir	192725-17-0	-
Darunavir	206361-99-1	-
Saquinavir mesylate	149845-06-7	-
Atazanavir	198904-31-3	HY-17367A
Indinavir sulphate	157810-81-6	-
Nelfinavir	159989-64-7	-
Fosamprenavir	226700-81-8	CAY21609

Table 2.2.4.3.1: List of HIV protease inhibitors. Unspecified catalogue molecules were obtained from Prestwick Chemical Library.

Chemical name	CAS ID	Catalogue number
Bromfenac	91714-93-1	CAY23763
Etodolac	41340-25-4	-
Diclofenac	15307-86-5	-
Lornoxicam	70374-39-9	CAY70220
Etoricoxib	202409-33-4	-
Lumiracoxib	220991-20-8	CAY27645
Celecoxib	169590-42-5	-
Ketorolac	74103-06-3	-
Polmacoxib	301692-76-2	CAY17509
Meloxicam	71125-38-7	-
Valdecoxib	181695-72-7	-
Parecoxib	198470-84-7	HY-17474
Zaltoprofen	74711-43-6	CAY14662
Imrecoxib	395683-14-4	HY-114200
Parecoxib Na	198470-85-8	CAY29017

Table 2.2.4.3.1: List of selective COX-2 inhibitors. Unspecified catalogue molecules were obtained from Prestwick Chemical Library.

2.3 Organotypic raft culture

2.3.1 Preparation of dermal equivalents

Fibroblasts were harvested and resuspended in $1x10^7$ cells per ml. To make six dermal equivalent, collagen premix (25ml) was prepared with 10x DMEM (2.5ml), 1N NaOH (460μl), 100x Penicillin/Streptomycin (250μl), FBS (2.5ml) and Corning® Collagen I Rat Tail 100mg (20ml). The entire procedure was carried out on ice to prevent collagen cross-linking. Then 1ml of the collagen premix was plated in each transwell insert (Costar Transwell Permeable Supports 6-well). Allow the collagen to solidify into a gel by incubation at 37℃ for at least 5min. Meanwhile, $6x10^6$ EF-1F cells were added into the collagen premix. After the acellular layer of collagen is solidified, plate 2.6ml of collagen with fibroblast on top into each transwell insert. Allow the gel to set for at least 2hr at 37℃. Finally, add enough DMEM media to immerse the whole dermal equivalent. Collagen gels can be used 24hr after and no longer than 7 days.

2.3.2 Seeding and differentiation of NIKS

On the day of plating keratinocyte onto dermal equivalent (Day 0), transwell inserts were transferred into 6-well deep well plates (DISCOVERY LABWARE). After the collagen gel surface dried, a total number of $1x10^6$ NIKS cells were resuspended in 100μl of FC-media and loaded onto each dermal equivalent. Incubation in 37℃ for 2hr allowed the keratinocyte to attach. To lift the raft culture to an air-liquid interface, at least 4 cotton pads were placed under the transwell insert and 12ml of FC-media was added. 24hr after lifting the transwell insert, media was changed to the cornification media (FI-media with 1.88mM $CaCl_2$ and 10μM of C8:0). Cornification media was changed every 48hr.

2.3.3 Harvesting raft culture samples

On Day 15, raft culture was soaked in 4% PFA in PBS overnight for fixation at 4℃. After fixing, raft cultures were kept in PBS at 4℃. Paraffin-embedded rafts were then transversely sectioned (sections of 5μm thickness) on to Superfrost plus slides.

2.4 Phenotypic analysis

2.4.1 Western blot

2.4.1.1 Buffer and Reagents

Buffer	Constituents
RIPA lysis buffer	25mM Tris-HCl pH=7.6, 150mM NaCl, 1% Triton-X100, 1% sodium deoxycholate, 0.1% SDS
Running buffer	NuPAGE MOPS SDS Running Buffer (1x, stock at 20x)
Transfer Buffer A	0.3 M Tris, 20% Methanol, 0.02% SDS
Transfer Buffer B	25 mM Tris, 20% Methanol, 0.02% SDS
Transfer Buffer C	25 mM Tris, 20% Methanol, 6-amino-caproic acid, 0.02% SDS

Table 2.4.1.1: Western blot buffer and constituents.

2.4.1.2 Protein extraction

NIKS cells were seeded at $1x10^5$ per well on 6-well plates, with $5x10^4$ feeder cells. Media changed every 48 hours. Cell lysates were collected by scraping with RIPA buffer (added protease inhibitor (ab201111) / phosphatase inhibitor (ab201112)). Sonicate cell lysate on water-ice mixture 4℃ for 10min and spin down at maximum speed for 10min at 4℃. Transfer the supernatant into new Eppendorf tubes. Protein samples were kept at -20℃.

2.4.1.3 Protein quantification

Protein concentration was characterized by Pierce™ BCA Protein Assay Kit (Thermo Scientific™), following manufacturers' protocol.

2.4.1.4 Protein gel electrophoresis

Total protein amount of 20-25μg was firstly denatured with appropriate volume of 4X buffer and β-Mercaptoethanol (must make up 2.5% of total volume) and boiled at 80℃ for 15min. After cooling down on ice, the samples were loaded onto NuPageTM 4-12% Bis-Tris Protein Gels and resolved with 60mA for 90min. Amersham™ ECL™ Rainbow™ Marker (Full Range, Merck) was used as ladder.

2.4.1.5 Semi-dry membrane transfer

PVDF membrane (BioRad) was primed in 100% methanol for 10-15 seconds and soaked in buffer C for 10min. Protein gels were soaked in buffer C for 10min before transfer. Filter papers were soaked with different buffers to make the semi-dry transfer sandwich. Protein was transferred to membrane using Trans-Blot® SD Semi-Dry Transfer Cell (Bio-Rad), at 130mA for 45-60min per gel.

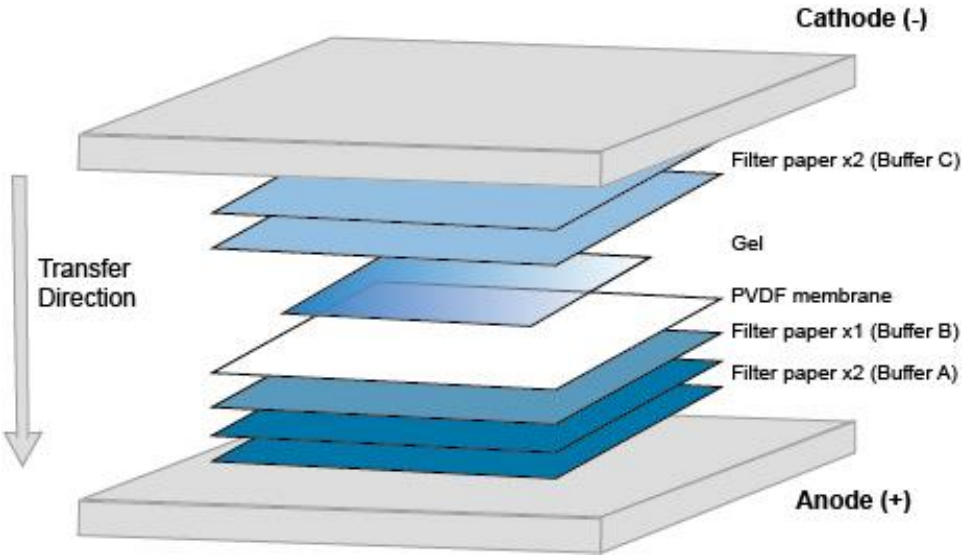


Figure 2.4.1.5: Graphic illustration of Semi-dry transfer sandwich.

2.4.1.6 Protein detection

Membranes were incubated in 1% milk/PBS Tween-0.01% (v/v) for one hour at room temperature after transfer. Membranes were then cut if necessary to allow probing with multiple antibodies. Membranes were next incubated with primary antibodies diluted in 1% milk/PBS Tween-0.01% at 4°C overnight.

Membranes were washed in PBST for three times (5min each time) on shaker to remove unspecific primary antibody bindings. Then probed with secondary antibodies diluted in 1% milk/PBST in dark for 1hr at room temperature with shaking. After final washes in PBST for five times (5min each time) on shaker, membrane left to completely dry between filter papers before imaging with Odyssey imaging system (Licor).

2.4.2 Flow cytometry

2.4.2.1 Colour intensity sorting

Cells were sorted for fluorescent intensity with a BD Aria Fusion sorter. A 488nm laser was used to detect GFP/mAG positive cells in the 530/30nm channel. A 561nm laser was used to detect mCherry/mKO2 positive cells in the 610/20nm channel. Cells were sorted into four quartiles based on fluorescent intensity to select for the highest expressing cells.

2.4.2.2 Analytical FACS

For analysis of colour expressing NIKS cells (FUCCI), cells were seeded at different densities on 6-well plates with feeders depending on experiment purpose. Media was changed 24hr after seeding and cells were collected 48hr post changing media. For harvest, after wash in 1ml sterile PBS twice, feeder cells were firstly removed by trypsinization for 2min at 37℃. Then replaced with 1ml of trypsin-EDTA and incubated at 37℃ for 15-20min. Cells are neutralized with 2ml of FC medium and centrifuged at 1500rpm for 5min at 4℃. After aspirating the

supernatant, cell pellets were resuspended in 1ml of FACS running buffer (2% FBS and 0.5mM EDTA in PBS) and kept on ice.

For DNA content analysis, NIKS cells were stained with live DNA stain DRAQ5 (1:250, ab108410, Abcam) for 10min on ice, and used for Flow Cytometry analysis without washing.

For staining of intracellular antigens, NIKS cells were seeded in T25 flasks starting at different densities, with feeder cells. Media was changed 24hr after seeding and cells were collected 48hr post changing media. Cells were harvested by trypsinization and spin down same as described before. Then fixed in 1ml 0.1% PFA in PBS for 10min at room temperature. NIKS cells were washed with 2ml of permeabilization buffer (4% FBS, 0.1% Triton-X in PBS) by spin down at 2000rpm for 5min, before treating with permeabilization buffer for 15min at room temperature. Both primary and secondary antibodies were prepared in permeabilization buffer, with the same dilution as immunofluorescence staining. After permeabilization and spin down at 2000rpm for 5min, cells were resuspended in 100μl of primary antibody and incubated for at least 30min in dark on ice. Then the cells were washed with 1ml of permeabilization buffer by centrifugation at 2000rpm for 5min. Cells were then incubated with 100μl of secondary antibody for 20min in dark on ice. After washing with permeabilization buffer and spin down, cells were finally resuspended in 1ml of the FACS running buffer for analysis.

Before FACS analysis, all cell suspensions were filtered through Falcon® 70μm cell strainer (CORNING) to remove doublets. Flow cytometry results were obtained with the Cytek® DxP8 machine and analysed with FlowJo software (Version 10). Filter setting 530/30 was used for mAG signal, 615/25 for mKO2 and 666/27 for Alexa Fluor 647/DRAQ5.

2.4.3 Immunofluorescence

NIKS were seeded at different densities on Imaging Chamber 8 (MoBiTec) with feeders. Media were changed 24hr after seeding. 48hr post changing media, cells were fixed with 4% PFA in

PBS for 10min at room temperature, with wash in PBS twice before and after fixation. Cells were stored in PBS at 4℃ or directly used for permeabilization with 0.1% Triton X-100 (Promega) for 30min at room temperature. After washing twice with PBS, cells were blocked in 10% normal horse/goat serum (depending on secondary antibody species) in PBS for 1hr at room temperature. Primary and secondary antibodies were diluted with blocking buffer.

Cells were incubated with the primary antibody in dark at 4℃ overnight. After washing with PBS 5min for 3 times, the secondary antibody was added to the cells together with nucleus staining and incubated for 1hr at room temperature in dark. For amplification of primary antibody, cells were incubated with ImmPRESS® HRP Polymer Detection Kit for 1hr at room temperature in dark. After washing with PBS 5min for 3 times, chromogenic staining was developed using fluorescent tyramide/DAB IHC Detection Kit for 2min and terminated in water. Finally, slides are mounted with CitiFluorTM AF1 (Electron Microscopy Sciences) and kept at 4℃ in dark. Immunofluorescence images were taken with Carl Zeiss Microscopy AxioCam MRm/c camera.

2.4.4 Immunohistochemistry

Raft samples embedded in paraffin went through a process of dewaxing and rehydration. Slides were incubated at 60℃ overnight and dewaxed by submerged in two portion of xylene for 10min and 5min sequentially. Then tissues were rehydrated with graded alcohols for 2 minutes at a time (twice in 100% ethanol, followed by once in 80%, 50% and 30% ethanol). Finally soaked in dH_2O for 5min.

For epitope exposure, slides were equilibrated in DAKO Target Retrieval Reagent (Aligent, USA) for 10 minutes at room temperature, before being heated to 110°C for 15 minutes in a Decloaking Chamber (Biocare Medical, USA). After cooled down to room temperature, slides were equilibrated with 1x PBS for 5min on shaker. ImmEDGE™ Hydrophobic Barrier Pen (Vector Lab) was used to mark around the tissue and same process as immunofluorescence were carried out to stain for target protein.

2.4.5 Array Scan/Opera Phenix

NIKS cells were seeded at different densities with 1x represent $6x10^3$ cells/well (Appendix File 2) on Corning® 96 well black polystyrene microplates (Perkin Elmer), on top of $3x10^3$ feeder cells. Media was changed 24 hours after seeding. On Day3 (48hr after changing media), cells were washed twice with PBS and fixed with 4% paraformaldehyde in PBS for 10min. FUCCI NIKS were permeabilized and stained for K10 follow the same protocol as immunofluorescence. Fixed cells were kept in 0.5% Procline in PBS to prevent dry out and send for ArrayScan (ThermoFisher)/Opera Phenix (Perkin Elmer) for data collection.

2.4.6 Competition assay

For high density competition assay, eGFP/mCherry NIKS (1:1 ratio) were seeded on Imaging Chamber 8 (MoBiTec) or Corning® 96 well black polystyrene microplates (Perkin Elmer) with feeders. Cells were maintained in FC media for up to 9-day, with media changed every 48hr after seeding. Cells were fixed with 4% PFS in PBS and stained with DAPI (1:1000 dilution in PBS) for imaging by LSM700 Confocal Microscope (Zeiss) or Opera Phenix (Perkin Elmer).

Protein target	Serial number	Western blot	Immuno-fluorescence	Immuno-histochemistry	FACS
GAPDH	MAB374	1/1000			
p53	DO-1	1/200	1/200	1/200	
NICD	Val1744	1/1000	1/200	1/200	
K-10	ms611p	1/200	1/200-300	1/200	
K-10	ab76318	1/10000	1/150	1/150	
Notch1	EP1238Y			1/200	
Notch1	C44H11	1/1000	1/200		
HES1	D6P2U	1/1000			
HES1	ab119776			1/150	
HA-tag	ab130275	1/1000	1/1000		1/1000
involucrin	MA5-11803			1/200	

involucrin	sc28557	1/200			
RBP-JK	sc-8213	1/200			
K-18	ab32118		1/200		
p-EGFR	sc12351	1/200			
active-YAP1	ab205270	1/1000	1/2000	1/500	
flagrrin	sc66192	1/200	1/50	1/50	
MCM7	ab2360		1/50	1/100	
E-cadherin	sc21791	1/200		1/100	
COX2	12282T	1/1000			1/200
DSG2	ab150372	1/10000			
Geminin	sc-74456		1/200		
MCM7	ab52489			1/100	
p63a	4892S	1/500			
dNP63	Poly6190	1/500			
p63	4A4			1/400	
DLG1 (SAP97)	2D11			1/200	
CALML-5	HPA040725			1/400	
K-14	GTX76603			1/200	
DLG1 (SAP97)	2D11			1/200	
DLG4 (PSD95)	sc-32290			1/200	
MAGI	sc-100326	1/200		1/200	
Scrib	sc-11049			1/200	
FAP (PTPN13)	NBP2-94207			1/200	
hPAR3	sc-53819			1/200	
PATJ	PA5-76397			1/200	
EGFP	CAB4211	1/400		1/200	
mCherry	M11217	1/400		1/200	

Table 2.4.1: Primary antibody usage.

Detection Probe	Serial number	Species	Dilution	Application
IRDye 680RD	926-68071	Goat anti-rabbit IgG	1/5000	WB
IRDye 800CW	926-32210	Goat anti-mouse IgG	1/5000	WB
DAPI	11926621	-	1/1000	IF/IHC
DRAQ5	ab108410	-	1/250	FACS
DRAQ7	ab109202	-	1/500	IF
Alaxa Flour 350	A21068	Goat anti-rabbit IgG (H+L)	1/200	IF/IHC
Alaxa Flour 488	A11029	Goat anti-mouse IgG (H+L)	1/200	IF/IHC
Alaxa Flour 488	A11008	Goat anti-rabbit IgG (H+L)	1/200	IF/IHC
Alaxa Flour 594	A11032	Goat anti-mouse IgG (H+L)	1/200	IF/IHC
Alaxa Flour 594	A11012	Goat anti-rabbit IgG (H+L)	1/200	IF/IHC
Alaxa Flour 594	ab150160	Goat anti-rat IgG (H+L)	1/200	IF/IHC
Alaxa Flour 647	A21235	Goat anti-mouse IgG (H+L)	1/200	IF/IHC
Alaxa Flour 647	A21244	Goat anti-rabbit IgG (H+L)	1/200	IF/IHC
ImmPRESS HRP Polymer Detection Kit, Peroxidase	MP-7402	Horse anti-mouse IgG	-	IF/IHC
ImmPRESS HRP Polymer Detection Kit, Peroxidase	MP-7401	Horse anti-rabbit IgG	-	IF/IHC
CF350 tyramide	92170	-	1/400	IF/IHC
CF488A tyramide	92171	-	1/400	IF/IHC
CF555 tyramide	96021	-	1/400	IF/IHC
CF640R tyramide	92175	-	1/400	IF/IHC

Pierce™ DAB Substrate Kit	34002	-	-	IF/IHC

Table 2.4.2: Secondary (detection) antibody usage.

2.5 Data analysis

2.5.1 Analysis software

Fluorescent images were analysed with Fiji (ImajeJ). ArrayScan data analysed with pairing software HCS Studio Cell Analysis Software (ThermoFisher). Opera Phenix data analysed with pairing software Harmony (Perkin Elmer).

2.5.2 Statistical methods

Unless specified, all datasets were analyzed with Prism (GraphPad) and carried out One-way ANOVA with GP-style P-values. All figures were shown as Mean +/- standard deviation (SD).

2.6 Single cell Transcriptome sequencing

2.6.1 Sample preparation and sequencing service

NIKS and NIKS LXSN 16E6 are seeded on 6-well plates at $8x10^5$ cells per well with $1x10^5$ irradiated J2-3T3 feeder cells seeded 24hr in prior. Media (FC) was changed 24hr after seeding and cells were then allowed to expand for another 48hr until collection.

When collecting single cell suspension, feeder fibroblasts were firstly removed by 1ml Trypsin-EDTA treatment for 1min at 37℃ after PBS wash. Then NIKS cells were detached by incubating 1ml Trypsin-EDTA at 37℃ for 15-20min. Cell suspension is neutralized by equal amount of FC media and transferred to FALCON® tube to pellet the cells by centrifuge at 1500rpm for 5min at 4℃. NIKS pellet are resuspended by pipetting and washed in PBS + 0.04% BSA twice by centrifuge. Single cell suspension was then filtered through 70μm cell strainer to remove

doublets. Finally, before sample submission, cell viability and concentration were determined using a Countess® II Automated Cell Counter (to have final 6-10,000 cells in 45μl of PBS + 0.04% BSA with cell viability >90%).

Single cell RNA sequencing was conducted by CRUK, Cambridge; using 10X Genomics service prepared with 3' gene expression library and sequenced on SLX_NovaSeq6000_S1 chip.

2.6.2 Sequencing result analysis

Adapter trimmed sequences were aligned against human reference genome (refdata-gex-GRCh38-2020-A) using cellranger count. Output raw count matrix were processed with R studio, analysed with Seurat (Hao *et al.* 2021) & Cellchat (Jin *et al.* 2021).

3. *In vitro* modelling of HPVE6-mediated viral persistence at the basal epithelium

3.1 Introduction

As the HPV life cycle is closely linked with the epithelial terminal differentiation programme, it is important to understand how the tissue is regulated. For epithelial homeostasis, contact inhibition of proliferation, exiting the basal layer (including cell division directionalities and apical extrusion) and commitment to differentiation are the main processes proposed for multi-layered epithelial development and maintenance.

Initially, skin cell cultures were primarily 2D with keratinocytes (Rheinwald & Green 1975). The skin consists of epidermal and dermal layers, in which keratinocytes and fibroblasts are the main cell types contributing to the two-layer, respectively. With the monolayer tissue culture system using NIKS and 3T3 fibroblasts, the aim was to mimic the behaviour of the basal layer cells of the epithelia. NIKS was chosen for the *in vitro* system with its high reproducibility for experiments and transcript expression profile comparable to primary keratinocytes (Allen-Hoffmann *et al.* 2000a). Although the 2D tissue culture system provided limitations with respect to testing as the results obtained from *in vitro* 2D cell culture studies could not always be translated to *in vivo* studies, it is still traditionally used as the first step of testing (Edmondson *et al.* 2014).

Organotypic epithelial "raft" cultures are tissue culture systems that permit the full differentiation of keratinocyte monolayers via the culturing of the cells on collagen bases (i.e. dermal equivalent) at the air-liquid interface (Chow & Broker 1997). NIKS cells have the differentiation properties conserved from primary keratinocyte, which has the potential to form epithelial tissue (Allen-Hoffmann *et al.* 2000b). The raft culture is, therefore, a good model system for mimicking the epidermis. Our results first established with 2D *in vitro* system were further supported by organotypic raft culture, which suggests the system is a reliable technique for investigating basal epithelial homeostasis.

Patients with HPV-related infections suffer from symptoms that are highly refractory to treatment, lasting for years and have a high occurrence rate (Hoffman *et al.* 2017; Maw 2004; Woodhall *et al.* 2011). Persistent infection is characterised as the virus is not cleared but remains in specific cells of infected individuals. The mechanisms by which persistent infections are maintained involve both modulation of virus and cellular gene expression and modification of the host immune response. Various stimuli may trigger the reactivation of persistent infection, and host immunosuppression is one of the key factors for viral reactivation. Persistent high-risk HPV infection has been shown to constitute the most critical risk factor for the development of cervical precancers cervical intraepithelial neoplasia (CIN) grade 2 and 3 (Koshiol *et al.* 2008). To achieve effective HPV clearance, it is important to identify the infected reservoir and specifically target these cells for terminal differentiation. Based on previous research (Brimer *et al.* 2012; Kranjec *et al.* 2017) and unpublished data from our lab, we hypothesised that the HPV E6 protein plays a major role in maintaining infected cells at the basal layer of epithelia.

This section describes the construction and validation of the *in vitro* basal epithelial model system in detail. With different HPV E6 expressing cell lines, the project suggests an evolutionarily conserved role of E6 in viral persistent infection and reservoir formation.

3.2 Optimisation and Validation of FUCCI Cell Cycle Sensor System

The FUCCI system, named after the Florescent Ubiquitination-based Cell Cycle Indicator, was first described by Sakaue-Sawano *et al.* Based on the oscillating expression pattern of Cdt-1 and Geminin during the cell cycle, the red fluorescent protein mKO2 and green fluorescent protein mAG were conjugated to the two proteins respectively (Sakaue-Sawano *et al.* 2008). The activity of E3 ligase complexes SCF^{Skp2} and APC^{Cdh1} are regulated in a cell-cycle-dependent manner, and Cdt-1 and Geminin are two direct targets of these E3 ligases, respectively (Vodermaier 2004). The FUCCI-expressing NIKS cell line was previously constructed by Dr

Nagayasu Egawa, using the pBOB_EF1_FastFUCCI plasmid. Cells are fluorescently red when the cell is in the G0/G1 phase and green in the S/G2/M phase (Figure 3.2a). Despite the neomycin antibiotic selection process being successfully carried out during the FUCCI NIKS cell line construction, the initial overall expression level of FUCCI was lower than 20% (Figure 3.2b). This low expression level of FUCCI after selection might be due to the randomness of the retrovirus' integration into the keratinocyte genome, resulting in the generation of a heterogeneous group of various expression levels of FUCCI proteins. To increase the FUCCI expression level, FUCCI NIKS cells were sorted by positive mAG expression (mAG+ population in Figure 3.2b). The FUCCI expression level was checked one passage after sorting, and the total expression level had increased to around 90% (Figure 3.2c).

To validate the cell cycle indicated by the FUCCI system, FUCCI NIKS cells were stained by DRAQ5, a cell-permeable far-red fluorescent DNA dye that can be used on live cells in combination with mKO2 and mAG (Figure 3.2d). The DNA content was analysed through flow cytometry and gated by 2N and 4N. Within the 4N gating, which are cells in the late S/G2/M phase, 88.3% of cells were green. Roughly 10% of cells with 2N DNA content (cells are in the G1/early S phase) were mAG positive, which indicated that the cells were in the early S phase. Therefore, we are confident that the FUCCI system could be used for studying live cell proliferation and sorted FUCCI expressing NIKS cells were transduced with retroviral vectors to generate consistent expression cell lines.

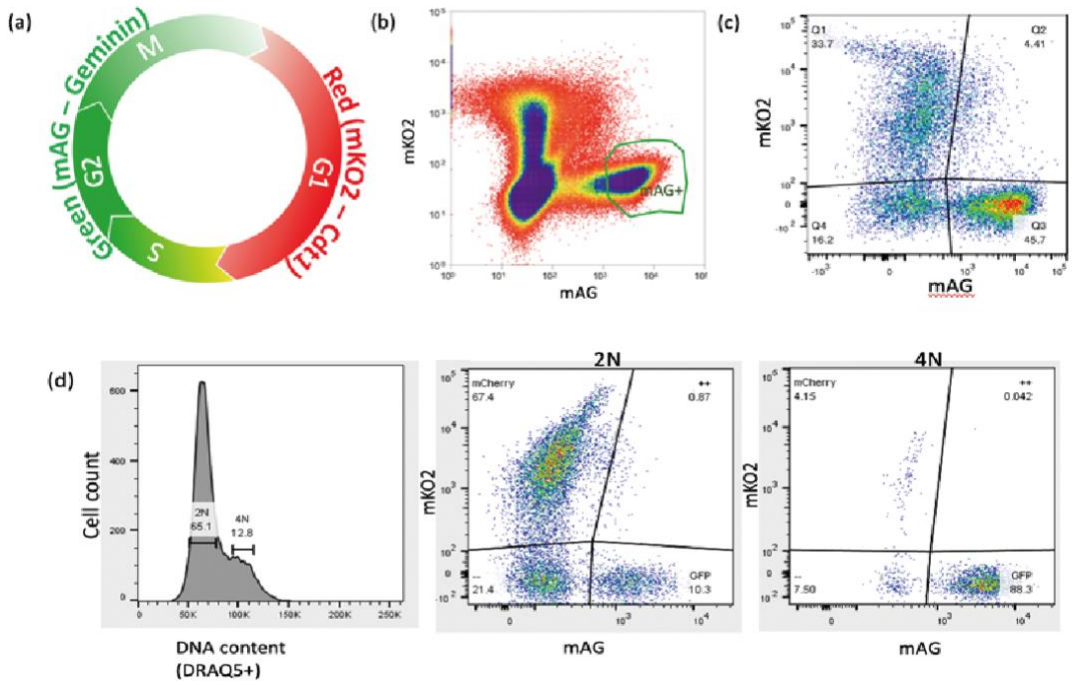


Figure 3.2: FUCCI cell line optimisation and validation. (a) Oscillating expression scheme of FUCCI cell cycle indicator. Red fluorescent protein mKO2 is conjugated with Cdt-1 marking cells in the G1 phase as red; green fluorescent protein mAG is conjugated with Geminin marking cells in the S/G2/M phase as green. (b) Flow cytometry expression level of FUCCI positive cells before sorting. The sorted population (mAG+) is roughly 10% of total cells. (c) FUCCI expression level in NIKS cells passage one after sorting for mAG+ cells. (d) Live cell DNA content was stained with DRAQ5. DNA content was gated by 2N and 4N, separating the G1/early S phase and G2/M phase. The FUCCI expression was checked after gating by DNA content.

3.3 Keratinocyte (NIKS) response to cell density (FUCCI+K10)

Although the monolayer tissue culture of NIKS cells cannot fully represent the epithelium, when NIKS cells reach confluence, they receive contact inhibition signals from adjacent cells similar to the *in vivo* setting (Allen-Hoffmann *et al.* 2000b). Therefore, monolayer assays can provide insight into how the basal layer homeostasis is affected by HPV E6 expression.

Firstly, to elucidate the normal NIKS response to cell density, FUCCI NIKS samples were collected at a different time point after seeding at the same density ($1x10^5$ cells per well on a 6-well plate). As expected from the growth assay, when the cell density increases, NIKS cell proliferation slows down. With the FUCCI system, the slowing down of cell proliferation is indicated by the increasing proportion of mKO2+ cells and the decreasing of the mAG+ population (Figure 3.3a). Even though FUCCI is a dual colour system presenting cells in G1 and S/G2/M phases differentially, for an analytic purpose, we focused on the proportion of cycling cells (%S/G2/M), which is the mAG+ population.

Meanwhile, as the cell gets confluent, NIKS start to delaminate from the culture container and pile up to form multi-layered structures as the stratified epithelium. To identify this subpopulation committed to differentiation, the suprabasal early differentiation marker Keratin 10 was used as a positive marker (Figure 3.3b). NIKS samples from different time points were intracellular stained with K10, and the Cy5 (far-red) channel was used for detection on top of FUCCI. From the result, K10 only starts after cells reach confluency.

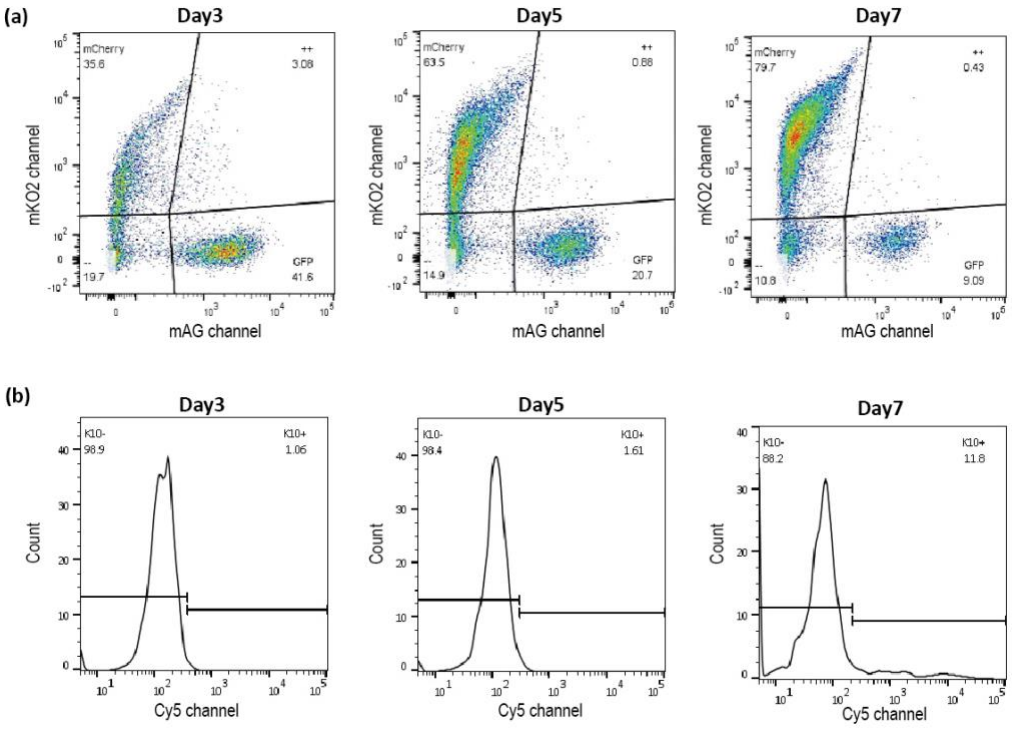


Figure 3.3: NIKS response regarding proliferation and differentiation as cell density increases. Flow Cytometry data of FUCCI (a) and K10 (b) expression when NIKS were seeded at the same density and collected at different time points.

3.4 HPV16E6 Functions in Keratinocyte by Promoting Basal Cell Proliferation and a Delay in Commitment to Differentiation

Previously, a NIKS LXSN 16E6 growth advantage over NIKS LXSN was described in our lab (Kranjec *et al.* 2017). During cell line construction, NIKS cells are transduced with retroviral vectors containing a pLXSN_16E6 plasmid. The 16E6 protein was expressed from the 5'LTR promoter in the pLXSN vector (Supplementary Figure 1). Previous data from our lab has shown that the *in vivo* HPV16 early promoter p97, and the 5'LTR promoter in the pLXSN vectors, both generated similar levels of E6 mRNA transcripts in the NIKS cells (6-12 copies per GAPDH transcript) (previous unpublished data from Doorbar's lab). Therefore, our lab uses the pLXSN plasmid vector to generate NIKS cell lines to mimic the natural expression level of 16E6 in keratinocytes.

Initially, we looked at the saturation density of NIKS harbouring empty LXSN vector and LXSN 16E6 (3.1a). To investigate the function of HPV16E6 in the basal layer of the epithelium, we first looked into the total cell count when cells are post-confluence.

To investigate this, NIKS cells were seeded at a density of $6x10^5$ cells per well in 6-well plates, allowing cells to reach confluence 48 hours after seeding. Cell numbers were counted on Day 1/2/4/6 after plating. Once cells had become confluent, the growth of NIKS LXSN slowed down as the cell numbers continued to rise (Figure 3.4.1a). For NIKS LXSN 16E6, the cell doubling rate also slowed down, but to a lesser extent when compared to NIKS LXSN. On Day 6, the cell number of NIKS LXSN 16E6 was significantly higher when compared to NIKS LXSN (p-value < 0.01). This result suggests that a function of 16E6 in keratinocytes promotes cell cycle entry in the basal epithelium even when cells are at high cell density.

For long-term HPV persistent infection, infected cells are prone to stay at the basal epithelium layer; therefore, we hypothesised that E6 also functions in the keratinocyte differentiation scheme. NIKS LSXSN and NIKS LXSN 16E6 were seeded at the same cell density and cultured until post-confluence for analysis. Immunocytochemistry for K10 shows that when cells are at post-confluence, NIKS LXSN starts to express K10, indicating cells are committing to terminal differentiation. Conversely, for NIKS LXSN 16E6, no K10-positive cells were observed (Figure 3.4.1b), suggesting a delay in terminal differentiation. To quantitatively analyse this, we did a K10 intracellular staining analysis with Flow cytometry. NIKS LXSN were firstly gated by K10 positive and negative populations. However, from flow cytometry, we observed that 23.0% of NIKS LXSN and 14.0% of NIKS LXSN 16E6 population were K10 positive at post-confluence (Figure 3.4.1c). This result was much higher than seen with immunofluorescence. The AxioVision fluorescent microscope only generates 8-bit images, and the dynamic range is very narrow compared with FACS. Therefore, we decided to have an additional gate which distinguished between high K10+ and low K10+ populations. In NIKS LXSN, about 4.5% of cells express K10 at a high level, similar to our observation with the fluorescent microscope. HPV16E6 can delay NIKS cells' commitment to differentiation, and this phenotype is revealed by the downregulation of both total K10+ and high K10+ populations.

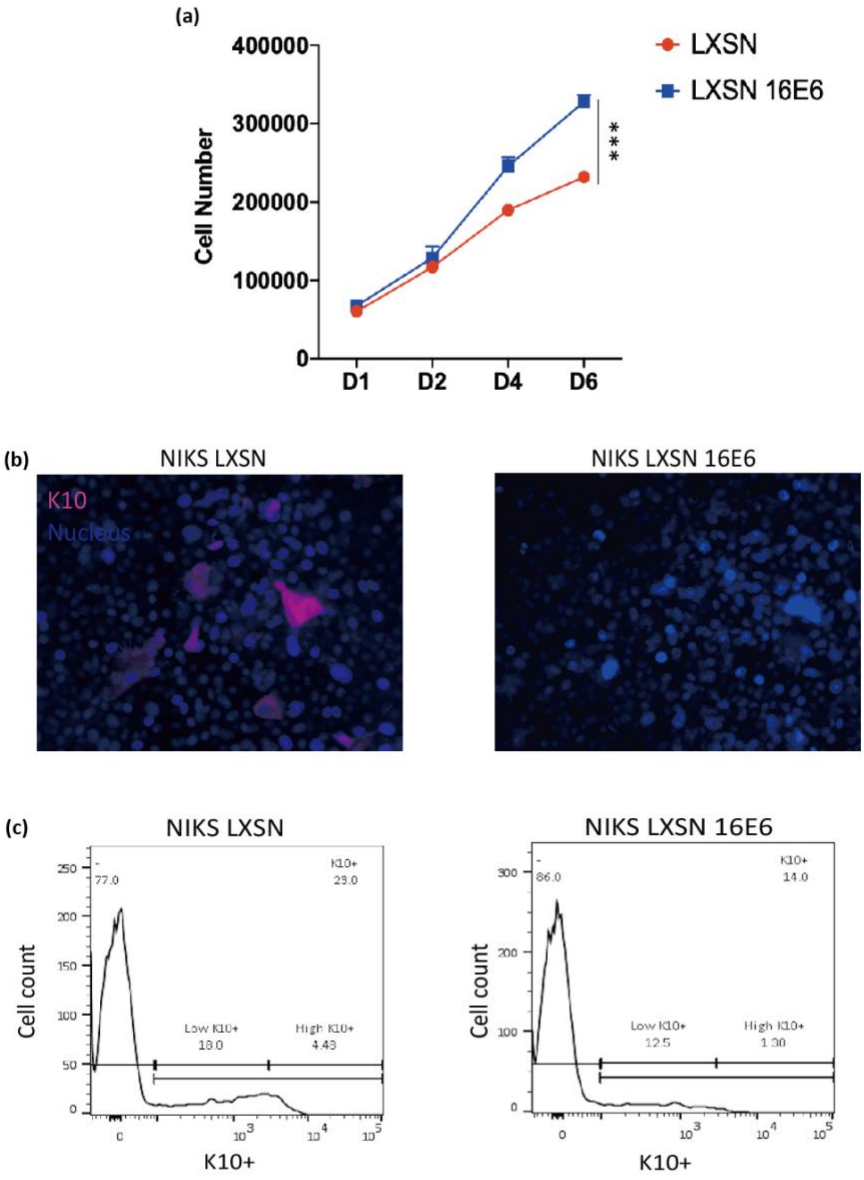


Figure 3.4.1: Confirmation of HPV16E6 function on NIKS cells in cell growth rate and differentiation. (a) Growth assay over six days. Cells were seeded at $6x10^5$ on each well on 6-well plates. (b) Immunofluorescence of NIKS LXSN and NIKS LXSN 16E6 at post-confluence,

stained for Keratin 10 and nucleus. (c) Flow cytometry of post-confluence NIKS LXSN and NIKS LXSN 16E6, gated for total K10+ and low/high K10+ populations.

HPV16E6 effects on keratinocytes depend on the downregulation of p53 to modify the host cell environment and facilitate the viral life cycle (Thomas *et al.* 1999b). To confirm this function of my NIKS LXSN 16E6, cells were seeded at $1x10^5$ cells per well on 6-well plates, and the lysate was collected in a RIPA buffer at different times (Day 3/5/7/9) after seeding. Downregulation of the total p53 level by 16E6 was seen in NIKS expressing 16E6 compared with NIKS LXSN at different cell densities (Figure 3.4.2a). The total Notch receptor protein level was also reduced in NIKS LXSN 16E6, and the p53-dependent expression of Notch was confirmed in our lab previously, with the 16E6 mutant unable to degrade p53 (Kranjec *et al.* 2017). Notch is considered an essential pathway governing epithelial homeostasis, the activation of which triggers the terminal differentiation scheme (Watt *et al.* 2008). With the 16E6 mutant, the expression of K10 was salvaged in NIKS to a similar level as the NIKS LXSN (Kranjec *et al.* 2017). Therefore, the degradation of p53 and downstream Notch receptors by HPV16E6 is considered to be important for the phenotypes of driving keratinocyte proliferation and preventing cellular commitment to differentiation.

Additionally, from the result, p53 and Notch protein level is naturally regulated by cell density (reflected as time after seeding in Figure 3.4.2a). With the growth assay (Figure 3.4.1a), a longer culture time for NIKS LXSN and NIKS LXSN 16E6 will lead to a greater difference in cell density. To reduce variation in cell density and for experimental reproducibility, we decided to seed the FUCCI NIKS cells at $1x/4x/8x10^5$ cells per well in 6-well plates and collect all samples at the same time. The culture media was changed at 24 hours, and samples were collected 72 hours after plating. The different seeding density 1x/4x/8x represents cells at pre-/sub-/post-confluence, respectively (Figure 3.4.2b). With the range of seeding density, we aim to mimic the wound-healing (pre-confluent) and homeostatic (post-confluent) regulated by HPV E6 expression.

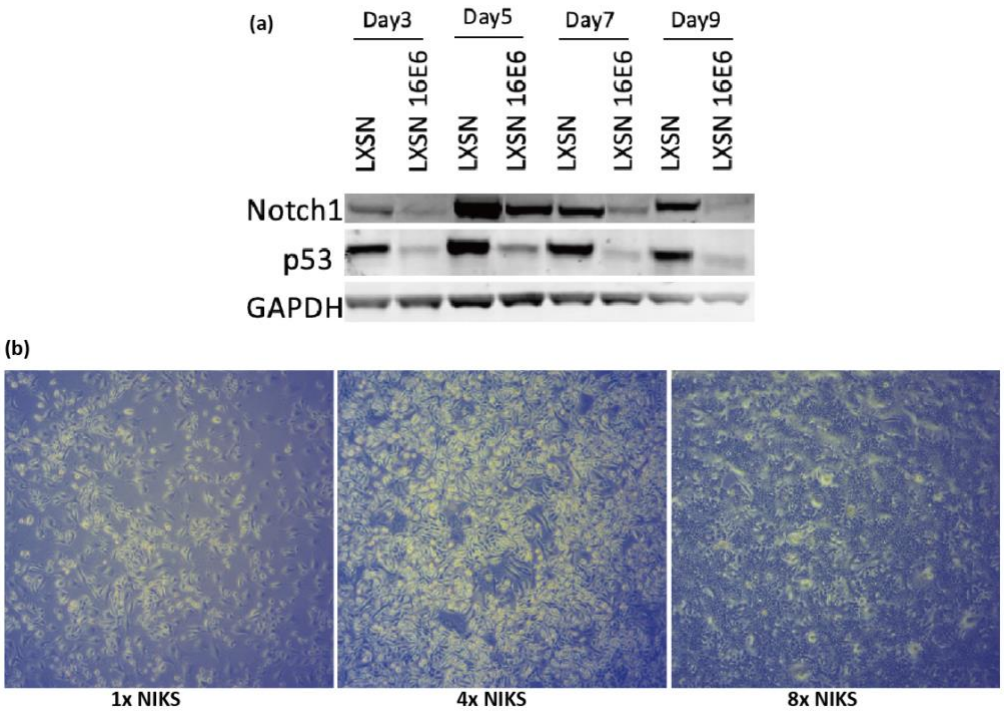


Figure 3.4.2: Density regulated NIKS phenotype and experimental setting. (a) Western blot of FUCCI NIKS LXSN and NIKS LXSN 16E6, stained for protein content of Notch1, p53 and GAPDH. Cells were seeded at $1x10^5$ cells per well on 6-well plates, and samples were collected at different time points. (b) Light microscope images of NIKS LXSN before collection for analysis, which is 72 hours post seeding. Seeding density of 1x/4x/8x represents the pre-/sub-/post-confluence status of keratinocytes.

3.5 Finalising the *in vitro* model of basal epithelium and choosing a data analyser

Assembling these results, we finalised our *in vitro* basal epithelial layer model with NIKS with cell density; the populational proportion of proliferation (FUCCI) and commitment for differentiation (K10) can be measured simultaneously to indicate keratinocyte homeostasis and modulations by HPV E6. Also, experiment conditions were set as seeding at a range of cell densities; data was collected 72hr after seeding for more comparable data between samples. Under this condition, cells were cultured in FC media (which contains a high amount of EGF) for 48hr which cells consumed most of the EGF and reached a steady state. For large-scale

data acquisition and analysis, especially for High-Throughput Screening (HTS) of potential anti-viral agents, FACS and cell count assay was not suitable. In this study, we applied high-content analysers ArrayScan XTI (ThermoFisher) and Opera Phenix (Perkin Elmer) and carried out further experiments in 96-well plates.

3.5.1 ArrayScan

With ArrayScan, our data was acquired with a 10x SemiApoChromat objective lens. The lens has a basic correction for red, green and blue index mismatch and is made of Calcium Fluorite, giving low chromatic aberration. A total image set of 21-fields was taken from each well on 96-well plates (Figure 3.5.1a/b).

The nucleus was marked with DAPI and used for cell density measurements. Data from each cell was collected, and the total intensity of red (TotalIntenR) and green (TotalIntenG) fluorescence levels for each cell were plotted using R studio (RStudio Team 2015) with a log axis (Figure 3.5.1c). Two distinct populations were seen in the scatter plot, indicating that the cells were in the G0/G1 phase and S/G2/M phase, respectively. Since we are only focusing on the mAG+ cycling cells, the proportion of green positive cells was calculated by pairing HCS Studio Cell Analysis Software with the threshold of TotalIntenR<1e+04 and TotalIntenG>1e+04. The analysis was carried out by normalising the percentage of mAG+ cells in FUCCI LXSN 16E6 by the percentage in FUCCI LXSN. Additionally, log2 was taken to show the fold-changes for better data representation and future analysis. The ArrayScan gives a similar FUCCI result to the FACS after normalisation (Figure 3.5.1d). The K10 positivity was obtained as the total intensity at the Cy5 channel per well (data shown in Section 4).

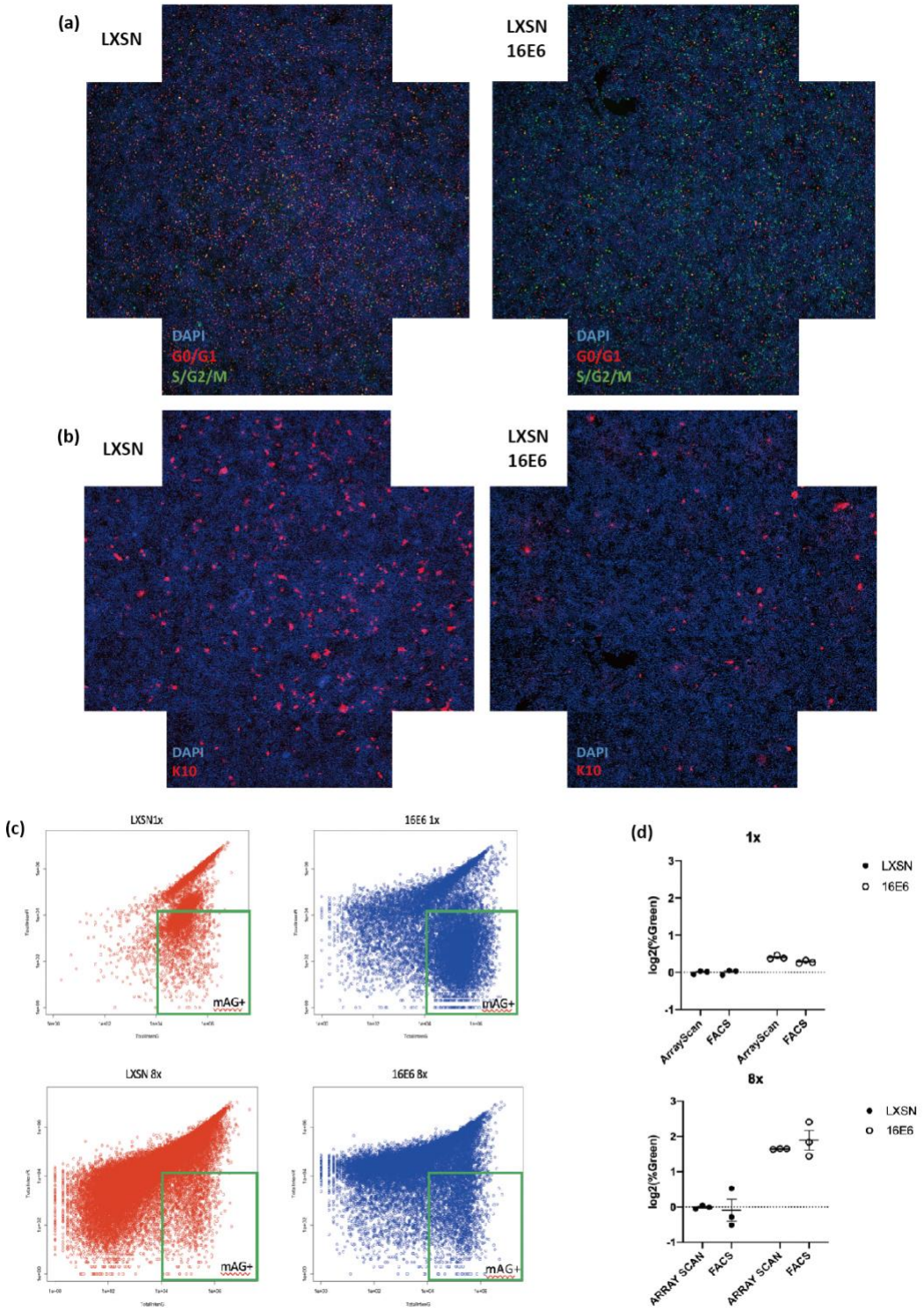


Figure 3.5.1: Comparing Flow Cytometry and Array Scan results. (a) Representative image taken from ArrayScan results refers to 21-field per well (with 10x objectives). (b) The total intensity of green and red positivity was generated by R studio. Every single dot represents a cell. Green positive cells (mAG+) were gated by TotalIntenG>1e+04 and TotalIntenR<1e+04,

proportion of mAG+ cells were also generated by R studio. (c) Summary normalised analysis data comparing the proportion of Green positive cells in LXSN and LXSN 16E6 from FACS and ArrayScan.

3.5.2 Opera Phenix

Later, as our analysis requires further cellular sensitivity, we moved our analyser to Opera Phenix, which allows high-content confocal imaging. It is equipped with sCMOS cameras, which have additional filters when collecting fluorescence signals to achieve a higher signal-to-noise ratio.

Experiments were carried out in 96-well plates and collected with 40x objective; z-stack images were obtained when necessary. All data were analysed with pairing Harmony Software. The dataset from Opera Phenix is similar to ArrayScan but presents a better separation of different fluorescent signals from individual cells (Figure 3.5.2). In addition to FUCCI/K10 analysis, Opera Phenix is also used for competition assay analysis (details described in section 3.6).

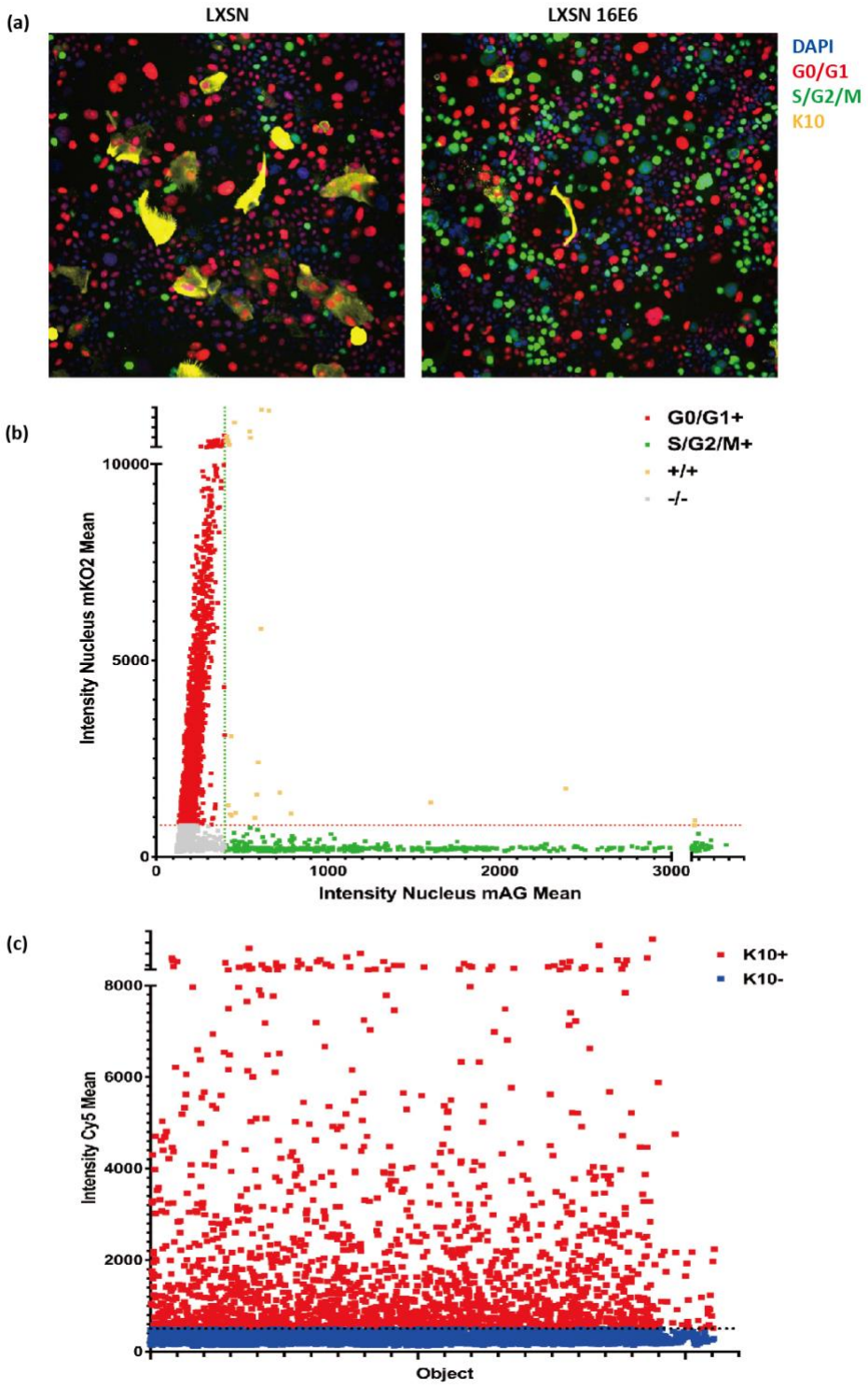


Figure 3.5.2: Opera Phenix raw data and analysis. (a) Representative image of post-confluent FUCCI NIKS LXSN/LXSN 16E6 stained with K10 (with 40x objectives). (b) Scatter plot of FUCCI signal measured as mean intensity per nucleus. Each point represents a single nucleus, with a threshold identifying the G0/G1 and S/G2/M population marked in a dotted line. (c) Scatter

plot of Cy5 (K10) fluorescent signal measured as mean intensity per nucleus. Each point represents a single nucleus, with a threshold identifying the K10 positive population marked in a dotted line. Gaps on the y-axis represents the high expression level of the fluorescent signal, which separation from the majority of signals shows better graphical illustration.

3.6 HPV16E6 facilitates viral reservoir formation

To further evaluate the significance of 16E6-mediated homeostatic regulation for the viral life cycle and resemble the real infection where uninfected and infected cells are actively interacting with each other, we developed a mixture growth assay referred to as the competition assay.

NIKS LXSN and LXSN 16E6 cell lines were seeded at high density (8x seeding density) on 96-well plates to mimic the confluent basal layer at the starting point.

3.6.1 The tissue culture system is comparable with the basal/suprabasal layer of epithelium

Expression of keratin markers K5 and K14 are restricted to the basal layer, while K1, K10, and involucrin indicate a differentiated phenotype (Fuchs 2003). Firstly, we confirmed that NIKS is a comparable model for studying basal to spinous transition by K14 and K10 expression profiles in our 3D system that can be maintained for up to 9-day *in vitro* and raft culture (Figure 3.6.1). In later experiments, we mostly focused on the bottom layer of the *in vitro* culture, which represents the basal epithelium layer, to discuss HPV E6-mediated viral persistence.

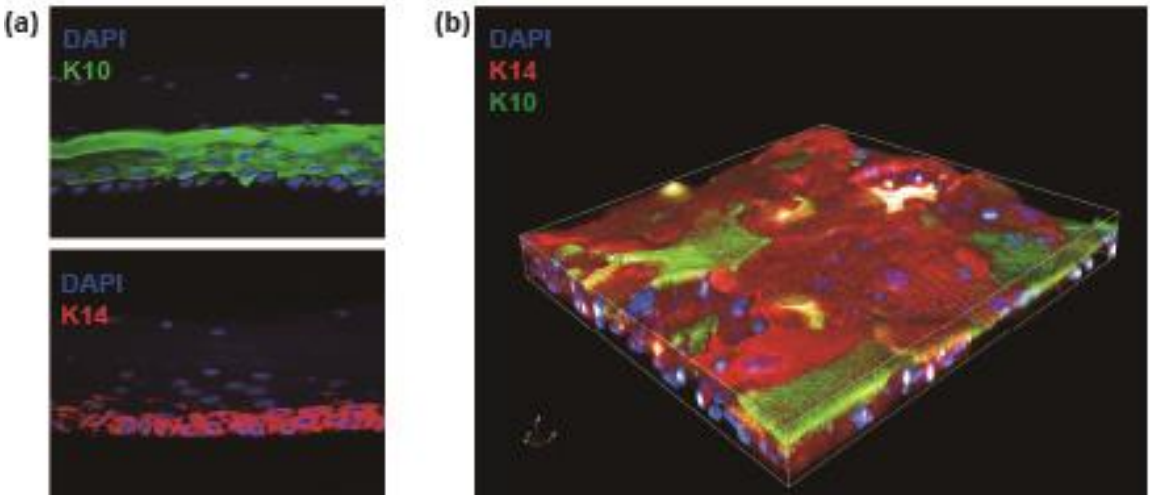


Figure 3.6.1: Organotypic raft culture images (a) and z-stack of multi-layered *in vitro* culture of NIKS seeded at very high density maintained for 7-day (b) stained for K14 and K10. Images captured with 40x objectives by AxioVision fluorescent microscope and Opera Phenix, respectively.

3.6.2 HPV16E6 expressing NIKS preferentially occupies the basal layer of epithelium

Fluorescent markers eGFP/mCherry were used to distinguish different NIKS populations. When seeded at a 1:1 ratio, eGFP LXSN and mCherry LXSN has no growth advantage over each other; random distribution of eGFP and mCherry signal is observed in the bottom layer after 7-day culture (Figure 3.6.2a). As a result of no difference between populational proliferation and differentiation probability of eGFP LXSN and mCherry LXSN, the percentage of eGFP and mCherry NIKS remains the same as Day1. However, when seeding eGFP LXSN and mCherry LXSN 16E6 at the same ratio, the mCherry population gradually expands and colonises the bottom layer of the mixed culture, which after 7-day culture, mCherry LXSN 16E6 expands from 46.36% to 92.70% of the total basal population. The competitive advantage phenotype of mCherry LXSN 16E6 is further confirmed with organotypic raft culture, where the eGFP and mCherry NIKS were seeded at a 10:1 ratio (Figure 3.6.2b), where the mCherry 16E6 expands and occupies the basal layer after 14-day culture. When there is no competition between the eGFP and mCherry population, the mCherry cells acquire the cell fate to proliferate and differentiate equally to maintain epithelial homeostasis. As a result, mCherry LXSN forms a

column surrounded by eGFP LXSN cells. In contrast, as mCherry LXSN 16E6 cells have increased proliferation, less delamination and delayed differentiation, they occupy the basal layer of epithelium and eGFP LXSN were outcompeted to differentiate to upper layers.

To minimise the technical variations between independent replicates, the percentage of mCherry NIKS was normalised by the Day1 result depending on cell type (Figure 3.6.2c). By collecting data every 24hr after seeding, when seeding at confluency, the two NIKS LXSN cell lines together reached a steady state. While if the E6 expressing cells are cultured together with empty vector expressing cells, the basal homeostasis changes until they reached a new steady state set by E6 modulation.

Moreover, with the Harmony software, we were able to calculate cell body size based on the nucleus stain and cell fluorescence (Figure 3.6.2d). Firstly, as expected, there was no significant difference in the average cell body size between eGFP LXSN and mCherry LXSN on Day7. While in the eGFP LXSN:mCherry LXSN 16E6 group, the 16E6 population colonised the bottom layer and cell density increased (data not shown). As a result, the average cell body size of mCherry LXSN 16E6 is significantly lower than the mCherry NIKS LXSN cells (Figure 3.6.2e). It is also interesting to note that the eGFP LXSN is significantly smaller than the control group eGFP LXSN and mCherry LXSN 16E6. The phenotypic change is considered due to the eGFP LXSN cells being extruded towards the upper layers by mCherry 16E6; therefore, the attachment area at the bottom layer will be smaller.

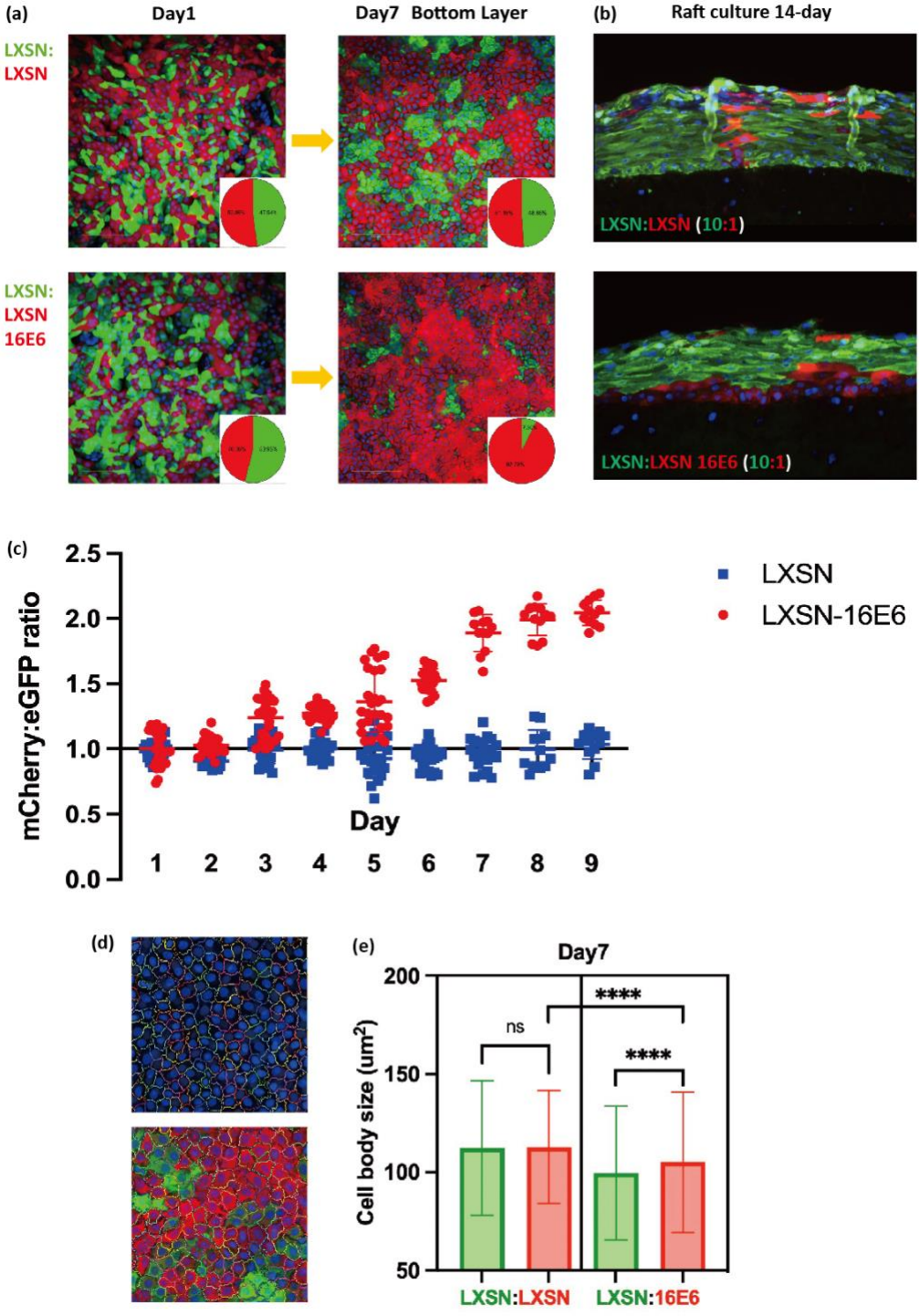


Figure 3.6.2: HPV16E6 expressing NIKS preferentially occupies the basal layer of epithelium. (a) The same number of eGFP LXSN and mCherry LXSN/LXSN 16E6 were seeded at confluency, with representative images of Day1 and Day7 bottom layer. The pie chart shows relative eGFP and mCherry cell percentages in the images. Images captured with 40x objectives. (b)

Organotypic raft culture of eGFP LXSN and mCherry LXSN/LXSN 16E6 seeded at 10:1 ratio. Paraffin-embedded slices stained for eGFP and mCherry expression. Images captured with 20x objectives. (c) Competition growth assay of mCherry:eGFP ratio, each point represents a replicate. All data normalised by Day1 mCherry/eGFP ratio of the corresponding cell line. Data are shown as mean±SD. (d) Representative image of cell recognition for analysing cell body size, with nucleus (DAPI) only (top) and eGFP/mCherry fluorescence (bottom). (e) Mean +/- SD of cell body size of the different populations at Day7 bottom layer (One-way ANOVA, P-value style as GP, **** indicates p-value <0.00009). Data are shown as mean±SD.

3.7 Functionally conserved E6 role in HPV persistence

While the carcinogenic high-risk HPV only accounts for a small subgroup of the total HPV types, HPVs have evolved to infect and thrive at many different body sites by adapting to local epithelial conditions, which adopted distinctive epithelial niches and mechanisms of site-specific immune regulation (Egawa & Doorbar 2017). Generally, both low- and high-risk HPV diseases can persist for years, are highly refractory to treatment and are recurrent. Hence, we proposed that E6 might be an evolutionarily conserved basal epithelial homeostasis regulator that has co-evolved with the host to adapt unique epithelial niche. Therefore, studying HPV E6 protein has great potential to reveal specific molecular pathways governing local keratinocyte proliferation and differentiation.

3.7.1 Construction of HPV E6 Expressing Cell Lines

To test our hypothesis that different HPV types utilise a common E6-mediated mechanism in altering the keratinocyte life cycle, cell lines expressing E6s from significant clinical representatives of different HPV groups were generated (low-risk cutaneous μ-HPV1, low-risk cutaneous α-HPV27, low-risk mucosal α-HPV11, high-risk mucosal α-HPV16, low-risk cutaneous β-HPV8 and low-risk cutaneous γ-HPV65).

HPV E6 sequences were amplified from HPV genomic plasmid or previously constructed E6 harbouring plasmids and incorporated into destination vectors pLXSN and pQCXIN respectively by the Gateway® cloning system (Figure 3.7.1). Genomic HPV E6 DNA size is around 500bp, which was confirmed in all recombination steps by digestion of restriction enzyme BsrGI. The sequence was checked with entry vectors (pENTR) to confirm the correct insertion of HPV E6 sequences. The endogenous expression level of E6 was achieved by cloning the E6 sequence into the pLXSN vector backbone, which uses a promoter which resides in the retrovirus 5'LTR region. Meanwhile, constructing pQCXIN_E6 plasmids with a CMV promoter allows for a higher expression level in target cell lines (Qin *et al.* 2010). The reason for designing two different expression vector backbones (pLXSN and pQCXIN) was to enable an evaluation of the effect of expression levels of different E6 proteins on keratinocytes. Although it is not quantified in NIKS cells, the expression level of the CMV promoter in the pQCXIN vector is generally higher than the 5'LTR promoter used in the pLXSN vector (Qin *et al.* 2010). To date, all my experiments were carried out with pLXSN_E6 vectors to examine the effect of endogenous expression of E6 protein on NIKS cells. Empty vector pLXSN was used as a negative control.

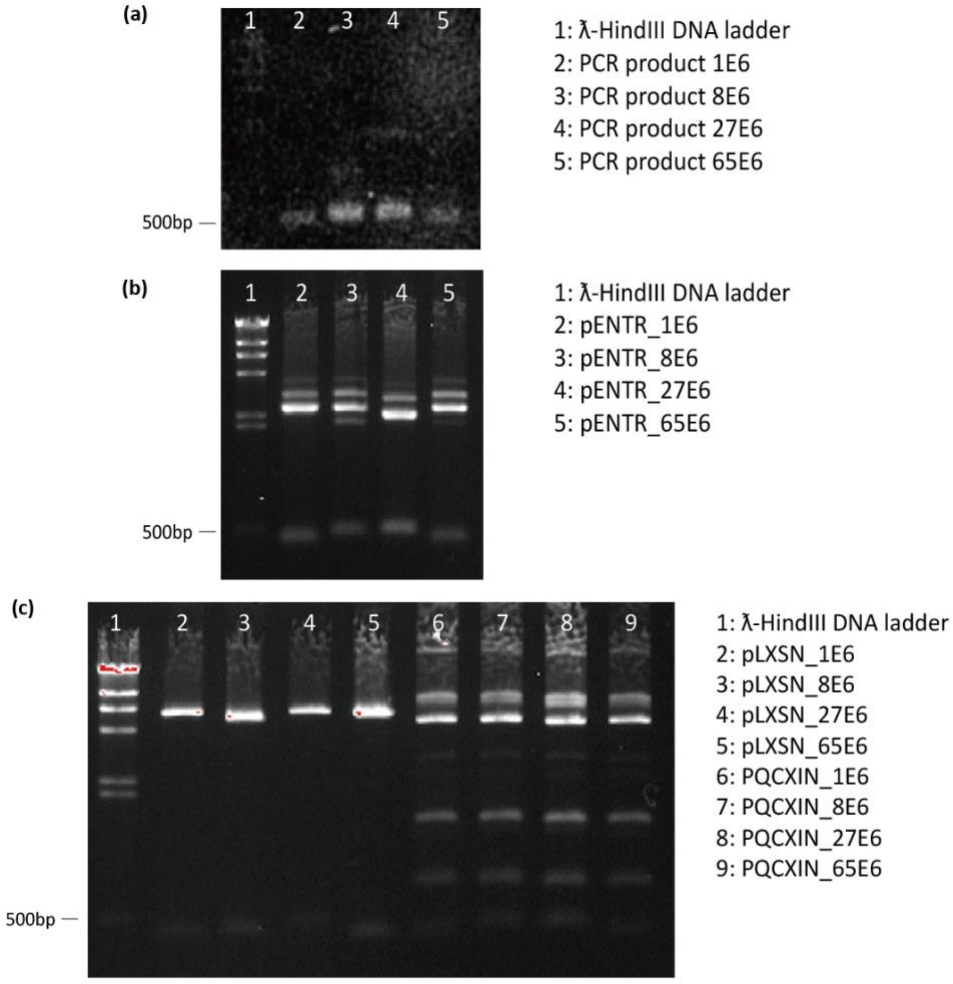


Figure 3.7.1: Gel electrophoresis of (a) PCR amplicon of HPV E6s; (b) restriction enzyme BsrGI digested entry vector (pENTR) after BP recombination reaction; (c) restriction enzyme BsrGI digested expression vector (pLXSN and pQCXIN) after LR recombination reaction. The expected DNA size of E6 is around 500bp.

3.7.2 HPV E6 generally alters basal epithelial homeostasis and promotes viral persistence

FUCCI NIKS LXSN HPVE6 cell lines were seeded at a range of different densities and allowed to grow for 72hr before collecting data (Figure 3.7.2a/b/c). Generally, NIKS LXSN has 30% of the population that is active in the cell cycle, with the percentage decreasing when cell density increases (Figure 3.7.2a). This is referred to as contact inhibition of proliferation (Ribatti 2017).

When NIKS LXSN fully occupies the culture container, the population reaches a balance between proliferating and delaminating cells indicated by the number of cells remaining the same at the bottom layer while still 5% of the population is in the cycle (S/G2/M phase) (Figure 3.7.2a/c). Then the NIKS LXSN start to express early differentiation marker K10 after reaching confluency (Figure 3.7.2b).

In general, with HPV E6 expression NIKS has an increased proliferation rate at low- and/or high-cell density resulting in higher saturation density and delayed differentiation with different strategies adopted by different HPV types according to their niche.

For the α-group of HPV E6, it is obvious that the high-risk HPV16E6 has the most significant function in driving cell cycle entry at all densities and resulted in a roughly 20% increase in the population saturation density; this data is compatible with our growth assay results from section 3.4.1. Moreover, the differentiation scheme is largely impaired by HPV16E6, shown as a massive delay in K10 expression. The low-risk α-HPV11 and α-HPV27 E6 possessed similar functions, which also drove overall NIKS proliferation and delayed differentiation after confluence but to a lesser extent than the high-risk HPV16E6. It is also notable that the saturation density of HPV27E6 did not significantly differ from the LXSN group while the HPV11E6 has a roughly 8% increase; this is likely due to the viruses having specific regulatory adaptions to cutaneous and mucosal epithelial sites respectively.

Previously the μ-/β-/γ-HPV groups are often compared together due to similar infectious sites, this homogenous model using NIKS helped us to dissect the individual functions in more detail. Firstly, for the μ-1E6, it is more likely that the virus-infected host cells expand during the wound healing process as the proliferation rate is significantly higher in low-density conditions but not after confluence. Also, the differentiation modulation by 1E6 is less compared to other HPV types under the same conditions. Although the low-risk γ-65E6 reached a similar saturation density as high-risk α-16E6, it failed to maintain its advantage and reached a balanced proliferative potential similar to the LXSN (roughly 8% of the S/G2/M population). For β-8E6, which belongs to the only subgroup that has carcinogenic relations among the

cutaneous low-risk HPV types, the virus does not affect cell proliferation or saturation density. The most pathogenic function is thought to be done by modulating the normal keratinocytes differentiation scheme, which is observed as basically no K10 positivity after confluence; this phenotype is further elucidated in section 3.7.3 with organotypic raft culture.

Next, competition assays were used to compare the relative fitness of infected HPVE6 expressing cells versus uninfected LXSN cells. Differently fluorescently labelled eGFP LXSN and mCherry LXSN HPVE6 cells were seeded at a 1:1 ratio to confluency. Mixed cultures were maintained for 7-day *in vitro,* and bottom layer mCherry percentage was counted. By normalising %mCherry cells at the bottom layer on Day7 with %mCherry cells on Day1 of the same cell group, a general hypothesis is that HPV E6 expressing NIKS have a significant growth advantage over uninfected cells (Figure 3.7.2d).

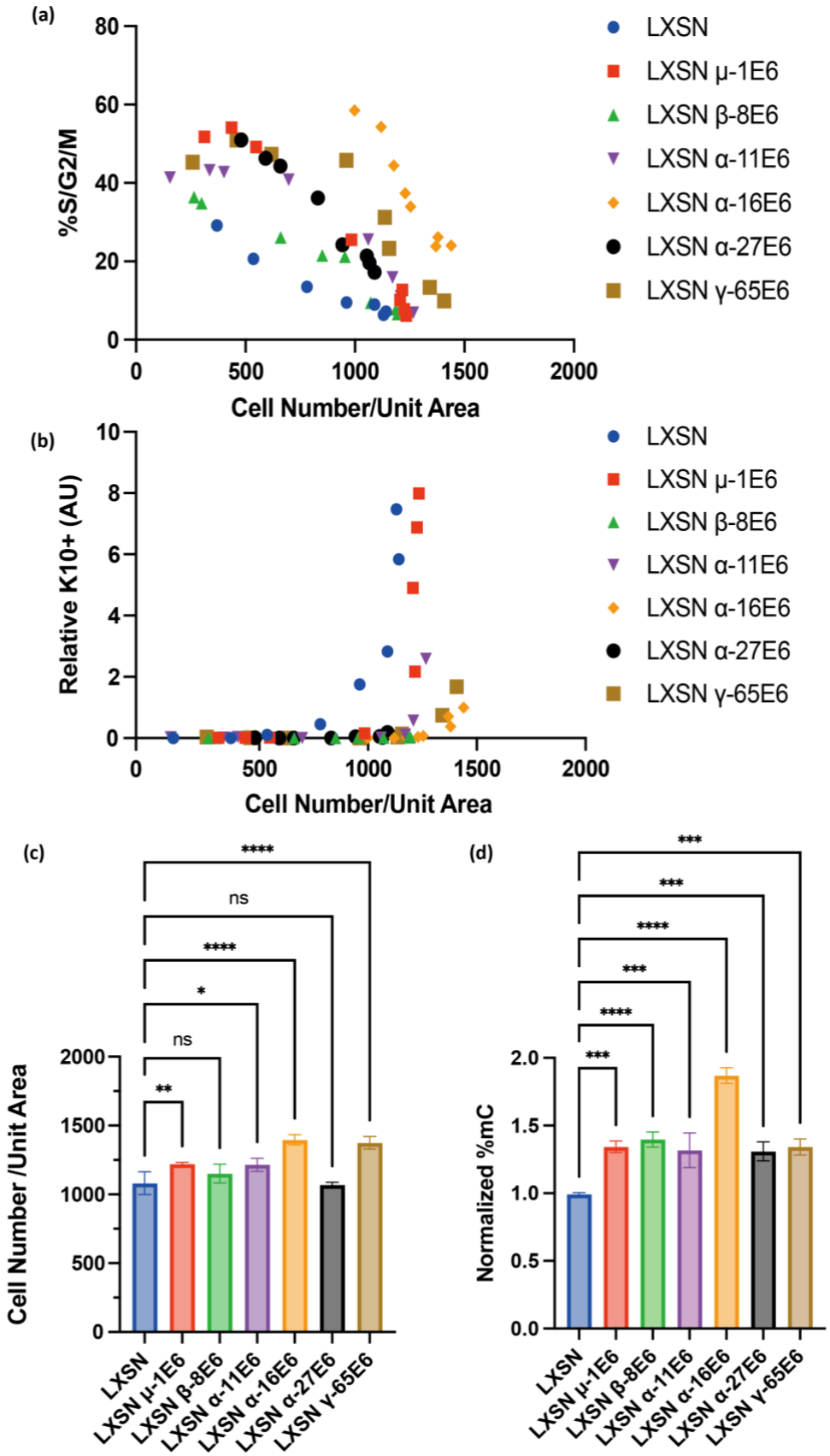


Figure 3.7.2: HPVE6 modulation to basal epithelial homeostasis. (a) Percentage of S/G2/M positive cells of FUCCI NIKS with different HPV E6 expression in response to cell number per

unit area (420,000μm^2). (b) Relative populational K10 positivity of FUCCI NIKS with different HPV E6 expression in response to cell number per unit area (420,000μm^2). (c) Saturation density (cell number per unit area (420,000μm^2)) of FUCCI NIKS with different HPV E6 expression with One-way ANOVA analysis in relation to FUCCI NIKS LXSN (P-value style as GP). Data are shown as mean±SD. (d) The same number of mCherry LXSN HPVE6 and eGFP LXSN were seeded at confluency. Relative mCherry LXSN HPVE6: eGFP LXSN ratio at Day7 bottom layer, all data normalised by Day1 of each experimental setting. One-way ANOVA analysis in relation to LXSN group (P-value style as GP). Data are shown as mean±SD.

3.7.3 Organotypic raft culture shows delayed keratinocyte differentiation induced by HPVE6

Although the results obtained from *in vitro* 2D cell culture studies may not always be translated to *in vivo* studies, it is still used as the first step of testing (Edmondson *et al.* 2014). Thereafter, we expanded our results to a 3D organotypic raft culture representing the tissue level.

The results were first analysed under a light microscope to see the morphological differences between epithelia expressing different HPV E6s (Figure 3.7.3a). For most of the NIKS cell lines expressing HPV E6s (except 8E6), the formation of a multi-layered epithelium was successful, with all different layers ranging from the basal to cornified layers being easily distinguishable by morphology with Hematoxylin and Eosin (H&E) staining. In a normal NIKS LXSN raft, cells from the basal layer become differentiated as they move into the upper layers. Eventually, in the cornified layer, very few cell nuclei can be seen as all the cells are fully differentiated and cell death is activated. The morphology of the cornified layer was altered in the NIKS raft with HPVE6 expression, and a significant increase in the number of nuclei present in the cornified layer was observed. This retention of nuclei in the cornified layer is termed parakeratosis, which is usually associated with an increased cell turnover in disease (including non-viral papilloma). While for NIKS LXSN 8E6, the differentiation scheme is completely disrupted with

no obvious differentiated layers can be observed. The results from raft culture are compatible with our FUCCI/K10 system, indicating increased general thickness and less cell terminal fate by HPV E6 expression. It is also important to note that the basal layer cell arrangement is affected by HPV E6 expressions. For NIKS LXSN cells, basal layer cells are mostly distributed evenly, and the cell body spreads along the basal lamina. While with HPV E6 expression, the basal layer composition became dysregulated, cell density increased, and cell morphology changed to vertical in shape.

The expression of K10 in NIKS LXSN starts from the suprabasal layer and diminishes before the cornified layer (Figure 3.7.3b). The most outstanding NIKS LXSN 8E6 hardly generate any K10 expressing cells, in agreement with the FUCCI/K10 results. With the other HPV E6 types, as the complete stratified epithelium structures are formed, K10 expression begins at the suprabasal layer of individual settings. For instance, in NIKS LXSN 1E6 raft, it is first noticed that the K10 negative layer is thicker compared to the NIKS LXSN. However, when considering the H&E result, the 'delayed' K10 expression is explained by the dysregulation of the basal layer. This is important as a low level of HPV during persistent infection does not completely abort the infected cell's ability to differentiate, as the viral life cycle is dependent on the host differentiation scheme, but E6 function to preferably stay at the basal layer for the establishment of long-term infections.

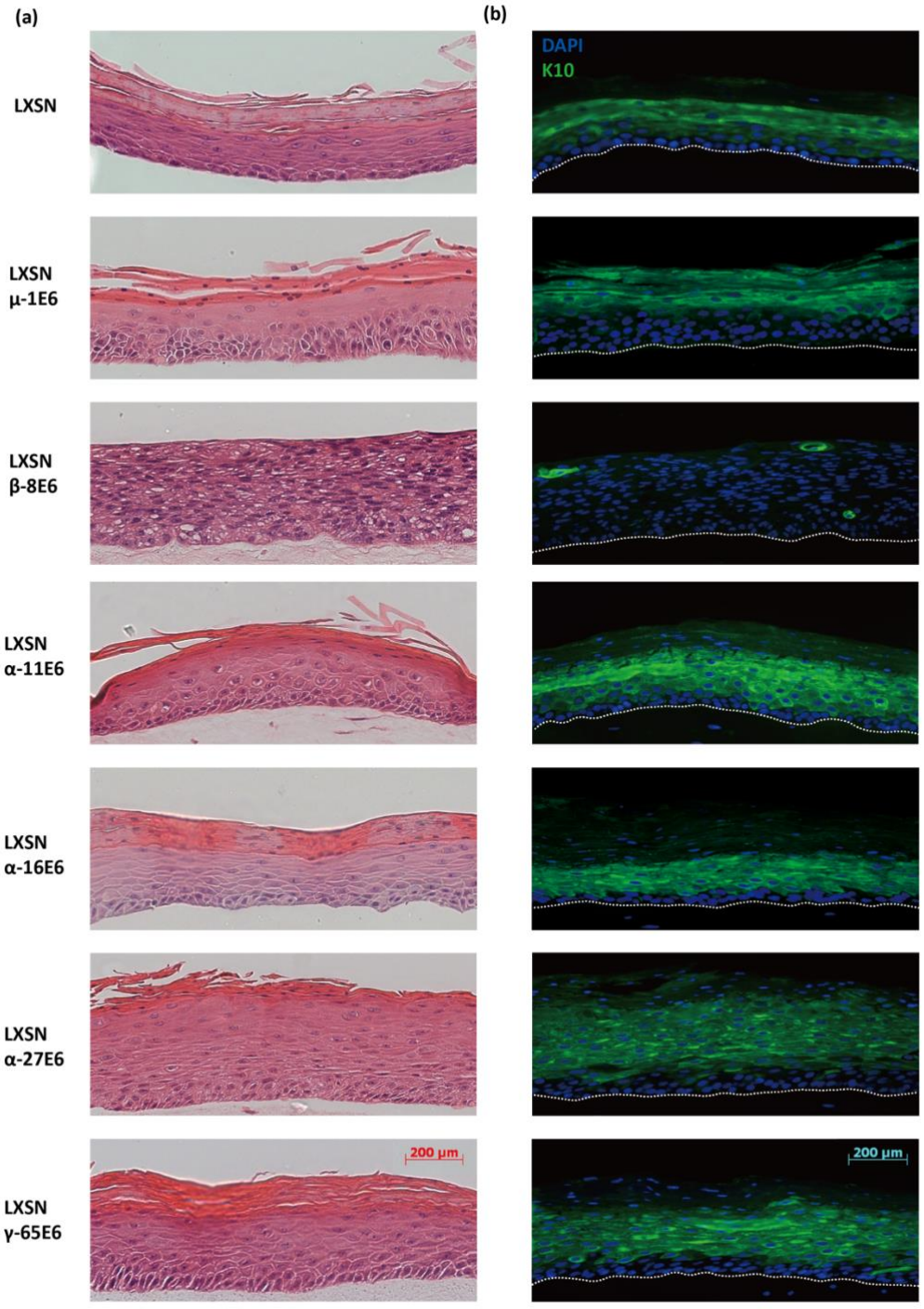


Figure 3.7.3: Organotypic raft culture of NIKS HPVE6s. (a) H&E sections; (b) Paraffin-embedded slices stained for K10 and DAPI (dotted line indicates the basal lamina attachment). All images were collected by 20x objectives.

3.8 Discussion

In this section, we aimed to develop a model system aiming to investigate basal epithelial homeostasis *in vitro* with Normal Immortalized Keratinocytes (NIKS) as a standard cell type to dissect HPV viral modulation during persistent infections. A preliminary cell count growth assay indicated the NIKS population doubling rate decreased when cell density increased, and finally reached a steady state. To further elucidate this phenotype, in addition to cell density counting, we applied the FUCCI system to distinguish between proliferating (S/G2/M phase) and resting (G0/G1 phase) cells. Consistent with the growth assay result, the percentage S/G2/M population decreased when cell density increased and was finally maintained at a 5% populational proliferation rate. Since the population is still cycling at the saturated rate, excess NIKS are delaminated from the bottom and pushed to the upper layers, forming a multi-layered structure in dish. This subpopulation is committed to differentiation and can be marked with K10 expression.

Based on the growth advantage HPV16E6 confers on NIKS cells, a higher cell number in NIKS LXSN 16E6 from the growth assay experiment can be shown as an increase in the proportion of proliferating cells (%mAG+) using the FUCCI system. FUCCI and K10 expression levels were firstly quantified by Flow cytometry, and then with two high-content analysers ArrayScan and Opera Phenix all exhibited similar levels in terms of the ratio proportion of mAG+ and K10+ cells between NIKS LXSN and NIKS LXSN 16E6, suggesting that the FUCCI/K10 NIKS will be suitable as a standardised system for detailed analysis of HPV E6 modulation in host homeostasis.

Under real infection situations, infected cells are in contact with uninfected cells at the basal layer, therefore it is also important to model the relative competition between two populations *in vitro*. With the confocal imaging system, we were able to dissect the multi-layered structure formed after confluence. By seeding eGFP and mCherry cells expressing either LXSN or 16E6 at confluency, the relative percentage of each population at the bottom layer of the mixed culture can be calculated along a 9-day timeline. While the eGFP LXSN and

mCherry LXSN cells have no statistically significant difference between cell bottom layer occupancy, mCherry LXSN 16E6 gradually expands to colonise the bottom layer until it reaches a new balance.

HPV-related symptoms normally persist for years and are highly refractory to treatments; patients who suffer from disease recurrence suggest there is a viral reservoir in the epithelial that has 'stem-like' properties. Then we tested the E6 proteins of representatives from different HPV genera, which have important clinical relevance (high-risk mucosal α-HPV16, low-risk mucosal α-HPV11, low-risk cutaneous α-HPV27, low-risk cutaneous β-HPV8, low-risk cutaneous γ-HPV65 and low-risk cutaneous μ-HPV1). The effect of different HPV E6 proteins in driving cell proliferation and delaying the commitment to differentiation at different densities was validated by the FUCCI expression and K10 staining on monolayer NIKS culture. Competition assay results also support the hypothesis that HPV E6 expressing cells have significantly increased fitness to occupy the basal layer of epithelium. Morphological changes observed in organotypic raft culture, which mimics the fully differentiated epithelia, further support our hypothesis that different HPV E6 functions alter cell processes *in vivo*. Despite their difference in carcinogenicity, our results show for the first time that the low- and high-risk HPV types have evolved similar molecular strategies to regulate their maintenance and their vegetative replication to complete the viral life cycle, but that they have important differences in their regulatory control.

A comprehensive analysis has identified a distinct interaction pattern between host proteins and different types of HPV E6s (White *et al.* 2012). Specific tissue tropism was described in HPV-induced diseases (Egawa *et al.* 2015). One example is HPV1, which causes typical Myrmecia warts on the palms and soles and very rare lesions at extra-palmoplantar sites (Jablonska *et al.* 1985). Although primarily infecting genital mucosal epithelial and causing cervical cancer, infection with HPV16 also occurs at cutaneous sites resulting in invasive cancer progression (Eliezri *et al.* 1990). Collective data has been published discussing HPV16 infections at different mucosal/cutaneous epithelial sites resulting in a variety of different disease phenotypes and distinct cellular-HPV protein interaction patterns among different

diseases (Gulbahce *et al.* 2012). Therefore, under our standardised model system with NIKS, we were able to compare the phenotypic changes but were also limited to showing detailed molecular pathways which is closely related to its natural infection site.

However, based on previous literature, some general regulatory mechanisms are adopted by HPVs from similar evolutionary branches. Notch signalling promotes the commitment of human keratinocytes to differentiation and restricts their growth potential (Nguyen *et al.* 2006). For β-8E6 and μ-1E6, binding with the transcriptional co-activator MAML protein with the LXXLL sequence motif inhibits Notch-induced gene transcription (Brimer *et al.* 2012; Meyers *et al.* 2013, 2017). However, high-risk α-16E6 indirectly regulates total Notch receptor expression level through p53. In normal keratinocytes, Notch receptor expression levels are partially co-regulated by the induction or knock-down of cellular p53 (Lefort *et al.* 2007). High-risk α-16E6 caused delays in keratinocyte commitment to differentiation by downregulating the total Notch receptor expression level through p53 degradation. However, this hypothesis does not apply to other low-risk α-HPV E6 types, which also exhibited a significant inhibitory effect on the expression of differentiation markers (Figure 3.7) but did not affect total Notch expression (data not shown). One hypothesis is that low-risk α-E6 could indirectly affect cellular p53 levels and localisation by regulating cell density. Western blots had shown that the level of keratinocyte p53 decreased when cells were post-confluent. From my results, α-11E6 and α-27E6 can promote cell proliferation when cells are at a high cell density, which may indirectly downregulate cellular p53 levels and then partially affect the downstream total Notch transcriptional expression. Mass spectrometry analysis has identified the Recombination Signal Binding Protein for Immunoglobulin Kappa J Region (RBPJ) as a novel host target by both low- and high-risk α-HPVE6 proteins (White *et al.* 2012). RBPJ is a key component for the transcription factor to bind to the Notch intracellular domain (NICD) and cellular DNA regulating gene transcription upon Notch activation (Kopan & Ilagan 2009a). Although there is currently no established function for the α-HPVE6 and RBPJ, this interaction might explain low-risk α-HPVE6 inhibition of the Notch-signalling pathway resulting in a delay in the keratinocyte commitment to differentiation. Moreover, p300 is also a co-activator involved in the RBPJ DNA-binding complex and downstream gene transcription

(Oswald *et al.* 2001). While the association of α-11E6 and p300 could indirectly downregulate the cellular p53 level (Zimmermann *et al.* 2000), it may also serve a role in Notch-induced differentiation. Only a few studies have explored the biological activities of γ-HPVs in cell-based or animal models. Isolation of γ-HPVs from cutaneous skin was often regarded as normal skin viral flora (Foulongne *et al.* 2012), and total γ-HPV relevance in common warts is less than 3% with HPV65 as one of the most frequent types associated with cutaneous lesions (Bruggink *et al.* 2012). However, recent sequence alignment demonstrated the similarity of γ-HPV E6 with α/β-HPV E6s, and putative binding sites for E6AP and MAML1 were suggested (Grace & Munger 2017).

By investigating the broadly evolutionary function of HPV E6 protein at the basal epithelial level, we hope to develop a model of epithelial homeostasis using HPV as a tool and possibly discover a therapeutic target that can widely target the HPV persistent population for effective clearance. Amongst all HPV E6 types tested, α-16E6 has the most robust function in the modulation of basal keratinocytes homeostasis. In the next sections, for high-throughput screening for small molecules targeting persistent HPV infections (Section 5) and molecular analysis identifying critical basal-suprabasal switch (Section 6), it is more technically feasible to work with the α-16E6 population as a starting point.

4. High-Throughput Screening (HTS) of small molecule inhibitors against HPV16E6

4.1 Introduction

Generally, three major HTS systems are used for HPV drug discovery studies; *in silico*, RNAi and small molecule screenings. Among these methods, *in silico* fragment-based drug design is also known as virtual high throughput screening, which incorporates computational methods to help the design of new drug candidates more rapidly and at lower costs. *In silico* drug design consists of a collection of tools helping to make rational decisions at the different steps of the drug discovery process, such as the identification of a biomolecular target of therapeutical interest, the selection or the design of new lead compounds and their modification to obtain better affinities, as well as pharmacokinetic and pharmacodynamic properties (Phatak *et al.* 2009; Zoete *et al.* 2009). To date, most *in silico* studies aimed to identify new inhibitors for disrupting high-risk E6-E6AP interactions and the elevation of p53 expression level was used as the marker for successful inhibition (Celegato *et al.* 2020). The HPV E6 protein has been proposed as a good drug target because it poses a special structure that cellular LXXLL-binding proteins do not have (Zanier *et al.* 2013b). Homology modelling and molecular dynamic simulations were used to obtain a set of HPV-16 E6 protein structures. Then, *in silico* screening methods, including absorption, distribution, metabolism, and excretion properties prediction and Structural-based Virtual Screening, were carried out to identify new chemicals that disrupt E6-E6AP interactions with high affinity and efficacy (Ricci-López *et al.* 2019). There is a clear potential for the *in silico* HTS for HPV drug discovery and optimisation, as it provides a more subjective way to study the interaction between viral proteins and drug molecules. However, the computational-based modelling results inevitably require experimental validations in cells, which may result in a longer timeframe to progress to clinical studies and eventual drug licensing compared to conventional drug discovery methods, which rely on molecules that have already been discovered.

RNA interference (RNAi) is an ancient intra-cellular mechanism that regulates gene expression and cell function. It is a powerful tool for analysis of the effect by gene knockdown and has also been used as a positive control in new drug molecule discoveries (Celegato *et al.* 2020). Large-scale gene silencing using RNAi high-throughput screening (HTS) has opened an exciting frontier to study gene function in mammalian cells systematically. There are two major types of HTS RNAi, small interference RNA (siRNA) and short hairpin RNA (shRNA). The former allows gene knockdown at a well-controlled introductory titre for a large variety of cell types for a short period, as the siRNA will be diluted as the cell divides; while the latter allows a longer-term gene knockdown effect, but it is harder to balance different shRNA concentrations and have poor transfection efficiency in primary cell lines (Gao *et al.* 2014). RNA analysis doesn't directly lead to new drugs, but it offers insight into HPV's life cycle through the identification of new host cell factors and processes (Aydin *et al.* 2014; Lipovsky *et al.* 2013). As a result, assisting the discovery of new druggable targets for HPV-associated diseases and targeting host cellular mechanisms might lead us to find wide-spectrum antiviral drugs. Although RNAi has been a successful tool for *in vitro* studies, the major limitation is the delivery method for *in vivo* systems. RNAi HTS assays were described in mice (Beronja *et al.* 2013; Fellmann & Lowe 2014) and are likely to become an essential technology advance that facilitates gene function identifications in a physiological context.

High throughput screening using small molecules has been exploited for a longer period compared with other methods mentioned above. One of the main advantages of small molecule HTS over RNAi HTS is the effect size of the data, which is important for drug discovery. RNAi knockdowns of genes are typically less effective than those achieved with small molecules, resulting in considerable noise in data. Moreover, the off-target effect of siRNA on similar sequences will result in more false-positive results compared with small molecules.

For HPV drug discovery, the Luciferase-reporter system has been utilised in a variety of ways for small molecule HTS. To screen for small molecule inhibitors for viral vegetative replication, *Fradet-Turcotte et al.* have developed a model system where they use heterologous E1 and E2 expression vectors and monitor the replication of an HPV origin-containing plasmid using

luciferase as a reporter (Fradet-Turcotte *et al.* 2010). The method allows scientists to identify potential inhibitors of E1-E2 interactions, but does not consider the regulatory roles of other proteins in HPV infection. Organotypic raft culture experiment has suggested that efficient HPV genome amplification is also dependent on E4 and E5 proteins (Hebner & Laimins 2006). Moreover, as E1 and E2 are expressed from heterologous vectors, both protein and replication levels may significantly differ from HPV genome replication. Using luciferase reporter genes engineered with full-length HPV genomes, Toots et al. developed an HTS system that correlates expression level with viral genome copy number in moderately differentiated osteosarcoma cells (U2OS). HPV disease treatment poses a considerable problem due to its highly evolved and specific niches to various epithelial sites that make finding a general antiviral drug hard, particularly when targeting the amplification stage of the viral genome. A similar method can also be used to measure cancer cell viability (the activity of cellular ATP), either by transfecting Luciferase-reporter plasmids or by using a commercially available kit (Alkhilaiwi *et al.* 2019; Walhart *et al.* 2020). Luciferase-reporter assays are quick and easy ways to identify HPV inhibitors. However, they have limitations, including type-specific inhibition, lack of consideration of cytotoxicity for healthy epithelial cells, and potential for recurrence.

In the previous sections, we proposed that HPV E6 proteins modulate the epithelial pathways to keep the infected cells at the basal layer, which is a new antiviral drug target for the clearance of infected cells by antagonising E6 function. Because older drugs can potentially be repurposed for new disease treatments, the process is widely adopted due to the low development costs and shorter development timelines associated with repurposing 'old' drugs (Pushpakom *et al.* 2018). With the support from Maruho Co., Ltd, Japan, we performed a high-throughput screening for small molecule inhibitors using FDA-approved drugs (Prestwick® Chemical Library) against HPV16E6 function, primarily from cell cycle regulation and extended to commitment to differentiation and competition with adjacent uninfected cells.

4.2 Quality control and statistics of HTS system

In section 3.5, the high-content imaging machines were explained, and, in this section, ArrayScan was used for 1st screening, and then confirmation analyses were carried out with Opera Phenix. Effective analytic quality control methods serve as a gatekeeper for excellent quality assays. In a typical HTS experiment, the distinction between positive control and negative control is an index for good quality. Greater separation and lower deviation of the positive and negative control are desired for a good HTS. Several methods have been used to evaluate the HTS system, including signal-background ratio (s/b ratio), signal-noise ratio (s/n ratio), signal window (SW) and Z-factor (Table 4.2). While the s/b and s/n ratios do not consider the variability of both negative/positive control datasets, they can be misleading in terms of accessing screening quality. The signal window (SW) provides a degree of separation between signals (Sittampalam *et al.* 1997), and the Z' factor is a useful tool used to assess the robustness of the assay for screening (Zhang *et al.* 1999a). Although both parameters measure that assay signal adjusted for assay variability, it is described that Z' has better precision properties than SW to measure assay performance (Iversen *et al.* 2006).

Quality control methods	Formula	Criteria	Comments
Signal-background ratio (s/b ratio)	$\frac{\bar{x}_P}{\bar{x}_N}$		Interpretation based on graphics does not contain data variability
Signal-noise ratio (s/n ratio)	$\frac{\bar{x}_P - \bar{x}_N}{sd_N}$		Interpretation based on graphics consider data variability of negative control only

Signal window (SW)	$\frac{\lvert\bar{x}_P - \bar{x}_N\rvert - 3(sd_P + sd_N)}{sd_N}$	SW≥2 (Recommended), SW≥1 (Acceptable)	Interpretation based on graphics accounts for data variability of both control
Z-factor	$\frac{\lvert\bar{x}_P - \bar{x}_N\rvert - 3(sd_P + sd_N)}{\lvert\bar{x}_P - \bar{x}_N\rvert}$	Z-factor=1 (ideal), 1>Z-factor≥0.5 (Excellent), 0.5>Z-factor>0 (Doable), Z-factor=0 (Yes/No)	Interpretation based on graphics accounts for data variability of both control

Table 4.2: Quality control methods, their formula and characteristics. $\bar{x}_P$ and sd_P are the sample mean and standard deviation of positive control; $\bar{x}_N$ and sd_N are the sample mean and standard deviation of negative control, respectively. Adapted from (Zhang 2011).

Based on the optimized assay protocol, the assay performed well when seeding at high confluency, simulating the basal epithelial homeostasis *in vitro.* FUCCI NIKS LXSN and 16E6 cells were seeded at 8x of seeding density on 96-well plates; media changed 24hr after seeding and allowed to grow for another 48hr before collecting for data analysis (Figure 4.2a). Since the aim of the drug discovery is to modulate keratinocyte homeostasis with the presence of viral protein, less intervention to the assay cell line is desirable. Although cell cycle status is the major effector measured, synchronising cells with chemicals is not used for the screening, and instead, we allow the cell line to reach a saturated steady state by controlling media supply (especially environmental EGF level, which high-level presence can drive cell proliferation). Following this idea, small molecules/vehicles were added 24hr before data collection. With the optimised conditions, we calculated the parameters (listed in Table 4.2) commonly used as tools to measure assay performance during validation of an assay for HTS.

In the previous section, the *in vitro* assay evaluating HPV E6 function in basal layer keratinocytes is measured in three different parameters, namely cell density (No. of nuclei), cell proliferation (%S/G2/M) and commitment to differentiation (Total populational K10 intensity). While for the purpose of HTS, the effect with the small molecule is evaluated at a

single parameter individually. Therefore, %S/G2/M is used as the main selection parameter for HPV16E6 inhibitors as its regulation is more direct by virus protein compared to the other phenotypic measurements. Using the fully optimised assay, we generated a Z-factor and signal window of assay running 80 positive controls (LXSN 16E6) and 16 negative controls (LXSN) (Figure 4.2b). Raw data of %S/G2/M from LXSN and LXSN 16E6 were plotted (Figure 4.2c), with statistical measurements of mean and standard deviation (sd) indicated by solid and dotted lines, respectively. From the control result, we achieved a Z-factor of 0.64 and a signal window of 21.89. A Z-factor≥0.5 and a signal window ≥2 are appropriate for the HTS assay; therefore, our assay is deemed to be highly robust and reproducible and hence suitable for HTS applications. Additionally, the 3xsd range, which by definition covers 99.7% of the confidence range of the mean in a normally distributed dataset, is shown in the graph. For identifying 'HIT' molecules, the small molecule effect that lies outside of the 3xsd range is picked up.

To screen for HPV16E6 inhibitors for the elimination of persistently infected basal keratinocytes, a graphical illustration is better for assessing small molecule effects (Figure 4.2d). By seeding LXSN and 16E6 at the same high density at starting point, after being treated for 72hr under the same condition, the 16E6 population exhibited higher cell density, a higher proportion of cycling cells and lower K10 expression compared to LXSN. Normalisation and log2 calculation were performed with cell density and %cycling cells for easier interpretation. For candidate small molecule which affects on 16E6 population is represented in the red star in Figure 4.2d; it should target 16E6 specifically with no negative effect on the NIKS LXSN population. To achieve a successful antiviral effect, NIKS LXSN 16E6 should respond to small molecule treatment and show lower cell number, less %cycling cells and increased differentiation by K10 expression. Ideally, after small molecule treatment, NIKS LXSN 16E6 become less competitive than NIKS LXSN to stay at the basal layer, and then persistently infected cells can be cleared by extruded towards terminal differentiation in the epithelium.

Although the solvent vehicle for small molecules DMSO (0.1% of final concentration) shows some degree of toxicity to NIKS cells in general, it is affecting both LXSN and LXSN 16E6 cell

line in the same way, and it does not affect the normalised ratio change. Therefore, this level of toxicity is tolerated by the assay system for the screening project.

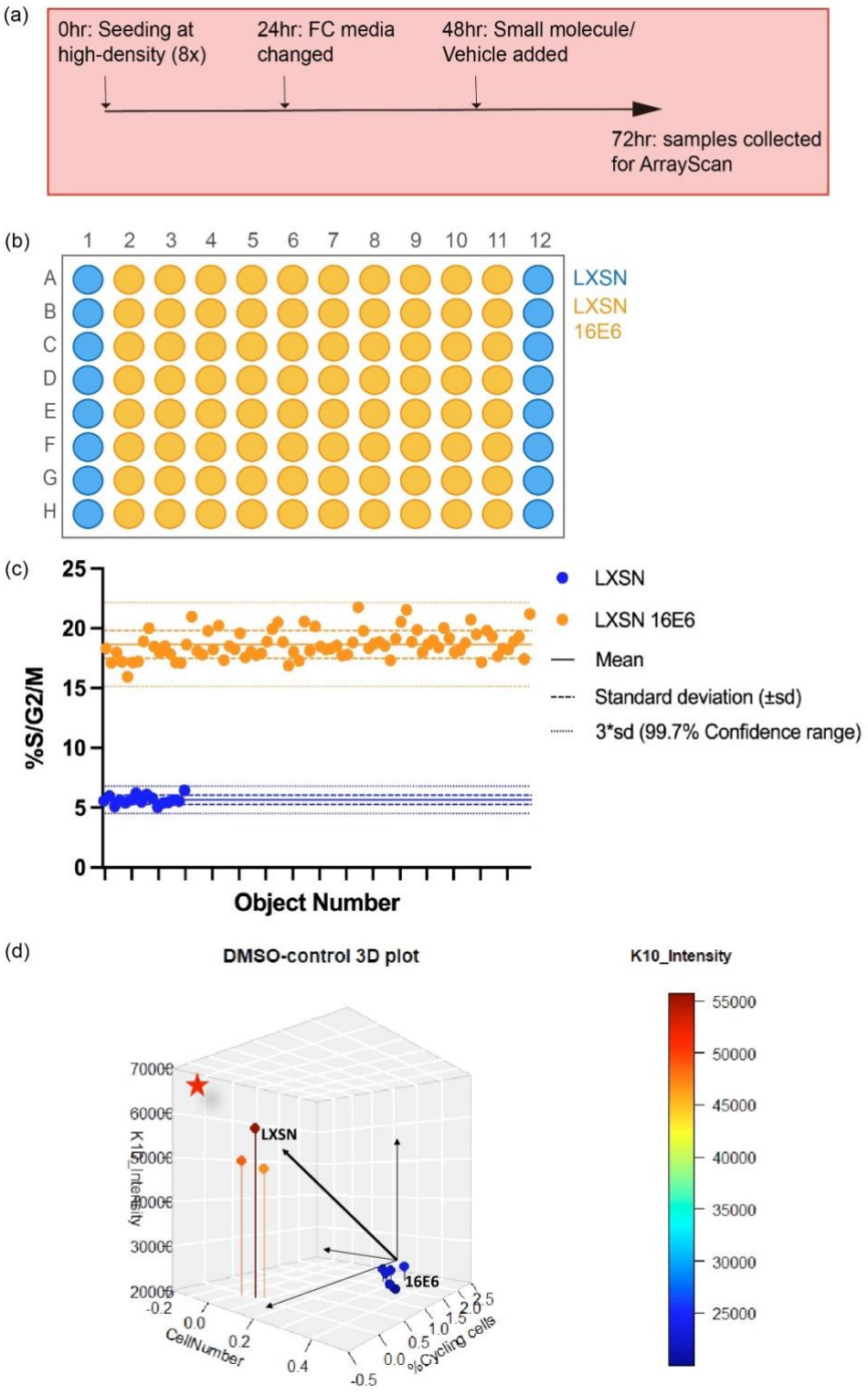


Figure 4.2: Quality control, statistics and design of HTS assay. (a) Timeline of HTS assay. (b) Plate layout of control plate for HTS quality control assessment. (c) Scatter plot of %S/G2/M measured by FUCCI of negative (LXSN) and positive controls (LXSN 16E6) with statistical interpretations. (d) 3D scatter plot of control plates, CellNumber and %Cycling cells (%S/G2/M) were normalised by LXSN data, and log2 calculation was carried out, indicating fold change. K10 intensity was measured with total plate immunofluorescence intensity. Arrows suggest small molecule inhibitory effects on LXSN 16E6 function in NIKS, and the red star represents the ideal small molecule effect on LXSN 16E6 population.

4.3 First screening and HIT selection

4.3.1 Screening with Prestwick Chemical Library (PCL)

With support from Maruho Co., Ltd, we decided to repurpose current drugs for potential anti-HPV effects using the Prestwick® Chemical Library (PCL). There is a total of 1,269 FDA-approved drug molecules from the library, and all small molecules were supplied in 10mM of DMSO irrespective of their differential solubility to solvents. By applying a topical treatment to the epithelium, we are able to achieve a higher local concentration of the drug compared with systemic treatments. We carried out our primary HTS assay with 10uM of the small molecule in DMSO. Hence, it will be a 1000x dilution from the stock, and a final concentration of DMSO will be 0.1% in FC media. The total chemical list can be found in Appendix File 1.

Our aim was to identify a small molecule that specifically inhibits HPV16E6 function with minimal toxicity to normal keratinocytes. Therefore, treatment of each small molecule for the HTS is tested in both LXSN and LXSN 16E6 cell lines. An example plate layout is shown in Figure 4.3.1a with detailed plate layout can be found in Appendix File 2. For each experiment set, column 1 is eight replicates of DMSO-treated negative controls, with 80 small molecules treated samples from columns 2 to 11 and additionally, column 12 was used for testing several anti-HPVE6 small molecules proposed in the literature (more details explained in Section 4.3.2).

Next, an example graphical illustration of one plate from the 1st HTS assay is shown in Figure 4.3.1b. The vehicle control (DMSO-treated) LXSN and LXSN 16E6 cell lines were plotted in solid dots distinguished by two different colour schemes. Then with the replicates from each control group, the mean and standard deviation is calculated, and the $\pm$3xsd range around the mean is indicated as the coloured boxes. Individual small molecule treated samples were plotted as hollow dots and categorised by colour/cell type. Any small molecule effects that lie outside of the $\pm$3xsd from the mean are identified as 'HIT' molecules. However, we have additional criteria to compare the effect of the same small molecule on two different cell lines, which LXSN should fit inside the control range, with LXSN 16E6 moving towards LXSN controls after treatment. Here we show a Drug L (shown in red triangles), which meets our criteria and is picked up as our 'HIT' molecules for further evaluation.

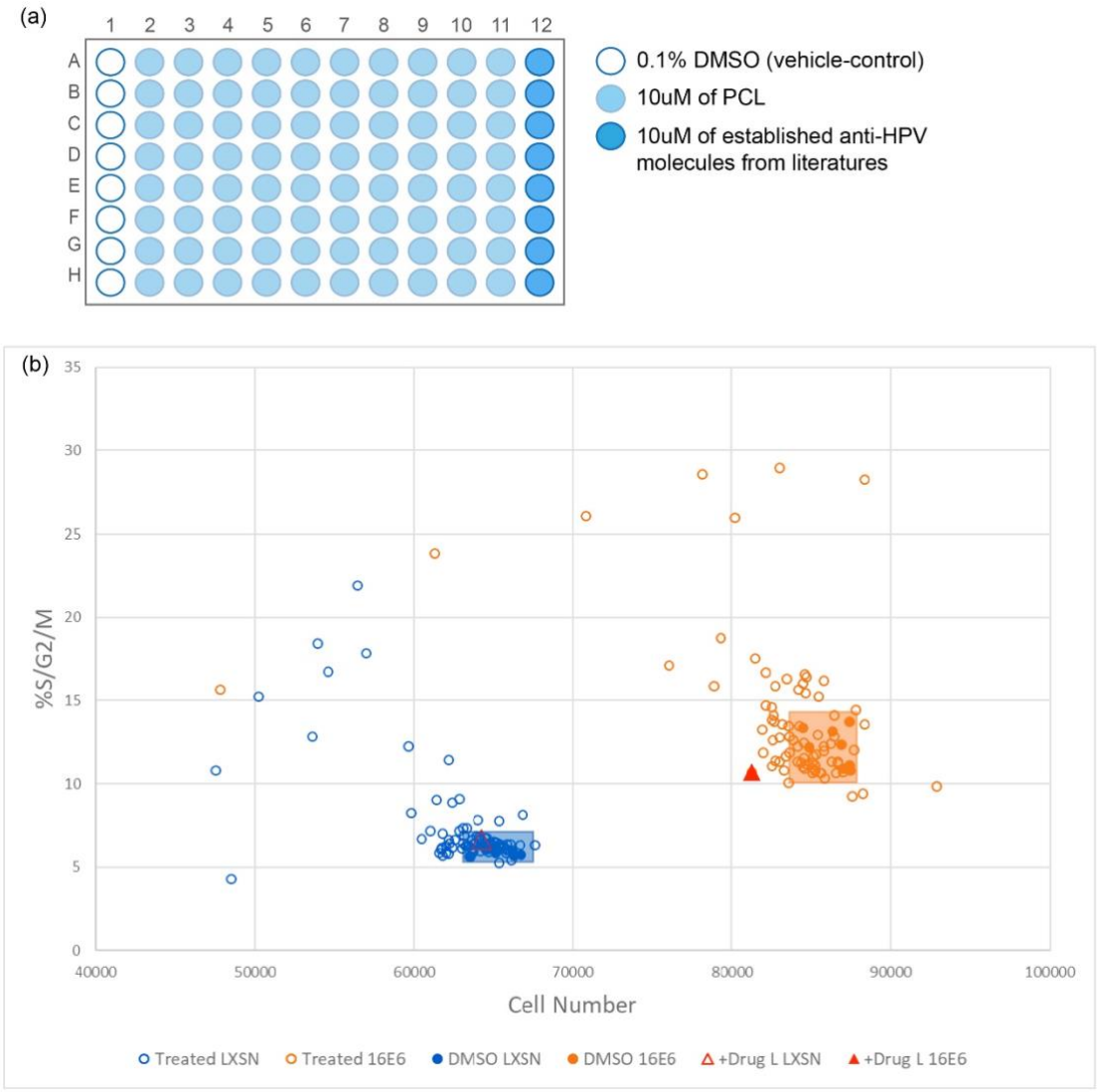


Figure 4.3.1: Graphical representation of 1st high-throughput screening assay. (a) Example 96-well plate layout of HTS, including eight wells of DMSO-treated negative control and different small molecule treated wells. Both cell line LXSN and LXSN 16E6 responses to small molecules were tested at the same time with the same plate layout. For a more detailed plate layout, check the Appendix File 2. (b) Example of 80 small molecules (one plate) effect on cell number and %S/G2/M of LXSN and LXSN 16E6 cells, respectively. Solid dots represent the DMSO-treated control data, and hollow dots represent the data after 24hr of 10uM small molecule treatment. A colour scheme was used for identifying different cell lines. With DMSO-treated control data, the mean$\pm$3xsd range is calculated and indicated as the boxed area on the scatter plot. The triangle shows an example of the HIT molecule effect (marked in red), in

which the solid red triangle is LXSN 16E6 treated with Drug L, and the hollow red triangle is LXSN treated with Drug L.

Although the scattered plot is a good way of presenting the HTS data, for larger-scale analysis, we applied a numerical method to calculate the small molecule effect size regarding to vehicle control. The data for analysis was calculated as raw data subtracted by the mean of the control and divided by the standard deviation of the control.

$$effect\ size = \frac{x - \bar{x}}{sd}$$

The transformed data is listed in Appendix File 3, with colour formatting to show small molecule effect $\geqslant$3xsd marked in green and $\leqslant$-3xsd marked in red. With the criteria, a final candidate list consisting of 167 potential HPV E6 inhibitors is picked which are small molecules have no negative effect on NIKS LXSN while causing a significant decrease in the growth of NIKS LXSN 16E6 (Table 4.3.1). Then each small molecule is categorised by its current established molecular targets to show similarities or differences, which may provide some hints for any future analysis of E6-modulated molecular pathways.

Target Type	LXSN (Cell Number)	LXSN (%S/G2/M)	16E6 (Cell Number)	16E6 (%S/G2/M)
Carrier				
Na+				
Clopamide	-2.976830282	-1.028814058	-3.187188814	-0.701598023
Serotonin transporter				
Zimelidine dihydrochloride monohydrate	-0.588509406	-0.781578465	1.041842137	-3.087388447
Uric acid uptake				
Benzbromarone	-2.465494588	10.30610596	-3.237330589	1.243524743
Zoxazolamine	-1.139907657	3.681112914	-3.357160885	1.889397552
Noradrenaline release				
Bretylium tosylate	-0.404618936	1.78190509	-4.36024437	-0.659161555
Solute carrier family 12 members 3				
Indapamide	-1.856628306	-0.845491454	-2.989767648	-3.888444252
DNA				
Alkylating-agent				
Trioxsalen	-2.02939874	7.120876756	-3.079181119	3.417397805
Undefined				
Nifuroxazide	-1.87802179	1.484256943	-4.608076471	2.460111486
Nitrofurantoin	-0.322879426	0.929302247	-4.581089447	2.832655025
Ornidazole	0.43212549	-0.71148822	-3.828852526	2.042923838
Enzyme				
5-AMP activated protein kinase				
Phenformin hydrochloride	-0.785790848	9.038126213	-5.172991678	8.350398522
Alcohol dehydrogenase				
Disulfiram	-0.654605382	1.843451699	1.274227809	-3.317710815
Bacterial transpeptidase				
Cloxacillin sodium salt	0.275757418	-1.524667706	-4.412738023	-2.567425457
Carbonic Anhydrase 2				
Merbromin	-2.563519848	36.21277827	-3.003803952	-65.48845274
Cyclooxygenase				
Etofenamate	-1.171290185	0.198540695	-3.388381347	1.649681419
Fenoprofen calcium salt dihydrate	-2.672105373	1.238582747	-3.78884775	0.461199731
Meclofenamic acid sodium	-0.029753799	-1.144903968	-3.419999066	-3.2887141
Mefenamic acid	-1.166587097	-1.094927893	0.523923	-3.98276367
Nimesulide	-1.403522143	6.864978608	-5.279881472	4.858097297
Tolfenamic acid	-0.928225131	-0.031292406	-5.614580138	0.911702712
Nabumetone	-0.980303744	7.623395203	-4.087887162	-0.550107671
Aceclofenac	-0.253210483	-0.709516434	-3.211563969	-3.746661219
Ketoprofen	-2.02305967	-0.868258712	-0.912097288	-3.997925847
Cyclooxygenase 1				
Meloxicam	-2.274737248	2.880293412	-3.296820814	0.349411039
Cyclooxygenase 2				
Rofecoxib	-2.09040763	3.726995766	-3.560716615	6.865402241
DNA-dependent RNA polymerase				
Rifapentine	-1.511429446	3.16314179	-4.64272982	1.32097291
Folic acid synthesis				
Mafenide hydrochloride	-0.274111166	-0.043950176	-3.496826105	0.684298766

Sulfadiazine	-2.498990977	-2.644241632	-0.076403441	-4.211259781
Sulfadimethoxine	-1.419627502	3.964267233	-3.01837293	1.256316305
Sulfamethazine sodium salt	-2.906125902	0.115367682	-3.608976902	0.71558155
Trimethoprim	-1.074271879	0.422461293	-3.640459265	-2.427164047
Sulfamethoxazole	0.201185845	0.382138015	-2.086548893	-3.517522256
Fumarate reductase				
Oxantel pamoate	0.007285895	2.321753851	-3.691156034	3.264667496
Pyrantel tartrate	-1.657684768	-0.859551531	-2.504087149	-4.643969971
Heme polymerase				
Mefloquine hydrochloride	-2.532778158	1.440304342	0.857627527	-3.001126953
HIV-1 protease				
Lopinavir	-0.819329303	2.867485958	-4.455180509	-1.396214228
HMG-CoA reductase				
Simvastatin	-1.954400252	-0.005985787	-4.757437193	-0.058507391
Monoamine oxidase				
Iproniazide phosphate	-0.298624313	-0.932116065	-3.822171914	-2.788786652
Tranylcypromine hydrochloride	-0.557173502	-1.306781971	-3.608392327	-5.19639087
Phenelzine sulfate	0.105971623	-1.04708414	-3.027179075	1.362917453
NO-dependent activation of soluble guanylate cyclase				
Ambroxol hydrochloride	-0.549089752	3.300918383	-3.81098633	-0.306649227
Phosphodiesterases				
Etofylline	0.099778828	-1.084722899	-3.558287736	-3.233380438
Pentoxifylline	-0.602329298	1.776815587	-4.594450672	-1.395454653
Renin				
Aliskiren hemifumarate	-1.062306302	-0.706017176	1.912362041	-3.47419092
Topoisomerase I				
Ciprofloxacin hydrochloride monohydrate	-2.200989448	-1.529398146	0.914797579	-3.573150163
Transcriptase reverse				
Zidovudine, AZT	-1.676854231	5.557584857	-3.250078495	1.528164094
Peptidoglycan synthesis				
Bacitracin	-1.539470498	2.507476424	-4.196830045	-0.032041074
DOPA decarboxylase				
Benserazide hydrochloride	-0.877524735	6.822860325	-4.017074287	0.338669124
Cholinesterase				
Pralidoxime chloride	-1.63368459	1.900649991	-3.192221027	-1.156179364
Plasminogen				
Tranexamic acid	0.463089464	-2.656663997	-0.521949542	-4.458667771
Alkaline phosphatase				
Levamisole hydrochloride	-0.050912515	-1.021726839	-0.491218726	-3.5640015
Kinase				
Aminopurine, 6-benzyl	-0.260177378	1.040455294	-3.036531932	-2.317823495
Topoisomerase II				
Norfloxacin	-2.838960403	0.458072587	-2.945675608	-4.044798627
GPCR				
Bradykinin receptor				
Fenspiride hydrochloride	-2.843088933	-0.534037673	-3.861587525	-0.103996896
Catecholaminergic receptor				

Asenapine maleate	-0.195411064	2.866025678	-4.902426889	1.883994975
Mirtazapine	0.010695144	2.374492588	-3.175140876	2.536560079
Dopaminergic D1 / D2 receptor				
Apomorphine hydrochloride hemihydrate	-1.428124623	1.177948497	1.610631214	-3.392593103
Dopaminergic receptor				
Fluphenazine dihydrochloride	-1.339553628	0.822164305	-4.20994993	3.784638749
Histaminergic H1 receptor				
Brompheniramine maleate	-1.572780597	1.127548018	-5.122973653	2.094292625
Chlorpheniramine maleate	-1.744207358	-0.99406034	0.169469501	-3.488849374
Doxylamine succinate	-1.139440102	-1.624634646	0.501558524	-3.66086513
Orphenadrine hydrochloride	-2.788901977	2.299570695	-6.589281441	-8.198704609
Pheniramine maleate	-0.40427012	-2.014262666	0.548523924	-4.058726777
Histaminergic H2 receptor				
Famotidine	-2.331872568	-1.013538867	1.636834154	-4.093563451
Nizatidine	-0.412572073	1.302902421	-3.066936711	-0.975420816
Opioid receptor				
Naltrexone hydrochloride dihydrate	-1.360719816	-1.238568643	0.689293487	-3.157524873
Sigma1 receptor				
Carbetapentane citrate	-1.381003742	2.8741172	-3.365659488	1.323557773
Histaminergic H3 receptor				
Thioperamide maleate	-1.090424111	0.59677103	-4.193717391	-0.599025178
Adrenergic Beta 1 receptor				
Xamoterol hemifumarate	-0.740486056	-0.079437972	-3.291047785	-0.662332906
Adrenergic receptor				
Metaraminol bitartrate	-0.429705139	0.019508086	-3.070603054	-2.809529945
Adrenergic beta-2 receptor				
Salbutamol	-0.413707085	0.791150976	-3.099329686	-1.744679073
Ion channel				
Ca2+ channel				
Prenylamine lactate	-2.5953658	14.1780962	-4.449605401	4.846917906
Ca2+ channel voltage dependant				
Ethosuximide	-0.140450009	0.271089563	-4.33925129	1.63545488
K+ channel voltage dependant				
Gliclazide	-0.236365935	-2.271946522	-9.527933614	-11.85927547
Na+				
Riluzole hydrochloride	-0.407514291	5.730112463	-7.99755447	7.465892303
Primidone	-0.382594862	0.637042013	-3.7314883	-0.693136731
Voltage-dependent N-type Ca2+ channel				
Levetiracetam	-1.19477467	0.570042462	1.707125831	-3.787154128
Voltage-gated Na+ channel				
Prilocaine hydrochloride	-0.286238723	-0.546613155	-5.593202179	-0.920810139
Voltage-gated L-type Ca2+ channel				
Nitrendipine	-0.998657139	5.643909879	-3.060711403	-0.364013539
Nimodipine	-1.551706094	13.68031221	-3.044369971	0.48027976
Ionotropic receptor				
Nicotinic receptor				
Hexamethonium dibromide dihydrate	-1.785832072	-2.324254608	0.151075232	-3.908876983
Pentolinium bitartrate	-0.475377002	-0.744935762	-3.32246213	-2.158997245

Pipenzolate bromide	-1.779083035	-1.38962957	-3.519881995	-0.069506433
GABA-A receptor				
Propofol	-1.182802864	-2.896582086	-3.132302441	-1.349046442
Membrane				
Undefined				
Monensin sodium salt	-2.614068282	21.83586971	-5.366780182	4.260058521
Nuclear receptor				
Androgen receptor				
Clomiphene citrate (Z,E)	-2.041589349	10.86445763	-6.072464795	4.498581989
Nilutamide	-0.463349798	0.690332239	-3.536947437	-1.153899293
Opioid receptor delta				
Nalbuphine hydrochloride	-2.125077238	0.549056766	0.453202349	-3.183829085
Peroxisome proliferator-activated alpha receptor				
Gemfibrozil	-2.80980266	-0.923570172	0.504860536	-5.727908438
Peroxisome proliferator-activated gamma receptor				
Pioglitazone	-1.706905669	1.235127636	0.888594638	-3.370381941
Thyroid hormone receptor				
Liothyronine	-1.953463114	3.805985866	-3.722933672	0.547150679
Tiratricol	-2.990167812	7.321599442	-13.54045832	6.783157628
Undefined				
Tibolone	-1.269344919	2.092128869	-3.778079874	-1.119259973
Estrogen beta receptor				
Estrone	-1.641637727	2.789768073	-4.52910584	-0.503817583
Protein				
Tubulin beta chain				
Oxibendazol	-1.557446466	0.760952754	-7.073777684	3.050172884
Pump				
ATP/Mg2+ Pump				
Reserpine	-2.122147864	1.287297908	-3.879983085	1.128851726
Receptor				
Glucocorticoid receptor				
Flumethasone	-2.230274614	5.046751324	-3.682382139	1.368332268
Progesterone receptor				
Mometasone furoate	-0.846966472	8.115217658	-3.176068548	1.564526949
Ribosome				
30S unit				
Cytarabine	-1.539312877	-0.245265194	-4.113429556	2.291704486
Oxytetracycline dihydrate	-2.665072711	1.863376446	-3.73534449	1.856930916
Ribostamycin sulfate salt	-1.048938984	2.361147879	-3.100743693	-0.567154858
Tigecycline	-1.121147451	-1.739753854	-3.287345975	-0.243968007
50S unit				
Amikacin hydrate	-1.542617617	1.556397474	-4.24271542	3.803085697
Transmembrane receptor				
Sulfonylurea receptor K-ATP channel				
Glibenclamide	-2.413329236	1.622555149	-3.040587658	1.726062362
Tolbutamide	0.091779801	-1.781822343	-1.872769306	-3.542322477
Cholecystokinin receptor antagonist				

Lorglumide sodium salt	-0.447443523	0.824901129	-4.165703507	-1.258291963
kappa receptor				
Idazoxan hydrochloride	-0.415019193	1.549969122	-3.553288869	-1.160170785
Undefined				
Copper chelating agent				
Tetraethylenepentamine pentahydrochloride	0.11680064	1.099172079	-3.05331448	-0.14010032
Iron chelating agent				
Deferoxamine mesylate	-0.622085539	1.540326972	-3.343373218	0.573646514
Undefined				
Alverine citrate salt	0.057461396	0.807513533	-4.981258112	-2.488442304
Carbadox	-1.723918766	0.156976567	-3.773375223	2.640108835
Chlorzoxazone	-0.705026481	0.11585876	-6.495752872	2.960527612
Flunixin meglumine	-2.718909479	12.33353887	-3.670054542	1.722395545
Glafenine hydrochloride	-2.723833473	7.134877897	-3.565879117	3.603962049
Isoniazid	-0.339135513	-0.65954672	-3.789436915	1.673397337
Khellin	-0.042726795	0.128717992	1.719677793	-3.153078853
Mefexamide hydrochloride	-2.665304111	-0.305915972	-3.946431299	0.300235963
Thalidomide	-0.544013812	-1.032096441	-3.392608557	-1.000325198
Fluspirilen	-0.28838077	12.35223814	-3.3439629	-0.803766065
Bephenium hydroxynaphthoate	-0.178872183	3.67573091	-3.55795785	-1.319854345
Lacosamide	-0.098648643	-1.743406905	-2.09055726	-3.331563544
Unknown				
Paliperidone	-0.059575119	1.068066854	-3.47235987	-0.319347186
Chelating agent				
Pentetic acid	-1.360830791	-1.794674833	-1.914476636	-3.123037686

Table 4.3.1: List of HIT molecules from 1st screening. Data grouped by the established molecular targets of each drug, not representing its molecular pathway in our model system. Individual effects were calculated as the deviation from their mean and divided by the standard deviation set by DMSO-treated data of each set. Data highlighted in green as the effect is ⩾3xsd from mean, and data in red as the effect is ⩽-3xsd from mean.

4.3.2 Screening with established anti-E6 molecules from literature

At the same time as performing unbiased screening with FDA-approved drugs, we also explored literature about established small molecule inhibitors against HPV16E6 and some basal epithelial differentiation modulators based on our hypothesis (Table 4.3.2). These additional small molecules were compared in parallel with the PCL small molecules.

Small molecule	Pathway targeted	Reference
17-AAG	HSP90 inhibitor	(Ajiro & Zheng 2015)
GDC-0449	Sonic hedgehog SMO inhibitor	(Vishnoi *et al.* 2016)
LDE-225	Sonic hedgehog SMO inhibitor	(Vishnoi *et al.* 2016)
DAPT	Notch/NICD, γ-secretase inhibitor	(Kranjec *et al.* 2017; Weijzen *et al.* 2003)
Luteolin	E6/E6AP binding inhibitor from computational modelling	(Cherry *et al.* 2013; Rietz *et al.* 2016)
Gefitinib	EGFP pathway inhibitor	(Chen *et al.* 2018; He *et al.* 2015)
ROCK-inhibitor (Y-27632)	S1P-ROCK/Non-canonical Notch inhibitor	(Vaezi *et al.* 2002; Yugawa *et al.* 2013)

Table 4.3.2: List of potential HPV16E6 inhibitors and basal epithelial pathway modulators.

For each small molecule, independent triplicates were performed with FUCCI NIKS LXSN and FUCCI NIKS LXSN 16E6. The results were calculated the same as effect size in the HTS assay and graphed as mean with standard deviation (Figure 4.3.2). Firstly, focusing on LXSN 16E6 population only, 17-AAG/LDE-225/Luteolin all have a significant inhibitory effect on cell growth, shown as lower cell number. The increase of %S/G2/M is likely due to the natural FUCCI NIKS response to lower cell density. However, when compared with LXSN, 17-AAG/LDE-225/Luteolin are negatively affecting NIKS growth generally. For the rest of the list, there is no obvious function of the small molecule on both populations.

The discrepancy between our results and previous literature suggested the fundamental difference in how experimental conditions could affect results. The majority of the published research focused on the oncogenicity and transformative ability of HPV16E6 to keratinocytes, and therefore the antiviral agents were developed in cancer cell lines. While our project primarily focuses on the natural homeostasis of basal epithelial cells with a low level of HPVE6 expression, which has a different cellular environment compared to cancer cells. Additionally,

it is interesting to note that the different effects of GDC-0449 and LDE-225 on NIKS, which claimed to target the same protein (Smoothened) of the Sonic Hedgehog pathway, suggesting off-target effects in keratinocytes affecting cell proliferation.

Although none of these small molecules tested fits our aim for specifically inhibiting the HPV16E6 effect on keratinocyte homeostasis, the results provide some hints for finding the molecular pathway of our 'HIT' molecules.

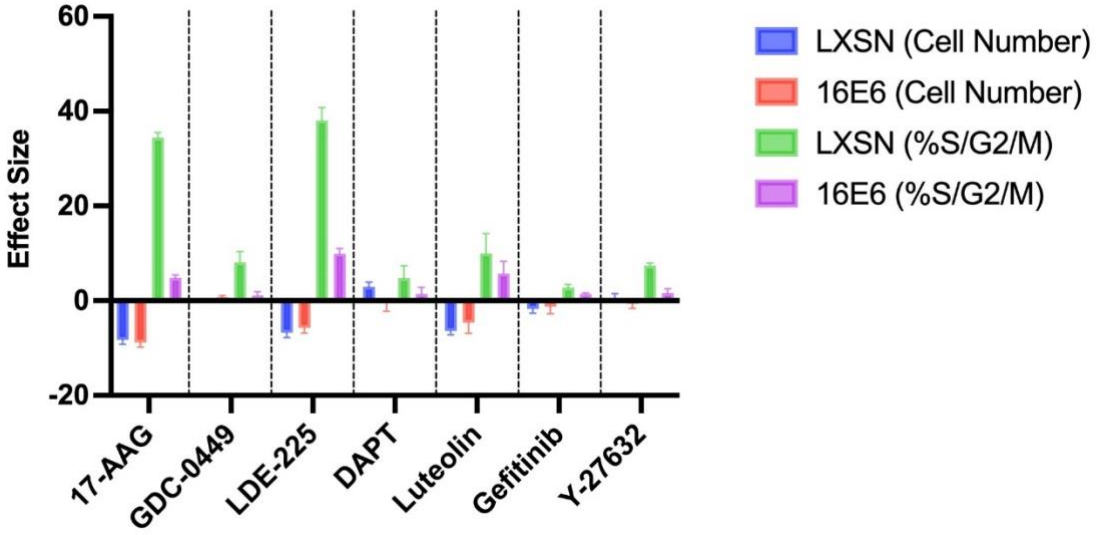


Figure 4.3.2: Effect of anti-HPV molecules from the literature on FUCCI NIKS LXSN and LXSN 16E6 cell density and %S/G2/M. Triplicates were performed with each small molecule and graphed with the effect size (calculated as the number of standard deviations from the mean).

4.4 Confirmation screening with HIT molecules only

A total of 167 small molecules were selected as candidates for inhibitors of HPV16E6 function from the 1st screening assay. In the previous section, the E6 function on NIKS is observed in a variety of densities indicating HPV16 has multiple functions in modulating normal epithelial homeostasis. Therefore, we carried out a confirmation screening assay with the 'HIT' small molecules to verify their results and by slightly adjusting the seeding densities of LXSN and LXSN 16E6 to further categorise the candidates (Figure 4.4.1).

The graphical illustration provides more detail compared to numerical analysis. Firstly, by plotting the control values of DMSO-treated LXSN and LXSN 16E6 represented as black and red dots, respectively, we created an estimated trend of how normal %S/G2/M is responding to cell density changes in both cell lines indicated as black/red-coloured circled area. For our ideal HPV16E6 inhibitors, treated LXSN 16E6 cells should move towards the LXSN control area, and the big blue circle shows our expected top priority small molecule results. Next, we plotted how the LXSN 16E6 cells' behaviour changed in response to 24hr 'HIT' molecule treatment, indicated as hollow dots in the graph. The LXSN data is not shown as, from the first screening result, these 'HIT' molecules have no general cell toxicity. With the scattered plot, we were able to further categorise the 'HIT' molecules with different priorities for future experiments, with the highest priority candidates lying within the estimated blue circle area. Results shown as different colours and individual small molecules details can be found in Figure 4.4.2.

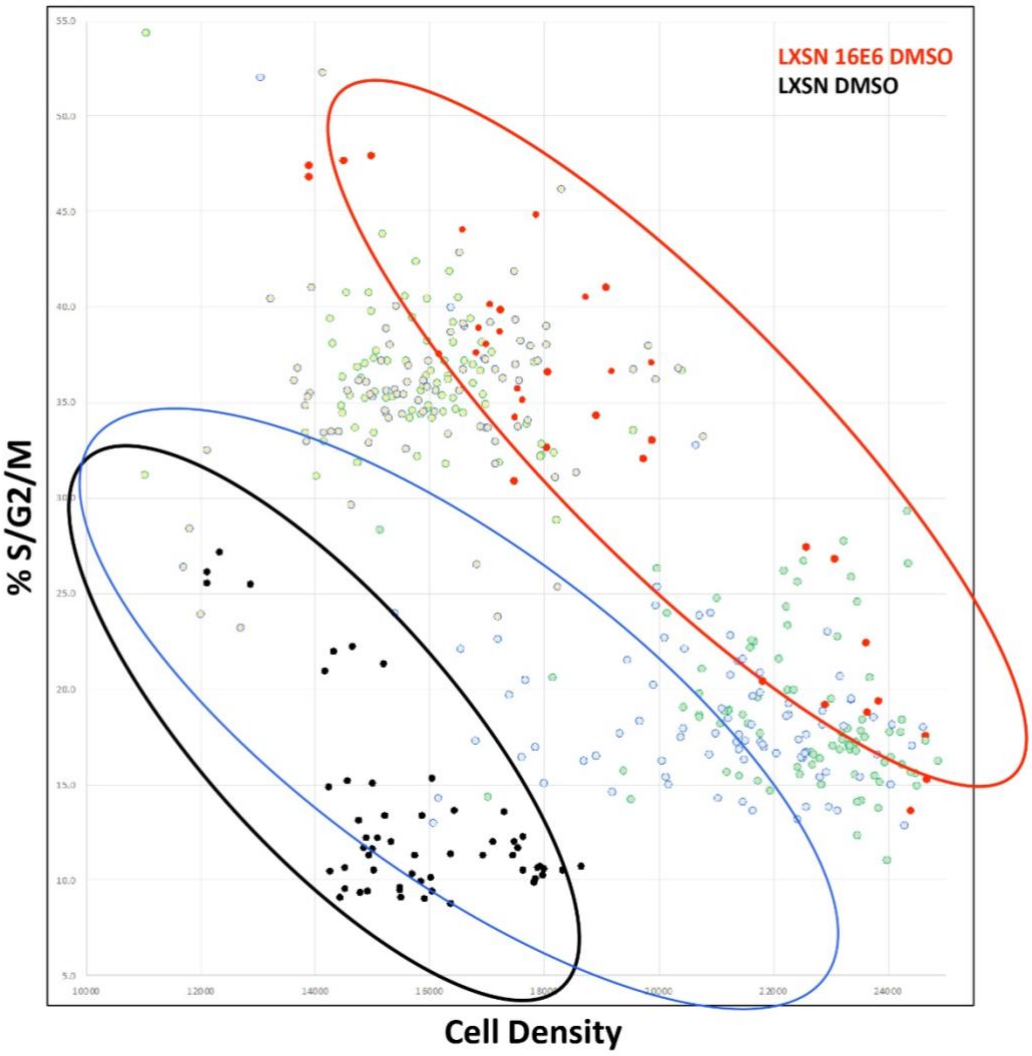
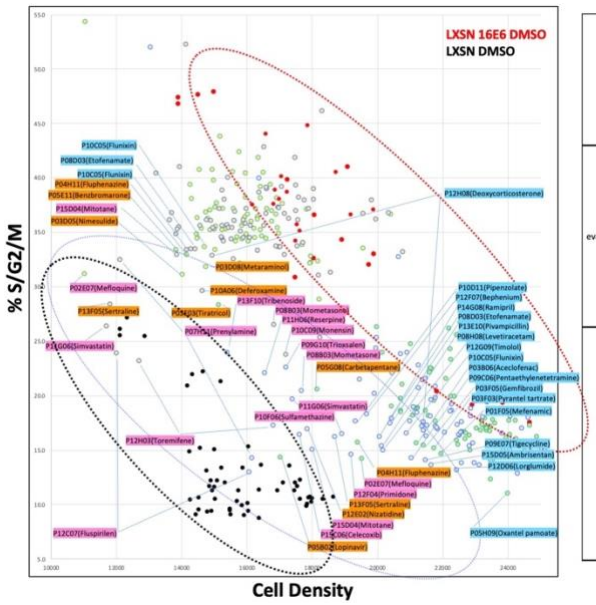
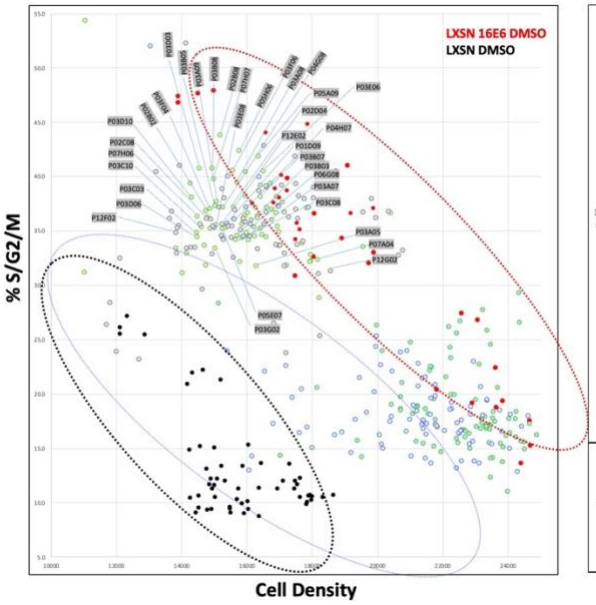


Figure 4.4.1: Scatter plot of confirmation assay of cell density and %S/G2/M of LXSN and LXSN 16E6. DMSO-treated control samples with an adjusted seeding density of LXSN and LXSN 16E6 are plotted as black and red solid dots, respectively, and the estimated control response of %s/G2/M to a cell density of each cell line is indicated as a big circle with matching colours. Small molecule treated LXSN 16E6 samples are shown in hollow dots, which colour only refers to experimental settings. The big blue circle shows the estimated area of the ideal effect with small molecule treatment.



1st	P05B02	Lopinavir	HIV-1 protease
	P13F05	Sertraline	Serotonin transporter
	P03E03	Tiratricol	Thyroid hormone receptor
	P05G08	Carbetapentane citrate	Sigma1 receptor
	P10A06	Deferoxamine mesylate	Iron chelating agent
	P03D05	Nimesulide	Cyclooxygenase
	P03D08	Metaraminol bitartrate	Adrenergic receptor
	P04H11	Fluphenazine dihydrochloride	Dopaminergic receptor
	P12E02	Nizatidine	Histaminergic H2 receptor
	P05E11	Benzbromarone	Uric acid uptake
1st* (need to re-evaluate cell toxicity on NIKS)	P12C07	Fluspirilen	Undefined
	P02E07	Mefloquine hydrochloride	Heme polymerase
	P11G06	Simvastatin	HMG-CoA reductase
	P12H03	Toremifene	Estrogen receptor
	P15D04	Mitotane	suppress adrenal steroid production
	P15C06	Celecoxib	Cyclooxygenase 2
	P07H11	Prenylamine lactate	Ca2+ channel
	P08B03	Mometasone furoate	Undefined
	P09G10	Trioxsalen	Undefined
	P10C09	Monensin sodium salt	Undefined
	P11H06	Reserpine	Undefined
	P10F06	Sulfamethazine sodium salt	Folic acid synthesis
	P12F04	Primidone	Na+ channel
	P13F10	Tribenoside	Vitamin
2nd	P01F05	Mefenamic acid	Cyclooxygenase
	P03B06	Aceclofenac	Cyclooxygenase
	P03F05	Gemfibrozil	Peroxisome proliferator-activated alpha receptor
	P05H09	Oxantel pamoate	Fumarate reductase
	P10C05	Flunixin meglumine	Undefined
	P12H08	Deoxycorticosterone	Mineralocorticoid receptor
	P12F07	Bephenium hydroxynaphthoate	Undefined
	P14G08	Ramipril	Angiotensin-converting enzyme
	P08H08	Levetiracetam	Undefined
	P10D11	Pipenzolate bromide	Nicotinic receptor
	P12G09	Timolol maleate salt	Adrenergic beta receptor
	P13E10	Pivampicillin	Undefined
	P08D03	Etofenamate	Cyclooxygenase
	P09C06	Pentaethylenetetramine pentahydrochloride	Bile acid sequestrant
	P15D05	Ambrisentan	Endothelin receptor A
	P09E07	Tigecycline	Undefined
	P12D06	Torglumide sodium salt	Cholecystokinin receptor antagonist
	P03F03	Pyrantel tartrate	Fumarate reductase



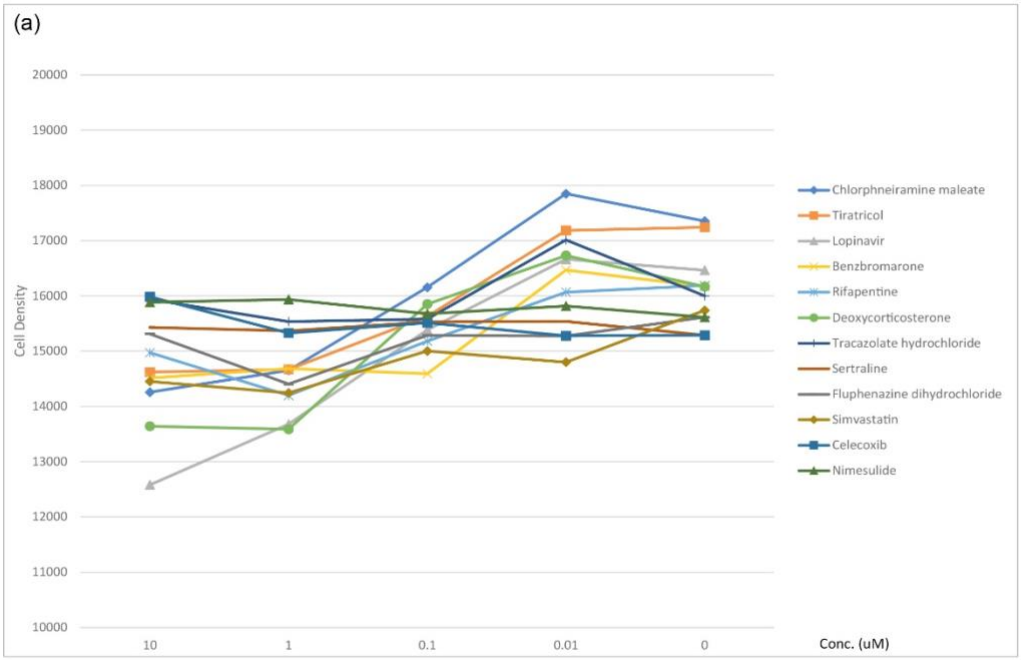
3rd (have some effect mainly in low density cells)	P06G08	Rifapentine	DNA-dependent RNA polymerase
	P03A05	Ornidazole	Undefined
	P03A07	Mafenide hydrochloride	Folic acid synthesis
	P03B07	Iproniazide phosphate	Monoamine oxidase
	P05A09	Zoxazolamine	Uric acid uptake
	P05H06	Amikacin hydrate	50S unit
	P07A04	Rofecoxib	Cyclooxygenase 2
	P12E02	Nizatidine	Histaminergic H2 receptor
	P12G02	Pentetic acid	Chelating agent
	P01D09	Hexamethonium dibromide dihydrate	Nicotinic receptor
	P02B02	Khellin	Undefined
	P02B08	Disulfiram	Alcohol dehydrogenase
	P02C08	Pioglitazone	Peroxisome proliferator-activated gamma receptor
	P02D04	Ciprofloxacin hydrochloride monohydrate	Topoisomerase I
	P03A08	Riluzole hydrochloride	Na+ channel
	P03A09	Nitrofurantoin	Undefined
	P03B03	Etofylline	Phosphodiesterases
	P03B05	Alverine citrate salt	Undefined
	P03B08	Sulfamethoxazole	Folic acid synthesis
	P03C03	Levamisole hydrochloride	Alkaline phosphatase
	P03C08	Lacosamide	Undefined
	P03C10	Aminopurine, 6-benzyl	Kinase
	P03D03	Thalidomide	Undefined
	P03D06	Asenapine maleate	Catecholaminergic receptor
	P03D10	Prilocaine hydrochloride	Voltage-gated Na+ channel
	P03E06	Tolfenamic acid	Cyclooxygenase
	P03E08	Tibolone	Undefined
	P03F04	Fenspiride hydrochloride	Bradykinin receptor
	P03F06	Mefexamide hydrochloride	Undefined
	P03G02	Norfloxacin	Topoisomerase II
	P04G08	Oxytetracycline dihydrate	30S unit
	P04H07	Glibenclamide	Sulfonylurea receptor K-ATP channel
	P05E07	Ambroxol hydrochloride	NO-dependent activation of soluble guanylate cyclase
	P07H06	Nifuroxazide	Undefined
	P07H07	Mirtazapine	Catecholaminergic receptor
	P12F02	Propofol	GABA-A receptor
4th (might have high cell toxicity)	P08E06	Carbadox	Undefined
	P08E09	Oxibendazol	Tubulin beta chain
	P08G06	Cytarabine	30S unit
	P10A09	Sulfadimethoxine	Folic acid synthesis
	P12D07	Nitrendipine	Voltage-gated L-type Ca2+ channel
	P05H08	Clomiphene citrate (Z,E)	Androgen receptor
	P12G06	Salmeterol	Adrenergic beta-2 receptor
	P13C05	Oxfendazol	Undefined
	X01	Verteporfin	YAP inhibitor
	P14F07	Estriol	Estrogen receptor
	P13B03	Rosuvastatin	HMG-CoA reductase

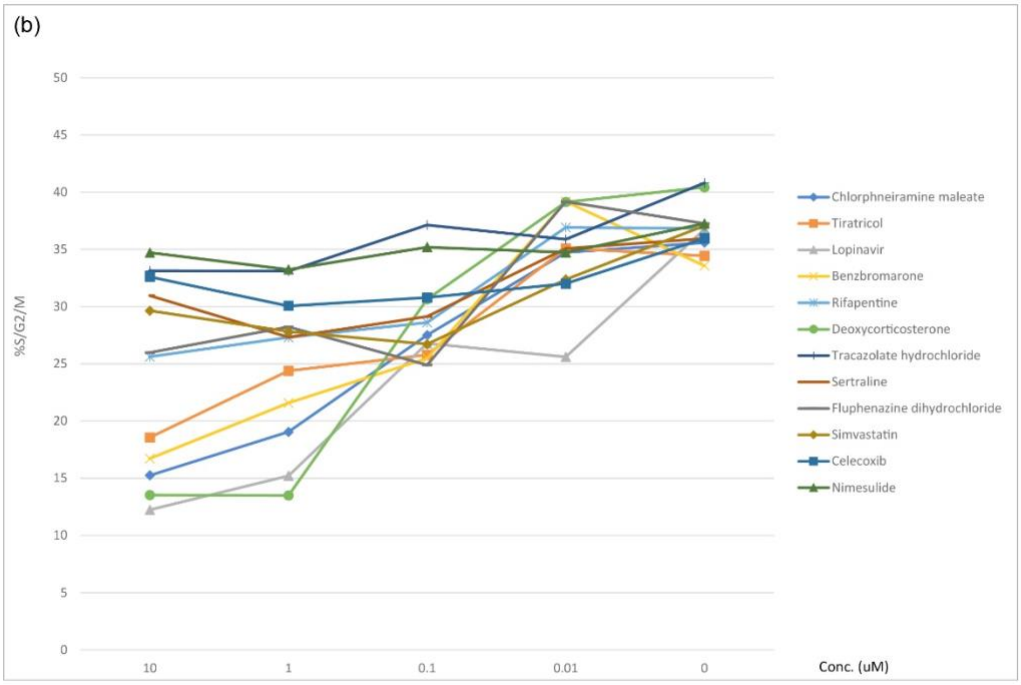
Figure 4.4.2: Results of confirmation screening with 'HIT' small molecules and further categorisation of drug candidates by priority. DMSO-treated control samples with an adjusted seeding density of LXSN and LXSN 16E6 are plotted as black and red solid dots, respectively, and the estimated control response of %s/G2/M to a cell density of each cell line is indicated as a big circle with matching colours. The big blue circle shows the estimated area of the ideal effect with small molecule treatment. Individual small molecules were then categorised with their effect closer to the blue circle area as higher priority candidates. The table summarises the small molecule and their previously proposed cellular targets with different colours distinguishing the priority classes.

4.5 Dose-dependent assay of top priority candidates

After running the primary and confirmation screen, in which a single concentration of the compound is used, a subset of compounds is selected for a more quantitative assessment. To start, we proceeded with 12 top priority candidates, namely Lopinavir, Celecoxib, Chlorpheniramine maleate, Tiratricol, Benzbromarone, Rifapentine, Deoxycorticosterone, Tracazolate hydrochloride, Sertraline, Simvastatin, Nimesulide and Fluphenazine dihydrochloride.

Experiments were carried out with a 10-fold dilution of the small molecules starting from 10uM, 1uM, 100nM and 10nM with 0.1% DMSO as vehicle control. Results were graphed as Cell Number, %S/G2/M and K10 intensity in response to drug concentrations on FUCCI NIKS LXSN 16E6 only (Figure 4.5.1). The experiment was also carried out with LXSN in the same condition as a control, while the results for LXSN is not shown as all concentrations did not affect normal NIKS function in the three phenotypes we measured. Different small molecules were distinguished by colour schemes. Generally, in Cell Number and %S/G2/M results, the candidates follow the dose-dependent trend, which confirms our HTS results of these small molecules are reliable inhibitors to HPV16E6 function. The results of concentration-dependency of the drugs in K10 expression are less obvious. This is likely due to our measurement of K10 intensity being collected as whole well data, and it is not as sensitive as cell number and %S/G2/M. However, there is a general trend that with the treatment of the small molecules, LXSN 16E6 cells express more K10 compared with vehicle control.


(a)
Cell Density
20000
19000
18000
17000
16000
15000
14000
13000
12000
11000
10000
10
1
0.1
0.01
0
Conc. (uM)
Chlorphneiramine maleate
Tiratricol
Lopinavir
Benzbromarone
Rifapentine
Deoxycorticosterone
Tracazolate hydrochloride
Sertraline
Fluphenazine dihydrochloride
Simvastatin
Celecoxib
Nimesulide


(b)
%S/G2/M
50
45
40
35
30
25
20
15
10
5
0
10
1
0.1
0.01
0
Conc. (uM)
Chlorphneiramine maleate
Tiratricol
Lopinavir
Benzbromarone
Rifapentine
Deoxycorticosterone
Tracazolate hydrochloride
Sertraline
Fluphenazine dihydrochloride
Simvastatin
Celecoxib
Nimesulide

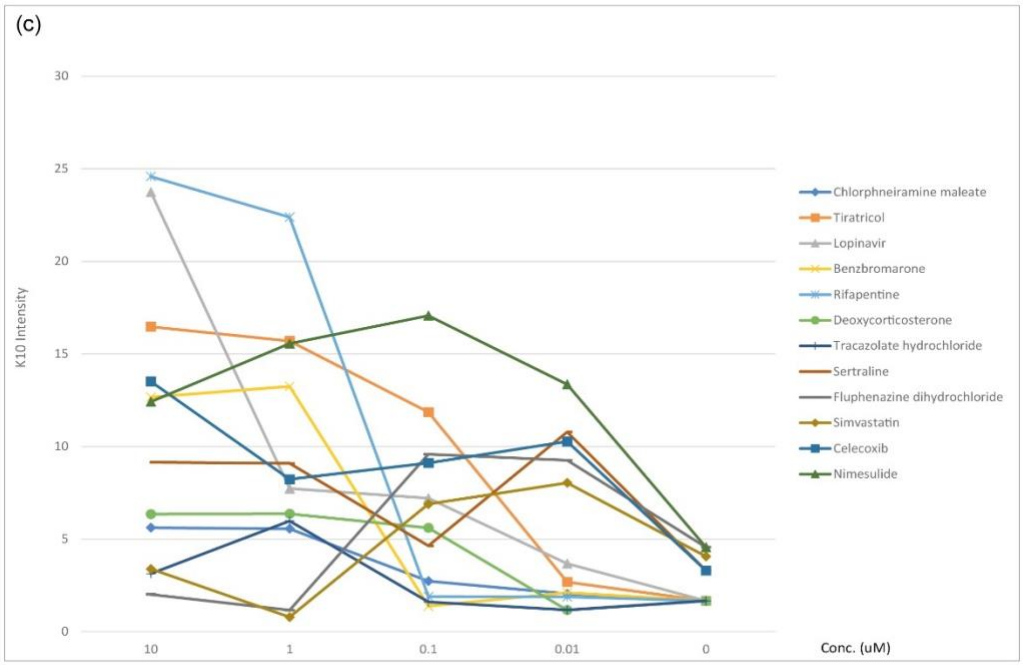


Figure 4.5.1: Dose-dependent response of top priority HPV16E6 inhibitors on NIKS LXSN 16E6, with small molecule concentrations of 10-fold dilutions (10uM, 1uM, 100nM, 10nM) and vehicle control of 0.1% DMSO effect on (a) Cell Number, (b) %S/G2/M and (c) Total K10 intensity. The colour scheme indicates different small molecules.

Among the 12 top priority candidates, Lopinavir and Celecoxib exhibited the most promising dose-dependent response on inhibition of HPV16E6 function in NIKS in the follow up analysis (Figure 4.5.2). Based on these results, Lopinavir and Celecoxib were chosen for further analysis in competition assays and to find the specific molecular pathways of action.

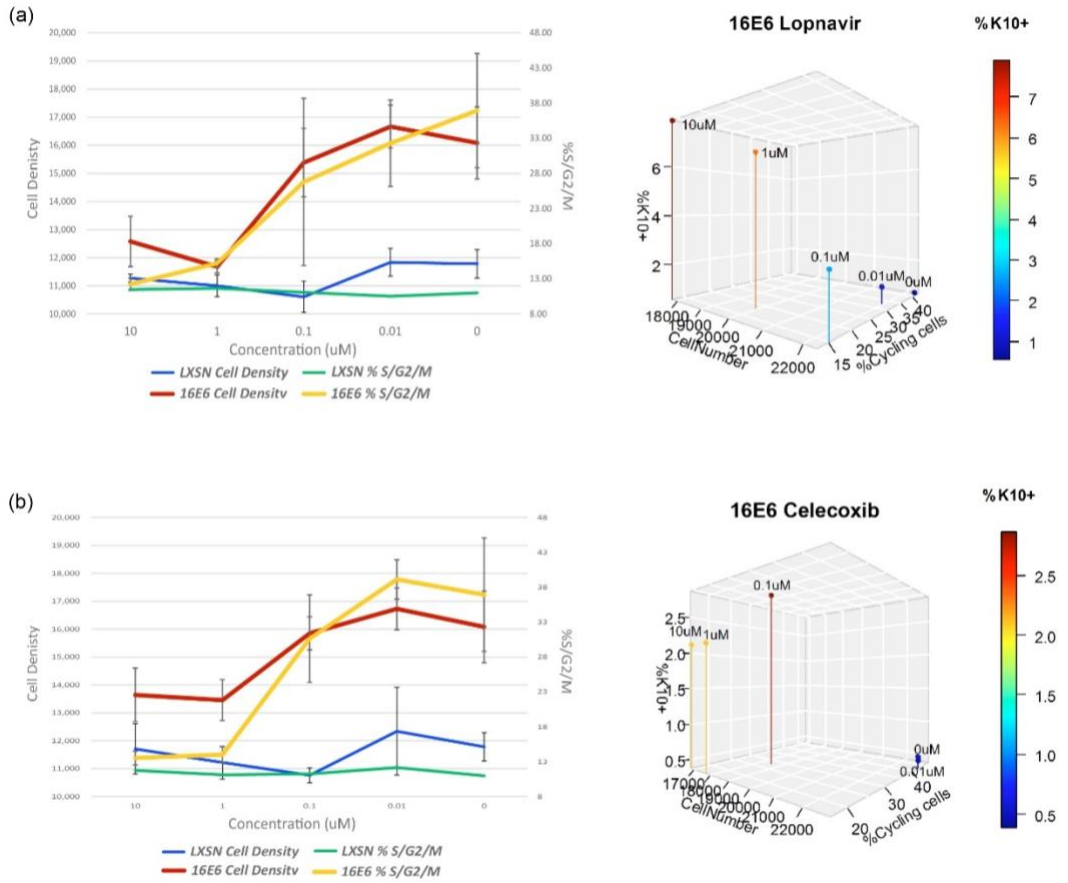
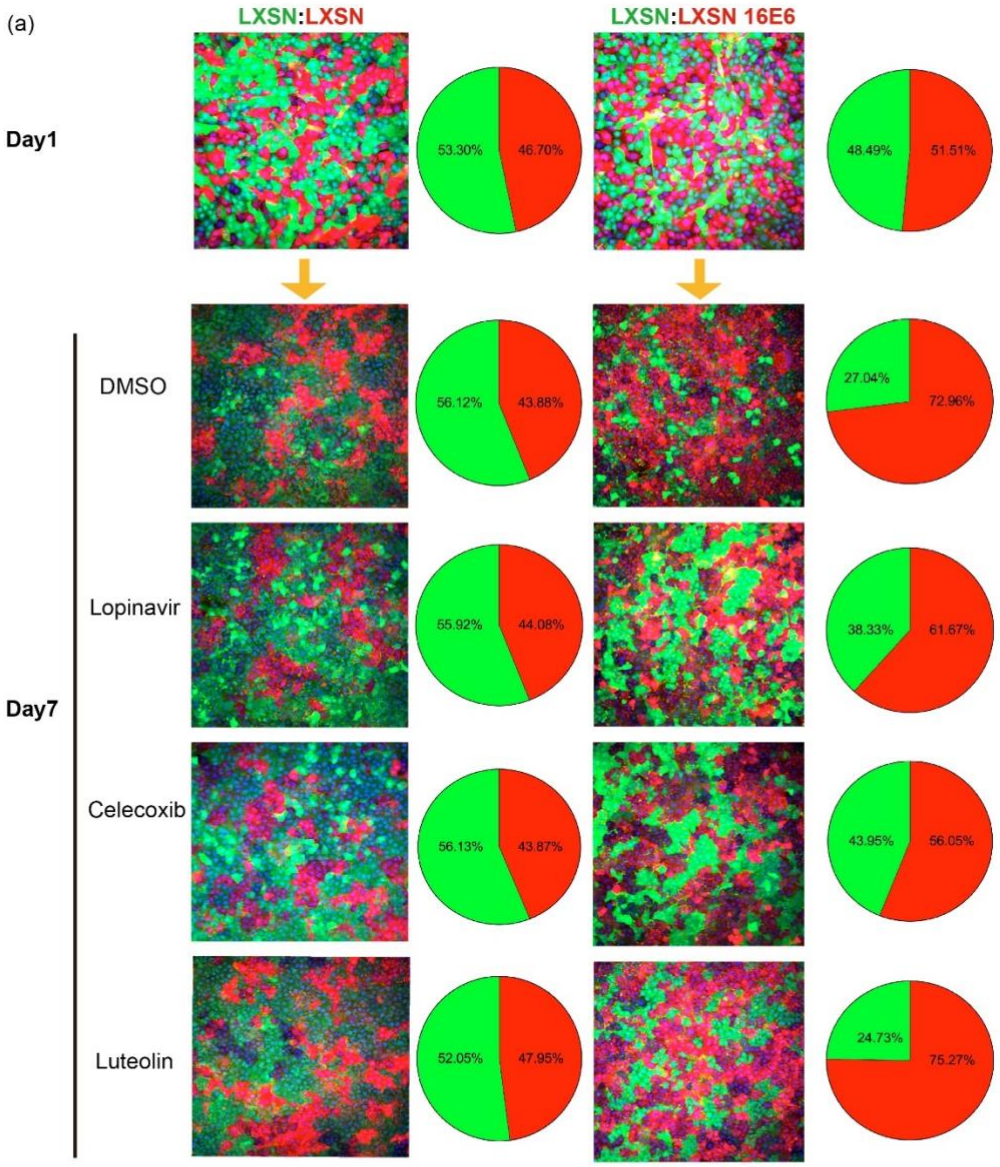


Figure 4.5.2: Dose-dependent response of NIKS 16E6 in Cell Density, %S/G2/M and K10 intensity treated with 10-fold dilutions starting with 10uM of (a) Lopinavir and (b) Celecoxib. Graph plotted with mean±standard deviation.

4.6 Competition assay screening of top priority candidates

So far, the effect of Lopinavir and Celecoxib in specific inhibition of HPV16E6 modulations in NIKS on cell density, %S/G2/M and K10 expression is reliable and reproducible. While in the epithelial tissue, infected and uninfected cells are in contact and interact with each other. From Section 3.6.2, the competition assay model suggested HPV16E6 expressing NIKS preferentially stay at the bottom layer in *in vitro* culture. Thus, with antiviral agents, the aim is to eliminate the infected cell population from the bottom layer of the epithelium. NIKS EGFP LXSN and NIKS mCherry LXSN were seeded at a 1:1 ratio to confluency (8x of total seeding

density on 96-well plates). Samples were collected 24hr after seeding to represent the starting condition, in which 10uM small molecules in DMSO were added every 48hr with FC media change for a continuous 7-day period (Figure 4.6a). At Day7, mixed cultures were fixed and used a confocal high-content imaging system (Opera Phenix) to zoom in the bottom layer only. Representative images were shown for vehicle control (DMSO), Lopinavir and Celecoxib. Additionally, Luteolin was also tested in parallel as it has been proposed to be HPV16E6 & E6AP binding inhibitor by computational modelling (Cherry *et al.* 2013). The proportion of EGFP and mCherry population is calculated and shown in pie charts. In the control group of EGFP LXSN with mCherry LXSN, the proportion of two coloured populations was constant over 7-day of *in vitro* culture and treatment of small molecules did not distinguish the two fluorescent protein expressions in NIKS LXSN. While when seeding EGFP LXSN versus mCherry LXSN 16E6 together at the same ratio, at Day7, the vehicle control group (DMSO) have a significantly higher proportion of mCherry cells at the bottom layer agrees with our previous results. With Lopinavir and Celecoxib treatment, the proportion of mCherry was less compared to the DMSO control suggesting a 16E6-mediated function in persisting at the basal epithelium layer is antagonised by these two small molecules. However, the putative E6&E6AP binding inhibitor Luteolin did not behave as a good anti-HPV candidate in the competition assay model. This is likely due to it affecting both NIKS LXSN and NIKS LXSN 16E6 in cell growth in the same manner through inhibition of normal E6AP functions (Figure 4.3.2). Therefore, it may lack specificity against E6-mediated homeostatic modulations. A total of 18 replicates were measured for each condition, and the proportion of mCherry LXSN 16E6 at the Day7 bottom layer is calculated and plotted (Figure 4.6b). Data shown as mean$\pm$standard deviation, and One-Way ANOVA analysis was used for measuring statistical differences between small molecule treated samples against the DMSO-treated control group. Both Lopinavir and Celecoxib treatments showed statistically significant downregulation of the mCherry 16E6 population after 7-day with p-values<0.0001. In comparison, Luteolin slightly shifts the balance of LXSN and 16E6 compared with DMSO control, with a marginal p-value of 0.0425.


(a)
LXSN:LXSN
LXSN:LXSN 16E6
Day1
53.30%
46.70%
48.49%
51.51%
Day7
DMSO
56.12%
43.88%
27.04%
72.96%
Lopinavir
55.92%
44.08%
38.33%
61.67%
Celecoxib
56.13%
43.87%
43.95%
56.05%
Luteolin
52.05%
47.95%
24.73%
75.27%

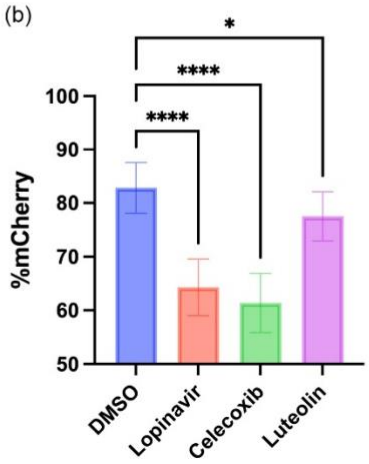

(b)
*
****
****
%mCherry
100
90
80
70
60
50
DMSO
Lopinavir
Celecoxib
Luteolin

Figure 4.6: Competition assay of EGFP LXSN: mCherry LXSN/16E6 treated with Lopinavir, Celecoxib and Luteolin for 7-day. (a) The same number of EGFP LXSN and mCherry LXSN/LXSN 16E6 were seeded at confluency, with representative images of Day1 and Day7 bottom layer. The pie chart shows relative EGFP and mCherry cell percentages in the images. Images captured with 40x objectives. Small molecules (10uM in DMSO) were added every 48hr with media change starting from 24hr after seeding for a total of 7-day. (b) Percentage of mCherry LXSN 16E6 at the bottom layer of Day7 mixed culture, with small molecule treatments. Data represented as mean±standard deviation, statistical significance calculated using One-way ANOVA analysis (P-value style as GP).

4.7 Pathway analysis of small molecule effects (HIV protease inhibitor and COX-2 inhibitor)

The phenotypic analysis has shown Lopinavir and Celecoxib are reliable HPV16E6 inhibitors on NIKS. Next, we aimed to find out the molecular pathways governing the effect. Lopinavir was used as anti-retroviral therapy targeting HIV-1 protease for inhibition. Its design is based on the "peptidomimetic" principle, wherein the molecule contains a hydroxy-ethylene scaffold which mimics the normal peptide linkage (cleaved by HIV protease) but which itself cannot be cleaved, thus resulting in the production of immature, non-infectious viral particles (de Clercq 2009). Celecoxib is a nonsteroidal anti-inflammatory drug (NSAID) which selectively inhibits cyclooxygenase-2 (COX-2) in a non-competitive manner. COX-2 is expressed heavily in inflamed tissues, where it is induced by inflammatory mediators (Rajakariar *et al.* 2006).

These established pathways of Lopinavir and Celecoxib are not suggestive of their specificity against HPV16E6 functions. Firstly, by exploring the PCL list zoomed in on the anti-retroviral and COX-inhibitors classes' results from 1st screening assay (Table 4.7), the function of Lopinavir and Celecoxib is distinct from the general category. Next, by searching the literature, we purchased some extra small molecules which have more specificity compared with Lopinavir and Celecoxib for their current usage. By testing these pharmacologically similar

molecules, we aim to identify whether the anti-HPVE6 effect can be found generally in the same class of drugs.

(a)

Target Names	LXSN (Cell density)	LXSN (%S/G2/M)	16E6 (Cell density)	16E6 (%S/G2/M)
NS5B palm polymerase				
Dasabuvir	-5.009350611	10.83501187	-4.241126962	4.72898686
DNA polymerases				
Acyclovir	-1.90453918	0.763533777	0.45584911	-1.228058416
Ganciclovir	0.186490362	0.912258924	-1.962076616	-1.011184143
DNA viral polymerase				
Penciclovir	-1.63057537	0.640796939	-1.187051475	0.454243018
Valacyclovir hydrochloride	3.527253003	-0.645962172	0.119470452	-1.208520913
DNA viral synthesis				
Famciclovir	-0.076716392	-1.569350753	0.02018559	-1.786604533
HIV integrase				
Dolutegravir	-2.736931988	3.454659558	0.120837928	0.254891297
HIV-1 protease				
Lopinavir	-0.819329303	2.867485958	-4.455180509	-1.396214228
Nelfinavir mesylate	-5.079337768	-18.28885737	-7.326527048	-0.46975865
Inosine-5-monophosphate dehydrogenase				
Ribavirin	-1.127620151	-1.022684096	-2.584505352	-0.162253028
Non-nucleoside reverse transcriptase				
Nevirapine	-1.939672247	-0.428306988	-1.900999079	0.709715121
Protease				
Indinavir sulfate	-1.71708704	1.706020294	0.234931092	-0.768873523
Ritonavir	-3.413591018	3.19962607	0.604805211	0.078750783
Saquinavir mesylate	-3.146692463	1.974596299	-5.143335894	4.012240578
Reverse transcriptase				
Abacavir Sulfate	-1.384690204	4.396055243	-2.339353882	-0.265176798
Efavirenz	-3.106559391	8.344472111	-3.309125618	0.622932256
Undefined				
Amprenavir	-0.911334913	1.70263529	-0.806689938	-1.021747097
Darunavir	-1.738294236	-0.280576954	-3.461898608	-0.11819643
Delavirdine	-1.817308759	0.587365817	-2.425456062	0.09713969

(b)

Target Names	LXSN (Cell density)	LXSN (%S/G2/M)	16E6 (Cell density)	16E6 (%S/G2/M)
Cyclooxygenase				
(R) -Naproxen sodium salt	-0.824279276	-0.719391375	-1.31212142	1.587980865
(S)-Naproxen	-3.26571017	-3.854964762	-0.945103597	-3.399044658

Aceclofenac	-0.253210483	-0.709516434	-3.211563969	-3.746661219
Acemetacin	-2.260551255	0.26093029	-1.067928799	0.384489015
Acetaminophen	-1.050999704	-2.17870571	-0.019779898	-0.174229169
Acetylsalicylic acid	0.229780457	2.742857442	0.915080084	1.252288737
Acetylsalicylsalicylic acid	-1.203660071	-0.434658513	1.246436812	-2.728674217
Ampiroxicam	-1.603770934	0.902933529	-0.169436309	0.540862474
Azapropazone	-0.806781937	-0.466638615	-3.092331527	0.158142939
Bufexamac	-3.706066	10.18727776	-3.394365824	5.583149844
Carprofen	-0.927840473	-2.401336508	1.458969594	-2.509904371
Diclofenac sodium	-1.830044858	0.470179543	-1.497028162	-0.623057693
Diflunisal	-4.156425099	2.571370616	-1.950227058	-1.590346098
Etodolac	-4.727762859	-0.734002655	-6.888572863	-8.753380321
Etofenamate	-1.171290185	0.198540695	-3.388381347	1.649681419
Felbinac	0.224902497	-0.143632219	-0.520785891	-1.882494476
Fenbufen	-2.534481318	0.547805269	-2.947679792	-0.772126247
Fenoprofen calcium salt dihydrate	-2.672105373	1.238582747	-3.78884775	0.461199731
Fentiazac	-2.109164953	1.503321888	-1.897658161	1.443204009
Flufenamic acid	-0.857524056	-1.135340077	-2.957032648	-2.272448147
Flurbiprofen	-0.421136991	0.220419462	-2.786797868	-1.214360843
Fosfosal	0.217884468	-0.442106457	-0.089976226	-0.938503532
Indomethacin	-1.457693686	0.803925024	-0.773028331	0.964254714
Indoprofen	-0.345335168	0.460340483	-1.022159674	-1.471025895
Isopyrin hydrochloride	0.232557433	-2.361615736	-1.371518024	-1.704097119
Isoxicam	-1.876811192	-3.4258955	-0.632639914	-3.451283993
Ketoprofen	-2.02305967	-0.868258712	-0.912097288	-3.997925847
Ketorolac tromethamine	-0.45662022	1.317718964	-2.664626205	-1.58264609
Meclofenamic acid sodium	-0.029753799	-1.144903968	-3.419999066	-3.2887141
Mefenamic acid	-1.166587097	-1.094927893	0.523923	-3.98276367
Mesalamine	0.072012684	-2.836442926	-1.447925571	1.185151499
Nabumetone	-0.980303744	7.623395203	-4.087887162	-0.550107671
Nifenazone	-4.012495043	2.97418287	-2.67043439	-1.816556185
Niflumic acid	-2.593941263	1.387480452	-1.078680378	1.265963096
Nimesulide	-1.403522143	6.864978608	-5.279881472	4.858097297
Oxaprozin	0.443872027	-0.557546244	0.057491395	0.149425841
Oxyphenbutazone	0.356121853	0.749599368	-1.832534725	0.360114692
Phenacetin	-4.526343182	-1.372248752	-3.747645518	4.644142904
Phenylbutazone	-0.521271975	-0.683941424	-0.932230335	-0.761451311
Piroxicam	-2.0264141	1.088669074	-2.751269796	-2.090293911
S-(+)-ibuprofen	-0.095058348	3.296722956	-1.758843946	-1.155083931
Sulindac	-0.244322958	-1.153867477	-0.58886944	-0.89564567
Suprofen	0.265678481	-0.271720082	0.723230394	-0.871398706
Tenoxicam	-4.211015226	0.30511713	-3.103717295	4.204644043
Tiaprofenic acid	-3.117661591	-0.19167083	-0.801980796	1.530850279
Tolfenamic acid	-0.928225131	-0.031292406	-5.614580138	0.911702712
Tolmetin sodium salt dihydrate	-0.320969593	1.014178436	-1.035247126	-0.682882344
Triflusal	-2.641960653	-0.988848045	-2.420171304	2.061117976

Zomepirac sodium salt	0.188561573	-0.098208542	-1.156905765	-1.068222921

Target Names	LXSN (Cell density)	LXSN (%S/G2/M)	16E6 (Cell density)	16E6 (%S/G2/M)
Cyclooxygenase-1				
Meloxicam	-2.274737248	2.880293412	-3.296820814	0.349411039

Target Names	LXSN (Cell density)	LXSN (%S/G2/M)	16E6 (Cell density)	16E6 (%S/G2/M)
Cyclooxygenase 2				
Celecoxib	-1.839649394	0.394839515	-3.245394317	0.443380412
Etoricoxib	-1.841935736	10.21868761	1.053910119	-0.461080394
Rofecoxib	-2.09040763	3.726995766	-3.560716615	6.865402241
Valdecoxib	-3.425732029	0.437043327	-1.08536562	-0.308335235

Table 4.7: List of small molecules from PCL categorised by molecular targets of (a) Anti-retroviral therapy and (b) Non-steroid anti-inflammatory drugs (NSAIDs)/Cyclooxygenase inhibitors. The top candidates, Lopinavir and Celecoxib from the HTS assay, are highlighted in orange colour. Individual effects were calculated as the deviation from their mean and divided by the standard deviation set by DMSO-treated data of each set. Data highlighted in green as the effect is $\geqslant$3xsd from mean, and data in red as the effect is $\leqslant$-3xsd from mean.

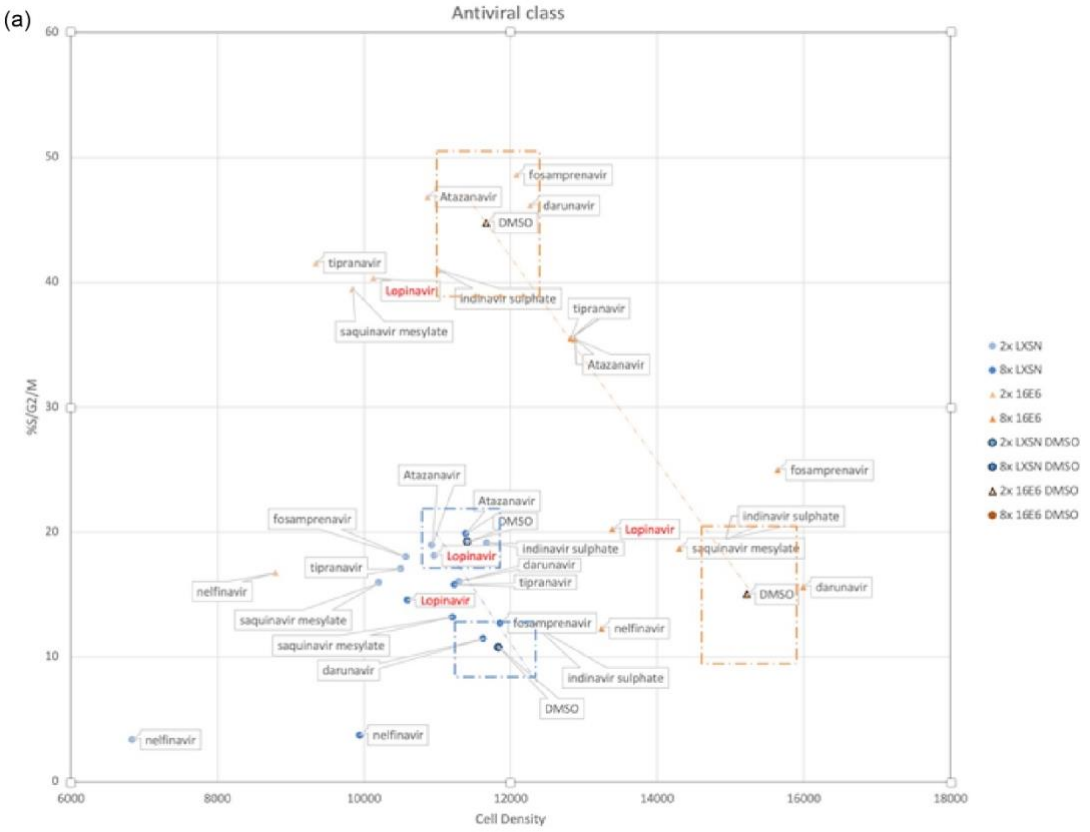
For more detailed analysis, we seeded the LXSN and LXSN 16E6 at low- (2x) and high-cell density (8x) and plotted the mean of DMSO-vehicle control values in blue and orange dots with black lining distinguished by cell types (Figure 4.7.1a/b). Then the mean$\pm$3xstandard deviation range is calculated and graphed as dotted boxes, with a trend line representing the natural %S/G2/M response with cell density changes. Individual effects on cell density and %S/G2/M of anti-retroviral drugs (Figure 4.7.1a) and COX-2 selective inhibitors (Figure 4.7.1b) are plotted with Lopinavir and Celecoxib highlighted in red texts. For the anti-retroviral class, individual molecules have distinct functions on NIKS cell growth and proliferation with no obvious similarities that can be concluded from their proposed targets. While the COX-2 selective inhibitors similar to Celecoxib generally have similar effects in inhibition of NIKS LXSN 16E6 cell growth, and as a result, the %S/G2/M was higher in response to lower cell density. However, Celecoxib is still more promising amongst the class as it only significantly impacts on post-confluent (8x) LXSN 16E6 population with no cell toxicity to LXSN at all densities. K10 expression further supports the observations from cell density and %S/G2/M (Figure 4.7.1c),

as the Lopinavir and Celecoxib treated LXSN 16E6 population express a similar level of K10 compared with the DMSO control group but at 10% lower cell density. While the nature of K10 expression analysis is less quantitative compared with the other phenotypes, it only serves a supportive role for the small molecule analysis.

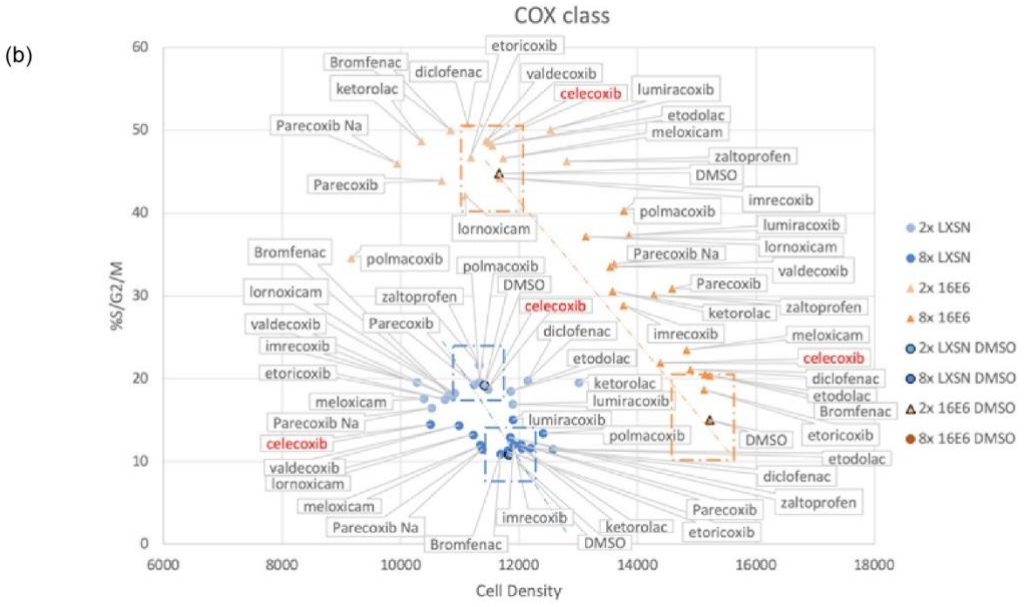
In order to compare the small molecule effect on LXSN and LXSN 16E6 more directly, the competitive assay was performed for all the HIV-protease inhibitors and COX-2 selective inhibitors. NIKS EGFP LXSN and NIKS mCherry LXSN 16E6 were seeded at a 1:1 ratio (50% of each population) at high-cell density, and mixed cultures were maintained up to 7-day with small molecules added every 48hr starting 24hr after seeding. NIKS EGFP LXSN and NIKS mCherry LXSN group were compared in parallel as control (data not shown). On Day7, cells were fixed, and the proportion of mCherry LXSN 16E6 cells at the bottom layer of the mixed culture was counted (Figure 4.7.1d). When there is no competitive advantage between the EGFP and mCherry populations, the proportion of mCherry at the bottom layer of the Day7 sample should stay similar to the seeding ratio of 50%. When the mCherry population have viral protein 16E6 expression, the population preferentially stays at the bottom layer, and as a result, 86% of the bottom layer was occupied by the E6 expressing cells after 7-day. With small molecule treatment, we aim to reduce the proportion of mCherry LXSN 16E6 population at the bottom layer. Amongst all the small molecules from the same class, there was no general effect in each group, with Lopinavir and Celecoxib still the most promising candidates for eliminating infected cells during viral persistence.

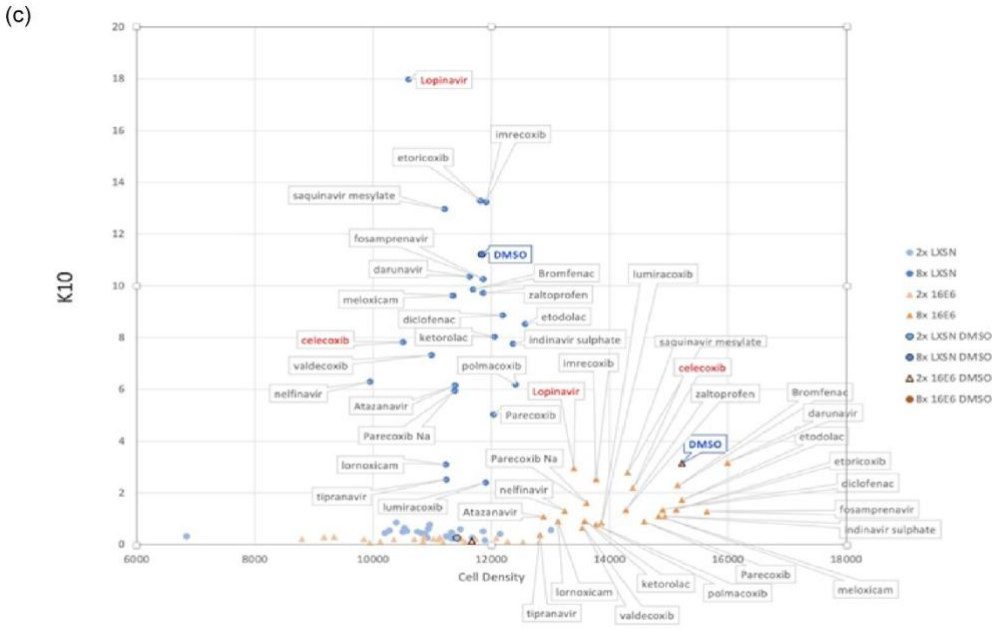
Combing these results, the anti-HPVE6 function of Lopinavir and Celecoxib is unique amongst those other small molecules which have similar drug targets and mechanisms under current usages. The next step was to evaluate their specificity against HPV16E6 modulated molecular pathways.

(a)


Antiviral class
%S/G2/M
Cell Density
2x LXSN
8x LXSN
2x 16E6
8x 16E6
2x LXSN DMSO
8x LXSN DMSO
2x 16E6 DMSO
8x 16E6 DMSO
fosamprenavir
Atazanavir
darunavir
DMSO
tipranavir
Lopinavir
indinavir sulphate
saquinavir mesylate
nelfinavir

(b)


COX class
%S/G2/M
Cell Density
2x LXSN
8x LXSN
2x 16E6
8x 16E6
2x LXSN DMSO
8x LXSN DMSO
2x 16E6 DMSO
8x 16E6 DMSO
etoricoxib
Bromfenac
diclofenac
valdecoxib
ketorolac
celecoxib
lumiracoxib
etodolac
meloxicam
Parecoxib Na
zaltoprofen
Parecoxib
DMSO
imrecoxib
polmacoxib
lornoxicam



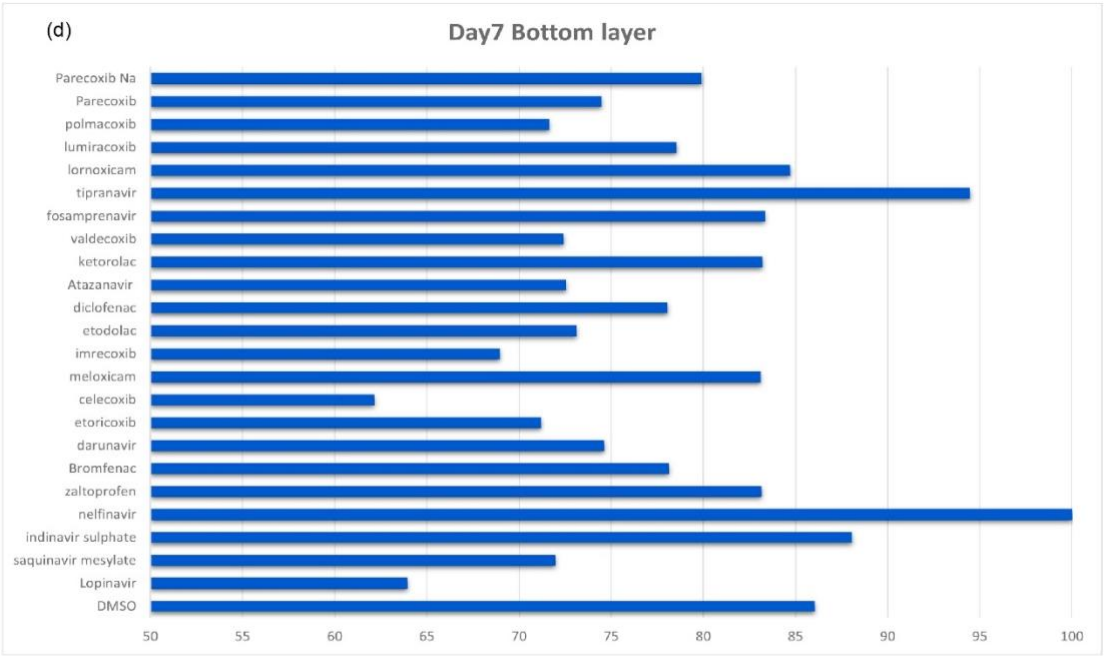


Figure 4.7.1: Effect of same class small molecules of HIV-protease inhibitor and COX-2 selective inhibitors on NIKS. (a) Cell density and %S/G2/M response to HIV-protease inhibitor treatment for 24hr. (b) Cell density and %S/G2/M response to COX-2 selective inhibitor treatment for 24hr.With DMSO-treated control data, the mean$\pm$3xsd range is calculated and

indicated as the boxed area on the scatter plot and trendline, suggesting %S/G2/M response with cell density changes. (c) K10 expression in response to cell density with/out small molecules treatments. Individual small molecules are indicated in texts, with the top priority candidates lopinavir and Celecoxib highlighted in red. (d) Competition assay of small molecules treated LXSN vs LXSN 16E6. EGFP LXSN and mCherry LXSN 16E6 were seeded at equal proportions (50% as control). The proportion of mCherry LXSN 16E6 cells at the bottom layer were calculated on Day7 and plotted.

Since the mode of action of Lopinavir and Celecoxib against HPV E6 phenotypes remains unclear, I tried to evaluate the established pathogenic functions of 16E6 and how it changed by these two small molecules. One of the major pathogenic functions of HPV16E6 is through downregulation of p53 in keratinocytes (vande Pol & Klingelhutz 2013). Homeostatic regulation of p53 oscillates during the cell cycle and directs cellular fates by different pulse intensity and frequency regulations (Purvis *et al.* 2012). After cells acquired the terminal fate, p53 expression decreased (Wawryk-Gawda *et al.* 2014). Individual cell behaviour is better observed with immunofluorescence than with the Western blot technique (Figure 4.7.2). The experiment was designed as seeding NIKS LXSN and NIKS LXSN 16E6 cells at high-density (8x) on 8-chambered slides; media changed 24hr after seeding. Then 10uM small molecules in DMSO were added in 24hr after changing the media. Samples were collected 6hr and 24hr after adding the small molecules by fixation. Previous results show after 24hr of small molecule treatment, some of the LXSN 16E6 cell fate has changed from proliferation to differentiation marked by exiting cell cycle (lower %S/G2/M with FUCCI) and K10 expression. Thus, if the small molecule function in antagonising 16E6 by alteration of p53, the phenotype in p53 expression change is more likely to be observed prior to 24hr of drug added in. On post-confluent NIKS samples, nuclear-localised p53 can be observed in the LXSN population but not the LXSN 16E6. With the Lopinavir and Celecoxib treatment, there were no obvious changes in the p53 level in NIKS LXSN under all experimental conditions. For Lopinavir, there were no changes in the expression level of p53 in LXSN 16E6 cells at 6hr and 24hr after adding the drug. While a slight upregulation of nuclear p53 is observed in Celecoxib-treated NIKS LXSN 16E6 after 6hr, then this phenotype diminished after 24hr.

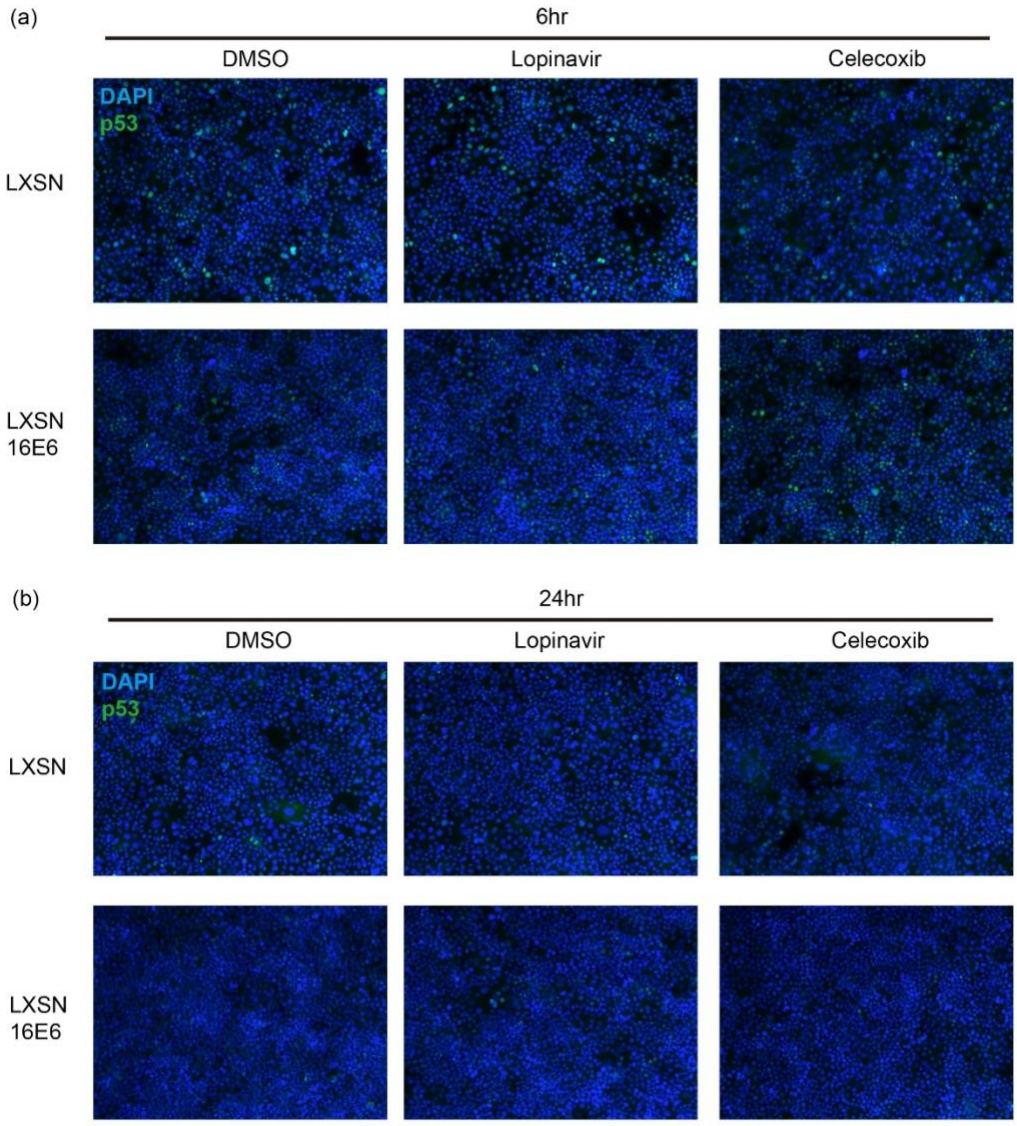


Figure 4.7.2: Immunofluorescence of p53 on NIKS LXSN and LXSN 16E6 treated with Lopinavir and Celecoxib for 6hr (a) or 24hr (b) with DMSO as the negative control.

4.8 Discussion

In this section, by utilising the high-content imaging machines (ArrayScan and Opera Phenix) and our *in vitro* modelling system of FUCCI with K10 expression in NIKS, I carried out a high-throughput screening project against HPV16E6 by repurposing 1200+ the current FDA-approved small molecules (Prestwick Chemical Library). The phenotypic changes in saturation cell density and %S/G2/M in NIKS with 16E6 expression by LXSN vector were shown to be reliable HTS assay systems. With the 1st screening, we filtered a total of 167 small molecules

that have potential anti-HPV functions in downregulating saturation density and/or %S/G2/M, specifically in the 16E6 population without cellular toxicity to NIKS LXSN. With further confirmation analysis with a range of cell density and concentrations and competition assay system, Lopinavir and Celecoxib appeared to be the top priority candidates. By analysing a range of drugs from the same class of HIV-protease inhibitors and COX-2 selective inhibitors, Lopinavir and Celecoxib still behave distinctly amongst the structurally/functionally similar small molecules in inhibition of HPV16E6 modulated changes in NIKS phenotypes, suggests undiscovered molecular targets of these two drugs.

The most successfully exploited antiviral targets have been viral enzymes, specifically the viral nucleotide kinases and polymerases. Lopinavir selectively inhibits with high potency the HIV-1 protease-mediated cleavage of gag and gag-pol polyproteins, resulting in the production of immature HIV virions which are non-infectious (Croxtall & Perry 2010). However, most HPV proteins lack intrinsic enzymatic activity. The potential of using lopinavir against HPV-associated lesions was initially found as effective and adherent anti-retroviral therapies were associated with reduced oncogenic HPV prevalence and incidence (Minkoff *et al.* 2010) and increased regression of cervical lesions (Minkoff *et al.* 2001). Meanwhile, *in vitro* studies have shown that the introduction of high-dose (15mM) lopinavir into E6-transfected C33A cells (HPV-negative cervical cancer) results in an increase in cellular p53 level (Hampson *et al.* 2006). Some evidence suggests the anti-HPV function is considered to be the selective inhibitory effect on proteasomes and prevent E6-E6AP mediated p53 degradation via proteasome pathway in SiHa cells (Batman *et al.* 2010; Zehbe *et al.* 2011). A chemical mapping study suggested that the main changes in E6 expressing cells treated with HIV protease inhibitors happen in the nucleus and might have other undiscovered functions (Kim *et al.* 2010). However, my experiments suggested controversial results whether Lopinavir targets HPV16E6 function via the proteasome pathway in NIKS. One of the major reasons is our cell cycle indicator FUCCI expression is also dependent on cellular proteasomal degradation of Geminin and Cdt-1 proteins, while the FUCCI signal behaves normally with lopinavir treatment at a range of concentrations. It is undoubted that Lopinavir's anti-HPV effect is reproducible both

in *in vitro* and *in vivo* conditions, but the specificity against the HPVE6-modulated molecular pathway remains unclear.

Celecoxib belongs to the Non-Steroid Anti-Inflammatory Drugs (NSAIDs) family, is a selective inhibitor of Cyclooxygenase-2 (COX-2),and has been administered for the treatment of rheumatoid arthritis and osteoarthritis since the last century (Goldenberg 1999). COX-2 is an inducible form of cyclooxygenase expressed at sites of inflammation. Deregulation of COX-2 expression leads to an increased abundance of its principal metabolic product, prostaglandin E2 (PGE2), affecting multiple aspects of cell physiology required for tumour development and maintenance (Greenhough *et al.* 2009). COX-2 expression level has been associated with early cervical carcinogenesis (Saldivar *et al.* 2007) and related to disease severity and HPV detection in squamous lesions of the cervix (Sarian *et al.* 2006). There was little evidence of the direct relationship between HPV viral protein expression and COX-2 induction. Here, with a low level of HPV16E6 expression by the LXSN vector, we found nuclear expression of COX-2 is elevated at the tissue level (Figure 4.8). COX-2 is primarily localised to the cytoplasmic membrane where it functions to metabolise PGE2. Nuclear localisation of COX-2 was associated with the stemness of several cancer cells while lacking evidence for its regulation and activity (Bocca *et al.* 2011; Cheng *et al.* 2018; Thanan *et al.* 2012). An experiment with HeLa cells treated with 25uM Celecoxib suggested that COX-2 has a potential role in binding p53, so inhibiting COX-2 and releasing p53 together with inhibiting E6 transcription could elevate cellular p53 levels (Saha *et al.* 2011). However, in my project, the expression system for HPVE6 protein is different, so this finding is less inspiring. Overall, there are some uncertainties about Celecoxib's anti-HPV mechanism.

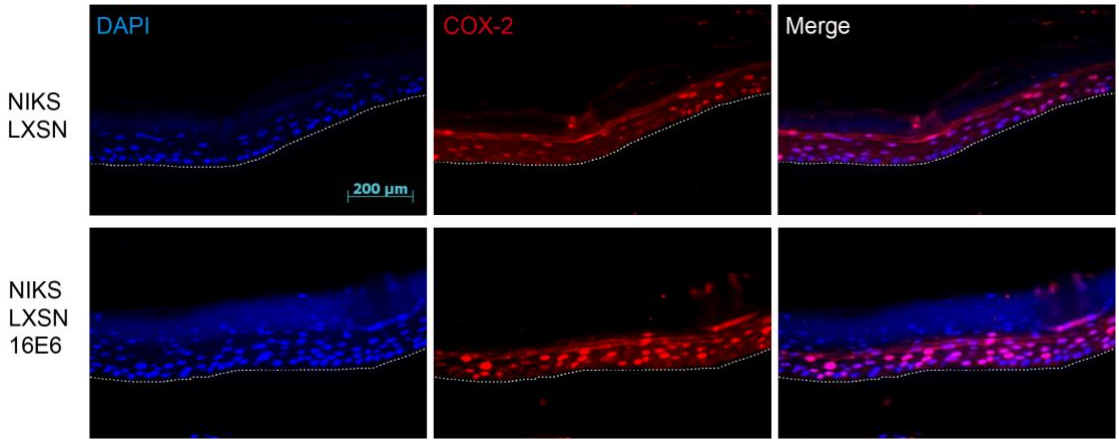


Figure 4.8: Immunohistochemistry of COX-2 on NIKS LXSN and NIKS LXSN 16E6 organotypic raft cultures. The dotted line represents the basal lamina. Images captured with 20x objectives.

With antiviral small molecules, I examined the p53 expression regulation in NIKS LXSN 16E6 during persistence and was reminded of the transient changes in molecular targets, which remain elusive. Therefore, I decided to study 16E6 and its molecular targets for modulating normal keratinocyte homeostasis and combine the results with the drug discovery project in the future.

5. Identification of epithelial basal-suprabasal transition stimuli and 16E6 molecular targets

5.1 Introduction

The previous two chapters focused on HPV16E6-mediated changes in basal epithelial homeostasis and HTS-based anti-HPV agents targeting persistently infected papillomavirus reservoirs for clearance. However, since the datasets were based primarily on phenotypic analysis, it is unclear which host molecules 16E6 modulates is important for virus' basal layer persistence, so progress in identifying the molecular targets of small molecule inhibitors is slow.

Most HPV E6 research has been targeted towards understanding the role of the E6 protein in carcinogenesis and the viral life cycle. The most well-established cellular targets for HPV16E6 protein are p53 and PDZ-domain-containing proteins, whose degradations are facilitated by E6 and responsible for cancer cell transformation and metastasis, respectively (vande Pol & Klingelhutz 2013). The next aim would be to understand the importance of these cellular targets for normal basal epithelial homeostasis. With the same reliable *in vitro* model system using NIKS described, I generated HPV16E6 functional mutant and corresponding RNA knockdown cell lines to explore the natural functions of p53 and PDZ-domain-containing proteins in the epidermis.

Since the epithelium's basal layer is composed of heterogeneous cells, a small proportion has proliferative potential and undergoes constant turnover. The transition from basal to suprabasal is a complex process, including regulation in the cell cycle, extracellular matrix attachment with the basal lamina, cell-cell junction reformation and initiation of terminal differentiation fate. For the *in vitro* model system, NIKS in tissue culture also possess the same differentiation potential and pathways to form stratified epithelium. Currently, all the experiments were created based on broadly categorising the subpopulations as proliferating or committed to differentiation. It is not clear how the transition triggers the basal cell to

delaminate and acquire the terminal differentiation cell fate. Particularly for my interest, based on the previous results, the expression of 16E6 itself appeared to delay differentiation of keratinocytes. The virus may have achieved these functions by altering the threshold of cell fate decision or triggering redundant pathways.

Single-cell RNA sequencing presents a powerful technology to simultaneously quantify transcriptional variability for thousands of genes across hundreds of cells. With scRNA sequencing, the resolution of our experiment increased. Under the experimental setting of the post-confluent NIKS population with or without HPV16E6 expression, the population can be finely subclustered to explore novel basal-suprabasal transition markers.

This chapter assesses HPV16E6 targeted cellular proteins' functions in terms of basal epithelial homeostasis based on the known functions of HPV16E6 targeted cellular proteins (p53 and PDZ-containing proteins). Next, I modelled for potential host-virus interactions based on recent advances in scRNA sequencing and bioinformatic analysis.

5.2 p53 and PDZ proteins both contribute to basal cell maintenance

Previously, the function of HPV16E6 protein in maintaining infected basal cell 'stemness' were described and can be reverted by small molecule treatments. The next aim is to identify host proteins targeted by HPV16E6 for modulation of the cell environment. To start with, the most well-established functions for HPV16E6 pathogenicity are the degradation of p53 and some PDZ-domain proteins in the epidermis (van de Pol and Klingelhutz, 2013). Mutants of 16E6, which are deficient for p53 or PDZ protein degradations, are described by Murakami et al.

NIKS harbouring 16E6 wild-type (16E6) and mutants unable to degrade p53 (16E6SAT), PDZ-domain proteins (16E6dPDZ) and double mutants (16E6SATdPDZ) expressed by LXSN vector were generated by lentiviral transduction with empty vector LXSN as a negative control. Cell lines were then validated by western blot of sub-confluent NIKS lysates (Figure 5.2a). Staining

with p53 and MAGI (as one of the main PDZ protein targets of 16E6 from a systemic analysis by Kranjec and Banks) showed that the mutants SAT/dPDZ can restore p53/PDZ-protein levels respectively in 16E6 expressing background without affecting the other functions.

FUCCI NIKS LXSN/16E6/mutant cells were seeded at a range of densities on 96-well plates, and media was changed 24hr post-seeding and cells were then allowed to expand for another 48hr before analysis. Samples were paraformaldehyde-fixed and immunostained with DAPI for nucleus recognition and Keratin 10 (K10) as an early differentiation marker. With FUCCI, the populational proliferation rate is indicated as %S/G2/M and plotted against the number of nuclei of individual cell lines (Figure 5.2b). NIKS proliferation negatively responds to cell density increase. It reaches a saturation density (confluency) when the cell number of the bottom layer remains constant with a baseline proliferation rate of roughly 5-8%. With HPV16E6 expression, cell proliferation is increased at all densities, and the population reaches confluency at higher cell numbers. Results of the mutant analysis suggested the proliferation and density modulation largely depend on the degradation of p53, as the 16E6SAT cells behaved similar to the empty vector control LXSN. The 16E6dPDZ mutant also affects 16E6 function in driving cell cycle entry and modulating contact inhibition sensing mechanisms but to a much lesser extent compared with the 16E6SAT mutant. The double mutant 16E6SATdPDZ further impaired the 16E6 functions, with the proliferation rate at low cell density restored to the same level as the empty vector control LXSN. Early differentiation marker staining of K10 of FUCCI NIKS cells showed similar results (Figure 5.2c), in which p53 seems to play a more important role than PDZ proteins in cell density sensing and cell fate determination. The saturation density of each cell line is further analysed by One-Way ANOVA in relation to LXSN (Figure 5.2d) or LXSN 16E6 (Figure 5.2e) populations. Although the restoration of p53 and/or PDZ-domain proteins affects 16E6 function in cell cycle entry, there was still a statistically significant increase in saturation densities of the 16E6 mutant cell lines, suggesting there are unidentified 16E6 targets in modulating keratinocytes homeostasis.

Following this, I used the competition assay to determine how 16E6 expressing cells preferentially stay at the basal layer. Two fluorescent cell lines EGFP and mCherry cells, were

transduced to express either empty vector LXSN or 16E6/mutant. With the saturation density known from the FUCCI experiment, the EGFP and mCherry cells were seeded at confluency with a 1:1 ratio to start with (Day0). Then the mixed cell culture was maintained for 7-day with media change every other day from Day1. Samples were collected on Day1 and Day7, and the bottom layer is visualised by Confocal microscopy (Figure 5.2f). From the images, DAPI was used for nucleus recognition, and each nucleus was then categorised by its colour. The ratio of EGFP and mCherry cells were then calculated and shown as pie charts. The Day1 results indicated the starting condition of each experiment setting, with roughly 50% of the EGFP cells and 50% of the mCherry cells. After 7-day of growth, there is a great competitive advantage of 16E6 expressing cells (mCherry LXSN 16E6) over the LXSN control (EGFP LXSN) in occupying the bottom layer. As expected of two populations with different proliferation and differentiation rates, there is competition, and by altering these phenotypes as with 16E6SAT mutant, the competitive advantage is greatly impaired. However, in the 16E6dPDZ dataset, the population retained considerably higher proliferation and lower differentiation rate; it failed to compete with the empty vector LXSN population. The results suggested that HPV16E6 targets PDZ-domain proteins for viral persistence at the basal epithelial layer via a different pathway other than regulating cell cycle entry.

Many PDZ-domain-containing substrates targeted by HPV16E6 for degradation are membranous proteins involved in the regulation of cell polarity, assembly of cell junctions and cell attachments (Thomas *et al.* 2008), which are actively regulated during basal cell apical extrusion as a result of cell competition. Therefore, our next aim is to narrow down the host PDZ-domain-containing targets, which are most important for the viral-induced competitive advantage.

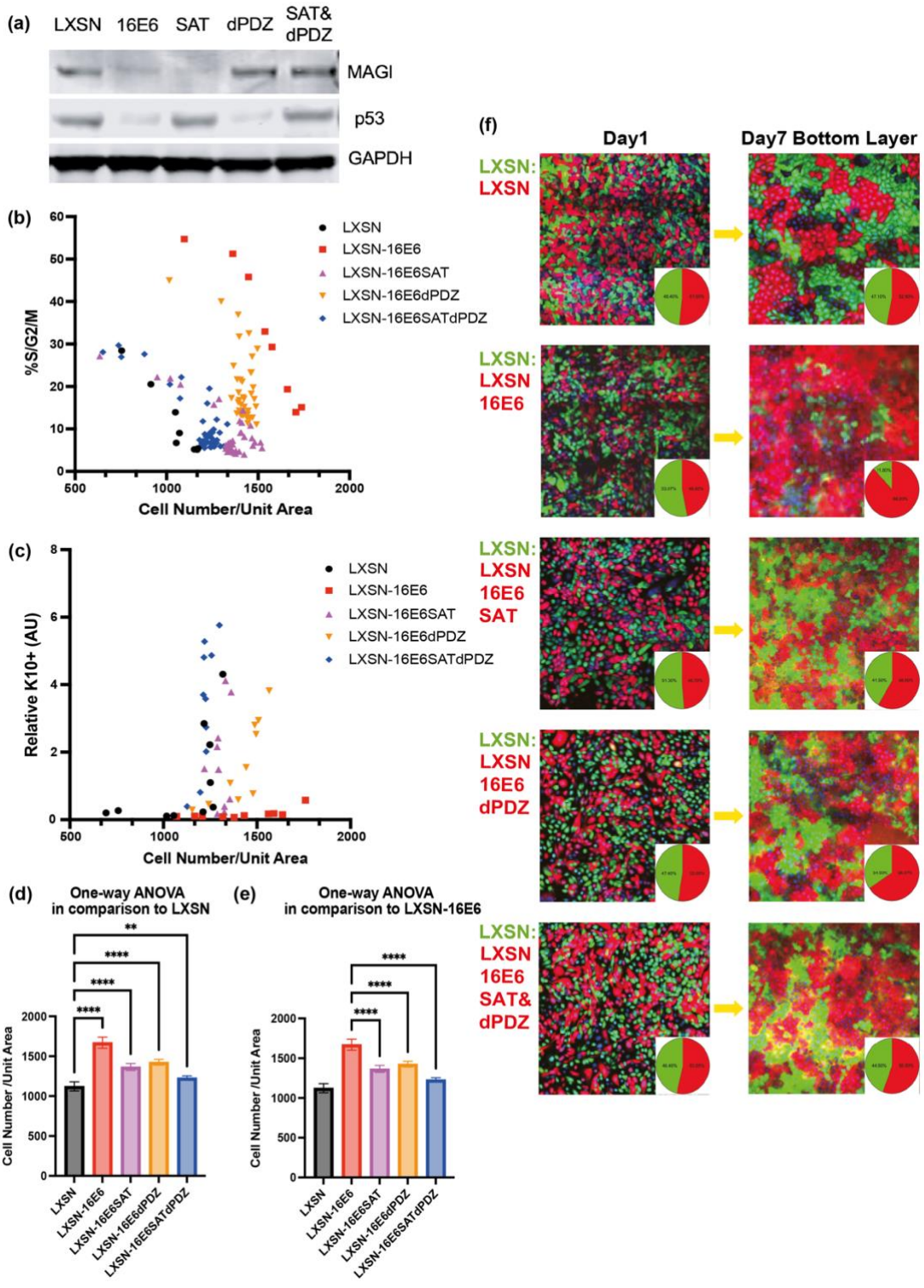


Figure 5.2: Analysis of HP16E6 mutants suggests both p53 and PDZ-proteins are important targets for viral persistence. (a) Western blot of sub-confluent mCherry NIKS lysates of 16E6 and mutants, with empty vector LXSN as the negative control. Immunoblot membrane stained with MAGI, p53 for validating mutant functions and GAPDH was used as a loading control. (b)

Percentage of S/G2/M positive cells of FUCCI NIKS with HPV16E6 and mutant expression in response to cell number per unit area (420,000μm^2). (c) Relative populational K10 positivity of FUCCI NIKS with HPV16E6 and mutant expression in response to cell number per unit area (420,000μm^2). (d) Saturation density (cell number per unit area (420,000μm^2)) of FUCCI NIKS with HPV16E6 and mutants expression with One-way ANOVA analysis in relation with FUCCI NIKS LXSN (P-value style as GP). Data are shown as mean±SD. (e) Saturation density (cell number per unit area (420,000μm^2)) of FUCCI NIKS with HPV16E6 and mutant expression with One-way ANOVA analysis in relation with FUCCI NIKS LXSN 16E6 (P-value style as GP). Data are shown as mean±SD. (f) The same number of eGFP LXSN and mCherry LXSN/16E6/SAT/dPDZ/SATdPDZ were seeded at confluency, with representative images of Day1 and Day7 bottom layer. The pie chart shows relative eGFP and mCherry cell percentages in the images. Images captured with 40x objectives.

5.3 Systemic analysis of PDZ-domain-containing substrates of HPV16E6 involved in modulating basal epithelial homeostasis

5.3.1 MAGI and hPAR3 are the main targets for HPV16E6 for viral persistence

A number of PDZ-domain-containing substrates of HP16E6 for degradation were identified, namely DLG1 (Gardiol *et al.* 1999), DLG4 (Handa *et al.* 2007), MAGI (Glaunsinger *et al.* 2000), hPAR3 (Yoshimatsu *et al.* 2017), FAP/PTPN13 (Spanos *et al.* 2008), Scrib (Nakagawa & Huibregtse 2000) and PATJ (Latorre *et al.* 2005). RNA interference (shRNA) plasmids were gifted by Dr. Tohru Kiyono. Specificity to individual PDZ-domain-containing proteins was confirmed by plasmid DNA sequencing, and stable expression of RNA knockdown in NIKS was generated by lentiviral transduction. Protein level knockdown was validated by immunostaining on raft cultures (Appendix Figure 2). Luciferase knockdown was used as a negative control.

FUCCI NIKS expressing different shRNAs were seeded at different densities in 96-well plates and cultured for 72hr with media change at 24hr after seeding. Samples were then fixed with

paraformaldehyde and immunostained with DAPI for nuclear recognition and K10 as an early differentiation marker. By plotting the proportion of proliferating cells (%S/G2/M) and relative K10 positivity against cell density, respectively (Figure 5.3.1a/b), amongst the PDZ-domain proteins tested in this experiment, the MAGI/hPAR3/PATJ reduction in NIKS drive cell cycle entry and caused delay in commitment to differentiation similar to HPV16E6 function. Meanwhile, it is interesting to note that the knockdown of the DLG family (DLG1 and DLG4) negatively impacts cell 'fitness' as in decreased proliferation and early induced differentiation. FAP and Scrib play a minimal role in cell cycle regulation. Quantitative One-Way ANOVA analysis of saturation density of individual cell lines further supported the results observed in %S/G2/M and K10 (Figure 5.3.1c).

In the previous experiments, increased proliferation and/or decreased differentiation suggested the cell line's competitive advantage to stay at the bottom layer in the long-term mixed culture assay. However, the PDZ-domain containing protein knockdown suggested additional phenotypes in defining basal epithelial cell 'fitness'. Similar to previous competition assay experiments, EGFP or mCherry NIKS either expressing the negative control or the targeted RNAi were seeded at confluency at a 1:1 ratio (Day0). Samples were collected for analysis and Day1 and Day7, with media change every 48hr starting from Day1. Cell number was calculated by DAPI nucleus recognition, and each nucleus was then categorised into different coloured populations. The percentage of mCherry cells (%mC) was calculated by dividing the number of mCherry cells by the total cell number at the bottom layer. In Day1 results, roughly 50% of mCherry cells are in the mixed culture (data not shown). Then the effect of each individual PDZ-domain-containing protein RNAi in competition with negative control Luciferase RNAi was processed as normalising the %mC at Day7 to the %mC at Day1 with individual experiment settings (Figure 5.3.1d). Since there was no competition between EGFP LuciRi and mCherry LuciRi, the relative ratio of %mCherry remains the same and normalised as 1. Firstly, as expected from previous experiments, the absence of MAGI and hPAR3 resulted in the preferential expansion of the population at the basal layer. Although PATJRi in NIKS actively drives cell proliferation, it is actually outcompeted by the normal NIKS (as the pseudo-RNAi to Luciferase), with statistically significant lower occupancy after 7-day

of growth at the bottom layer. Also, the DLG-family (DLG1 and DLG4) seems to have a positive role in cell attachment to the basal layer, as impaired cell proliferation and early onset of differentiation did not result in their elimination from the basal layer.

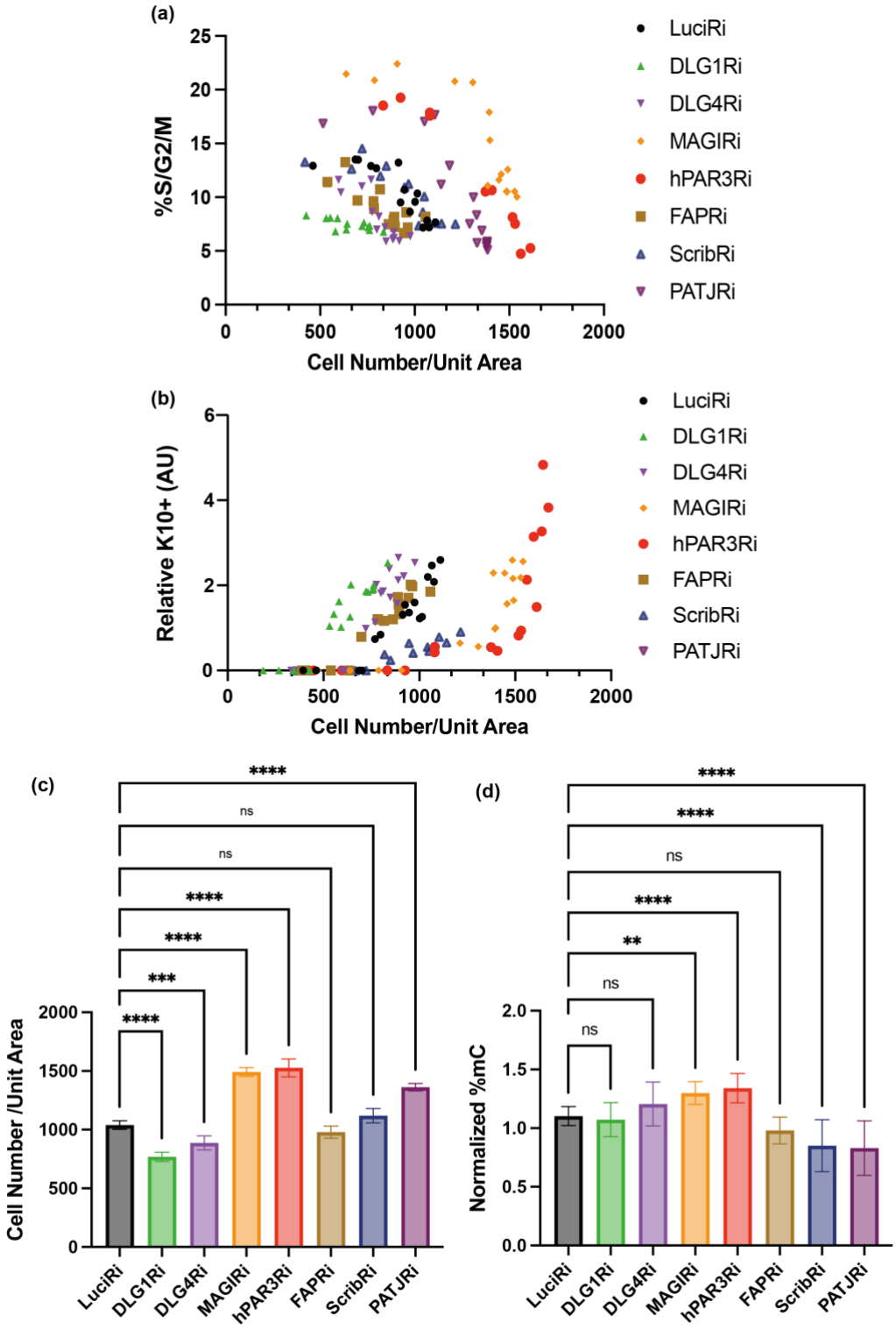


Figure 5.3.1: Systemic analysis of functions of PDZ-domain containing proteins in NIKS proliferation, differentiation, and competition. (a) Percentage of S/G2/M positive cells of FUCCI NIKS with RNA knockdown of Luciferase (LuciRi) and PDZ-domain containing proteins in response to cell number per unit area (420,000μm^2). (b) Relative K10 positivity of FUCCI

NIKS with RNA knockdown of Luciferase (LuciRi) and PDZ-domain containing proteins in response to cell number per unit area (420,000μm^2). (c) Saturation density (cell number per unit area (420,000μm^2)) of FUCCI NIKS with RNA knockdown of Luciferase (LuciRi) and PDZ-domain containing proteins with One-way ANOVA analysis in relation with FUCCI NIKS LuciRi (P-value style as GP). Data are shown as mean±SD. (d) The same number of EGFP LuciRi and mCherry LuciRi/PDZ-proteinRi were seeded at confluency. Percentage of mCherry cells (%mC) were calculated on Day1 and Day7. Then Day7 %mC data were normalised by Day1 %mC of different experiment settings and plotted. One-way ANOVA analyses were carried out in relation to normalised %mC in EGFP LuciRi with mCherry LuciRi (indicated as LuciRi). Data are shown as mean±SD and P-value style as GP.

5.3.2 The importance of MAGI and hPAR3 in regulating epithelial homeostasis is further elucidated with 3D organotypic raft culture

Next, we expanded our results to a 3D organotypic raft culture representing the tissue level. PDZ-containing proteins targeted by HPV16E6 for degradation are largely junctional cell compartments which are tightly regulated during the keratinocyte differentiation scheme; therefore, it is essential to consider their natural expression profile in the normal stratified epithelium.

Generally, organotypic raft culture results matched the *in vitro* tissue culture model. NIKS harbouring shRNA of different PDZ-domain-containing proteins were cultured for 14-day to form fully stratified epithelial tissue. Firstly, by examining the appearance of Hematoxylin and Eosin staining (H&E) with a light microscope (Figure 5.3.2 left panel), distinctive multilayered structures were formed with most of the NIKS shRNA cell lines. However, MAGI and hPAR3 knockdown NIKS failed to establish all epithelial layers within the 14-day culture. Then I applied immunostaining of early differentiation marker Keratin 10 (K10) to focus on the basal-to-suprabasal layers (Figure 5.3.2 right panel). For MAGI and hPAR3 knockdown NIKS, early differentiation is triggered with slight delay shown as few suprabasal layer cells have lower

K10 expression levels compared to other NIKS samples. Another notable feature is the structural unevenness at upper layers of multilayered structures, it is less obvious in the H&E sections, but cell appearance and K10 immunostaining indicated MAGI and hPAR3 knockdown in NIKS changed the appearance of cells.

The raft culture results strengthened my claim that MAGI and hPAR3 are the major PDZ-containing protein targets for HPV16E6 function, impairing normal keratinocyte differentiation. It is important to note that keratinocyte differentiation potential is not completely abolished by HPV16E6 expression or single protein knockdown, as there might be redundant pathways that facilitate the completeness of epithelial fate. My next aim would be to identify the molecular basis for basal keratinocyte commitment to differentiation. Thus far, our experimental systems rely on the markers for basal and differentiated epithelial layers to determine the function of HPV E6 and their target proteins.

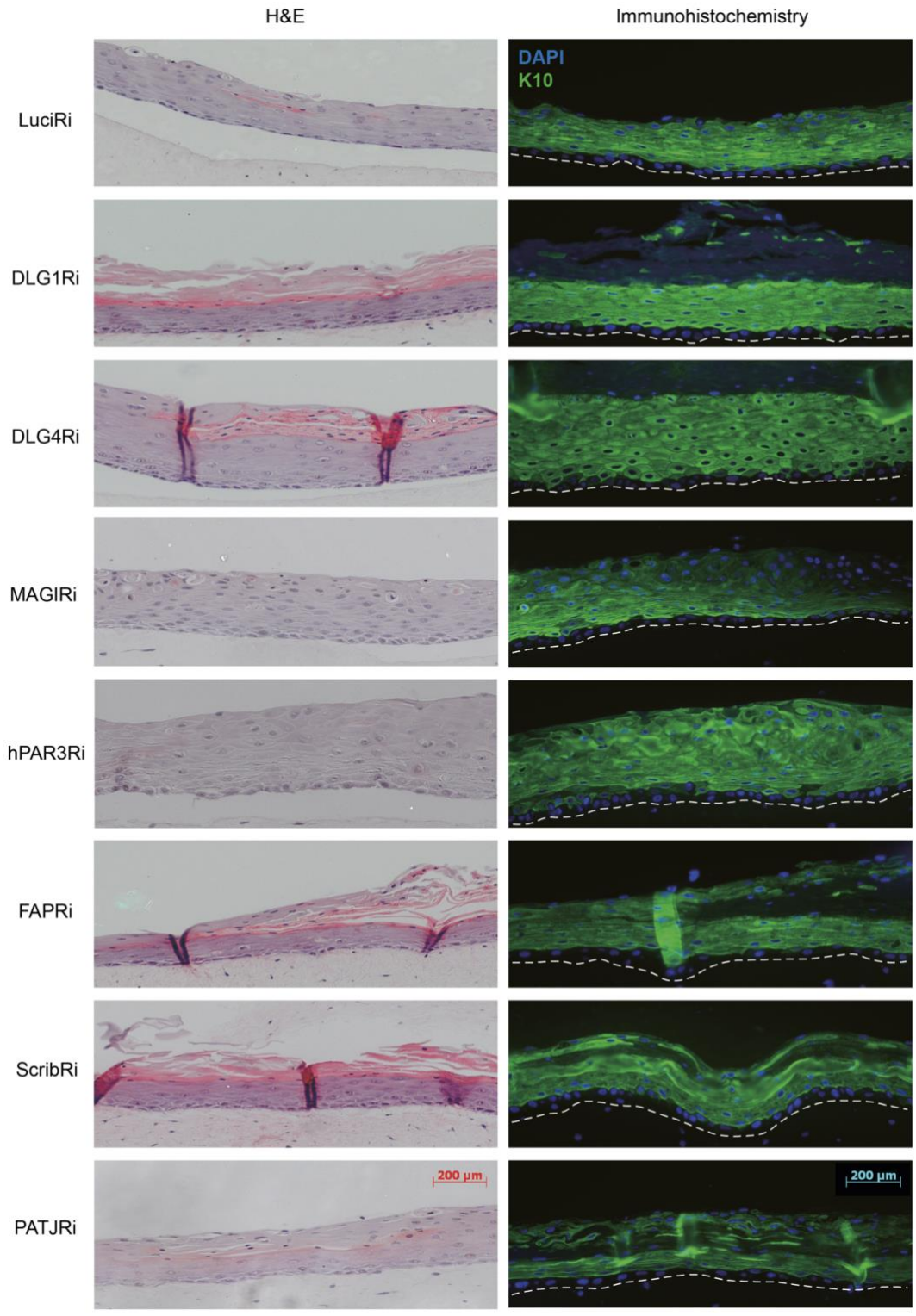


Figure 5.3.2: Organotypic raft culture of NIKS with shRNA knockdowns. (a) H&E sections; (b) Paraffin-embedded slices stained for K10 and DAPI (dotted line indicates the basal lamina attachment). All images were collected by 20x objectives.

5.4 Single-cell RNA analysis identifies subpopulations for keratinocytes Basal-Suprabasal transition

5.4.1 Quality control of scRNA sequencing dataset

Two individual populations, NIKS and NIKS 16E6, were seeded at confluency (8x of seeding density depending on the container) and expanded for 72hr in tissue culture with media change 24hr after seeding. Single-cell suspensions were librariedand sequenced with 10x Genomics service by CRUK Cambridge. Individual sequences were mapped to the reference human genome (Ghc38). Output matrixes were then analysed with Seurat (S4, R studio 4.2.0).

Single-cell RNA information on the number of features (genes mapped to reference human genome), number of counts (reads by Illumina sequencing), and percentage of mitochondrial and ribosomal RNA contents were plotted (Figure 5.4.1a). The majority of the cells have roughly 5,000 genes expressed detected by 50,000 reads per cell and contain 5% of mitochondrial and 28% of ribosomal-related RNA expression. The total expression profile by the number of genes expressed by NIKS 16E6 is slightly higher compared with NIKS (Figure 5.4.1a' nFeature_RNA'). Cells that have a low number of features/counts and high mitochondrial RNA contents are considered low-quality cell debris. Based on the violin plots, datasets were filtered by nFeatures_RNA > 2,500, nCount_RNA > 10,000 and percent.mt < 25 (Figure 5.4.1b).

Based on reference human genome RNA and out datasets, there are a total of 36,603 variable features for NIKS/NIKS 16E6 detected. The variance and expression level of individual features amongst the NIKS/NIKS 16E6 population is plotted (Figure 5.4.1c). The top 5,000 varied features were chosen for further analysis (in red dots), and the top 30 varied features in each dataset are indicated as text.

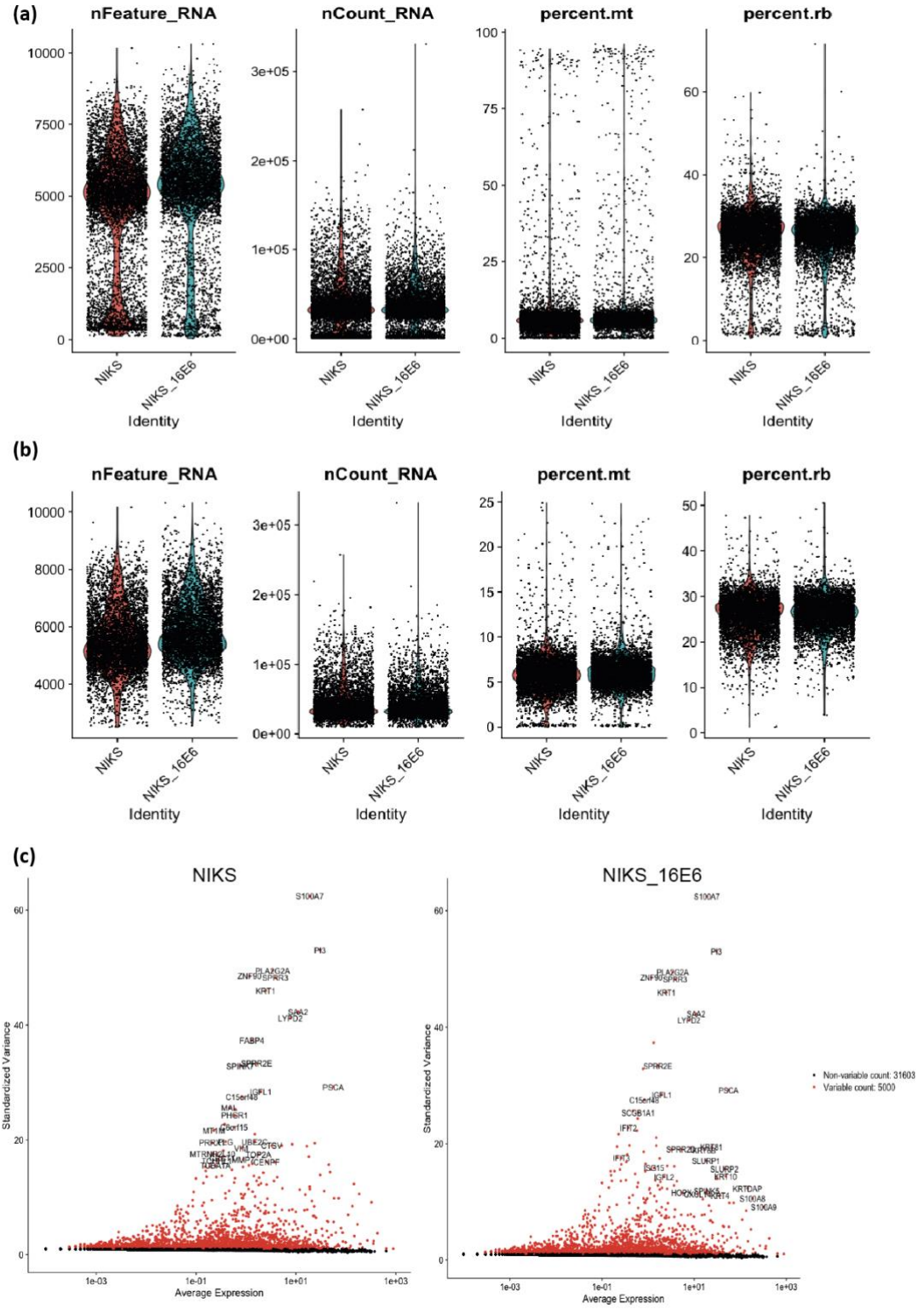


Figure 5.4.1: Quality control of scRNA sequencing dataset. (a) Violin plot of cellranger count raw matrixes of NIKS and NIKS 16E6 mapped to reference human genome (Ghc38). Each dot represents a single cell, showing the number of genes mapped to the human genome (nFeature_RNA), the number of reads by Illumina sequencing (nCount_RNA), percentage of

mitochondrial (percent.mt) and ribosomal (percent.rb) RNA contents. (b) Violin plot of filtered data matrixes by nFeatures_RNA > 2,500, nCount_RNA > 10,000 and percent.mt < 25. (c) Scatter plot of total variable features from NIKS and NIKS 16E6. The top 30 varied genes within the population are indicated in the text. A total of 5,000 top varied features from each dataset were chosen for further analysis.

5.4.2 Determining dimensions and resolution of clustering

After filtering out the low-quality cells, the initial separation of subpopulations (clustering) in the dataset is carried out as linear dimension reduction with Principal Component Analysis (PCA). This is an unbiased mathematical program to group cells into different clusters based on their RNA expression pattern. It is required to set the threshold of the difference of RNA expression to define the cell cluster, which also defines the resolution of the grouping method for proper analysis. The JackStraw package was developed for this purpose, to statistically estimate the variations of the standard deviation of RNA expression level within each cluster after running PCA. The result is plotted as standard deviations against number of clusters for PCA in the combined dataset of NIKS and NIKS 16E6 together (Figure 5.4.2). Basically, higher standard deviations suggest greater differences of RNA expression level within the clusters reflecting the small number of clusters and low resolution. When the standard deviation did not decrease further by increasing the number of clusters, it is considered as the saturation of clustering, suggesting that the number cluster is appropriate for further analysis. From the figure, a total of 12 clusters provides a good threshold for identifying subpopulations in NIKS and NIKS 16E6 datasets.

The PCA method provided reliable, unbiased clustering of subpopulations, but it fails to measure the relative distance between the individual clusters for the single cell RNA dataset (Appendix Figure 3). To better visualise the relationships of different clusters in the 2D graph, the non-linear reduction method Uniform Manifold Approximation and Projection (UMAP) was then applied based on PCA clusters. In the UMAP method, high-dimensional data is

compressed using topological mathematical algorithms to map local correlations. All the figures shown in later sections are based on the UMAP plot.

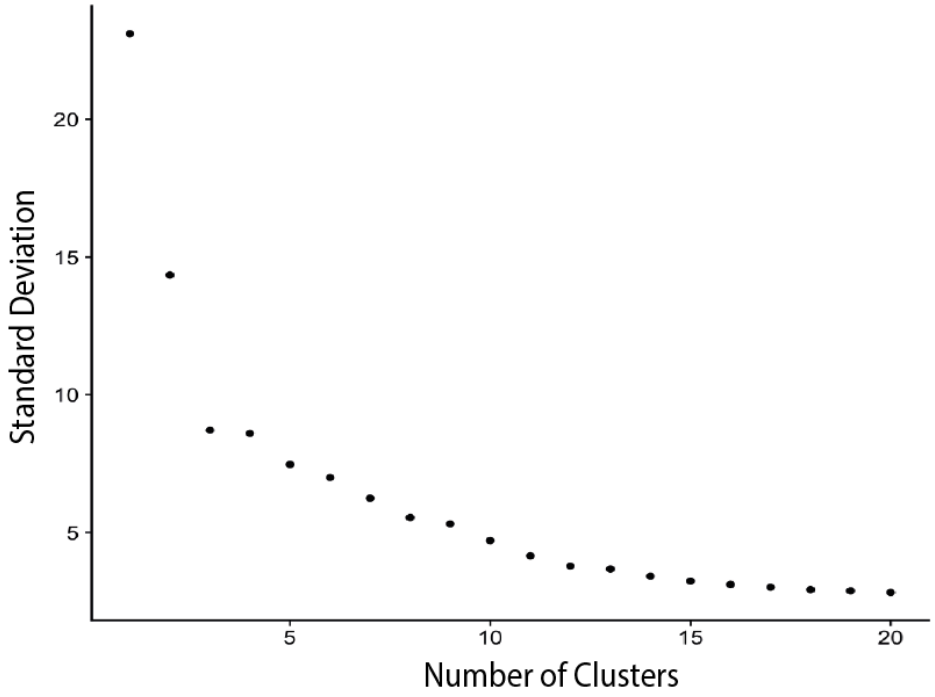


Figure 5.4.2: Analysis of variations in the number of clusters in the combined dataset of NIKS and NIKS 16E6. Mathematic methods of PCA and JackStraw were used to analyse the number of clusters and standard deviation between clusters.

5.4.3 Data structuring of NIKS and NIKS 16E6 (Merge vs Integration)

To compare the similarity and differences in NIKS and NIKS 16E6 populations, I combined the two datasets. Initially simply by merging the two datasets into one, a total of 12 clusters were identified in the combined dataset (Figure 5.4.3a). When grouped by the data origin (orig.ident) of either NIKS and NIKS 16E6, the two populations are nicely separated based on unbiased grouping methods (clustering). While by visualising the expression profile of established basal, early differentiation and cell cycle G2/M phase markers, namely integrin-α6 (ITGA6), involucrin (IVL) and cyclin B1 (CCNB1); the separated clusters of NIKS and NIKS 16E6 exhibited similar cell composition but spatially distant by the current "Merge" clustering method. Although the HPV16E6 expression alters the total NIKS expression profile, based on

our previous knowledge of the virus life cycle, persistently infected basal keratinocytes retain the proliferation and differentiation schemes as the normal keratinocytes (Doorbar 2005).

To address the issue in data presentation, by identifying the conserved markers (anchors) in each cluster of NIKS and NIKS 16E6, the cells were re-positioned on the UMAP plot. The method is known as Harmony integration, which regresses the total expression profile variation between the two datasets. As expected, the two populations have the same data structure (Figure 5.4.3b) in the basal-to-suprabasal transitioning stage, confirmed by conserved marker expressions. It is important to note that although the clusters from the two datasets are plotted overlapped, the fundamental difference in 16E6 expression separated them into distinct clusters.

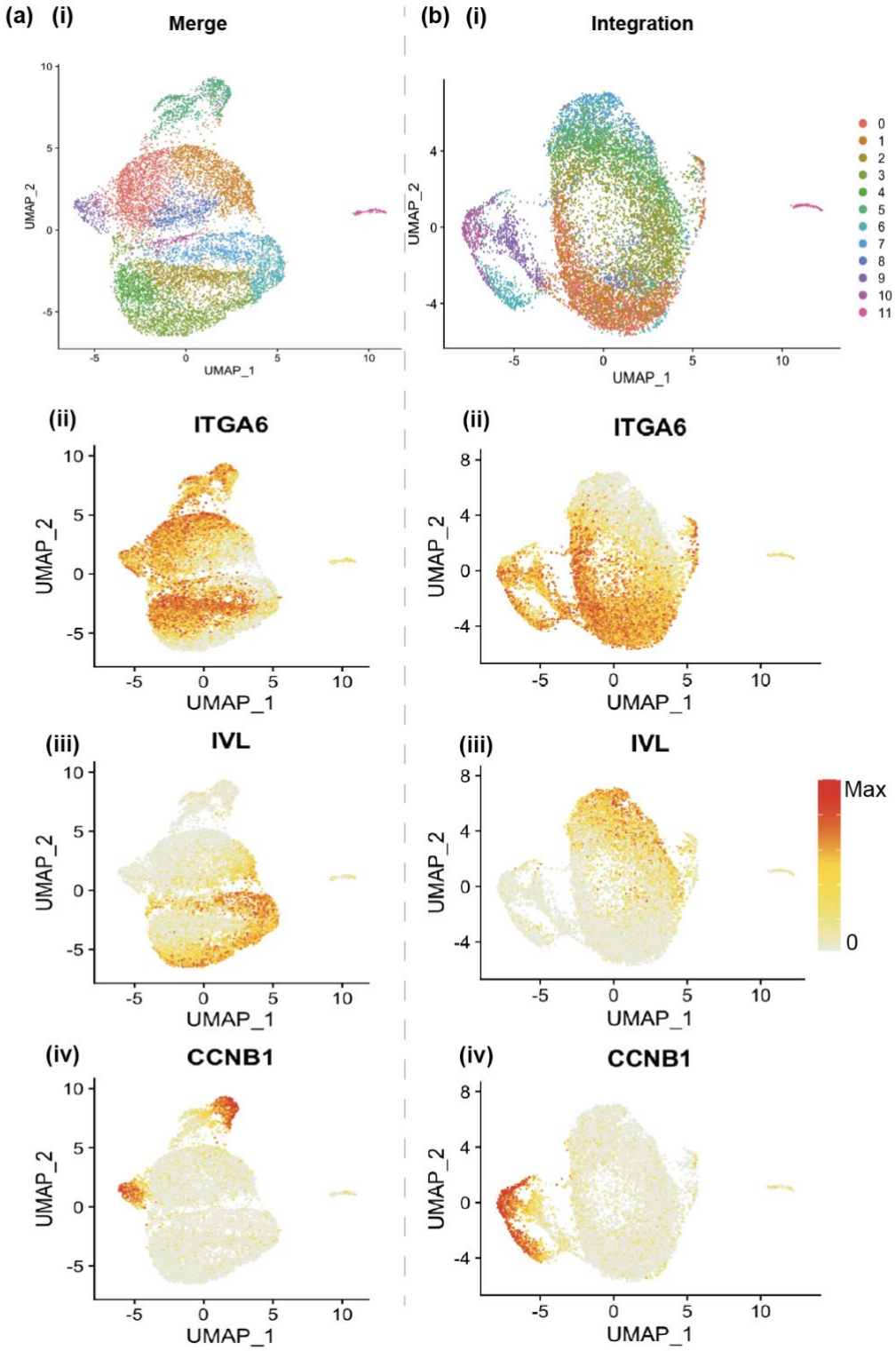


Figure 5.4.3: UMAP plots of the data structure of NIKS and NIKS 16E6 merge (a) and integration with Harmony (b), with the same dimensionality resolution as PCA. Data shown as dimension UMAP plot of 12 different clusters from the combined datasets of NIKS and NIKS 16E6 (i), and RNA expression profile of markers for basal keratinocyte (ITGA6) (ii), committed

to differentiation (IVL) (iii) and G2/M phase cell cycle (CCNB1) (iv). Expression level is shown as log2 fold change.

5.4.4 Categorising individual clusters with established epithelial markers

After pre-processing the dataset and adjusting the resolution of the integrated dataset, a total of 12 individual clusters were separated. NIKS and NIKS 16E6 populations exhibited high homology in terms of general cell type compositions. The integrated dataset was then split into two individual datasets, based on the sample identity of the original NIKS and 16E6 samples. A total of 7 clusters were generated for each NIKS and NIKS 16E6 population. By probing the clusters with the established cellular markers for cells in S/G2/M phase (CCNB1 and CDK1), basal layer keratinocytes (ITGA6 and COL17A1), early epidermal differentiation (IVL) and fibroblasts (VIM). Individual clusters were then renamed into three basal groups (Basal-1/2/3), three differentiating groups (Diff-1/2/3) and a small cluster of fibroblasts (Fibroblast) for *in vitro* culture of NIKS (Figure 5.4.4a).

By calculating the relative proportion of each group in NIKS and NIKS 16E6, the pie chart shows that the Basal groups generally increase in NIKS 16E6 population (Figure 5.4.4b). The proportion of the Basal-1 group which expresses the S/G2/M phase markers agrees with our previous results from Chapter 3, that when keratinocytes reach confluent and start committed to differentiation in tissue culture condition, roughly 8% of NIKS and 25% of NIKS 16E6 remains proliferative by FUCCI expression. It is interesting to note that Basal-3 is unique for NIKS and almost absent in the NIKS 16E6 population, suggesting a possibly undefined population in basal keratinocytes (further analysed in section 5.4.8). The Diff-3 cluster (suggested to be the same sub-population as K10+ from Chapter 3) is also largely reduced in the NIKS 16E6 population, which has the highest differentiation marker (IVL) expression level, this is suggested due to the 'delayed' commitment to differentiation by 16E6 expression, and the commitment is triggered at higher cell density. Detailed clustering and expression profiles of

basal and early differentiation markers (ITGA6 and IVL) for NIKS and NIKS 16E6 populations are shown as UMAP plots (Figure 5.4.4c).

Overall, the scRNA data is comparable with previous phenotypic observations and provides an increased resolution to identify the molecular determinants for basal cell fate decisions.

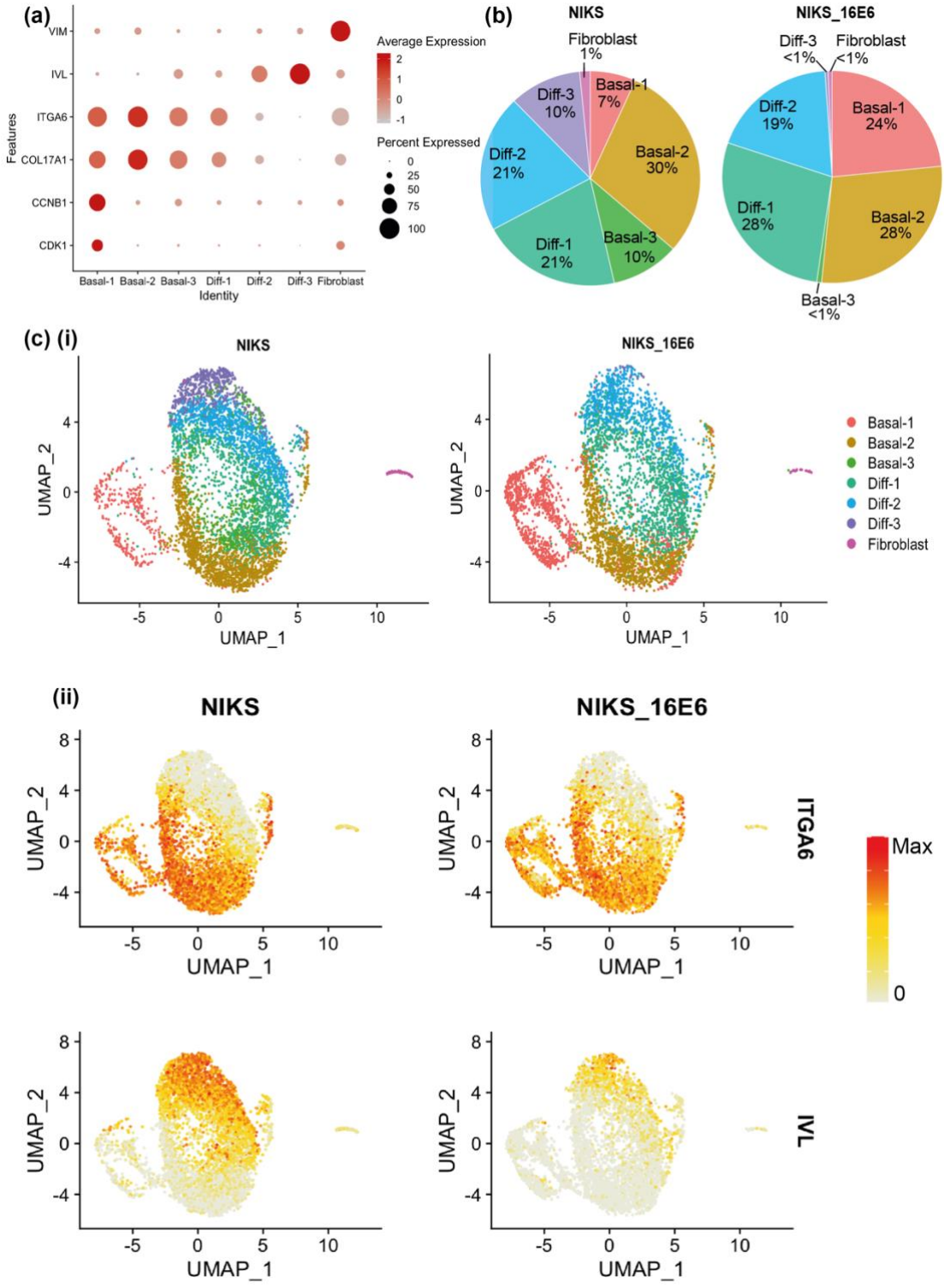


Figure 5.4.4: Categorising individual clusters of NIKS and NIKS 16E6. (a) Dot plot of conserved markers for cell cycle S/G2/M phase (PCNA, CDK1 and CCNA2), basal keratinocytes (ITGA6), commitment to differentiation (IVL) and Fibroblast (VIM). The colour scale represents the average expression level of genes (shown as log2 fold change), and dot sized indicates the

proportion of cells with positive gene expression within the cluster. (b) Pie chart of the percentage of the individual cluster in each single cell dataset of NIKS or NIKS 16E6. (c) (i) Individual UMAP cluster plot separated by the original identity of the population. (ii) Feature plot of the expression level of ITGA6 and IVL. Expression level is shown as log2 fold change.

5.4.5 Building pseudo-time trajectory based on conserved markers

After categorising the clusters into basal and differentiating groups, the UMAP plot provided some basic information about the relative distance between each group. Based on experiments and clinical observations, the majority of basal keratinocytes were in a quiescent state (G0/G1 phase) with constant integrin expressions for attachment to the basal lamina. The epithelium is constantly experiencing turnover from the loss of the cornified layer; therefore, the basal cells are required to determine their cell fate towards either division to recapitulate the basal layer or commit to differentiate to form the fully stratified epithelium.

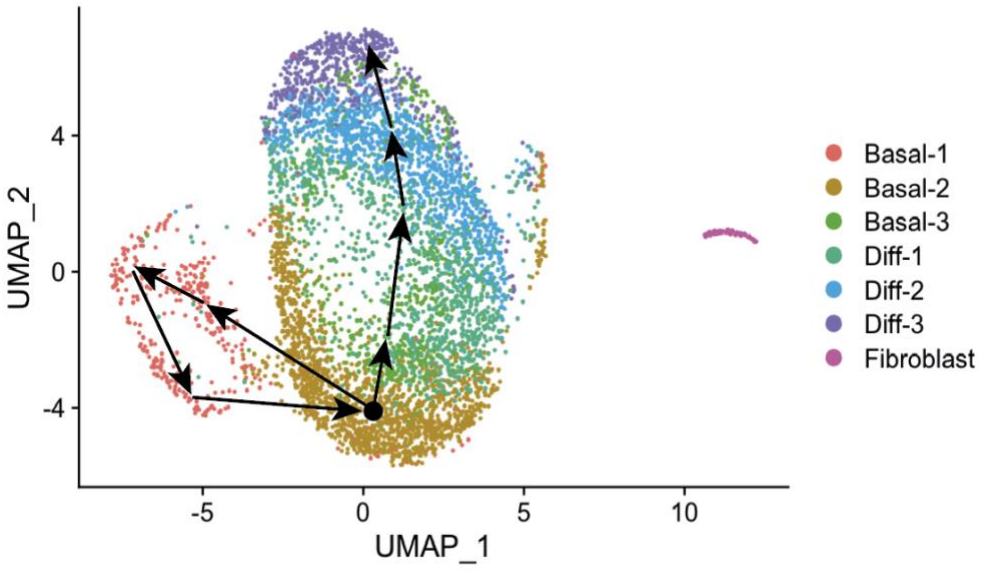


Figure 5.4.5: Pseudo-time trajectory analysis of post-confluent NIKS population. The black dot represents the starting point, and the arrow indicates the pseudo-time flow.

From the scRNA dataset, the majority of the basal group cells were grouped into the Basal-2 cluster, while the Basal-1 cluster highly expresses the cell cycle S/G2/M phase markers (Figure 5.4.4a). Then Basal-2 was set as the starting point for the pseudo-time trajectory analysis for NIKS (black dot in Figure 5.4.5). The quiescent basal keratinocytes could either enter the cell cycle for proliferation and division as the Basal-2 cells move towards the left side on the UMAP plot. The scRNA sequencing results nicely distinguished the cells in S, G2 and M phases, but it is not our interest for this analysis, so I grouped them into the same Basal-1 group. While the grouping did not separate the S/G2/M cells, the spatial UMAP plot shown the cell cycle status and cells after the M phase returned to the Basal-2 group.

The basal keratinocytes would also commit to differentiate and delaminate to form the fully stratified epithelium; therefore, the Basal-2 cells can develop towards differentiation as arrow indicated towards the Diff-1/2/3 groups (Figure 5.4.5). This trajectory is built based on the phenotypic observations and our current knowledge to the epithelium.

5.4.6 Identify molecular markers differentially expressed in individual clusters

After categorising individual clusters, differential expression analysis, using DESeq2 (Love *et al.* 2014), was carried out comparing each cluster against the rest of the single cell population. Total upregulated and downregulated genes of each cluster are supplied in the Appendix File 4 as NIKS.all.markers or NIKS16E6.all.markers.

After rearranging by cluster and log2 fold change, top50 upregulated genes from NIKS clusters were highlighted and plotted as a heatmap (both NIKS and NIKS 16E6 are plotted with the same gene list in Figure 5.4.6, the list is provided in the Appendix File 4 as NIKS.top50.markers).

Cluster tree is generated by Seurat's 'BuildClusterTree' algorithm. NIKS and NIKS 16E6 have the same development cluster tree. Fibroblasts are firstly separated from the keratinocytes population, and then NIKS is further divided into the basal and differentiating populations.

Basal-1 and Basal-2 are clustered together and share the same development distance as Basal-3. Diff-1 is more closely connected with the Basal groups compared with the Diff-2/3 suggesting it may be the early transitioning population from Basal to Suprabasal cells.

In the NIKS population, HPV16E6 expression generally reduces Diff-1 and Diff-2 top-regulated markers, which is consistent with the hypothesis that 16E6 cells are more resistant to differentiation.

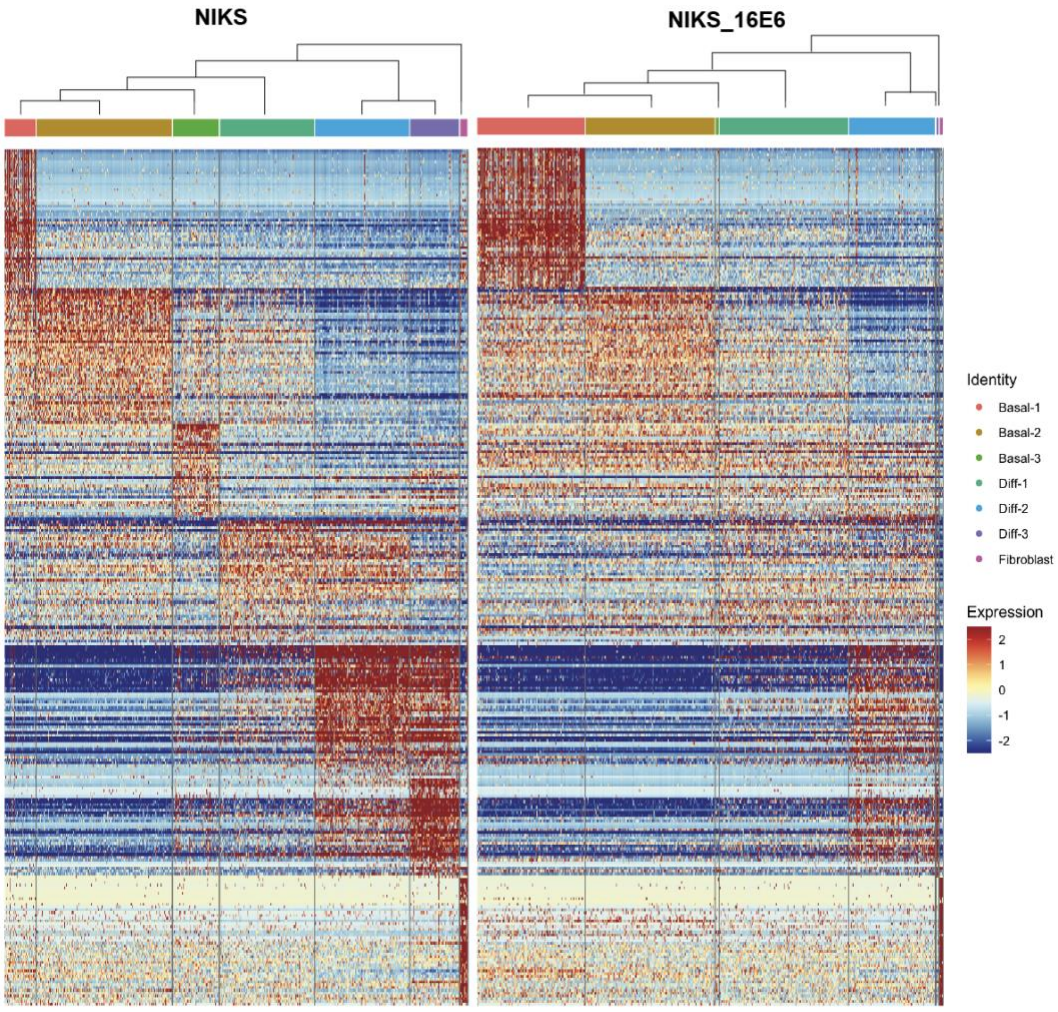


Figure 5.4.6: Heatmap of top50 upregulated genes in each cluster compared to the whole NIKS population. The dendrogram is generated by the 'BuildClusterTree' algorithm in Seurat. Expression level is shown as log2 fold change.

5.4.7 Ligand-Receptor (LR) expression predicts cell signalling model

Signalling crosstalk via soluble and membrane-bound factors is critical for informing diverse cellular decisions, including decisions to activate cell cycle or programmed cell death, undergo migration or differentiate along the lineage (Jamora & Fuchs 2002). Characterising cellular interactions in vivo and at high resolution remains challenging. As scRNA-seq approaches provide high resolution, bioinformatics has developed towards constructing models to predict cell signalling based on the ligand-receptor expression. Here I applied CellChat (Jin *et al.* 2021), an algorithm that incorporates signalling molecule interaction information from the KEGG Pathway database (Kanehisa *et al.* 2017) to identify intercellular communications from the NIKS scRNA-seq data.

5.4.7.1 HPV16E6 expression generally downregulates NIKS LR expression profile

Basal-1 clusters were excluded from the ligand-receptor (LR) prediction as it forms a separate trajectory path compared with the differentiating groups, and Fibroblast is not considered for LR prediction as our *in vitro* model system could not represent the Fibroblast-Keratinocyte interaction network.

Seurat analysed NIKS, or NIKS 16E6 scRNA expression profiles were aligned against KEGG-established LR pairs. The total number of predicted interactions was calculated (Figure 5.4.7.1.1a), with NIKS 16E6 having 32% lower interactions compared with NIKS (851 and 1245 interactions, respectively).

By dissecting LR interactions from each cluster, the heatmap generally shows there is an increased number of cell-cell signalling pathways in the Basal groups (Figure 5.4.7.1.1b). Moreover, the Basal groups are mainly signal senders, and the Diff groups are signalling

receivers. The interaction probability for each sender-receiver pair is calculated by the percentage of cells expressing the LR pair within the target clusters. This interaction pattern is similar between NIKS and NIKS 16E6 populations. The Basal-3 and Diff-3 groups in NIKS 16E6 population have only 1% abundancy; the small sample size makes it less convincing to draw meaningful conclusions on cell-cell signalling predictions.

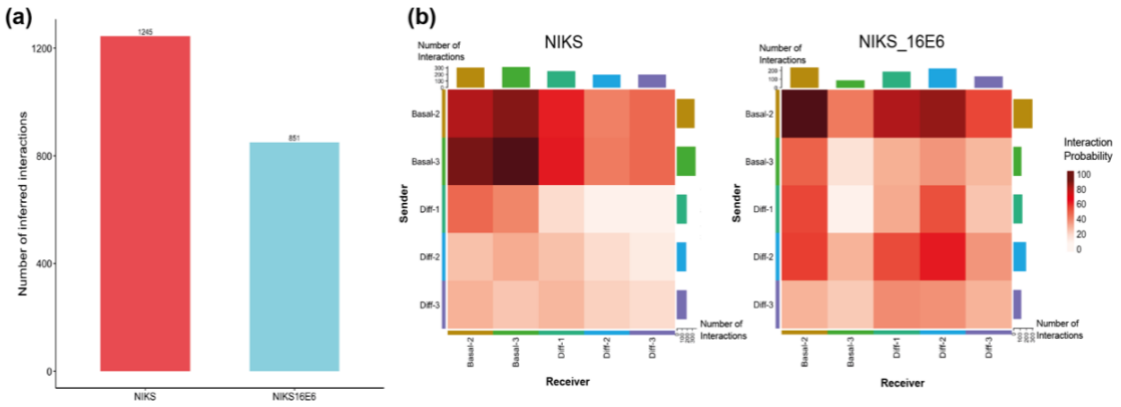


Figure 5.4.7.1.1: HPV16E6 expression downregulates NIKS LR expression profile. (a) Histogram of the number of inferred interactions of NIKS and NIKS 16E6 based on KEGG-pathways. (b) Heatmap of cell-cell signalling interactions of individual clusters of NIKS and NIKS 16E6; column height indicates the number of inferred interactions, and colour scale indicates interaction probability calculated based on the percentage of cells in the cluster expressing ligand or receptor.

The LR pairs were then evaluated as ligand expression for the outgoing signal and receptor expression for the incoming signal of each cluster. Heatmaps of all significantly upregulated signal interactions (p-value threshold <0.01) in individual clusters of either NIKS or NIKS 16E6 population are plotted (Figure 5.4.7.1.2). The colour scale represents the interaction strength which is calculated by the number of cells in the cluster and the expression level of each LR pair, with all the data sets normalised to a range of 0-1. The top coloured column plots the total signalling strength of a cell group by summarising all signalling pathways displayed in the

heatmap, and the right grey column plots the total signalling strength of a signalling pathway by summarising all cell groups displayed in the heatmap.

Despite the large amount of information provided, I chose to focus on several pathways relevant to the goal of this project: the basal-suprabasal cell fate decision. There is a transition of cell attachment pattern from Basal groups to Diff groups changing from Laminin and Collagen to Desmosomes and Cadherins (CDH). Notch and Wnt pathways have been studied as different LR pairs direct distinct cell fate during development, and their role in epithelial homeostasis has been proposed (Clevers & Nusse 2012; Kopan & Ilagan 2009b; Siebel & Lendahl 2017). The following sections dissect these two signalling pathways for modelling future experiments.

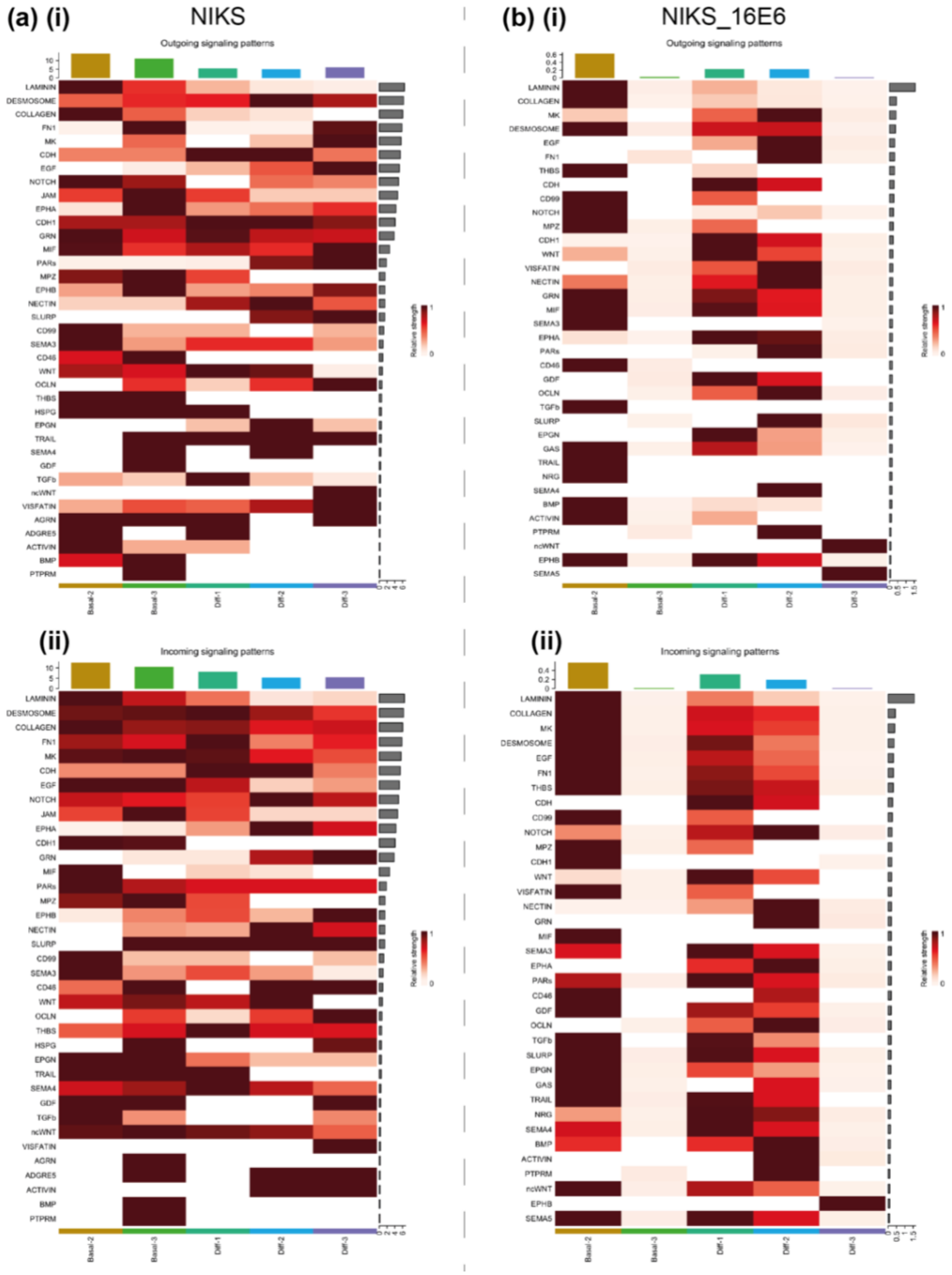


Figure 5.4.7.1.2: Heatmap of outgoing (i) and incoming (ii) signals in individual clusters of NIKS (a) and NIKS 16E6 (b). The colour scale represents the interaction strength which is calculated by the number of cells in the cluster and the expression level of each LR pair, with all the data sets normalised to a range of 0-1. The top coloured column plots the total signalling strength of a cell group by summarising all signalling pathways displayed in the

heatmap, and the right grey column plots the total signalling strength of a signalling pathway by summarising all cell groups displayed in the heatmap.

5.4.7.2 NIKS 16E6 attenuates Notch-signalling by downregulating Notch1 transcription

Notch pathway is highly conserved for versatile signalling transductions across species and biological functions. The one-to-one interaction between each LR pair in Notch signalling emphasises the importance of regulation at the expression level (Bray 2016). The communication probability (as colour scale) is calculated with the number of cells in the cluster and the expression level of each LR pair, with the signal sending cells expressing ligand and signal receiving cells expressing receptors (Figure 5.4.7.2a). The figure is plotted as each LR pair in relation to each signal sending/receiving cluster pair, and NIKS and NIKS 16E6 are plotted next to each other distinguishable by coloured text. Only interactions with a p-value <0.01 are plotted, and all other groups were excluded from plotting. Together with the violin plot of the Notch pathway genes (expression detected by scRNA sequencing), in which the y-axis represents the relative expression level, and the width of the violin represents the relative proportion of cells at a certain expression level (Figure 5.4.7.2b), there are some predictions in 16E6 functions in Notch pathway can be drawn.

There is a large reduction in interaction probability of NIKS 16E6 clusters compared with the NIKS population, and this is highly likely caused by the downregulation of Notch-receptor expression levels (especially Notch1). The Jagged-ligand expression profile changed when NIKS committed to differentiation, as JAG-2 is more dominant in Basal groups while JAG-1 expression increases when NIKS differentiates. The JAG-1 and JAG-2 transcription regulation is not targeted by HPV16E6. Also, the expression pattern of Notch receptors changed when NIKS committed to differentiation, with a change from Notch1 to Notch2/3 when the cell cluster changed from Basal to Diff. It is special to note that the Delta-like ligands (DLL1 and DLL3) are only expressed by NIKS 16E6 clusters, with the regulation of DLL1 transcription restricted to the Basal groups.

Heatmap provides clear information on the relative relationship of signal transduction in the Notch pathway, as the Basal groups behave as signal senders and the Diff groups as signal receivers (Figure 5.4.7.2c).

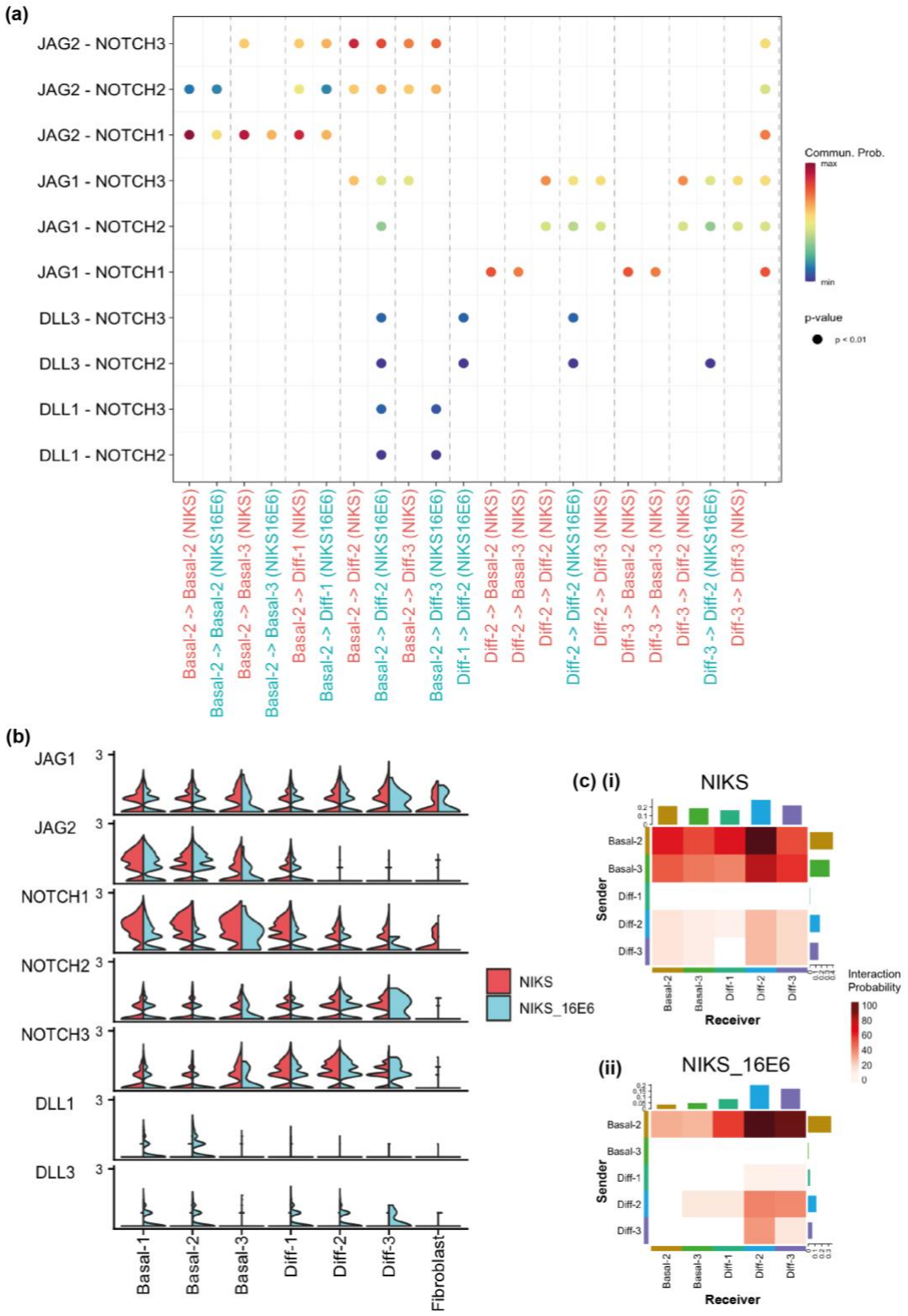


Figure 5.4.7.2: NIKS 16E6 attenuates Notch-signalling by downregulating Notch1 transcription. (a) Bubble plot of individual LR pair in Notch pathway, categorised by different signal sending and receiving cluster pair and NIKS or NIKS 16E6 population. The colour scale is calculated with the number of cells in the cluster and the expression level of each LR pair, with

the signal sending cells expressing ligand and signal receiving cells expressing receptors. Only data with a p-value<0.01 is shown. (b) Violin plot of Notch pathway LR gene expressions. The y-axis represents the relative expression level, and the width of the violin represents the relative proportion of cells at a certain expression level. NIKS and NIKS 16E6 are plotted next to each other distinguishable by colour. (c) Heatmap of predicted interaction probability, shown as a colour scale, between each cluster pair of NIKS (i) and NIKS 16E6 (ii). The column plot on the top and right represents the sum of interaction probability or each column or row, respectively.

5.4.7.3 HPV16E6 alters the Wnt signalling pathway expression profile

The Wnt-signalling pathway can be predicted with the same method. Wnt is a secreted ligand that binds to heterodimer receptors. The results need to be interpreted more carefully in order to obtain feasible hypotheses to guide future research. The figures are plotted based on the same criteria as the Notch pathway analysis.

Similar to the Notch pathway, NIKS 16E6 generally has fewer interactions predicted for the Wnt pathway, and especially the Basal groups play a minimal role in both signals sending and receiving patterns (Figure 5.4.7.3). When NIKS cells are committed to differentiation, there is a negative relationship in Wnt ligands (Wnt6 and Wnt4) transcription activities. The Wnt7B ligand is generally expressed by all NIKS clusters with no significant modulation by 16E6 expression, but the corresponding receptors (FZD6, FZD3, LRP5 and LRP6) pairs mainly expressed at the Diff groups suggest the differentiated cells are the signal receivers. The receptor protein gene expression generally does not change with the clusters, except the LRP5 mRNA level decreases as NIKS differentiates. The Diff groups are more likely to receive the Wnt signals than the Basal groups. This may strengthen our understanding as HPV only utilise E6 to modulate the threshold of basal cell homeostasis but does not transform the persistently infected cells.

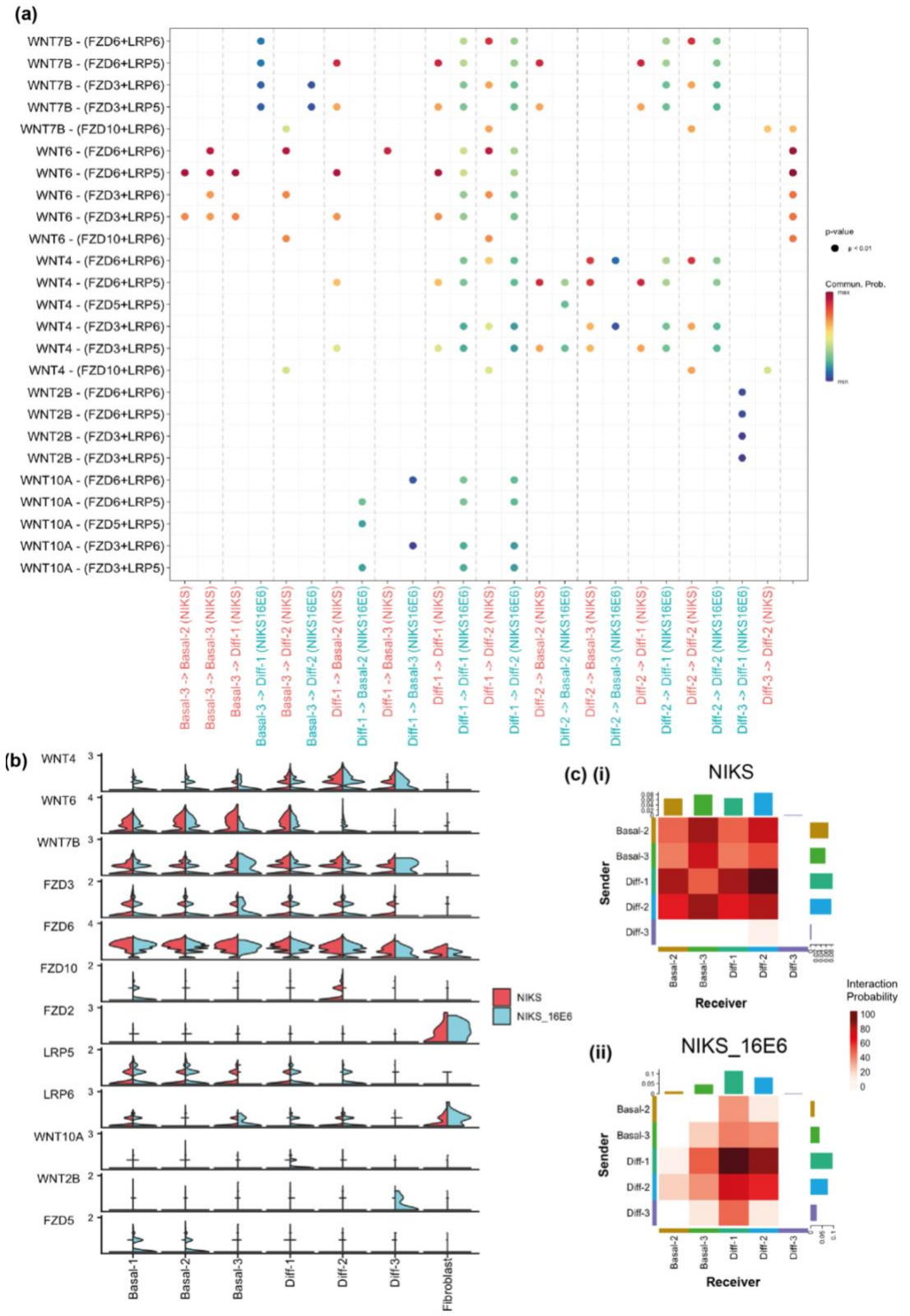


Figure 5.4.7.3: HPV16E6 alters the Wnt signalling pathway expression profile. (a) Bubble plot of individual LR pair in Wnt-signalling pathway, categorised by different signal sending and receiving cluster pair and NIKS or NIKS 16E6 population. The colour scale is calculated with the number of cells in the cluster and the expression level of each LR pair, with the signal sending

cells expressing ligand and signal receiving cells expressing receptors. Only data with a p-value<0.01 is shown. (b) Violin plot of Wnt pathway LR gene expressions. The y-axis represents the relative expression level, and the width of the violin represents the relative proportion of cells at a certain expression level. NIKS and NIKS 16E6 are plotted next to each other distinguishable by colour. (c) Heatmap of predicted interaction probability, shown as a colour scale, between each cluster pair of NIKS (i) and NIKS 16E6 (ii). The column plot on the top and right represents the sum of interaction probability or each column or row, respectively.

5.4.8 Identification of a unique cluster in NIKS which is absent in the NIKS 16E6 population

As previously mentioned, Basal-3 is a unique cluster expressed by NIKS only. By extracting the top upregulated genes from the Basal-3 cluster, ordered by log2 fold change from the most upregulated against the rest of the population, a heatmap is plotted suggesting this population is not only special against 16E6 expression but also against the NIKS population (Figure 5.4.8a). The top regulated gene is metalloproteinase-7 (MMP7), with log2 fold change of 2.38 (p-value = 5.92E-137).

Not all cells in the Basal-3 cluster express the same level of MMP7 mRNA; the Basal-3 can be further divided into two subclusters (Basal-3a and 3b) based on their expression profile (Figure 5.4.8b). When plotting the expression level of MMP7 on Basal-3 sub-populations (Figure 5.4.8c-i), Basal-3a is low-MMP7 expressing while Basal-3b has a higher MMP7 expression level. Then ITGA6 and IVL expression suggest that MMP7 expression level correlates with NIKS commitment to differentiation (Figure 5.4.8c), while the MMP7 mRNA expression is not sustained to the differentiated clusters.

Preliminary immunostaining results on NIKS raft culture sections confirmed our scRNA results (Figure 5.4.8d) that MMP7 protein expression is mainly restricted to the basal layer and NIKS 16E6 downregulates MMP7 expression at the transcription level.

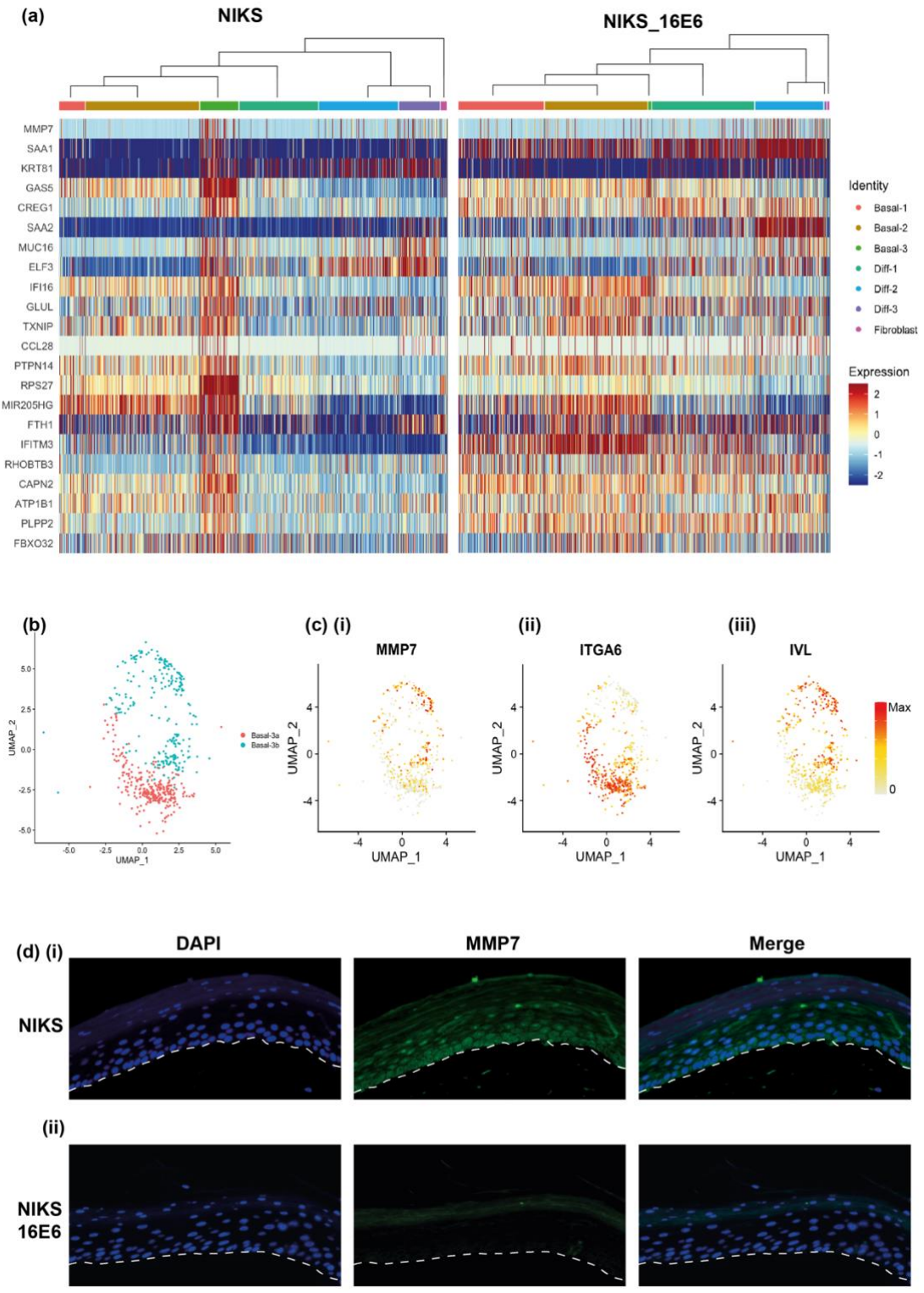


Figure 5.4.8: Basal-3 is a unique cluster expressed by NIKS but not NIKS 16E6. (a) Heatmap of top22 upregulated genes in Basal-3 compared to the whole NIKS population. Both NIKS and NIKS 16E6 are plotted against the same gene list. (b) UMAP plot of subcluster of Basal-3 (Basal-3a and Basal-3b). (c) Feature plot of the expression level of MMP7 (i), ITGA6 (ii) and IVL (iii).

Expression level is shown as log2 fold change. (d) H&E (i) and immunohistochemistry of MMP7 (ii) on NIKS and NIKS 16E6 raft culture sections.

5.5 Discussion

5.5.1 HPV16E6 mutants and shRNA analysis identify p53 and PDZ-domain containing proteins as having a complementary role in basal layer persistence

The primary objective of this chapter is to investigate mutants of HPV16E6 based on existing functions, and to determine whether p53 is the primary cell cycle regulator in basal keratinocytes, which also affects NIKS LXSN 16E6 colonization of the bottom layer in competition assays. The PDZ-containing proteins targeted by 16E6 are less important for basal cell cycle entry. The restoration of PDZ-containing proteins expression level in the 16E6dPDZ mutant cells still exhibited a similar level of proliferating cells as the 16E6 cells. When co-cultured with normal NIKS, the dPDZ mutant failed to colonise the bottom layer, suggesting that the populational delamination and differentiation balanced with the high proliferation rate. The mutant analysis dissects the process of basal-suprabasal transition, of which proliferation, contact inhibition, delamination and commitment to differentiation are independent stages for a basal cell to acquire terminal fate. It also strengthens my understanding of the evolutionary similarity and differences in different types of HPV E6 protein (from Chapter 3). The differences and similarities in the results on the NIKS model may reflect the regulation mechanisms of natural tropism epithelial sites in keratinocyte behaviour.

Tumour suppressor p53 typically functions as a specific DNA binding protein regulating transcription of the genes involved in cell cycle arrest, apoptosis and senescence (Sherr 2004). Induction of p53 by cellular stress triggers multiple cellular programs, including DNA damage repair, cell cycle arrest and apoptosis. In epithelial tissue, expression of p53 is mainly restricted to the basal layer (Botchkarev & Flores 2014), and it has been shown that sustained p53 induction but not pulse induction can upregulate the expression of genes associated with cellular terminal fate (Purvis *et al.* 2012).

P53 functions through several mechanisms. It has been best characterised as a transcription factor that activates target genes, including the cyclin-dependent kinase (CDK) inhibitor p21 (reviewed in Vousden and Prives, 2009). Based on the results, restoration of p53 levels in NIKS 16E6 cells largely impairs virus protein function in driving cell cycle entry. Moreover, unpublished data from an MPhil student from our lab suggest that treatment of small molecule Lopinavir (screened antiviral candidates from Chapter 4) in NIKS 16E6 increases the p21 expression level without any obvious elevation of p53 level. Together this suggests that p21 is the actual effector of p53 regulating NIKS cell cycle progression. It is not necessary to target viral protein or its direct cellular partners, but downstream cell effector pathways can also achieve successful clearance of infected viral reservoirs.

The long co-evolution history of HPV with *Homo Sapiens* provides us with a great opportunity to utilise the viral proteins as tools for understanding normal homeostasis pathways. By investigating the molecular targets for 16E6 degradation, the importance of PDZ-domain containing proteins during basal keratinocyte commitment to differentiation is emphasised with new insights. Using the exogenous expression of shRNA to knockdown 16E6 targeted PDZ-containing proteins in NIKS, I further narrow down the most important basal epithelial regulators targeted by HPV16E6 are MAGI and hPAR3. The results are consistent with previous research in which MAGI appears to be the main degradation target for the 16E6 PDZ-binding motif (Araujo-Arcos *et al.* 2022; Kranjec & Banks 2011).

Reduced extracellular matrix adhesion (hemidesmosomes/integrin α6β4) stimulates cultured keratinocyte terminal differentiation (Watt 2002). Also, as the cells differentiate towards the apical epithelium, paracellular interactions change from desmosomes to adherent junctions and, finally, tight junctions. In response, membranous protein composition is tightly regulated (reviewed in Matter et al., 2005). Both differentiation and cancer metastasis requires cell junction protein rearrangements to delaminate and migrate, the high-risk E6 function in degrading PDZ-domain-containing proteins has been studied for carcinogenesis, and the type-specific PDZ-interaction profiles could potentially predict their oncogenic potential (Thomas

et al. 2016b). While in this project, evaluating 16E6 targeted PDZ-domain containing proteins for normal epithelial homeostasis suggest divergent functions for viral targeting different proteins during its life cycle.

The shRNA results broadly separate the PDZ-domain-containing proteins into two groups, which MAGI1 and hPAR3 knockdown promote basal cell persistence in opposite to DLG/Scrib/PATJ. These findings agree with the previously reported polarity protein complex conformation, in which MAGI and PAR3 form a complex for cell adhesion (Barmchi *et al.* 2016) while DLG colocalises with Scrib at the basolateral sites (Bilder & Perrimon 2000; Dow *et al.* 2003).

Both MAGI1 and hPAR3 knockdowns result in higher NIKS confluence density and delayed differentiation commitment. Also, similar to the NIKS 16E6, shRNA MAGI1/hPRA3 NIKS preferably colonises the bottom layer in the competition assay. Silencing of MAGI1 increases the protein level of β-catenin, cyclin D1 and phospho-Akt while reducing the expression of PTEN, proposing that the Wnt/β-catenin and PTEN/Akt signalling pathways are the effectors of MAGI1-modulated cell proliferation (Lu *et al.* 2019). MAGI1 can bind with N-cadherin-β-catenin complexes as well as the Notch pathway ligand DLL1 (Mizuhara *et al.* 2005). The observations suggest both Wnt and Notch pathways are potentially important regulators of cell fate decisions and will be discussed further with scRNA seq analysis.

Knockdown of Scrib only impairs MDCK cells from being extruded when cultured together with normal cells (Norman *et al.* 2012). Consistent with my result, shRNA Scrib NIKS cells did not show phenotypic changes in terms of cell density, %S/G2/M or relative K10+ when cultured alone, but they failed to compete with the pseudo-KD NIKS to stay at the bottom layer in the mixed culture. In addition, other oncogenic viral proteins, such as Tax protein of human T-cell leukaemia virus-1, also binds to both Scrib and DLG1, inducing the miss-localisation of Scrib in infected cells and disrupting the growth inhibitory effects of DLG1 (Arpin-André & Mesnard 2007). Taken together, targeting MAGI/PAR3 complex could modulate infected cell homeostasis, possibly by affecting Notch/Wnt and downstream pathways. While inevitably,

Scrib/DLG degradation by HPV16E6 facilitates viral oncogenesis, its role in the normal HPV viral life cycle is yet to be discovered.

5.5.2 scRNA sequencing results models for basal-suprabasal transition and cell crosstalk pathways

The results on the heterogeneity of keratinocytes *in vitro* provide the basis for potential future investigations. Transcriptional phenotyping is a proven method to capture a snapshot of cell state under different conditions and to compare alterations in cell state following exposure to environmental stimuli (Kalluri 2019). The pilot study of single-cell RNA sequencing with post-confluent NIKS and NIKS 16E6 could serve as a blueprint for future analysis of identifying and categorising transitioning basal keratinocytes. The high-resolution analysis allowed me to build a detailed trajectory path for basal cell fate towards proliferation or differentiation and identify potential markers for the transitioning population.

From the scRNA seq results, the Basal-3 cluster, which is positioned between the quiescent Basal-2 and early Differentiated clusters, is unique for NIKS and absent for NIKS 16E6. The most identifiable marker for Basal-3 is the overexpression of the matrix metalloproteinases 7 (MMP7). The MMP family was discovered for their role in amphibian metamorphosis (Brinckerhoff & Matrisian 2002), and their function during tissue re-modelling is being explored (Page-McCaw *et al.* 2007). MMP7, for our interest, functions during the epithelial wound healing process by cleaving the extracellular matrix (ECM) and ECM-associated proteins to facilitate cell migration (reviewed in Rohani and Parks, 2015). Within the Basal-3 cluster, MMP7 expression is negatively correlated with integrin-α6 (ITGA6) and positively correlated with early differentiation commitment marker involucrin (IVL). Differentiation of basal keratinocytes requires delamination from the ECM-integrin adhesion matrix. The MMP7 protein is hypothesised to be critical for this delamination process. For the infected cells, viral E6 protein may directly or indirectly inhibit MMP7 transcription from refraining from differentiation. A knockdown experiment of MMP7 in NIKS is ongoing to investigate the

importance of NIKS delamination and commitment to differentiation and also try to identify cellular sensors/effectors of the MMP7 pathway for normal basal epithelial homeostasis.

Under real infection cases, the persistently infected cells are in contact with the healthy uninfected cells; therefore, cell-cell signalling plays an important role in cell fate decisions. Two of the essential pathways proposed for basal epithelial differentiations are Notch and Wnt pathways. Here with the scRNA seq data of the corresponding ligand-receptor expression level, some prediction models can be drawn to provide hints for future experiments.

The Notch pathway has pleiotropic functions in adult tissue homeostasis, determined by ligand-receptor pairs. Notch receptors (Notch1/2/3/4), ligands Delta-like ligands (DLL1/3/4), and Jagged-like ligands (JAG1/2) are all transmembrane proteins that functionally separate signal-sender and receiver cells. Notch1 is predominantly expressed in the basal cells of normal squamous epithelium (Sakamoto *et al.* 2012), and more specifically, its transcription is downregulated by 16E6 in p53-degradation dependent manner (Kranjec *et al.* 2017). Similar ligands have distinct biological roles in Notch-pathway activation, such as DLL1/DLL4 (Nandagopal *et al.* 2018) or JAG1/JAG2 (Choi *et al.* 2009). From the scRNA seq results, the JAG ligands expressions are negatively regulated with JAG2 predominantly expressed in the Basal groups, and JAG1 level increased as cells committed to differentiation. There is no significant regulation of JAG1/2 ligands transcripts level by 16E6 expression, but the decreased interaction probability is caused by inhibition of Notch1 expression. Viral E6 also increases the transcription of DLL1, which sends the differentiation signal to the adjacent cells (Khelil *et al.* 2021). This may suggest viral infected cells not only avoid itself to receive the Notch signals but also instruct the healthy cells to commit to differentiation. As mentioned before, Delta-like ligands can interact with PDZ-domain-containing proteins (e.g., MAGI1) (Mizuhara *et al.* 2005). The PDZ-binding sequences are dispensable for ligand activation and inhibition of Notch signalling. They are required for ligands to affect cell adhesion and migration (Glittenberg *et al.* 2006; Mizuhara *et al.* 2005; Six *et al.* 2004; Wright *et al.* 2004).

A close ally to Notch is the Wnt signalling pathway. In fact, Wnt plays a decisive role at multiple steps during the development of the skin (reviewed in Widelitz, 2008). Wnt is a large family of secreted signalling molecules with homology to the fly Wingless protein. The canonical Wnt pathway consists of secreted Wnt ligands binding with the heterodimer receptor of the Frizzled (Fz) and the lipoprotein receptor-related proteins (LRP), which leads to the accumulation of β-catenin and translocation to the nucleus to activate the downstream pathways as a transcription factor (Barker 2008). It is less clear to identify the signal sender/receivers for the Wnt-pathway signalling. However, the Wnt signalling pathway is more restricted to the basal layers compared with Notch pathways, as in the Diff-3 groups, there is barely any expression of Wnt receptors. From the scRNA seq results, when NIKS differentiates, there is a switch from Wnt6 to Wn4 signalling. Other research suggested that Wnt6 overexpression leads to increased cell proliferation (Bonnet *et al.* 2021), and loss of Wnt6 contributes to differentiation (Beaton *et al.* 2016). All these supports viral E6 expressing cells have increased basal property, as the Wnt6 expression of NIKS 16E6 is extended to the Diff-2 cluster. Wnt4 is generally considered for activating the non-canonical pathway, which acts independently to β-catenin. Many individual reports have suggested Wnt4 functions in embryonic development, ageing and malignancies, but it remains an open question of what regulative mechanisms of Wnt4 in different cells are (reviewed in Zhang et al., 2021). Wnt is a very complex pathway, and the functions of ligand-receptor pairs are cell type-specific, even within the skin appendages (Widelitz 2008).

Although scRNA seq provides convincing molecular insights, there are also some technical limitations that need to be considered while translating the bioinformatic analysis to experiments. We used the 10X Genomics library which probes are designed from 3' of the RNA to avoid picking up the prokaryotic RNA transcripts and only have 100-200 base pairs of the PCR product generated. While it gives low noise gene expression profiling, the data do not contain any transcript-splicing information. Particularly for our purpose of detecting viral transcripts expression, even driven by a constant retroviral expression vector LXSN, not all infected cells can be detected. The detection site is after the poly-A signal, so we failed to extract any information on the relative abundance of E6 and its splicing forms. However, on

the good side, even with the low expression level of viral E6 protein, the infected cell expression profile is distinct from the uninfected cells (Figure 5.4.3a).

As scRNA sequencing is being more widely adopted, the issue of whether the 'zero' expression is actual biological or technical has become realised and debated (Jiang *et al.* 2022). Biological zero is defined as the true absence of a gene's transcripts in a cell. The technical zero reflects the loss of information about truly expressed genes due to the inefficiencies of the technologies for sample preparation and sequencing. One cause of technical zeros is the imperfect mRNA capture efficiency in the reverse transcription (RT) step from mRNA to cDNA. The efficiency has a considerable variation across protocols and may be as low as 20% (Schwaber *et al.* 2019).

The abundance of RNA transcripts is only one of the steps of protein expression level and activity. The function of protein is regulated by multiple factors. The most extensively pathological functions of 16E6 are through downregulation of p53 and PDZ-domain containing proteins in a ubiquitination-dependent degradation manner. Therefore, there is no decrease in RNA transcripts levels of these genes in NIKS 16E6 (Figure 5.5.2). Especially to note that MAGI1 has negligible RNA expression level, but western blot and immunostaining detections suggest the relatively high abundance of the MAGI1 protein in NIKS. Regarding this issue, recently bioinformaticians have proposed that the proteins with low turnover rates, such as cell adhesion molecules have a greater functional impact compared with the high turnover proteins such as metabolic-associated genes (Smail *et al.* 2021). As the proteomic analysis develops, more protein modification information becomes available, the prediction model will be more realistic in the foreseeable future, and bioinformatic analysis will be more instructive.

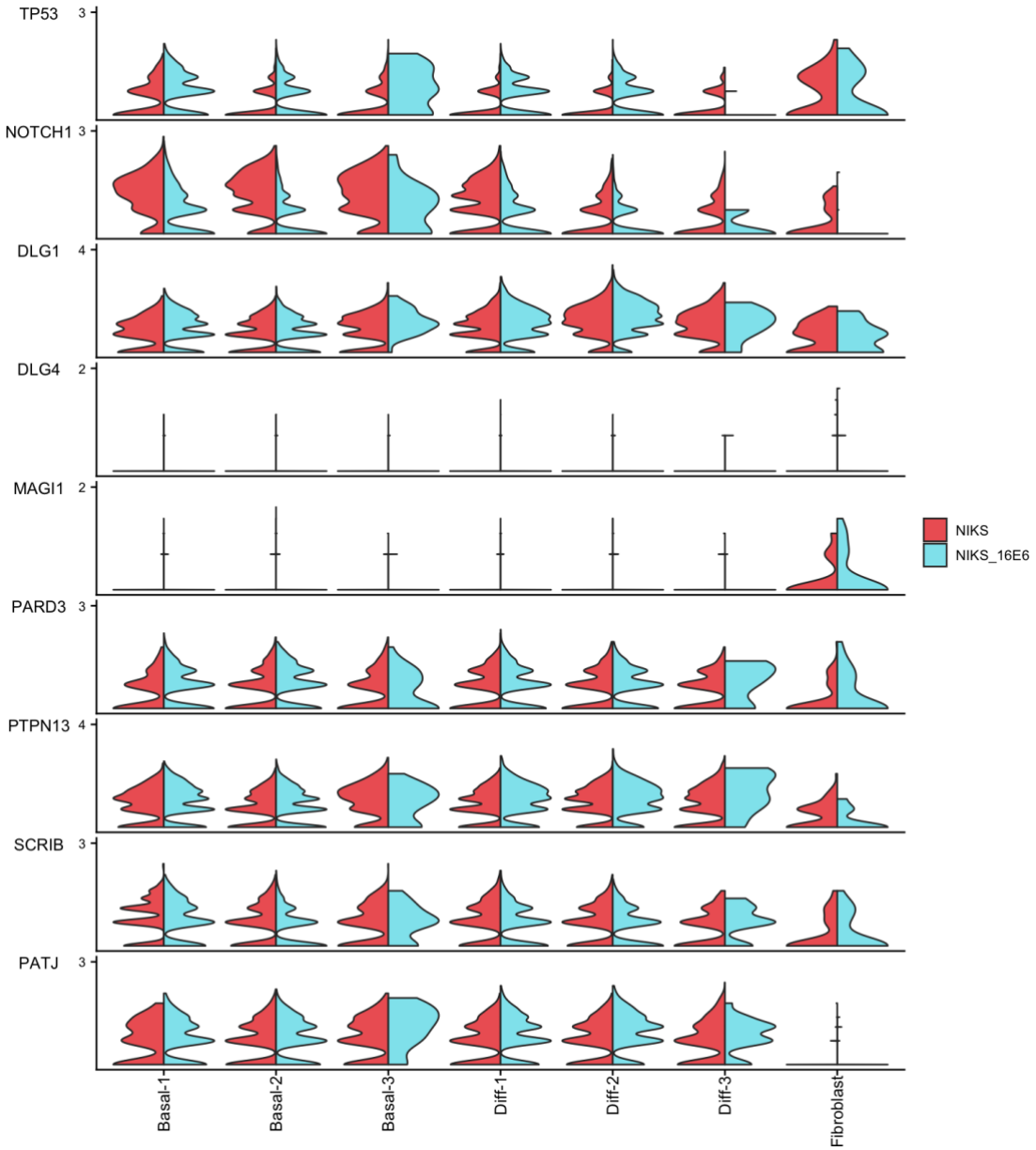


Figure 5.5.2: HPV16E6 pathological downregulation in cellular protein expression level does not depend on transcription regulation. Violin plot of TP53, Notch1, and PDZ-domain containing proteins targeted by 16E6 (DLG1, DLG4, MAGI1, PARD3/hPAR3, PTPN13/FAP, SCRIB, PATJ). The y-axis represents the relative expression level, and the width of the violin represents the relative proportion of cells at a certain expression level. NIKS and NIKS 16E6 are plotted next to each other distinguishable by colour.

6. Discussion and Future Scope

6.1 Development of *in vitro* model systems to understand persistent HPV infections

In this thesis, two-layered *in vitro* model systems have been developed to understand persistent HPV infections. HPV generally infects the fully stratified epithelium. The establishment of successful infections requires the viral particle to gain access to the proliferative basal cells to allow persistence and propagation of the virus (Doorbar 2006; Schiller *et al.* 2010). Basal cells move up through the epithelium and undergo terminal differentiation to form the multi-layered structure.

The preliminary cell count growth assay revealed that the NIKS population doubling rate reduced with increasing cell density and eventually achieved a stable state. In order to better understand this phenotype, we used the FUCCI method to distinguish between proliferating (S/G2/M phase) and resting (G0/G1 phase) cells in addition to cell density counts. Consistent with the results of the growth experiment, the percentage S/G2/M population decreased as cell density increased and was eventually maintained at a 5% populational proliferation rate. Because 5% of the population is still cycling at post-confluence, an equal amount of NIKS is delaminating from the bottom of the tissue culture plate and being pushed to the upper layers of the dish, generating a multi-layered structure. This subpopulation is committed to differentiation and can identified by K10 expression.

Proliferation of HPV-infected cells is enhanced by early viral protein expression. Several HPV proteins have either direct or indirect effects on cells (reviewed in Hamid, Brown and Gaston, 2009). The accessory proteins E6 and E7 promote the growth of HPV-infected epithelial cells. The action of these proteins subverts the normal process of terminal differentiation, causing an increase in both the number of infected cells and virus genome replication that will ultimately release infectious viral particles. Both E6 and E7 increase the progression of the cell cycle, and both of these proteins associate with cell cycle regulators (Münger *et al.* 1989b; Scheffner *et al.* 1990). More recent research has dissected the functions of E6 and E7 in the virus life cycle, in which E6 is the major basal epithelium regulator while E7 acts in the upper layers of the epithelium (Kranjec *et al.* 2017).

The cell count assay confirmed the increase of NIKS proliferation when expressing HPV16 E6 protein. The E6 protein expression is driven by the retroviral vector LXSN, which uses a 5'LTR promoter, resulting in a comparable E6 transcript level compared with the HPV16 early promoter p97 (unpublished data from Doorbar's lab).

Based on the growth advantage HPV16E6 provides to NIKS cells, there was an increase in the proportion of proliferating cells (%S/G2/M) in various cell densities, and the saturated proliferation rate was also significantly higher compared with NIKS. Staining with K10 indicated that NIKS 16E6 cells commit to differentiation at a higher density, and this is referred to as delayed differentiation. Utilising the advancement of high-content imaging machines, ArrayScan (Thermo Fisher) and Opera Phenix (Perkin Elmer), the NIKS and FUCCI/K10 can be used as a reliable and reproducible model system to study basal-to-suprabasal cell transitions and how it is regulated by persistent HPV infections.

Because infected cells come into contact with uninfected cells at the basal layer in real infection scenarios, it is important to mimic the relative competition between the two populations *in vitro*. We were able to dissect the multi-layered structure created after confluence using a confocal imaging device (Opera Phenix). By seeding the eGFP and mCherry cells expressing either empty vector LXSN or LXSN 16E6 at confluency, the relative percentage of each population at the bottom layer of the mixed culture may be measured over a 9-day period. While there is no difference between eGFP LXSN and mCherry LXSN cells, mCherry LXSN 16E6 eventually preferably grow to colonise the bottom layer.

6.2 E6-mediated epithelial homeostasis is a common strategy shared by the different types of HPV to establish persistent infections

HPV-related lesions can persist for years, for both low- and high-risk types. In general, persistent infection is an indication that the pathogens have successfully adapted to their host, and that extinction of the pathogen is unlikely to occur as long as the host continues to exist. Evidence suggests all HPV types existed before *Homo Sapiens* evolved (Bernard *et al.* 2006; Chan *et al.* 1997; Deau *et al.* 1991; Heinzel *et al.* 1995). Hence, based on Fahrenholz's rule,

the co-evolution of HPV and host could reflect the diversity and niche of epithelium. Large-scale sequencing analysis identified that most HPVs are skin commensals for immunocompetent people (Antonsson *et al.* 2000). Due to the specificity between HPV and their tropism, we were able to study the specific homeostasis at different epithelium sites.

In this project, we hypothesised that E6 is a common evolutionary conserved basal epithelial homeostasis regulator for persistent infections. To test this idea, representatives from each HPV group that is associated with clinical symptoms were chosen, namely α-HPV16, α-HPV11, α-HPV27, β-HPV8, γ-HPV65 and μ-HPV1. The E6 protein from each type is cloned and expressed by the same retroviral vector LXSN. With the *in vitro* model systems, the findings confirmed that despite their inherent variations in carcinogenicity, low- and high-risk HPV strains have independently evolved similar molecular strategies to regulate viral maintenance and vegetative reproduction to complete the viral life cycle (as suggested by Doorbar et al. 2015). The next steps will be trying to identify the commonality among the epithelial sites, and identify molecular pathways that governs basal-suprabasal transition targeted by HPV E6 proteins.

6.3 Targeting viral reservoir clearance

Persistent infection is not a unique strategy for HPV pathogenesis. Actually, approximately 12% of the worldwide cancer incidence is caused by a persistently infected virus (White *et al.* 2014), such as Human Papillomavirus (HPV), Epstein-Barr virus (EBV), Hepatitis B virus (HBV), Hepatitis C virus (HCV), Human Herpesvirus-8 (HHV-8), Human T-cell Leukaemia virus type-1 (HTLV-1) and Human Immunodeficiency virus (HIV) (Boldogh *et al.* 1996; Hausen 1991; Nicoll *et al.* 2012). While only a small proportion of persistently infected HPV developed into cancer, the viral reservoir at the basal epithelium is essential for lesion formation, disease progression and recurrence (Doorbar 2013; Gravitt & Winer 2017).

Rabbit Papillomavirus models (CRPV and ROPV) suggest latent papillomavirus infection persists after initial lesion regression, in which no viral particle or late protein E4 expression detectable in tissue but viral DNA positively stained in the lower epithelial layers (Amella *et*

al. 1994; Maglennon *et al.* 2011b). Additionally, only a subset of basal cells is capable of persistent PV infections as less than one copy of viral DNA detected from the basal samples analysed (Maglennon *et al.* 2011b). This persistent latent infected subset of basal cells is hypothesised to have 'stem-like' properties (Zhang *et al.* 1999b). Once established, the latent reservoir is thought to be undetectable by the molecular regulation of persistence and host immune surveillance (Mbonye & Karn 2017). Therefore, understanding the persistently HPV-infected cells will provide us new insights into basal epithelial homeostasis, and more importantly, targeting the viral reservoir will facilitate complete viral clearance both for productive and latent infections.

Drug repurposing is a cost-effective method for new drug development by maximising the usage of currently approved drugs. Over 30% of drugs marketed since 1997 have been proposed to have multiple therapeutic targets over 20 years (Sahragardjoonegani *et al.* 2021). A High-Throughput Screening of FDA-approved drugs against persistent HPV infections was carried out. Initially, with the readout of NIKS cell density and %S/G2/M, roughly 13% (167/1265) of the small molecules (HIT molecules) have specific negative regulation to NIKS 16E6 cells with no obvious change in NIKS groups. The HIT molecules are further analysed with the competition assay model, and finally narrowed down to 12 top candidates. The top candidates are confirmed to have a dose-dependent inhibition effect, specifically target 16E6 expressing NIKS only. All of the 12 candidate small molecules were proposed to target different cellular pathways, while none of their current usages is directly related to the virus or stratified epithelium. The results provided a huge potential to discover the virus-targeted basal epithelial pathways.

Two of the top candidates, Lopinavir and Celecoxib, have some evidence against HPV-related diseases previously published (Batman *et al.* 2011; JS *et al.* 2017; Lahiri *et al.* 2015; Rosenthal *et al.* 2012). Screening with the same family drugs of Anti-Retroviral Therapies (ART) and COX-2 selective inhibitors further confirmed the specificity of Lopinavir and Celecoxib against HPV16 E6 modulated homeostasis, as the rest of the families failed to achieve similar anti-HPV functions.

Lopinavir, an HIV protease inhibitor, was originally developed as a combination therapy to reduce the serum intervention and the ability of HIV to evolve resistance to the drug based on the previous protease inhibitor Ritonavir (Sham *et al.* 1998). For antiretroviral treatment, Lopinavir alone is not biologically active as a HIV protease inhibitor, so it's prescribed as a combination therapy together with Ritonavir (FDA, 2000). However, the HTS result shown Ritonavir failed to exhibit the same function for anti-HPV targets. Depletion of viral E6/E7 transcripts (Park *et al.* 2021) or inhibition or E6-dependent p53 degradation (Hampson *et al.* 2006) has been proposed for cervical carcinomas cells *in vitro*. In comparison, the p53 protein expression level or its localization does not have obvious changes in our NIKS 16E6 model, data collected by time-course immunofluorescence. Unpublished data from Yuwen Chen, a current PhD student in Doorbar's lab, suggested that Lopinavir increases p53 transcriptional activity as seen as upregulation of p21 in NIKS 16E6 cells. Interestingly, Lopinavir is also effective against the low-risk cutaneous α-HPV27 E6 protein, suggesting it may target a common pathway conserved among the α-group.

Celecoxib is a COX-2 inhibitor and nonsteroidal anti-inflammatory drug (NSAID). COX-2 is a cyclooxygenase that is induced at sites of inflammation. The level of COX-2 is positively related to disease severity of HPV-positive cervical squamous lesions and early cancers (Greenhough *et al.* 2009; Saldivar *et al.* 2007), yet no solid conclusions of whether HPV E6 protein directly interacts with the proteins in the COX-2 pathway. Celecoxib can downregulate E6 expression and therefore promote p53-dependent cell apoptosis in the HPV18+ HeLa cells (Saha *et al.* 2011). However, this phenotype is not observed in our results. Organotypic raft culture immunohistochemistry has shown increased nuclear localisation or COX-2, which is an indicator of stemness in several cancers (Bocca *et al.* 2011; Cheng *et al.* 2018; Thanan *et al.* 2012). The nuclear COX-2 belongs to the non-canonical COX pathway to act as a transcriptional factor, and the function of nuclear COX-2 for HPV infections can be explored. A study in mice supported the potential of using celecoxib against HPV16 infections, which K14-HPV16 transgenic mice have significantly higher CD8+ T cell infiltration upon celecoxib treatment (Santos *et al.* 2016).

Next, the project aim was to dissect the molecular pathways targeted by HPV16 E6, which are essential for basal persistence and competition, which will later fuel into our understanding of the molecular basis for the anti-HPV drugs.

6.4 HPV16 E6 and its cellular targets

Most of the high-risk HPV16 E6 research focused on its function during pathogenesis and viral life cycle. The well-studied functions of HPV16 E6 are the degradation of cellular p53 and PDZ-domain-containing proteins in the epidermis. Both functions are indispensable for high-risk E6 malignant transformation (Mantovani & Banks 2001). The absence of p53 contributes to cell cycle progression (Thomas *et al.* 1999b), while PDZ-domain containing protein degradation disrupts the organisation of epithelial architecture and promote epithelial-mesenchymal transition (EMT) in the transformed cells (Ganti *et al.* 2015). The transformed infected cells are unable to support viral vegetative genome amplification and new virion production (Graham 2017a), and it is a dead-end for the viral life cycle which accounts for a small proportion of HPV infections from clinical observations (Szymonowicz & Chen 2020). The importance of HPV16E6 targeting p53 and PDZ-domain-containing proteins for degradation in modulating epithelial homeostasis and persisting viral reservoir need be explored.

HPV16 E6 mutants that are deficient in p53 (16E6SAT) or PDZ-domain containing protein (16E6dPDZ) degradation were described previously (Murakami *et al.* 2019). The results suggested p53 as the main cell cycle regulator targeted by HPV16 E6, in driving cell cycle entry at all cell densities. As a result, NIKS 16E6SAT cells differentiated at similar densities as the uninfected cells and largely impaired its competitive advantage in colonising the bottom layer of the mixed cultures. The NIKS 16E6dPDZ cells, in contrast, still have increased proliferation and delayed commitment to differentiation in comparison to NIKS. However, interestingly, NIKS with dPDZ mutant also failed to compete with the uninfected cells as the wildtype E6 harbouring cells. The findings proposed new functions of targeting PDZ proteins during the viral life cycle, which can promote HPV basal persistence independent to cell cycle regulations.

6.4.1 p53

The tumour suppressor protein p53 is an attractive target for many viruses to establish persistent infections. In addition to HPV E6 (Scheffner *et al.* 1990), the SV40 large T antigen (LT) (Jiang *et al.* 1993), the HIV Tat (Li *et al.* 1995), the HLTV-1 Tax (Mulloy *et al.* 1998) and the HCV NS3 (Kwun *et al.* 2001) proteins are also known to target p53 and inhibit its function. Once it is activated by the unscheduled induction of DNA replication, the tumour suppressor p53 is able to promote either the cell cycle arrest or the death of the infected cell, both of which are major constraints in the process of viral replication (Rivas *et al.* 2010).

P53, discovered as a tumour suppressor protein, governs cell cycle progression in G1 to S phase (Kuerbitz *et al.* 1992). DNA damage response upregulates wildtype p53 and induces cell cycle arrest at the G1 phase. The level of p53 naturally oscillates during the cell cycle. Several models have been developed considering both amplitude and frequency of the p53 level change affect cell cycle progression and cell fate decisions (Jun-Feng & Ya 2010; Sun *et al.* 2009; Wang *et al.* 2019). Experimentally, only sustained p53 induction can upregulate the expression of genes associated with terminal cellular fates (Purvis *et al.* 2012). The elevation of the p53 baseline level also induces cell competition by mechanical stress, in which cells have high p53 levels have lower population confluency density and become extruded when in contact with the normal cells (Wagstaff *et al.* 2016).

Loss or functional deficient form of p53, on the other hand, creates a pro-disease cell competition environment. In the progenitor cells, there is a trend for constant selection for the least damaged cells that are p53 dependent. Unlike the classical p53 pathways in which the p53 level is rapidly changing during the cell cycle, this progenitor competition could persist for months, and it depends on the relative instead of the absolute p53 level of its competitors (Bondar & Medzhitov 2010). This phenomenon may explain the low expression level of HPV E6 at the basal layers and the ability of HPV to establish latent infections. HPV infects the basal cells with the proliferative potential to establish successful infections, which are considered ‘stem-like’ basal cells (Doorbar 2013). With a low expression of E6 protein in the basal viral reservoir, it may be sufficient to outcompete the uninfected cells by sustained transient lowering of cellular p53 level. Also, loss of p53 favours competition for cells that are early in differentiation during development (Dejosez *et al.* 2013).

6.4.2 PDZ-domain-containing proteins

The C-terminus PDZ-binding motif is unique but highly conserved among high-risk HPV E6 proteins. Targeting PDZ-domain-containing proteins is also conserved among the pathogenic viruses (Table 6.4.2) (Javier & Rice 2011), while from the commonality and differences between different types of viruses, we can broadly differentiate that some PDZ-domain-containing proteins plays distinct roles for HPV infections. For example, the DLG1 protein is targeted by the epithelium (HPV and Adenovirus), leukocytes (HTLV-1, HIV) and the nervous system (Rabies) infectious viruses suggesting the DLG1 has multiple functions and might be the highly conserved oncogenic targets.

PDZ targets
CAL/GOPC
Papillomaviridae
HPV E6 proteins
DLG1
Adenoviridae, Rhabdoviridae, Orthomyxoviridae, Papillomaviridae, Retroviridae
Human Ad E4-ORF1 proteins, Rabies G proteins, Influenza A NS1 proteins, HPV E6 proteins, HTLV-1 Tax protein, HTLV-1 Env protein, HIV Gag protein
DLG4
Papillomaviridae, Retroviridae
HPV E6 proteins, HTLV-1 Tax protein
hPAR3
Papillomaviridae
HPV E6 proteins
MAGI1
Adenoviridae, Orthomyxoviridae, Papillomaviridae
Human Ad E4-ORF1 proteins, Influenza A NS1 proteins, HPV E6 proteins
MAGI2
Orthomyxoviridae, Papillomaviridae
Influenza A NS1 proteins, HPV E6 proteins
MAGI3
Orthomyxoviridae, Papillomaviridae, Retroviridae
Influenza A NS1 proteins, HPV E6 proteins, HTLV-1 Tax protein
MUPP1
Adenoviridae, Rhabdoviridae, Papillomaviridae
Human Ad E4-ORF1 proteins, Rabies G proteins, HPV E6 proteins
PALS1
Coronaviridae
SARS E protein
PATJ
Adenoviridae, Papillomaviridae
Human Ad E4-ORF1 proteins, HPV E6 proteins
PDLIM2
Orthomyxoviridae, Retroviridae
Influenza A NS1 proteins, HTLV-1 Tax protein
PTPN13
Papillomaviridae
HPV E6 proteins
PTPN3
Papillomaviridae
HPV E6 proteins
PTPN4
Rhabdoviridae
Rabies G proteins
Scribble
Flaviviridae, Orthomyxoviridae, Papillomaviridae, Retroviridae
TBEV NS5 protein, Influenza A NS1 proteins, HPV E6 proteins, HTLV-1 Tax protein
TIP-1
Papillomaviridae, Retroviridae
HPV E6 proteins, HTLV-1 Tax protein
TIP-2
Hepadnaviridae, Papillomaviridae, Retroviridae
HBV Core protein, HPV E6 proteins, HTLV-1 Tax protein
ZO-1
Flaviviridae, Papillomaviridae
TBEV NS5 protein, DV NS5 protein, HPV E6 proteins
ZO-2
Adenoviridae
Human Ad9 E6-ORF1 protein

Table 6.4.2: PDZ-containing proteins are common targets for pathogenic viruses (adapted from Javier & Rice 2011).

The dPDZ mutant results can be well explained for viral persistence. The cellular PDZ-domain containing proteins targeted by HPV16 E6 is further dissected to evaluate individual function

during basal cell homeostasis. NIKS with shRNA knockdown of HPV16 E6 targeted PDZ-domaining-containing proteins, namely DLG1 (Gardiol *et al.* 1999), DLG4 (Handa *et al.* 2007), MAGI (Glaunsinger *et al.* 2000), hPAR3 (Yoshimatsu *et al.* 2017), FAP/PTPN13 (Spanos *et al.* 2008), Scrib (Nakagawa & Huibregtse 2000) and PATJ (Latorre *et al.* 2005), were analysed for their cell cycle status, commitment to differentiation and relative competitive ability against the pseudo-knockdown NIKS (shRNA Luciferase). Among these PDZ-domain-containing proteins, they can be further grouped into three cell-to-cell junction complexes. The DLGs and Scrib form the lateral adherent junction, the hPAR3 and MAGI belong to the same tight junction complex, and PATJ is normally located on the apical surface (James & Roberts 2016; Javier & Rice 2011). The cell junction complex grouping is compatible with the shRNA results, in which primarily the DLGs/Scrib and hPAR3/MAGI have a distinct role in NIKS homeostasis, with the former knockdowns promoting cell detachment and the latter knockdown in NIKS behaving like E6 persistently infected cells.

The DLGs/Scrib knockdown NIKS exhibited lower cell confluence density, fewer cycling cells, more likely to differentiate and therefore being outcompeted by the normal cells. Controversial data have been reported on whether DLG1 knockdown affects epithelial cellular junctional protein compositions and downstream cellular pathways, such as E-cadherin and β-catenin, leaves the function of DLG1 in epithelium remains obscure (Laprise *et al.* 2004; Stucke *et al.* 2007; Yamanaka & Ohno 2008). The Scrib knockdown cells are extremely sensitive to crowd mechanical stresses, and it is believed that this is because of an increase in baseline levels of p53. Specifically, the compaction of Scrib knockdown cells activates the Rho-associated kinase (ROCK) and p38, which ultimately leads to further elevation of p53 and cell death (Wagstaff *et al.* 2016). The high p53 activity is central to Scrib knockdown cell phenotypes as low cellular density and hypersensitivity to the crowd, which are the two main features of the mechanical competition mechanisms.

On the opposite, the MAGI/hPAR3 knockdown NIKS have increased proliferation rate, higher confluence density, delayed commitment to differentiation and are more competitive to colonise the basal layers *in vitro*, similar to the wildtype HPV16 E6 expressing cells. MAGI1 is expressed in both polarised and non-polarised cells, and its expression is distributed around the body (reviewed in Excoffon et al., 2022). MAGI1 has been shown to potentially modulate

cell proliferation via its interaction with β-catenin (Dobrosotskaya & James 2000), the Notch-ligand DLL1 (Mizuhara *et al.* 2005) and the MAPK pathway proteins (Boldt *et al.* 2016). Additionally, MAGI1 recruits cellular E-cadherin for cell-to-cell junction formation (Dobrosotskaya & James 2000). The E6-dependent downregulation of E-cadherin level is directly associated with lowering the antigen-presenting Langerhans cells number in the infected tissue (Matthews *et al.* 2003). This active immune evasion strategy for HPV persistent infection may be achieved by the degradation of MAGI-1. In *Drosophila*, Par3 accumulates during mitosis, and its localisation and abundance are correlated with the daughter cell fate (Lechler & Fuchs 2005). In mammals, loss of Par3 homolog in the mouse mammary gland induces cell proliferation and cell apoptosis simultaneously while disrupting the normal differentiation pathway and causing the accumulation of mammary progenitors *in vitro* (McCaffrey & Macara 2009).

Apart from Scrib, the function of these cell junctional proteins in normal tissue competition (non-malignant) remains unclear. The shRNA NIKS results provided some new insights for cell-to-cell junction proteins in basal epithelial homeostasis. It will be worth examining the expression and localisation of the functional, structural binding partners of the PDZ-domain containing protein knockdown NIKS to reach a confident conclusion on their importance for HPV persistent infections. The combinations of a variety PDZ-domain containing proteins degraded by different high-risk HPV E6 proteins have been suggested as indicators of their oncogenicity (Thomas *et al.* 2016b). Therefore, the high-risk HPV E6 proteins targeting the PDZ-domain containing proteins are more complex than promoting cell migration for carcinogenesis, of which individual cellular targets have their own roles in the viral life cycle to be explored.

6.5 Trigger for basal cell fate commitments

The basal epithelial layer is a heterogeneous population, in which cells constantly experience turnover and cell fate decisions to maintain the stratified epithelium from loss of the cornified layer. Basal keratinocytes need to experience regulation in the cell cycle, decreased extracellular matrix attachment with the basal lamina, rearrangement of cell-cell junction

conformation and finally, initiation of terminal differentiation fate. To understand the different processes during the basal-to-suprabasal layer transition and dissect how individual process is controlled by the HPV persistent infection, a high-resolution molecular technique, single-cell RNA sequencing (scRNA seq), was applied to the post-confluent NIKS and NIKS 16E6 population. The results finely separated our *in vitro* model system into six different cell clusters, namely Basal-1/2/3 and Differentiation-1/2/3 (Diff-1/2/3). The NIKS and NIKS 16E6 post-confluent populations have the same clusters, but the relative abundance of each individual cluster is changed. The scRNA seq results are fully compatible with the phenotypic analysis with FUCCI and K10.

6.5.1 MMP7 is predicted to be the basal-to-suprabasal transition determinant targeted by HPV16 E6

One of the striking features of NIKS 16E6 scRNA seq is the loss of the Basal-3 cluster compared to NIKS. The Basal-3 cluster is positioned between the quiescent Basal-2 cluster and the early Diff-1 cluster, suggesting its potential function in triggering differentiation. Differentiation expression analysis identifies several molecular markers unique to the Basal-3 cluster. Among the Basal-3 markers, upregulations of ribosomal protein RPS27 and long non-coding RNA GAS5 suggest the metabolic profile of this cluster is greatly altered. When ordered by log2 fold change of these Basal-3 markers, MMP7 is the most upregulated transcript. The positive correlation of MMP7 with the early differentiation marker IVL expression level within the Basal-3 cluster further indicates that MMP7 is highly likely to be a basal-to-suprabasal transition determinant.

The Matrix Metalloproteinase-7 (MMP7) is a secreted zinc- and calcium-dependent endopeptidase that degrades a broad range of extracellular matrix substrates and additional substrates. MMP7 is one of the few MMPs expressed in the polarised epithelium, which functions are partially determined by its translocation to either apical or basolateral sites (Nelson *et al.* 2000). The protein is secreted in an inactive form. The pre-domain is cleaved when it is released from the cell, and then the pro-domain is autocatalytically removed upon activation of its enzymatic activity (Ii *et al.* 2006). Transcription of MMP7 is induced by the Wnt pathway β-catenin activation (van Es *et al.* 2005) and acts as an effector for Sox9-

dependent differentiation of the intestinal epithelium (Bastide *et al.* 2007). MMP7 appears to be involved in the shedding process in epithelial cells, especially on the apical surface of cells (Nishi *et al.* 2003). The specificity of MMP7 to its substrates is ranked highest among the MMP family (reviewed in (Okada 2013). Cleavage of integrin-β4 (von Bredow *et al.* 1997)and E-cadherin (McGuire *et al.* 2003) by MMP7 are suspected to be the most relevant functions in the basal-to-suprabasal transition stage.

6.5.2 Cell-cell communication pathways Notch and Wnt are common targets by HPV E6 proteins

While MMP7 may function as an effector for basal cell delamination and differentiation, the cell signalling pathways upstream of MMP7 activation need to be studied. Literature proposed Notch and Wnt signalling pathways are essential for epithelial homeostasis (Okuyama *et al.* 2008; Rangarajan *et al.* 2001b; Slavik *et al.* 2007; Widelitz 2008), and targeted by various HPV E6 proteins (Brimer *et al.* 2012; Drews *et al.* 2019; Kranjec *et al.* 2017; Meyers *et al.* 2013, 2017; Sominsky *et al.* 2017).

Notch signalling is conducted by the transmembrane ligands (Delta-like or Jagged) binds to the receptors (Notch) expressed on adjacent cells. The cleavage of the intracellular domain of Notch receptor (NICD) by γ-secretase resulted in the translocation of NICD to the nucleus and activated downstream gene transcriptions. Notch1 is the primary Notch receptor expressed by the basal epithelium (Sakamoto *et al.* 2012). The downregulation of Notch1 expression in a p53-dependent manner (Kranjec *et al.* 2017), is considered to be the main strategy for high-risk HPV16 E6-modulated basal persistence. In addition, differential expression of the Delta-like ligands (DLL1 or DLL4) in the signal-sending cells will direct the signal-receiving cells towards opposite cell fates (Khelil *et al.* 2021; Nandagopal *et al.* 2018). When DLL1 bind with Notch1, the cell receiving the signal will trigger a differentiation scheme. Here, the scRNA seq show, at the same time as downregulating Notch1 transcription, induction of DLL1 transcription by HPV16 E6 expression suggesting the infected cells are sending signals to the uninfected neighbours to differentiate while the infected cells generally receive very low Notch signals.

The regulation of the Wnt pathway is more complex. Generally, the pathway is activated by the binding of secreted Wnt ligands to the transmembrane heterodimer receptor comprised of the Frizzled (Fz) and the lipoprotein receptor-related proteins (LRP). The unphosphorylated form of β-catenin is then translocated into the nucleus and serves as a transcriptional factor to activate the downstream pathways (Barker 2008). The secreted Wnt ligands could not distinguish signal sending and receiving cells as the Notch signalling pathway. From the scRNA seq results, the Wnt pathway receptors are restricted to the undifferentiated clusters and upregulation of LRP5/FZD5 receptor expression by 16E6 suggests their potential function in proliferation. Individual Wnt ligands are proven to exhibit cell-specific functions (Clevers & Nusse 2012). As NIKS 16E6 has extended Wnt and increased receptor expressions, Wnt6 functions to stimulate cell proliferation (Bonnet *et al.* 2021), and its loss contribute to differentiation (Beaton *et al.* 2016) could be well integrated with the theory of HPV persistent infections.

6.6 Future scopes

Thus far, the project described the molecular insights into which HPV E6 proteins are essential for the establishment of persistent infections. High-throughput screening of small molecules suggested the possibility of targeting clearance of persistently infected basal cells. Progress have been made towards dissecting the heterogenicity of the basal epithelial cells. Cell trajectories and interaction models have been predicted based on the previous literature and direct future experiments. There are two major aims remain to be solved: defining the basal epithelial cell competition mechanisms targeted by the E6 protein; successful development of anti-HPV therapies requires all viral proteins for the completeness of the viral life cycle.

6.6.1 Exploring mechanisms for cell competition modulated by HPV E6

Apoptotic and proliferative regimes define the homeostatic cell density for each epithelium (Matamoro-Vidal & Levayer 2019). It is likely that the cells with a higher homeostatic density will expand at the expense of the cells with a lower homeostatic density if two cell populations in a mixture have different homeostatic densities (Baker 2020). A mixture is likely to enter a

state in which one population is proliferating, but the other population is being lost. By discovering specific markers for cell competition, we would be able to identify the process more easily.

The stretch-activated cation channel Piezo, which is Ca^{2+}-permeable, triggers cell extrusion independently of caspase activation (Eisenhoffer *et al.* 2012b). In epithelial tissues, differential growth could lead to the accumulation of mechanical stress (Shraiman 2005). Those cells that are proliferating more rapidly will push on their neighbours, leading to an increase in local pressure. This will lead to preferential elimination of the slower-proliferating cells in the vicinity of the fast-growing population if the fast-growing population is less sensitive to pressure than the neighbouring cells. Essentially, this mechanism is based on three hypotheses: the epithelial tissue must behave as a solid (i.e., stress does not dissipate through cell movement or rearrangement); compaction must be capable of causing cell death; and the sensitivity to compaction must be different among cell types (Matamoro-Vidal & Levayer 2019).

6.6.2 The importance of other HPV proteins and the completeness of viral life cycle

The high-risk E6 proteins are transcribed as bi-cistronic RNA, which by splicing produce the viral E7 protein. Although our results emphasised the importance of E6 protein during viral persistence, the E7 protein should not be ignored. High-risk E6 and E7 are together sufficient to induce immortalisation of primary human keratinocytes (Hawley-Nelson *et al.* 1989b; Münger *et al.* 1989a). Individual expression of each oncogene induced epithelial hyperplasia and skin tumours, however E7 was found to primarily cause benign, highly differentiated tumours(Herber *et al.* 1996), whereas those promoted by E6 were mostly malignant (Song *et al.* 1999). The two oncoproteins may play different roles during carcinogenesis, and their cooperative actions may also be a key factor in transformation. (Mantovani & Banks 2001).

Differentiated epithelia, which no longer produce cellular DNA, are inherently incapable of supporting HPV replication (McLaughlin-Drubin & Münger 2009). The E7 protein is mainly responsible for maintaining the competence of infected cells leaving the basal layer to support DNA synthesis (Cheng *et al.* 1995). E7 proteins seem to play a specific role in the viral

life cycle differently depending on the HPV type. HPV11 and 31 E7 have been shown to be necessary for episomal maintenance of the viral genome (Oh *et al.* 2004; Thomas *et al.* 1999a), whereas HPV16 and 18 E7 are required to support the productive stage of a viral life cycle (Flores *et al.* 2000; McLaughlin-Drubin *et al.* 2005). A study showed HPV16 E7 bind to the pRB family members is to allow the suprabasal cells to synthesize DNA, instead of inducing their degradation (Collins *et al.* 2005).

A rapid proteasomal degradation mediated by the ubiquitin ligase MDM2 results in a short half-life for p53 tumour suppressor protein in normal cells (Honda & Yasuda 2000). E7 expression in HPV cells is associated with higher steady-state levels of p53 (Demers *et al.* 1994), and p53 half-lives are longer in E7 expressing cells, suggesting that p53 degradation may be perturbed by E7 (Jones *et al.* 1997a). It has been shown that p53 stabilisation in E7 cells is independent of $p14^{ARF}$, an inhibitor of MDM2-mediated p53 degradation (Bates *et al.* 1998; Münger *et al.* 2001; Y *et al.* 1998). High-risk HPV E6/E7 expressing cervical carcinoma cells lines do not undergo normal MDM2-mediated p53 turnover (Hengstermann *et al.* 2001), and MDM2 is less efficiently bound to p53 in E7 cells compared to normal cells (Seavey *et al.* 1999). A correlation has been demonstrated between E7-mediated pRB degradation and p53 stabilization (Jones *et al.* 1997b) and MDM2 has been implicated in pRB degradation (Uchida *et al.* 2005). The increase in cellular p53 level by E7 expression is implicated as an advantage for viral clearance once the E6 functions are inhibited.

There is a need for HPVs to complete its productive viral life cycle, which requires infected cells to differentiate. By understanding and antagonizing viral strategy in maintaining infected reservoir at the basal epithelium, the project opens the way of the future investigation of how HPV regulates normal epithelial homeostasis and will facilitate the development of HPV clearance therapies.

Bibliography

Ajiro, M., & Zheng, Z. M. (2015). E6^E7, a novel splice isoform protein of human papillomavirus 16, stabilizes viral E6 and E7 oncoproteins via HSP90 and GRP78. *MBio*, **6**(1). doi:10.1128/MBIO.02068-14/SUPPL_FILE/MBO001152166ST3.PDF

Alcolea, M. P., & Jones, P. H. (2015). Cell competition: winning out by losing notch. *Cell Cycle (Georgetown, Tex.)*, **14**(1), 9–17.

Alkhilaiwi, F., Paul, S., Zhou, D., … Yuan, H. (2019). High-throughput screening identifies candidate drugs for the treatment of recurrent respiratory papillomatosis. *Papillomavirus Research*, **8**, 100181.

Allen-Hoffmann, B. L., Schlosser, S. J., Ivarie, C. A. R., Meisner, L. F., O'Connor, S. L., & Sattler, C. A. (2000a). Normal Growth and Differentiation in a Spontaneously Immortalized Near-Diploid Human Keratinocyte Cell Line, NIKS. *Journal of Investigative Dermatology*, **114**(3), 444–455.

Allen-Hoffmann, B. L., Schlosser, S. J., Ivarie, C. A. R., Meisner, L. F., O'Connor, S. L., & Sattler, C. A. (2000b). Normal Growth and Differentiation in a Spontaneously Immortalized Near-Diploid Human Keratinocyte Cell Line, NIKS. *Journal of Investigative Dermatology*, **114**(3), 444–455.

Amella, C. A., Lofgren, L. A., Ronn, A. M., Nouri, M., Shikowitz, M. J., & Steinberg, B. M. (1994). Latent infection induced with cottontail rabbit papillomavirus. A model for human papillomavirus latency. *The American Journal of Pathology*, **144**(6), 1167.

Antonsson, A., Forslund, O., Ekberg, H., Sterner, G., & Hansson, B. G. (2000). The Ubiquity and Impressive Genomic Diversity of Human Skin Papillomaviruses Suggest a Commensalic Nature of These Viruses. *Journal of Virology*, **74**(24), 11636–11641.

Aragona, M., Panciera, T., Manfrin, A., … Piccolo, S. (2013). A Mechanical Checkpoint Controls Multicellular Growth through YAP/TAZ Regulation by Actin-Processing Factors. *Cell*, **154**(5), 1047–1059.

Araujo-Arcos, L. E., Montaño, S., Bello-Rios, C., Garibay-Cerdenares, O. L., Leyva-Vázquez, M. A., & Illades-Aguiar, B. (2022). Molecular insights into the interaction of HPV-16 E6 variants against MAGI-1 PDZ1 domain. *Scientific Reports 2022 12:1*, **12**(1), 1–14.

Arpin-André, C., & Mesnard, J.-M. (2007). The PDZ Domain-binding Motif of the Human T Cell Leukemia Virus Type 1 Tax Protein Induces Mislocalization of the Tumor Suppressor hScrib in T cells *. *Journal of Biological Chemistry*, **282**(45), 33132–33141.

Aydin, I., Weber, S., Snijder, B., … Schelhaas, M. (2014). Large Scale RNAi Reveals the Requirement of Nuclear Envelope Breakdown for Nuclear Import of Human Papillomaviruses. *PLoS Pathogens*, **10**(5). doi:10.1371/journal.ppat.1004162

Baker, N. E. (2020). Emerging mechanisms of cell competition. *Nature Reviews Genetics 2020 21:11*, **21**(11), 683–697.

Balda, M. S., Garrett, M. D., & Matter, K. (2003). The ZO-1-associated Y-box factor ZONAB regulates epithelial cell proliferation and cell density. *Journal of Cell Biology*, **160**(3), 423–432.

Baleja, J. D., Cherry, J. J., Liu, Z., … Androphy, E. J. (2006). Identification of inhibitors to papillomavirus type 16 E6 protein based on three-dimensional structures of interacting proteins. *Antiviral Research*, **72**(1), 49–59.

Barker, N. (2008). The Canonical Wnt/β-Catenin Signalling Pathway. *Methods in Molecular Biology*, **468**, 5–15.

Barmchi, M. P., Samarasekera, G., Gilbert, M., Auld, V. J., & Zhang, B. (2016). Magi Is Associated with the Par Complex and Functions Antagonistically with Bazooka to Regulate the Apical Polarity Complex. *PLOS ONE*, **11**(4), e0153259.

Bastide, P., Darido, C., Pannequin, J., … Jay, P. (2007). Sox9 regulates cell proliferation and is required for Paneth cell differentiation in the intestinal epithelium. *Journal of Cell Biology*, **178**(4), 635–648.

Basu, P., Taghavi, K., Hu, S. Y., Mogri, S., & Joshi, S. (2018). Management of cervical premalignant lesions. *Current Problems in Cancer*, **42**(2), 129–136.

Bates, S., Phillips, A. C., Clark, P. A., … Vousden, K. H. (1998). p14ARF links the tumour suppressors RB and p53. *Nature 1998 395:6698*, **395**(6698), 124–125.

Batman, G., Oliver, A. W., Zehbe, I., Richard, C., Hampson, L., & Hampson, I. N. (2010). Lopinavir Up-Regulates Expression of the Antiviral Protein Ribonuclease L in Human Papillomavirus-Positive Cervical Carcinoma Cells: *Https://Doi.Org/10.3851/IMP1786*, **16**(4), 515–525.

Batman, G., Oliver, A. W., Zehbe, I., Richard, C., Hampson, L., & Hampson, I. N. (2011). Lopinavir up-regulates expression of the antiviral protein ribonuclease L in human papillomavirus-positive cervical carcinoma cells. *Antiviral Therapy*, **16**(4), 515–525.

Beaton, H., Andrews, D., Parsons, M., … Crean, J. (2016). Wnt6 regulates epithelial cell differentiation and is dysregulated in renal fibrosis. *Https://Doi.Org/10.1152/Ajprenal.00136.2016*, **311**(1), F35–F45.

Bergvall, M., Melendy, T., & Archambault, J. (2013). The E1 proteins. *Virology*, **445**(1–2), 35–56.

Bernard, H.-U., Calleja-Macias, I. E., & Dunn, S. T. (2006). Genome variation of human papillomavirus types: Phylogenetic and medical implications. *International Journal of Cancer*, **118**(5), 1071–1076.

Beronja, S., Janki, P., Heller, E., … Fuchs, E. (2013). RNAi screens in mice identify physiological regulators of oncogenic growth. *Nature*, **501**(7466), 185–190.

Bilder, D., & Perrimon, N. (2000). Localization of apical epithelial determinants by the basolateral PDZ protein Scribble. *Nature 2000 403:6770*, **403**(6770), 676–680.

Blanpain, C., Lowry, W. E., Pasolli, H. A., & Fuchs, E. (2006). Canonical notch signaling functions as a commitment switch in the epidermal lineage. *Genes & Development*, **20**(21), 3022–35.

Bocca, C., Bozzo, F., Bassignana, A., & Miglietta, A. (2011). Antiproliferative effects of COX-2 inhibitor celecoxib on human breast cancer cell lines. *Molecular and Cellular Biochemistry*, **350**(1–2), 59–70.

Boldogh, I., Albrecht, T., & Porter, D. D. (1996). Persistent Viral Infections. *Medical Microbiology*. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK8538/

Boldt, K., van Reeuwijk, J., Lu, Q., … Roepman, R. (2016). An organelle-specific protein landscape identifies novel diseases and molecular mechanisms. *Nature Communications 2016 7:1*, **7**(1), 1–13.

Bondar, T., & Medzhitov, R. (2010). p53-mediated hematopoietic stem and progenitor cell competition. *Cell Stem Cell*, **6**(4), 309.

Bonnet, C., Oh, D., Mei, H., … Deng, S. X. (2021). Wnt6 plays a complex role in maintaining human limbal stem/progenitor cells. *Scientific Reports 2021 11:1*, **11**(1), 1–10.

Bordignon, V., Di Domenico, E. G., Trento, E., … Ensoli, F. (2017). How Human Papillomavirus Replication and Immune Evasion Strategies Take Advantage of the Host DNA Damage Repair Machinery. *Viruses*, **9**(12). doi:10.3390/v9120390

Botchkarev, V. A., & Flores, E. R. (2014). p53/p63/p73 in the epidermis in health and disease. *Cold Spring Harbor Perspectives in Medicine*, **4**(8), a015248.

Boyer, S. N., Wazer, D. E., & Band, V. (1996). E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Research*, **56**(20), 4620–4.

Bravo, I. G., & Alonso, Á. (2004). Mucosal Human Papillomaviruses Encode Four Different E5 Proteins Whose Chemistry and Phylogeny Correlate with Malignant or Benign Growth. *Journal of Virology*, **78**(24), 13613.

Bravo, I. G., & Félez-Sánchez, M. (2015). Papillomaviruses: Viral evolution, cancer and evolutionary medicine. *Evolution, Medicine, and Public Health*, **2015**(1), 32–51.

Bray, S. J. (2016). Notch signalling in context. *Nature Reviews Molecular Cell Biology 2016 17:11*, **17**(11), 722–735.

Brimer, N., Lyons, C., & Vande Pol, S. B. (2007). Association of E6AP (UBE3A) with human papillomavirus type 11 E6 protein. *Virology*, **358**(2), 303–310.

Brimer, N., Lyons, C., Wallberg, A. E., & vande Pol, S. B. (2012). Cutaneous papillomavirus E6 oncoproteins associate with MAML1 to repress transactivation and NOTCH signaling. *Oncogene*, **31**, 4639.

Brinckerhoff, C. E., & Matrisian, L. M. (2002). Matrix metalloproteinases: a tail of a frog that became a prince. *Nature Reviews Molecular Cell Biology 2002 3:3*, **3**(3), 207–214.

Bruggink, S. C., De Koning, M. N. C., Gussekloo, J., … Eekhof, J. A. H. (2012). Cutaneous wart-associated HPV types: Prevalence and relation with patient characteristics. *Journal of Clinical Virology*, **55**, 250–255.

Butts, B. N., Fischer, P. R., & MacK, K. J. (2017). Human Papillomavirus Vaccine and Postural Orthostatic Tachycardia Syndrome: A Review of Current Literature. *Journal of Child Neurology*, **32**(11), 956–965.

Campo, M. S., Graham, S. V., Cortese, M. S., … Man, S. (2010). HPV-16 E5 down-regulates expression of surface HLA class I and reduces recognition by CD8 T cells. *Virology*, **407**(1), 137–142.

Carifi, M., Napolitano, D., Morandi, M., & Dall'Olio, D. (2015). Recurrent respiratory papillomatosis: current and future perspectives. *Therapeutics and Clinical Risk Management*, **11**, 731.

Castanon, A., Landy, R., Brocklehurst, P., … Raghavan, R. (2014). Risk of preterm delivery with increasing depth of excision for cervical intraepithelial neoplasia in England: nested case-control study. *BMJ (Clinical Research Ed.)*, **349**. doi:10.1136/BMJ.G6223

Celegato, M., Messa, L., Goracci, L., … Loregian, A. (2020). A novel small-molecule inhibitor of the human papillomavirus E6-p53 interaction that reactivates p53 function and blocks cancer cells growth. *Cancer Letters*, **470**, 115–125.

Chan, S. Y., Chew, S. H., Egawa, K., … Bernard, H. U. (1997). Phylogenetic Analysis of the Human Papillomavirus Type 2 (HPV-2), HPV-27, and HPV-57 Group, Which Is Associated with Common Warts. *Virology*, **239**(2), 296–302.

Chaturvedi, A. K. (2010). Beyond Cervical Cancer: Burden of Other HPV-Related Cancers Among Men and Women. *Journal of Adolescent Health*, **46**(4), S20–S26.

Chen, X., Loo, J. X., Shi, X., … Li, F. (2018). E6 protein expressed by high-risk hpv activates super-enhancers of the EGFR and c-met oncogenes by destabilizing the histone demethylase KDM5C. *Cancer Research*, **78**(6), 1418–1430.

Cheng, L., Wang, Y., & Du, J. (2020). Human Papillomavirus Vaccines: An Updated Review. *Vaccines*, **8**(3), 1–15.

Cheng, S., Schmidt-Grimminger, D. C., Murant, T., Broker, T. R., & Chow, L. T. (1995). Differentiation-dependent up-regulation of the human papillomavirus E7 gene reactivates cellular DNA replication in suprabasal differentiated keratinocytes. *Genes & Development*, **9**(19), 2335–2349.

Cheng, T. M., Chin, Y. T., Ho, Y., … Davis, P. J. (2018). Resveratrol induces sumoylated COX-2-dependent anti-proliferation in human prostate cancer LNCaP cells. *Food and Chemical Toxicology*, **112**, 67–75.

Cherry, J. J., Rietz, A., Malinkevich, A., … Androphy, E. J. (2013). Structure based identification and characterization of flavonoids that disrupt human papillomavirus-16 E6 function. *PLoS ONE*, **8**(12), e84506.

Choi, K., Ahn, Y.-H., Gibbons, D. L., … Kurie, J. M. (2009). Distinct Biological Roles for the Notch Ligands Jagged-1 and Jagged-2 *. *Journal of Biological Chemistry*, **284**(26), 17766–17774.

Chow, L. T., & Broker, T. R. (1997). In vitro experimental systems for HPV: Epithelial raft cultures for investigations of viral reproduction and pathogenesis and for genetic analyses of viral proteins and regulatory sequences. *Clinics in Dermatology*, **15**(2), 217–227.

Claudy, A. L., Touraine, J. L., & Mitanne, D. (1982). Epidermodysplasia verruciformis induced by a new human papillomavirus (HPV-8). *Archives of Dermatological Research*, **274**(3–4), 213–219.

Clevers, H., & Nusse, R. (2012). Wnt/β-Catenin Signaling and Disease. *Cell*, **149**(6), 1192–1205.

Coles, V. A. H., Chapman, R., Lanitis, T., & Carroll, S. M. (2016). The costs of managing genital warts in the UK by devolved nation: England, Scotland, Wales and Northern Ireland. *International Journal of STD and AIDS*, **27**(1), 51–57.

Collins, A. S., Nakahara, T., Do, A., & Lambert, P. F. (2005). Interactions with Pocket Proteins Contribute to the Role of Human Papillomavirus Type 16 E7 in the Papillomavirus Life Cycle. *Journal of Virology*, **79**(23), 14769–14780.

Combita, A. L., TouzÃ©, A., Bousarghin, L., Sizaret, P.-Y., MuÃ±oz, N., & Coursaget, P. (2001). Gene transfer using human papillomavirus pseudovirions varies according to virus genotype and requires cell surface heparan sulfate. *FEMS Microbiology Letters*, **204**(1), 183–188.

Croxtall, J. D., & Perry, C. M. (2010). LopinavirRitonavir: A review of its use in the management of HIV-1 infection. *Drugs*, **70**(14), 1885–1915.

Day, P. M., Thompson, C. D., Schowalter, R. M., Lowy, D. R., & Schiller, J. T. (2013). Identification of a role for the trans-Golgi network in human papillomavirus 16 pseudovirus infection. *Journal of Virology*, **87**(7), 3862–70.

de Clercq, E. (2009). Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV. *International Journal of Antimicrobial Agents*, **33**(4), 307–320.

De La Cruz-Hernández, E., Pérez-Cárdenas, E., Contreras-Paredes, A., … Dueñas-González, A. (2007). The effects of DNA methylation and histone deacetylase inhibitors on human papillomavirus early gene expression in cervical cancer, an in vitro and clinical study. *Virology Journal*, **4**. doi:10.1186/1743-422X-4-18

de Villiers, E.-M., Fauquet, C., Broker, T. R., Bernard, H.-U., & zur Hausen, H. (2004). Classification of papillomaviruses. *Virology*, **324**(1), 17–27.

Deau, M. C., Favre, M., & Orth, G. (1991). Genetic heterogeneity among human papillomaviruses (HPV) associated with epidermodysplasia verruciformis: evidence for multiple allelic forms of HPV5 and HPV8 E6 genes. *Virology*, **184**(2), 492–503.

Dejosez, M., Ura, H., Brandt, V. L., & Zwaka, T. P. (2013). Safeguards for Cell Cooperation in Mouse Embryogenesis Shown by Genome-Wide Cheater Screen. *Science*, **341**(6153), 1511–1514.

Demers, G. W., Halbert, C. L., & Galloway, D. A. (1994). Elevated Wild-Type p53 Protein Levels in Human Epithelial Cell Lines Immortalized by the Human Papillomavirus Type 16 E7 Gene. *Virology*, **198**(1), 169–174.

DiGiuseppe, S., Luszczek, W., Keiffer, T. R., Bienkowska-Haba, M., Guion, L. G. M., & Sapp, M. J. (2016). Incoming human papillomavirus type 16 genome resides in a vesicular compartment throughout mitosis. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(22), 6289–94.

Dobrosotskaya, I. Y., & James, G. L. (2000). MAGI-1 Interacts with β-Catenin and Is Associated with Cell–Cell Adhesion Structures. *Biochemical and Biophysical Research Communications*, **270**(3), 903–909.

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

Doorbar, J. (2006). Molecular biology of human papillomavirus infection and cervical cancer. *Clinical Science*, **110**(5), 525–541.

Doorbar, J. (2013). Latent papillomavirus infections and their regulation. *Current Opinion in Virology*, **3**. doi:10.1016/j.coviro.2013.06.003

Doorbar, J., Egawa, N., Griffin, H., Kranjec, C., & Murakami, I. (2015). Human papillomavirus molecular biology and disease association. *Reviews in Medical Virology*, **25**, 2–23.

Doorbar, J., & Griffin, H. (2007). Intrabody strategies for the treatment of human papillomavirus-associated disease. *Expert Opinion on Biological Therapy*, **7**(5), 677–689.

Doorbar, J., Quint, W., Banks, L., … Stanley, M. A. (2012). The Biology and Life-Cycle of Human Papillomaviruses. *Vaccine*, **30**, F55–F70.

Doorbar, J., Zheng, K., Aiyenuro, A., … Griffin, H. M. (2021). Principles of epithelial homeostasis control during persistent human papillomavirus infection and its deregulation at the cervical transformation zone. *Current Opinion in Virology*, **51**, 96–105.

Dorji, T., Nopsopon, T., Tamang, S. T., & Pongpirul, K. (2021). Human papillomavirus vaccination uptake in low-and middle-income countries: a meta-analysis. *EClinicalMedicine*, **34**. doi:10.1016/J.ECLINM.2021.100836/ATTACHMENT/83D3842F-04E6-48DB-A60C-0C0385AE78FE/MMC2.PDF

Dow, L. E., Brumby, A. M., Muratore, R., … Humbert, P. O. (2003). hScrib is a functional homologue of the Drosophila tumour suppressor Scribble. *Oncogene 2003 22:58*, **22**(58), 9225–9230.

Dreer, M., van de Poel, S., & Stubenrauch, F. (2017). Control of viral replication and transcription by the papillomavirus E8^E2 protein. *Virus Research*, **231**, 96–102.

Drews, C. M., Case, S., & Pol, S. B. V. (2019). E6 proteins from high-risk HPV, low-risk HPV, and animal papillomaviruses activate the Wnt/ß-catenin pathway through E6AP-dependent degradation of NHERF1. *PLoS Pathogens*, **15**(4), e1007575.

Duffy, C. L., Phillips, S. L., & Klingelhutz, A. J. (2003). Microarray analysis identifies differentiation-associated genes regulated by human papillomavirus type 16 E6. *Virology*, **314**(1), 196–205.

Durst, M., Gissmann, L., Ikenberg, H., & zur Hausen, H. (1983). A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proceedings of the National Academy of Sciences*, **80**(12), 3812–3815.

EAGLE, H., & LEVINE, E. M. (1967). Growth Regulatory Effects of Cellular Interaction. *Nature*, **213**(5081), 1102–1106.

Edmondson, R., Broglie, J. J., Adcock, A. F., & Yang, L. (2014). Three-Dimensional Cell Culture Systems and Their Applications in Drug Discovery and Cell-Based Biosensors. *ASSAY and Drug Development Technologies*, **12**(4), 207–218.

Edwards, T. G., Koeller, K. J., Slomczynska, U., … Fisher, C. (2011). HPV episome levels are potently decreased by pyrrole-imidazole polyamides. *Antiviral Research*, **91**(2), 177–186.

Egawa, K., Delius, H., Matsukura, T., Kawashima, M., & de Villiers, E.-M. (1993). Two Novel Types of Human Papillomavirus, HPV 63 and HPV 65: Comparisons of Their Clinical and Histological Features and DNA Sequences to Other HPV Types. *Virology*, **194**(2), 789–799.

EGAWA, K., INABA, Y., YOSHIMURA, K., & ONO, T. (1993). Varied clinical morphology of HPV-1-induced warts, depending on anatomical factors. *British Journal of Dermatology*, **128**(3), 271–276.

Egawa, N., & Doorbar, J. (2017). The low-risk papillomaviruses. *Virus Research*, **231**, 119–127.

Egawa, N., Egawa, K., Griffin, H., & Doorbar, J. (2015). Human Papillomaviruses; Epithelial Tropisms, and the Development of Neoplasia. *Viruses*, **7**(7), 3863–3890.

Eisenhoffer, G. T., Loftus, P. D., Yoshigi, M., … Rosenblatt, J. (2012a). Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia. *Nature*, **484**(7395), 546–549.

Eisenhoffer, G. T., Loftus, P. D., Yoshigi, M., … Rosenblatt, J. (2012b). Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia. *Nature*, **484**(7395), 546–549.

Eliezri, Y. D., Silverstein, S. J., & Nuovo, G. J. (1990). Occurrence of human papillomavirus type 16 DNA in cutaneous squamous and basal cell neoplasms. *Journal of the American Academy of Dermatology*, **23**(5), 836–842.

Elston, R. C., Doorbar, J., & Napthine, S. (1998). The identification of a conserved binding motif within human papillomavirus type 16 E6 binding peptides, E6AP and E6BP. *Journal of General Virology*, **79**(2), 371–374.

Excoffon, K. J. D. A., Avila, C. L., Alghamri, M. S., & Kolawole, A. O. (2022). The magic of MAGI-1: A scaffolding protein with multi signalosomes and functional plasticity. *Biology of the Cell*, **114**(7), 185–198.

Faucher, A. M., White, P. W., Brochu, C., Grand-Maître, C., Rancourt, J., & Fazal, G. (2004). Discovery of Small-Molecule Inhibitors of the ATPase Activity of Human Papillomavirus E1 Helicase. *Journal of Medicinal Chemistry*, **47**(1), 18–21.

FDA, U. S. F. & D. A. (n.d.). Drug Approval Package: Kaletra (Lopinavir/Ritonavir) NDA #21-226 & 21-251. Retrieved August 18, 2022, from https://www.accessdata.fda.gov/drugsatfda_docs/nda/2000/21-226_Kaletra.cfm

Fellmann, C., & Lowe, S. W. (2014, January 24). Stable RNA interference rules for silencing. *Nature Cell Biology*, Nature Publishing Group, , pp. 10–18.

Finzer, P., Krueger, A., Stöhr, M., … Rösl, F. (2004). HDAC inhibitors trigger apoptosis in HPV-positive cells by inducing the E2F-p73 pathway. *Oncogene*, **23**(28), 4807–4817.

Fischer, A., & Gessler, M. (2007). Delta-Notch--and then? Protein interactions and proposed modes of repression by Hes and Hey bHLH factors. *Nucleic Acids Research*, **35**(14), 4583–96.

Flores, E. R., Allen-Hoffmann, B. L., Lee, D., & Lambert, P. F. (2000). The Human Papillomavirus Type 16 E7 Oncogene Is Required for the Productive Stage of the Viral Life Cycle. *Journal of Virology*, **74**(14), 6622–6631.

Flores, E. R., Allen-Hoffmann, B. L., Lee, D., Sattler, C. A., & Lambert, P. F. (1999). Establishment of the Human Papillomavirus Type 16 (HPV-16) Life Cycle in an Immortalized Human Foreskin Keratinocyte Cell Line. *Virology*, **262**(2), 344–354.

Forman, D., de Martel, C., Lacey, C. J., … Franceschi, S. (2012). Global Burden of Human Papillomavirus and Related Diseases. *Vaccine*, **30**, F12–F23.

Foulongne, V., Sauvage, V., Hebert, C., … Eloit, M. (2012). Human Skin Microbiota: High Diversity of DNA Viruses Identified on the Human Skin by High Throughput Sequencing. *PLoS ONE*, **7**(6), e38499.

Fradet-Turcotte, A., & Archambault, J. (2007). Recent advances in the search for antiviral agents against human papillomaviruses. *Antiviral Therapy*, , pp. 431–451.

Fradet-Turcotte, A., Morin, G., Lehoux, M., Bullock, P. A., & Archambault, J. (2010). Development of quantitative and high-throughput assays of polyomavirus and papillomavirus DNA replication. *Virology*, **399**(1), 65–76.

Fu, L., van Doorslaer, K., Chen, Z., … Burk, R. D. (2010). Degradation of p53 by Human Alphapapillomavirus E6 Proteins Shows a Stronger Correlation with Phylogeny than Oncogenicity. *PLoS ONE*, **5**(9), 1–8.

Fuchs, E. (2003). Keratins and the Skin. *Http://Dx.Doi.Org/10.1146/Annurev.Cb.11.110195.001011*, **11**(1), 123–154.

Fujii, T., Saito, M., Iwasaki, E., … Aoki, D. (2006). Intratumor injection of small interfering RNA-targeting human papillomavirus 18 E6 and E7 successfully inhibits the growth of cervical cancer. *International Journal of Oncology*, **29**(3), 541–548.

Ganti, K., Broniarczyk, J., Manoubi, W., … Banks, L. (2015). The Human Papillomavirus E6 PDZ Binding Motif: From Life Cycle to Malignancy. *Viruses*, **7**(7), 3530–3551.

Gao, S., Yang, C., Jiang, S., … Wang, H. (2014). Applications of RNA interference high-throughput screening technology in cancer biology and virology. *Protein and Cell*, **5**(11), 805–815.

Gardiol, D., Kühne, C., Glaunsinger, B., Lee, S. S., Javier, R., & Banks, L. (1999). Oncogenic human papillomavirus E6 proteins target the discs large tumour suppressor for proteasome-mediated degradation. *Oncogene 1999 18:40*, **18**(40), 5487–5496.

Ghosh, A. K., Osswald, H. L., & Prato, G. (2016, June 9). Recent Progress in the Development of HIV-1 Protease Inhibitors for the Treatment of HIV/AIDS. *Journal of Medicinal Chemistry*, American Chemical Society, , pp. 5172–5208.

Glaunsinger, B. A., Lee, S. S., Thomas, M., Banks, L., & Javier, R. (2000). Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins. *Oncogene 2000 19:46*, **19**(46), 5270–5280.

Glittenberg, M., Pitsouli, C., Garvey, C., Delidakis, C., & Bray, S. (2006). Role of conserved intracellular motifs in Serrate signalling, cis-inhibition and endocytosis. *The EMBO Journal*, **25**(20), 4697.

Goldenberg, M. M. (1999). Celecoxib, a selective cyclooxygenase-2 inhibitor for the treatment of rheumatoid arthritis and osteoarthritis. *Clinical Therapeutics*, **21**(9), 1497–1513.

Goon, P., Sonnex, C., Jani, P., Stanley, M., & Sudhoff, H. (2007). Recurrent respiratory papillomatosis: an overview of current thinking and treatment. *European Archives of Oto-Rhino-Laryngology 2007 265:2*, **265**(2), 147–151.

Gottschling, M., Goker, M., Stamatakis, A., Bininda-Emonds, O. R. P., Nindl, I., & Bravo, I. G. (2011). Quantifying the Phylodynamic Forces Driving Papillomavirus Evolution. *Molecular Biology and Evolution*, **28**(7), 2101–2113.

Grace, M., & Munger, K. (2017). Proteomic analysis of the gamma human papillomavirus type 197 E6 and E7 associated cellular proteins. *Virology*, **500**, 71–81.

Graham, S. V. (2017a). The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. *Clinical Science*, **131**(17), 2201–2221.

Graham, S. V. (2017b). Keratinocyte Differentiation-Dependent Human Papillomavirus Gene Regulation. *Viruses*, **9**(9). doi:10.3390/v9090245

Graham, S. V. (2017c). The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. *Clinical Science (London, England : 1979)*, **131**(17), 2201–2221.

Gravitt, P. E., & Winer, R. L. (2017). Natural History of HPV Infection across the Lifespan: Role of Viral Latency. *Viruses 2017, Vol. 9, Page 267*, **9**(10), 267.

Greenhough, A., Smartt, H. J. M., Moore, A. E., … Kaidi, A. (2009). The COX-2/PGE 2 pathway: key roles in the hallmarks of cancer and adaptation to the tumour microenvironment. *Carcinogenesis*, **30**(3), 377–386.

Griffin, H., Elston, R., Jackson, D., … Doorbar, J. (2006). Inhibition of papillomavirus protein function in cervical cancer cells by intrabody targeting. *Journal of Molecular Biology*, **355**(3), 360–378.

Gudipaty, S. A., & Rosenblatt, J. (2017). Epithelial cell extrusion: Pathways and pathologies. *Seminars in Cell & Developmental Biology*, **67**, 132–140.

Gulbahce, N., Yan, H., Dricot, A., … Barabási, A.-L. (2012). Viral Perturbations of Host Networks Reflect Disease Etiology. *PLoS Computational Biology*, **8**(6), e1002531.

Guo, A. (1999). *Huangdi Neijing Suwen Jiao Zhu Yu Yi (Yellow Emperor's Inner Classic: Plain Questions – Critically Compared, Annotated and Translated)*, Tianjin, China: Tianjin Kexue Jishu Chubanshe (Tianjin Science and Technology Press).

H. Adler, D. (2010). The Impact of HAART on HPV-Related Cervical Disease. *Current HIV Research*, **8**(7), 493–497.

Hagiwara, K., Uezato, H., Arakaki, H., … Hattori, A. (2005). A genotype distribution of human papillomaviruses detected by polymerase chain reaction and direct sequencing analysis in a large sample of common warts in Japan. *Journal of Medical Virology*, **77**(1), 107–112.

Hamid, N. A., Brown, C., & Gaston, K. (2009). The regulation of cell proliferation by the papillomavirus early proteins. *Cellular and Molecular Life Sciences 2009 66:10*, **66**(10), 1700–1717.

Hampson, L., Kitchener, H. C., & Hampson, I. N. (2006). Specific HIV protease inhibitors inhibit the ability of HPV16 E6 to degrade p53 and selectively kill E6-dependent cervical carcinoma cells in vitro. *Antiviral Therapy*, **11**(6), 813–825.

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, **100**(1), 57–70.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. *Cell*, **144**(5), 646–674.

Handa, K., Yugawa, T., Narisawa-Saito, M., Ohno, S., Fujita, M., & Kiyono, T. (2007). E6AP-Dependent Degradation of DLG4/PSD95 by High-Risk Human Papillomavirus Type 18 E6 Protein. *Journal of Virology*, **81**(3), 1379–1389.

Hao, Y., Hao, S., Andersen-Nissen, E., … Satija, R. (2021). Integrated analysis of multimodal single-cell data. *Cell*, **184**(13), 3573-3587.e29.

Harari, A., Chen, Z., & Burk, R. D. (2014). Human papillomavirus genomics: past, present and future. *Current Problems in Dermatology*, **45**, 1–18.

Hausen, H. zur. (1991). Viruses in Human Cancers. *Science*, **254**(5035), 1167–1173.

Hawley-Nelson, P., Vousden, K. H. H., Hubbert, N. L. L., Lowy, D. R. R., & Schiller, J. T. T. (1989a). HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes., **8**(12). doi:10.1002/j.1460-2075.1989.tb08570.x

Hawley-Nelson, P., Vousden, K. H., Hubbert, N. L., Lowy, D. R., & Schiller, J. T. (1989b). HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes. *The EMBO Journal*, **8**(12), 3905.

He, C., Mao, D., Hua, G., … Wang, C. (2015). The Hippo/YAP pathway interacts with EGFR signaling and HPV oncoproteins to regulate cervical cancer progression. *EMBO Molecular Medicine*, **7**(11), 1426–1449.

He, G., Vasilieva, E., Harris, G. D., Koeller, K. J., Bashkin, J. K., & Dupureur, C. M. (2014). Binding studies of a large antiviral polyamide to a natural HPV sequence. *Biochimie*, **102**(1), 83–91.

Hebner, C. M., & Laimins, L. A. (2006, March). Human papillomaviruses: Basic mechanisms of pathogenesis and oncogenicity. *Reviews in Medical Virology*, , pp. 83–97.

Heinzel, P. A., Chan, S. Y., Ho, L., … Bernard, H. U. (1995). Variation of human papillomavirus type 6 (HPV-6) and HPV-11 genomes sampled throughout the world. *Journal of Clinical Microbiology*, **33**(7), 1746–1754.

Hengstermann, A., Linares, L. K., Ciechanover, A., Whitaker, N. J., & Scheffner, M. (2001). Complete switch from Mdm2 to human papillomavirus E6-mediated degradation of p53 in cervical cancer cells. *Proceedings of the National Academy of Sciences*, **98**(3), 1218–1223.

Herber, R., Liem, A., Pitot, H., & Lambert, P. F. (1996). Squamous epithelial hyperplasia and carcinoma in mice transgenic for the human papillomavirus type 16 E7 oncogene. *Journal of Virology*, **70**(3), 1873.

Hirata, H., Tomita, K., Bessho, Y., & Kageyama, R. (2001). Hes1 and Hes3 regulate maintenance of the isthmic organizer and development of the mid/hindbrain. *The EMBO Journal*, **20**(16), 4454–4466.

Hoffman, S. R., Le, T., Lockhart, A., … Smith, J. S. (2017). Patterns of persistent HPV infection after treatment for cervical intraepithelial neoplasia (CIN): A systematic review. *International Journal of Cancer*, **141**(1), 8–23.

Honda, R., & Yasuda, H. (2000). Activity of MDM2, a ubiquitin ligase, toward p53 or itself is dependent on the RING finger domain of the ligase. *Oncogene 2000 19:11*, **19**(11), 1473–1476.

Hougardy, B. M. T., Maduro, J. H., van der Zee, A. G. J., … de Jong, S. (2006). Proteasome inhibitor MG132 sensitizes HPV-positive human cervical cancer cells to rhTRAIL-induced apoptosis. *International Journal of Cancer*, **118**(8), 1892–1900.

Howie, H. L., Koop, J. I., Weese, J., … Galloway, D. A. (2011). Beta-HPV 5 and 8 E6 Promote p300 Degradation by Blocking AKT/p300 Association. *PLoS Pathogens*, **7**(8), e1002211.

Howley, P. M., & Pfister, H. J. (2015). Beta genus papillomaviruses and skin cancer. *Virology*, **479–480**, 290–296.

HPV and Cancer - National Cancer Institute. (2019). *NIH*.

Ii, M., Yamamoto, H., Adachi, Y., Maruyama, Y., & Shinomura, Y. (2006). Role of Matrix Metalloproteinase-7 (Matrilysin) in Human Cancer Invasion, Apoptosis, Growth, and Angiogenesis: *Https://Doi.Org/10.1177/153537020623100103*, **231**(1), 20–27.

Iversen, P. W., Eastwood, B. J., Sittampalam, G. S., & Cox, K. L. (2006). A comparison of assay performance measures in screening assays: Signal window, Z' factor, and assay variability ratio. *Journal of Biomolecular Screening*, **11**(3), 247–252.

Jablonska, S., Orth, G., Obalek, S., & Croissant, O. (1985). Cutaneous warts clinical, histologic, and virologic correlations. *Clinics in Dermatology*, **3**(4), 71–82.

James, C. D., & Roberts, S. (2016). Viral Interactions with PDZ Domain-Containing Proteins—An Oncogenic Trait? *Pathogens 2016, Vol. 5, Page 8*, **5**(1), 8.

Jamora, C., & Fuchs, E. (2002). Intercellular adhesion, signalling and the cytoskeleton. *Nature Cell Biology 2002 4:4*, **4**(4), E101–E108.

Jansen, E. E. L., Zielonke, N., Gini, A., … Priaulx, J. (2020). Effect of organised cervical cancer screening on cervical cancer mortality in Europe: a systematic review. *European Journal of Cancer*, **127**, 207–223.

Javier, R. T., & Rice, A. P. (2011). Emerging Theme: Cellular PDZ Proteins as Common Targets of Pathogenic Viruses. *Journal of Virology*, **85**(22), 11544–11556.

Jiang, D., Srinivasan, A., Lozano, G., & Robbins, P. (1993). SV40 T antigen abrogates p53-mediated transcriptional activity. *Oncogene*, **8**(10), 2805–2812.

Jiang, R., Sun, T., Song, D., & Li, J. J. (2022). Statistics or biology: the zero-inflation controversy about scRNA-seq data. *Genome Biology 2022 23:1*, **23**(1), 1–24.

Jin, S., Guerrero-Juarez, C. F., Zhang, L., … Nie, Q. (2021). Inference and analysis of cell-cell communication using CellChat. *Nature Communications 2021 12:1*, **12**(1), 1–20.

Jones, D. L., Thompson, D. A., & Münger, K. (1997a). Destabilization of the RB Tumor Suppressor Protein and Stabilization of p53 Contribute to HPV Type 16 E7-Induced Apoptosis. *Virology*, **239**(1), 97–107.

Jones, D. L., Thompson, D. A., & Münger, K. (1997b). Destabilization of the RB Tumor Suppressor Protein and Stabilization of p53 Contribute to HPV Type 16 E7-Induced Apoptosis. *Virology*, **239**(1), 97–107.

Jonson, A. L., Rogers, L. M., Ramakrishnan, S., & Downs, L. S. (2008). Gene silencing with siRNA targeting E6/E7 as a therapeutic intervention in a mouse model of cervical cancer. *Gynecologic Oncology*, **111**(2), 356–364.

JS, R., MW, S., JH, B., … D, A. (2017). A stratified randomized double-blind phase II trial of celecoxib for treating patients with cervical intraepithelial neoplasia: The potential predictive value of VEGF serum levels: An NRG Oncology/Gynecologic Oncology Group study. *Gynecologic Oncology*, **145**(2), 291–297.

Jun-Feng, X., & Ya, J. (2010). A mathematical model of a P53 oscillation network triggered by DNA damage. *Chinese Physics B*, **19**(4), 040506.

Kalluri, A. S. (2019). Defining endothelial cell functional heterogeneity and plasticity using single-cell RNA-sequencing. Retrieved from https://dash.harvard.edu/handle/1/42013060

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, **45**(D1), D353–D361.

Kantang, W., Chunsrivirot, S., Muangsin, N., Poovorawan, Y., & Krusong, K. (2016). Design of peptides as inhibitors of human papillomavirus 16 transcriptional regulator E1-E2. *Chemical Biology & Drug Design*, **88**(4), 475–484.

Karlsson, O. A., Ramirez, J., Öberg, D., … Jemth, P. (2015). Design of a PDZbody, a bivalent binder of the E6 protein from human papillomavirus. *Scientific Reports*, **5**(1), 1–7.

Katoh, M., & Katoh, M. (2007). Integrative genomic analyses on HES/HEY family: Notch-independent HES1, HES3 transcription in undifferentiated ES cells, and Notch-dependent HES1, HES5, HEY1, HEY2, HEYL transcription in fetal tissues, adult tissues, or cancer. *International Journal of Oncology*, **31**(2), 461–6.

Kaushik, S., Pepas, L., Nordin, A., Bryant, A., & Dickinson, H. O. (2011). Surgical interventions for high grade vulval intraepithelial neoplasia. *The Cochrane Database of Systematic Reviews*, (1), CD007928.

Kawai, K., Egawa, N., Kiyono, T., & Kanekura, T. (2009). Epidermodysplasia-Verruciformis-Like Eruption Associated with Gamma-Papillomavirus Infection in a Patient with Adult T-Cell Leukemia. *Dermatology*, **219**(3), 274–278.

Khelil, M., Griffin, H., Bleeker, M., … Jordanova, E. (2021). Delta-Like Ligand–Notch1 Signaling Is Selectively Modulated by HPV16 E6 to Promote Squamous Cell Proliferation and Correlates with Cervical Cancer Prognosis. *Cancer Research*, **81**, canres.CAN-20.

Kim, D. H., Jarvis, R. M., Allwood, J. W., … Goodacre, R. (2010). Raman chemical mapping reveals site of action of HIV protease inhibitors in HPV16 E6 expressing cervical carcinoma cells. *Analytical and Bioanalytical Chemistry*, **398**(7–8), 3051–3061.

Kines, R. C., Thompson, C. D., Lowy, D. R., Schiller, J. T., & Day, P. M. (2009). The initial steps leading to papillomavirus infection occur on the basement membrane prior to cell surface binding. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(48), 20458–20463.

Klug, A., & Finch, J. T. (1965). Structure of viruses of the papilloma-polyoma type: I. Human wart virus. *Journal of Molecular Biology*, **11**(2), 403-IN44.

Kopan, R., & Ilagan, Ma. X. G. (2009a). The Canonical Notch Signaling Pathway: Unfolding the Activation Mechanism. *Cell*, **137**(2), 216–233.

Kopan, R., & Ilagan, Ma. X. G. (2009b). The Canonical Notch Signaling Pathway: Unfolding the Activation Mechanism. *Cell*, **137**(2), 216–233.

Koshiol, J., Lindsay, L., Pimenta, J. M., Poole, C., Jenkins, D., & Smith, J. S. (2008). Persistent human papillomavirus infection and cervical neoplasia: a systematic review and meta-analysis. *American Journal of Epidemiology*, **168**(2), 123–37.

Kranjec, C., & Banks, L. (2011). A systematic analysis of human papillomavirus (HPV) E6 PDZ substrates identifies MAGI-1 as a major target of HPV type 16 (HPV-16) and HPV-18 whose loss accompanies disruption of tight junctions. *Journal of Virology*, **85**(4), 1757–64.

Kranjec, C., Holleywood, C., Libert, D., … Doorbar, J. (2017). Modulation of basal cell fate during productive and transforming HPV-16 infection is mediated by progressive E6-driven depletion of Notch. *The Journal of Pathology*, **242**(4), 448–462.

Kuerbitz, S. J., Plunkett, B. S., Walsh, W. v, & Kastan, M. B. (1992). Wild-type p53 is a cell cycle checkpoint determinant following irradiation. *Proceedings of the National Academy of Sciences*, **89**(16), 7491–7495.

Kunitoki, K., Funato, M., Mitsunami, M., Kinoshita, T., & Reich, M. R. (2021). Access to HPV vaccination in Japan: Increasing social trust to regain vaccine confidence. *Vaccine*, **39**(41), 6104–6110.

Kwun, H. J., Jung, E. Y., Ahn, J. Y., Lee, M. N., & Jang, K. L. (2001). p53-dependent transcriptional repression of p21waf1 by hepatitis C virus NS3. *Journal of General Virology*, **82**(9), 2235–2241.

Lahiri, C. D., Dugan, K. B., Xie, X., … Ofotokun, I. (2015). Oral Lopinavir Use and Human Papillomavirus Infection in HIV-positive Women. *Journal of Acquired Immune Deficiency Syndromes (1999)*, **70**(2), e63.

LaMontagne, D. S., Bloem, P. J. N., Brotherton, J. M. L., Gallagher, K. E., Badiane, O., & Ndiaye, C. (2017). Progress in HPV vaccination in low- and lower-middle-income countries. *International Journal of Gynecology & Obstetrics*, **138**, 7–14.

Laprise, P., Viel, A., & Rivard, N. (2004). Human Homolog of Disc-large Is Required for Adherens Junction Assembly and Differentiation of Human Intestinal Epithelial Cells. *Journal of Biological Chemistry*, **279**(11), 10157–10166.

Latorre, I. J., Roh, M. H., Frese, K. K., Weiss, R. S., Margolis, B., & Javier, R. T. (2005). Viral oncoprotein-induced mislocalization of select PDZ proteins disrupts tight junctions and causes polarity defects in epithelial cells. *Journal of Cell Science*, **118**(18), 4283–4293.

Lechler, T., & Fuchs, E. (2005). Asymmetric cell divisions promote stratification and differentiation of mammalian skin. *Nature 2005 437:7056*, **437**(7056), 275–280.

Lefort, K., Mandinova, A., Ostano, P., … Dotto, G. P. (2007). Notch1 is a p53 target gene involved in human keratinocyte tumor suppression through negative regulation of ROCK1/2 and MRCK kinases. *Genes & Development*, **21**(5), 562–577.

Letian, T., & Tianyu, Z. (2010). Cellular receptor binding and entry of human papillomavirus. *Virology Journal*, **7**(1), 2.

Li, C. J., Wang, C., Friedman, D. J., & Pardee, A. B. (1995). Reciprocal modulations between p53 and Tat of human immunodeficiency virus type 1. *Proceedings of the National Academy of Sciences of the United States of America*, **92**(12), 5461.

Lin, Z., Bazzaro, M., Wang, M. C., Chan, K. C., Peng, S., & Roden, R. B. S. (2009). Combination of proteasome and HDAC inhibitors for uterine cervical cancer treatment. *Clinical Cancer Research*, **15**(2), 570–577.

Lipovsky, A., Popa, A., Pimienta, G., … DiMaio, D. (2013). Genome-wide siRNA screen identifies the retromer as a cellular entry factor for human papillomavirus. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(18), 7452–7457.

Liu, Y., Liu, Z., Androphy, E., Chen, J., & Baleja, J. D. (2004). Design and characterization of helical peptides that inhibit the E6 protein of papillomavirus. *Biochemistry*, **43**(23), 7421–7431.

Lolo, F.-N., Casas-Tintó, S., & Moreno, E. (2012). Cell Competition Time Line: Winners Kill Losers, which Are Extruded and Engulfed by Hemocytes. *Cell Reports*, **2**(3), 526–539.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology 2014 15:12*, **15**(12), 1–21.

Lowell, S., Jones, P., Le Roux, I., Dunne, J., & Watt, F. M. (2000). Stimulation of human epidermal differentiation by delta-notch signalling at the boundaries of stem-cell clusters. *Current Biology : CB*, **10**(9), 491–500.

Lu, Y., Sun, W., Zhang, L., & Li, J. (2019). Silencing Of MAGI1 Promotes The Proliferation And Inhibits Apoptosis Of Glioma Cells Via The Wnt/β-Catenin And PTEN/AKT Signaling Pathways. *OncoTargets and Therapy*, **12**, 9639.

Luna, A. J., Sterk, R. T., Griego-Fisher, A. M., … Ozbun, M. A. (2021). MEK/ERK signaling is a critical regulator of high-risk human papillomavirus oncogene expression revealing therapeutic targets for HPV-induced tumors. *PLOS Pathogens*, **17**(1), e1009216.

Macara, I. G., Guyer, R., Richardson, G., Huo, Y., & Ahmed, S. M. (2014a). Epithelial homeostasis. *Current Biology*. doi:10.1016/j.cub.2014.06.068

Macara, I. G., Guyer, R., Richardson, G., Huo, Y., & Ahmed, S. M. (2014b, September 8). Epithelial homeostasis. *Current Biology*, Cell Press, , pp. R815–R825.

Maglennon, G. A., McIntosh, P., & Doorbar, J. (2011a). Persistence of viral DNA in the epithelial basal layer suggests a model for papillomavirus latency following immune regression. *Virology*, **414**(2), 153–163.

Maglennon, G. A., McIntosh, P., & Doorbar, J. (2011b). Persistence of viral DNA in the epithelial basal layer suggests a model for papillomavirus latency following immune regression. *Virology*, **414**(2), 153–163.

Malecka, K. A., Fera, D., Schultz, D. C., … Marmorstein, R. (2014). Identification and characterization of small molecule human papillomavirus E6 inhibitors. *ACS Chemical Biology*, **9**(7), 1603–1612.

Mantovani, F., & Banks, L. (2001). The Human Papillomavirus E6 protein and its contribution to malignant progression. *Oncogene 2001 20:54*, **20**(54), 7874–7887.

Marchiando, A. M., Graham, W. V., & Turner, J. R. (2010). Epithelial Barriers in Homeostasis and Disease. *Annual Review of Pathology: Mechanisms of Disease*, **5**(1), 119–144.

Marques-Pereira, J. P., & Leblond, C. P. (1965). Mitosis and differentiation in the stratified squamous epithelium of the rat esophagus. *American Journal of Anatomy*, **117**(1), 73–87.

Martinez-Zapien, D., Ruiz, F. X., Poirson, J., … Zanier, K. (2016). Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature*, **529**(7587), 541–545.

Matamoro-Vidal, A., & Levayer, R. (2019). Multiple Influences of Mechanical Forces on Cell Competition. *Current Biology*, **29**(15), R762–R774.

Matter, K., Aijaz, S., Tsapara, A., & Balda, M. S. (2005). Mammalian tight junctions in the regulation of epithelial differentiation and proliferation. *Current Opinion in Cell Biology*, Elsevier Ltd, , pp. 453–458.

Matthews, K., Leong, C. M., Baxter, L., … Hibma, M. (2003). Depletion of Langerhans Cells in Human Papillomavirus Type 16-Infected Skin Is Associated with E6-Mediated Down Regulation of E-Cadherin. *Journal of Virology*, **77**(15), 8378–8385.

Maw, R. (2004). Critical appraisal of commonly used treatment for genital warts. *International Journal of STD & AIDS*, **15**(6), 357–64.

Mbonye, U., & Karn, J. (2017). The Molecular Basis for Human Immunodeficiency Virus Latency. *Https://Doi.Org/10.1146/Annurev-Virology-101416-041646*, **4**, 261–285.

McBride, A. A. (2008). Chapter 4 Replication and Partitioning of Papillomavirus Genomes. *Advances in Virus Research*, **72**, 155–205.

McBride, A. A. (2013). The papillomavirus E2 proteins. *Virology*, **445**(1–2), 57–79.

McBride, A. A. (2017). Mechanisms and strategies of papillomavirus replication. *Biological Chemistry*, **398**(8), 919–927.

McCaffrey, L. M., & Macara, I. G. (2009). The Par3/aPKC interaction is essential for end bud remodeling and progenitor differentiation during mammary gland morphogenesis. *Genes & Development*, **23**(12), 1450–1460.

McGuire, J. K., Li, Q., & Parks, W. C. (2003). Matrilysin (Matrix Metalloproteinase-7) Mediates E-Cadherin Ectodomain Shedding in Injured Lung Epithelium. *The American Journal of Pathology*, **162**(6), 1831–1843.

McLaughlin-Drubin, M. E., Bromberg-White, J. L., & Meyers, C. (2005). The role of the human papillomavirus type 18 E7 oncoprotein during the complete viral life cycle. *Virology*, **338**(1), 61–68.

McLaughlin-Drubin, M. E., & Münger, K. (2009). The human papillomavirus E7 oncoprotein. *Virology*, **384**(2), 335–344.

Meyers, J. M., Spangle, J. M., & Munger, K. (2013). The Human Papillomavirus Type 8 E6 Protein Interferes with NOTCH Activation during Keratinocyte Differentiation. *Journal of Virology*, **87**(8), 4762–4767.

Meyers, J. M., Uberoi, A., Grace, M., Lambert, P. F., & Munger, K. (2017). Cutaneous HPV8 and MmuPV1 E6 Proteins Target the NOTCH and TGF-β Tumor Suppressors to Inhibit Differentiation and Sustain Keratinocyte Proliferation. *PLOS Pathogens*, **13**(1), e1006171.

Minkoff, H., Ahdieh, L., Massad, L. S., … Palefsky, J. (2001). The effect of highly active antiretroviral therapy on cervical cytologic changes associated with oncogenic HPV among HIV-infected women. *AIDS (London, England)*, **15**(16), 2157–2164.

Minkoff, H., Zhong, Y., Burk, R. D., … Strickler, H. D. (2010). Influence of adherent and effective antiretroviral therapy use on human papillomavirus infection and squamous intraepithelial lesions in human immunodeficiency virus-positive women. *The Journal of Infectious Diseases*, **201**(5), 681–690.

Mizuhara, E., Nakatani, T., Minaki, Y., Sakamoto, Y., Ono, Y., & Takai, Y. (2005). MAGI1 Recruits Dll1 to Cadherin-based Adherens Junctions and Stabilizes It on the Cell Surface *. *Journal of Biological Chemistry*, **280**(28), 26499–26507.

Moody, C. (2017a). Mechanisms by which HPV Induces a Replication Competent Environment in Differentiating Keratinocytes. *Viruses*, **9**(9), 261.

Moody, C. A. (2017b). Mechanisms by which HPV Induces a Replication Competent Environment in Differentiating Keratinocytes. *Viruses 2017, Vol. 9, Page 261*, **9**(9), 261.

Moody, C. A., & Laimins, L. A. (2009). Human Papillomaviruses Activate the ATM DNA Damage Pathway for Viral Genome Amplification upon Differentiation. *PLoS Pathogens*, **5**(10), e1000605.

Moody, C. A., & Laimins, L. A. (2010). Human papillomavirus oncoproteins: pathways to transformation. *Nature Reviews Cancer*, **10**(8), 550–560.

Müller, M., Gissmann, L., Cristiano, R. J., … Zhou, J. (1995). Papillomavirus capsid binding and uptake by cells from different tissues and species. *Journal of Virology*, **69**(2), 948–54.

Mulloy, J. C., Kislyakova, T., Cereseto, A., … Franchini, G. (1998). Human T-Cell Lymphotropic/Leukemia Virus Type 1 Tax Abrogates p53-Induced Cell Cycle Arrest and Apoptosis through Its CREB/ATF Functional Domain. *Journal of Virology*, **72**(11), 8852–8860.

Münger, K., Basile, J. R., Duensing, S., … Zacny, V. L. (2001). Biological activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene 2001 20:54*, **20**(54), 7888–7898.

Münger, K., Phelps, W. C., Bubb, V., Howley, P. M., & Schlegel, R. (1989a). The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. *Journal of Virology*, **63**(10), 4417.

Münger, K., Werness, B. A., Dyson, N., Phelps, W. C., Harlow, E., & Howley, P. M. (1989b). Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. *The EMBO Journal*, **8**(13), 4099–4105.

Murakami, I., Egawa, N., Griffin, H., … Doorbar, J. (2019). Roles for E1-independent replication and E6-mediated p53 degradation during low-risk and high-risk human papillomavirus genome maintenance. *PLOS Pathogens*, **15**(5), e1007755.

Nagata, R., & Igaki, T. (2018). Cell competition: Emerging mechanisms to eliminate neighbors. *Development, Growth & Differentiation*, **60**(9), 522–530.

Nakagawa, S., & Huibregtse, J. M. (2000). Human Scribble (Vartul) Is Targeted for Ubiquitin-Mediated Degradation by the High-Risk Papillomavirus E6 Proteins and the E6AP Ubiquitin-Protein Ligase. *Molecular and Cellular Biology*, **20**(21), 8244–8253.

Nandagopal, N., Santat, L. A., LeBon, L., Sprinzak, D., Bronner, M. E., & Elowitz, M. B. (2018). Dynamic Ligand Discrimination in the Notch Signaling Pathway. *Cell*, **172**(4), 869-880.e19.

Nelson, A. R., Fingleton, B., Rothenberg, M. L., & Matrisian, L. M. (2000). Matrix Metalloproteinases: Biologic Activity and Clinical Implications. *Https://Doi.Org/10.1200/JCO.2000.18.5.1135*, **18**(5), 1135–1149.

Nguyen, B.-C., Lefort, K., Mandinova, A., … Dotto, G. P. (2006). Cross-regulation between Notch and p63 in keratinocyte commitment to differentiation. *Genes & Development*, **20**(8), 1028–42.

NHS. (2008). The Human papillomavirus vaccine: The virus, the diseases and the new HPV vaccine. *NHS*.

Nicoll, M. P., Proença, J. T., & Efstathiou, S. (2012). The molecular basis of herpes simplex virus latency. *FEMS Microbiology Reviews*, **36**(3), 684–705.

Nishi, E., Higashiyama, S., & Klagsbrun, M. (2003). Heparin-Binding Epidermal Growth Factor-like Growth Factor (HB-EGF). *Encyclopedia of Hormones*, 235–241.

Niu, X. Y., Peng, Z. L., Duan, W. Q., Wang, H., & Wang, P. (2006). Inhibition of HPV 16 E6 oncogene expression by RNA interference in vitro and in vivo. *International Journal of Gynecological Cancer*, **16**(2), 743–751.

Norman, M., Wisniewska, K. A., Lawrenson, K., … Fujita, Y. (2012). Loss of Scribble causes cell competition in mammalian cells. *Journal of Cell Science*, **125**(1), 59–66.

Odenwald, M. A., Choi, W., Kuo, W. T., … Turner, J. R. (2018). The scaffolding protein ZO-1 coordinates actomyosin and epithelial apical specializations in vitro and in vivo. *Journal of Biological Chemistry*, **293**(45), 17317–17335.

Oh, S. T., Longworth, M. S., & Laimins, L. A. (2004). Roles of the E6 and E7 Proteins in the Life Cycle of Low-Risk Human Papillomavirus Type 11. *Journal of Virology*, **78**(5), 2620–2626.

Ohsawa, S., Vaughen, J., & Igaki, T. (2018). Cell Extrusion: A Stress-Responsive Force for Good or Evil in Epithelial Homeostasis. *Developmental Cell*, **44**(3), 284–296.

Okada, Y. (2013). Proteinases and Matrix Degradation. *Kelley's Textbook of Rheumatology*, 97–115.

Okuyama, R., Tagami, H., & Aiba, S. (2008). Notch signaling: Its role in epidermal homeostasis and in the pathogenesis of skin diseases. *Journal of Dermatological Science*, **49**(3), 187–194.

Orenstein, W. A., Gellin, B. G., Beigi, R. H., … Zettle, M. (2016). Overcoming Barriers to Low HPV Vaccine Uptake in the United States: Recommendations from the National Vaccine Advisory Committee: Approved by the National Vaccine Advisory Committee on June 9, 2015. *Public Health Reports*, **131**(1), 17.

Orr, S. E., Gokulan, K., Boudreau, M., Cerniglia, C. E., & Khare, S. (2019). Alteration in the mRNA expression of genes associated with gastrointestinal permeability and ileal TNF-α secretion due to the exposure of silver nanoparticles in Sprague-Dawley rats. *Journal of Nanobiotechnology*, **17**(1). doi:10.1186/S12951-019-0499-6

Orth, G. (2006). Genetics of epidermodysplasia verruciformis: Insights into host defense against papillomaviruses. *Seminars in Immunology*, **18**(6), 362–374.

Orth, G., Favre, M., & Croissant, O. (1977). Characterization of a New Type of Human Papillomavirus That Causes Skin Warts. *Journal of Virology*, **24**(1), 108.

Ostrow, R. S., Shaver, M. K., Turnquist, S., … Faras, A. J. (1989). Human Papillomavirus-16 DNA in a Cutaneous Invasive Cancer. *Archives of Dermatology*, **125**(5), 666.

Oswald, F., Tauber, B., Dobner, T., … Schmid, R. M. (2001). p300 Acts as a Transcriptional Coactivator for Mammalian Notch-1. *Molecular and Cellular Biology*, **21**(22), 7761–7774.

Packard, A., Georgas, K., Michos, O., … Costantini, F. (2013). Luminal Mitosis Drives Epithelial Cell Dispersal within the Branching Ureteric Bud. *Developmental Cell*, **27**(3), 319–330.

Page-McCaw, A., Ewald, A. J., & Werb, Z. (2007). Matrix metalloproteinases and the regulation of tissue remodelling. *Nature Reviews Molecular Cell Biology 2007 8:3*, **8**(3), 221–233.

Pal, A., & Kundu, R. (2020). Human Papillomavirus E6 and E7: The Cervical Cancer Hallmarks and Targets for Therapy. *Frontiers in Microbiology*, **10**, 3116.

Park, S., Auyeung, A., Lee, D. L., Lambert, P. F., Carchman, E. H., & Sherer, N. M. (2021). HIV-1 Protease Inhibitors Slow HPV16-Driven Cell Proliferation through Targeted Depletion of Viral E6 and E7 Oncoproteins. *Cancers*, **13**(5), 1–17.

Patel, R. v., Yanofsky, V. R., & Goldenberg, G. (2012). Genital Warts: A Comprehensive Review. *The Journal of Clinical and Aesthetic Dermatology*, **5**(6), 25.

Phatak, S. S., Stephan, C. C., & Cavasotto, C. N. (2009, September). High-throughput and in silico screenings in drug discovery. *Expert Opinion on Drug Discovery*, , pp. 947–959.

Phelps, W. C., & Alexander, K. A. (1995, September 1). Antiviral therapy for human papillomaviruses: Rationale and prospects. *Annals of Internal Medicine*, American College of Physicians, , pp. 368–382.

Pollock, K. G. J., Kavanagh, K., Potts, A., … Donaghy, M. (2014). Reduction of low- and high-grade cervical abnormalities associated with high uptake of the HPV bivalent vaccine in Scotland. *British Journal of Cancer 2014 111:9*, **111**(9), 1824–1830.

Poulson, D. F. (1937). Chromosomal Deficiencies and the Embryonic Development of Drosophila Melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, **23**(3), 133–7.

Purvis, J. E., Karhohs, K. W., Mock, C., Batchelor, E., Loewer, A., & Lahav, G. (2012). p53 Dynamics Control Cell Fate. *Science*, **336**(6087), 1440–1444.

Pushpakom, S., Iorio, F., Eyers, P. A., … Pirmohamed, M. (2018). Drug repurposing: progress, challenges and recommendations. *Nature Reviews Drug Discovery 2018 18:1*, **18**(1), 41–58.

Qin, J. Y., Zhang, L., Clift, K. L., Hulur, I., & Xiang, A. P. (2010). Systematic Comparison of Constitutive Promoters and the Doxycycline-Inducible Promoter. *PLoS ONE*, **5**(5), 10611.

Rajakariar, R., Yaqoob, M. M., & Gilroy, D. W. (2006). COX-2 in Inflammation and Resolution. *Molecular Interventions*, **6**(4), 199.

Ramirez, J., Poirson, J., Foltz, C., … Travé, G. (2015). Targeting the Two Oncogenic Functional Sites of the HPV E6 Oncoprotein with a High-Affinity Bivalent Ligand. *Angewandte Chemie International Edition*, **54**(27), 7958–7962.

Rangarajan, A., Talora, C., Okuyama, R., … Dotto, G. P. (2001a). Notch signaling is a direct determinant of keratinocyte growth arrest and entry into differentiation. *The EMBO Journal*, **20**(13), 3427–3436.

Rangarajan, A., Talora, C., Okuyama, R., … Dotto, G. P. P. (2001b). Notch signaling is a direct determinant of keratinocyte growth arrest and entry into differentiation. *The EMBO Journal*, **20**(13), 3427–3436.

Rheinwald, J. G., & Green, H. (1975). Serial cultivation of strains of human epidermal keratinocytes: the formation of keratinizing colonies from single cells. *Cell*, **6**(3), 331–43.

Ribatti, D. (2017). A revisited concept: Contact inhibition of growth. From cell biology to malignancy. *Experimental Cell Research*, **359**(1), 17–19.

Ricci-López, J., Vidal-Limon, A., Zunñiga, M., … Aguila, S. (2019). Molecular modeling simulation studies reveal new potential inhibitors against HPV E6 protein. *PLOS ONE*, **14**(3), e0213028.

Rietz, A., Petrov, D. P., Bartolowits, M., DeSmet, M., Davisson, V. J., & Androphy, E. J. (2016). Molecular Probing of the HPV-16 E6 Protein Alpha Helix Binding Groove with Small Molecule Inhibitors. *PLOS ONE*, **11**(2), e0149845.

Rivas, C., Aaronson, S. A., & Munoz-Fontela, C. (2010). Dual Role of p53 in Innate Antiviral Immunity. *Viruses 2010, Vol. 2, Pages 298-313*, **2**(1), 298–313.

Roberts, J. N., Buck, C. B., Thompson, C. D., … Schiller, J. T. (2007). Genital transmission of HPV in a mouse model is potentiated by nonoxynol-9 and inhibited by carrageenan. *Nature Medicine*, **13**(7), 857–861.

Rohani, M. G., & Parks, W. C. (2015). Matrix remodeling by MMPs during wound repair. *Matrix Biology*, **44–46**, 113–121.

Rosenthal, D. W., DeVoti, J. A., Steinberg, B. M., Abramson, A. L., & Bonagura, V. R. (2012). TH2-like Chemokine Patterns Correlate with Disease Severity in Patients with Recurrent Respiratory Papillomatosis. *Molecular Medicine*, **18**(1), 1338.

RStudio Team. (2015). RStudio: Integrated Development for R., Boston, MA: RStudio, Inc.

Saha, B., Adhikary, A., Ray, P., … Das, T. (2011). Restoration of tumor suppressor p53 by differentially regulating pro- and anti-p53 networks in HPV-18-infected cervical cancer cells. *Oncogene 2012 31:2*, **31**(2), 173–186.

Sahragardjoonegani, B., Beall, R. F., Kesselheim, A. S., & Hollis, A. (2021). Repurposing existing drugs for new uses: a cohort study of the frequency of FDA-granted new indication exclusivities since 1997. *Journal of Pharmaceutical Policy and Practice 2021 14:1*, **14**(1), 1–8.

Sakakibara, N., Chen, D., & McBride, A. A. (2013). Papillomaviruses Use Recombination-Dependent Replication to Vegetatively Amplify Their Genomes in Differentiated Cells. *PLoS Pathogens*, **9**(7), e1003321.

Sakamoto, K., Fujii, T., Kawachi, H., … Yamaguchi, A. (2012). Reduction of NOTCH1 expression pertains to maturation abnormalities of keratinocytes in squamous neoplasms. *Laboratory Investigation 2012 92:5*, **92**(5), 688–702.

Sakaue-Sawano, A., Kurokawa, H., Morimura, T., … Miyawaki, A. (2008). Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. *Cell*, **132**(3), 487–498.

Saldivar, J. S., Lopez, D., Feldman, R. A., … Baldwin, W. S. (2007). COX-2 overexpression as a biomarker of early cervical carcinogenesis: A pilot study. *Gynecologic Oncology*, **107**(1), S155–S162.

Santos, C., Neto, T., Ferreirinha, P., … Gil Da Costa, R. M. (2016). Celecoxib promotes degranulation of CD8+ T cells in HPV-induced lesions of K14-HPV16 transgenic mice. *Life Sciences*, **157**, 67–73.

Sarian, L. O., Derchain, S. F., Yoshida, A., Vassallo, J., Pignataro, F., & de Angelo Andrade, L. A. L. (2006). Expression of cyclooxygenase-2 (COX-2) and Ki67 as related to disease severity and HPV detection in squamous lesions of the cervix. *Gynecologic Oncology*, **102**(3), 537–541.

Sarrazin, S., Lamanna, W. C., & Esko, J. D. (2011). Heparan sulfate proteoglycans. *Cold Spring Harbor Perspectives in Biology*, **3**(7). doi:10.1101/cshperspect.a004952

Schaal, T. D., Mallet, W. G., McMinn, D. L., … Parekh, B. S. (2003). Inhibition of human papilloma virus E2 DNA binding protein by covalently linked polyamides. *Nucleic Acids Research*, **31**(4), 1282–91.

Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J., & Howley, P. M. (1990). The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell*, **63**(6), 1129–1136.

Schiller, J. T., Day, P. M., & Kines, R. C. (2010). Current understanding of the mechanism of HPV infection. *Gynecologic Oncology*, **118**(1), S12–S17.

Schwaber, J., Andersen, S., & Nielsen, L. (2019). Shedding light: The importance of reverse transcription efficiency standards in data interpretation. *Biomolecular Detection and Quantification*, **17**, 100077.

Seavey, S. E., Holubar, M., Saucedo, L. J., & Perry, M. E. (1999). The E7 Oncoprotein of Human Papillomavirus Type 16 Stabilizes p53 through a Mechanism Independent of p19ARF. *Journal of Virology*, **73**(9), 7590–7598.

Serrano, B., Brotons, M., Bosch, F. X., & Bruni, L. (2018). Epidemiology and burden of HPV-related disease. *Best Practice & Research Clinical Obstetrics & Gynaecology*, **47**, 14–26.

Sham, H. L., Kempf, D. J., Molla, A., … Norbeck, D. W. (1998). ABT-378, a Highly Potent Inhibitor of the Human Immunodeficiency Virus Protease. *Antimicrobial Agents and Chemotherapy*, **42**(12), 3218.

Sherr, C. J. (2004). Principles of tumor suppression. *Cell*, **116**(2), 235–46.

Shraiman, B. I. (2005). Mechanical feedback as a possible regulator of tissue growth. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(9), 3318–3323.

Siebel, C., & Lendahl, U. (2017). Notch Signaling in Development, Tissue Homeostasis, and Disease. *Https://Doi.Org/10.1152/Physrev.00005.2017*, **97**(4), 1235–1294.

Silverberg, M. J., Thorsen, P., Lindeberg, H., Grant, L. A., & Shah, K. v. (2003). Condyloma in pregnancy is strongly predictive of Juvenile-Onset recurrent respiratory papillomatosis. *Obstetrics & Gynecology*, **101**(4), 645–652.

Simms, K. T., Hanley, S. J. B., Smith, M. A., Keane, A., & Canfell, K. (2020). Impact of HPV vaccine hesitancy on cervical cancer in Japan: a modelling study. *The Lancet Public Health*, **5**(4), e223–e234.

Sittampalam, G. S., Iversen, P. W., Boadt, J. A., … Lister, M. D. (1997). Design of Signal Windows in High Throughput Screening Assays for Drug Discovery: *Http://Dx.Doi.Org/10.1177/108705719700200306*, **2**(3), 159–169.

Six, E. M., Ndiaye, D., Sauer, G., … Logeat, F. (2004). The Notch Ligand Delta1 Recruits Dlg1 at Cell-Cell Contacts and Regulates Cell Migration * [boxs]. *Journal of Biological Chemistry*, **279**(53), 55818–55826.

Slattum, G., McGee, K. M., & Rosenblatt, J. (2009). P115 RhoGEF and microtubules decide the direction apoptotic cells extrude from an epithelium. *The Journal of Cell Biology*, **186**(5), 693–702.

Slavik, M. A., Allen-Hoffmann, B. L., Liu, B. Y., & Alexander, C. M. (2007). Wnt signaling induces differentiation of progenitor cells in organotypic keratinocyte cultures. *BMC Developmental Biology*, **7**(1), 9.

Smail, M. A., Reigle, J. K., & McCullumsmith, R. E. (2021). Using protein turnover to expand the applications of transcriptomics. *Scientific Reports 2021 11:1*, **11**(1), 1–9.

Smotkin, D., Prokoph, H., & Wettstein, F. O. (1989). Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms. *Journal of Virology*, **63**(3), 1441–7.

Sominsky, S., Shterzer, N., Jackman, A., Shapiro, B., Yaniv, A., & Sherman, L. (2017). E6 proteins of α and β cutaneous HPV types differ in their ability to potentiate Wnt signaling. *Virology*, **509**, 11–22.

Song, S., Pitot, H. C., & Lambert, P. F. (1999). The Human Papillomavirus Type 16 E6 Gene Alone Is Sufficient To Induce Carcinomas in Transgenic Animals. *Journal of Virology*, **73**(7), 5887–5893.

Spanos, W. C., Hoover, A., Harris, G. F., … Lee, J. H. (2008). The PDZ Binding Motif of Human Papillomavirus Type 16 E6 Induces PTPN13 Loss, Which Allows Anchorage-Independent Growth and Synergizes with Ras for Invasive Growth. *Journal of Virology*, **82**(5), 2493–2500.

Spayne, J., & Hesketh, T. (2021). Estimate of global human papillomavirus vaccination coverage: analysis of country-level indicators. *BMJ Open*, **11**(9). doi:10.1136/BMJOPEN-2021-052016

Stanley, M. A. (2012). Epithelial cell responses to infection with human papillomavirus. *Clinical Microbiology Reviews*, **25**(2), 215–22.

Šterbenc, A., Hošnjak, L., Chouhy, D., … Poljak, M. (2017). Molecular characterization, tissue tropism, and genetic variability of the novel Mupapillomavirus type HPV204 and phylogenetically related types HPV1 and HPV63. *PLOS ONE*, **12**(4), e0175892.

Stucke, V. M., Timmerman, E., Vandekerckhove, J., Gevaert, K., & Hall, A. (2007). The MAGUK Protein MPP7 Binds to the Polarity Protein hDlg1 and Facilitates Epithelial Tight Junction Formation. *Https://Doi.Org/10.1091/Mbc.E06-11-0980*, **18**(5), 1744–1755.

Sun, T., Chen, C., Wu, Y., Zhang, S., Cui, J., & Shen, P. (2009). Modeling the role of p53 pulses in DNA damage- induced cell death decision. *BMC Bioinformatics 2009 10:1*, **10**(1), 1–12.

Szymonowicz, K. A., & Chen, J. (2020). Biological and clinical aspects of HPV-related cancers. *Cancer Biology & Medicine*, **17**(4), 864.

Tang, S., Tao, M., McCoy, J. P., & Zheng, Z.-M. (2006). 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. *Journal of Virology*, **80**(9), 4249–63.

Thanan, R., Murata, M., Ma, N., … Kawanishi, S. (2012). Nuclear localization of COX-2 in relation to the expression of stemness markers in urinary bladder cancer. *Mediators of Inflammation*, **2012**. doi:10.1155/2012/165879

Thomas, J. T., Hubert, W. G., Ruesch, M. N., & Laimins, L. A. (1999a). Human papillomavirus type 31 oncoproteins E6 and E7 are required for the maintenance of episomes during the viral life cycle in normal human keratinocytes. *Proceedings of the National Academy of Sciences*, **96**(15), 8449–8454.

Thomas, M. C., & Chiang, C.-M. (2005). E6 Oncoprotein Represses p53-Dependent Gene Activation via Inhibition of Protein Acetylation Independently of Inducing p53 Degradation. *Molecular Cell*, **17**(2), 251–264.

Thomas, M., Laura, R., Hepner, K., … Banks, L. (2002). Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation. *Oncogene*, **21**(33), 5088–5096.

Thomas, M., Myers, M. P., Massimi, P., Guarnaccia, C., & Banks, L. (2016a). Analysis of Multiple HPV E6 PDZ Interactions Defines Type-Specific PDZ Fingerprints That Predict Oncogenic Potential. *PLOS Pathogens*, **12**(8), e1005766.

Thomas, M., Myers, M. P., Massimi, P., Guarnaccia, C., & Banks, L. (2016b). Analysis of Multiple HPV E6 PDZ Interactions Defines Type-Specific PDZ Fingerprints That Predict Oncogenic Potential. *PLOS Pathogens*, **12**(8), e1005766.

Thomas, M., Narayan, N., Pim, D., … Banks, L. (2008). Human papillomaviruses, cervical cancer and cell polarity. *Oncogene 2008 27:55*, **27**(55), 7018–7030.

Thomas, M., Pim, D., & Banks, L. (1999b). The role of the E6-p53 interaction in the molecular pathogenesis of HPV. *Oncogene*, **18**(53), 7690–7700.

Tomiyama, L., Sezaki, T., Matsuo, M., Ueda, K., & Kioka, N. (2015). Loss of Dlg5 expression promotes the migration and invasion of prostate cancer cells via Girdin phosphorylation. *Oncogene*, **34**(9), 1141–1149.

Tomson, N., Sterling, J., Ahmed, I., Hague, J., & Berth-Jones, J. (2011). Human papillomavirus typing of warts and response to cryotherapy. *Journal of the European Academy of Dermatology and Venereology*, **25**(9), 1108–1111.

Toots, M., Ustav, M., Männik, A., … Ustav, M. (2017). Identification of several high-risk HPV inhibitors and drug targets with a novel high-throughput screening assay. *PLoS Pathogens*, **13**(2). doi:10.1371/journal.ppat.1006168

Tsuboi, A., Ohsawa, S., Umetsu, D., … Fujimoto, K. (2018). Competition for Space Is Controlled by Apoptosis-Induced Change of Local Epithelial Topology. *Current Biology*, **28**(13), 2115-2128.e5.

Uchida, C., Miwa, S., Kitagawa, K., … Kitagawa, M. (2005). Enhanced Mdm2 activity inhibits pRB function via ubiquitin-dependent degradation. *The EMBO Journal*, **24**(1), 160–169.

Uhlen, M., Fagerberg, L., Hallstrom, B. M., … Ponten, F. (2015). Tissue-based map of the human proteome. *Science*, **347**(6220), 1260419–1260419.

Vaezi, A., Bauer, C., Vasioukhin, V., & Fuchs, E. (2002). Actin Cable Dynamics and Rho/Rock Orchestrate a Polarized Cytoskeletal Architecture in the Early Steps of Assembling a Stratified Epithelium. *Developmental Cell*, **3**(3), 367–381.

van Es, J. H., Jay, P., Gregorieff, A., … Clevers, H. (2005). Wnt signalling induces maturation of Paneth cells in intestinal crypts. *Nature Cell Biology 2005 7:4*, **7**(4), 381–386.

vande Pol, S. B., & Klingelhutz, A. J. (2013). Papillomavirus E6 oncoproteins. *Virology*, **445**(1–2), 115–137.
Venuti, A., Paolini, F., Nasir, L., … Borzacchiello, G. (2011). Papillomavirus E5: the smallest oncoprotein with many functions. *Molecular Cancer*, **10**(1), 140.
Vishnoi, K., Mahata, S., Tyagi, A., … Bharti, A. C. (2016). Cross-talk between Human Papillomavirus Oncoproteins and Hedgehog Signaling Synergistically Promotes Stemness in Cervical Cancer Cells. *Scientific Reports 2016 6:1*, **6**(1), 1–15.
Vodermaier, H. C. (2004). APC/C and SCF: Controlling Each Other and the Cell Cycle. *Current Biology*, **14**(18), R787–R796.
von Bredow, D. C., Nagle, R. B., Bowden, G. T., & Cress, A. E. (1997). Cleavage of β4 Integrin by Matrilysin. *Experimental Cell Research*, **236**(1), 341–345.
Vousden, K. H., & Prives, C. (2009). Blinded by the Light: The Growing Complexity of p53. *Cell*, **137**(3), 413–431.
Wagstaff, L., Goschorska, M., Kozyrska, K., … Piddini, E. (2016). Mechanical cell competition kills cells via induction of lethal p53 levels. *Nature Communications 2016 7:1*, **7**(1), 1–14.
Walhart, T., Isaacson-Wechsler, E., Ang, K.-H., Arkin, M., Tugizov, S., & Palefsky, J. M. (2020). A Cell-Based Renilla Luminescence Reporter Plasmid Assay for High-Throughput Screening to Identify Novel FDA-Approved Drug Inhibitors of HPV-16 Infection. *SLAS Discovery : Advancing Life Sciences R & D*, **25**(1), 79–86.
Wang, D.-G., Wang, S., Huang, B., & Liu, F. (2019). Roles of cellular heterogeneity, intrinsic and extrinsic noise in variability of p53 oscillation. *Scientific Reports 2019 9:1*, **9**(1), 1–11.
Wang, N. J., Sanborn, Z., Arnett, K. L., … Cho, R. J. (2011). Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. *Proceedings of the National Academy of Sciences*, **108**(43), 17761–17766.
Wang, Q., Griffin, H., Southern, S., … Doorbar, J. (2004). Functional analysis of the human papillomavirus type 16 E1=E4 protein provides a mechanism for in vivo and in vitro keratin filament reorganization. *Journal of Virology*, **78**(2), 821–33.
Wang, S. H. J., Celic, I., Choi, S. Y., … Kolodkin, A. L. (2014). Dlg5 regulates dendritic spine formation and synaptogenesis by controlling subcellular N-cadherin localization. *Journal of Neuroscience*, **34**(38), 12745–12761.
Watt, F. M. (2002). NEW EMBO MEMBER'S REVIEW: Role of integrins in regulating epidermal adhesion, growth and differentiation. *The EMBO Journal*, **21**(15), 3919–3926.
Watt, F. M., Estrach, S., & Ambler, C. A. (2008). Epidermal Notch signalling: differentiation, cancer and adhesion. *Current Opinion in Cell Biology*, **20**(2), 171–179.
Wawryk-Gawda, E., Chylińska-Wrzos, P., Lis-Sochocka, M., … Jodłowska-Jędrych, B. (2014). P53 protein in proliferation, repair and apoptosis of cells. *Protoplasma*, **251**(3), 525–33.
Weijzen, S., Zlobin, A., Braid, M., Miele, L., & Kast, W. M. (2003). HPV16 E6 and E7 oncoproteins regulate Notch-1 expression and cooperate to induce transformation. *Journal of Cellular Physiology*, **194**(3), 356–362.
Weinberg, R. A. (1995). The retinoblastoma protein and cell cycle control. *Cell*, **81**(3), 323–330.
White, E. A. (2019). Manipulation of Epithelial Differentiation by HPV Oncoproteins. *Viruses*, **11**(4). doi:10.3390/v11040369

White, E. A., Kramer, R. E., Tan, M. J. A., Hayes, S. D., Harper, J. W., & Howley, P. M. (2012). Comprehensive analysis of host cellular interactions with human papillomavirus E6 proteins identifies new E6 binding partners and reflects viral diversity. *Journal of Virology*, **86**(24), 13174–86.

White, M. K., Pagano, J. S., & Khalili, K. (2014). Viruses and Human Cancers: a Long Road of Discovery of Molecular Paradigms. *Clinical Microbiology Reviews*, **27**(3), 463.

White, P. W., Titolo, S., Brault, K., … Archambault, J. (2003). Inhibition of human papillomavirus DNA replication by small molecule antagonists of the E1-E2 protein interaction. *Journal of Biological Chemistry*, **278**(29), 26765–26772.

WHO | Cervical cancer. (2018). *WHO*.

Widelitz, R. B. (2008). Wnt signaling in skin organogenesis. *Organogenesis*, **4**(2), 123.

Williams, S. E., Beronja, S., Pasolli, H. A., & Fuchs, E. (2011). Asymmetric cell divisions promote Notch-dependent epidermal differentiation. *Nature*, **470**(7334), 353–358.

Woodhall, S., Jit, M., Soldan, K., … Lacey, C. (2011). The impact of genital warts: Loss of quality of life and cost of treatment in eight sexual health clinics in the UK. *Sexually Transmitted Infections*, **87**, 458–463.

Woodhall, S., Ramsey, T., Cai, C., … Lacey, C. (2008). Estimation of the impact of genital warts on health-related quality of life. *Sexually Transmitted Infections*, **84**, 161–166.

Wright, G. J., Leslie, J. D., Ariza-McNaughton, L., & Lewis, J. (2004). Delta proteins and MAGI proteins: an interaction of Notch ligands with intracellular scaffolding molecules and its significance for zebrafish development. *Development*, **131**(22), 5659–5669.

Y, Z., Y, X., & WG, Y. (1998). ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. *Cell*, **92**(6), 725–734.

Yamanaka, T., & Ohno, S. (2008). Role of Lgl/Dlg/Scribble in the regulation of epithelial junction, polarity and growth. *Frontiers in Bioscience : A Journal and Virtual Library*, **13**(17), 6693–6707.

Yamato, K., Yamada, T., Kizaki, M., … Yoshinouchi, M. (2008). New highly potent and specific E6 and E7 siRNAs for treatment of HPV16 positive cervical cancer. *Cancer Gene Therapy*, **15**(3), 140–153.

Yamayoshi, A., Kato, K., Suga, S., … Wake, N. (2007). Specific apoptosis induction in human papillomavirus-positive cervical carcinoma cells by photodynamic antisense regulation. *Oligonucleotides*, **17**(1), 66–79.

Yoakim, C., Ogilvie, W. W., Goudreau, N., … White, P. W. (2003). Discovery of the first series of inhibitors of human papillomavirus type 11: inhibition of the assembly of the E1-E2-Origin DNA complex. *Bioorganic & Medicinal Chemistry Letters*, **13**(15), 2539–2541.

Yoshimatsu, Y., Nakahara, T., Tanaka, K., … Kiyono, T. (2017). Roles of the PDZ-binding motif of HPV 16 E6 protein in oncogenic transformation of human cervical keratinocytes. *Cancer Science*, **108**(7), 1303–1309.

Yoshinouchi, M., Yamada, T., Kizaki, M., … Yamato, K. (2003). In vitro and in vivo growth suppression of human papillomavirus 16-positive cervical cancer cells by e6 siRNA. *Molecular Therapy*, **8**(5), 762–768.

Young, J. M., Abidine, A. Z. el, Gómez-Martinez, R. A., … Ozbun, M. A. (2021). Protamine Sulfate Is a Potent Inhibitor of Human Papillomavirus Infection In Vitro and In Vivo. *Antimicrobial Agents and Chemotherapy*, **66**(1). doi:10.1128/AAC.01513-21

Yugawa, T., Nishino, K., Ohno, S.-I., … Kiyono, T. (2013). Noncanonical NOTCH signaling limits self-renewal of human epithelial and induced pluripotent stem cells through ROCK activation. *Molecular and Cellular Biology*, **33**(22), 4434–47.

Zanier, K., Charbonnier, S., Sidi, A. O. M. hamed O., … Travé, G. (2013a). Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science*, **339**(6120), 694–698.

Zanier, K., Charbonnier, S., Sidi, A. O. M. O., … Travé, G. (2013b). Structural Basis for Hijacking of Cellular LxxLL Motifs by Papillomavirus E6 Oncoproteins. *Science*, **339**(6120), 694–698.

Zanier, K., Stutz, C., Kintscher, S., … Hoppe-Seyler, F. (2014). The E6AP Binding Pocket of the HPV16 E6 Oncoprotein Provides a Docking Site for a Small Inhibitory Peptide Unrelated to E6AP, Indicating Druggability of E6. *PLoS ONE*, **9**(11), e112514.

Zehbe, I., Richard, C., Lee, K. F., Campbell, M., Hampson, L., & Hampson, I. N. (2011). Lopinavir shows greater specificity than zinc finger ejecting compounds as a potential treatment for human papillomavirus-related lesions. *Antiviral Research*, **91**(2), 161–166.

Zhang, J. H., Chung, T. D. Y., & Oldenburg, K. R. (1999a). A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *Journal of Biomolecular Screening*, **4**(2), 67–73.

Zhang, P., Nouri, M., Brandsma, J. L., Iftner, T., & Steinberg, B. M. (1999b). Induction of E6/E7 Expression in Cottontail Rabbit Papillomavirus Latency Following UV Activation. *Virology*, **263**(2), 388–394.

Zhang, Q., Pan, Y., Ji, J., Xu, Y., Zhang, Q., & Qin, L. (2021). Roles and action mechanisms of WNT4 in cell differentiation and human diseases: a review. *Cell Death Discovery 2021 7:1*, **7**(1), 1–10.

Zhang, X. D. (2011). Quality Control in Genome-Scale RNAi Screens. *Optimal High-Throughput Screening*, 42–61.

Zheng, K., Egawa, N., Shiraz, A., … Doorbar, J. (2022). The Reservoir of Persistent Human Papillomavirus Infection; Strategies for Elimination Using Anti-Viral Therapies. *Viruses*, **14**(2). doi:10.3390/V14020214

Zheng, Y., Zhang, J., & Rao, Z. (2004). Ribozyme targeting HPV16 E6E7 transcripts in cervical cancer cells suppresses cell growth and sensitizes cells to chemotherapy and radiotherapy. *Cancer Biology and Therapy*, **3**(11), 1129–1134.

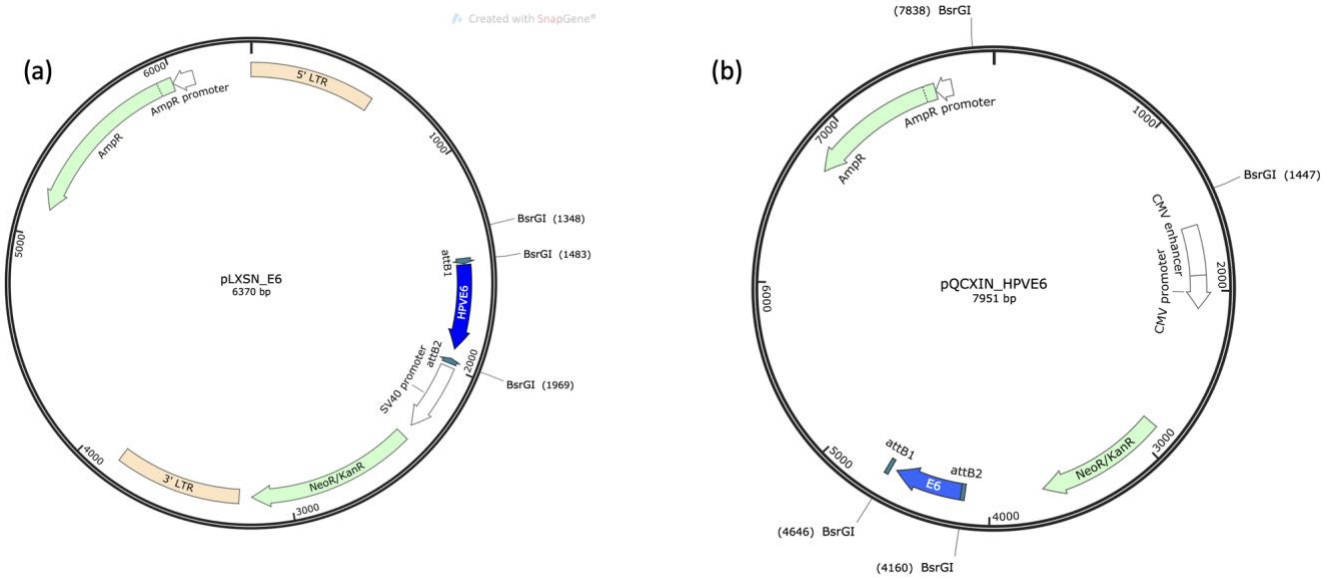
Zhou, Z., Guo, Y., Liu, Y., … Qiu, J. (2015). Methylation-mediated silencing of Dlg5 facilitates bladder cancer metastasis. *Experimental Cell Research*, **331**(2), 399–407.

Zimmermann, H., Koh, C.-H., Degenkolbe, R., … Bernard, H.-U. (2000). *Printed in Great Britain Interaction with CBP/p300 enables the bovine papillomavirus type 1 E6 oncoprotein to downregulate CBP/p300-mediated transactivation by p53*. *Journal of General Virology*, Vol. 81.

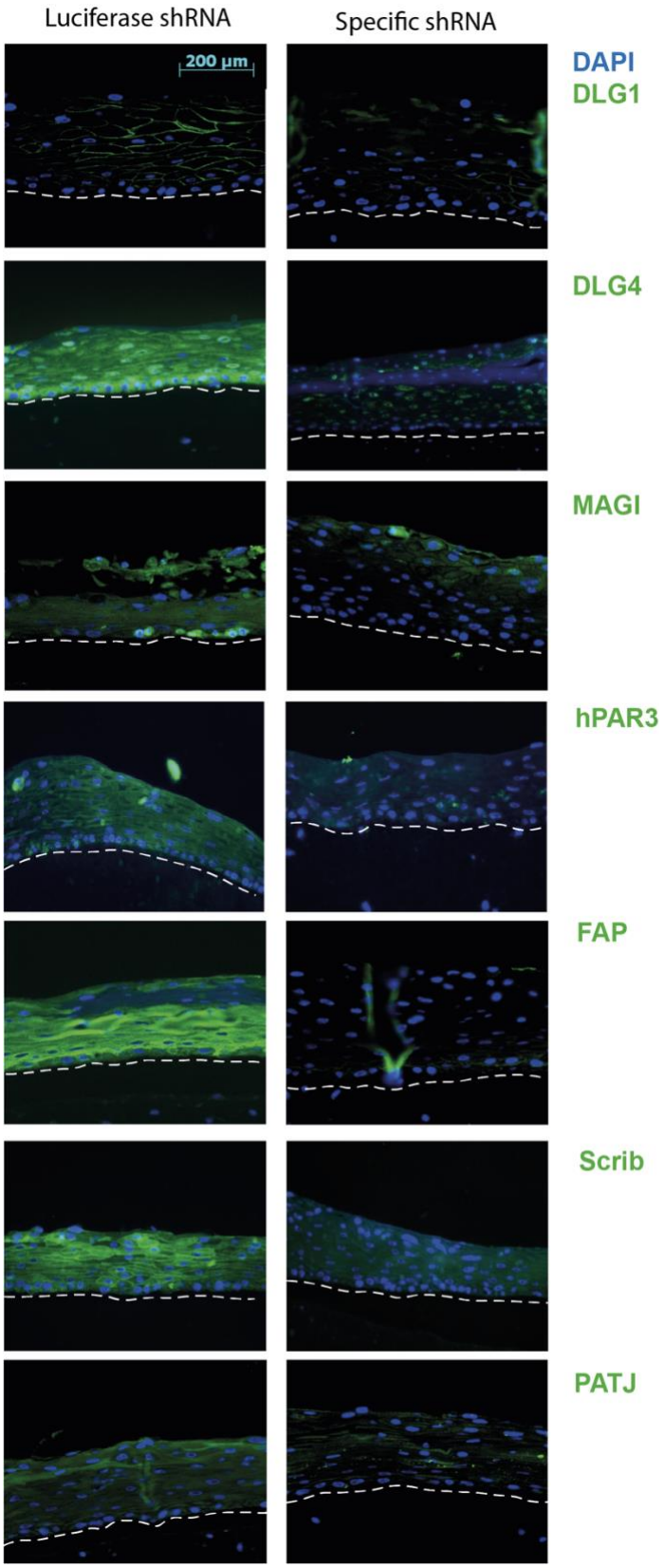
Zoete, V., Grosdidier, A., & Michielin, O. (2009). Docking, virtual high throughput screening and in silico fragment-based drug design. *Journal of Cellular and Molecular Medicine*, **13**(2), 238–248.

zur Hausen, H. (1977). Human papillomaviruses and their possible role in squamous cell carcinomas. *Current Topics in Microbiology and Immunology*, **78**, 1–30.

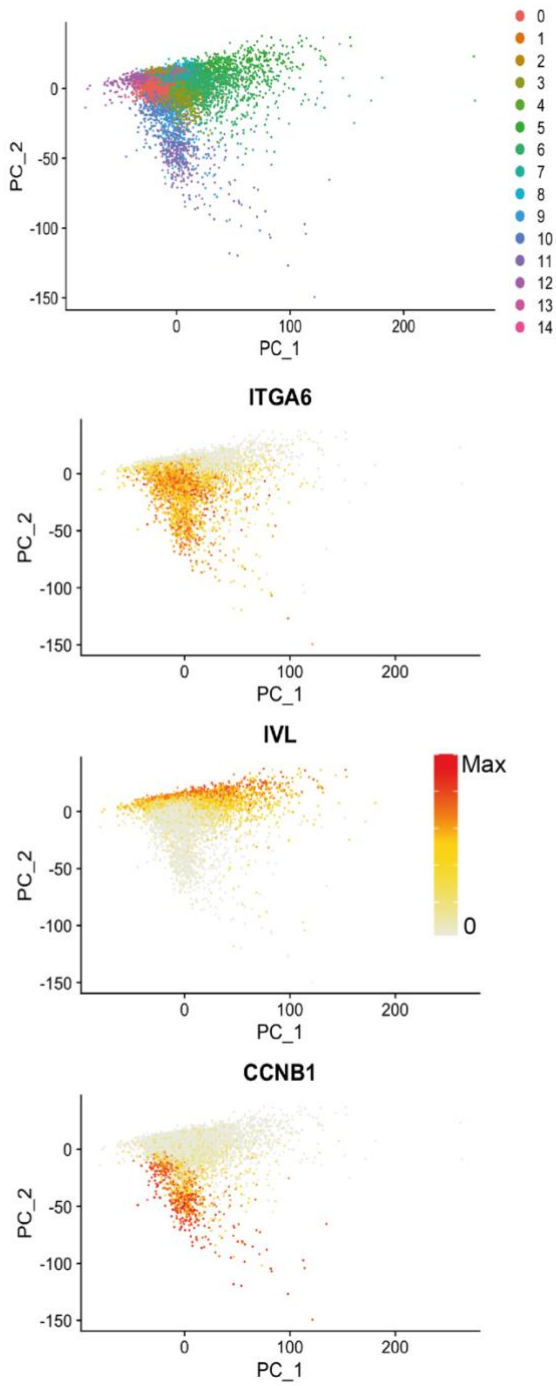
Appendix



Appendix Figure 1: Plasmid backbone design of (a) pLXSN_E6; (b) pQCXIN_E6



Appendix Figure 2: Validation of shRNA knockdown of PDZ-domain containing proteins. Raft culture of NIKS shRNA cells (right panel) were generated with Luciferase shRNA (LuciRi) (left panel) as pseudo-knockdown control. Immunostainings used to validate the downregulation of targets at protein level. Images collected with 40x objective.



Appendix Figure 3: **PCA plots of the data structure of NIKS and NIKS 16E6** Data shown as dimension PCA plot of 15 different clusters from the combined datasets of NIKS and NIKS 16E6, and RNA expression profile of markers for basal keratinocyte (ITGA6), committed to differentiation (IVL) and G2/M phase cell cycle (CCNB1). Expression level is shown as log2 fold change.

Appendix File 1: Prestwick Chemical Library®

Appendix File 2: High-Throughput Screening plate layout

Appendix File 3: High-Throughput Screening results

Appendix File 4: scRNA seq markers