Pull-down of Human papillomavirus 16 E6 variants for Interactome Analysis: A Methods Development and Analysis

A MSc thesis presented to

The Faculty of Graduate Studies and Department of Biology

of

Lakehead University

By: MEHRAN MASOOM

Supervisor: Dr. Ingeborg Zehbe

Committee Members

Dr. Simon Lees and Dr. Heidi Schraft

External Reviewer

Dr. Michael Campbell

February 29, 2020

Abstract

Human Papillomavirus 16 (HPV16) is a double-stranded DNA virus known as a causative agent in almost all cervical cancers and an increasing number of oropharyngeal cancers. Variants of HPV16, such as the Asian-American (AA) and L83V, have been found to have increased abilities to promote carcinogenesis. Even though previous interactome studies identified which proteins interact with HPV16 E6, few have looked at interactions between variants of E6 in particular AAE6 and the European Prototype (EPE6). This thesis had two objectives: develop a method to co-immunoprecipitate host cellular proteins that interact with E6 variants; and identify potential differences between cellular host proteins and variant E6. We were successful in developing a method that not only could pull-down and identify E6 variant interacting proteins but the E6 variant proteins themselves. Using liquid chromatography-mass spectrometry (LC MS/MS), we identified 13 proteins that interact with both AAE6 and EPE6, along with six unique AAE6 interacting proteins and six unique EPE6 interacting proteins. Of the interactors we found, seven were of particular interest: TRIP12, GNL2, INO80B, CHMP4B, MX2, RPSK6K4A, and PROK2. These proteins affect a variety of cellular functions, including DNA replication and repair, telomere maintenance, cellular proliferation, ERK1/2 signaling, signal transduction, and immune response. Identification of different proteins using different bioinformatic analyses further provide evidence that AAE6 and EPE6 may have unique interactions with their host cells resulting in varied abilities to promote carcinogenesis. The identification of these proteins has furthered our understanding of potential mechanisms that allow AAE6 to promote carcinogenesis more than EPE6. More wet lab work is still required to confirm these interactions and determine their exact effects on the host cell.

Lay Summary

Faculty and students in the Department of Biology are bound together by a common interest in explaining the diversity of life, the fit between form and function, and the distribution and abundance of organisms. Human papillomavirus 16 (HPV16) is a common virus that researchers have found to be the cause of almost all cervical cancers as well as the majority of head and neck cancers in both men and women. This thesis looks explicitly at two variants of the E6 protein in HPV16: the Asian-American (AA) and European Prototype (EP). Previous research conducted within our lab, along with other groups, found that AAE6 causes more invasive cancers compared to EPE6. This thesis aimed to develop a method to identify proteins that interact with E6 variants. This thesis also aimed to identify whether there is a difference in proteins that interact between AAE6 and EPE6 such that it would uncover new and novel understandings for why AAE6 can better cause cancer. The information uncovered by this study can help improve our understanding of how HPV causes cancer as well as potentially discover new ways to treat HPV16 cancers.

Dedication

This is for you bro, you have never given up on me just like I will never give up on you!

Acknowledgements

Professional

The most important individual throughout this entire journey has been my supervisor and mentor, Dr. Ingeborg Zehbe. You have spent countless hours with me, assisting me, and pushing me to become the best version of myself. You have helped me understand how to be humble and never give up when the going gets tough. I am always amazed by how you help me look at research with an open mind and helped me realize that there is a multitude of ways to approach a single task or problem. You have given me so many opportunities to connect and collaborate with researchers throughout Canada and the United States. For that, I am eternally grateful.

To Dr. Guillem Dayer, my co-supervisor! You have spent countless hours helping me be successful. The countless laughs we had every day I have no words that can describe how thankful I am for all that you have done for me. When things became challenging you helped me to continue pushing forward. Not only were you a fantastic mentor and fellow scientist, but you are a fantastic life-long friend as well.

To my committee members Dr. Heidi Schraft and Dr. Simon Lees, thank you for the time you have taken out of your busy lives to help guide me throughout this journey. I also would like to thank Dr. Michael Campbell, my external reviewer, for taking the time to evaluate my thesis.

Dr. James Knockleby and Dr. Hoyun Lee from the Health Sciences North Research Institute, I cannot thank you enough for making me feel at home during my time in Sudbury, Ontario. Dr. Hoyun Lee, thank you once again for your hospitality and outstanding kindness. Dr. Knockleby, thank you for spending dozens of hours with me,

from late nights to early mornings with the mass spectrometer! Dr. Melissa Togtema, Dr. Robert Jackson, I cannot thank you two enough for all of your support. From getting me settled in the lab and supporting me. You both taught me so many skills and spent so many hours reviewing my work and helping me improve every day. I hope you both know that I always remember the moments where you helped me. From learning how to culture cells to designing figures and reviewing my experiment designs. You both are brilliant and excellent scholars! I would also like to thank my fellow lab members and friends, Meagan, and Yuri. For amazing memories and helping me throughout my early days at the Thunder Bay Regional Health Research Institute.

Personal

I am so grateful to have my mother (Lucja) and father (Shoaa) in my life. You keep me striving to be the best I can be every day. Matthew, I don't know how you haven't gotten tired of me talking about my thesis, but I'm glad you are there to listen and encourage me. Adrian, I appreciate every time you check up on me to see how things are going. I don't always express that I am thankful, but please know every day I knew you were making sure I was making progress.

Table of Contents

Table of Figures

Table of Tables

1 Introduction

1.1 Historical Overview of Human Papillomavirus

Since the late seventies, human papillomavirus (HPV) is known to be responsible for almost all (99.7 %) worldwide cases of cervical cancer (Walboomers et al. 1999; zur Hausen 1977). Recently, approximately 1 350 women in Canada and 569 800 women across the globe were diagnosed with cervical cancer, making it the fourth most common female cancer worldwide (Smith et al. 2019; Ferlay et al. 2019). Primarily, HPV affects intraepithelial cells of the cervix and other squamous cells within the ano-genital region of males and females causing a substantial number of cervical and ano-genital lesions. However, HPV also has the potential to infect squamous cells of the pharynx in both males and females causing oropharyngeal squamous cell carcinomas (OPSCCs) (Burk et al. 2003; Chaturvedi et al. 2011; Wakeham et al. 2014). From 1989 to 2004, worldwide HPV prevalence in OPSCCs increased drastically from 16 % to over 70 %. In Canada alone, the annual incidence of HPV related OPSCCs has risen on average by 3.4 % and 1.1 % per year for males and females respectively (Nuttall et al. 2016). Researchers believed that the number of HPV positive OPSCC's would outnumber HPV positive cervical cancers by 2020 (Chaturvedi et al. 2011). In 2019 their hypotheses were confirmed when the number of HPV positive OPSCC's in the United States (n=13 300) was greater than the number of HPV positive cervical cancers (n=10 933) (OPSCC incidence (n=19 000) and 70.1% HPV positive incidence of OPSCC's obtained from the United States Cancer Statistics and Saraiya et al. 2016).

As of 2019, researchers have annotated 225 HPV types while, over 240 are awaiting classification as oncogenic "high-risk" (HR) and "low-risk" (LR), making this virus a

global burden for humans (Papillomavirus Episteme (PaVE) from Van Doorslaer et al. 2016; Pastrana et al. 2019). HR HPV types such as 16 and to a lesser extent 18, are responsible for over half of all HPV positive cervical carcinomas (reviewed in zur Hausen 2009). In 2007, Australia initiated the first government-funded National HPV Vaccination Program (NHVP), to reduce the incidence of HPV infections in youth (Tabrizi et al. 2012). The program was aimed at observing genotype-specific HPV infections in females aged 18 - 24 over a span of four years prior to and after the vaccination period (2005 – 2007 and 2010-2011 respectively). The study concluded that the introduction of an HPV vaccination program reduced HPV 6, 11, 16, 18 infections by over 20 %. The Australian NHVP is so successful that researchers believe cervical cancers will no longer be a public health concern (less than 4 new cases per 100 000 females) in Australia within the next 20 years (Hall et al. 2019). Surprisingly, with HPV vaccines such as Gardasil9 and Cervarix available to the public, less than 2 % of the global female population have completed vaccinations (Harper and DeMars 2017). Although studies suggest one dose of HPV bivalent vaccines is enough to develop long-lasting immunity to HPV16 and 18 types, uptake rate and availability is low and next to none in developing countries (Kreimer et al. 2018). In Canada, HPV vaccine uptake is far below the average compared to other developed countries such as Australia (Tabrizi et al. 2012; Bird et al. 2017). Canada has set a goal of 85 % vaccine uptake for all female populations but as of 2017 only managed slightly under 56 % (Bird et al. 2017).

1.2 HPV16 Genome and Functions of Early Genes

HPV16 is a 7.9 kb double-stranded DNA virus that exists as circular episomes or integrated within the host genome (Smith et al. 2011; Cheung et al. 2013). There are four unique regions within the HPV16 genome: a long control region (LCR), an early region, a late region and a non-coding region (NCR) (Smith et al. 2011) (*Figure 1*). There are six early genes present within the early genome region. Early genes are numerically numbered and denoted by the letter 'E' (E1, E2, E4, E5, E6, E7). E1 is a helicase/ATPase while E2 behaves as a transcriptional regulator of E6 and E7 via binding to one of four binding sites in the LCR (McBride et al. 1991; Hughes et al. 1993; Bouvard et al. 1994). HPV16 transfected keratinocytes possess a E8^E2 spliced mRNA that is found to repress E1 replicase expression limiting early gene expression (Lace et al. 2008). E4 is present within the open reading frame (ORF) of E2. Transcription can initiate through the ORF within E2 or by an E1^E4 spliced mRNA that originates within the E1 ORF (Doorbar 2013). E4 plays a vital role in releasing the virus from the host environment as well as disrupting cellular keratin networks within cells (Doorbar 2013). E5 reduces tumour suppressor proteins p21 and p27 expression in HPV16 infected cells, increasing cellular proliferation (Venuti et al. 2011). E5 was recently discovered to be heavily involved in reducing miRNA-196a expression in HPV16 infected cells aiding in cellular transformation (Liu et al. 2015). E6 and E7 are of particular interest in HPV research. These genes are responsible for production of their respective proteins E6 and E7. These proteins are extensively studied because of their ability to transform and immortalize host cells in the absence of other early or late genes (reviewed in Vande Pol et al. 2013). E6 is involved with a multitude of functions including degradation of cellular proteins such as tumour suppressor protein P53

via the E6 protein binding to an E6 associated protein (E6AP) causing cellular immortalization. E7 most notably inhibits the retinoblastoma tumour suppressor protein (pRB) leading to loss of cell cycle regulation and increased cellular proliferation (reviewed in zur Hausen 2009).

1.3 E6 Oncoprotein

For over 40 years, researchers extensively studied and described the functions of E6 (zur Hausen et al. 1977; Scheffner et al. 1990; Huibregtse et al. 1991; Kiyono et al. 1997; Zhang et al. 2005). There are several biological functions regulated by the E6 protein, most notably as described above in *Section 1.2* the degradation of P53 through E6AP (reviewed in Vande Pol et al. 2013). Transcription is also affected by the expression of the E6 protein by interacting with histone acetyltransferases (HAT) such as human alteration/deficiency in activation protein (hAda3) (Neveu et al. 2012, reviewed in Vande Pol et al. 2013).

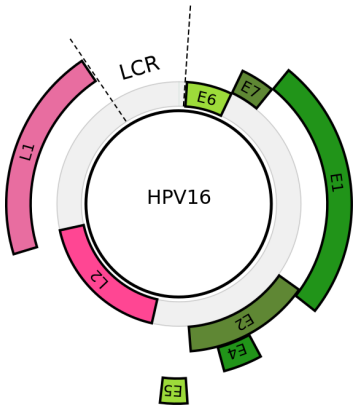

Figure 1 – Schematic representation of the circular HPV16 genome developed using the Pathogen Host Analysis Tool (PHAT) (Gibb et al. 2019). Green regions are early genes expressed during initial stages of HPV16 viral infection. Pink regions are late genes which are expressed in later stages of HPV16 infection. E6 is the oncogene of interest because of its ability to promote cell immortalization, immune evasion and malignancy (reviewed in Vande Pol et al. 2013).

This interaction is dependent on the interaction between E6 and E6AP whereas E6's ability to target another HAT, tat-interacting protein 60 kDa (TIP60) is independent of the formation of the E6-E6AP-P53 protein complex (Jha et al. 2010; reviewed in Vande Pol et al. 2013). Interestingly E6's ability to interact with E6AP is required to activate human telomerase reverse transcriptase (hTERT). Therefore, E6 is capable of immortalizing cells through maintaining telomeres within the genome (Liu et al. 2005).

Within the carboxy-terminus of high-risk E6 oncoproteins, a PDZ binding domain (PBM) is present (contains amino acid motif ETQL) allowing for the interaction of E6 with proteins containing a PDZ domain. These interactions result in the degradation or inhibition of PDZ-containing partners involved in cellular adhesion and polarity control such as: human homologue of Drosophila discs large (hDlg), human homologue of Drosophila scribble (hScrib), membrane-associated guanylate kinase with an inverted arrangement of protein-protein interaction domains (MAGI-1) and multi-PDZ-containing protein (MUPP1) (Nguyen et al. 2003). It should be noted that there is no PBM present within low-risk HPV E6 proteins suggesting that the PBM is integral for promoting epithelial hyperplasia (Nyugen et al. 2003; reviewed in Vande Pol et al. 2013). Interestingly, E6 in HPV16 contains a bimodal half-life with one segment having a half-life of approximately 4 hours, the other less than 30 minutes (Androphy et al. 1987). It is worth noting that E6 is stabilized in the presence of E6AP (Tomaić et al. 2009).

Recently, HPV16 E6 has been found to target the toll like receptor (TLR) pathway (Oliveira et al. 2018). E6 is capable of targeting six different proteins involved in this pathway: Inhibitor of nuclear factor kappa-B kinase subunits beta and epsilon, Interleukin-1 receptor-associated kinase-like 2, myeloid differentiation primary response protein

MyD88, TIR domain containing adapter molecule 1, and, TNF receptor associated factor 6. The result of E6's interactions with these TLR pathway proteins is an increase in NF-κB activation (Oliveira et al. 2018) It is important to note that many of these experiments were conducted using mutants of the prototype HPV16 E6 (*Table 1 Refs Within*). These proteins do not contain all of the mutations observed in naturally occurring variants, whereas our research aims to fill in this gap by using the full genome of HPV16 E6 variants AAE6 and EPE6. Even though so much is known about E6, many mechanisms underlying its ability to promote carcinogenesis remain unknown.

1.4 HPV16 E6 Variants

There are nucleotide variations within HPV16 E6 genes known as single nucleotide polymorphisms (SNPs). Analysis of SNPs within intratypes of HPV16 E6 has led to more specific classifications called variants. Many intratype SNPs are used to define variants of HPV16 E6 and initially were thought to be geographically distributed (Zuna et al. 2009). Since the first HPV16 genome sequenced was from a woman of European descent, the term "European Prototype" (EP) was used to define this variant (Seedorf et al. 1985; Zuna et al. 2009). The number of HPV16 E6 variants have increased steadily since the 1980s. In 1993, Ho et al. identified five distinct HPV16 geographical variants from cervical cancer biopsies named: European (EP-1 and EP-2), Asian (As), African-1 (Af-1), African-2 (Af-2) and Asian American (AA-1, AA-2, North-American (NA)) (Ho et al. 1993). Each of the geographical variants was thought to have evolved over 200 000 years through genetic drift likely originating in Africa. Variant nomenclature has largely remained the same (except for the Asian variant which sometimes is included as a European sub-variant) until 2013

(Ho et al. 1993; Smith et al. 2011; Burk et al. 2013). Unfortunately, geographical variant nomenclature is misleading due to factors such as breeding between humans, and increased migration rates (Clifford et al. 2019). The introduction of lineage variants in 2013 was a significant shift in HPV nomenclature moving away from geographic variants and using four alphabetic lineages from 'A' to 'D' and four numeric sub-lineages from '1 - 4' to classify and categorize intratype HPV SNPs (Burk et al. 2013). Current lineage classification of HPV16 is as follows: Lineage A is the most common worldwide with sub-lineages A1 - 2 being the most prevalent. Sub-lineages A3 and 4 are localized to East Asia and pose increased cancer risks to East Asia and South American Populations. Lineages B and C are localized to populations within Africa with sub-lineages B1 and C1 being the most common in Sub Saharan and Northern Africa respectively. Interestingly although sub-lineages B1 - 3 are primarily located in Sub-Saharan Africa, sub-lineage B4 is most common within Northern Africa. Lineage D HPVs are most commonly found in individuals from South/Central America. D4 is a unique sub-lineage that is almost exclusively found within Northern Africa. (PaVE from Van Doorslaer et al. 2016; Clifford et al. 2019).

HPV16 E6 variants have been reported to vary in their ability to produce malignant lesions within intraepithelial cells of the cervix and squamous cells of the head and neck (Zuna et al. 2009). Specifically, the Asian American variant (DNA sequence: sub-lineage D2 from GenBank ID AY686579.1; Amino acid sequence: sub-lineage D2 and D3 from GenBank ID: AAV91644.1 and AAO85408.1) of HPV16 E6 (AAE6) has been shown to possess significantly greater abilities to produce malignant lesions of the cervix compared to their European counterpart (EPE6) (DNA sequence: sub-lineage A1 from GenBank ID:

K02718.1; Amino acid sequence: A1 from GenBank ID: AAA46939.1) (Sichero et al. 2007; Zehbe et al. 2009 refs within). AAE6 and EPE6 genes differ by only six SNPs, half of which are non-synonymous (Zehbe et al. 1998). These SNPs result in AAE6 differing by three amino acids from EP to AA beginning form the second start codon of the 151-residue E6: Q14H, H78Y, and L83V (Figure 2) (Zehbe et al. 1998; Zehbe et al. 2009, Jackson et al. 2014).

This research shall focus on only the Prototype (A1) and Asian-American (D2/D3) sub-lineages because of their varied abilities to promote tumourigenesis (Richard et al. 2010; Niccoli et al. 2012). Research from the late 1990's has demonstrated that although the Prototype sub-lineage is equally present in Cervical intraepithelial neoplasia grade I – III (CIN I – III), the majority (94%) of invasive cervical carcinomas (ICC) were variants such as AAE6 (Zehbe et al. 1998). AAE6 has increased invasive and transforming potential resulting in increased carcinogenic abilities (Richard et al. 2010; Niccoli et al. 2012; Jackson et al. 2016).

1.5 Current Discoveries of HPV16 E6 Variant Studies

AAE6 possesses greater abilities to produce malignant lesions of the cervix compared to EPE6. The AAE6 variant of E6 is associated with increased viral persistence and increased cervical intraepithelial neoplasia (CIN) grade II to III development since 2000 (Villa et al. 2000; Berumen et al. 2001). Recently, Jackson et al. (2014) found that within three-dimensional keratinocyte raft models AAE6 produced moderate hyperplasia compared to EPE6 epithelium which only developed a mild dysplasia within epithelium (*Figure 3*). Researchers also found the Asian-American variant of HPV16 E6 to increase

PI3K/AKT and MAPK pathway expression through increased activation of MEK1, ERK2 and AKT2 proteins (Hochmann et al. 2016; Cuninghame et al. 2017). To further explain the mechanisms behind such functional observations, Zacapala-Gomez et al. completed a global transcriptome study of five E6 variants (AA-a, AA-c, E-A176/G350, E-C188/G350 and E-G350 (L83V)) expressed in C33A cells and found 387 differently expressed genes compared to cells transfected with the European Prototype (Zacapala-Gomez et al. 2016). Zacapala-Gomez et al. identified that Asian-American variants and the L83V (E-G350) variant of E6 upregulate cell-cell adhesion, protein and tyrosine kinase activity. As described in the section above, many of the proteins E6 interacts with are involved with adhesion and kinase activities. Another important finding was L83V's upregulation of N-cadherin (cadherin 2), AA-a's upregulation of cadherins 18 and 6, as well as AA-c's upregulation of cadherin 9. Furthermore, recent clinical studies confirm that the D sub-lineage (including AA) is most associated with cervical carcinomas in the Americas (Ortiz-Ortiz et al. 2015; Clifford et al. 2019).

In recent years, our lab has made significant strides to unravel physiological, epidemiological and molecular differences further in AAE6 and EPE6 transduced cell lines (Zehbe et al. 2009; Richard et al. 2010; Niccoli et al. 2012; Togtema et al. 2015; Jackson et al. 2016; Cuninghame et al. 2017). Our group previously found signaling pathway differences between variant E6 cell lines such as increased hypoxia inducing factor 1-alpha (HIF-1α) expression within Asian-American E6 cells resulting in dysregulation of cellular metabolism similar to the Warburg effect. Increased HIF-1α levels were caused by the activation of MAPK/ERK1/2 signaling pathway (Richard et al. 2010; Cuninghame et al. 2017). Our research has also found that AAE6 possesses more effective migration ability

compared to EPE6 (Niccoli et al. 2012). Such abilities could further explain increased viral persistence and increased CIN grade II to III development as described previously. Our lab also identified that AAE6 was capable of immortalizing and transforming PHFKs in the absence of E7 whereas EPE6 could immortalize but not transform PHFKs without E7 (Togtema et al. 2015).

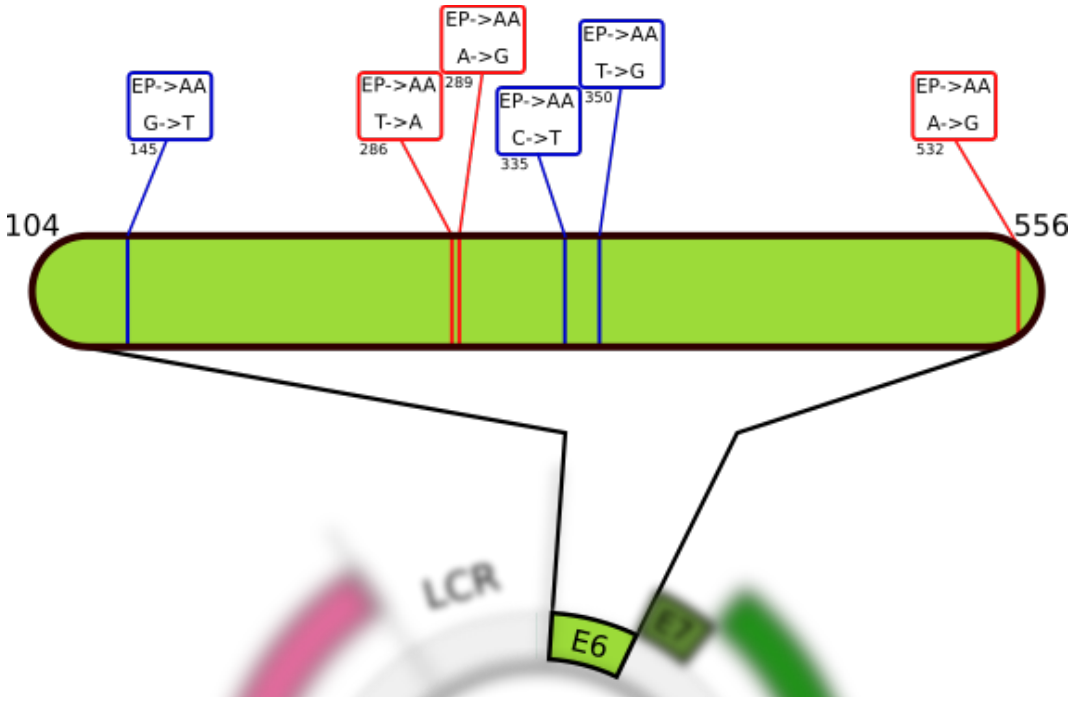

Figure 2 – To scale depiction of HPV16 E6 SNPs between AAE6 and EPE6. Lines in blue indicate a SNP that results in an amino acid change (missense mutation), whereas lines in red indicate no amino acid changes at that particular site (nonsense mutation). SNPs resulting in amino acid changes from EPE6 to AAE6 (Q14H, H78Y, and L83V) are found at nucleotide (NT) positions G145T, C335T, and T350G. SNPs that do not result in any change in amino acids are found at NT positions: T286A, A289G, and G532A. This correlates with SNPs that are found in Zehbe et al. 1998. Developed using Inkscape™ version 0.92.2 on original image in *Figure 1.*

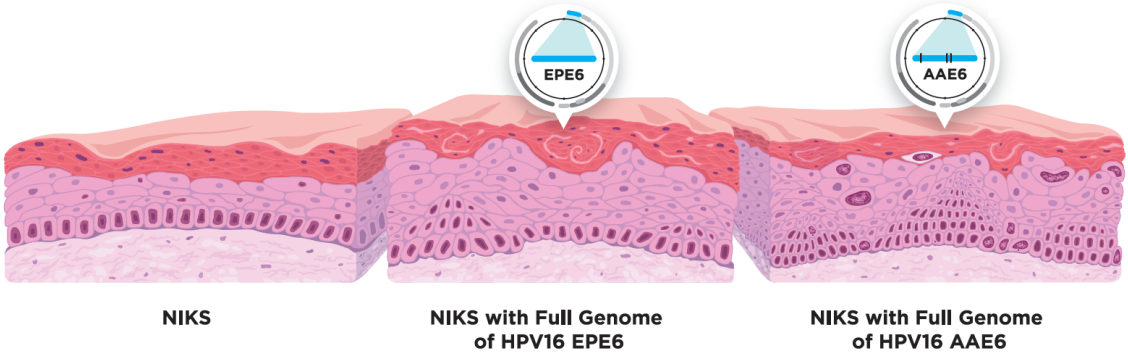

Figure 3 – Near-diploid immortalized keratinocytes (NIKS) grown on top of a dermal equivalent (consists of a collagen matrix and embedded fibroblasts) demonstrating differences in cellular dysplasia due to 16 E6 variant expression. AAE6 (right) shows increased dysplasia compared to raft culture of EPE6 (centre). Adapted from Jackson et al. 2016, originally published in *BMC Genomics*, and free-to-use with attribution (CC BY 4.0).

At the transcriptome level, our lab has identified a unique transcriptional profile for the AAE6 variant of HPV16 E6 allowing for a proliferating phenotype to be exhibited within the epithelium. From this, our group identified increased levels of proliferation, resulting in a significant reduction in chromosomal stability within the AAE6 variant, a common characteristic of cancer (Jackson et al. 2016).

1.6 Protein-Protein Interaction Methodology

The most common methods used to experimentally identify potential protein-protein interactions are through binary (direct) and co-complex (direct and indirect) interactions (BI and CCI respectively) (De Las Rivas, 2010). Binary interactome methods detect physical and direct protein-protein interactions (PPI). Examples of Binary techniques are: yeast two hybrid (Y2H) or membrane yeast two hybrid (MYTH), Luminescence-based mammalian interactome mapping (LUMIER), mammalian protein-

protein interaction trap (MAPPIT) and mammalian membrane two hybrid (MaMTH, Snider et al. 2015). The most common binary method used in E6 interactome studies is the Y2H assay. The Y2H assay works by developing a bait (target) protein fused to a transcriptional activating domain (AD) and a library of complementary DNA (cDNA) fused to a DNA binding domain (BD). If a protein containing a BD comes in contact with a bait protein bearing the AD, expression of a reporter gene will occur, and the interaction is detected (Snider et al. 2015). Unfortunately, the Y2H method may cause a significant number of false positives and negatives. One reason for this is posttranslational modifications for specific proteins required to be functional which does not occur in an environment where Y2H are conducted (Rao et al. 2014). A reason for false negative findings is due to the requirement that both proteins must have access to the nucleus. Since many proteins can only be found in cellular organelles such as the mitochondria it is unlikely that any expression of a reporter gene will be observed (Snider et al. 2015). As such, co-complex approaches are progressively becoming preferred over binary methods to conduct interactome studies (Ronco et al. 1997, White et al. 2012; Oliveira et al. 2018).

Co-complex methods consist of tandem affinity purification-mass spectrometry (TAP-MS), and affinity purification mass spectrometry (AP-MS) (Howie et al. 2011, White et al. 2012, Rao et al. 2014). TAP-MS is an affinity purification method consisting of protein complex interaction with a TAP-tagged protein and two purification steps. TAP-tagged proteins initially are bound to an IgG coated support via *Staphylococcus aureus* protein A (ProtA) and washed before cleavage of the TEV protease cleavage site. A subsequent purification step of the eluate is completed via binding to calmodulin coated supports in the presence of calcium (Ca^{2+}). A final elution is done using ethylene glycol-

bis(ß-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA) and the resultant eluate can be analyzed by MS (Puig et al. 2001).

One of the most common affinity purification methods is co-immunoprecipitation (Co-IP). Co-IP is a common *in vitro* protein interaction method between protein complexes and an antibody. The protein complex contains an antigen enabling the binding of a target protein to a solid support such as magnetic or agarose beads (Yaciuk 2007). The collection of agarose beads is done by centrifugation while collection of magnetic beads is done by using a strong magnetic stand. Subsequent washes remove any nonspecific protein interactions. Finally, protein complexes (containing bait and prey) are eluted from the solid support by different elution methods. Harsh elutions such as SDS buffer can strip beads of all bound proteins but suffer from antibody contamination as those are also released from the beads. Milder elutions such as synthetic peptide, or mild acidic (0.2 M Glycine pH <3.0) elutions can reduce antibody contamination while removing protein complexes from the beads (ThermoFisher Sci. Overview of the Immunoprecipitation (IP) Technique). Eluting with mild solutions does have downsides however, as stronger bait/prey interactions may not release from the beads leading to loss of interacting protein. One advantage of the Co-IP method is that interactions can be identified from proteins present in their native form or allow for posttranslational modifications (PTM). However, one should acknowledge the potential for non-specific binding between proteins and the antibody, or solid support. Therefore, a control such as non-transduced cell lysate preferably is used (Yaciuk 2007, Rao et al. 2014). Once Co-IP is complete, pulled-down proteins can be identified by mass spectrometric methods.

In this thesis, the interactome of HPV16 E6 will be studied. To do this, a co-complex method known as co-immunoprecipitation was used to identify potential PPIs. A list of all known interactors to HPV16 E6 can be found in *Table 1*. This table also mentions the methods used to identify and confirm each interactor that will be discussed in further detail below.

MS is a rapid, highly sensitive method to identify complex protein samples (Aebersold and Mann 2003). The most common approach to mass spectrometric protein identification is "bottom-up". Bottom-up proteomic approaches first digest the protein into small oligopeptide fragments by proteases. The most common protease is trypsin which cleaves the carboxy terminus of lysine or arginine. Depending on the proteins being analyzed, it may be favourable to use another protease such as: ArgC, chymotrypsin, LysC or LysN. Each protease has a unique specificity for a particular cleavage site within an amino acid sequence (Rogers and Bomgarden 2016). The resulting oligopeptides can then be separated and analyzed by two common methods explored in this thesis: liquid chromatography electrospray ionization (LC MS/MS) and matrix assisted laser desorption ionization mass spectrometry (MALDI MS/MS).

The principle of both MS methods is similar. That is "to measure the mass-to-charge (m/z) ratio of ionized peptides and identify the number of peptides that possess this m/z value" (Aebersold and Mann 2003). The way how each technique accomplishes this, is different. MALDI MS/MS utilizes a high-powered laser to ionize peptide samples. Samples must be dried and placed in a crystalline matrix to allow for proper analysis. The samples will then travel through a single or multiple time of flight mass separators (TOF) to calculate the mass of the ion produced. This is accomplished by separating ionized

peptides based on their differences in kinetic energies. Heavier ionized peptides will travel slower than ionized peptides with a low molecular weight (Aebersold and Mann 2003). the charge of the ionized sample is calculated by the detector and analyzer producing a m/z ratio (Aebersold and Mann 2003). LC MS/MS first separates proteins within a mobile liquid phase using liquid chromatography (LC) or high-performance liquid chromatography (HPLC). Separated proteins are converted into ions using electrospray ionization. Electrospray ionization is the process of ionizing the samples using a high voltage source to convert the solution of proteins into ions (Abersold and Mann 2003). For greater sensitivity, a nanoelectrospray can be used (Maziarz et al. 2000; Michalski et al. 2011). The ions then travel to a mass detector which can be a TOF sensor or an orbitrap analyzer (Aebersold and Mann 2003; Michalski et al. 2011). The benefit of an orbitrap analyzer is that it is capable of accurately analyzing most peptides large or small. Finally. the samples continue to a detector which will calculate the final m/z ratio (Aebersold and Mann 2003; Michalski et al. 2011).

Both MS methods have advantages and some drawbacks. For example, MALDI MS/MS is excellent for analyzing simple samples with high resolution and accuracy, while LC MS/MS can better handle complex protein mixtures (Aebersold and Mann 2003). One problem that all mass spectrometers face and continuously attempt to overcome is under sampling. Under sampling occurs when not every peptide injected into the mass spectrometer is detected and analyzed. This causes many proteins (~30 %) to be missed and fail to be identified (Pagala et al. 2015). Many experiments perform analyses with multiple replicates to reduce under sampling (Pagala et al. 2015). Biological replicates are commonly used for analysis of proteomic datasets. Many groups use at least two biological

replicates to produce accurate mass spectrometric datasets (White et al. 2012; Mellacheruvu et al. 2013).

Table 1 – Reported cellular targets of the human papillomavirus type 16 E6 protein (1991 to 2018)

Protein Name	Accession ID	Binding Mechanism	Identification Method	Mass Spectrometry	Outcome of Interaction	HPV16 E6 Sub-lineage	Reference
(1) Transcriptional adapter 3 (TADA3) alias hADA3	**O75528**	Targets hADA3 to prevent co-activation of P53	Yeast two-hybrid; confirmed with GST pull-down	None	Inactivated by high risk E6; coactivator for P53-mediated transactivation of target promoters and P53 stabilization	Not specified; most likely EP/A1 following publications back to Halbert et al. 1992; used E6 mutants F2V, 8S9, A10T, Y54H	**Kumar et al. 2002**
(2) Bcl-2 homologous antagonist/killer (BAK1)	**Q16611**	Targets BAK1 through E6AP (UBE3A) known to interact with it; has **zinc**-binding properties	**GST pull-down; originally described in Thomas and Banks, 1998**	None	Part of the pro-apoptotic Bcl-2 family; degraded by E6 thereby preventing BAK-induced apoptosis.	Not specified	**Thomas and Banks 1999**
(3) BRCA1-associated RING domain protein 1 (BARD1)	**Q99728**	BARD1 binds to E6 through E6's two zinc finger motifs with zinc finger 1 (AA 30-66) being the most important region for binding; has **zinc** finger, and **metal ion**-binding properties	Yeast two-hybrid; confirmed with immunoprecipitation	None	Tumour suppressor through apoptotic signaling; inhibits E6, and stabilizes P53	Not specified; mutants Δ9-13, Δ101-105, Δ111-115, Δ138-142, 45Y/47Y/49H, R39G, L50G, D120G, K115E, C66G, C136G, C66G/C136G	**Yim et al. 2007**
(4) Breast cancer type 1 susceptibility protein (BRCA1)	**P38398**	Targeted via zinc finger domains of E6; has **zinc** finger, and **zinc ion**-binding properties	GST pull-down	None	Inhibits telomerase activity; inactivated by E6	Not specified; most likely EP/A1 based on BLAST comparison; gift from Peter Howley; they used mutants at both zinc finger domains: C66, 136G	**Zhang et al. 2005**
(5) Golgi-associated PDZ and coiled-coil motif-containing protein (GOPC) alias CAL	**Q9HD26**	The **PDZ** domain of CAL interacts with the **PDZ**-binding motif of E6; interacts with the E6/E6AP complex enhancing proteasome degradation	GST pull-down	**MALDI-TOF**	Involved in the vesicular trafficking pathway between membranous organelles; E6 mediates proteasome degradation of CAL through E6AP	Not specified; most likely EP/A1 based on BLAST comparison; gift from Peter Howley; used "wildtype and mutants E6 T149I, L151V, ΔC1, ΔC4, L50G, C66G/C136G, D120G	**Jeong et al. 2007**
(6) Histone-arginine methyltransferase CARM (CARM1)	**Q86X55**	Not acquired	*In vitro* methyltransferase assay	None	Activates P53 downstream genes by histone methylation; inhibited by E6; results in prevention of P53-responsive promotors and downregulation of P53 downstream gene expression	Not specified; His-tagged E6 and mutants E66C/6S F47R	**Hsu et al. 2012**
(7) CREB-binding Protein (CREBBP) alias CBP	**Q92793**	E6 binds to 3 regions on CBP and p300: C/H1, C/H3 and C-terminus. Binding is independent of P53; has **zinc**-binding properties	Co-immunoprecipitation	None	Acts as a co-activator for cell cycle regulation and differentiation; E6 inhibits the activation of P53 and NF-kB by CBP/p300	Not specified; E6 "wildtype and mutants Δ106-110 Δ133-137, 45Y/47H/49H, Δ123-127, Δ128-132, C66G/C136G	**Patel et al. 1999**
(8) Ubiquitin carboxyl-terminal hydrolase CYLD (CYLD)	**Q9NQC7**	Not acquired; has **zinc**-binding properties	Determined effect based on Electrophoretic mobility shift assay (EMSA) and NF-kB reporter gene assay	None	Tumour suppressor through negative regulation of the NF-kB pathway; E6 mediated ubiquitination and proteasomal degradation of CYLD resulting in hypoxia-induced NF-kB activation	Not specified; E6 nucleotide positions 32-617 based on Halbert et al. 1991	**An et al. 2008**
(9) Discs large homologue 1 (DLG1)	**Q12959**	E6 binds to second **PDZ** domain via C-terminal XS/TXV/L motif	Maltose-binding protein and GST pull-down	None	Involved in cell junctions and tumour suppression; E6 binds to hDLG promoting the transformation of cells	Not specified; "prototype" and mutants: E6L151del, E6L151V, E6L151I, E6L151P, E6L151LP, E6L151ETQV	**Kiyono et al. 1997**
(10) Discs large homologue 4 (DLG4)	**P78352**	Binds to C-terminus of E6 mainly through its second **PDZ** motif. Last amino acid change from leucine to valine changes affinity	Maltose-binding protein and GST pull-down	None	Suggested tumour suppressor function; E6 binds induces proteolytic degradation of DLG4	Not specified; "prototype" and mutants E6SDΔ151, E6-151V, E6-151, E6SAT (8S9A10T unable to target P53 but can bind E6AP), E6SAT-Δ151, E6SAT-Δ151V	**Handa et al. 2007**
(11) Ubiquitin-protein ligase E3A (UBE3A) alias E6AP	**Q05086**	Not stated in original paper; has **zinc** finger, and **metal ion**-binding properties	Co-immunoprecipitation and GST pulldown	None	E6 forms a stable complex with E6AP and mediates numerous downstream interactions such as proteolytic degradation of P53	Not specified; sequence based on Werness et al. 1990	**Huibregtse et al. 1991**

Protein Name	Accession ID	Binding Mechanism	Identification Method	Mass Spectrometry	Outcome of Interaction	HPV16 E6 Sub-lineage	Reference
(12) Reticulocalbin-2 (RCN2) alias E6BP, ERC55	**Q14257**	Chen et al. 1998 determined that E6BP residues 18 – 29 (AA Sequence: VSLEEFLGDY) is the binding site for E6	Yeast two-hybrid; confirmed with GST pull-down	None	A truncated form of ERC55; a putative calcium-binding protein; interacts with E6 to form a complex and is thought to be involved with E6 induced transformation and degradation of cellular proteins	Not specified	**Chen et al. 1995**
(13) Signal-induced proliferation-associated 1-like protein 1 (SIPA1L1) alias E6TP1	**O43166**	Binds to E6TP1's C-terminus at residue 194; **PDZ** domain within E6TP1 has little effect on binding with E6	Yeast two-hybrid; confirmed with GST pull-down	None	Possible tumour suppressor protein; degradation by E6 potentially alters G-associated protein signaling pathways	Not specified	**Gao et al. 1999**
(14) FAS-associated death domain protein (FADD)	**Q13158**	Targets the N-terminus with major binding contributions at Serine residue 10, 14, 16 and 18 and Glutamic acid residue 19; site-directed mutants enabled localization of E6-binding to N-terminal end of FADD	Mammalian two-hybrid to identify E6 binding to FADD and in vitro GST pull-down to confirm interaction deletion	None	Triggers apoptotic pathway and cell death; E6 accelerates depredation of FADD preventing transmission of apoptotic signals through FAS pathway	Not specified	**Filippova et al. 2004**
(15) Fibulin-1 (FBLN1)	**P23142**	No consensus binding motif identified; has **calcium**-binding properties	Yeast two-hybrid; confirmed with GST pull-down	None	Hypothesized E6 binds to and inhibits Fibulin-1 allowing for invasion and metastasis	Not specified; same plasmids as Chen et al. 1995	**Du et al. 2002**
(16) E3 ubiquitin-protein ligase HERC2 (HERC2)	**O95714**	Interaction with E6 is E6AP dependent; has **zinc f**inger, and **zinc ion**-binding properties	Liquid chromatography with tandem mass spectrometry (LC-MS/MS) and confirmed using co-immunoprecipitation with anti-HA magnetic beads	**LC MS/MS**	Can only interact with E6 through formation of complex with E6AP to result in degradation of HERC2	Not specified; E6 "wild type" and mutants E6 8S9A10S (AA changes at position 8, 9, and 10 (R-P-R to S-A-T)); E6 I128T (AA change at position 128 (I to T)); E6 Δ146-151 (deletion of final 5 AA in 151 residue sequence)	**White et al. 2012**
(17) Protein scribble homolog (SCRIB) alias hScrib	**Q14160**	Interacts with E6AP in the presence of E6, C-terminus of E6 recognizes **PDZ** domain of hScrib	GST pull-down	**MALDI MS/MS**	Tumour suppressor protein; ubiquitinated in the presence of both E6 and E6AP resulting in degradation and reducing integrity of tight junctions	Not specified; N-terminal FLAG tagged E6 described in Tong et al. 1997, E6 SAT (8-10) mutants used	**Nakagawa et al. 2000**
(18) Telomerase reverse transcriptase (hTERT)	**O94807**	Proximal promotor/regulatory regions (nt position -251 to -88 and +5 to +40) involved with 60% of E6-induced hTERT activity; has **DNA**-binding properties	Identified E6 as potential up-regulator for hTERT using telomerase activity assay; confirmed E6 upregulates hTERT mRNA using mRNA protection assay	None	E6 induces increased hTERT activity resulting in maintenance of telomere length	Not specified	**Veldman et al. 2001**
(19) Inhibitor of nuclear factor kappa-B kinase subunit beta (IKKB)	**O14920**	Not acquired; has **nucleotide (ATP)**-binding properties	Identified and confirmed by co-immunoprecipitation assay	None	E6 binding potentially intervenes with NF-kB activation during bacterial or viral infection or DNA damage	CaSki cells and plasmids containing HPV16 E6 however sub-lineage and mutants not specified	**Oliveira et al. 2018**
(20) Inhibitor of nuclear factor kappa-B kinase subunit epsilon (IKKE)	**Q14164**	E6 binds to a location within the first 160 residues of IKKε; has **nucleotide (ATP)**-binding properties	Identified and confirmed by co-immunoprecipitation assay	None	E6 binding potentially intervenes with type I IFN, NF-kB and STAT signaling	CaSki cells and plasmids containing HPV16 E6 however sub-lineage and mutants not specified	**Oliveira et al. 2018**
(21) Interleukin-1 receptor-associated kinase-like 2 (IRAK2)	**O43187**	Not specified, has **nucleotide (ATP)**-binding properties	Identified and confirmed by co-immunoprecipitation assay	None	E6 binding potentially intervenes with NF-kB activation during TLR9 signaling	CaSki cells and plasmids containing HPV16 E6 however sub-lineage and mutants not specified	**Oliveira et al. 2018**
(22) Interferon regulatory factor 3 (IRF3) alias IRF-3	**Q14653**	IRF-3 residues 109-149 contain the ELLG sequence most likely attributing to interaction with E6; ELLG is also present in the E6 binding domain of E6AP; has **DNA**-binding properties	Yeast two-hybrid; confirmed with GST pull-down	None	Transcriptional activator; interacts with E6 inhibiting transactivation of IFN-ß	Not specified	**Ronco et al. 1998**
(23) Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 1 (MAGI1) alias MAGI-1	**Q96QZ7**	E6 PBM interacts with **PDZ**1 of MAGI-1; has **nucleotide (ATP)**-binding properties	GST pull-down	None	Functions in signal transduction and likely a tumour suppressor protein; E6 targets MAGI-1 for proteasomal degradation	Not specified; "wild type" E6 and E6-T149D/L151A mutants	**Glaunsinger et al. 2000**

Protein Name	Accession ID	Binding Mechanism	Identification Method	Mass Spectrometry	Outcome of Interaction	HPV16 E6 Sub-lineage	Reference
(24) Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 2 (MAGI2) alias MAGI-2	**Q86UL8**	**PDZ**1 domain interacts with E6 most likely through the PDZ Binding Motif (PBM)	*In vitro* degradation assay used varying concentrations to observe degradation of MAGI-2	None	Functions in signal transduction; E6 targets MAGI-2 for proteasomal degradation	Not specified	**Thomas et al. 2002**
(25) Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 3 (MAGI3) alias MAGI-3	**Q5TCQ9**	**PDZ**1 domain interacts with E6 most likely through the PBM. PDZ1 domain of MAGI-3 alone is capable of inhibiting the interaction between E6 and MAGI-3; has **nucleotide (ATP)**-binding properties	*In vitro* degradation assay used varying concentrations to observe degradation of MAGI-3	None	Functions in signal transduction; E6 targets MAGI-3 for proteasomal degradation	Not specified	**Thomas et al. 2002**
(26) Mastermind-like protein 1 (MAML1)	**Q92585**	E6 interacts through acidic carboxy terminal LXXLL motif on MAML1	Yeast two-hybrid	None	Transcriptional co-activator; E6 represses transcriptional activity and downstream Notch signaling pathways	Not specified	**Brimer et al. 2012**
(27) DNA replication licensing factor MCM7 (MCM7) alias hMCM7	**P33993**	Deletion analysis found E6's N-terminal residues 1-91 bind to hMCM7 C-terminal residues 572-719; has **nucleotide (ATP)**-binding properties	Yeast two-hybrid	None	Component of replication licensing factors; E6 potentially interferes with its ability to associate with chromatin avoiding G1-phase arrest point	Not specified	**Kukimoto et al. 1998**
(28) Methylated-DNA--protein-cysteine methyltransferase (MGMT)	**P16455**	MGMT interacts with E6AP through L2G box sequence LLGXXXS/T; **PDZ** domain present shows potential binding with E6; has **zinc**-binding properties	GST pull-down; confirmed with immunoprecipitation and immunodepletion	None	DNA repair protein that protects against mutations; E6 promotes ubiquitination-dependent degradation	Not specified	**Srivenugopal et al. 2002**
(29) Myc proto-oncogene protein (MYC)	**P01106**	Binds to E6 in an E6AP dependent manner along with E2F1 transcription factor; has **DNA**-binding properties	GST pull-down	None	Transcription factor involved in cellular regulatory processes; binding of E6 results in shortening of MYC's half-life and accelerated degradation	Not specified	**Mesilaty et al. 1998**
(30) Myeloid differentiation primary response protein MyD88 (MYD88)	**Q99836**	Not acquired	Identified and confirmed by co-immunoprecipitation	None	E6 binding potentially prevents innate immune receptor signaling	CaSki cells and plasmids containing HPV16 E6 however sub-lineage and mutants not specified	**Oliveira et al. 2018**
(31) Nuclear factor kappa B subunit 1/2 (NFKB1/2) alias NF-kB	**P19838**/**Q00653**	Not acquired; has **DNA**-binding properties	Immortalization assays and dual luciferase assays	None	E6 increased NF-kB levels for baseline and TNF-α by 2- to 3-fold	Not specified	**Vandermark et al. 2012**
(32a) Transcriptional repressor (NFX1) alias NFX1-91	**Q12986-3 (NFX1-isoform 3)**	NFX1-91 is destabilized by the E6/E6AP complex at NFX-91's C-terminus; has **zinc**-finger and **DNA**-binding properties	Yeast two-hybrid; confirmed with co-immunoprecipitation and RT-qPCR	None	Transcriptional repressor of hTERT promoter; E6/E6AP complex destabilizes NFX1-91 through ubiquitination	Not specified; E6 mutants F2V, 8S/9A/10T	**Gewin et al. 2004**
(32b) Transcriptional repressor (NFX1) alias NFX1-123	**Q12986-1 (NFX1-isoform 1)**	NFX1-123 is stabilized in the presence of E6; Gewin et al. 2004 demonstrated its interaction with E6 in vitro using a yeast two-hybrid screen and *in vitro* binding assays; has **zinc**-finger and **DNA**-binding properties	Yeast two-hybrid; confirmed with co-immunoprecipitation and GST pull-down	**LC MS/MS**	Transcriptional activator of hTERT promoter; E6 may bring NFX1-123 to the hTERT promoter allowing for increased hTERT activation and overexpression	Not specified	**Katzenellenbogen et al. 2007**
(33) InaD-like protein (PATJ)	**Q8NI35**	E6 binds to the PDZ domain of PATJ (ETQL AA sequence)	Yeast two-hybrid; confirmed with co-immunoprecipitation	None	E6 binds to and targets PATJ for degradation independently of E6AP preventing the formation of a TJ-associated complex Par6-aPKC-PAR3 responsible for regulation of kinase activity and formation of tight junctions in polarized cells	Not specified	**Storrs and Silverstein 2007**
(34) Paxillin (PAXI)	**P49023**	Not acquired; E6 effect likely occurs downstream of paxillin tyrosine phosphorylation and is sensitive to status of actin polymerization; has **zinc**-binding properties	GST pull-down	None	Transduces signals from plasma membrane to focal adhesions and the actin cytoskeleton; E6 disrupts paxillin-mediated actin fibre formation	Not specified; E6 "wildtype" and mutants 141T, C128T, C17P, C20S, C53R, C90S, C93H, C93S, H105D, R116S, C124V, Δ127-137, Δ134-137	**Tong et al. 1997**

Protein Name	Accession ID	Binding Mechanism	Identification Method	Mass Spectrometry	Outcome of Interaction	HPV16 E6 Sub-lineage	Reference
(35) E3 ubiquitin-protein ligase PDZRN3 (PDZRN3)	**Q9UPQ7**	**PDZ**RN3 interacts with E6 within the PBM; **zinc** finger, and has **zinc**-binding properties	Yeast two-hybrid	None	When interacting with E6, PDZRN3 is targeted for degradation increasing STAT5-ß activation	Not specified	**Thomas and Banks 2014**
(36) Serine/threonine-protein kinase N1 (PKN1) alias PKN	**Q16512**	E6 binds to C-terminal region of PKN; has **nucleotide (ATP)**-binding properties	Yeast two-hybrid; confirmed with GST pull-down	None	Protein kinase that contributes to oncogenic transformation signal transduction pathways; PKN phosphorylates E6, which may allow for E6 to influence Rho-mediated signaling	Not specified; E6 "wildtype" and mutants F2L/Y43C, K34E, Q35R, I101V, Y84C, C103R/D120G/I128M/R131P, C111Y/Q116H/R117K/R124STOP, R124G/H126R, C63S	**Gao et al. 2000**
(37) Protein arginine N-methyltransferase 1 (PRMT1)	**Q99873**	Not acquired; has **amino acid**-binding properties	*In vitro* methyltransferase assay	None	Coactivates and methylates proteins; E6 reduces PRMT1-induced methylation of histone H4 at R3; resulted in reduced P53 transactivation	Not specified; His-tagged E6; used mutants E66C/6S F47R	**Hsu et al. 2012**
(38) Caspase-8 (CASP8) alias Procaspase-8	**Q14790**	Not acquired; both E6 full-length and truncated E6* can bind to procaspase 8	Mammalian two-hybrid; confirmed with GST pull-down, immunoprecipitation and co-immunoprecipitation	None	Initiator caspase in several apoptotic pathways; full-length E6 targets procaspase 8 for degradation and also decreases caspase 8 interaction with FADD and procaspase 8 dimerization; a truncated form of E6 was found to stabilize procaspase 8	Not specified	**Filippova et al. 2007**
(39) Tyrosine-protein phosphatase non-receptor type 3 (PTPN3)	**P26045**	Interaction between C-terminus of E6 and the PDZ domain of PTPN3; has **substrate**-binding properties	GST pull-down	None	Membrane-associated phosphatase; degraded by E6 which prevents tyrosine phosphorylation of growth factor receptors	Not specified; E6 "wildtype" and mutants E6ΔC, E6F2V	**Jing et al. 2007**
(40) Cellular tumour antigen P53 (P53)	**P04637**	Not acquired; has **zinc- and DNA**-binding properties	Immunoprecipitation	None	Tumour suppressor protein and negative regulator of proliferation; E6 binds to and degrades P53	Not specified	**Werness et al. 1990**
(41) Histone acetyltransferase p300 (EP300)	**Q09472**	Binding domain on E6 between residues 100-147. Binds to 3 regions on CBP/p300: C/H1, C/H3 and C-terminus; has **zinc**-binding properties	Co-immunoprecipitation	None	Coactivator important for cell differentiation and cell cycle progression; E6 prevents the activation of P53 and NF-kB by CBP/p300	Not specified; E6 "wildtype" and mutants Δ133-137, 45Y/47H/49H, Δ123-127, Δ128-132 and C66G/C136G	**Patel et al. 1999**
(42) Histone-lysine N-methyltransferase SETD7 (SETD7)	**Q8WTS6**	Not acquired; has **substrate**-binding properties	*In vitro* methyltransferase assay	None	Methylation of histones and non-histone substrates such as P53; inhibition by E6 results in the decrease of P53 stability and activity	Not specified; His-tagged E6 and mutants E66C/6S F47R	**Hsu et al. 2012**
(43) Telomerase reverse transcriptase (TERT)	**O14746**	Not acquired; has **magnesium**-binding properties	Modified telomere repeat amplification protocol assay (TRAP)	None	Synthesizes telomere repeat sequences; E6 causes ubiquitin-mediated degradation of a telomerase repressing protein; most likely E6 interaction with hTERT causes telomerase activation	Not specified; E6 "wildtype" and mutants Δ146-151, Δ118-122, and 8S/9A/10T	**Klingelhutz et al. 1996**
(44) Tax-1 binding protein 3 (TX1B3) alias TIP-1	**O14907**	TIP-1 interacts with E6 by binding to the PDZ binding region at E6's terminus	Yeast two-hybrid; confirmed with co-immunoprecipitation	None	TIP-1 interacts with E6 but rather than being degraded, it results in increased activation of RhoA kinase	Not specified	**Hampson et al. 2004**

Protein Name	Accession ID	Binding Mechanism	Identification Method	Mass Spectrometry	Outcome of Interaction	HPV16 E6 Sub-lineage	Reference
(45) Histone acetyltransferase KAT5 (KAT5) alias TIP60	**Q92993**	Charged residues of the N-terminus of E6; has **zinc** finger, and **metal ion**-binding properties	GST pull-down	None	Tumour suppressor protein; E6 destabilizes and degrades TIP60 promoting cell proliferation and cell survival	Not specified; mutant C33G	**Jha et al. 2010**
(46) Tumor necrosis factor receptor superfamily member 1A (TNR1A)	**P19438**	E6 binds to the C-terminal cytoplasmic tail of TNF R1	Immunoprecipitation; confirmed with mammalian two-hybrid	None	Involved in apoptotic signaling; E6 inhibits TNF induced apoptosis and formation of the death induced signaling complex (DISC)	Not specified	**Filippova et al. 2002**
(47) TNF receptor-associated factor 6 (TRAF6)	**Q9Y4K3**	Not acquired; has **zinc**-finger, and **zinc**-binding properties	Identified and confirmed by co-immunoprecipitation	None	E6 binding potentially de-regulates DNA damage response and host immunity	CaSki cells and plasmids containing HPV16 E6 however sub-lineage and mutants not specified	**Oliveira et al. 2018**
(48) TIR domain-containing adapter molecule 1 (TICAM1) alias TRIF	**Q8IUC6**	Not acquired	Identified and confirmed by co-immunoprecipitation	None	E6 binding potentially inhibits innate immune functions and antiviral responses	CaSki cells and plasmids containing HPV16 E6 however sub-lineage and mutants not specified	**Oliveira et al. 2018**
(49a) Tuberin (TSC2)	**P49815**	Residues 1-175 and 1251-1807 of TSC2 are required for binding to residues of 260-316 and 428-500 of E6AP	GST pull-down; confirmed with co-immunoprecipitation	None	Tumour suppressor; E6 binds to E6AP/TSC2 complex and targets TSC2 for degradation	Not specified; most likely EP/A1 based on BLAST comparison; gift from Peter Howley	**Zheng et al. 2008**
(49b) Tuberin alias TSC2	**P49815**	DILG and ELVG domains of Tuberin bind to E6 residues 78-104	Yeast two hybrid; confirmed with GST pull-down	None	Tumour suppressor protein; negatively controls proliferation; E6 binding causes ubiquitin mediated degradation of Tuberin	Not specified	**Lu et al. 2004**
(50) Ubiquitin carboxyl-terminal hydrolase 15 (UBP15) alias USP15	**Q9Y4E8**	Not acquired	Targeted mass spectrometry	**Targeted MS**	Interaction of E6 with active form of USP15 results in increased stability and increased E6 half-life	Not specified	**Vos et al. 2009**
(51) DNA repair protein XRCC1 (XRCC1)	**P18887**	E6 interacts with the N-terminus of XRCC1 (residues 107-170)	Yeast-two hybrid; confirmed with co-immunoprecipitation	None	Required for DNA repair and genetic stability; E6 inhibits XRCC1. and prevents ability to maintain genetic integrity and utilize DNA strand break repair mechanisms	Not specified	**Iftner et al. 2002**

1.7 Bioinformatic Interactome Analysis Software

Once proteins have been identified by MS, relationships between samples must be made. There are numerous tools available for researchers to analyse relationships between identified proteins. For instance, Reactome is an online database that can allow researchers to build biological process maps with identified proteins within the human genome. This database is constantly updated and monitored to provide users with the most up to date information possible (Fabregat et al. 2018). Mapped proteins can be analyzed using a graphical user interface enabling easy identification of what role a protein may play in the biological system (Fabregat et al. 2018). Such a tool can be used to make sense of differences in PPI between a viral protein and a host cell. Another useful database that helps visualize interacting proteins is the STRING database (Szklarczyk et al. 2017).

The STRING database enables researchers to identify PPIs that may exist within a particular dataset. These PPIs can be classified as direct or indirect based on a variety of experimental databases such as: the IMEx consortium and BioGRID. This database can provide insight into experimental, or interlog (similarity between orthologs) evidence as well as curated databases citing existence of an interaction between two proteins (Szklarczyk et al. 2017). Such a database can simplify the process of pathway building and PPI identification. Finally, the Protein Analysis Through Evolutionary Relationships (PANTHER) database builds functional relationships between proteins based on over 900 different genomes (Mi et al. 2013).

1.8 Hypothesis, Aims and Rationale

Hypothesis: It is hypothesized that the Asian American (D2/D3) variant will possess several different virus-host PPIs resulting in increased activity of carcinogenic molecular pathways.

Specific Aims: To identify differential PPIs, the pull-down of E6 binding partners using co-immunoprecipitation (Co-IP) was optimized and LC MS/MS was used to identify pulled-down proteins. Next, literature searches and bioinformatic methods were used to identify potential differing interactors between each variant through a combination of mass spectrometric and bioinformatic processes. Once differences between variant E6 interacting host proteins were identified, the potential effects of E6 on each interactor was investigated.

Decades of research helped unravel what makes E6 capable of immortalizing primary human foreskin keratinocytes (PHFKs) without the presence of other HPV16 early or late genes. We have a better understanding of E6's ability to activate and alter numerous molecular pathways such as: phosphatidylinositol-3-kinase (PI3K) and extracellular signal-regulated kinase/ mitogen-activated protein kinase (ERK/MAPK) (Reviewed in Vande Pol et al. 2013). However, the mechanisms underlying AAE6's efficiency at the protein level to promote carcinogenesis compared to EPE6 continues to elude researchers. This thesis attempted to develop a method to effectively pull down HPV16 E6 oncoproteins along with interacting host proteins regardless of variant type. This method will be optimized to perform the pulldown, sample preparation (lysis, inhibitors etc.), MS analysis (MALDI vs. LC MS/MS) and utilization of bioinformatic systems to identify novel binders or different binders. The combination of Co-IP, the most common method for identifying

new PPIs, and significant advances in MS technologies ensured that data produced was accurate and meaningful (Berggård et al. 2007). Co-IP was conducted on retrovirally transduced primary human foreskin keratinocytes (PKFKs) possessing E6 protein linked to a HA tag developed by a previous member of our lab (Niccoli et al. 2012). PHFKs were appropriate because as the host keratinocyte differentiates, upregulation of viral genes increases, and production of viral and capsid proteins occurred. Eventually, the cell cycle in non-dividing cells restarts and dysregulated growth begins, resulting in cellular immortalization, a hallmark of cancer (Stanley 2010; Hanahan et al. 2011). HA-tagged E6 proteins and their interacting protein partners were eluted using anti-HA tagged magnetic beads. SDS-PAGE electrophoresis and western blotting were used to visually confirm the presence of E6. As a control, the Co-IP was completed in tandem with PHFKs retrovirally transduced with only the HA tag.

To identify the binding partners of the different E6 variants, our group collaborated with Dr. Hoyun Lee (Health Sciences North, Sudbury, ON) whom we have worked extensively with in the past to conduct the peptide mass fingerprinting analysis (Richard et al. 2010; Niccoli et al. 2012) and purchased LC MS/MS services from Harvard University because of their outstanding work on *White et al. 2012*. By developing this method, I was be able to identify potential differences in E6 host protein interactions. Identifying differences in PPI between each variant, it will be possible to unravel potential tumourigenic mechanisms between AAE6 (D2/D3) and EPE6 (A1) variants resulting in altered abilities to promote cancer development within PHFKs.

2 Materials and Methods

As mentioned in *Section 1.8.* numerous methods were added, altered or removed throughout the duration of this thesis to optimize the pull-down of E6 along with associated binding partners by Co-IP. *Section 2.1* discusses the methods for cell culture to prepare samples for downstream applications. The initial cell culture protocol (*Section 2.8.1*) was used for MALDI MS/MS trials at the Health Sciences North Research Institute (HSNRI) while the optimized cell culture protocol (*Section 2.9.2*) was used for optimization and LC MS/MS trials at Harvard University. The protocol for Co-immunoprecipitation varied based on three main characteristics: Lysis, IP and Elution. Initial protocols for the immunoprecipitation, 2-D electrophoresis, 1-D electrophoresis and Tryptic in-gel digestion were conducted for all trials at HSNRI using MALDI MS/MS while all optimized protocols for immunoprecipitation were conducted at the Thunder Bay Regional Health Research Institute (TBRHRI) for all LC MS/MS trials at Harvard University.

2.1 Cell Lines

As described in Niccoli et al. 2012, PHFKs from the same donor were used for all experiments. All PHFKs were transduced with one of the following: AAE6 (variant D2/D3) or EPE6 (variant A1). To successfully observe differences in virus host interactions between HPV16 E6 variants, PHFKs were retrovirally transduced with one of the following: hemagglutinin (HA) tag (PHFK-HA), PHFK's transduced with carboxy terminus (C-terminus) HA-AAE6 (AAE6) and PHFK's transduced with C-terminus HA-EPE6 (EPE6). Transducing cells expressing the HA-tag, allows for specific immunoprecipitation of a target protein bound to the affinity tag (Free et al. 2009). Here,

the target protein was E6 and was tagged at the carboxy terminus with HA. The affinity tag was placed at the C-terminus because we wanted to reduce any disruption in interactions present at or near SNP sites for both AAE6 and EPE6. The consequence of placing an affinity tag at the carboxy terminus is the potential to disrupt PDZ interacting proteins (Personal communications with Dr. Elizabeth White). PHFK cells containing AAE6 or EPE6 vary in appearance from their primary cell counterparts (*Figure 4*). However, PHFK's transduced with C-terminus HA-L83VE6 (L83V) were initially used for troubleshooting experiments as they express constant levels of E6.

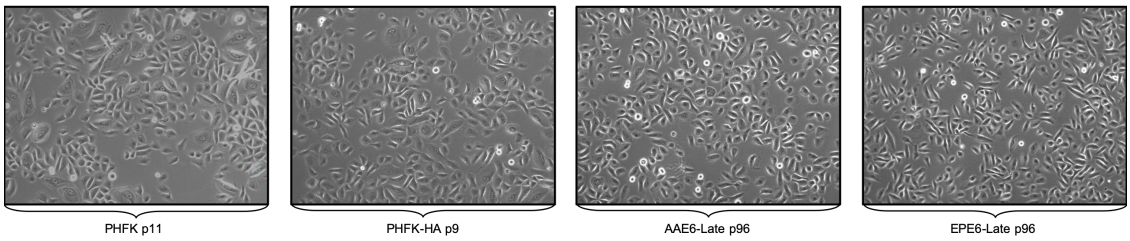

Figure 4 – Phase contrast images of AAE6 and EPE6 late passage samples and non-transduced PHFK cell samples (PHFK and PHFK-HA). All cell samples are at 75-80 % confluency and as mentioned on the previous page. There is a visual difference between HPV16 E6 transduced cell lines and PHFK/PHFK-HA primary cells. E6 transduced cell lines are smaller and have little variation in their shape and size compared to non-HPV16 E6 transduced PHFKs. Cells were imaged using an Axiovert 200 inverted phase-contrast microscope with halogen bulb on. Objective magnification was set to 10 X, photon of 1,0,4, and clear top filter. Exposure time for each image was 5 ms. Contrast of images were edited using Photoshop CC 2017 (Adobe, San Jose CA).

2.2 Cell Culture

All cells used in this thesis were cryogenically preserved in liquid nitrogen. To begin culturing cells for Co-IP and MS experiments, samples were removed from liquid nitrogen, thawed and placed in tissue culture flasks containing completed keratinocytes growth media (KGM, EpiLife medium completed with 60 μM calcium (EpiLife, Fisher Sci., Waltham MA, Cat# MEPI500CA); 1 X antibiotic/antimycotic solution (Anti/Anti,

Fisher Sci., Cat# SV3007901); and 1 X Human Keratinocyte Growth Hormone (HKGS, Fisher Sci., Cat# S0015)). Depending on the number of cells seeded, different flasks were used. For < 4 X 10^5 cells a sterile 25 cm^2 Falcon™ Tissue culture flasks (T25, Fisher Sci., Waltham MA, Cat# 10-126-10) containing 5 mL KGM was used. For 400 000 – 1 X 10^6 cells, a sterile 75 cm^2 Falcon™ Tissue culture flasks (T75, Fisher Sci., Cat# 1368065) containing 11 mL of KGM was used. When seeding 1 million or more cells, a sterile Nunc™ EasYFlask™ 225 cm^2 culture flasks (T225, Fisher Sci., Cat# 12-565-221) containing 33 mL of complete KGM was used. Cells were incubated at 37.0 °C in an atmosphere containing 5.0 % CO_2 and monitored daily to ensure constant growth and samples were free of any possible contamination (*Section 2.3*). Every 48 hours, growth media was removed and replaced with fresh room temperature KGM. Once cells reached 75-80 % confluency (*Figure 4*), KGM was removed and cells were treated with trypsin (Trypsin, Fisher Sci., Cat# SH3023602) and incubated for 10 minutes to allow all cells to detach from the flask. Trypsin neutralizing solution (TNS) (Cedarlane, Burlington ON, Cat# 080-100) was then added and cells were collected and centrifuged at 750 rpm for 5 minutes to pellet cells. Cells were either collected to be cryopreserved for future experiments or reseeded into new flasks for subsequent pull-down experiments. Although cells typically proliferate better when reseeded into their original flasks, the seeding of cells into new flasks was done to reduce risks of contamination and ensure similar growth rates when expanding.

2.3 Contamination Tests

As mentioned in *Section 2.2*, cells were observed daily for potential contamination. *Mycoplasma sp.* unfortunately, is a common contaminant in cell cultures. Mycoplasma contamination could alter the physical and physiological characteristics of cells as well as reduce accuracy and reproducibility of MS results (Nikfarjam et al. 2012). It was of utmost importance to ensure no such contamination occurred. Since mycoplasma cannot be seen with typical light microscopes, 100 000 cells were seeded every second passage into a sterile 35 mm Falcon™ tissue culture dish (Fisher Sci., Cat# CA25382-064) with 3 mL of complete KGM and one autoclaved cover slip (Fisher Sci., Cat# 12-550-15). Cells were cultured as described in *Section 2.2* and once 75-80 % confluent, were incubated in Carnoy's fixative (3 : 1 methanol to glacial acetic acid) twice (1st incubation – 5 minutes, 2nd incubation – 10 minutes) and allowed to air dry. Cover slips were gently removed from the culture dish and placed cell side down on a clean microscope slide with 25 µL mounting medium containing 4',6-diamidino-2-phenylindole (DAPI, Cedarlane, Burlington ON, Cat# H-1200). Slides were visualized and imaged (*Figure 5*) to confirm the presence or absence of mycoplasma.

2.4 PCR

Variants required confirmation of SNPs prior to use in any Co-IP or MS procedure to ensure no cross-contamination occurred across cell lines. As mentioned in *Section 1.4* there are six SNPs between the EP and AA variants. To confirm the presence of SNPs, polymerase chain reaction (PCR) was utilized similarly for all experiments. DNA extraction was conducted using the DNeasy Blood and Tissue kit (Qiagen, Hilden

Germany, Cat# 69504) 1 X 10^6 cells were thawed from -80 °C and resuspended in 200 µL sterile Dulbecco's PBS (dPBS, Fisher Sci., Cat# SH3002802) with 20 µL proteinase K. Addition of proteinase K was used to remove proteins which may result in degradation of DNA. RNase A (4 µL of 100 mg/mL) was added to the suspension of cells followed by addition of 200 µL Buffer AL (contains guanidine hydrochloride). Samples were vortexed thoroughly for 1 minute and incubated at 56 °C for 10 minutes. Following incubation, 200 µL of anhydrous ethyl alcohol was added to each sample and vortexed again for 1 minute. The solution containing DNA was pipetted into a DNeasy spin column and centrifuged for 1 minute at 6 000 X g. Two wash solutions resuspended in anhydrous ethanol labelled AW1 and AW2 were prepared. Next, 500 µL of wash buffer AW1 was added to the spin column and centrifuged for 1 minute at 6 000 X g. The flow-through was discarded, 500 µL of wash buffer AW2 was added to each sample and centrifuged for 3 minutes at 20 000 X g. To remove any residual flow-through, the spin column was placed into a new collection tube and centrifuged once more for 1 minute at 20 000 X g. The spin column was placed into a new 1.5 mL microcentrifuge tube prior to the addition of 100 µL buffer AE. After addition of buffer AE, the sample was incubated at room temperature for 2 minutes and centrifuged for 1 minute at 6000 X g. This elution step was done twice. Eluted DNA samples were pooled together and stored at 4 °C overnight labelled 'template DNA'.

DNA concentration and purity were measured using BioTek® Gen5™ software version 2.06. Master mixes for PCR contained the following: 1 U of Platinum Taq, 1 X PCR buffer, 1 mM $MgCl_2$ (Taq, PCR buffer and $MgCl_2$ came from same package, Fisher Sci., Cat# 10966018), 200 µM dNTP (10 mM, Fisher Sci., Cat# N8080260), and 0.5 µM

of both forward and reverse primers (Millipore Sigma). The sequences for the forward and reverse primers are shown beneath *figure 5*.

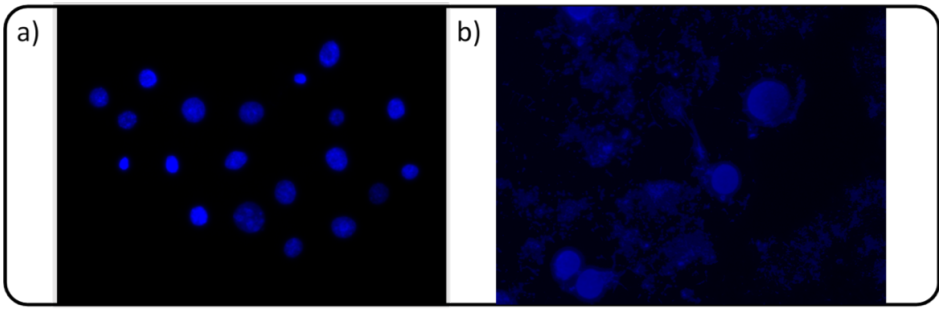

Figure 5 – Fluorescence microscopy image of E6 variant nuclei stained with 4',6-diamidino-2-phenylindole (DAPI). Cell samples in a) are AAE6 transduced PHFK cells that are mycoplasma free and no other signs of contamination are present. The absence of granular debris that is stained blue with DAPI indicates no mycoplasma contamination. In image b), a previous example from our laboratory shows EPE6 transduced cells with significant mycoplasma contamination surrounding most nuclei. Mycoplasma appears as small granular sized particles at 40 X magnification. Images visualized with an Axiovert 200 fluorescence inverted microscope at 40 X with an exposure of 10 ms.

Forward Primer (Initial Concentration 100 µM): 5' – CAATGTTTCAGGACCCACA – 3'

Reverse Primer (Initial Concentration 100 µM): 5' – GTTTCTCTACGTGTTCTTGA – 3'

Into the master mix was added 100 ng of template DNA and filled to a final volume of 25 µL using nuclease free dH_2O (General Electric Healthcare Life Sciences, Mississauga ON, Cat# SH3053802). Samples were run using an Applied Biosciences 2720 Thermal Cycler (ThermoFisher Sci., Cat# 4359659) using the cycles outlined in *Table 2*. Once the thermocycler entered the 'hold' step, 10 µL of post-PCR product was immediately run on a 2 % PCR certified agarose gel (BioRad, Cat# 161-3104) containing 0.01 % (v/v) 10 000 X GelRed Nucleic Acid gel stain (Biotium, Fremont CA, Cat# 41003) at a constant voltage of 100 V for 1 hour and 15 minutes with the remaining samples stored at 4 °C overnight.

Post electrophoresis, the gel was imaged using UVP imager for 0.1 seconds, and 0.15 seconds with 1 X 1 binning.

Table 2 – Procedure for Touchdown PCR for the identification of SNPs for all HPV16 E6 variants. The protocol was developed by a previous member of the lab and serves as a general template for touchdown PCR protocols. This protocol contains the same steps as a traditional PCR; however, the annealing step occurs at a higher temperature than the greatest melting point of the primer and slowly decreases until several degrees below the lowest melting point of the primers.

Number of Cycles	Cycle	Temperature (°C)	Duration (Minutes)
1	Polymerization	94	4
40 - (* Annealing temperature of 64 °C (cycle 1) decreased by 1 °C per cycle until 48 °C)	Denaturation	94	1
	Annealing*	64	1
	Extension	72	2
1	Final Extension	72	7
1	Hold	4	∞

Once the identified bands were in detected by electrophoresis (*Supplemental Figure S1*) purification of post-PCR product could take place. This was done using a QIAquick PCR Purification Kit (Qiagen, Cat# 28104). The remaining 15 µL of PCR product was combined with 75 µL of buffer PB and mixed by pulse vortexing sample before placing into a QIAquick spin column. The Spin column was inserted into a 2 mL collection tube and incubated at room temperature for 1 minute before centrifuging at 17 900 X g for 1 minute. Once flow-through was discarded, 750 µL of buffer PE containing anhydrous ethanol was added to the spin column and centrifuged at 17 900 X g for 1 minute. To remove any residual ethanol, a subsequent centrifugation was done for 1 minute. The spin column was then placed into a new 1.5 mL microcentrifuge tube and incubated for 2 minutes at room temperature with 30 µL of elution buffer (Buffer EB) added onto the membrane. After incubation, the DNA was eluted by centrifuging for 1 minute at 17 900 X g. For quality control, 10 µL of eluted DNA was run on a 2 % agarose gel for 1 hour and

15 minutes at a constant voltage of 100 V. The image was taken using the same methods as the unpurified post-PCR product. An image of the gel can be seen in supplemental data (*Figure S2*). The purified samples were then sent to the Lakehead University Paleo DNA laboratory for sequencing and SNPs were confirmed using MEGA7 (Kumar et al. 2016)

2.5 Cell Lysis

After collected cells were stored at -80 °C for up to 24 hours, the samples underwent protein extraction. Protein extraction begins with cell lysis using Mammalian Protein Extract Reagent (MPER, Thermo Sci., Waltham MA, Cat# 78501), supplemented with 1 X HALT™ phosphatase inhibitor cocktail (ThermoFisher Sci., Cat# 78428), 1 X cOmplete EDTA free protease inhibitor (Millipore Sigma., Cat# 4693159001), 1 mM phenylmethylsulfonyl fluoride (PMSF, ThermoFisher Sci., Cat# 36978) and 150 mM NaCl (Fisher Sci., Cat# BP358212). To each pellet, 3 mL of ice-cold supplemented lysis buffer was added. Gently using a pipette, the frozen cell pellet was resuspended in the lysis buffer and immediately placed on a tilting table on ice for 30 minutes. After incubation, the lysed cells were centrifuged twice at 14 000 X g for 15 minutes at a temperature of 4 °C. In between each centrifugation, the supernatant was carefully removed ensuring to not disturb the pellet and placed in two new clean low protein binding micro centrifuge tube (Fisher Sci., Cat# PI88379). After centrifugation, 5 µL of the resulting supernatant was diluted in 45 µL of ultrapure ddH_2O for quantification of protein (*Section 2.6*). The remaining extracted protein was stored at 4 °C for 30 minutes during quantification of protein.

2.6 Bradford Assay

Magnetic beads have a limited capacity to bind target proteins which varies from product to product. The Pierce™ Anti HA-Magnetic Beads (Fisher Sci., Cat# PI88836) have a working range from 500 µg to 1 mg of input protein per 25 µL of beads. For this reason, prior to incubating lysed cells on anti-HA magnetic beads, the concentrations and protein quantity of samples needed to be determined. This was done using either the Pierce™ bicinchoninic acid (BCA) Protein Assay (Fisher Sci., Cat# PI23225) or the Detergent Compatible Protein Assay (*DC* Protein Assay Reagents Package, Bio-Rad Laboratories Inc., Hercules CA. Cat# 5000116). The BCA assay is a highly sensitive method for quantifying protein concentrations through the reduction of Cu^{2+} to Cu^{1+} within an alkaline medium (Smith et. al. 1985). Bovine serum albumin (BSA) standards were prepared with ratios according to the manufacturer's protocol. A total volume of 6 mL working reagent (WR) was prepared using 50 : 1 ratio of Reagent A (sodium carbonate, sodium bicarbonate, bicinchroninic acid and sodium tartrate in a 0.1 M sodium hydroxide solution) and B (4 % (v/v) cupric sulfate solution). Into a 96 well plate, 25 µL of protein extract and BCA standard was added to each well and immediately after, 200 µL of working reagent was added to each well and incubated for 30 minutes at 37 °C. The colour change is caused by the chelation of one Cu^{2+} to two molecules of BCA (Smith et. al. 1985). The assay was done with 1 : 10 diluted samples in MPER. Once again 25 µL of samples and standards were combined with 200 µL of WR and incubated for 30 minutes at 37 °C. Incubated samples were placed in a Synergy HTX Multi-Mode Reader (BioTek Instruments, Inc., Winooski VT, Cat# S1LFTA-1340001) and absorbance was measured

at 562 nm using BioTek Gen5 (BioTek Instruments, Inc., Cat# S1LFTA-1340001) software.

Gen 5 is a microplate reading software developed by BioTek Instruments Inc.™ that permits the analysis of nucleic acids, and quantification of protein samples. Gen 5 software was used to generate a linear curve of protein standards based on their absorbances from the BCA assay. A linear slope equation was calculated by protein standard absorbance values and presented in the form "Y=mX+B". The values for each variable were as follows: Y – sample/standard absorbance, X – Protein concentration (µg/mL), A – slope of curve generated, B – Y-intercept. By inserting the calculated absorbances ('Y') for each sample, the concentration was determined. This method of protein quantification was used for initial pull-down attempts at the Health Sciences North Research Institute.

The DC Assay is an alternative method to quantify proteins using a reaction similar to the Lowry assay. The Lowry assay is based on a two-step reaction where proteins tyrosine, tryptophan as well as cystine, cysteine and histidine initially react with copper in an alkaline (pH >7.0) solution (Lowry et al. 1951). The DC assay uses a proprietary alkaline copper tartate solution (Reagent A) to treat proteins with copper. In protein solutions containing detergents such as Tween 20, sodium dodecyl sulphate (SDS) or nonidet P-40, 10 µL of surfactant solution containing SDS must be added to every 500 µL of reagent A to create reagent A'. Copper-treated proteins obtain their maximum colour through via reduction of copper bound to proteins (Lowry et al. 1951). The DC assay utilizes a proprietary folin reagent (reagent B) to perform this reducing step. In a 96-well plate, 5 µL of dilute extracted protein and known bovine serum albumin protein standards (concentration range: 0.2 µg/µL – 1.5 µg/µL) was added to individual wells. Next, reagent

A' was prepared by combining 10 µL of reagent S and 500 µL of reagent A in a separate 1.5 mL microcentrifuge tube. For each well containing either protein extract or standard, 25 µL of reagent A' was added. Finally, 200 µL of reagent B was added to each well and the samples were incubated at room temperature for 15 minutes. After incubating, the 96-well plate was placed into a PowerWaveXS™ absorbance reader (BioTek Instruments, Inc.) The final protocol utilized the DC Assay as it was accurate and easily accessible within our laboratory.

KC Junior is a legacy software provided by BioTek upon purchase of the PowerWaveXS™. This software automatically calculates the concentration of proteins by analyzing the absorbance at a wavelength of 750 nm. Similarly, to Gen5 software, a linear concentration curve is generated by known protein standards and any outliers were removed. Unlike the Gen5 software, there was no need to calculate the equation of the curve because the concentration of protein was automatically determined.

2.7 Western Blotting

The following antibodies: anti-HA 1 : 1000 (monoclonal mouse (mAB), Abcam [HA, C5], Cat# ab181818, 1 mg/mL), anti-E6AP 1 : 1000 (polyclonal rabbit (pAB) [H-182], Santa Cruz Antibodies, Dallas TX, Cat# SC-25509, 200 µg/mL), anti-P53 1 : 500 and 1 : 1000 (monoclonal rabbit (RabmAB) [E26], Abcam, Cambridge UK, Cat# ab32389), anti-HA conjugated to horseradish peroxidase (HRP) 1 : 1000 (HA-HRP, rat mAB [3F10], Roche, Cat# 12013819001, 25 U/mL), anti-actin 1 : 1000 (goat polyclonal (pAB) [I-19], Santa Cruz Antibodies, Dallas TX, Cat# SC-1616, 100 µg/mL), anti-AIP2 (WWP2) 1 : 1000 (rabbit polyclonal (pAB) FisherSci., Cat# PA5-42294), as well as secondary

antibodies: anti-mouse conjugated to HRP 1 : 2000 (Jackson Immunoresearch Laboratories Inc., West Grove, PA, Cat# 115-035-062), anti-rabbit conjugated to HRP 1 : 1000 (ThermoFisher Sci., Cat# SA1-200, 40 μg/mL), anti-goat to HRP 1 : 1000 (Jackson Immunoresearch Laboratories Inc., Cat# 705-035-147) were used.

To each sample, 6 X SDS loading was added until a final concentration of 1 X SDS was reached. Samples were boiled for 10 minutes to complete protein reduction. Reduced samples were run on a mini 4-12% gradient gel (Bio-Rad, Cat# 4561094) for 1 hour and 15 minutes at a constant voltage of 120 V. Fifteen minutes prior to the completion of electrophoresis, a polyvinyl difluoride membrane (PVDF, Fisher Sci., Cat# PI88518) was cut to similar dimensions as the mini gel. Once the gel completed the run, it was carefully removed from the casing and immediately placed in ice-cold 1 X transfer buffer (100 mL 10 X transfer buffer (14.41 g glycine, 3.03 g Tris base, 200 mL methanol (Fisher Sci., Cat# A454-4), 700 dH_2O). At the same time, the PVDF membrane was transferred into ice-cold 1 X transfer buffer in a separate container and both the gel and membrane were incubated for 10 minutes. After incubation in transfer buffer, the gel and membrane were placed together in between blotting paper to assemble a transfer sandwich and inserted into an ice-cold tank containing 1 X transfer buffer continuously mixed with a magnetic stir bar. The protein was transferred to the membrane for 1 hour at a constant voltage of 100 V and maximum amperage of 0.35 A.

After the transfer, the membrane was removed from the sandwich and incubated for 10 minutes with 0.05% TBS-T at room temperature. During the incubation, blocking buffer was prepared (5% (w/v) powdered skim milk (Safeway Pleasanton CA) in 0.05% TBS-T). Post incubation in 0.05 % TBS-T, the membrane was transferred into a 50 mL conical tube

containing 15 mL of prepared blocking buffer. The sample was incubated on a rotating rack for 1 hour at room temperature before being transferred into another 50 mL conical cube containing primary antibody in 5 mL of blocking buffer. Once placed in primary antibody, the sample was stored on a rotating rack overnight at 4 °C.

The next day, samples were removed from primary antibody and washed three times for 5 minutes with 0.05% TBS-T at room temperature. If the primary antibody was not conjugated to HRP, the sample was placed in another 50 mL tube containing secondary antibody in 5 mL of blocking buffer for 1 hour at room temperature. After the final incubation in secondary antibody, the samples were washed once again three times in 0.05% TBS-T for 5 minutes. Immediately after the final wash, excess TBS-T was removed by gently placing on a dry paper towel and then placing on plastic wrap. To produce chemiluminescence, samples were incubated in equal volumes of peroxide reagent and luminol/enhancer reagent (Clarity Western ECL Substrate, BioRad, Cat# 1705061) for 1 minute. Immediately following incubation, the samples were imaged at 1 X 1 binning for 0.1 seconds, 1, 5, and 20-minute intervals.

2.8 MS1 Protocol

2.8.1 Cell culture

All primary cells (PHFK, PHFK-HA) and cell lines (AAE6-Late, EPE6-Late) were thawed from liquid nitrogen and seeded (600 000 cells) into sterile T75 flasks containing 11 mL of complete KGM. Cells were incubated in the same manner as described in *Section 2.2*. Once cells reached 75-80 % confluency (*Figure 4*), KGM was removed and cells were washed with 5 mL of dPBS prior to treatment with 3 mL trypsin. Once all cells detached

from the base of the flask, trypsin was neutralized with TNS and centrifuged for 5 minutes at a speed of 750 rpm and temperature of 4 °C. Post centrifugation, cells were resuspended in 2 mL of complete KGM and 500 000 cells were collected for storage in liquid nitrogen while 1 800 000 cells were seeded equally (600 000 cells per flask) into three new T75 flasks containing 11 mL of complete KGM. Reseeded cells were cultured in the same manner as above and centrifuged at the same speed for the same duration. However, instead of resuspending cells in complete KGM, 2 mL of dPBS was used to resuspend cells and 600 000 cells were reseeded in one new T75 while remaining cells were centrifuged again as described above. After centrifugation, cells had dPBS removed using suction and flash frozen using liquid nitrogen and stored at -80 °C for up to one month for subsequent Co-IP experiments.

2.8.2 Cell Lysis

Collected cells for co-immunoprecipitation (Co-IP) were stored at -80 °C prior to use. Immediately after cells were removed from -80 °C, mammalian protein extract reagent (MPER, Thermo Sci., Waltham MA, Cat# 78501) supplemented with protease inhibitors (1 X protease inhibitor cocktail (PIC, Millipore Sigma, Burlington MA, Cat# P8340-1 ml)), phosphatase inhibitors (2 mM Na_3VO_4 (Millipore Sigma, Cat# 22,059-0), 50 mM NaF (Millipore Sigma, Cat# S-7920), and 1 mM phenylmethylsulfonyl fluoride (PMSF, ThermoFisher Sci., Cat# 36978)) was added and gently mixed to increase cell lysis efficiency. Samples were incubated on ice with constant mixing on an orbital platform (The Belly Dancer, IBI Scientific, Dubuque IA, Cat# BDRAA115S) for 10 minutes and promptly centrifuged at 14 000 g for 15 minutes at 4 °C. The resulting supernatant was

stored at 4 °C for 30 minutes during quantification of protein concentration (*Section 2.6*) prior to incubation on magnetic beads.

2.8.3 Co-IP

Co-immunoprecipitation of beads will allow E6 to be selectively pulled down along with interacting proteins by binding of HA tag to the antibodies on a solid support (*Section 1.6).* Prior to binding of E6 and associated binding partners to Pierce anti-HA magnetic beads (Anti-HA magnetic bead, Fisher Sci., Cat# PI88836L), 25 µL of beads were pre-rinsed by incubating once in 125 µL of room temperature tris buffered saline with 0.05 % Tween - 20 (0.05% TBS-T, 3.03 g Tris base (Fisher Sci., Cat# BP1521), 8.76 g NaCl (Fisher Sci., Cat# BP358212), pH 7.5 in 1 L of dH_2O and 0.05 % Tween - 20 (Fisher Sci., Cat# BP337500)) for 30-seconds followed by a 1-minute incubation at room temperature using 1 mL of 0.05 % TBS-T. Following cleaning of beads, 600 µg of cell lysate was added to anti-HA magnetic beads, resuspended and incubated overnight on a rotating rack at 4 °C. Beads were gently washed with room temperature 0.05 % TBS-T 3 times for 30 seconds then eluted by incubation at 95 °C for 10 minutes in 1 X SDS loading dye with dithiothreitol (DTT, Fisher Sci., Cat# BP172-25) for trials using 1-D electrophoresis or in 8 M urea (BioRad, Cat# 161-0730) containing 400 µL Triton X-100 (4 % (v/v), United States Biological, Salem MA, Cat# T8655) 50 mM DTT and 0.0002 % bromophenol blue (Fisher Sci., Cat# FLB3925) for trials attempting to use 2-D electrophoresis. Eluted samples were stored in -80 °C overnight until ready for electrophoresis and western blot analysis.

2.8.4 2-D electrophoresis

Isoelectric focusing (IEF) has the potential to increase separation of pulled down proteins. This is done by first separating samples by their isoelectric point (PI) and then by their molecular weight. To prepare immobilized protein gradient (IPG) strips, six Immobiline DryStrip pH 3-10, 13 cm strips (General Electric Healthcare Life Sciences, Cat# 17600114) were trimmed to fit within the Protean IEF System cell (Bio-Rad Laboratories Inc., Cat# 165-4000). Prior to placement of IPG strips on IEF trays, eluted Co-IP samples were thawed from -80 °C and spread evenly into each IEF tray well. 25 µL of all eluted samples were added into individual wells. Rehydration solution containing 275 µL of 2D elution buffer, and 15 µL of Bio-Lyte 3/10 40 % (Bio-Rad Laboratories Inc., Cat# 165-1112) was applied on top of each well. IPG strips were carefully placed agarose side down such that the anode and cathode touched both sides of the sample. Any air bubbles underneath IPD strips were removed by applying gentle pressure to the plastic side of the gel. Finally, each strip within the wells was covered with 2 mL of mineral oil (Plus One dry strip cover fluid, General Electric Healthcare Life Sciences, Cat# 17-1335-01). The plastic tray cover was placed on the IEF tray and samples were left to incubate at room temperature overnight.

The following day, the IEF tray containing now rehydrated IPG strips with samples were placed onto the Protean IEF System. Sterile absorbent paper strips were dipped in MilliQ water and placed onto each anode and cathode that was in contact with an IPG strip. The paper was used to absorb salts that may accumulate at the electrodes. IEF was completed in several steps as instructed by the manufacturer (*Table 3*). The stepwise increase of voltage to each of the Immobiline strips prevented excessive heating and

denaturing of target proteins caused by minute concentrations of salts and ampholytes within the strip (Rabilloud et al. 2011). The extended voltage at 8 000 V was done to ensure finite focusing of all proteins within the sample containing an isoelectric point (PI) between 3 and 10. The final hold at 50 V was used to maintain pH gradient. Blot paper was changed every hour to maximize salt removal and improve protein migration.

Table 3 – Steps used during IEF of eluted samples. Gradual steps to increase the voltage were used to ensure that minimal heat was generated. Total runtime was 7 hours and 40 minutes.

Step	Voltage (V)	Increase method	Time (hours)
1	500	Step and Hold	1
2	1 000	Gradient	1
3	8 000	Gradient	3
4	8 000	Step and Hold	2 & 40 minutes
5	50 (hold)	Hold	99

SDS-PAGE gels (10 %) were cast and stored at 4 °C overnight until ready for use. 1 % Overlay agarose was produced using 25 mL dH_2O, 0.25 g agarose (Fisher Sci., Cat# BP161-100) and 0.01 g of Pierce Coomassie Brilliant Blue R-250 (Fisher Sci., Cat# PI20278). IPG strips were placed on the top of each gel and run for 2 hours at 300 V. Gels were subsequently stained using Thermo Scientific Pierce™ Silver Stain Kit (Fisher Sci., Cat# PI24612). To silver stain PAGE gels, they first were washed twice in 100 mL of MilliQ water for 5 minutes. Next, gels were fixed in a 30 : 10 ethanol to acetic acid solution with two 15-minute incubations. Post fixation, gels were washed in 10 % ethanol twice for 5 minutes and twice for 5 minutes using MilliQ water to remove any residual ethanol or fixative. A sensitizing working solution containing 1 : 500 silver stain sensitizer to autoclaved water was incubated on gels for exactly 1 minute. The sensitizing solution increases the proteins ability to bind silver ions (Steinberg, 2009). Proteins were impregnated with silver ions via 5-minute incubation in 1 : 50 silver stain working solution.

The development of silver stained gels is completed through the reduction of silver ions bound to proteins to metallic silver that is insoluble (Steinberg, 2009). Gels were developed by incubating in working developer solution (500 µL enhancer with 25 mL developer) using extended times as needed until desired bands appeared. To prevent over developing of silver stained proteins, samples were placed in a 5 % (v/v) acetic acid stop solution. Gels were imaged using a charge coupled device (CCD) camera as needed (FluorChemQ, Proteinsimple, San Jose CA, Cat# 92-14095-00). It should be noted that the process of 2-D electrophoresis was not continued following initial attempt as it was not an optimal method for downstream proteomic applications.

2.8.5 1-D SDS-PAGE Electrophoresis

Since 2-D electrophoresis did not successfully allow for efficient protein separation, 1-D electrophoresis was attempted. Eluted proteins were loaded into mini 12 % polyacrylamide gels with Precision Plus Protein Standard Dual Colour Ladder (Bio-Rad Laboratories Inc., Cat# 161-0374) and ran for 1 hour and 15 minutes at a constant voltage of 120 V. Gels were incubated twice for 30 minutes at room temperature with gentle mixing using a fixing solution (50 % (v/v) methanol and 7 % (v/v) acetic acid). Immediately following protein fixation, gels were stained overnight using Invitrogen™ Molecular Probes™ SYPRO™ Ruby Protein Gel Stain (Fisher Sci., Cat# S12000). SYPRO Ruby is a fluorescent stain with similar sensitivity to silver staining methods that provides increased downstream compatibility for MS analysis. Similarly, to coomassie colourimetric stains, the proteins non-covalently interact with chelate ruthenium (II) resulting in a luminescent stain that can be excited with ultraviolet (UV) light (Lauber et al. 2001; Steinberg 2009). Gels were washed in a solution composed of 10 % methanol and 7 % acetic acid twice for

15 minutes. The gels were then imaged using the FluorChemQ imager (Proteinsimple, San Jose CA, Cat# 92-14095-00) in the auto exposure and contrast setting with an excitation wavelength of 280 nm. Gel bands were cut using sterile Graham Field Single-Use Scalpels (Fisher Sci., Cat# 08-927-5A) on an ultraviolet (UV) transilluminator table and placed into autoclaved 1.5 mL microcentrifuge tubes. 1-D electrophoresis was utilized only for MALDI MS attempts since we could get improved identification of proteins by keeping the entire protein sample in elution buffer and not be limited by well size.

2.8.6 MALDI MS

Samples were digested using the Pierce™ In-Gel Tryptic Digestion Kit (ThermoFisher Sci., Waltham MA, Cat# PI89871). Before digestion, SYPRO stain was removed from gel slices by adding 200 µL of destaining solution (80 mg of ammonium bicarbonate in 40 mL 50 : 50 mixture of acetonitrile and ultrapure water) to each 1.5 mL tube and incubating twice for 30 minutes at 37 °C. As digestion solution (25 mM ammonium bicarbonate in ultrapure water) was able to be stored at 4 °C for up to two months, 15 mL was prepared. Next, gel slices were reduced in 30 µL of reducing buffer (3.3 µL tris[2-carboxyethyl]phosphine (TCEP) in 30 µL of digestion buffer per sample) and incubated for 10 minutes at 60 °C. The purpose of the reducing step by TCEP is to break disulfide bonds originally formed when cysteine is oxidized to form cystine (Muskal et al. 1990; Gundry et al. 2010). A 5 X alkylation solution (500 mM Iodoacetamide) was prepared and combined with digestion buffer at a 4 : 1 ratio immediately prior to the alkylation of reduced gel slices. To the cooled slices, 30 µL of alkylation buffer was added and incubated in the dark at room temperature for 1 hour. Alkylation of proteins prevents free sulfhydryl groups on cysteine amino acid residues from being reoxidized and form disulfide bonds

(Gundry et al. 2010). Post alkylation of samples, slices were washed two times with 200 µL of destaining buffer for 15 minutes at 37 °C. Gel slices were dehydrated by incubating in 50 µL of acetonitrile for 15 minutes at room temperature and air dried for 10 minutes. Trypsin working solution was prepared by ten-fold dilution in ultrapure water. Immediately prior to digestion, trypsin working solution was added to digestion buffer to create a 10 ng/µL activated trypsin solution. Gel pieces were swelled by submerging in activated trypsin and incubated at room temperature for 15 minutes. An additional 25 µL of digestion buffer was added to each sample and incubated at 30 °C overnight. The resulting digestion mixtures were placed in a clean 1.5 mL tube and prepared for analysis with MALDI MS.

Each spot on the MALDI plate had 1 µL of Pierce™ 0.1 % Trifluoroacetic acid (v/v) (ThermoFisher Sci., Waltham MA, Cat# PI85172) added as the matrix for samples. 1 µL of digested sample was added to each well ensuring to record which band was present at each spot on the metal plate. Plates were allowed to air dry prior to placement in the mass spectrometer. Metal plates were inserted into the vacuum chamber of a MicroMx MALDI TOF mass spectrometer (Waters, Milford MA, Cat# 201000178). Samples were run by Dr. James Knockleby and each sample had a minimum of 100 tandem MS/MS spectra obtained prior to continuing to the next sample. Raw spectra files were saved to a USB drive. Peak lists along with their respective intensities were saved in the form of a Microsoft Office word document.

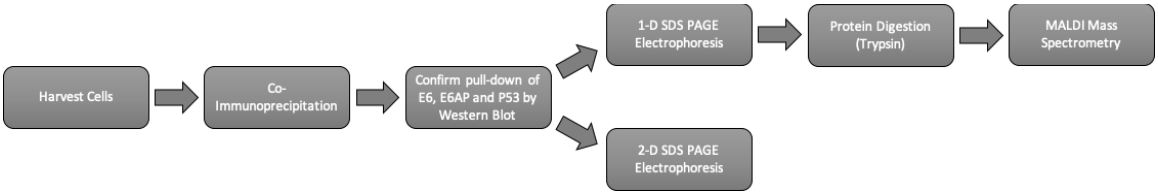

Figure 6 – Overview of main steps in MS1 protocol.

2.9 LC MS/MS

2.9.1 Cell Culture

Similarly, to *Section 2.8.1*, PHFK cells with or without transduced viral DNA genomes were used for LC MS/MS analysis. The following changes were made to accommodate increased cell production. 1.5 X 10^6 cells of each of the following: PHFK-HA, AAE6, EPE6 were seeded into sterile Nunc™ EasYFlask™ 225 cm^2 culture flasks containing 33 mL KGM. Media changes occurred every 48 hours as described previously while trypsinizing, pelleting and resuspension of pellets were done in the same manner as *Section 2.8.1* for each flask with a volume increase by a factor of 3 for KGM and a factor of two for trypsin and TNS (33 mL KGM, 6 mL trypsin and 18 mL TNS). After neutralizing trypsin activity, the cells were centrifuged in the same manner as *Section 2.8.1*, resuspended in 2 mL of complete KGM and seeded 50 % of each sample into two new T225 flasks. By expanding from one to 2 confluent T225 flasks, we had enough cells to seed 1.5 million cells into 9 new T225's (13.5 X 10^6 cells total). Cells would be once again grown until 75-80 % confluent as previously described and harvested using the methods described above. Once cells were centrifuged, they were resuspended in 2 mL of complete KGM. For each sample, 1.5 X 10^6 cells were seeded into nine sterile T225 containing 33 mL of fresh KGM. Since continued cell growth for subsequent pull-downs was necessary, one of the 9 T225 flasks were used for another expansion into two T225's then further expanded into 9 T225 flasks. Once cells seeded for pull-downs were 75-80 % confluent, four flasks from each sample were treated for 4 hours with 10 mL of fresh KGM containing 30 μM MG132 proteasome inhibitor in DMSO (MG132, Millipore Sigma., Cat# 474791-5MG or 474791-1MG). The other four flasks were treated with an equal volume of DMSO

(vehicle) in 10 mL of KGM. Following treatment, cells were collected as described above and stored in -80 °C overnight.

2.9.3 Cell Lysis

After collected cells were stored at -80 °C for up to 24 hours, the samples were lysed with the following changes. The composition of the lysis buffer changed to include MPER supplemented with 1 X HALT™ phosphatase inhibitor cocktail (ThermoFisher Sci., Cat# 78428), 1 X cOmplete EDTA free protease inhibitor (Millipore Sigma., Cat# 4693159001), 1 mM PMSF and 150 mM NaCl (Fisher Sci., Cat# BP358212). Each frozen pellet had 3 mL of ice-cold supplemented lysis buffer added. Gently using a pipette, the cell pellet was resuspended in lysis buffer and immediately placed on a tilting table on ice for 30 minutes. After incubation, the lysed cells were separated into two 2 mL low protein binding microcentrifuge tubes (Fisher Sci., Cat# PI88379). Samples were centrifuged twice at 14 000 X g for 15 minutes at a temperature of 4 °C. In between each centrifugation, the supernatant was carefully removed ensuring to not disturb the pellet and placed in new clean low protein binding microcentrifuge tubes. After centrifugation, 5 µL of the resulting supernatant was diluted in 45 µL of ultrapure ddH_2O for quantification (*Section 2.6*). The remaining extracted protein was stored at 4 °C for 30 minutes during quantification to reduce any protein degradation.

2.9.4 Co-IP

Co-IP was adjusted to allow samples to be eluted in a MS-compatible buffer at a high enough concentration to allow for identification, reduce leeching of antibodies from solid supports and maintain stability of proteins post elution. anti-HA magnetic beads were washed by resuspending twice in 1 mL ice-cold complete lysis buffer (Composition found

in *Section 2.9.3*) in sterile 15 mL conical tubes (Fisher Sci., Cat# 1495949B) followed by incubation on magnetic stands for 2 minutes and subsequent removal of supernatant. For each MG132 sample after beads were washed, 4 mg of extracted protein was loaded onto 80 µL of anti-HA magnetic beads and incubated on a rotating rack overnight at 4 °C. As the volume required to load 4 mg of protein onto the beads exceeded the volume capacity of the tubes, two aliquots containing 2 mg of protein and 20 µL of beads were used instead. Post-overnight incubation, the unbound protein was removed by incubating the sample on a magnetic stand for 4 minutes on ice and then carefully pipetting the supernatant. Using 1 mL ice cold supplemented lysis buffer without inhibitors, beads were washed three times for 5 minutes on a rotating rack at 4 °C. To remove all detergents and inhibitors, samples were subsequently washed with 1 mL sterile filtered 1 X PBS three times for 5 minutes on a rotating rack at 4 °C. In between all washes, samples were incubated on a magnetic stand for 2 minutes prior to removal of supernatant. This prevented aspiration of beads from tubes resulting in loss of sample. To elute proteins, samples were pooled together and incubated for 10 minutes in 100 µL 0.2 M glycine (Fisher Sci., Cat# BP3815) pH buffered to 2.5 at room temperature. After incubation in elution buffer, samples were placed on magnetic stands for 4 minutes and eluate was placed into a new 2 mL microcentrifuge tube. Immediately after elution, samples were neutralized using 12 µL of 1 M Tris-HCl (pH 7.5). Thirty µL of eluted samples were reduced using 7.5 µL 6 X SDS loading dye with DTT and heated at 95 °C for 10 minutes. All samples were then stored at -80 °C prior to western blotting (*Section 2.7*) or shipping for MS analysis.

2.9.5 LC MS/MS by Harvard Mass Spectrometry and Proteomics Facility (HMSPF)

Upon arrival at HMSPF, samples were reduced, alkylated and digested in preparation of LC MS/MS analysis. Similarly, to MALDI MS/MS sample preparation *(Section 2.8.6)* TCEP was used to reduce samples. To reduce samples using TCEP, 20 µL of each sample, 2.0 µL of 20 mM TCEP (Millipore Sigma., Cat# 75259-1G) in 50 mM triethylammonium bicarbonate [pH 8.0] (TEAB, Millipore Sigma., Cat# T7408) buffer was added. Samples were incubated at 37 °C for 1 hour and subsequently cooled to room temperature vortexed and pulse centrifuged to ensure samples remained at the bottom of each tube. Next, 2.2 µL of freshly prepared 40 mM iodoacetamide (Millipore Sigma., Cat# I1149-5G) in 50 mM TEAB was added to each sample for alkylation. Samples with added iodoacetamide were covered with tin foil and incubated at room temperature for 1 hour. Trypsin (1 : 50-1 : 100 [w/w], Promega Corp., Madison WI, Cat# V5111) was then added and incubated at 37 °C for 16 hours overnight. After incubation, samples had 1 µL of formic acid (Millipore Sigma., Cat# 56302-50ML) thoroughly vortexed and placed into a nanoACQUITY ultra performance liquid chromatography (UPLC) system (Waters, Cat# WTF-ARC-2998) to begin LC MS/MS protein identification. The mass spectrometer used for all LC MS/MS tests was an Orbitrap Elite™ Hybrid Ion Trap Orbitrap Mass Spectrometer (ThermoFisher Sci., Cat# IQLAAEGAAPFADBMAZQ).

After each sample was placed into the mass spectrometer the spectra obtained were input into Proteome Discoverer version: 2.4.0.292. A complete description of the workflow used at HMSPF can be found in *Supplemental Data S3 and S4*. Three databases were used for identification of proteins: Uniprot Human Proteome Database; a custom curated common contaminant database; and a user defined database containing the AA sequences

for HPV16 E6 (AAE6 and EPE6). Proteins were identified and emailed in the form of a heatmap within Microsoft Excel.

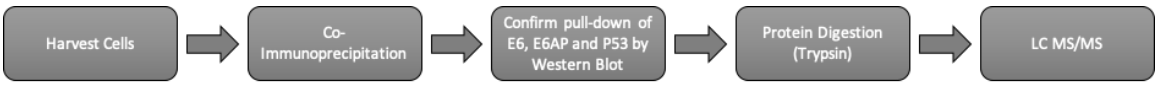

Figure 7 – Overview of main steps in MS2 protocol.

2.10 Identification of Proteins for MALDI MS/MS

2.10.1 MALDI MS/MS proteins using Peptide Mass Fingerprinting

Generated spectra peak lists from all MALDI MS/MS trials were entered into the Mascot peptide Mass Fingerprinting search (Matric Science Inc. Boston MA). Search parameters can be seen in *Figure 8*. Peak lists were entered into the query entry location (*Figure 8*) and searched for matching proteins within the contaminant database. All proteins found were downloaded as a .CSV file and added to the list of proteins found for the respective samples. All unmatched peptides were then searched once again for further contaminants in the cRAP database curated by the Global Proteome Machine using UNIPROT protein sequences. Once again, all proteins found were placed into the same .CSV file and unmatched peptides were searched for a third time in the NCBI database. All peak lists were analyzed in this manner and compiled into three separate trials for all six samples.

2.10.2 Filtering of protein and peptide results from Mascot

Using Microsoft Excel, Repeated proteins were removed for each sample using the "IF(COUNTIF)" function (*Figure 9*). Once repeated peptides were removed from each sample, all samples were compiled together, and ascension numbers were compared to determine proteins unique to PHFK and PHFK-HA samples. This was once again done

using the "IF(COUNTIF)" function. All proteins unique to PHFK in each sample proceeded to undergo several comparisons. First Individual trials for late passages of AAE6 and EPE6 were compared for unique and common proteins using the "IF(COUNTIF)" function. Unique proteins for all trials for each comparison (i.e. unique proteins from trial 1,2 and 3, for late AAE6 and late EPE6 comparison) were compiled and used for analysis in several databases.

2.10.3 Biological process of unique proteins in Panther DB

The Protein Analysis Through Evolutionary Relationships (PANTHER) is a database containing evolution of gene function and pathway relationships (Mi et al. 2013). Ascension numbers of identified proteins post filtering were entered into the manual search section separated by a space and analyzed for *Homo sapiens* gene functional classifications. This search was conducted for all three replicates of: AAE6 Late, EPE6 Late during MALDI MS/MS analysis and the first replicate of LC MS/MS identified proteins (*Section 2.9.6*).

2.10.4 Reactome Pathway Analysis

Reactome is a public database that enables researchers to cellular processes at the molecular level for over 50 % of human protein-coding genes. Unlike PANTHER (*Section 2.5.3*), which analyzes evolutionary relationships for gene functions and pathway relationships, Reactome generates systematically described molecular pathways allowing for a deeper understanding of mechanisms behind signal transductions, metabolism, DNA replication and more (Fabregat et al. 2018). By importing the ascension numbers of discovered proteins for all three replicates of: AAE6 Late and EPE6 Late, pathway analysis could be done. The results were viewed independently for each sample type and compared

for similarities. Results were compiled in a PDF file consisting of detailed pathway analysis and likelihood of any false positives. Detected pathways were exported as a .CSV file and filtered using 'IF(COUNTIF)' function described in *Section 2.10.2*.

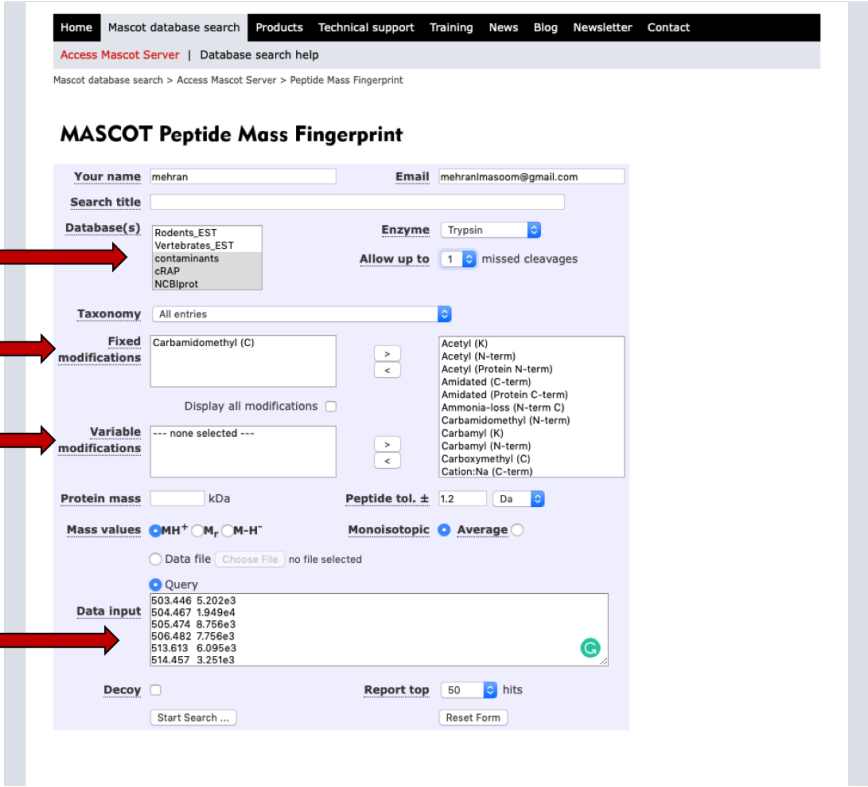

Figure 8 – Screenshot of the Mascot Peptide Mass Fingerprint input screen. The four arrows from top to bottom depict the following: Database selection(s), Fixed modifications, Variable modifications, and manual data input.

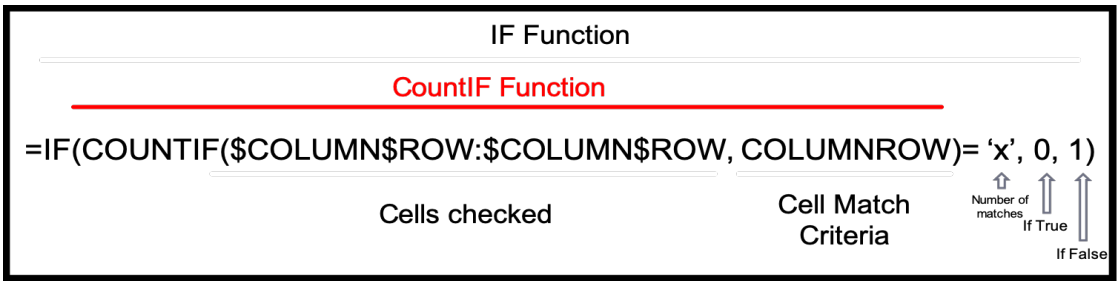

Figure 9 – Equation used to filter repeated proteins from unique proteins. All values for "COLUMN", "ROW", "COLUMNROW" could be modified to fit the user's needs. For example, if column 'A' from rows 1 to 1000 were to be checked to see if any cells contained the same value as cell "G3" the cells checked would be written as "A1: A1000, G3". If this value is equal to the number specified by 'x' the number '0' will be the output and if the value is not equal to 'x' the number '1' will be the output.

2.11 Protein Filtering for LC MS/MS

As HMSPF used Proteome Discoverer to identify proteins prior to delivery of the data, no identification of MS/MS spectra data was required. Proteins from each sample needed to be sorted and filtered for contaminants. The filtering of LC MS/MS data was done by first removing all proteins in AAE6 and EPE6 samples containing the names: ribosomal, ribosomal protein, or *CON* (n=50). The Contaminant Repository for Affinity Purification MS Data (CRAPome, Mellacheruvu D, et al. 2013) allowed for the removal of commonly found proteins in LC MS/MS experiments. Commonly identified proteins returned a score of at least 200 out of 411 experiments resulting in removal from further analysis (n=44). Finally, any remaining proteins had their respective UniProt Identifiers imported into Reactome, and removed any proteins involved with Metabolism of RNA from analysis (n=89). With all remaining proteins unique from common contaminants and ribosomal proteins, they were filtered in one of two ways: *Peptide Method* and *Protein-Pathway Method* (Summarized in *Figure 10*).

2.11.1 Protein Filtering "*Peptide Method*"

As proteins present within any PHFK sample most likely is a contaminant, they needed to be removed (n=586). Now that the remaining proteins were also unique to all negative control samples potential candidates needed to be selected. The criteria for shortlisting a protein by the *Peptide Method* was the following:

1. Any EPE6- or AAE6-targeted protein must have a sum of 2 or more peptides in either trial independent of treatment.

OR

2. Any EPE6- or AAE6-targeted protein must have at least one peptide in both trials independent of treatment.

Using this "*peptide method*" we identified 19 proteins as AAE6 binders and 19 proteins as EPE6 binders of which 13 proteins are found in both AAE6 and EPE6 (*Table S2*). *Figure 10* shows a flow chart depicting how proteins were filtered using the "*peptide method*".

2.11.2 Protein Filtering "*Protein-Pathway Method*"

An alternative and more inclusive filtering process was conducted due to a loss of known E6 interacting proteins using the stringent *Peptide Method*. This method reduces any reliance on the number of peptides present in a given sample but emphasizes the protein differences between two groups: AA and PHFK-HA as well as EP and PHFK-HA. First proteins from AA, EP and PHFK samples were filtered similarly to *Section 2.11.1* All AAE6 targeted proteins that appeared in PHFK-HA samples were removed. The same was done for EPE6 targeted proteins and PHFK-HA samples and. The final filtering step in the protein-pathway method was to import once again all remaining AA, EP and PHFK-HA proteins into Reactome individually (i.e. all AA proteins imported, then all EP proteins

then all PHFK-HA proteins). All proteins in AA or EP samples involved in a similar pathway to that of PHFK-HA proteins were removed from further analysis.

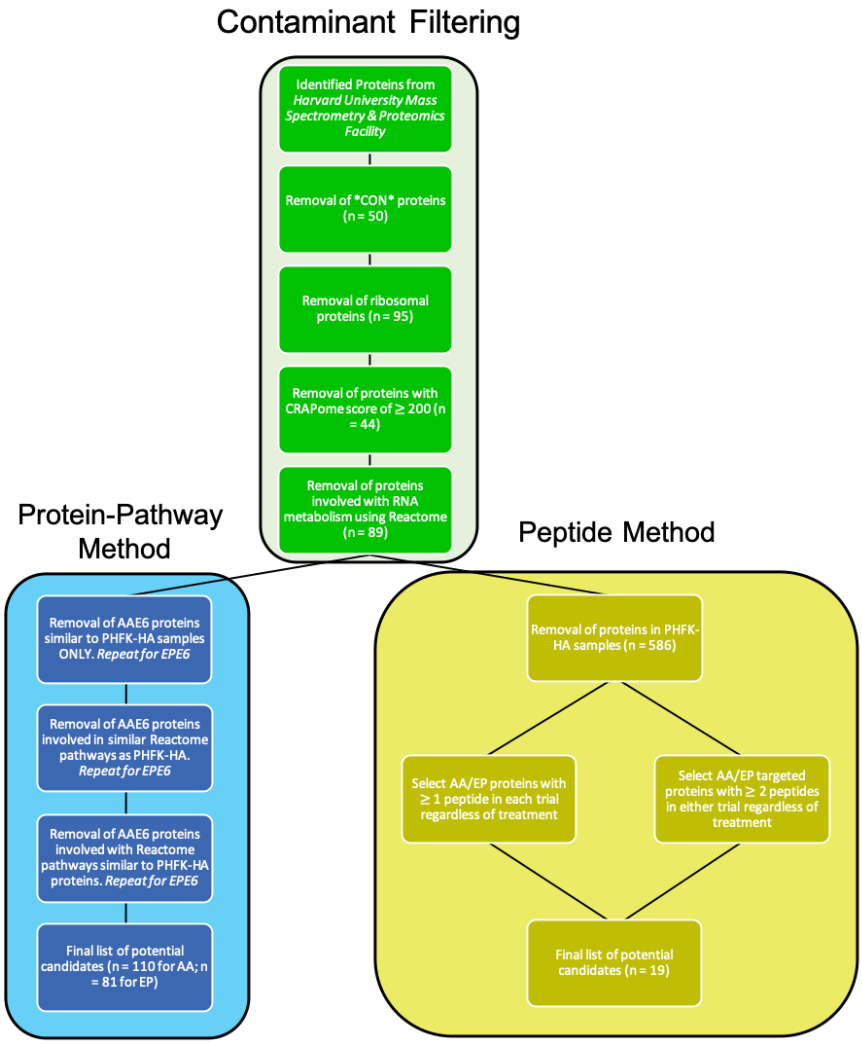

Figure 10 – Summary of both *Peptide* (yellow) and *Protein-Pathway* (blue) filtering methods for LC MS/MS Trials. Steps in green are part of both methods. PHFK=Primary Human Foreskin Keratinocytes; *CON*=technical contaminants (e.g. Trypsin).

2.11.3 Bioinformatic analysis of filtered proteins

Proteins filtered by the *Peptide method* were of greatest interest and remained the focus of bioinformatic analyses. All analyses began through investigation of each protein's involvement in various Reactome pathways. Proteins were individually imported into the

Reactome database using their respective UniProt identifiers. By importing each protein separately into Reactome, we simplified the output result of which pathway they were involved in. Once all proteins were analyzed through Reactome and their involvement in biological pathways were recorded, a short list of potential intriguing proteins was developed. For a protein to be included within the short-list, they needed to be involved in pathways involving hallmarks of cancer such as: angiogenesis, inflammation, DNA repair/destabilization, energy dysregulation; HPV; cervical cancer; or relationships with previously published literature within our lab group. For proteins filtered by the *Protein-Pathway Method*, the list of proteins identified for AAE6 and EPE6 were run through Reactome. The level and significance of enrichment in each pathway was identified and this information was used to compare the effects of AAE6 and EPE6.

Next, for *Peptide Method* selected proteins, they were imported into The BioGrid (Oughtred et al. 2019) using their UniProt Identifier. One of the benefits to BioGrid is that it lists all published interactions with any particular protein. The list of unique interactors for each protein were searched for their presence in the final *Peptide* Method table. Each unique interacting protein present within our datasets were noted and investigated for their respective function and potential for the E6 interacting protein's possible effect within the cell.

3 Results

The results are presented in three parts. The first section discusses the approaches conducted at the Health Sciences North Research Institute (HSNRI) in Sudbury Ontario with the generous assistance and hospitality of Dr. Hoyun Lee and Dr. James Knockleby but which had to be abandoned as discussed below. The second section discusses the conditions for each optimization done to the Co-IP protocol. All of the tests done for the second part of the thesis were conducted at the Thunder Bay Regional Health Research Institute (TBRHRI). The final part of the thesis results discusses the proteins identified by LC MS/MS through Harvard University's Center for Mass Spectrometry Proteomics, Cambridge Massachusetts, United States.

3.1 MALDI MS/MS at HSNRI

3.1.1 2-D Electrophoresis

HPV16 E6 interacts with a variety of cellular host proteins that vary in size. Some interacting proteins have similar masses. An example of two proteins are hDLG and E6AP. Both of these proteins have a mass around 100 KDa (hDLG and E6AP have masses of 100.4 KDa and 100.68 KDa respectively). Proteins with such similar masses may be difficult to differentiate on an SDS PAGE gel through conventional 1-D methods. Therefore, an attempt was made to separate pulled down proteins as much as possible through 2-D electrophoresis. Sample proteins were separated first by pI and then by their molecular weight. The added separation by pI should prove beneficial for proteins with similar masses. Due to the potential of low amounts of protein pulled down from the co-IP, silver staining was used. Unfortunately, no bands formed indicating that 2-D

electrophoresis was not an efficient method to increase protein separation and overall detectability for downstream applications.

3.1.2 Presence of Protein of Interest

To prevent unnecessary use of the mass spectrometer, western blotting was done to confirm the presence of the target protein E6. Samples were incubated as described (*Section 2.2.5 and 2.2.6*) and eluted with 20 µL 1 X SDS reducing buffer. From the total eluted sample, 4 µL was set aside for transfer to a PVDF membrane. Post transfer, samples were incubated overnight in anti-HA mouse mAb and a subsequent 1-hour incubation with goat anti-mouse Ab conjugated with HRP. As expected, samples transduced with any variant of HPV16 E6 showed an additional band at 18 KDa (*Figure 11*). This additional band was absent in all PHFK and PHFK-HA samples. Therefore, E6 positive samples analyzed for MALDI MS/MS contained E6 but our negative controls did not.

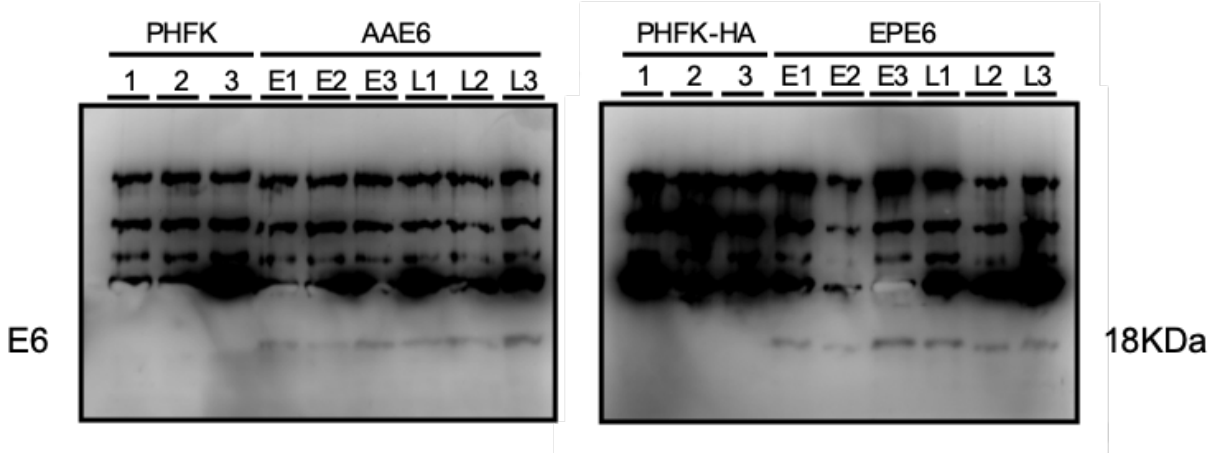

Figure 11 – Western blot of PHFK, PHFK HA, AA and EP for three biological replicates (E denotes early passage samples whereas L denotes late passage samples). Each sample was incubated overnight with 1°Ab and for 1 hour with 2°Ab. The lowest band (18 KDa) present in all AAE6 and EPE6 samples is E6. As expected, this band was absent for all PHFK and PHFK-HA samples. Images were taken with 2-minute exposure, automatic binning adjustment and the door closed.

3.1.3 Identified Proteins

Three biological replicates for each of AAE6, EPE6, PHFK and PHFK-HA were analyzed after in gel tryptic digestion with a MALDI MS/MS. Spectra from each sample were uploaded for Mass peptide fingerprinting analysis and compiled into an Excel spreadsheet. Samples were first organized based on sample type (i.e. AAE6 or EPE6), then filtered for the following: repeating peptides (within sample type) and repeating peptides and accession numbers (between sample type and negative controls (PHFK and PHFK-HA)). Once samples were filtered and sorted, AAE6 and EPE6 samples were compared for unique accession numbers, and unique peptides. The criteria to determine whether an identified protein was significant, a threshold of three or more peptides was chosen. Interestingly, there was no significant protein that appeared in all three biological replicates for a given sample (i.e. AAE6, EPE6 or PHFK-HA). When looking at the proteins present in both AAE6 and EPE6 samples, several proteins were involved in the cellular immune system (*Table 4*). Partial MHC class II antigens as well as interleukin 11 and 17 were identified as unique to AAE6. Therefore, there is the possibility that AAE6 can alter immune functions within the cell. Two ubiquitin protein ligases NEDD4-like E3 ubiquitin protein ligase WWP2 (WWP2) and E3 ubiquitin protein ligase NEURL3 were identified in only the AAE6 cells. The presence of WWP2 was of interest as there has not been data published of its interaction with E6. This protein not only was identified in AAE6 but was absent in all samples of EPE6 and the negative controls PHFK, PHFK-HA. Therefore, there is the possibility that E6 interacts with another E3 ligase in vitro.

Proteins unique to EPE6 samples obtained from the MS data generated other interesting findings (*Table 5*). A known HPV16 E6 interacting protein caspase 8 (CASP8,

Table 5) was identified in only EPE6. There was also a CASP8 associated protein indicating there may be more to the E6-CASP8 interaction than previously thought and will require more investigation. Another known interacting protein to E6 was c-MYC (*Table 1*). Unique to EPE6 samples was a c-MYC binding protein present in only one replicate. Interestingly, the tumour suppressor protein breast cancer type 2 susceptibility protein was identified as unique to only EPE6-transduced PHFKs. Unfortunately, the processed MS data failed to identify several key proteins known to bind to E6 such as E6AP, P53, and most importantly failed to identify E6 itself for both AAE6 and EPE6 samples. Therefore, another approach was required to allow for straight-forward bioinformatics analysis. In order to create such an approach, laboratory techniques were revisited and optimized such that the total quantity of protein extracted from each sample was maximized; minimize antibody leeching during Co-IP elutions; and optimize retention of E6 variants and their interacting proteins. The first technique revisited was lysing variant E6 transduced mammalian PHFK cells.

3.2 Optimization of Co-IP for LC MS/MS trials

3.2.1 Lysis of Mammalian PHFK

PHFK's transduced with HPV16 L83VE6 were lysed using several buffers. Buffers were cooled to 4 °C prior to lysis of cells. After incubation of cells in lysis buffer, protein concentration was determined for total protein amount obtained and cells were air dried onto a clean glass microscope slide and imaged (*Figure 12*). To have a baseline for lysis efficiency, a control lysis was conducted using 1 X PBS. The use of MPER has no effect on nuclear membrane lysis as the border of each nuclei visually appears intact. Even when

increasing the volume of lysis buffer from 10 : 1 (10 µL buffer to 1 µg of cells) to 20 : 1 there was no visual difference in lysis efficiency. The nuclei lysed with MPER have sharp edges and little nuclear debris throughout the sample. Interestingly, the addition of 150 mM NaCl to MPER resulted in increased nuclear disruption indicating improved lysis efficiency. Similarly, to the lysis of MPER with 150 mM NaCl, nuclear disruption efficiency increased when lysing with a Tris-HCl lysis buffer developed by Elizabeth White et al. 2012. Once again there is visual disruption of the nuclear membrane. However, some cells lysed with this buffer demonstrated improved lysis efficiency as nuclear debris is present in regions distant from lysed nuclei. Overall, we determined that the most efficient lysis method was using the Tris-HCl buffer. As, there were several steps to completing a Co-IP, I continued to test the incubation and washing efficiency of all three buffers.

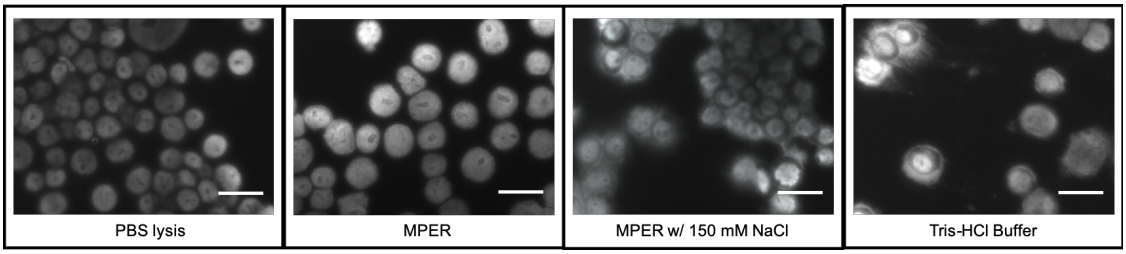

Figure 12 – DAPI stained nuclei of PHFK cells transduced with the L83V HPV16 E6 oncogene. Image was taken in greyscale. All lysis buffers show varying disruption of the nuclear membrane. Cells lysed with MPER alone experienced poor nuclear disruption and visually did not vary from the control. The addition of 150 mM NaCl to MPER resulted in a clear increase in nuclear membrane disruption. This can be identified by increased nuclear debris surrounding each nucleus. Finally, cells lysed with Tris-HCl/nonidet P-40 buffer seemed to have the greatest levels of nuclear membrane disruption.

Lysis efficiency was quantitatively and qualitatively compared via western blot using actin as a normalizer for densitometry (*Figure 13*). Since the original method used MPER to lyse PHFK cells for MALDI MS, this was considered our benchmark. Lysis

efficiency increased the most (28.0 %, n=1) when Tris-HCl was used as a buffer. When adding 150 mM of NaCl to MPER, the efficiency of lysing PHFK's increased by 2.82 % (n=1) compared to the benchmark. Overall, the buffer with the greatest lysis efficiency was Tris-HCl. Surprisingly, when a complete Co-IP using the same number and duration of washes with different lysis buffers, there visually appeared to be more E6 eluted using the MPER + NaCl lysis buffer (*Figure 13 C*). Therefore, going forward, the lysis buffer used was MPER with 150 mM NaCl.

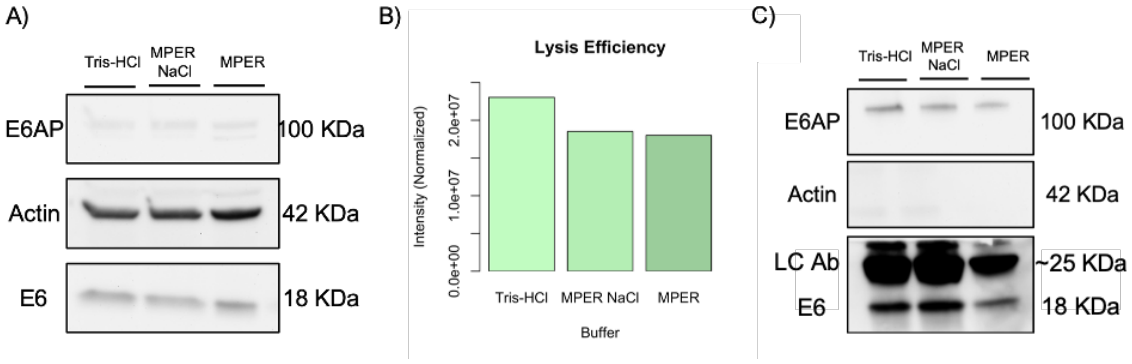

Figure 13 – A) Western blot of L83V transduced PHFK's lysed with Tris-HCl lysis buffer [pH 7.5] (Column 1), MPER with 150 mM NaCl (Column 2), and MPER (Column 3). In this figure, actin was used as a normalizer and the blotting of E6AP was done to confirm the ability of the antibody to bind to E6AP effectively. B) Densitometry of E6 post lysis using Tris-HCl [pH 7.5] (Column 1), MPER with 150 mM NaCl (Column 2), and MPER (Column 3). Data was plotted using R Studio Version 1.1.383 and densitometry completed using Image Lab Version 6.0.1. C) Comparison between total E6 and E6AP eluted from L83V cells lysed with either Tris-HCl (Column 1), MPER with 150 mM NaCl (Column 2) or MPER (Column 3). Western Blot images were taken at 20-minute exposure and 1 X 1 binning with gamma adjusted to 0.4 using Photoshop™.

3.2.2 Antibody Selection for Western Blotting/Co-IP

Proper choice of antibodies used for western blotting is important when trying to optimize the Co-IP protocol. When attempting to detect our E6 variants in transduced

PHFK's, two anti-HA antibodies were used. The first was a monoclonal antibody that requires the use of a secondary antibody such as anti-mouse conjugated to an HRP tag. The other option is to use a primary anti-HA antibody with a conjugated HRP tag already on the antibody. *Figures 14 – 19* show differences between the results of the western blots. When using the monoclonal anti-HA antibody with no HRP conjugated tag (*Figures 14-17*), a secondary antibody that interacts with IgG proteins that may leech from the beads can allow for dual detection of IgG and E6. Pierce anti-HA magnetic beads use mouse IgG and therefore it is possible to use the combination of mouse monoclonal anti-HA primary with anti-mouse conjugated to an HRP tag so E6 can be visualized at the same time as any leeched antibodies from beads. If E6 is solely to be detected, it is possible to use anti-HA antibody conjugated to HRP as shown in *Figures 18 and 19*.

3.2.3 Incubation and Wash of Input Protein

To determine the optimal time required to sufficiently bind E6 to our anti-HA magnetic beads, equal amounts of lysed proteins recovered post lysis of MPER were used. Most literature stated that they used an overnight incubation (4-16 h) however, due to the instability of E6 post-lysis, there was reason to believe a shorter incubation time would be beneficial. For each sample we used 600 µg of input protein with 25 µL of magnetic beads and incubated at 4 °C for 30 minutes and overnight (*Figure 14 A and B respectively*). Washes were concentrated using a 3K concentrator (Fisher Sci., Cat# 88512) until the volume was low enough for the entire sample to be loaded into each well. Western blotting was completed and imaged to identify how much E6 was lost while incubating and washing (*Figure 14*). Samples that were incubated for 30 minutes could not sufficiently bind all incubated E6 onto the beads and each wash caused further loss of E6. Interestingly when

incubating our input protein overnight, there appeared to be no E6 within the flow-through and no E6 in any washes. Therefore, incubating input protein overnight resulted in increased binding of E6 to the magnetic beads and overall reduced loss of E6 during washes. Most likely an increased incubation time will result in increased retention of binding partners to E6.

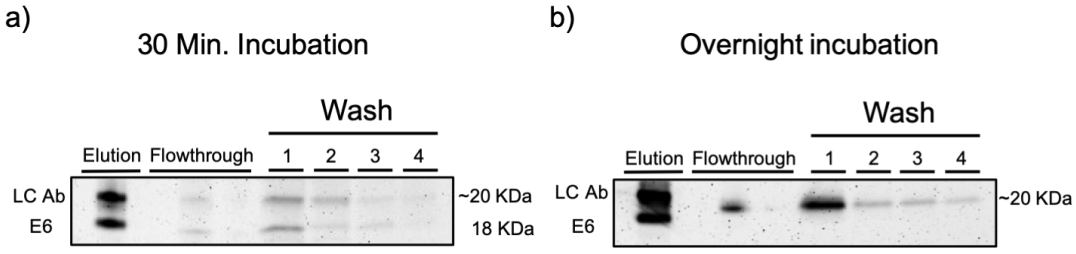

Figure 14 – Comparison between incubating 600 µg of lysed PHFK's transduced with L83V input protein on 25 µg of anti-HA magnetic beads for 30 minutes or overnight. A) Shows the effect washing samples with TBS containing 0.05 % tween 20 after 30 minutes of incubation. B) Shows the effect washing samples with TBS containing 0.05 % tween 20 after incubating samples on beads overnight. Images were taken with the same CCD camera with 4 X 4 binning, 10-minute exposure and gamma adjusted to 0.5 using Photoshop™; LC=light chain

As shown in *Supplemental Table 1*, a comprehensive literature search was done for all known interactome studies done for HPV16 E6. The results from this search provided a standard for how previous studies washed their samples to remove nonspecific binders and potential contaminants. White et al. 2012 demonstrated an efficient method to co-immunoprecipitate E6 along with known binders that consisted of three 5-minute washes in cold lysis buffer, and three 5-minute washes in cold 1 X PBS. Their positive results, other papers and the manufacturer's suggestion to use three washes provided enough rationale to wash our samples with three 5-minute washes with cold lysis buffer and three 5-minute washes with cold 1 X PBS.

3.2.4 Elution of Protein

To determine the efficiency of elution buffers on our samples of interest, several buffers were used and compared to our original method using SDS loading dye. As a baseline, equal volumes of SDS elution's were done on each sample post alternative elution buffer. A method to determine elution efficiency is either by decreased presence of E6 within the SDS elution buffer or increased presence of E6 in the alternative elution buffer. Initial elution methods used SDS reducing buffer as this would strip all proteins from the beads. However as seen in *Figure 15.* the elution is too harsh causing large quantities of antibody to leech from the beads. This left only two buffers available for elution. The first was HEPES Buffer (25 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) [pH 8.0], 8 M urea, 0.02 % (v/v) Triton X-100 and 5 % (v/v) glycerol, and the second was RIPA buffer (25 mM Tris-HCl [pH 7.6], 150 mM NaCl, 1 % sodium deoxycholate, 0.1% SDS and 8 M urea). When eluting protein using HEPES buffer (*Figure 15*), E6 failed to be detected on the western blot at any exposure time. It is worth noting that increasing the temperature of incubation to 37 °C or time of incubation did not result in any noticeable improvement of elution efficiency. When using RIPA buffer, elution efficiency improved compared to HEPES buffer. E6 was detected at a moderate (5-minute) exposure time with 4 X 4 binning. Unfortunately, the majority of E6 present within the sample remained on the beads as can be seen in the SDS lanes for *Figure 15 B*. A trial run of variant samples (AA and EP) using RIPA buffer to elute the protein was done (*Figure 15 A and B*) and sent for MS identification. Unfortunately, with poor results, once again troubleshooting was done using L83V cell lines and several other buffers were tested.

As a result of the literature search, the most popular elution method was using synthetic HA peptide to out compete the interaction between E6 with HA tag, associated binding partners and magnetic bead antibodies. As shown in *Figure 15 C* eluting protein with HA peptide failed to successfully elute E6 and associated binding partners. Changing the concentration of HA peptide (used 125 µg/mL to 1 mg/mL) as well as altering incubating temperature and duration failed to improve elution efficiency. Finally, two acidic elution buffers were used: Glycine-HCl [pH 2.5] and a commercial elution buffer from ThermoFisher Scientific (Acid EB, Cat# 1858606). Upon first glance there does not appear to be a significant difference between the two buffers ability to elute proteins. However, when analyzing the SDS elutions for both acidic buffers, it appears the Glycine buffer had removed all E6 from the beads while keeping the majority of antibody on the solid substrate. The commercial elution buffer did remove a substantial amount of E6 from the beads but there remained a detectable amount of target protein on the magnetic beads. Therefore, we continued to optimize elution buffers using Glycine-HCl. To determine the optimal pH of Glycine-HCl to elute our protein without leeching antibody off of the beads. Three pH levels were tested: 2.2, 2.5 and 2.8 (*Figure 15 D*). There did not appear to be a noticeable difference between the three samples for the amount of E6 eluted. Visually there appears to be more antibody leeched from the beads when comparing the pH 2.2 versus the 2.5 elution. As a result, further troubleshooting continued with 0.2 M Glycine-HCl pH 2.5. While determining which buffer would be best suited for elution of target proteins and MS analysis, an attempt was made to scale up the experiment. This was done by performing a Co-IP using twice the volume of magnetic beads (50 µL) and 1.2 mg of input protein.

(*Figure 15 D*). Elution conditions were the same as other acidic elution's and resulted in increased target protein elution with a moderate increase in antibody leeching.

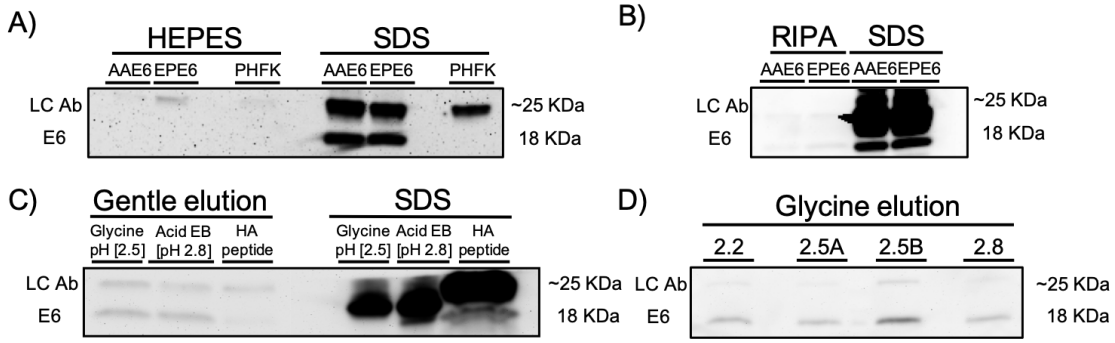

Figure 15 – Elution comparison between HEPES buffer and SDS elution A). Image was taken with 5-minute exposure, 4 X 4 binning, and gamma adjusted to 0.2. Elution of proteins using RIPA buffer was done and because there was improved elution, a shorter exposure time was used (1-minute) with 4 X 4 binning B). Gentle elution attempts were done on L83V cell lines as troubleshooting needed to be conducted once again C). Images were obtained using a 10-minute exposure due to decreased binning (1 X 1) and 0.2 gamma. Decreased binning was used to make it easier to identify E6 in the presence of substantial antibody when eluting with SDS loading buffer. Elution of samples using 0.2 M Glycine-HCl with varying pH. Images were taken with 20-minute exposure and 1 X 1 binning with 0.5 gamma correction.

3.2.5 Stability of Protein Within Elution Buffer

Once determining the most efficient elution buffer compatible with MS applications, another challenge was presented. As described in *section 1.3.0*, E6 has a short half-life (less than30 minutes) and the samples must be stable for at least a week to ensure accurate identification of proteins within the eluted sample. As such, the stability of E6 was tested with protein eluted by 0.2 M glycine [pH 2.5] and immediately neutralized. This was done by performing two identical pull-downs with equal amounts of protein (1.2 mg). To simulate immediate denaturation of protein, one sample was incubated in 6 X SDS loading dye with DTT immediately after elution. Once the sample was denatured, it was

stored at -80 °C for one week while the other sample was immediately stored at -80 °C without denaturing. Storage at -80 °C simulated shipping conditions on dry ice and handling conditions both in our lab and at the proteomic facility. Once one week had passed, both samples were thawed, the nondenatured sample was reduced with 6 X SDS loading dye containing DTT and 30 µL of both samples were run on a 15 % SDS PAGE gel. Subsequent western blotting was conducted as shown below (*Figure 16 A*). Samples were analyzed in triplicate (n=3) by comparing the ratio of E6 to the light chain antibody. *Figure 16 B* shows that there was no significant difference between the ratio of E6 to light chain antibody. Therefore, storage of Co-IP'd samples in neutralized 0.2 M Glycine [pH 2.5] provides a stable environment for E6.

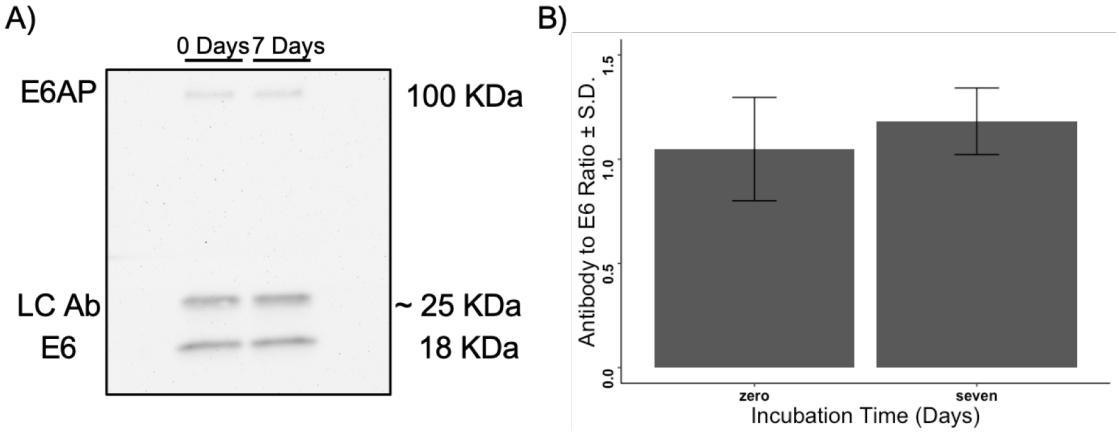

Figure 16 – Stability of protein eluted using 0.2 M glycine [pH 2.5] from Co-IP'd L83V cellular proteins. a) Western Blot of E6 and E6AP for samples with varying time between elution and denaturing using DTT. Denaturing samples immediately compared to 7 days post elution resulted in no significant difference in the total amount of E6 and E6AP pulled-down. b) Ratio of E6 to light chain antibody (n=3) also shows little difference between denaturing samples immediately or 7 days post elution (standard deviation (SD) bars overlap). Western Blotting image was taken with 1 X 1 binning, 20-minute exposure and gamma adjusted to 0.4 using Photoshop™.

3.2.6 Quality Control of IP Methodology

To ensure the IP methodology worked, a comparison was done between a transduced cell line containing E6 (L83V) and non-transduced control (PHFK) cells. Western blotting on eluted IP samples for both PHFK, and L83V samples was conducted to identify specificity of interactions with HA antibodies. A positive result would show differences between pulled down proteins for each sample (no protein should be pulled down for PHFK). A known interactor E6AP was used to determine if our IP could selectively pull-down proteins interacting with E6 (*Figure 17*). The figure shows a clear band representing E6 and E6AP in the lane containing L83V eluted proteins while there is an absence of E6 and E6AP in the control lane containing PHFK input protein. This indicates the Co-IP is able to selectively pull-down proteins bound to E6. Another example can be found in *Figure 18*.

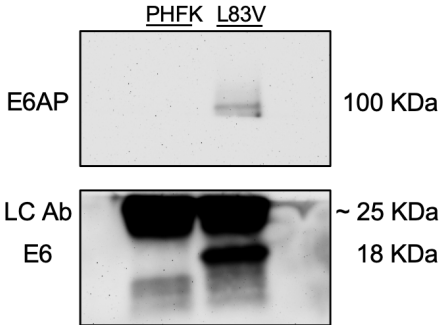

Figure 17 – Confirming the selective pull-down ability of the Co-IP. A positive cell line (L83V) and negative control (PHFK) were used. L83V cellular proteins were capable of pulling-down both E6 and E6AP while PHFK was unable to pull-down any known E6 binders as expected. This confirmed the method is able to selectively pull-down proteins that only interact with HPV16 E6. Image was taken with 1 X 1 binning, 20-minute exposure and gamma changed to 0.4 using Photoshop™.

3.2.7 Effect of Proteasome Inhibitor MG132 on HPV16 E6

To determine the effect of pulled-down proteins treated with proteasome inhibitor, AAE6 and EPE6 transduced cells were treated for 4 hours with either 30 μM MG132 in DMSO, or an equivalent volume of DMSO as a control. Samples were harvested and equal amounts of input protein were pulled down using the method discussed in *Section 2.1.2*. The ideal effect of the proteasome inhibitor would be to prevent the degradation of P53 a known target of E6. To confirm the presence of P53, a western blot using 35 μL (31.3 % of total elution volume) of eluted protein was done. For quality control the membrane was also blotted to detect E6 and E6AP in the samples. *Figure 18* clearly demonstrates successful increase in P53 for not only the input which was expected but in the elution of MG132 treated proteins. As expected, P53 was absent in the controls (DMSO treatment and PHFK). Therefore, the use of MG132 is beneficial to detect interacting proteins that may be degraded or present in low quantities.

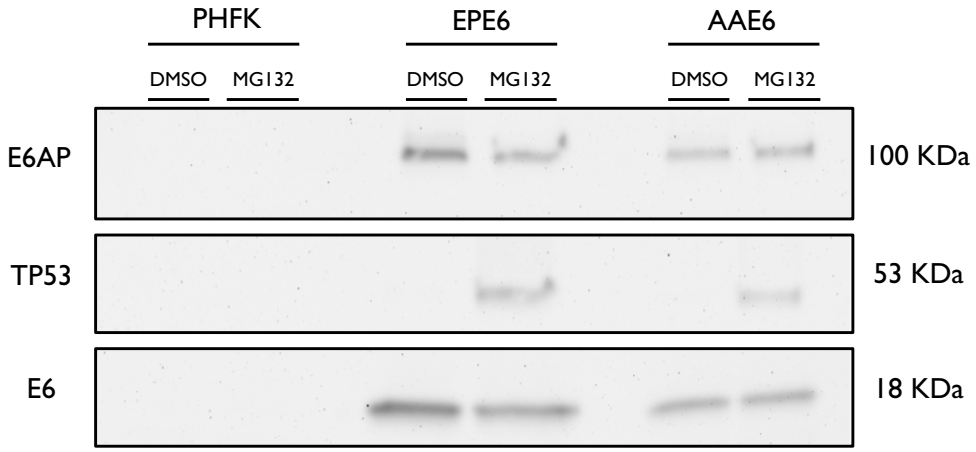

Figure 18 – Comparison of co-immunoprecipitated E6 variants and PHFK proteins treated for 4 hours with either 30 μM MG132 proteasome inhibitor or DMSO. As expected in the negative control (PHFK) there was no E6AP, P53, or E6 present within the elution. The DMSO control provided no detectable band for P53 while the MG132 treated sample contained a detectable level of P53 in E6 transduced cells. Image was taken with 1 X 1 binning, 20-minute exposure and gamma changed to 0.2 using Photoshop™.

3.2.8 Increasing Quantity of Protein Input

Although the Co-IP can selectively pull-down E6, E6AP and P53 (with MG132 treatment (*Section 3.2.7*)), the bands appear faint increasing the possibility that these proteins along with other interacting proteins will fail to be detected during MS analysis. Therefore, as briefly trialed in *Section 3.2.4* (*Figure 15 D)* pull-downs were attempted with increasing the quantity of both anti-HA magnetic beads and input protein from lysed samples. The Pierce Anti-HA magnetic beads have a capacity to bind at least 100 µg of protein and a few trials were required to identify a point in which we could saturate the binding of E6 and associated binding partners to the beads. The maximum capacity for our laboratory to grow cells for all three trials at the same time is eight T225 flasks used as described in *Section 2.1.2*. This level of cell culture allowed for 8 mg of input protein to be loaded for each sample. Since the volume required to input 8 mg of protein was greater than 2 mL, the sample was split in half and loaded with equal amounts of beads, as described in *Section 2.2.6*. We found the maximum volume of beads to sufficiently be saturated by 8 mg of input protein was 80 µL or 10 µL of beads per 1 mg of input protein. *Figure 19* demonstrates the difference in total E6 and binding partners pulled-down between 1.2 mg of input protein with 50 µL of magnetic beads and our optimized 8 mg of input protein and 80 µL of magnetic beads. Both western blots used equal volumes of eluted samples (35 µL or 31.5 % of total elution volume). In conclusion, 8.0 mg input protein and 80 µL of anti-HA magnetic beads allows for maximum pull-down of E6 and associated binding partners.

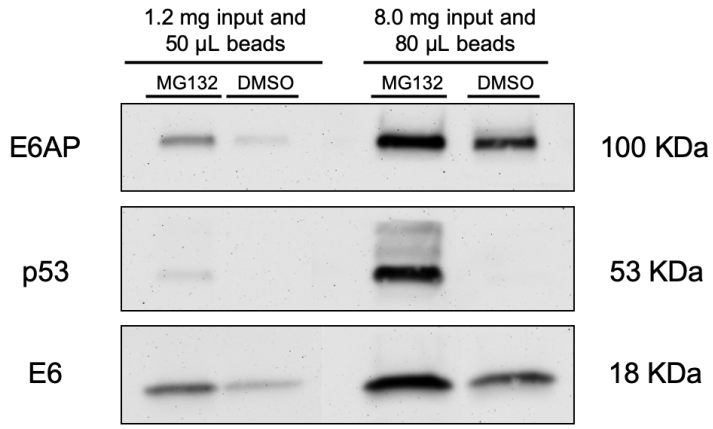

Figure 19 – Effectiveness of saturating input protein on Pierce Anti-HA Magnetic Beads. Using 1.2 mg of input protein with 50 µL of magnetic beads we were able to detect a single binding partner E6AP with preliminary LC MS/MS attempts. There are faint bands present for E6 and E6AP for samples treated with DMSO while slightly stronger bands are present for E6 and E6AP with MG132 treatment. As expected with treatment of MG132, P53 is present in low quantities for the MG132 treated sample with 1.2 mg of input protein. Increasing the quantity of input protein and beads to 8 mg and 80 µL respectively there is a noticeable increase in the amount of pulled-down proteins. All bands are present in large quantities for MG132 treated cells while there is an extremely faint band present for P53 with DMSO treated cells. Images were taken with 1 X 1 binning, 20-minute exposure and 0.2 gamma correction using Photoshop™.

3.3 LC MS/MS

3.3.1 Detection of HPV16 E6 and Known Interacting Proteins

By optimizing the Co-IP protocol and using a more sensitive mass spectrometer (LC MS/MS) successful AAE6 pull-down from AAE6 cell lines treated by both DMSO and MG132 was demonstrated. Two peptides were identified in AAE6 treated with DMSO while only one peptide was identified in AAE6 treated with MG132. Not only did the Mass spectrometer identify AAE6, it also identified five known E6 binding proteins in at least one of the following samples: AAE6 DMSO, AAE6 MG132, EPE6 DMSO, EPE6 MG132.

The most important known binder E6AP was identified in all E6 samples and absent in all PHFK-HA control samples. P53 was identified in only MG132 treated AAE6 and EPE6 samples. BCL2 was identified only in MG132 treated EPE6. Interestingly both MAGI2 and MAGI3 were detected in DMSO treated AAE6 while being absent in all EPE6 samples. It is worth noting that E6TP1 was identified in PHFK-HA MG132 treated cells and as this was a control sample the protein was excluded from subsequent analysis. Interestingly I still could not detect EPE6 in any of the treatment samples.

3.3.2 Identification of Significant Proteins

Many proteins within each MS sample were present with only identification of one or two peptides. Therefore, multiple approaches were used to look into possible unique interactions. The *Peptide Method* (*Section 2.11.1*) used a minimum identification criterion of either: three identified peptides in a sample replicate for example trial two AAE6 DMSO identified protein SDA1 homolog (SDAD1) with three peptides; one peptide in each sample replicate to be considered significant for example protein Retinitis pigmentosa 9 protein (RP9) had one peptide identified in both trial one and trial two EPE6 DMSO samples; or one peptide in two different variant treatment replicates. By using this approach, I identified 25 different proteins in AAE6 and EPE6 combined, of which 13 were common between AAE6 and EPE6 with six proteins unique to EPE6 and another six unique to AAE6 (*Table 4-5*).

Table 4 – Heat map of Peptide Method depicting potential candidates for AAE6-targeted proteins (greater than three peptides). Each column represents a single sample type and each row represents a protein (name of protein on right side). The gene name for each corresponding protein is in brackets beside the protein name. Bolded Acc# (Column 1) are common to EPE6-targeted proteins (*Table 5*). Numbers in each coloured box corresponds to the number of peptides identified.

Acc#	Description	CRAPome	AAE6 DMSO T1	AAE6 MG132 T1	AAE6 MG132 T2	AAE6 DMSO T2	Sum
Q8TDD1	ATP-dependent RNA helicase DDX54 OS=Homo sapiens GN=DDX54 PE=1 SV=2	49 / 411			3		3
P04637	Cellular tumor antigen p53 OS=Homo sapiens GN=TP53 PE=1 SV=4	52 / 411		1		1	2
Q9H444	Charged multivesicular body protein 4b OS=Homo sapiens GN=CHMP4B PE=1 SV=1	28 / 411	1	1	3	2	7
Q14669	E3 ubiquitin-protein ligase TRIP12 OS=Homo sapiens GN=TRIP12 PE=1 SV=1	29 / 411			3	1	4
Q14244	Ensconsin OS=Homo sapiens GN=MAP7 PE=1 SV=1	42 / 411	2		3	3	8
Q5T3I0	G patch domain-containing protein 4 OS=Homo sapiens GN=GPATCH4 PE=1 SV=2	37 / 411			3	3	6
Q9C086	INO80 complex subunit B OS=Homo sapiens GN=INO80B PE=1 SV=2	5 / 411		2		2	4
Q7Z5P9	Mucin-19 OS=Homo sapiens GN=MUC19 PE=1 SV=3	NA	1	1		1	3
Q8WTT2	Nucleolar complex protein 3 homolog OS=Homo sapiens GN=NOC3L PE=1 SV=1	42 / 411		1		1	2
Q13823	Nucleolar GTP-binding protein 2 OS=Homo sapiens GN=GNL2 PE=1 SV=1	45 / 411			11	6	17
Q9UMY1	Nucleolar protein 7 OS=Homo sapiens GN=NOL7 PE=1 SV=2	11 / 411	1		1	1	3
Q96GQ7	Probable ATP-dependent RNA helicase DDX27 OS=Homo sapiens GN=DDX27 PE=1 SV=2	41 / 411			9	6	15
Q9HC23	Prokineticin-2 OS=Homo sapiens GN=PROK2 PE=1 SV=2	1 / 411	3				3
Q9NVU7	Protein SDA1 homolog OS=Homo sapiens GN=SDAD1 PE=1 SV=3	23 / 411	1	1	4	3	9
O75676	Ribosomal protein S6 kinase alpha-4 OS=Homo sapiens GN=RPS6KA4 PE=1 SV=1	1 / 411	1		1		2
O95478	Ribosome biogenesis protein NSA2 homolog OS=Homo sapiens GN=NSA2 PE=1 SV=1	15 / 411		1	1	2	4
Q5JTH9	RRP12-like protein OS=Homo sapiens GN=RRP12 PE=1 SV=2	61 / 411			3	1	4
Q9NUQ6	SPATS2-like protein OS=Homo sapiens GN=SPATS2L PE=1 SV=2	20 / 411			8	4	12
Q05086	Ubiquitin-protein ligase E3A OS=Homo sapiens GN=UBE3A PE=1 SV=4	1 / 411	5	4	2	4	15

Table 5 – Heat map of Peptide Method depicting potential candidates for EPE6-targeted proteins (greater than three peptides). Each column represents a single sample type and each row represents a protein (name of protein on right side). The gene name for each corresponding protein is in brackets beside the protein name. Bolded Acc# (Column 1) are common to AAE6-targeted proteins (*Table 4*). Numbers in each coloured box corresponds to the number of peptides identified.

Acc#	Description	CRAPome	EPE6 DMSO T1	EPE6 MG132 T1	EPE6 MG132 T2	EPE6 DMSO T2	Sum
Q8TDD1	ATP-dependent RNA helicase DDX54 OS=Homo sapiens GN=DDX54 PE=1 SV=2	49 / 411			5		5
Q4AC94	C2 domain-containing protein 3 OS=Homo sapiens GN=C2CD3 PE=1 SV=4	1 / 411	3				3
P04637	Cellular tumor antigen p53 OS=Homo sapiens GN=TP53 PE=1 SV=4	52 / 411		4			4
Q9H444	Charged multivesicular body protein 4b OS=Homo sapiens GN=CHMP4B PE=1 SV=1	28 / 411			5		5
Q14669	E3 ubiquitin-protein ligase TRIP12 OS=Homo sapiens GN=TRIP12 PE=1 SV=1	29 / 411			3		3
Q14244	Ensconsin OS=Homo sapiens GN=MAP7 PE=1 SV=1	42 / 411		2	3	2	7
Q5T3I0	G patch domain-containing protein 4 OS=Homo sapiens GN=GPATCH4 PE=1 SV=2	37 / 411			2	3	5
P20592	Interferon-induced GTP-binding protein Mx2 OS=Homo sapiens GN=MX2 PE=1 SV=1	1 / 411	4	1			5
Q13823	Nucleolar GTP-binding protein 2 OS=Homo sapiens GN=GNL2 PE=1 SV=1	45 / 411		1	2	1	4
Q9UMY1	Nucleolar protein 7 OS=Homo sapiens GN=NOL7 PE=1 SV=2	11 / 411	1	1	1	1	4
Q96GQ7	Probable ATP-dependent RNA helicase DDX27 OS=Homo sapiens GN=DDX27 PE=1 SV=2	41 / 411			7	1	8
Q9NVU7	Protein SDA1 homolog OS=Homo sapiens GN=SDAD1 PE=1 SV=3	23 / 411	1		5	2	8
Q8TA86	Retinitis pigmentosa 9 protein OS=Homo sapiens GN=RP9 PE=1 SV=2	52 / 411	1			1	2
O95478	Ribosome biogenesis protein NSA2 homolog OS=Homo sapiens GN=NSA2 PE=1 SV=1	15 / 411			3	1	4
P84101	Small EDRK-rich factor 2 OS=Homo sapiens GN=SERF2 PE=1 SV=1	33 / 411			2	3	5
Q9NUQ6	SPATS2-like protein OS=Homo sapiens GN=SPATS2L PE=1 SV=2	20 / 411		2	7	1	10
O75683	Surfeit locus protein 6 OS=Homo sapiens GN=SURF6 PE=1 SV=3	38 / 411			3		3
P11441	Ubiquitin-like protein 4A OS=Homo sapiens GN=UBL4A PE=1 SV=1	19 / 411	1		2		3
Q05086	Ubiquitin-protein ligase E3A OS=Homo sapiens GN=UBE3A PE=1 SV=4	1 / 411	5	4	2	2	13

We then screened for functions related to HPV-related tumourigenesis and immune suppression, obtaining a short-list of 7 proteins. Screening for HPV related functions involved identifying peer-reviewed literature for each variant interacting protein using Google Scholar, and PubMed databases. Each literature search consisted of using the protein name followed by each of the following: HPV, HPV16, HPV16 E6, cancer, or cervical cancer (e.g. MX2 HPV16 E6). Among the short-list of proteins, three were found

exclusively in AAE6: the INO80 complex subunit B (INO80B), the Prokineticin-2 (PROK2) and Ribosomal protein S6 kinase alpha-4 (RPS6KA4), three were found in both AAE6 and EPE6: the E3 ubiquitin-protein ligase (TRIP12), the Nucleolar GTP-binding protein 2 (GNL2) and the charged multivesicular body protein 4b (CHMP4B), and interferon-induced GTP-binding protein (Mx2) was found only in EPE6 samples. By using several databases including Reactome (Fabregat et al. 2018), and, BioGrid (Oughtred et al. 2019) we were able to identify potential effects that E6 variants have on various host cellular proteins.

The nuclear INO80 complex subunit B (INO80B) is part of the ATP-dependent INO80 remodeling complex consisting of 12 proteins. It has key functions in transcription regulation, DNA replication and repair, telomere maintenance and chromosome segregation (Min et al. 2013; Seeber et al. 2013). Researchers observed higher than normal expression of INO80 in cervical cancer epithelial cells (Hu et al. 2016). By increasing INO80 expression, the homeobox protein Nanog became overexpressed, resulting in tumourigenesis promotion (Hu et al., 2016). In a previous study (Lee et al. 2014), researchers uncovered the necessity of INO80 in DNA replication fork progression. DNA replication progression by INO80 was accomplished through interaction with tumor suppressor BRCA1-associated protein-1 (BAP1) (Lee et al. 2014). Interestingly, the absence of BAP1 in a variety of cancer cell types resulted in destabilization and downregulation of INO80. (Lee et al. 2014; Hu et al. 2016). HPV16 E6 is known to interact (*Table 1*) with another protein involved in chromatin remodeling, Tip60, a component of the TRRAP/Tip 60 complex (Jha et al. 2010). In addition, E6 interacts with Myc which can recruit the Tip60 complex leading to histone acetylation (R Frank et al. 2003). Previous

reports indicate a link between Myc and INO80B. Indeed, among other proteins identified by the peptide method (Table 4-5), INO80B interacts with GNL2, RRP12 and RP9 based on Co-IP data (Cloutier et al. 2017; Kuroda et al. 2004), while proximity label-MS showed that Myc interacts with GNL2 and RRP12 (Kalkat et al. 2018). INO80B and RRP12 were present only in the AAE6 sample, GNL2 was found in both AAE6 and EPE6, whereas RP9 was unique to EPE6. Hence, even if EPE6 does not immunoprecipitate INO80B, EPE6 may alter INO80B functions indirectly.

Reactome analysis of INO80B yielded 8 pathways and for each of them are mentioned the protein present in the peptides method table that are involved in the same pathways: DNA Damage Recognition in GG-NER, Global Genome Nucleotide Excision Repair (GG-NER), UCH proteinases, Nucleotide Excision Repair, Deubiquitination (P53), DNA Repair (P53), Post-translational protein modification (P53 and Muc19), Metabolism of proteins (P53 and Muc19). Since Reactome analysis of GNL2, RRP12, and RP9 did not return any pathways, the roles of their interaction with INO80B and the pathways involving these proteins remain unclear. However, further literature research on nucleolar GTP-binding protein 2 (GNL2) revealed that the GNL2 protein can affect P53 levels and cell cycle regulator expression (Racevskis et al. 1996; Paridaen et al. 2011). GNL2 works similarly to nucleostemin (NS) by destabilizing P53 in zebrafish retinal cells (Paridaen et al. 2011). GNL2 is overexpressed in various cancers and promotes G_1/S phase transition. GNL2 increases cyclin-dependent kinase inhibitor expression and alters the P53/p21 pathway (Datta et al. 2015). Interactome study also indicates that in addition to INO80B, GNL2 also interacts with NSA2 (Huttlin et al. 2015) and GPATCH4 (Huttlin et al. 2017) both found in AAE6 and EPE6 within the peptide method table. Furthermore, Y2H assay

indicates that INO80B interacts with the other main HPV16 oncoprotein, E7 (Rozenblatt-Rosen et al. 2012).

In addition to INO80B, two other proteins out of the seven of interest, RPS6KA4 and PROK2, were found only in the AAE6 immunoprecipitated samples. RPS6KA4 regulates a variety of cellular functions such as: "cellular proliferation, motility, and survival" (Anjum and Blenis 2008). This serine/threonine-protein kinase is activated by $p38\alpha^{MAPK}$ and ERK1 (Pierrat et al. 1998). We previously identified increased signaling in the ERK1 pathway for AAE6 compared to EPE6, suggesting that ERK 1/2 plays a significant role in increasing H1F-1α levels in AAE6 cells. (Cuninghame et al. 2017). Consistent with this result, RPSKA4 was only detected in AAE6 samples in our current MS data. In fibroblasts, during mitogenic and stress stimuli, RPS6KA4 is activated by p38α or ERK1. In turn, RPS6KA4 phosphorylates the histone H3 leading to an increase in the promoter activity of certain cytokine and chemokines genes as well as an increase in the recruitment of NF-κB to its target promoters (Soloaga et al. 2003, Saccani et al. 2002). In addition, RPSK6A4 activates CREB causing activation of other anti-apoptotic proteins belonging to the Bcl-2 family of protein (Bfl-1/A1 and plasminogen activator inhibitor 2 (PAI-2)) (Park et al. 2005). None of the proteins in Table 5 were found in the Reactome pathway involving RPS6KA4 (Recycling pathway of L1, L1CAM interactions, Axon guidance, Developmental Biology) suggesting that these pathways are only altered by AAE6.

Prokineticin-2 (PROK2) is a chemokine-like protein usually expressed by a component of the innate immune system such as macrophages influencing host defence and angiogenesis in virus-related cancers (Lauttia et al. 2014, Kurebayashi et al. 2015).

Promotion of angiogenesis due to PROK2 was observed *in vitro* and *in vivo* (*Mus musculus*) which resulted in increased colon tumour mass (Kurebayashi et al. 2015). While PROK2 functions have not been fully characterized, Lauttia et al. 2014 observed an increased in PROK2 expression in human Merkel cell carcinomas caused by infection with Merkel cell polyomavirus. This expression was also correlated with a drastic increase of tumor infiltrating macrophages (Lauttia et al. 2014). So far little is known about PROK2's effects on angiogenesis. However, the protein is known to sequester the promoter of the HIF-1 and to alter the extracellular matrix (LeCouter et al. 2001, LeCouter et al. 2003). PEOK2 is found in 7 Reactome pathways: peptide ligand-binding receptors; G alpha (q) signalling events; Class A/1 (Rhodopsin-like receptors); GPCR ligand binding; GPCR downstream signalling; signaling by GPCR and signal transduction. Except for the latest also matching P53, no proteins in *tables 4 or 5* were associated to these pathways, suggesting that in a similar way as RPSK6K4A, binding of AAE6 to PROK2 could impact specific process not altered by EPE6.

In addition to GNL2, two proteins among the seven selected, TRIP12 and CHMP4B, are common between AA and EP. The thyroid hormone receptor interacting protein 12 (TRIP12) is an E3 ubiquitin-protein ligase that shares similarities with E6AP. The protein contains the conserved HECT domain (Homologous to E6AP Carboxy Terminus) as well as multiple LxxLL (where x denotes any amino acid) motifs that correspond to the E6 binding site on E6AP (Vande Pol and Klingelhutz 2013; Zanier et al. 2013; Larrieu et al. 2020). There are four motifs (LQALL AA position 402; LITLL AA position 485; LHFLL AA position 697; and LDQLL AA position 1862) present throughout TRIP12 that could potentially allow interaction with E6. TRIP12 triggers the ubiquitination

and degradation of several proteins including P53 activator proteins ARF (p14 in humans) or Brg-1-associated factor 57 (BAF57) (Haupt et al. 1997; Collado and Serrano 2010; Keppler and Archer 2010) potentially doubling the effect of P53 inactivation by E6 as ARF acts upstream of E6AP-mediated P53 degradation. Interestingly, TRIP12-dependant ARF degradation is inactivated upon Myc or TRADD binding to TRIP12 (Chen et al. 2010, Chio et al. 2012). This means that E6 has the potential to block TRIP12's interaction with Myc potentially allowing for degradation of ARF. In addition to Myc, TRIP12 also interacts with another known E6 binder, namely E6AP (Huttlin et al. 2015), but the consequence of this interaction is unknown. TRIP12 degradation of BAF57 could be inhibited when SMARCC1 is bound to TRIP12. BAF57 is a canonical component alongside BAF53 of the SWI/SNF chromatin remodeling complex (Martens and Winston 2003, Euskirchen et al., 2012). Interestingly BAF53 is essential for the expression of E6 and E7 when the viral genome has been integrated inf the host cell (Lee et al., 2011), but BAF57's requirement in this process is unknown. Based on Reactome, TRIP12 is involved in different pathways, and several are shared with some other cellular proteins targeted by the E6 proteins: antigen processing and ubiquitination/proteasome degradation (E6AP), class I MHC mediated antigen processing and presentation (E6AP), adaptive immune system (E6AP), immune system (P53, E6AP, MUC19, MX2).

The charged multi-vesicular body protein 4B (CHMP4B) was identified as a potential binder to HPV16 E6 for both AA and EP variants. The protein is a subunit of the endosomal sorting complex required for transport (ESCRT)-III complex during which it is involved in cytokinetic membrane abscission and a potential prognostic marker (Hu et al. 2014). In hepatocellular carcinomas, researchers determined overexpression of CHMP4B

and its involvement with cell cycle progression (Hu et al. 2014). Co-IP experiments also indicated that CHMP4B interacts with the inhibitory P53 isoform Δ133P53α that can block the activity of wild type P53 (Horikawa et al., 2014). Other interesting interactors of CHMP4B are IRF-2 (Hubel et al., 2019), BRCA2 (Malik et al., 2016) or E-cadherin (Guo et al., 2014). It appears that head and neck squamous cell carcinomas appear to display an increase in CHMP4B gene expression regardless if HPV was present or not, indicating a potential difference in the mechanism of action by HPV depending on the cells infected (Gollin 2014). CHMP4B and the ESCRT-III are important in membrane fission processes, including the budding of enveloped viruses (Strack et al., 2003). CHMP4B is associated with different Reactome pathways, and only one of them is shared with another protein present in *tables 4 and 5*: endosomal sorting complex required for transport (ESCRT), late endosomal microautophagy, budding and maturation of HIV virion, macroautophagy, autophagy, late phase of HIV life cycle, HIV life cycle, HCMV late events, HIV infection, HCMV infection, membrane trafficking, infectious disease, vesicle-mediated transport, disease (Muc19).

Our last protein of the interest is MX2, which is unique to EPE6. Proinflammatory signals are one of the hallmarks of cancer as cells could undergo necrosis recruiting inflammatory cells that in cancers can promote angiogenesis, proliferation, and invasiveness (Hanahan and Weinberg 2011). MX2 is a protein involved in innate immune response due to viral infections like human immunodeficiency virus (HIV). One of the main responsibilities of MX2 in HeLa cells is to permit G_1/S cell cycle progression (King et al. 2004). In HPV16 positive cells (W12 cell line), treatment with type I interferon (IFNα/β) inducible genes including MX2 resulted in loss of viral episomes. Interestingly

when viral DNA is integrated within host DNA, IFNα/β inducible genes like MX2 fail to inhibit E6 expression. Dysregulated E6 expression resulted in inhibition of IFNα/β inducible antiviral genes such as MX2 and activation of the TGF-β pathway (Pett et al. 2006). Based on Y2H experiments, MX2 interacts with the histone-lysine N-methyltransferase EHMT2 (Rolland et al., 2014), which increases P53-dependant expression of pro-apoptotic genes (i.e. Bax and Puma) and is a known interactor of p300/CBP (Rada et al., 2017). MX2 is associated with several Reactome pathways: ISG15 antiviral mechanism, antiviral mechanism by IFN-stimulated genes, IFNα/β signaling. IFN signaling is unique to the proteins in *table 4*. However, the immune system is shared with Muc19, TRIP12, P53 and E6AP, while cytokine signaling within the immune system also contains P53.

The "*Protein-Pathway Method"* provided a broader approach to identification of potentially interesting HPV16 E6 interacting proteins. Although seemingly broader—not all proteins identified using the "*Peptide Method"* were present in the final protein-pathway approach (*Table 6-7)* due to the fact that a final cleaning step was done with Reactome to eliminate any overlapping pathways also found non-specifically in the non-transduced control PHFKs. The reasoning for this was that in a given pathway, not all proteins would be targeted by E6 and that some may "stick" to the capturing beads during the Co-IP process and as a result, be detected in control PHFKs. In total, 171 unique proteins were identified in AAE6 and EPE6 collectively via the protein-pathway approach. AAE6 samples identified 110 proteins, while EPE6 samples identified 81 proteins. Of the 171 proteins identified, only 20 appeared in both AA and EP samples. Furthermore, when looking for protein overlap between the peptide method and the protein-pathway method

only three of proteins were identified namely the charged multivesicular body protein 4b (CHMP4B, Uniprot identifier: Q9H444), the cellular tumour antigen P53 (TP53, Uniprot identifier: P04637) and the ubiquitin-like protein 4A (UBL4A, Uniprot identifier: P11441).

Table 6 – Heat map of Protein-Pathway Method showing AAE6-targeted proteins unique to PHFK-HA. Number of peptides are shown in each coloured box. Bolded Proteins are also targeted by EPE6 (*Table 7*).

Protein Name	AAE6 DMSO T1	AAE6 DMSO T2	AAE6 MG132 T1	AAE6 MG132 T2
AAKG2			1	
ABCA6			1	
AFAD				1
AKAP9			1	
APC	1			
APC5	1			
APOD			1	
BDP1			1	
BUB1B			1	
CAR11	1			
CASPA			1	
CDK8		1		
CENPR			1	
CHM4B	1	2	1	3
CLCB	1			
CNPY3		2		2
CP19A			1	
CPE1A	2			
CPIN1			1	
CPSM	2			
CSRP1			1	
CTBP2	1			
CTND1	1			
CX6B1	1			
CXA9	1			
CYLD				1
DAND5	1			
DAPLE	1			
DLP1	1			
DNM1L	1			
DUS9	1			
DVL2		1		
EPC1	1		1	
4F2	1			
F263			1	
FBX32	1			
FBXL18	1			
FBXW7	1			
FIBA	1		1	
FOS	1		1	
FPGT			1	
GALK2	1			
GLCNE			1	
HBB		1		1
HCK	1			
HS105	1			
IF172	1			
IFNA1	1			
IL12R				1
IL1R1	1			
IPO5	1			
IPYR	1			
IPYR2			1	
ITA6	1			
KAT2A	1			

Protein Name	AAE6 DMSO T1	AAE6 DMSO T2	AAE6 MG132 T1	AAE6 MG132 T2
KCNB1	1			
KDM3B			1	
KGP2	1			
KIF5A			1	
KINH			1	
LEPR	1			
LIN37			1	
MASP2	1			
MRP3			1	
MTOR	1			
MTREX	1			
MUTYH			1	
NDUBA	1			
NEK7	1			
NFAC2			1	
NFYA		1		1
NMDE1			1	
NOSIP		1		
NPHP3	1			
NUBP1			1	
OTOF	1			
P53		1	1	
PCGF5			1	
PDE5A	1			
PDE6B			1	
PI2R	1			
PLA2R			1	
PLB1			1	
PRDM1			1	
RBNS5		1		
RBX1			1	
REV3L			1	
RN213	1			
ROBO3			1	
RPA2	2			
S12A6	1			
S22A4			1	
SFRP1	1			
SH3K1	1			
SLIT1	1		1	
SPSB2				1
TAP26				1
TCPA	1			
TEST			1	
TFDP1			1	
THIO	1		1	
TNR16	1			
TNR6C	1			
TRFL	1			
TT30B			2	
TTL10		1		
TTLL3			1	
UBA7	1			
UNC5B	1			
WDR35	1			

Table 7 – Heat map Protein-Pathway Method showing EPE6-targeted proteins unique to PHFK-HA proteins/pathways. Number of peptides are shown in each coloured box. Bolded Proteins are also targeted by AAE6 (*Table 6)*.

Protein Name	EPE6 DMSO T2	EPE6 DMSO T2	EPE6 MG132 T1	EPE6 MG132 T2
ACACB	1			
ADA12			1	
ADRB2	1			
AOFA			1	
BAK			1	
BDP1	1		1	
CAC1E				1
CAR11	1			
CBPA4			1	
CDK9	1			
CENPR	1		1	
CHM4B				5
CNPY3		1		2
CNTN4			1	
CTIP	1			
CUL3	1			
CX6B1			1	
DAP2B			1	
DDX58			1	
DNJC3				1
DPOE1	1			
DVL2				1
ENPP6			1	
EPC1	1			
FBW1A			1	
FIBA	1		1	
GALT3			1	
GROA			1	
HBA				1
HBB		1		1
ICAM3			1	
IF172	1			
IGF1			1	
IRAK4	1			
IRS1			1	
LIN37	1		1	
LRP1			1	
LSR				1
M3K1	1			
MRE11			1	
MUC20			1	

Protein Name	EPE6 DMSO T2	EPE6 DMSO T2	EPE6 MG132 T1	EPE6 MG132 T2
NDUAA			1	
NFYA				1
ODB2	1			
ODHX			1	
P53			4	
P85A		1		
PCNA			1	
PDE5A	1			
PDP1			1	
PGTA			1	
PIK3CG			1	
PLXA2	1			
PRGC1	1			
PTPRF			1	
RBP10			1	
ROBO2			1	
RPA34				1
RPE65			1	
S15A3	1			
SDK1	1			
SEMA3A	1			
SH3K1	1			
SMC2			1	
SNAPC1			1	
SNAPC4	1			
TAP26				1
TBB3		1		1
TBCA			1	
TBCD	1			
TEST			1	
TGO1	1			
THIO	1		1	
TLE1			1	
TRIPB	2			
TTL10				1
TTLL3	1			
UBL4A	1			2
WWTR1			1	
XRP2	1			
ZO2				1

Table 8 – Proteins targeted by both AAE6 and EPE6 identified using the Protein-Pathway method.

Protein Name	AAE6 DMSO T1	AAE6 DMSO T2	AAE6 MG132 T1	AAE6 MG132 T2	EPE6 DMSO T1	EPE6 DMSO T2	EPE6 MG132 T1	EPE6 MG132 T2
BDP1			1		1		1	
CAR11	1							
CENPR			1		1		1	
CHM4B	1	2	1	3				5
CX6B1	1							
DVL2		1						1
EPC1	1		1		1			
FIBA	1		1		1		1	
HBB		1		1		1		1
IF172	1				1			
LIN37			1		1		1	
NFYA		1		1				1
P53		1	1				4	
PDE5A	1				1			
SH3K1	1				1			
TAP26				1				1
TEST			1				1	
THIO	1		1		1		1	
TTL10		1						1
TTLL3			1		1			

Pathways targeted by AAE6 and EPE6 were independently analyzed using Reactome (*Table 9 and 10*). While the p-value was significant throughout for both sub-lineages, the false discovery rate (FDR) was statistically significant in 18/25 pathways for the AA lineage only. Findings will be discussed in the light of all 25 most significant pathways for both sub-lineages since values are relative only to currently known pathways in the scientific literature.

AAE6 was associated with 11 Notch1 signaling pathways (*Table 9 Rows 1, 2, 3, 9, 10, 11, 12, 13, 15, 16 and 22*) with the identification of several interacting proteins: cyclin-dependent kinase 8 (CDK8), histone acetyltransferase KAT2A, E3 ubiquitin-protein ligase RBX1 and two isoforms of the F-box/WD repeat-containing protein 7 FBW1B isoforms 1 and 4. CDK8 phosphorylates Notch rendering it for ubiquitination and degradation through the proteasome. Consequently, mastermind-like protein 1 (MAML1) does not acetylate Notch, and its transcription is not enhanced by p300 (Popko-Scibor et al. 2011). RBX1 and FBW1B isoforms were found to participate in Wnt signaling (*Table 9 Row 14*), providing evidence that these two cancer pathways communicate with one another. Wnt/beta-catenin signaling plays a role in development and adult homeostasis. In the latter, it is mostly inactive which is controlled by several kinases such as glycogen synthase kinase 3 (GSK3), casein kinase 1 (CK1), axin and adenomatous polyposis coli (APC), a tumour suppressor gene often mutated in colon cancer (Verheyen & Gottardi 2010 and references therein). Interestingly, APC was found once with just one unique peptide, yet it also appeared in Wnt signaling pathway targeted by AAE6 (*Table 9 Row 14*). Other notable findings identified from AAE6 proteins using the protein-pathway method were binders belonging to the Defective base excision repair (BER) associated with MUTYH pathway i.e., adenine

DNA glycosylase isoforms 3 and 6, TP53 regulates metabolic genes: P53, serine/threonine-protein kinase mTOR (MTOR), thioredoxin (THIO), trinucleotide repeat-containing gene 6C protein (TNR6C), cytochrome c oxidase subunit 6B1 (CX6B1) and 5'-AMP-activated protein kinase subunit gamma-2 (AAKG2). MUTYH germline mutations of the BER pathways cause MUTYH-associated polyposis (MAP), a disorder similar to familial adenomatous polyposis (FAP), caused by mutations in the APC gene (Mazzei et al., 2013). The TP53 regulates metabolic genes pathway also communicates with WNT signaling via Trinucleotide repeat-containing gene 6C protein (TNR6C). As such, AAE6-targeted cellular proteins derive from the axis of WNT and Notch1 signaling, as well as WNT signaling and TP53, regulates metabolic genes. The most striking candidates for this axis are RBX1, FBW1B, KAT2A, and CDK8 due to their presence in 8 to 11 pathways. The fact that the TP53 regulates metabolic genes pathway (*Table 9 Row 8*) is targeted by AAE6 in this study may explain our finding that this E6 variant deregulates cellular metabolism through the Warburg effect (Richards et al. 2010, Cuninghame et al. 2017). The presence of AAE6 targeted proteins in the Defective base excision repair (BER) associated with MUTYH strengthens our previous findings that the AA (D2/D3) sub-lineage seems to integrate earlier into the host genome than EP (A1) as evidenced in wet lab studies (Jackson et al. 2014, 2016). Indeed, another group reported that BER is essential for the HIV provirus DNA to integrate into the host genome, making the analogy with transposable elements (Yoder et al. 2011), which have a lot in common with viruses (Jackson et al. 2020, in preparation). Finally, Forkhead box protein O1 (FOXO)-mediated transcription also overlaps between the two variants via nuclear transcription factor Y subunit alpha NFYA. FOXO transcription factors act in pathways controlling cell survival, growth,

differentiation, and metabolism in various scenarios, such as growth factor deprivation, starvation, and oxidative stress (Eijkelenboom & Burgering 2013).

Table 9 – Most significant pathways found in AAE6 targeted proteins using Reactome. Pathways in bold have a significant false detection rate (FDR).

Pathway name	#Entities found	Entities pValue	Entities FDR*	Identified Entities	Protein Name
1. Loss of Function of FBXW7 in Cancer and NOTCH1 Signaling	3/6	3.31E-05	0.012	P62877 Q969H0-1 Q969H0-4	E3 ubiquitin-protein ligase *RBX1* F-box/WD repeat-containing protein 7 *FBXW7* isoform 1 F-box/WD repeat-containing protein 7 *FBXW7* isoform 4
2. FBXW7 Mutants and NOTCH1 in Cancer	3/6	3.31E-05	0.012	P62877 Q969H0-1 Q969H0-4	E3 ubiquitin-protein ligase *RBX1* F-box/WD repeat-containing protein 7 *FBXW7* isoform 1 F-box/WD repeat-containing protein 7 *FBXW7* isoform 4
3. NOTCH1 Intracellular Domain Regulates Transcription	5/48	1.30E-04	0.019	P49336 P62877 Q969H0-1 Q969H0-4 Q92830	Cyclin-dependent kinase 8 *CDK8* E3 ubiquitin-protein ligase *RBX1* F-box/WD repeat-containing protein 7 *FBXW7* isoform 1 F-box/WD repeat-containing protein 7 *FBXW7* isoform 4 Histone acetyltransferase *KAT2A*
4. Regulation of TP53 Expression	2/2	1.94E-04	0.019	P04637 O75626	Cellular tumour antigen *P53* PR domain zinc finger protein 1 *PRDM1*
5. Defective Base Excision Repair Associated with MUTYH	2/2	1.94E-04	0.019	Q9UIF7-3 Q9UIF7-6	Adenine DNA glycosylase *MUTYH* isoform 3 Adenine DNA glycosylase *MUTYH* isoform 6
6. Pyrophosphate hydrolysis	2/2	1.94E-04	0.019	Q9H2U2 Q15181	Inorganic pyrophosphatase 2, mitochondrial *IPYR2* Inorganic pyrophosphatase *IPYR*
7. Defective MUTYH substrate processing	2/2	1.94E-04	0.019	Q9UIF7-3 Q9UIF7-6	Adenine DNA glycosylase *MUTYH* isoform 3 Adenine DNA glycosylase *MUTYH* isoform 6
8. TP53 Regulates Metabolic Genes	6/88	2.70E-04	0.019	P42345 Q9UGJ0 P04637 P14854 P10599 Q9HCJ0	Serine/threonine-protein kinase mTOR MTOR 5'-AMP-activated protein kinase subunit gamma-2 *AAKG2* Cellular tumour antigen *P53* Cytochrome c oxidase subunit 6B1 *CX6B1* Thioredoxin *THIO* Trinucleotide repeat-containing gene 6C protein *TNR6C*
9. Constitutive Signaling by NOTCH1 HD+PEST Domain Mutants	5/59	3.34E-04	0.019	P49336 P62877 Q969H0-1 Q969H0-4 Q92830	Cyclin-dependent kinase 8 *CDK8* E3 ubiquitin-protein ligase *RBX1* F-box/WD repeat-containing protein 7 *FBXW7* isoform 1 F-box/WD repeat-containing protein 7 *FBXW7* isoform 4 Histone acetyltransferase *KAT2A*
10. Signaling by NOTCH1 HD+PEST Domain Mutants in Cancer	5/59	3.34E-04	0.019	P49336 P62877 Q969H0-1 Q969H0-4 Q92830	Cyclin-dependent kinase 8 *CDK8* E3 ubiquitin-protein ligase *RBX1* F-box/WD repeat-containing protein 7 *FBXW7* isoform 1 F-box/WD repeat-containing protein 7 *FBXW7* isoform 4 Histone acetyltransferase *KAT2A*

Pathway name	#Entities found	Entities pValue	Entities FDR*	Identified Entities	Protein Name
11. Signaling by NOTCH1 PEST Domain Mutants in Cancer	5/59	3.34E-04	0.019	P49336 P62877 Q969H0-1 Q969H0-4 Q92830	Cyclin-dependent kinase 8 *CDK8* E3 ubiquitin-protein ligase *RBX1* F-box/WD repeat-containing protein 7 *FBXW7* isoform 1 F-box/WD repeat-containing protein 7 *FBXW7* isoform 4 Histone acetyltransferase *KAT2A*
12. Constitutive Signaling by NOTCH1 PEST Domain Mutants	5/59	3.34E-04	0.019	P49336 P62877 Q969H0-1 Q969H0-4 Q92830	Cyclin-dependent kinase 8 *CDK8* E3 ubiquitin-protein ligase *RBX1* F-box/WD repeat-containing protein 7 *FBXW7* isoform 1 F-box/WD repeat-containing protein 7 *FBXW7* isoform 4 Histone acetyltransferase *KAT2A*
13. Signaling by NOTCH1 in Cancer	5/59	3.34E-04	0.019	P49336 P62877 Q969H0-1 Q969H0-4 Q92830	Cyclin-dependent kinase 8 *CDK8* E3 ubiquitin-protein ligase *RBX1* F-box/WD repeat-containing protein 7 *FBXW7* isoform 1 F-box/WD repeat-containing protein 7 *FBXW7* isoform 4 Histone acetyltransferase *KAT2A*
14. Signaling by WNT	10/299	8.48E-04	0.044	P25054 Q13237 P09497 P62877 P56545 Q9P219 P35913 Q8N474 O14641 Q9HCJ0	Adenomatous polyposis coli protein APC cGMP-dependent protein kinase 2 *KGP2* Clathrin light chain B *CLCB* E3 ubiquitin-protein ligase *RBX1* PC-terminal-binding protein 2 *CTBP2* Protein Daple *DAPLE* Rod cGMP-specific 3',5'-cyclic phosphodiesterase subunit beta *PDE6B* Secreted frizzled-related protein 1 *SFRP1* Segment polarity protein dishevelled homolog *DVL2* Trinucleotide repeat-containing gene 6C protein *TNR6C*
15. Signaling by NOTCH1	5/74	9.23E-04	0.045	P49336 P62877 Q969H0-1 Q969H0-4 Q92830	Cyclin-dependent kinase 8 *CDK8* E3 ubiquitin-protein ligase *RBX1* F-box/WD repeat-containing protein 7 *FBXW7* isoform 1 F-box/WD repeat-containing protein 7 *FBXW7* isoform 4 Histone acetyltransferase *KAT2A*
16. Signaling by NOTCH	8/205	0.0011	0.049	P04637 P49336 P62877 Q969H0-1 Q969H0-4 Q92830 Q14186 Q9HCJ0	Cellular tumour antigen *P53* Cyclin-dependent kinase 8 *CDK8* E3 ubiquitin-protein ligase *RBX1* F-box/WD repeat-containing protein 7 *FBXW7* isoform 1 F-box/WD repeat-containing protein 7 *FBXW7* isoform 4 Histone acetyltransferase *KAT2A* Transcription factor Dp-1 *TFDP1* Trinucleotide repeat-containing gene 6C protein *TNR6C*
17. Negative regulation of TCF-dependent signaling by DVL-interacting proteins	2/5	0.0012	0.049	Q9P219 O14641	Protein Daple *DAPLE* Segment polarity protein dishevelled homolog *DVL2*
18. Activation of NOXA and translocation to mitochondria	2/5	0.0012	0.049	P04637 Q14186	Cellular tumour antigen *P53* Transcription factor Dp-1 *TFDP1*

Pathway name	#Entities found	Entities pValue	Entities FDR*	Identified Entities	Protein Name
19. Diseases of Base Excision Repair	2/7	0.0023	0.087	Q9UIF7-3 Q9UIF7-6	Adenine DNA glycosylase *MUTYH* isoform 3 Adenine DNA glycosylase *MUTYH* isoform 6
20. Insulin processing	3/27	0.0026	0.091	P16870 Q12840 P33176	Carboxypeptidase E *CBPE* Kinesin heavy chain isoform 5A *KIF5A* Kinesin-1 heavy chain *KINH*
21. Oxidative Stress Induced Senescence	5/94	0.0026	0.091	P04637 P01100 P10599 Q14186 Q9HCJ0	Cellular tumour antigen *P53* Proto-oncogene c-Fos *FOS* Thioredoxin *THIO* Transcription factor Dp-1 *TFDP1* Trinucleotide repeat-containing gene 6C protein *TNR6C*
22. Pre-NOTCH Transcription and Translation	4/62	0.0036	0.113	P04637 Q92830 Q14186 Q9HCJ0	Cellular tumour antigen *P53* Histone acetyltransferase *KAT2A* Transcription factor Dp-1 *TFDP1* Trinucleotide repeat-containing gene 6C protein *TNR6C*
23. Activation of PUMA and translocation to mitochondria	2/9	0.0038	0.113	P04637 Q14186	Cellular tumour antigen *P53* Transcription factor Dp-1 *TFDP1*
24. Transcriptional Regulation by TP53	10/367	0.0038	0.113	Q9UGJ0 Q92851 P04637 P14854 O75626 P01100 P42345 P10599 Q14186 Q9HCJ0	5'-AMP-activated protein kinase subunit gamma-2 *AAKG2* Caspase-10 *CASP10* Cellular tumour antigen *P53* Cytochrome c oxidase subunit 6B1 *CX6B1* PR domain zinc finger protein 1 *PRDM1* Proto-oncogene c-Fos *FOS* Serine/threonine-protein kinase mTOR MTOR Thioredoxin *THIO* Transcription factor Dp-1 *TFDP1* Trinucleotide repeat-containing gene 6C protein *TNR6C*
25. FOXO-mediated transcription	4/66	0.0045	0.124	Q8N139 Q969P5 P23511 P10599	ATP-binding cassette sub-family A member 6 *ABCA6* F-box only protein 32 *FBX32* Nuclear transcription factor Y subunit alpha *NFYA* Thioredoxin *THIO*

Up to 8 EPE6 interactors are part of homology directed repair (HDR) DNA damage pathways (*Table 9 Rows 5,9,19, and 22*): P53, THIO, proliferating cell nuclear antigen (PCNA), CX6B1, double-strand break repair protein MRE11, CDK9, geranylgeranyl transferase type-2 subunit alpha PGTA and DNA endonuclease RBBP8. PCNA also interacts with the adenine DNA glycosylase MUTYH pathway (Parker et al. 2001) linking the two E6 variants under study. While EPE6 does target different cancer pathways than AAE6 namely the Hippo, PIK3/AKT, MET and epidermal growth factor receptor pathways (*Table 9 Rows 1,3,7,10,14, and 24*), Hippo communicates with the Wnt pathway through segment polarity protein dishevelled homolog DVL-2 (EPE6's Signaling by Hippo pathway and AAE6's pathways Signaling by WNT and Negative regulation of TGF-dependent signaling by DVL-interacting proteins). The most notable EPE6 target seems to be phosphatidylinositol 3-kinase regulatory subunit alpha PIK3R1 detected in 5 pathways related to PI3K, MET and EGFR (*Table 9 Rows 3,7,10,14, and 24*), which all communicate with each other. Wet lab studies are warranted for the above findings to finally establish any differential binding partners for the two sub-lineage E6s.

Table 10 – Most significant pathways of EPE6 targeted proteins using Reactome**.**

Pathway name	#Entities found	Entities pValue	Entities FDR*	Identified Entities	Mapped entities
1. Signaling by Hippo	3/20	4.21E-04	0.144	O14641 Q9UDY2 Q9GZV5	Segment polarity protein disheveled homolog *DVL-2* Tight junction protein *ZO-2* WW domain-containing transcription regulator protein 1 *WWTR1*
2. Post-chaperonin tubulin folding pathway	3/23	6.31E-04	0.144	Q13509 O75347 Q9BTW9	Tubulin beta-3 chain *TUBB3* Tubulin-specific chaperone A *TBCA* Tubulin-specific chaperone D *TBCD*
3. Activated NTRK3 signals through PI3K	2/6	8.81E-04	0.144	P35568 P27986	Insulin receptor substrate 1 *IRS1* Phosphatidylinositol 3-kinase regulatory subunit alpha *PIK3R1*
4. RNA Polymerase III Transcription Initiation from Type 3 Promoter	3/28	0.0011	0.144	Q16533 Q5SXM2 A6H8Y1	snRNA-activating protein complex subunit 1 *SNAPC1* snRNA-activating protein complex subunit 4 *SNAPC4* Transcription factor TFIIIB component B" homolog *BDP1*
5. HDR through Homologous Recombination (HRR)	4/66	0.0013	0.144	Q99708 Q07864 P49959 P12004	DNA endonuclease *RBBP8* DNA polymerase epsilon catalytic subunit *POLE* Double strand break repair protein *MRE11* Proliferating cell nuclear antigen *PCNA*
6. Erythrocytes take up oxygen and release carbon dioxide	2/8	0.0016	0.144	P69905 P68871	Hemoglobin subunit alpha *HBA1* Hemoglobin subunit beta *HBB*
7. PI3K/AKT activation	2/9	0.0020	0.144	P35568 P27986	Insulin receptor substrate 1 *IRS1* Phosphatidylinositol 3-kinase regulatory subunit alpha *PIK3R1*
8. RNA Polymerase III Transcription Initiation	3/36	0.0023	0.144	Q16533 Q5SXM2 A6H8Y1	snRNA-activating protein complex subunit 1 *SNAPC1* snRNA-activating protein complex subunit 4 *SNAPC4* Transcription factor TFIIIB component B" homolog *BDP1*
9. HDR through MMEJ (alt-NHEJ)	2/10	0.0024	0.144	Q99708 P49959	DNA endonuclease *RBBP8* Double strand break repair protein *MRE11*
10. Signaling by MET	4/80	0.0027	0.144	Q8N307 P27986 Q6VN20 Q96B97	Mucin-20 *MUC20* Phosphatidylinositol 3-kinase regulatory subunit alpha *PIK3R1* Ran-binding protein 10 *RANBP10* SH3 domain-containing kinase-binding protein 1 *SH3KBP1*

Pathway name	#Entities found	Entities pValue	Entities FDR*	Identified Entities	Mapped entities
11. Metabolism of proteins	25/2012	0.0028	0.144	O95786 P07550 P04637 Q8IWV2 Q13618 Q13217 Q9Y297 P02671 Q92696 Q6ZVT0 P05019 P15088 Q8N307 P23511 Q9UBK2 Q14435 P12004 Q9Y6M0 P10599 Q9P031 Q5JRA6 Q13509 Q9Y4R7 O75347 Q9BTW9	Antiviral innate immune response receptor RIG-I *DDX58* Beta-2 adrenergic receptor *ADRB2* Cellular tumour antigen *P53* Contactin-4 *CNTN4* Cullin-3 *CUL3* DnaJ homolog subfamily C member 3 *DNAJC3* F-box/WD repeat-containing protein 1A *BTRC* Fibrinogen alpha chain *FGA* Geranylgeranyl transferase type-2 subunit alpha *RABGGATA* Inactive polyglycylase *TTLL10* Insulin-like growth factor I *IGF1* Mast cell carboxypeptidase A *CPA3* Mucin-20 *MUC20* Nuclear transcription factor Y subunit alpha *NFYA* Peroxisome proliferator-activated receptor gamma coactivator 1-alpha *PPARGC1A* Polypeptide N-acetylgalactosaminyltransferase 3 *GLNT3* Proliferating cell nuclear antigen *PCNA* Testisin *PRSS21* Thioredoxin *THIO* Thyroid transcription factor 1-associated protein 26 *CCDC59* Transport and Golgi organization protein 1 homolog *MIA3* Tubulin beta-3 chain *TUBB3* Tubulin monoglycylase *TTLL3* Tubulin-specific chaperone A *TBCA* Tubulin-specific chaperone D *TBCD*
12. RNA Polymerase III Abortive and Retractive Initiation	3/41	0.0033	0.144	Q16533 Q5SXM2 A6H8Y1	snRNA-activating protein complex subunit 1 *SNAPC1* snRNA-activating protein complex subunit 4 *SNAPC4* Transcription factor TFIIIB component B" homolog *BDP1*
13. RNA Polymerase III Transcription	3/41	0.0033	0.144	Q16533 Q5SXM2 A6H8Y1	snRNA-activating protein complex subunit 1 *SNAPC1* snRNA-activating protein complex subunit 4 *SNAPC4* Transcription factor TFIIIB component B" homolog *BDP1*
14. Erythropoietin activates Phosphoinositide-3-kinase (PI3K)	2/12	0.0034	0.144	P27986 P48736	Phosphatidylinositol 3-kinase regulatory subunit alpha *PIK3R1* Thyroid transcription factor 1-associated protein 26 *CCDC59*
15. MET activates RAS signaling	2/12	0.0034	0.144	Q8N307 Q6VN20	Mucin-20 *MUC20* Ran-binding protein 10 *RANBP10*
16. Erythrocytes take up carbon dioxide and release oxygen	2/12	0.0034	0.144	P69905 P68871	Hemoglobin subunit alpha *HBA1* Hemoglobin subunit beta *HBB*
17. O2/CO2 exchange in erythrocytes	2/12	0.0034	0.144	P69905 P68871	Hemoglobin subunit alpha *HBA1* Hemoglobin subunit beta *HBB*
18. Carboxyterminal post-translational modifications of tubulin	3/43	0.0037	0.145	Q6ZVT0 Q13509 Q9Y4R7	Inactive polyglycylase *TTLL10* Tubulin beta-3 chain *TUBB3* Tubulin monoglycylase *TTLL3*

Pathway name	#Entities found	Entities pValue	Entities FDR*	Identified Entities	Mapped entities
19. DNA Double-Strand Break Repair	5/148	0.0043	0.146	P04637 Q99708 Q07864 P49959 P12004	Cellular tumour antigen *P53* DNA endonuclease *RBBP8* DNA polymerase epsilon catalytic subunit *POLE* Double strand break repair protein *MRE11* Proliferating cell nuclear antigen *PCNA*
20. TP53 regulates transcription of several additional cell death genes whose specific roles in P53-dependent apoptosis remain uncertain	2/14	0.0046	0.146	P04637 Q92696	Cellular tumour antigen *P53* Geranylgeranyl transferase type-2 subunit alpha *RABGGATA*
21. SEMA3A-Plexin repulsion signaling by inhibiting Integrin adhesion	2/14	0.0046	0.146	O75051 Q14563	Plexin-A2 *PLXNA2* Semaphorin-3A *SEMA3*
22. Transcriptional Regulation by TP53	8/367	0.0047	0.146	P04637 P50750 P14854 Q99708 P49959 Q92696 P12004 P10599	Cellular tumour antigen *P53* Cyclin-dependent kinase 9 *CDK9* Cytochrome c oxidase subunit 6B1 *COX6B1* DNA endonuclease *RBBP8* Double strand break repair protein *MRE11* Geranylgeranyl transferase type-2 subunit alpha *RABGGATA* Proliferating cell nuclear antigen *PCNA* Thioredoxin *THIO*
23. Toll-like Receptor Cascades	5/156	0.0053	0.146	Q9Y297 P02671 Q9NWZ3 Q13233 Q9BT09	F-box/WD repeat-containing protein 1A *BTRC* Fibrinogen alpha chain *FGA* Interleukin-1 receptor-associated kinase 4 *IRAK4* Mitogen-activated protein kinase kinase kinase 1 *MAP3K1* Protein canopy homolog 3 *CNPY3*
24. Signaling by EGFR	3/51	0.0060	0.146	O43184 P27986 Q96B97	Disintegrin and metalloproteinase domain-containing protein 12 *ADAM12* Phosphatidylinositol 3-kinase regulatory subunit alpha *PIK3R1* SH3 domain-containing kinase-binding protein 1 *SH3KBP1*
25. CRMPs in Sema3A signaling	2/16	0.0060	0.146	O75051 Q14563	Plexin-A2 *PLXNA2* Semaphorin-3A *SEMA3*

4 Discussion

I have successfully developed a method to selectively pull-down HPV16 E6 variants and their interacting partners. Using the peptide method, seven novel candidates may interact with E6 variants. Three proteins that potentially bind to AAE6 exclusively: INO80B, PROK2, and RPS6KA4. These proteins may provide insight into new mechanisms which E6 variants can affect chromosome remodelling, angiogenesis and dysregulated metabolism. One protein (MX2) potentially binds exclusively to EPE6 providing evidence that AAE6 may evade a component of the innate immune response. Finally, there were 3 proteins identified that may interact with both AA and EP. These proteins unravel new mechanisms for how E6 may degrade P53, and cause cell cycle progression. Using the Protein-Pathway method, AAE6 was found to interact with the NOTCH signalling pathway along with several cross-talking pathways (TP53 Regulates Metabolic Genes, Wnt signalling and Hypoxia signalling). This method also identified that AAE6 may interact with BER associated with MUTYH giving evidence that AAE6 can integrate earlier within the host genome compared to EPE6.

With my initial approach, Co-IP coupled to MALDI MS/MS, I successfully immunoprecipitated E6. However, the MS failed to identify the E6 protein itself as well as the classical known E6 interactors E6AP and P53. I therefore set out to optimize our Co-IP protocol and utilize an MS approach providing a user-friendly output amenable for bioinformatics using freely accessible software tools. The optimizations made to the original Co-IP protocol allowed for more effective isolation of 16 E6 variant interacting proteins. Many of the optimizations made resulted in a protocol similar to other groups (Brimer et al. 2007; White et al. 2012; Grace et al. 2017; Personal communication with Dr.

White). Each step of the protocol was optimized to isolate the maximum amount of E6 and interacting proteins in a MS compatible elution buffer. The elution of proteins from the beads was the step that needed to deviate from most of the literature protocols on E6 IP experiments as synthetic peptides did not succeed at releasing any antigens from the beads. It was interesting that most groups that used peptide elution methods needed to concentrate their samples using TCA precipitation. Our samples failed to be precipitated at most concentrations of TCA (10 % (w/v) – 100 % (w/v)). TCA precipitation resulted in no visible precipitation of proteins in any samples. This was most likely because our total eluted protein was below 0.016 mg of protein, the minimum mass of protein needed to see obvious protein precipitation (Ngo et al. 2015). Therefore, we needed to use an acidic buffer to elute proteins. This proved effective as we successfully minimized antibody leeching from the magnetic beads and eluted E6 (for AAE6 only) with its targeted proteins. Without MG132 treatment it is highly unlikely that we would see any evidence of P53 in eluted samples (Personal communication with Dr. White). The use of proteasome inhibitors only assisted to target sufficient quantities of P53 that we could visualize the protein in western blotting applications. In MS applications however, we found that we were able to see P53 in both samples at least once with MG132 treatment and once for AAE6 in DMSO treatment. Therefore, even though P53 appeared in both MG132 and DMSO treatments the use of a proteasome inhibitor increased the likelihood of identifying P53 (and other potential E6 targets) in MS. The final IP method is quick to prepare and has little variability in efficacy from using many premade reagents (MPER, cOmplete, HALT etc.) This method works well for the cell types we used and to generate 4 mg of protein extract and effectively target E6 binding proteins. It is important to remember that when using other cells this

method may be an effective starting point, but optimizations should be made to suit each researcher's needs. A complete comparison of the Co-IP optimization could not be done due to switching from MALDI MS to LC MS/MS. This was done because since the samples were quite complex and contained a large variety of proteins, we needed to use more sensitive equipment (Abersold and Mann 2003). By switching to LC MS/MS analysis our samples could be present in lower concentrations. The result of switching to LC MS/MS sample analysis was successful detection of E6 in AAE6 and consistent identification of known E6 interacting proteins. Using LC MS/MS we were able to generate confident and accurate results.

The process of filtering proteins poses a substantial limitation in bioinformatics as there are many ways to analyze MS datasets. Most groups filter datasets by removing proteins involved in negative controls, along with setting a minimum number of peptides required to be a valid result. The use of the protein-pathway method provided a more inclusive method for identifying potential E6 interacting proteins. Interestingly, however, the protein-pathway method also demonstrated that it is possible to manipulate datasets in a variety of ways. For instance, the input name for identified proteins was crucial. While using UniProt Identifiers allowed for obtaining the results seen in this thesis we obtained drastically different results when inputting the dataset into Reactome versus KEGG. We settled for the latter since this tool seemed the more recent. It is therefore necessary to state what format studies import their datasets for bioinformatic analysis because this can create issues in the replicability of results in the future.

This thesis aimed to develop a method to successfully pull-down and identify E6 variants along with associated binding partners. The protocol underwent several changes

culminating in a method similar to previous pull-downs by other research groups. I successfully identified AAE6 using this pull-down method and was able to identify several potential interactors that may help unravel new mechanisms utilized by HPV16 E6 to promote tumourigenesis. Further wet lab work is necessary to identify and unravel the validity of interacting proteins as well as E6's effects on each protein or pathway.

5 References

Aebersold R., Mann M. 2013. Mass spectrometry-based proteomics. Nature **422**: 198-207.

An J., Mo D., Liu H., Veena M. S., Srivastsan E. S., Massoumi R., Rettig M. B. 2008. Inactivation of the CYLD deubiquitinase by HPV E6 mediates hypoxia-induced NF-kB activation. Cancer Cell **14**(5): 394-407.

Androphy E. J., Hubbert N. L., Schiller J. T., Lowy D. R. 1987. Identification of the HPV-16 E6 protein from transformed mouse cells and human cervical carcinoma cell lines. EMBO **6**(4): 989-992.

Anjum R., Blenis J. 2008. The RSK family of kinases: emerging roles in cellular signalling. Nature Reviews Molecular Cell Biology **9**(10):

Berggård T., Linse S., James P. 2007. Methods for the detection and analysis of protein-protein interactions. Proteomics **7**: 2833-2842.

Berumen J., Ordoñez R. M., Lazcano E., Salmeron J., Galvan S. C., Estrada R. A., Yunes E., Garcia-Carranca A., Gonzalez-Lira G., Madrigal-de la Campa A. 2001. Asian-American variants of human papillomavirus 16 and risk for cervical cancer: a case-control study. Journal of the National Cancer Institute **93**(17): 1325-1330.

Bird Y., Obidiya O., Mahmood R., Nwankwo C., Moraros J. 2017 Human papillomavirus vaccination uptake in Canada: A systematic review and meta-analysis. International Journal of Preventive Medicine **8**: 71.

Bouvard V., Storey A., Pim D., Banks L. 1994. Characterization of the human papillomavirus E2 protein: evidence of *trans*-activation and *trans*-repression in cervical keratinocytes. EMBO **13**(22): 5451-5459.

Brimer N., Lyons Vande Pol S. B. 2007. Association of E6AP (UBE3A) with human papillomavirus type 11 E6 protein. Virology **358**(2007): 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**(43): 4639-4646.

Burk R. D., Harari A., Chen Z. 2013. Human papillomavirus genome variants. Virology **445**(0): 232-243.

Burk R. D., Terai M., Gravitt P. E., Brinton L. A., Kurman R. J., Barnes W. A., Greenberg M. D., Hadjimichael O. C., Fu L., McGowan L., Mortel R., Schwartz P. E., Hildesheim A. 2003. Distribution of Human Papillomavirus Types 16 and 18 Variants in Squamous Cell Carcinomas of the Cervix. Cancer Research **63**: 7215-7220.

Chaturvedi A. K., Engels E. A., Pfeiffer R. M., Hernandez B. Y., Xiao W., Kim E., Jiang B., Goodman M. T., Sibug-Saber M., Cozen W., Liu L., Lynch C. F., Wentzensen N., Jordan R. C., Altekruse S., Anderson W. F., Rosenberg P. S., Gillison M. L. 2011. Human Papillomavirus and rising oropharyngeal cancer incidence in the United States. Journal of Clinical Oncology **29**(32): 4294-4301.

Chen J. J., Reid C. E., Band V., Androphy E. J. 1995. Interaction of papillomavirus E6 oncoproteins with a putative calcium-binding protein. Science **269**(5223): 529-531.

Chen D., Shan J., Zhu WG., Qin J., Gu W. 2010. Transcription-independent ARF regulation in oncogenic stress-mediated P53 responses. Nature **464**(7288):624-7.

Cheung J. L. K., Cheung T. H., Yu M. Y., Chan P. K. S. 2013. Virological characteristics of cervical cancers carrying pure episomal form of HPV16 genome. Gynecologic Oncology **131**(2013): 374-379.

Chio I. I., Sasaki M., Ghazarian D., Moreno J., Done S., Ueda T., Inoue S., Chang Y. L., Chen N. J., Mak T. W. 2012. TRADD contributes to tumour suppression by regulating ULF-dependent p19Arf ubiquitylation. Nature Cell Biology **14**(6):625-33

Clifford G. M., Tenet V., Damien G., Alemany L., Pavón M. A., Chen Z., Yeager M., Cullen M., Boland J. F., Bass S., Steinberg M., Raine-Bennett T., Lorey T., Wentzensen N., Walker J., Zuna R., Schiffman M., Mirabello L. 2019. Human papillomavirus 13 sub-lineage dispersal and cervical cancer risk worldwide: whole viral genome sequences from 7116 HPV16-positive women. Papillomavirus Research **7**(2019): 67-74.

Cloutier P., Poitras C., Durand M., Hekmat O., Fiola-Masson É., Bouchard A., Faubert D., Chabot B., Coulombe B. 2017. R2TP/Prefoldin-like component RUVBL1/RUVBL2 directly interact with ZNHIT2 to regulate assembly of U5 small nuclear ribonucleoprotein. Nature Communications **8**(15615): 1-14.

Collado M., Serrano M. 2010. The TRIP from ULF to ARF. Cancer Cell **17**(4): 317-318.

Cuninghame S., Jackson R., Lees S. J., Zehbe I. 2017. Two common variants of human papillomavirus type 16 E6 differently dysregulate sugar metabolism and hypoxia signalling in permissive human keratinocytes. Journal of General Virology **98**(9): 2310-2319.

Datta D., Anbarasu K., Rajabather S., Priya R. S., Desai P., Mahalingam S. 2015. Nucleolar GTP-binding protein-1 (NGP-1) promotes G_2 to S phase transition by activating cyclin-dependent kinase inhibitor $p21^{Cip1/Waf1}$. The Journal of Biological Chemistry **290**(35): 21536-21552.

De Las Rivas J., Fontanillo C. 2010. Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. PLoS Computational Biology **6**(6): e1000807.

Deutsch E. W., Mendoza L., Shteynberg D., Slagel J., Sun Z., Moritz R. L. Trans-Proteomic Pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. Proteomics Clinical Applications **9**(0): 745-754.

Doorbar J. 2013. The E4 protein; structure, function and patterns of expression. Virology **445**(1): 80-98.

Du M., Fan X., Hong E., Chen J. J. 2002. Interaction of oncogenic papillomavirus E6 proteins with fibulin-1. Biochemical and Biophysical Research Communications **296**(2009): 962-969.

Eijkelenboom A., Burgering B. M. T. 2013. FOXOs: signalling integrators for homeostasis maintenance. Nature Reviews Molecular Cell Biology **18**(2017): 285-298.

Euskirchen G., Auerbach RK., Snyder M. 2012. SWI/SNF chromatin-remodeling factors: multiscale analyses and diverse functions. Journal of Biological Chemistry **287**(37): 30897-905

Fabregat A., Jupe S., Matthews L., Sifiropoulos K., Gillespie M., Garapati P., Haw R., Jassal B., Korninger F., May B., Milacic M., Roca C. D., Rothfels K., Sefilla C., Shamovsky V., Shorser S., Varusal T., Viteri G., Weiser J., Wu G., Stein L., Hermjakob H., D'Eustachio P. 2018. The reactome pathway knowledgebase. Nucleic Acids Research **46**: D649-D655.

Ferlay J., Colombet M., Soerjomataram I., Mathers C., Parkin D. M., Piñeros M., Znaor A., Bray F. 2019. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. International Journal of Cancer **144**(8): 1941 – 1953.

Filippova M., Johnson M. M., Bautista M., Filippov V., Fodor M., Tungteakkhun S. S., Williams K., Duerksen-Hughes P. J. 2007. The large and small isoforms of human papillomavirus type 16 E6 bind to and differentially affect procaspase 8 stability and activity. Journal of Virology **81**(8): 4116-4129.

Filippova M., Parkhurst L., Duerksen-Hughes P. J. 2004. The human papillomavirus 16 E6 protein binds to Fas-associated Death Domain and protects cells from Fas-triggered apoptosis. Journal of Biological Chemistry **279**(24): 25729-25744.

Filippova M., Song H., Connolly J. L., Dermody T. S., Duerksen-Hughes P. J. 2002. The human papillomavirus 16 E6 protein binds to tumor necrosis factor (TNF) R1 and protects cells from TNF-induced apoptosis. Journal of Biological Chemistry **277**: 21730-21739.

Free B. R., Hazelwood L. A., Sibley R. D. 2009. Identifying novel protein-protein interactions using co-immunoprecipitation and mass spectrometry. Current protocol neuroscience **0**(5): Unit-5.28.

Fukumoto C., Nakashima D., Kasamatsu A., Unozawa M., Shida-Sakazume T., Higo M., Ogawara K., Yokoe H., Shiiba M., Tanzawa H., Uzawa K. 2014. WWP2 is overexpressed in human oral cancer, determining tumor ize and poor prognosis in patients: downregulation of WWP2 inhibits the AKT signaling and tumor growth in mice. Oncoscience **1**(12): 807-820.

Gao Q., Kumar A., Srinivasan S., Singh L., Mukai H., Ono Y., Wazer D. E., Band V. 2000. PKN binds and phosphorylates human papillomavirus E6 oncoprotein. Journal of Biological Chemistry **275**: 14824-14830.

Gao. Q., Srinivasan S., Boyer S. N., Wazer D. E., Band V. 1999. The E6 oncoproteins of high-risk papillomaviruses bind to a novel putative GAP protein, E6TP1, and target it for degradation. Molecular and Cellular Biology **19**(1): 733-744.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M. R., Appel R. D., Bairoch A. 2005. *Protein identification and analysis tools on the ExPASy server* in John M. Walker ed. The proteomics protocols handbook. Humana Press. Full Text.

Gewin L., Myers H., Kiyono T., Galloway D. A. 2004. Identification of a novel telomerase repressor that interacts with the human papillomavirus type-16 E6/E6-AP complex. Genes and Development **18**: 2269-2282.

Gibb C. M., Jackson R., Mohammed S., Fiaidhi J., Zehbe I. 2019. Pathogen-Host Analysis Tool (PHAT): an integrative Platform to analyze next-generation sequencing data. Bioinformatics **35**(15): 2665-2667.

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 **19**: 5270-5280.

Gollin S. M. 2014. Cytogenetic alterations and their molecular genetic correlates in head and neck squamous cell carcinoma: A next generation window to the biology of disease. Genes Chromosomes and Cancer **53**(12): 972-990.

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

Greenberg R. A., Sobhian B., Pathania S., Cantor S. B., Nakatani Y., Livingston D. M. 2005. Multifactorial contributions to an acute DNA damage response by BRCA1/BARD1-containing complexes. Genes and Development **20**: 34-46.

Gross-Mesilaty S., Reinstein E., Bercovich B., Tobias K. E., Schwartz A. L., Kahana C., Ciechanover A. 1998. Basal and human papillomavirus E6 oncoprotein-induced degradation of Myc proteins by the ubiquitin pathway. PNAS **95**: 8058-8063.

Guo Z., Neilson LJ., Zhong H., Murray PS., Zanivan S., Zaidel-Bar R. 2014. E-cadherin interactome complexity and robustness resolved by quantitative proteomics. Science Signalling **7**(354): rs7.

Gundry R. L., White M. Y., Murray C. I., Kane L. A., Fu Q., Stanley B. A., Van Eyk J. E. 2010. Preparation of proteins and peptides for Mass spectrometry Analysis in a bottom-up proteomics workflow. Current Protocols in Molecular Biology. **90**(1): 10.25.1-10.25.23.

Hall M. T., Simms K. T., Lew J. B., Smith M. A., Brotherton J. M. L., Saville M., Frazer I. H., Canfell K. 2019. The projected timeframe until cervical cancer elimination in Australia: a modelling study. The Lancet Public Health **4**(1): E19-E27.

Hampson L., Li C., Oliver A. W., Kitchener H. C., Hampson I. N. 2004. The PDZ protein TIP-1 is a gain of function target of the HPV16 E6 oncoprotein. International Journal of Oncology **25**(2004):1249-1256.

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 degredation of DLG4/PSD95 by high-risk human papillomavirus type 18 E6 protein. Journal of Virology **81**(3): 1379-1389.

Harper D. M., DeMars L. R. 2017. HPV vaccines – a review of the first decade. Gynecologic Oncology **146**(1): 196-204.

Haupt Y., Maya R., Kazaz A., Oren M. 1997. Mdm2 promotes the rapid degradation of P53. Nature **387**: 296-299.

Heijink A. M., Talens F., Jae L. J., van Gijn S. E., Fehrmann R. S. N., Brummelkamp T. R., van Vugt M. A. T. M. 2019. BRCA2 deficiency instigates cGAS-mediated inflammatory signaling and confers sensitivity to tumor necrosis factor -alpha-mediated cytotoxicity. Nature Communications **10**(100): 1-14.

Higdon R., Kolker E. 2007. A predictive model for identifying proteins by a single peptide match. Bioinformatics **23**(3): 277-280.

Ho L., Chan S., Burk R. D., Das B. C., Fujinaga K., Icenocle J. P., Kahn T., Kiviat N., Lancaster W., Mavromara-Nazos P., Labropoulou V., Mitrani-Rosenbaum S., Norrild B., Pillai M. R., Stoerker J., Syrjaenen K., Syrjaenen S., Tay S., Villa L. L., Wheeler C. M., Williamson A., Bernard H. 1993. The genetic drift of human Papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. Journal of Virology **67**(11): 6413-6423.

Hochmann J., Sobrinho J. S., Villa L. L., Sichero L. 2016 The Asian-American variant of human papillomavirus type 16 exhibits higher activation of MAPK and PI3K/AKT signalling pathways, transformation, migration and invasion of primary human keratinocytes. Virology **492**(2016): 145-154.

Horikawa I., Fujita K., Jenkins LM., Hiyoshi Y., Mondal AM., Vojtesek B., Lane DP., Appella E., Harris CC. 2014. Autophagic degradation of the inhibitory P53 isoform Δ133P53α as a regulatory mechanism for P53-mediated senescence. Nature Communications **5**(4706): 1-11.

Howie H. L., Koop J. I., Weese J., Robinson K., Wipf G., Kim L., Galloway D. A. 2011. Beta-HPV5 and 8 E6 promote p300 degradation by blocking AKT/p300 association. PLoS Pathogens **7**(8): e1002211.

Hsu C. H., Peng K. L., Jhang H. C., Lin C. H., Wu S. Y., Chiang C. M., Lee S. C., Yu W. C., Juan L. J. 2012. The HPV E6 oncoprotein targets histone methyltransferases for modulating specific gene transcription. Oncogene **31**(18): 2335-2349.

Hu B., Jiang D., Chen Y., Wei L., Zhang S., Zhao F., Ni R., Lu C., Wan C. 2014. High CHMP4B expression is associated with accelerated cell proliferation and resistance to doxorubicin. Tumor biology **36**(4): 2569-2581.

Hu J., Liu J., Chen A., Lyu J., Ai G., Zheng Q., Sun Y., Chen C., Wang J., Qiu J., Wu Y., Cheng J., Shi X., Song L. 2016. Ino80 promotes cervical cancer tumorigenesis by activating Nanog expression. Oncotarget **7**(44): 72250 – 72262.

Hubel P., Urban C., Bergant V., Schneider WM., Knauer B., Stukalov A., Scaturro P., Mann A., Brunotte L., Hoffmann HH., Schoggins JW., Schwemmle M., Mann M., Rice CM., Pichlmair A. 2019. A protein-interaction network of interferon-stimulated genes extends the innate immune system landscape. Nature Immunology 20(4):493-502.

Hughes F. J., Romanos M. A., 1993. E1 protein of human papillomavirus is a DNA helicase/ATPase. Nucleic Acids Research **21**(25) 5817-5823.

Huibregtse J. M., Scheffner M., Howley P. M. 1991. A cellular protein mediates association of P53 with the E6 oncoprotein of human papillomavirus types 16 or 18. The EMBO Journal **10**(13): 4129-4135.

Huttlin EL., Ting L., Bruckner RJ., Gebreab F., Gygi MP., Szpyt J., Tam S., Zarraga G., Colby G., Baltier K., Dong R., Guarani V., Vaites LP., Ordureau A., Rad

R., Erickson BK., Wühr M., Chick J., Zhai B., Kolippakkam D., Mintseris J., Obar RA., Harris T., Artavanis-Tsakonas S., Sowa ME., De Camilli P., Paulo JA., Harper JW., Gygi SP. 2015. The BioPlex Network: A Systematic Exploration of the Human Interactome. Cell **162**(2):425-440.

Huttlin EL., Bruckner RJ., Paulo JA., Cannon JR., Ting L., Baltier K., Colby G., Gebreab F., Gygi MP., Parzen H., Szpyt J., Tam S., Zarraga G., Pontano-Vaites L., Swarup S., White AE., Schweppe DK., Rad R., Erickson BK., Obar RA., Guruharsha KG., Li K., Artavanis-Tsakonas S., Gygi SP., Harper JW. 2017. Architecture of the human interactome defines protein communities and disease networks. Nature **545**(7655):505-509

Iftner T., Elbel M., Schopp B., Hiller T., Loizou J. I., Caldecott K. W., Stubenrauch F. 2002. Interference of papillomavirus E6 protein with single-strand break repair by interaction with XRCC1. The EMBO Journal **21**(17): 4741-4748.

Jackson R., Togtema M., Lambert P. F., Zehbe I. 2014. Tumourigenesis Driven by the Human Papillomavirus Type 16 Asian-American E6 Variant in the Three-Dimensional Keratinocyte Model. PLOSone **9**(7): e101540.

Jackson R., Rosa B. A., Lameiras S., Cuninghame S., Bernard J., Floriano W. B., Lambert P. F., Nicolas A., Zehbe I. 2016. Functional variants of human papillomavirus type 16 demonstrate host genome integration and transcriptional alterations corresponding to their unique cancer epidemiology. BMC Genomics **17**(851): 1-16.

Jeong K. W., Kim H-Z., Kim S., Kim Y. S., Choe J. 2007. Human papillomavirus type 16 E6 protein interacts with cystic fibrosis transmembrane regulator associated ligand and promotes E6-associated protein-mediated ubiquitination and proteasomal degradation. Oncogene **26**: 487-499.

Jing M., Bohi J., Brimer N., Kinter M., Vande Pol S. B. 2007. Degradation of tyrosine phosphatase PTPN3 (PTPH1) by association with oncogenic human papillomavirus E6 proteins. Journal of Virology **81**(5): 2231-2239.

Jha S., Vande Pol S., Banerjee N. S., Duta a. B., Chow L. T., Dutta A. 2010. Destabilization of TIP60 by Human Papillomavirus E6 Results in Attenuation of TIP60-Dependent Transcriptional Regulation and Apoptotic Pathway. Molecular Cell **38**(5): 700-711.

Kalkat M., Resetca D., Lourenco C., Chan PK., Wei Y., Shiah YJ., Vitkin N., Tong Y., Sunnerhagen M., Done SJ., Boutros PC., Raught B., Penn LZ. 2018. MYC Protein Interactome Profiling Reveals Functionally Distinct Regions that Cooperate to Drive Tumorigenesis. Molecular Cell. **72**(5):836-848.e7.

Kanehisa M., Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Research **28**(1): 27-30.

Katzenellenbogen R. A., Egelkrout E. M., Vilet-Gregg P., Gewin L. C., Gafken P. R., Galloway D. A. 2007. NFX1-123 and poly(A) binding proteins synergistically augment activation of telomerase in human papillomavirus type 16 E6-expressing cells. Journal of Virology **81**(8): 3786-3796.

Keppler BR., Archer TK. 2010. Ubiquitin-dependent and ubiquitin-independent control of subunit stoichiometry in the SWI/SNF complex. Journal of Biological Chemistry **285**(46):35665-35674.

King M. C., Raposo G., Lemmon M. A. 2004. Inhibition of nuclear import and cell-cycle progression by mutated forms of the dynamin-like GTPase MxB. PNAS **101**(24): 8957-8962.

Kiyono T., Hiraiwa A., Fujita M., Hayashi Y., Akiyama T., Ishibashi M. 1997. Binding of high-risk human papillomavirus E6 oncoproteins to the human homologue of the Drosophila discs large tumor suppressor protein. PNAS **94**(21): 11612-11616

Klingelhutz A. J., Foster S. A., McDougall J. K. 1996. Telomerase activation by the E6 gene product of human papillomavirus type 16. Nature **380**: 79-32.

Kreimer A. R., Herrero R., Sampson J. N., Porras C., Lowy D. R., Schiller J. T., Schiffman M., Rodriguez A. C., Chanock S., Jimenez S., Schussler J., Gail M. H., Safaeian M., Kemp T. J., Cortes B., Pinto L. A., Hildesheim A., Gonzalez P., for the Costa Rica HPV Vaccine Trial (CVT) Group. 2018. Evidence for single-dose protection by the bivalent HPV vaccine-Review of the Costa Rica HPV vaccine trial and future research studies. Vaccine **18**: 30018-5.

Kukimoto I., Aihara S., Yoshiike K., Kanda T. 1998. Human papillomavirus oncoprotein E6 binds to the C-terminal region of human minichromosome maintenance 7 protein. Biochemical and Biophysical Research Communications **249**(1): 258-262.

Kumar A., Zhao Y., Meng G., Zeng M., Srinivasan S., Delmolino L. M., Gao Q., Dimri G., Weber G. F., Wazer D. E., Band H., Band V. 2002. Human papillomavirus oncoprotein E6 inactivates the transcriptional coactivator human ADA3. Molecular and Cellular Biology **22**(16): 5801-5812

Kumar S., Stecher G., Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution **33**(7): 1870-1874.

Kurebayashi H., Goi T., Shimada M., Tgai N., Naruse T., Nakazawa T., Kimura Y., Hirono Y., Yamaguchi A. 2015. Prokineticin 2 (PROK2) is an important factor for angiogenesis in colorectal cancer. Oncotarget **6**(28): 26242-26251.

Kuroda TS., Maita H., Tabata T., Taira T., Kitaura H., Ariga H., Iguchi-Ariga SM. 2004. A novel nucleolar protein, PAPA-1, induces growth arrest as a result of cell cycle arrest at the G1 phase. Gene **340**(1):83-98.

Lace M. J., Anson J. R., Thomas S. G., Turek L. P., Haugen T. H. 2008. The E8^E2 gene product of human papillomavirus type 16 represses early transcription and replication but is dispensable for viral plasmid persistence in keratinocytes. Journal of Virology **82**(21): 10841-10853.

Larrieu D., Brunet M., Vargas C., Hanoun N., Ligat L., Dagnon L., Lulka H., Pommier R. M., Selves J., Jády B. E., Bartholin L., Cordelier P., Dufrense M., Torrisani J. 2020. The E3 ubiquitin ligase TRIP12 participates in cell cycle progression and chromosome stability. Scientific Reports **10**(789): 1-17.

Lauber W. M., Carroll J. A., Dufield D. R., Kiesel J. R., Radabaugh M. R., Malone J. P. 2001. Mass spectrometry compatibility of two-dimensional gel protein stains. Electrophoresis **22**(5): 906-918.

Lauttia S., Sihto H., Kavola H., Koljonen V., Böhling T., Joensuu H. 2014. Prokineticins and Merkel cell polyomavirus infection in Merkel cell carcinoma. British Journal of Cancer **110**(6): 1446-1455.

LeCouter J., Kowalski J., Foster J., Hass P., Zhang Z., Dillard-Telm L., Frantz G., Rangell L., DeGuzman L., Keller GA., Peale F., Gurney A., Hillan KJ., Ferrara N. 2001. Identification of an angiogenic mitogen selective for endocrine gland endothelium. Nature **412**(6850):877-84.

LeCouter J., Lin R., Tejada M., Frantz G., Franklin Peale F., Hillan KJ., Ferrara N. 2003. The endocrine-gland-derived VEGF homologue Bv8 promotes angiogenesis in the testis: Localization of Bv8 receptors to endothelial cells. PNAS **100**(5): 2685–2690.

Lee K., Lee AY., Kwon YK., Kwon H. 2011. Suppression of HPV E6 and E7 expression by BAF53 depletion in cervical cancer cells. Biochemical and Biophysical Research Communications **412**(2):328-33.

Lee H., Lee S., Hur S., Seo J., Kwon J. 2014. Stabilization and targeting of INO80 to replication forks by BAP1 during normal DNA synthesis. Nature Communications **5**(5128): 1-14.

Liu C., Lin J., Li L., Zhang Y., Chen W., Cao Z., Zuo H., Chen C., Kee K. 2015. HPV16 early gene E5 specifically reduces miRNA-196a in cervical cancer. Scientific Reports **5**: 21-25.

Liu X., Yuan H., Fu B., Disbrow G. L., Apolinario T., Tomaic V., Kelley M. L., Bakers C. C., Huibregtse J., Schlegel R. 2005. The E6AP ubiquitin ligase is required for transactivation of the hTERT promoter by the Human Papillomavirus E6 Oncoprotein. Journal of Biological Chemistry **280**(11): 10807-10816.

Lowry O. H., Rosebrough N. J., Farr A. L., Randall R. J. 1951. Protein measurement with the folin phenol reagent. Journal of Biological Chemistry **193**: 265-275.

Lu Z., Hu X., Li Y., Zheng L., Zhou Y., Jiang H., Ning T., Basang Z., Zhang C., Ke Y. 2004. Human papillomavirus 16 E6 oncoprotein interferences with insulin signaling pathway by binding to tuberin. Journal of Biological Chemistry **279**: 35664-35670.

Maddika S., Kavela S., Rani N., Palicharla V. R., Pokorny J. L., Sarkaria J. N. Chen J. 2011. WWP2 is an E3 ubiquitin ligase for PTEN. Nature Cell Biology **13**(2011): 728-733.

Malik S., Saito H., Takaoka M., Miki Y., Nakanishi A. 2016. BRCA2 mediates centrosome cohesion via an interaction with cytoplasmic dynein. Cell Cycle **15**(16):2145-2156.

Martens JA., Winston F. 2003. Recent advances in understanding chromatin remodeling by Swi/Snf complexes. Current Opinions in Genetics & Development **13**(2):136-42.

Martin R. W., Orelli B. J., Yamazoe M., Minn A. J., Takeda S., Bishop D. K. 2007. RAD51 up-regulation bypasses BRCA1 function and is a common feature of BRCA1-deficient breast tumors. Cancer research **67**(20): 9658-9665.

Matthews K., Leong C. M., Inglis E., Yun K., Bäckstrom B. T., Doorbar J., 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.

Maziarz E. P., Baker G. A., Mure J. V., Woodab T. D. 2000. A comparison of electrospray versus nanoelectrospray ionization Fourier transform mass spectrometry for the analysis of synthetic poly(dimethylsiloxane)/poly(ethylene glycol) oligomer blends. International Journal of Mass Spectrometry **202**(2000): 241-250.

Mazzei F., Viel A., Bignami M. 2013. Role of MUTYH in human cancer. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis **743-744**(2013): 33-43.

McBride A. A., Romanczuk H., Howley P. M. 1991. The Papillomavirus E2 regulatory proteins Journal of Biological Chemistry. **266**(28): 18411-18414.

Mellacheruvu D., Wright Z., Couzens A. L., Lambert J., St-Denis N. A., Li T., Miteva Y. V., Hauri S., Sardiu M. E., Low T. Y., Halim V. A., Bagshaw r. D., Hubner N. C., al-Hakim A., Bouchard A., Faubert D., Fermin D., Dunham W. H., Goudreault M., Lin Z., Badillo B., Pawson T., Durocher D., Coulombe B., Abersold R., Superti-Furga G., Colinge J., Heck A. J. R., Choi H., Gstaiger M., Mohammed S., Cristea I. M., Bennett K. L., Washburn M. P., Raught B., Ewing R. M., Gingras A., Nesvizhskii A. I. 2013. The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nature Methods **10**(8): 730-736.

Mi H., Muruganujan A., Thomas P. D. 2013. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Research. **41**: D377-D386.

Michalski A., Damoc E., Hauschild J. P., Lange O., Wieghaus A., Makarov A., Nagaraj N., Cox J., Mann M., Horning S. 2011. Mass spectrometry-based proteomics using Q exactive, a high-performance benchtop quadruple orbitrap mass spectrometer. Molecular & Cellular Proteomics **10**(9): M111.011015.

Min J., Tian Y., Xiao Y., Wu L., Li L., Chang S. 2013. The mINO80 chromatin remodeling complex is required for efficient telomere replication and maintenance of genome stability. Cell Research **23**(12): 1396-1413.

Muskal S. M., Holbrook S. R., Kim S. 1990. Prediction of disulfide-bonding state of cysteine in proteins. Protein Engineering **3**(8): 667-672.

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.

Neveu G., Cassonnet P., Vidalain P. O., Rolloy C., Mendoza J., Jones L., Tangy F., Muller M., Demeret C., Tafforeau L., Lotteau V., Rabourdin-Combe C., Travé G., Dricot A., Hill D. E., Vidal M., Favre M., Jacob Y. 2012. Comparative analysis of virus-host interactomes with a mammalian high-throughput protein complementation assay based on *Gaussio princeps* luciferase. Methods **58**(4): 349-359.

Niccoli S., Abraham S., Richard C., Zehbe I. 2012. The Asian-American E6 Variant Protein of Human Papillomavirus 16 Alone is Sufficient to Promote Immortalization, Transformation, and Migration of Primary Human Foreskin Keratinocytes. Journal of Virology **86**(22): 12384-12396.

Nikfarjam L., Farzaneh P. 2012. Prevention and detection of mycoplasma contamination in cell culture. Cell Journal **13**(4): 203-212.

Nuttall R., Bryan S., Dale D., De P., Demers A., Ellison L., Rahal R., Shaw A., Smith L., Weir H. K., Woods R., Zakaria D. 2016. Canadian Cancer Society Statistics 2016 Canadian Cancer Society. **2016**.

Ngo A. N., Ezoulin M. J. M., Youm I., Youan B. C. 2015 Optimal Concentration of 2,2,2,-trichloroacetic acid protein precipitation based on response surface methodology. Journal of Analytical & Bioanalytical Techniques **5**(4): 198 – 211.

Nguyen M. L., Nguyen M. M., Lee D., Greip A. E., Lambert P. F. 2003. The PDZ Ligand Domain of the Human Papillomavirus Type 16 E6 Protein Is Required for E6' s Induction of Epithelial Hyperplasia In Vivo. Journal of Virology **77**(12): 6957-6964.

OliveiraL. B., Haga I. R., Villa L. L. 2018. Human papillomavirus (HPV)16 E6 oncoprotein targets the Toll-like receptor pathway. Journal of General Virology **99**(2018): 667-675.

Omenn G. S., States D. J., Adamski M., Blackwell T. W., Menon R., Hermjakob H., Apweiler R., Haab B. B., Simpson R. J., Eddes J. S., Kapp E. A., Mortiz R. L., Chan D. W., Rai A. J., Admon A., Aebersold R., Eng J., Hancock W. S., Hefta S. A., Meyer H., Paik Y., Yoo J., Ping P., Pounds J., Adkins J., Qian X., Wang R., Wasinger V., Wu C. Y., Zhao X., Zeng R., Archakov A., Tsugita A., Beer I., Pandey A., Pisano M., Andrews P., Tammen H., Speigher D. W., Hanash S. M. 2005. Overview of the HUPO plasma proteome project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. Proteomics **5**(2005): 3226-3245.

Ortiz-Ortiz J., Alarcón-Romero L. C., Jiménez-López M. A., Garzón-Barrientos V. H., Calleja-Macías I., Barrera-Saldaña H. A., Leyva-Vázquez M. A., Illades-Aguiar B. 2015. Association of human papillomavirus 16 E6 variants with cervical carcinoma and precursor lesions in women from Southern Mexico. Virology Journal **12**(29).

Oughtred R., Stark C., Breikreutz B. J., Rust J., BoucherL., Chang C., Kolas N., O'Donnell L., Leung G., McAdam R., Zhang F., Dolma S., Willems A., Coulombe-Huntington J., Chatr-Aryamontri A., Dolinski K., Tyers M. 2019. The BioGrid interaction database: 2019 update. Nucleic Acids Research **47**(D1): D529-D541.

Pagala V. R., High A. A., Wang X., Tan H., Kodali K., Mishra A., Kavdia K., Xu Y., Qu Z., Peng J. 2015. Quantitative protein analysis by mass spectrometry In: Meyerkord C. and Fu H. (editors), Protein-protein interactions: methods and applications. Methods in Molecular Biology **1278**: 281-305.

Park J. M., Greten F. R., Wong A., Westrick R. J., Arthur J. S. C., Otsu K., Hoffmann A., Montminy M., Kerin M. 2005. Signaling pathways and genes that inhibit pathogen-induced macrophage apoptosis—CREB and NF-kappaB as key regulators. Immunity **23**(3): 319-329.

Parker A., Gu Y., Mahoney W., Lee S. H., Singh K. K., Lu A. L. 2001. Human homolog of the MutY repair protein (hMYH) physically interacts with proteins involved in long patch DNA base excision repair. The Journal of Biological Chemistry **276**(8): 5547–5555.

Paridaen J. T. M. L., Janson E., Utami K. H., Pereboom T. C., Essers P. B., van Rooijen C., Zivkovic D., MacInnes A. W. 2011. The nucleolar GTP-binding proteins Gnl2 and nucleostemin are required for retinal neurogenesis in developing zebrafish. Developmental Biology **355**(2011): 286-301.

Pastrana D. V., Peretti A., Welch N. L., Borgogna C., Olivero C., Badolato R., Notarangelo L. D., Gariglio M., FitzGerald P. C., McIntosh C. E., Reeves J., Starrett G. J., Bliskovsky V., Velez D., Brownell I., Yarchoan R., Wyvill K. M., Uldrick T. S., Maldarelli F., Lisxo A., Sereti I., Gonzalez C. M., Androphy E. J., McBride A. A., Van Doorslaer K., Garcia F., Dvoretzky I., Liu J. S., Han J., Murphy P. M., McDermott D. H., Buck C. B. 2019. Metagenomic discovery of 83 new human papillomavirus types in patients with immunodeficiency. mSphere **3**(6): e00645-18.

Patel D., Huang S-M., Baglia L. A., McCance D. J. 1999. The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. The EMBO Journal. **18**(18): 5061-5072.

Pett M. R., Herdman M. T., Palmer R. D., Yeo G. S. H., Shivji M. K., Stanley M. A., Coleman N. 2006. Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response. PNAS **103**(10): 3822-3827.

Pierrat B., Correia J. S., Mary J. L., Tomás-Zuber M., Lesslauer W. 1998. RSK-B, a novel ribosomal S6 kinase family member, is a CREB kinase under dominant control of p38α mitogen-activated protein kinase ($p38\alpha^{MAPK}$). Journal of Biological Chemistry **273**(45): 29661-29671.

Popko-Scibor A., Lindberg M. J., Hansson M. L., Holmlund T., Wallberg A. E. 2011. Ubiquitination of Notch1 is tregulated by MAML1-mediated p300 acetylation of Notch1. Biochemical and Biophysical Research Communications **416**(3-4): 300-306.

Puig O., Caspary F., Rigaut G., Rutz B., Bouveret E., Bragado-Nilsson E., Wilm M., Séraphin B. 2001. The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods **24**(3): 218-229.

R Frank S., Parisi T., Taubert S., Fernandez P., Fuchs M., Chan H-M., Livingston D.M., Amati B. 2003. MYC recruits the TIP60 histone acetyltransferase complex to chromatin. EMBO Rep. **4**(6): 575–580.

Rabilloud T., Lelong C. 2011. Two-dimensional gel electrophoresis in proteomics: A tutorial. Journal of Proteomics **74**(10): 1829-1841.

Racevskis J., Dill A., Stockert R., Fineberg S. A. 1996. Cloning of a novel nucleolar guanosine 5'-triphosphate binding proteins autoantigen from a breast tumor. Cell Growth and Differentiation **7**(2): 271-280.

Rada M., Vasileva E., Lezina L., Marouco D., Antonov AV., Macip S., Melino G., Barlev NA. 2017. Human EHMT2/G9a activates P53 through methylation-independent mechanism. Oncogene **36**(7):922-932.

Rao V. S., Srinivas K., Sujini G. N., Sunand Kumar G. N. 2014. Protein-protein interaction detection: methods and analysis. International Journal of Proteomics **2014**: 1-12.

Richard C., Lanner C., Naryzhny S. N., Sherman L., Lee H., Lambert P. F., Zehbe I. 2010. The immortalizing and transforming ability of two common human papillomavirus 16 E6 variants with different prevalences in cervical cancer. Oncogene **29**(2010): 3435-3445.

Rogers J. C., Bomgarden R. D. 2016. Sample preparation for mass spectrometry-based proteomics; from proteomes to peptides. Advances in Experimental Medicine and Biology **919**: 43-62.

Rolland T., Taşan M., Charloteaux ., Pevzner SJ., Zhong Q., Sahni N., Yi S., Lemmens I., Fontanillo C., Mosca R., Kamburov A., Ghiassian SD., Yang X., Ghamsari L., Balcha D., Begg BE., Braun P., Brehme M., Broly MP., Carvunis AR., Convery-Zupan D., Corominas R., Coulombe-Huntington J., Dann E., Dreze M., Dricot A., Fan C., Franzosa E., Gebreab F.., Gutierrez BJ., Hardy MF., Jin M., Kang S., Kiros R., Lin GN., Luck K., MacWilliams A., Menche J., Murray RR., Palagi A., Poulin MM., Rambout X., Rasla J., Reichert P., Romero V., Ruyssinck E., Sahalie JM., Scholz A., Shah AA., Sharma A., Shen Y., Spirohn K., Tam S., Tejeda AO., Trigg SA., Twizere JC., Vega K., Walsh J., Cusick ME., Xia Y., Barabási AL., Iakoucheva LM., Aloy P., De Las Rivas J., Tavernier J., Calderwood MA., Hill DE., Hao T., Roth FP., Vidal M. 2014. A proteome-scale map of the human interactome network. Cell **159**(5):1212-1226.

Ronco L. V., Karpova A. Y., Vidal M., Howley P. M. 1998. Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity. Genes and Development **12**(13): 2061-2072.

Rozenblatt-Rosen O., Deo RC., Padi M., Adelmant G., Calderwood MA., Rolland T., Grace M., Dricot A., Askenazi M., Tavares M., Pevzner SJ., Abderazzaq F., Byrdsong D., Carvunis AR., Chen AA., Cheng J., Correll M., Duarte M., Fan C., Feltkamp MC., Ficarro SB., Franchi R., Garg BK., Gulbahce N., Hao T., Holthaus AM., James R., Korkhin A., Litovchick L., Mar JC., Pak TR., Rabello S., Rubio R., Shen Y., Singh S., Spangle JM., Tasan M., Wanamaker S., Webber JT., Roecklein-Canfield J., Johannsen E., Barabási AL., Beroukhim R., Kieff E., Cusick ME., Hill DE., Münger K., Marto JA., Quackenbush J., Roth FP., DeCaprio JA., Vidal M. 2012. Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. Nature **487**(7408):491-495.

Roy R., Chun J., Powell S. N. 2012. BRCA1 and BRCA2: Different roles in a common pathway of genome protection. Nature Reviews Cancer **12**: 68-78.

Saccani S., Pantano S., Natoli G. 2001. p38-dependent marking of inflammatory genes for increased NF-kB recruitment. Nature Immunology **3**: 69-75.

Salmi J., Nyman T. A., Nevalainen O. S., Aittokallio T. 2009. Filtering strategies for improving protein identification in high-throughput MS/MS studies. Proteomics **9**(4): 848-860.

Saraswathy N., Ramalingam P. 2011 Protein identification by Peptide Mass Fingerprinting (PMF). Concepts and Techniques in Genomics and Proteomics Chapter 13: 185-192.

Saraiya M., Unger E. R., Thompson T. D., Lynch C. F., Hernandez B. Y., Lyu C. W., Steinae M., Watson M., Wilkinson E. J., Hopenhayn C., Copeland G., Cozen W., Peters E. S., Huang Y., Saber M. S., Altekruse S., Goodman M. T. 2016. US assessment of HPV types in cancers: Implications for current and 9-valent HPV Vaccines. Journal of the National Cancer Institute **107**(6): djv086.

Scheffner M., Werness B. A., Huibergtse 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.

Seeber A., Hauer M., Gasser S. M. 2013. Nucleosome remodelers in double-strand break repair. Current Opinion in Genetics & Development 23(2): 174-184.

Seedorf K., Krammer G., Durst M., Suhai S., Rowekamp W. G. 1985. Human papillomavirus type 16 DNA sequence. Virology **145**(1): 181-185.

Sichero L., Ferreira S., Trottier H., Duarte-Franco E., Ferenczy A., Franco E. L., Villa L. L. 2007. High grade cervical lesions are caused preferentially by non-European variants of HPVs 16 and 18. International Journal of Cancer **120**(2007): 1763 – 1768.

Smith B., Chen Z., Reimers L., van Doorslaer K., Schiffman M., DeSalle R., Herrero R., Yu K., Wacholder S., Wang T., Burk R. D. 2011. Sequence Imputation of HPV16 Genomes for Genetic Association Studies. PLoS ONE **6**(6): e21375.

Smith L., Woods R., Brenner D., Bryan S., Louzado C., Shaw A., Turner D., Weir H. K. 2019**.** Canadian Cancer Society Statistics 2019 Canadian Cancer Society. **2019**.

Smith P. K., Krohn R. I., Hermanson G. T., Mallia A. K., Gartner F. H., Provenzano M. D., Fujimoto E. K., Goeke N. M., Olson B. J., Klenk D. C. 1985. Measurement of protein using bicinchoninic acid. Analytical Biochemistry **150**: 76-85.

Snider J., Kotlyar M., Saraon P., Yao Z., Jurisica I., Stagljar I. 2015. Fundamentals of protein interaction network mapping. Molecular Systems Biology **2015**(11): 848 – 868.

Soloaga A., Thomson S., Wiggin G. R., Rampersaud N., Dyson M. H., Hazzalin C. A., Mahadevan L. C., Arthur J. S. C. 2003. MSK2 and MSK1 mediate the mitogen- and stress-induced phosphorylation of histone H3 and HMG14. The EMBO Journal **22**(11): 2788-2797.

Srivenugopal K. S., Ali-Osman F. 2002. The DNA repair protein O^6-methylguanine-DNA methyltransferase is a proteolytic target for the E6 human papillomavirus oncoprotein. Oncogene **21**: 5940-5945.

Stanley M. 2010 Pathology and epidemiology of HPV infection in females. Gynecologic Oncology **117**(2): S5-S10.

Steinberg T. H. 2009. Chapter 31 protein gel staining methods: An introduction and overview. Methods in Enzymology **463**(2009): 541-563.

Storrs C. H., Silverstein S. J. 2007. PATJ, a tight junction-associated PDZ protein, is a novel degradation target of high-risk human papillomavirus E6 and the alternatively spliced isoform 18E6*∇. Journal of Virology **81**(8): 4080-4090.

Strack B., Calistri A., Craig S., Popova E., Göttlinger HG. 2003. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. Cell **114**(6):689-99.

Strickland S. W., Brimer N., Lyons C., Vande Pol S. B. 2018. Human papillomavirus E6 interaction with cellular PDZ domain proteins modulates YAP nuclear localization. Virology **516**(2018): 127-138.

Szklarczyk D., Morris J. H., Cook H., Kuhn M., Wyder S., Simonovic M., Santos A., Doncheva N. T., Roth A., Bork P., Jensen L. J., von Mering C. 2017. The string database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic acids Research **45**: D362-D368.

Tabrizi S. N., Brotherton J. M. L., Kaldor J. M., Skinner S. R., Cummins E., Liu B., Bateson D., McNamee K., Garefalakis M., Garland S. M. 2012. Fall in Human Papillomavirus prevalence following a national vaccination program. The Journal of Infectious Diseases **206**(11): 1645-1651.

The UniProt Consortium. 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Research **46**(D1): D158-D169.

Thomas M., Banks L. 1999. Human papillomavirus (HPV) E6 interactions with Bak are conserved amongst E6 proteins form high and low risk HPV types. Journal of General Virology **80**: 1513-1517.

Thomas M., Laura R., Hepner K., Guccione E., Sawyers C., Lasky L., Banks L. 2002. Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation. Oncogene **21**: 5088-5096.

Thomas M., Banks L. 2014. PDZRN3/LNX3 is a novel target of human papillomavirus type 16 (HPV-16) and HPV-18 E6. Journal of Virology **89**(2): 1439-1444.

Togtema M., Jackson R., Richard C., Niccoli S., Zehbe I. 2015. The human papillomavirus 16 European-T350G E6 variant can immortalize but not transform keratinocytes in the absence of E7. Virology **485**: 274-282.

Tomaić V., Pim D., Banks L. 2009. The stability of the human papillomavirus E6 oncoprotein is E6AP dependent.Virology **393**(2009): 7-10.

Tong X., Howley P. M. 1997. The bovine papillomavirus E6 oncoprotein interacts with paxillin and disrupts the actin cytoskeleton. PNAS **94**(9): 4412-4417.

United States Cancer Statistics - Incidence: 1999 - 2016, WONDER Online Database. United States Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2019. Accessed at http://wonder.cdc.gov/cancer-v2016.html on Feb 24, 2020.

Van Doorslaer K., Li Z., Xirasagar S., Maes P., Kaminsky D., Liou D., Sun Q., Kaur R., Huyen Y., McBride A. 2016. The Papillomavirus Episteme: a major update to the papillomavirus sequence database Nucleic Acids Research **45**(D1): D499-D506.
Database URL: https://pave.niaid.nih.gov/

Vande Pol S. B., Klingelhutz A. J. 2013. Papillomavirus E6 oncoproteins. Virology **445**(0): 115-137.

Vandermark E. R., Deluca K. A., Gardner C. R., Marker D. F., Schreiner C. N., Strickland D. A., Wilton K. M., Mondal S., Woodworth C. D. 2012. Human papillomavirus type 16 E6 and E7 proteins alter NF-kB in cultured cervical epithelial cells and inhibition of NF-kB promotes cell growth and immortalization. Virology **425**(1): 53-60.

Veldman T., Horikawa I., Barrett J. C., Schlegel R. 2001. Transcriptional activation of the telomerase hTERT gene by human papillomavirus Type 16 E6 oncoprotein. Journal of Virology **75**(9): 4467-4472.

Venuti A., Paolini F., Nasir L., Corteggio A., Roperto S., Campo M. S., Borzacchiello G. 2011. Papillomavirus E5: the smallest oncoprotein with many functions. Molecular Cancer **10**: 140.

Verheyen E. M., Gottardi C. J. 2010. Regulation of Wnt/bets-catenin signaling by protein kinases. Developmental Dynamics **239**(1): 34-44.

Villa L. L., Sichero L., Rahal P., Caballero O., Ferenczy A., Rohan T., Franco E. L. 2000. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. Journal of General Virology **81**: 2959–2968.

Vos R. M., Altreuter J., White E. A., Howley P. M. 2009. The Ubiquitin-specific peptidase USP15 regulates human papillomavirus type 16 E6 protein stability§∇. Journal of Virology **83**(17): 8885-8892.

Wakeham K., Kavanaugh K. 2014. The burden of HPV-associated anogenital cancers. Current Oncology Reports **16**(9): 402-413.

Walboomers J. M., Jacobs M. V., Manos M. M., Bosch F. X., Kummer J. A., Shah K. V., Snijders P. J., Peto J., Meijer C. J., Muñoz N. 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide The Journal of Pathology **189**(1): 12 – 19.

Werness B. A., Levine A. J., Howley P. M. 1990. Association of human papillomavirus types 16 and 18 E6 proteins with P53. Science **248**(4951): 76-79.

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): 13174-13186.

White E. A., Sowa M. E., Tan M. J. A., Jeudy S., Hayes S. D., Santha S., Münger K., Harper J. W., Howley P. M. 2012 Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses. PNAS **109**(5): E260-E267.

Yaciuk P. 2007. Co-Immunoprecipitation of protein complexes. In: Wold W.S.M., Tollefson A. E. (eds) Adenovirus Methods and Protocols. Methods in Molecular Medicine **131**: 103-111.

Yates J. R. 2000. Mass spectrometry: from genomics to proteomics. Trends in Genetics **16**(1): 5-8.

Yim E. K., Lee K. H., Myeong J., Tong S. Y., Um S. J., Park J. S. 2007. Novel interaction between HPV E6 and BARD1 (BRCA1-associated ring domain 1) and its biologic roles. DNA and Cell Biology **26**(10): 753-761.

Yoder K.E., Espeseth A., Wang X. H., Fang Q., Russo M.T., Lloyd R. S., Hazuda D., Sobol R.W., Fishel R. 2011. The base excision repair pathway is required for efficient lentivirus integration. PLoS ONE **6**(3): e17862.

Zacapala-Gómez A. E., Moral-Hernández O. D., Villegas-Sepúlveda N., Hidalgo-Miranda A., Romero-Córdoba S. L., Beltrán-Anaya F. O., Leyva-Vázquez M. A., Alarcón-Romero L. D. C., Illades-Aguiar B. 2016. Changes in global gene expression profiles induced by HPV 16 E6 oncoprotein variants in cervical carcinoma C33-A cells. Virology **488**: 187-195.

Zanier K., Charbonnier S., Sidi A. O., McEwen A. G., Ferrario M. G., Poussin-Courmontagne P., Cura V., Brimer N., Babah K. O., Ansari T. 2013. Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. Science **339**: 694-698.

Zehbe I., Wilander E., Delius H., Tommasino M. 1998. Human Papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. Cancer Research **58**: 829-833.

Zehbe, I., Richard C., DeCarlo C. A., Shai A., Lambert P. F., Lichtig H., Tommasino M., Sherman L. 2009. Human papillomavirus 16 E6 variants differ in their dysregulation of human keratinocyte differentiation and apoptosis. Virology **383**(1): 69-77.

Zhang Y., Fan S., Meng Q., Ma Y., Katiyar P., Schregel R., Rosen E. M. 2005. BRCA1 Interaction with Human Papillomavirus Oncoproteins. Journal of Biological Chemistry **280**(39): 33165-33177.

Zheng L., Ding H., Lu Z., Li Y., Pan Y., Ning T., Ke Y. 2008. E3 ubiquitin ligase E6AP-mediated TSC2 turnover in the presence and absence of HPV16 E6. Genes to Cells **13**: 285-294.

Zou X., Levy-Cohen G., Blank M. 2015. Molecular functions of NEDD3 E3 ubiquitin ligases in cancer. Biochimica et Biophysica Acta (BBA) Reviews on Cancer **1856**(1): 91-106.

Zuna R. E., Moore W. E., Shanesmith R. P., Dunn S. T., Wang S. S., Schiffman M., Blakey G. L., Teel T. 2009. Association of HPV16 E6 variants with diagnostic severity in cervical samples of 354 women in a US population. International Journal of Cancer **125**: 2609-2613.

zur Hausen H. 1977. Human papillomaviruses and their possible role in squamous cell carcinomas. Current Topics in Microbiology and Immunology **78**: 1-30.

zur Hausen H. 2009. Papillomaviruses in the causation of human cancers – a brief historical account. Virology **384**(2009): 260-265.

6 Supplemental Data

Figures

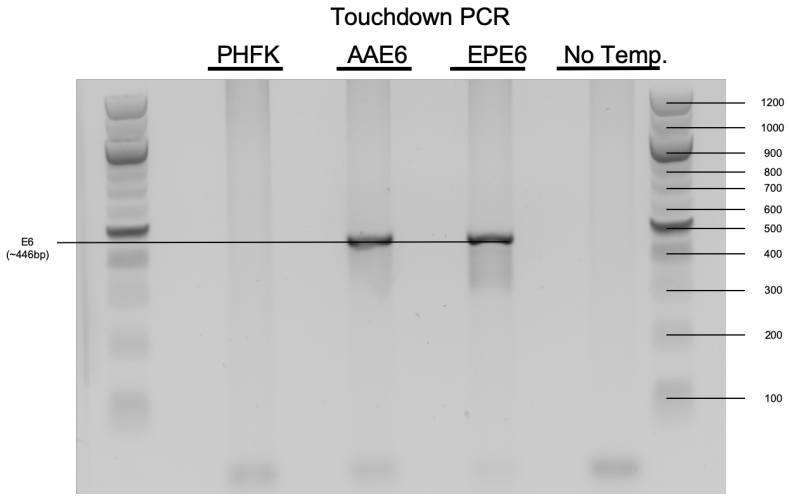

Figure S1 – Pre-PCR purification of late-passage AAE6, EPE6 as well as PHFK-HA and no template by touchdown PCR. Two prominent bands at approximately 446 bp can be seen in only the AAE6 and EPE6 samples indicating a successful amplification of viral DNA for sequencing. Lanes on the far left and right sides are DNA ladders used to identify sample sizes.

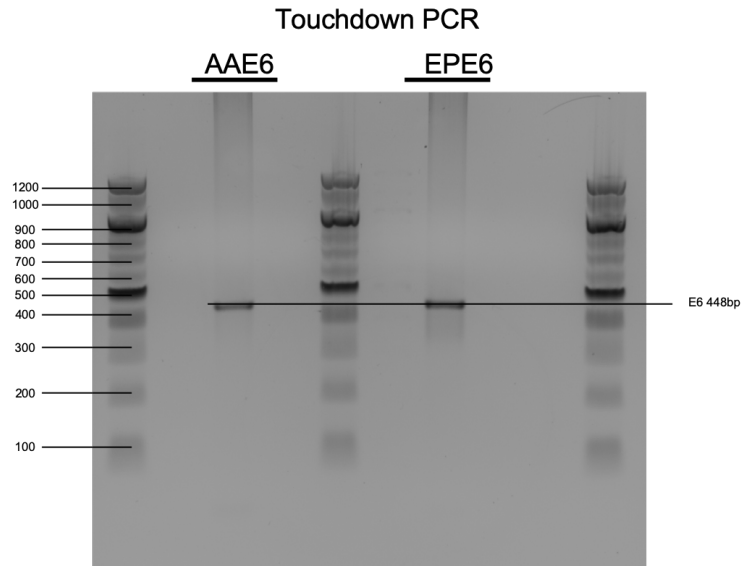

Figure S2 – Post-PCR purification of late-passage AAE6, EPE6 as well as PHFK-HA and no template. AAE6 and EPE6 bands remain and the bands present in pre-purification (*Figure S1*) samples are removed. Lanes on the far left and right sides are DNA ladders used to identify sample sizes.

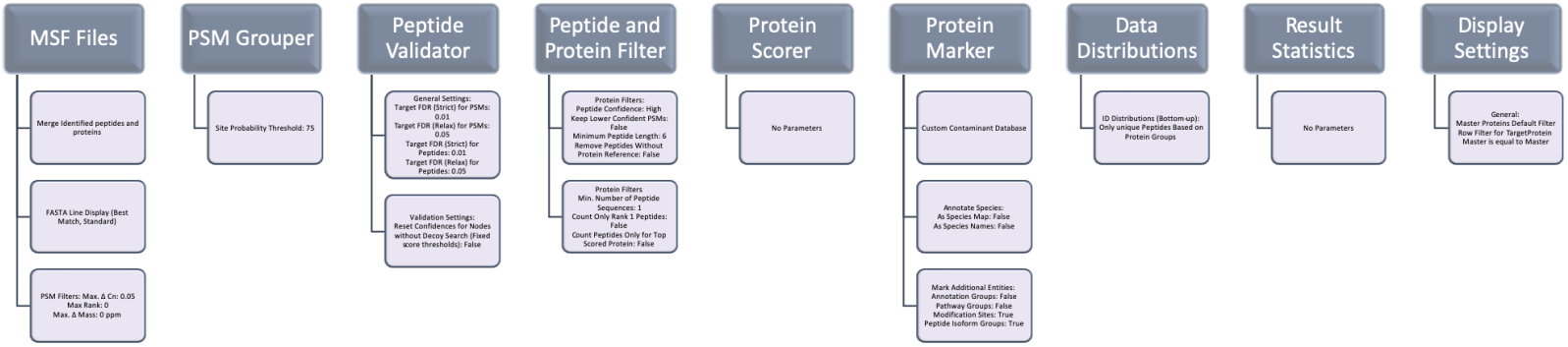

Figure S3 – Initial workflow for processing of mass spectrometry spectra files by HSMPF. All parts of the workflow were completed in Proteome Discoverer Version: 2.4.0.292. Workflow was done prior to analysis of data to ensure all files were in the correct format.

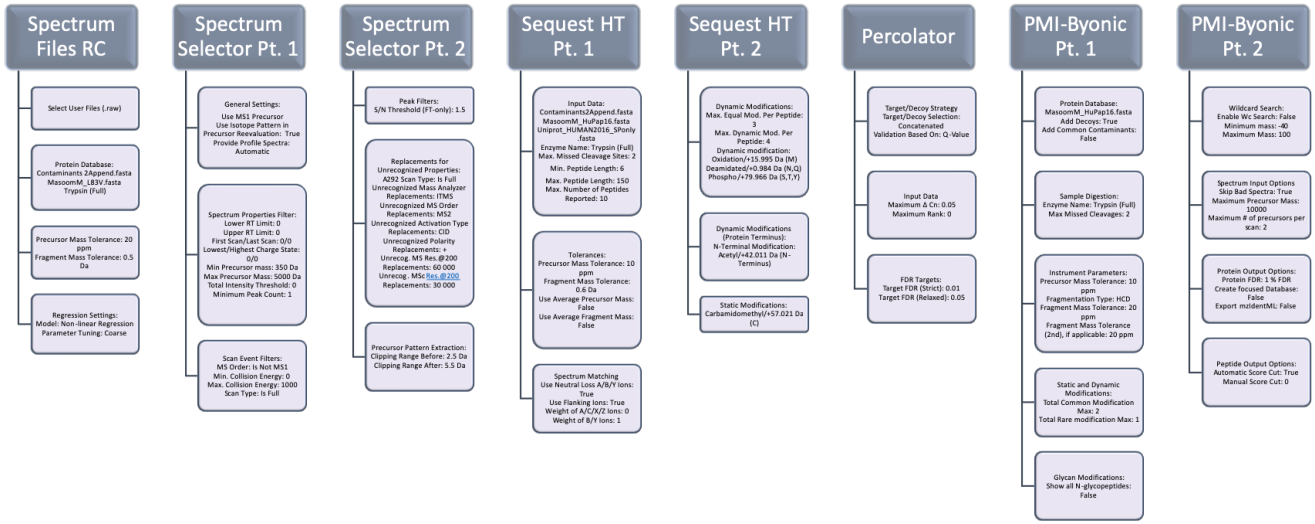

Figure S4 – Final workflow for processing data in Proteome Discoverer Version 2.4.0.292 at HSMPF. Major components of each workflow are present at the top of each column and further detail of each step is described in boxes beneath the headers.

Tables

Table S1 – Results from literature search for articles that performed Co-IP experiments targeting E6 or another HPV protein

Author	Lysis Buffer	Protease Inhibitor	Phosphatase Inhibitor	Proteasome Inhibitor	Lysis Incubation	Amount of cells	Cell Type	Volume Buffer	Filter (Y/N)	IP (Y/N)	Incubation Time (h)	Wash Buffer	Elution Buffer	MS Prep	MS (Y/N)	Sonication (Y/N)	E6 targeted (Y/N)
Subbiah V. K., Massimi P., boon S. S., Myers M. M., Sharek L., Garcis-Mata R., Banks L. 2012. The Invasive Capacity of HPV Transformed Cells Requires the hDlg Dependent Enhancement of SGEF/RhoG Activity. PLoS Pathogens **8**(2): e1002543.	HEPES Buffer (50 mM HEPES [pH 7.0], 500mM NaCl, 0.1% NP-40)	Yes (ND)	Yes (ND)	No	20 min on ice	7X10^5	HEK 293	ND	N	Y	3 @ 4C	3 washes with Lysis buffer	ND	N	N	N	Y (HPV18)
Bentley P., Tan M. J. M., McBride A. A., White E. A., Howley P. M. 2018. The SMC5/6 complex interacts with the papillomavirus E2 protein and influences maintenance of viral episomal DNA. Journal of Virology **92**(15): e00356-18.	Tris-HCl (50mM Tris-HCl [pH 7.5], 150mM NaCl, 0.5% NP-40, 1mM EDTA)	Yes (ND)	No	No	ND	90% confluence in 4 15cm dishes	293T	ND	N	Y	16h @ 4C	ND	250µg/mL HA peptide	N	Y (LC-MS/MS)	Y	N
Meyers J. M., Uberoi A., Grace M., Lambert P. F., Munger K. 2017. Cutaneous HPV8 and MmuPV1 E6 proteins target the NOTCH and TGF-ß tumor suppressors to inhibit differentiation and sustain keratinocyte proliferation. PLoS Pathogens **13**(1): e1006171.	Tris-HCl (50mM Tris-HCl [pH 8.0], 120mM NaCl, 1% NP-40)	No	No	No	ND	ND	U2OS and HCT116	ND	N	Y	ND	ND	ND	N	N	N	Y (HPV -8, -16)
J. Meyers PhD Dissertation 2017	Tris-HCl (50mM Tris-HCl [pH 8.0], 120mM NaCl, 1% NP-40)	No	No	No	ND	ND	Unknown	ND	N	Y	ND	ND	ND	N	N	Y	Y
Grace M. and Munger K. 2017. Proteomic analysis of the gamma human papillomavirus type 197 E6 and E7 associated cellular proteins. Virology **500**(2017): 71-81.	Tris-HCl (50mM Tris-HCl [pH 8], 150mM NaCl, 0.5% NP-40, 0.5mM EDTA)	Yes (Complete EDTA-free protease inhibitor per 50mL)	No	No	ND	4 X 15 cm dishes after 24h growth with 8 million cells seeded	HCT116	4 mL	Y (0.45µm spin filter)	Y	4-16h @ 4C	Several Ice Cold Lysis Buffer, then several washes in PBS (to get rid of detergent)	3 X 50mL of 250µg/mL HA Peptide (Sigma: 12149) in PBS	Combined and precipitated in 20% TCA for 25 min on ice. Centrifuged at 20 000g for 25 min. Pellets washed once with 500µL 10% TCA then 3 1 mL washes of cole acetone to remove HA peptide	Y (Orbitrap LC MS/MS)	N	Y (HPV197 E6 and E7)
Jang M. K., Anderson D. E., van Doorslaer K., McBride A. A. 2015. A proteomic approach to discover and compare interacting partners of papillomavirus E2 proteins from diverse phylogenetic groups. Proteomics **15** (12): 2038-2050.	HEPES Buffer (10mM [pH 7.9], 10mM KCl, 1.5mM MgCl2, 0.2mM EDTA, 0.5mM DTT)	Yes (cOmplete protease inhibitor cocktail, PMSF)	No	No	30 min at 4C, 20 stroke of Dounce Homogenizer	1X10^9	C33A	ND	N	Y	16h @ 4C	Extensive washing Buffer (20mM HEPES [pH 7.9], 20% glycerol, 150mM KCl, 0.2mM EDTA, 0.5mM DTT, 0.1% NP-40, 1mM PMSF)	2 x 0.25mL of (0.1mg/mL) FLAG Peptide	ND	Y (LC-MS/MS) LQTF	homogenizer	N (E2)
Martinez-Nöel G., Galligan J. T., Sowa M. E., Arndt V., Overton T. M., Harper J. W., Howley P. M. 2012. Identification and proteomic analysis of distinct UBE3A/E6AP protein complexes. Molecular and Cellular Biology **32**(15): 3095-3106.	Tris-HCl (50mM Tris-HCl [pH 7.5], 150mM NaCl, 0.5% NP-40)	Yes (cOmplete protease inhibitor cocktail)	Yes (12.5 mM NaF, 1mM Na3VO4, 12.5mM ß-glycerophosphate)	No	ND	80%-90% confluent in 5 X 15 cm dishes	T-Rex	5 mL ice cold lysis buffer	Y (20µm SFCA-PF syringe filter)	Y	2h @ 4C	2 times with 10mL Lysis buffer. 2 times with 10mL PBS	3 X 50µL of 500µg/mL HA Peptide (Sigma) in PBS		Y	N	N (E6-AP complex)
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): 13174-13186.	Tris-HCl (50mM Tris-HCl [pH 7.5], 150mM NaCl, 0.5% NP-40, 1mM EDTA)	Yes (cOmplete protease inhibitor cocktail)	No	No	ND	90% confluent in 4 X 15 cm dishes	293T, N/Tert	ND	Y (0.2µm syringe filter)	Y	16h @ 4C	5 times with lysis buffer. Exchange into PBS	eluted with SIGMA 250µg/mL HA peptide at RT	TCA precipitation, wash with acetone.	Y (LTQ, orbitrap, MS/MS)	Y	Y
White E. A., Sowa M. E., Tan M. J. A., Jeudy S., Hayes S. D., Santha S., Münger K., Harper J. W., Howley P. M. 2012 Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses. PNAS **109**(5): E260-E267.	Tris-HCl (50mM Tris-HCl [pH 7.5], 150mM NaCl, 0.5% NP-40)	Yes (cOmplete protease inhibitor cocktail)	Yes (25mM NaF, 1mM Na3VO4, 5mM ß-glycerophosphate)	30µM MG132 4 hours before harvesting	ND	ND	N/Tert-1	ND	N	Y	ND	3 times PBS	eluted with HA peptide at RT	TCA precipitation, wash with acetone.	Y (ND)	N	Y
Brimer N., Lyons Vande Pol S. B. 2007. Association of E6AP (UBE3A) with human papillomavirus type 11 E6 protein. Virology **358**(2007): 303-310.	0.5X Tris-HCl (50mM Tris-HCl [pH 7.5], 150mM NaCl, 1% NP-40)	0.01% PMSF, 1µg/mL leupeptin/aprotinin	Y (50mM NaF, 5mM NAPPI7, 1mM NaVO3	No	ND	5X10^8 CV1	CV1, HA CAT, NIKS	ND	N	Y	ND	Wash 3 times with Lysis Buffer	3 X 2µg FLAG peptide in 0.25X NP-40 Lysis Buffer	Band excision	Y	N	N (E6-AP)
Jeong K. W., Kim H. Z., Kim S., Choe J. 2007. Human papillomavirus type 16 E6 protein interacts with cystic fibrosis transmembrane regulator-associated ligand and promotes E6-associated protein-mediated ubiquitination and proteasomal degradation. Oncogene **26**(2017): 487-499.	Tris-HCl (50mM Tris-HCl [pH 7.5], 120mM NaCl, 0.5% NP-40)	1mM PMSF, and protease inhibitor	Y (50mM NaF, 200µM Na3VO4	No	ND	ND	293T	ND	N	Y	3h @ 4C	Lysis Buffer 3 times	SDS Sample Buffer	N/A	N	N	Y
Jing M., Bohl J., Brimer N., Kinter M., Vande Pol S. B. 2007. Degradation of tyepsine phosphatase PTPN3 (PTPH1) by association with oncogenic human papillomavirus E6 proteins. Journal of Virology **81**(5): 2231-2239.	0.5X Tris-HCl (50mM Tris-HCl [pH 7.5], 150mM NaCl, 1% NP-40)	Y (0.01% PMSF, 5mM EDTA, 1µg/mL leupeptin/aprotinin)	Y (50mM NaF, 5mM NAPPI7, 1mM NaVO3	No	ND	1X10^7	CV1	ND	N	Y	30 min @ 4C	Wash 3 times with 1X Lysis Buffer, 2 more washes with 0.25X NP-40	3 X 2µg FLAG peptide in 0.25X NP-40 Lysis Buffer	Band Excision	Y	Y (LC MS/MS)	Y

Table S2 – Common identified proteins targeted by AAE6 and EPE6 using the *Peptide Method*. Number in columns for AAE6 and EPE6 trials correspond to the number of peptides identified.

Acc#	Description	CRAPome	AAE6 DMSO T1	AAE6 MG132 T1	AAE6 MG132 T2	AAE6 DMSO T2	EPE6 DMSO T1	EPE6 MG132 T1	EPE6 MG132 T2	EPE6 DMSO T2
Q05086	Ubiquitin-protein ligase E3A OS=Homo sapiens GN=UBE3A PE=1 SV=4	1 / 411	5	4	2	4	5	4	2	2
Q9NUQ6	SPATS2-like protein OS=Homo sapiens GN=SPATS2L PE=1 SV=2	20 / 411			8	4		2	7	1
Q96GQ7	Probable ATP-dependent RNA helicase DDX27 OS=Homo sapiens GN=DDX27 PE=1 SV=2	41 / 411			9	6			7	1
Q9NVU7	Protein SDA1 homolog OS=Homo sapiens GN=SDAD1 PE=1 SV=3	23 / 411	1	1	4	3	1		5	2
Q14244	Ensconsin OS=Homo sapiens GN=MAP7 PE=1 SV=1	42 / 411	2		3	3		2	3	2
Q9H444	Charged multivesicular body protein 4b OS=Homo sapiens GN=CHMP4B PE=1 SV=1	28 / 411	1	1	3	2			5	
Q5T3I0	G patch domain-containing protein 4 OS=Homo sapiens GN=GPATCH4 PE=1 SV=2	37 / 411			3	3			2	3
Q8TDD1	ATP-dependent RNA helicase DDX54 OS=Homo sapiens GN=DDX54 PE=1 SV=2	49 / 411			3				5	
Q13823	Nucleolar GTP-binding protein 2 OS=Homo sapiens GN=GNL2 PE=1 SV=1	45 / 411			11	6		1	2	1
O95478	Ribosome biogenesis protein NSA2 homolog OS=Homo sapiens GN=NSA2 PE=1 SV=1	15 / 411		1	1	2			3	1
Q9UMY1	Nucleolar protein 7 OS=Homo sapiens GN=NOL7 PE=1 SV=2	11 / 411	1		1	1	1	1	1	1
P04637	Cellular tumor antigen p53 OS=Homo sapiens GN=TP53 PE=1 SV=4	52 / 411		1		1		4		
Q14669	E3 ubiquitin-protein ligase TRIP12 OS=Homo sapiens GN=TRIP12 PE=1 SV=1	29 / 411			3	1			3	

Table S3 – Common identified proteins targeted by AAE6 and EPE6 using the *Protein-Pathway Method*. Number in columns for AAE6 and EPE6 trials correspond to the number of peptides identified during mass spectrometry.

Accession	Description	AAE6 DMSO T1	AAE6 MG132 T1	EPE6 DMSO T1	EPE6 MG132 T1	AAE6 MG132 T2	AAE6 DMSO T2	EPE6 MG132 T2	EPE6 DMSO T2
Q9BXL7	Caspase recruitment domain-containing protein 11 OS=Homo sapiens GN=CARD11 PE=1 SV=3	1		1					
P04637	Cellular tumor antigen p53 OS=Homo sapiens GN=TP53 PE=1 SV=4		1		4		1		
Q13352	Centromere protein R OS=Homo sapiens GN=ITGB3BP PE=1 SV=2		1	1	1				
Q9H444	Charged multivesicular body protein 4b OS=Homo sapiens GN=CHMP4B PE=1 SV=1	1	1			3	2	5	
P14854	Cytochrome c oxidase subunit 6B1 OS=Homo sapiens GN=COX6B1 PE=1 SV=2	1			1				
Q9H2F5	Enhancer of polycomb homolog 1 OS=Homo sapiens GN=EPC1 PE=1 SV=1	1	1	1					
P02671	Fibrinogen alpha chain OS=Homo sapiens GN=FGA PE=1 SV=2	1	1	1	1				
P68871	Hemoglobin subunit beta OS=Homo sapiens GN=HBB PE=1 SV=2					1	1	1	1
Q6ZVT0	Inactive polyglycylase TTLL10 OS=Homo sapiens GN=TTLL10 PE=1 SV=2						1	1	
Q9UG01	Intraflagellar transport protein 172 homolog OS=Homo sapiens GN=IFT172 PE=1 SV=2	1		1					
P23511	Nuclear transcription factor Y subunit alpha OS=Homo sapiens GN=NFYA PE=1 SV=2					1	1	1	
Q9BT09	Protein canopy homolog 3 OS=Homo sapiens GN=CNPY3 PE=1 SV=1					2	2	2	1
Q96GY3	Protein lin-37 homolog OS=Homo sapiens GN=LIN37 PE=1 SV=1		1	1	1				
O14641	Segment polarity protein dishevelled homolog DVL-2 OS=Homo sapiens GN=DVL2 PE=1 SV=1						1	1	
Q96B97	SH3 domain-containing kinase-binding protein 1 OS=Homo sapiens GN=SH3KBP1 PE=1 SV=2	1		1					
Q9Y6M0	Testisin OS=Homo sapiens GN=PRSS21 PE=1 SV=1		1		1				
P10599	Thioredoxin OS=Homo sapiens GN=TXN PE=1 SV=3	1	1	1	1				
Q9P031	Thyroid transcription factor 1-associated protein 26 OS=Homo sapiens GN=CCDC59 PE=1 SV=2					1		1	
A6H8Y1	Transcription factor TFIIIB component B" homolog OS=Homo sapiens GN=BDP1 PE=1 SV=3		1	1	1				
Q9Y4R7	Tubulin monoglycylase TTLL3 OS=Homo sapiens GN=TTLL3 PE=1 SV=2		1	1					