Comparative mapping of E2-host interactions unravels new roles of E2 in human papillomavirus-induced pathogenesis

Muller Mandy

UFR Sciences du Vivant
Bâtiment Lamarck, Pièce RH71
Rue Hélène Brion,
case courrier 7136
75205 PARIS Cedex

Unité Génétique, Papillomavirus et Cancer Humain
Institut Pasteur
25 rue du Dr Roux
75724 Paris Cedex 15

UNIVERSITÉ DENIS DIDEROT - PARIS VII

École Doctorale B3MI – Biochimie, Biothérapie, Biologie Moléculaire et Infectiologie

Thèse de doctorat

Option : **Virologie**

Comparative mapping of E2-host interactions unravels new roles of E2 in human papillomavirus-induced pathogenesis

---

présentée par **Mandy Muller**

pour obtenir le titre de Docteur de l'Université Paris Diderot

Thèse préparée sous la direction de Caroline DEMERET

Soutenue le 28 juin 2013 devant le jury composé de :

Pierre-Emmanuel Ceccaldi	Président du jury
Cheng-Ming Chiang	Rapporteur
Antoine Touzé	Rapporteur
Françoise Bachelerie	Examinateur
Murielle Masson	Examinateur
Caroline Demeret	Directrice de thèse

If I have seen farther it is only by standing on the shoulders of giants.
I. Newton

La cartographie comparative des interactions E2-hôte souligne le rôle de E2 dans la pathogénie associée aux papillomavirus humain

Résumé - Les papillomavirus humains (HPV) sont de petits virus non-enveloppés contenant un ADN circulaire d'environ 8000 paires de bases. Ces virus sont strictement epithéliotropes mais peuvent infecter autant la peau que les muqueuses génitales ou encore orales, démontrant ainsi une grande diversité de tropisme. De même, la diversité pathogénique est très importante puisque les HPV sont connus pour être à l'origine d'infection asymptomatiques, de lésions bénignes hyperprolifératives telles que les verrues, mais aussi de lésions pouvant progresser vers le développement de cancers. Les HPV à bas-risque sont à l'origine de lésions bénignes tandis que les HPV à Haut-risque sont également responsables des lésions cancereuses. Notamment, ces HPV sont à l'origine des cancers du col de l'utérus, le second cancer de la femme, ainsi que nombreux autres cancers de la sphère ano-génitale. Les HPV sont également connus pour être impliqués dans le développement de certains cancers de la tête et du cou, occupant une place de plus en plus importante dans cette niche depuis la baisse des cancers induits par le tabac. Du côté de la peau, les HPV sont responsables de cancers chez les personnes immunodéprimées et chez des patients souffrant d'une maladie génétique rare : l'épidermodysplasie verruciforme. Dans la population générale, un rôle dans les cancers de la peau est pour l'heure suspecté mais fait encore débat.

L'organisation génétique est globalement conservée parmi les différents types de HPV. Le génome viral peut ainsi être séparé en trois parties : une région de régulation appellée LCR (pour Long Control Region) contenant l'origine de réplication de même qu'un certain nombre de séquences de régulation transcriptionelle ; une région tardive codant pour les protéines structurales L1 et L2 constituant la capside et particulièrement importantes pour les étapes d'entrée du virus dans la cellule hôte ; et une région précoce codant pour les protéines exprimés dans les premières étapes du cycle viral aboutissant à la réplication virale. Parmi les protéines précoces, E6 and E7 jouent un rôle central dans la déregulation du cycle cellulaire et de la prolifération, faisant de ces deux protéines les oncogènes viraux majeurs des HPV à haut-risque. La protéine E1 est l'hélicase virale impliquée dans la réplication des génomes viraux. La protéine E2, qui fait l'objet de cette thèse, intervient dans de nombreuses étapes du cycle viral. E2 est une protéine constituée de trois domaines : les domaines N-terminal et C-terminal sont bien conservés à la fois au niveau de la séquence primaire que de la structure 3D et sont séparés par une région charnière flexible non structurée et non-conservée. Le domaine N-terminal est un domaine de transactivation tandis que le domaine C-terminal est un domaine de dimérization et un domaine de liaison à l'ADN. Par sa capacité à se lier à l'ADN, E2 est un régulateur majeur du cycle viral. En effet, la région de régulation du génome des HPV contient des séquences particulières reconnues par les protéines E2 et appelées E2BS (pour E2 Binding Sites). En se liant à ces E2BS, E2 empêche le recrutement de facteurs de transcription essentiels et régule négativement l'expression des gènes viraux précoces E6 et E7. De même, en se fixant à ces E2BS, E2 aide le recrutement et la fixation de E1, l'hélicase virale, sur l'origine de réplication, faisant de E2 un élement essentiel à l'initiation de la réplication des HPV. Enfin, E2 est aussi connu depuis de nombreuses années comme étant le facteur viral responsable de la ségrégation des génomes viraux lors de la division cellulaire. En se liant à la fois à l'ADN viral et au chromosome de l'hôte, E2 agit comme un pont et permet ainsi de conserver un réservoir de genome HPV dans les cellules proliférantes de l'épithelium. E2 est donc une protéine essentielle à la fois au cycle viral productif et à la persistance virale, la persistance étant un facteur clef dans le risque de développement de cancer. De plus, depuis la dernière décennie, il émerge qu'au delà de ses fonc-

tions dépendantes de la liaison à l'ADN viral, E2 serait aussi capable de moduler directement la cellule hôte, principalement en établissant des interactions avec des protéines cellulaires. Il est ainsi envisagé que E2 pourrait directement participer à l'établissement de conditions cellulaires permissives au cycle viral le long de épithélium. De manière intéressante, certaines de ces fonctions de E2 sont spécifiques des HPV à haut-risque, tels que l'induction d'arrêts du cycle cellulaire, menant généralement à un certain degré d'instabilité genomique ou encore l'induction de l'apoptose, un mécanisme envisagé comme contre-balançant les propriétés transformantes de E6 et E7. Ceci a mené à l'hypothèse que E2 pourrait être directement impliquée dans des mécanismes menant au développement de cancer.

Depuis 2006, deux vaccins sont disponibles pour lutter contre les infections à HPV, cependant, la couverture vaccinale reste faible, le coût global est élevé et ces vaccins ne sont pas à visée thérapeutique. Il semble donc important de développer de nouvelles molécules et stratégies anti-virales. De par son rôle dans de nombreuses étapes du cycle et dans la persistance virale, sa bonne conservation de séquence et le fait qu'elle est exprimée précocement durant l'infection, la protéine E2 est envisagée comme une bonne cible pour développer une drogue anti-HPV. Cependant, il semble nécessaire de mieux connaitre l'impact de cette protéine virale sur la cellule hôte, en particulier, d'approfondir les notions de spécificité de fonction selon le pouvoir oncogène du HPV.

Le but de cette thèse était de mieux cerner le rôle de E2 dans la pathogenèse et la carcinogenèse associées aux infections par les HPV. Étant donné que les interactions protéine-protéine constituent un moyen efficace pour les protéines virales d'agir sur la cellule hôte, nous avons décidé d'aborder le problème des fonctions de E2 à travers la cartographie de son réseau d'interaction avec les protéines cellulaires. Néanmoins, contrairement à la plupart des études similaires qui se concentrent principalement sur les HPV les plus significatifs d'un point de vue clinique, nous avons choisi de tirer profit de la grande diversité de tropisme et de pouvoir pathogène de cette famille virale afin d'extraire des spécificités d'interaction propres à un type d'HPV. Nous avons donc sélectionné un panel de protéines E2 provenant de 12 génotypes HPV représentatifs de leur diversité : HPV cutanés (HPV1, 3, 5, 8, 9) ou muqueux (HPV6, 11, 16, 18, 32, 33, 39) ; HPV haut-risque (HPV5, 8, 16, 18, 33, 39) ou bas-risque (HPV1, 3, 5, 8, 9). Ces 12 protéines E2 ont été utilisées comme proies dans un criblage par double hybride en levure d'une banque d'ADN complémentaires issues de cellules HaCaT, une approche non-biaisée pour la détection d'interactions. Environ 200 protéines cellulaires ont été identifiées dans ce crible, dont seule une faible proportion interagissait avec plusieurs protéines E2. Cependant, en raison des caractéristiques intrinsèques des approches en levure telles que le taux d'interactions fausse-positives et surtout fausse-négatives, il est communément admis que les interactions identifiées doivent être validées par une autre technique. Nous avons donc sélectionné une centaine de protéines parmi les cibles les plus pertinentes identifiées dans le crible auxquelles nous avons ajouté un certain nombre de contrôles positifs, des protéines cellulaires connues dans la littérature comme étant des interacteurs de E2. Le tout combiné, nous avions 121 protéines cellulaires à tester dans notre deuxième étape de validation. Afin d'estimer le spécificité des interactions vis à vis des différents génotypes, il nous a semblé important de re-tester chaque protéine cellulaire avec l'ensemble des 12 protéines E2. Il nous fallait donc tester plus de 1400 interactions (121 protéines cellulaires contre 12 protéines E2). Pour cela nous avons utilisé la technique haut-débit développée dans notre laboratoire appelée HT-GPCA (pour High-Throughput *Gaussia princeps* Complementation Assay). Dans cette technique utilisée en cellule de mammifère, les protéines de l'interaction à tester sont chacune fusionnées à une moitié de l'enzyme Gaussia luciferase. Si les deux protéines interagissent, les fragments de la luciferase sont amenés à se rapprocher, ce qui est suffisant pour reconstituer l'activité enzymatique. Pour détecter une interaction, il suffit donc de mesurer un

signal luciferase. Cette technique a été montrée comme générant un taux d'interactions fausse-positives particulièrement bas ainsi qu'un très bon taux de recouvrement d'interactions connues, indiquant un faible taux d'interaction fausses-négatives. L'avantage de ce type d'analyse à large échelle est que ce n'est pas seulement des interactions qui sont comparées, mais des profiles d'interaction, ce qui permet de faire ressortir des informations générales sur les protéines virales étudiées. Dans notre cas, l'ensemble des données a été traité par « Hierarchical Clustering », une méthode permettant de regrouper les profiles d'interaction par similarité. Un dendrogramme a été généré - un arbre hiérarchique - classant les profiles d'interaction par ressemblance maximum. Nous avons comparé cet arbre basé sur les profiles d'interaction de E2 à un arbre phylogénétique basé sur les séquences des différentes protéines E2. De manière intéressante, nous avons ainsi pu observer que les deux arbres étaient très similaires, avec dans les deux cas, une première ségrégation des E2 basée sur une différence de tropisme, puis un second regroupement basé sur le pouvoir pathogène des différents HPV. Ceci signifie qu'en regardant simplement les interactions des protéines E2, il est possible de distinguer un HPV à haut-risque d'un HPV à bas-risque, suggérant deux hypothèses : soit les interactions sont le résultat de différence d'infection (différentes niches d'infection, modification globale du protéome de l'hôte...), soit, par ces interactions, E2 contribue directement au pouvoir pathogène des HPV. E2 étant de plus en plus associé à des fonctions autonomes pouvant potentiellement influencer la cellule hôte et le développement de cancer, il est tentant de penser que ce résultat est plutôt favorable à la deuxième hypothèse.

En utilisant toutes les interactions détectées dans cette étude, nous avons construit les réseaux d'interactions des protéines E2. L'analyse des degrés (nombre d'interaction connues) des cibles de E2 a mis en évidence que la protéine E2, comme de nombreuses autres protéines virales, cible préférentiellement des protéines hautement connectées dans la cellule hôte. Cibler des protéines cellulaires qui sont centrales à de nombreuses voies de signalisation permettrait aux protéines virales d'avoir un effet très large sur la cellule infectée en un minimum d'interactions. Finalement, pour avoir une vision plus fonctionnelle du réseau d'interactions, les cibles de E2 ont été classées en familles fonctionnelles en se basant sur leur classification en GO termes (Gene Ontology). Il a ainsi émergé que les cibles de E2 peuvent principalement être regroupées en cinq grandes familles fonctionnelles. La première famille émergeant de cette analyse correspond à des protéines impliquées dans la régulation de la transcription, ce qui corrobore le rôle principal de E2 comme facteur de transcription viral. Ainsi, en adoptant une approche haut-débit d'identification non-biasée d'interaction par double-hybride suivi d'une étape de validation comparative en cellule de mammifère, la principale fonction de E2 en tant que régulateur transcriptionel ressort en priorité, ce qui démontre la grande fiabilité de l'approche pour détecter les interactions importantes des protéines E2. Nous avons également mis en évidence un ciblage de protéines impliquées dans des mécanismes d'apoptose, d'ubiquitination, et de régulation des ARN, ce qui avait déjà été lié au rôle de E2 dans la cellule infectée, avec ici, un élargissement du spectre d'interactions connues. Finalement, nous avons mis en évidence un ciblage de protéines impliquées dans le transport intra-cellulaire et particulièrement de vésicules cytoplasmiques, ce qui correspond à une fonction potentiellement nouvelle pour les protéines E2. Le détournement des méchanismes de transport du système vésiculaire de l'hôte par les HPV a lieu à deux moments clefs lors de l'infection : lors de l'entrée du virus et de la translocation du génome viral au noyau, et lors de l'évasion immunitaire pour prévenir l'exposition d'antigènes viraux à la surface de la cellule infectée. E2 étant capable de se lier fortement au génome viral et à L2, une des deux protéines de capside, il est possible que E2 soit présent dans la particule virale, ce qui laisserait supposer que E2 pourrait avoir un rôle dans les mécanismes d'entrée du virus plutôt que dans la régulation de la présentation d'antigène.

Cette approche comparative d'étude des réseaux d'interactions de 12 protéines E2 a donc permis d'améliorer la compréhension globale des fonctions de cette protéine virale. L'aspect comparatif de cette étude en fait une approche de choix pour identifier des interactions spécifiques à un sous type de HPV qui pourraient être reliées à des caractéristiques pathologiques. En particulier, l'identification d'interactions spécifiques aux HPV à haut-risque pourrait être le premier pas vers l'identification de biomarqueurs permettant de détecter précocement l'apparition de cancers. Dans ce contexte, une interaction spécifique parmi celles identifiées a particulièrement attiré notre attention. Il s'agit de l'interaction entre une protéine cellulaire dénommée CCHCR1 et la protéine E2 spécifiquement de HPV16. CCHCR1 est impliquée dans des mécanismes de régulation de proliferation de kératinocytes, les cellules cibles des HPV. Étant donné l'aspect très important de la régulation de la prolifération cellulaire pour le développement de cancers et le caractère très spécifique de cette interaction pour HPV16, le HPV le plus représenté dans les cancers associés aux HPV, nous avons décidé d'explorer plus en détail l'impact potentiel de cette interaction sur la cellule infectée. Cette interaction est considérée comme spécifique puisque d'une part, les autres protéines E2 testées ne peuvent pas, ou alors que de manière très marginale, interagir avec CCHCR1 et que d'autre part, parmi toutes les interactions testées avec la protéine E2 de HPV16, l'interaction avec CCHCR1 est de loin la plus forte. Cette interaction avait précédemment été identifiée dans une étude en levure et les auteurs avait pu déterminer que l'interaction dépendait du domaine N-terminal de E2. Afin de cartographier de manière plus précise le domaine d'interaction, nous avons introduit des délétions en série des hélices alpha du domaine N-terminal de E2. Dès que la première hélice de E2 est retirée, l'interaction avec CCHCR1 est complètement abolie, ce qui est similaire à l'interaction avec BRD4, un partenaire majeur de E2. L'interaction entre E2 et BRD4 est très documentée et son domaine d'interaction sur E2 recouvre les trois hélices N-terminales de E2, il est donc envisageable que l'interaction entre CCHCR1 et HPV16 E2 soit similairement dépendante de ces trois hélices. Pour déterminer quelle surface des hélices est importante pour l'interaction, nous avons muté des acides aminés dans E2 exposés d'un côté ou de l'autre des hélices. Ici encore, les mutations des mêmes acides aminés affectant la fixation de BRD4 à E2 inhibent aussi l'interaction avec CCHCR1. Pris dans leur ensemble, ces résultats de mutagenèse semblent indiquer que BRD4 et CCHCR1 partagent la même surface d'interaction au niveau du N-terminal de E2. Nous avons donc voulu déterminer s'il y avait compétition entre ces deux protéines cellulaires pour interagir avec la protéine E2 de HPV16. Et en effet, nous avons pu mettre en évidence qu'en présence de CCHCR1, l'interaction entre BRD4 et HPV16 E2 est réduite d'un facteur cinq, et que cet effet est le résultat de l'interaction directe entre HPV16 E2 et CCHCR1 puisqu'en utilisant une autre protéine E2 capable d'interagir avec BRD4 mais non avec CCHCR1, aucune compétition n'est observable. Cette compétition a également une répercussion fonctionnelle puisqu'en présence de CCHCR1, l'effet activateur de BRD4 sur l'activité transcriptionnelle de E2 est réduit. Ces résultats tendent à montrer que l'interaction entre CCHCR1 et HPV16 E2 affecte la liaison de E2 avec BRD4, ce qui pourrait affecter les fonctions de HPV16 E2 en tant que régulateur transcriptionel du génome viral.

Afin de poursuivre l'étude de l'interférence de CCHCR1 sur les fonctions transcriptionelles de E2, nous avons mené des études de fluorescence. Il en ressort que CCHCR1 s'exprime dans le cytoplasme dans de petites structures rondes tandis que E2 d'HPV16 est plutôt diffus, à la fois dans le cytoplasme et le noyau, mais principalement dans le noyau. Cependant, lorsque les deux protéines sont co-exprimées, E2 est quasiment entièrement relocalisée dans les mêmes structures que CCHCR1, l'empéchant ainsi d'atteindre le noyau. Le fait que CCHCR1 piège E2 dans le cytoplasme renforce probablement la compétition observée avec BRD4 mais doit également avoir un fort impact sur les autres fonctions nucléaires de E2.

Nous nous sommes finalement penchés sur le lien entre cette interaction spécifique et la régulation de la prolifération des kératinocytes. La régulation de la prolifération est un mécanisme fortement détourné par les HPV pour leur propre profit. En effet, le site d'initiation de l'infection se situe à la base des épithélia, au niveau d'une couche de cellules proliférantes. Dans le cadre d'un epithelium non infecté, à un certain point, une des cellules de cette couche basale va se détacher et entrer dans un processus de différenciation qui l'amènera jusqu'au pôle opposé de l'épithélium où elle finira par mourir et être libérée dans le milieu environnant. C'est donc une véritable balance qui s'établit dans l'épithélium entre prolifération et différenciation et qui va déterminer le destin de la cellule. Cependant, en dépit des mécanismes de régulation de la différenciation, les HPV ont développé des mécanismes capables de maintenir la prolifération cellulaire dans les différentes couches de l'épithélium afin de subvenir à leur besoin pour la réplication virale. Nous avons montré que CCHCR1 favorise la prolifération, tandis que HPV16 E2 stimule la différenciation précoce des cellules. Mais lorsque les deux protéines sont co-exprimées, l'effet de E2 en tant qu'inducteur de la différenciation est fortement réduit. E2 est le facteur viral contre-balaçant les propriétés oncogèniques de E6 et E7. Or les protéines E6 et E7 maintiennent la cellule sous un stimuli permettant la prolifération des cellules. Il n'est donc pas étonnant de voir E2 s'opposer aux effets prolifératif de E6 et E7 en induisant la différenciation cellulaire. Cependant ici, il semblerait que E2 de HPV16, en interagissant avec CCHCR1, ait un désavantage dans ce mécanisme, ce qui pourrait augmenter le risque de prolifération incontrôlé stimulé par E6 et E7.

L'étude de l'interaction spécifique entre HPV16 E2 et CCHCR1 met donc en exergue un mécanisme qui pourrait potentiellement participer à un processus propre à HPV16 expliquant sa forte propension à progresser vers la conversion maligne.

L'ensemble des résultats obtenus durant cette thèse améliore la compréhension générale des fonctions de E2 lors d'une infection aux HPV. E2 apparaît comme un facteur critique qui participe au détournement des fonctions de l'hôte pour permettre au virus de se développer en dépit des moyens mis en oeuvre par la cellule pour se défendre. E2 émerge aussi comme un composant viral susceptible d'influencer directement l'issue de l'infection et de prendre part aux étapes préliminaires de conversion maligne.

Mots clefs : HPV, E2, Réseaux, Interaction Virus-Hôte, Détournement cellulaire, Différenciation de kératinocytes, HT-GPCA

Comparative mapping of E2-host interactions unravels new roles of E2 in human papillomavirus-induced pathogenesis

Abstract - Papillomaviruses are responsible for widespread infections in humans, causing pathegenesis ranging from inapparent infections to benign lesions, hyperplasia or cancers. Given the major public health concern due to HPV-associated cancers, most studies have focused on the early proteins expressed by the most clinically relevant HPVs most frequently found in cancers. Among the early proteins encoded by HPVs, the E2 protein regulates viral transcription, replication and mitotic segregation of the viral genome, mainly through the recruitment of host factors to the HPV regulatory region. E2 is therefore pivotal for both the viral productive cycle and for viral persistence, which is a major risk factor for cancer development. In addition, the E2 proteins have been shown to engage interactions important to directly modulate the host cell, thereby contributing to create suitable cell conditions for the successive stages of the HPV life cycle. Interestingly, some E2's roles have been demonstrated to be specific to the oncogenic HPVs, raising the idea that beyond its role in the general HPV regulation, E2 could also directly influence the fate of cancer development.

This thesis aimed at providing an overview of E2's functions across multiple HPV genotypes and at identifying specific features that distinguish the different HPV pathological traits. We mapped the virus-host interaction networks of the E2 proteins from a panel of 12 HPVs selected to be representative of the HPV diversity. Clustering of E2's interaction profiles correlated with the HPV phylogeny, raising the notion that E2 could directly contribute to the HPV pathogenesis. This work also emphasizes that the E2 proteins, like many other viral proteins, tend to target highly connected cellular proteins (cellular hubs), which is presumed to be an evolutionary way to maximize viral impacts on the host. E2 predominantly targets a subset of key cellular processes, like transcriptional regulation, apoptosis, RNA metabolism, ubiquitination or intracellular transport, which both confirms already known E2's functions and points to potential new functions. In addition, this large-scale comparative approach offers a framework to pinpoint interactions that are specifically associated with the most represented HPVs in cancers and therefore can be used as targets for the development of new therapeutics. In particular, we identified a specific interaction between the E2 protein from HPV16 and a cellular protein, CCHCR1, involved in the regulation of keratinocyte proliferation. We determined that CCHCR1's interaction domain on E2 overlaps with that of BRD4, a major interactor of E2, inducing a physical competition between the two cellular proteins. This competitive binding affects BRD4-mediated enhancement of E2's transcriptional activity, suggesting that the interaction with CCHCR1 might have an impact on the role of E2 in the infected cell. In addition, we showed that CCHCR1 induces the docking of HPV16 E2 into the cytoplasm which could further affect E2's nuclear functions. We also demonstrated that CCHCR1 impairs HPV16 E2's induction of keratinocytes early differentiation, presumably resulting from the negative effect of CCHCR1 on the nuclear functions of E2. This effect could have drastic consequences on the oncogenic potential of HPV16 and could participate to high prevalence in cancers of HPV16.

Taken together, these results enhance the general understanding of the impact of E2 during HPV infections and highlights its contribution in the HPV pathogenesis. E2 appears as a critical factor that participates in the global hijacking of the host cell to allow the virus to replicate despite the hostile environment. E2 also emerges as a viral component susceptible to directly influence the outcome of an HPV infection and to potentially impact on the preliminary steps of carcinogenic conversion.

Key words: HPV, E2, Network, Virus-host interactions, Cellular Hijacking, Keratinocyte differentiation, HT-GPCA.

Acknowledgments

I want to express my gratitude to the member of the jury and especially professor Chen-Ming Chiang for its careful reading of the manuscript and who came all the way from Texas to help me in the final steps of this thesis.

Tout d'abord, je souhaiterais remercier les membres de mon jury qui ont accepté de juger ce travail de thèse. Les professeurs Chiang et Touzé qui, par la lecture minutieuse et critique de ce manuscrit, m'ont permis de l'améliorer. J'exprime également toute ma reconnaissance aux docteurs Masson et Bachelerie pour avoir accepté d'en être les examinateurs. Je remercie également le professeur Ceccaldi d'avoir accepté de présider ce jury de thèse.

Je tiens à remercier toutes les personnes avec qui j'ai partagé le quotidien pendant ces années à l'Institut Pasteur. Tout d'abord Caroline qui a encadré mon travail de thèse, pour son éternel soutien, sa force, sa motivation, sa connaissance sans fin des papillomavirus, sa gentillesse et sa générosité. Je la remercie de m'avoir mené là où je suis aujourd'hui. Yves également, pour sa joie de vivre perpétuelle et son optimisme débordant, et Françoise pour sa grande rigueur et son amitié. Je remercie aussi Michel Favre et Sylvie Van Der Werf de m'avoir accueillit dans leurs laboratoires. Je remercie également Patricia pour son aide tout au long de ce travail de thèse. Un grand merci aussi toutes les personnes qui ont croisé mon chemin autour des papillomavirus, Christian, Hélène, Laurence, Guillaume, Delphine, Valentine, Caroline. Merci également à tous les membres de l'unité de Génétique Moléculaire des Virus ARN qui m'accueillent dans leur laboratoire sur cette dernière ligne droite.

Je remercie également la fondation l'Oréal pour son soutien financier qui m'a permis de finir cette thèse sereinement.

Je remercie tout particulièrement ma famille, mes plus grands fans et fervents supporters. Ma mère, qui savait qu'à coup de livres, je franchirais tous les murs ; mon père, qui ne dira plus jamais qu'il n'y a pas de quoi écrire une thèse ; mon frère et sa fille, ma nièce, que j'aime tant ; Cathy et Philippe pour tout l'amour qu'ils m'ont toujours donné. Une grosse pensée également pour mon Pépé et ma Mémé, mes oncles et mes tantes, mes cousins et cousines, parce que je suis fière de faire partie de cette grande famille.

Un grand merci à tous mes amis, particulièrement Alexandre, Cédric et Wahb qui me font partager leur vie de physiciens mais également à celles et ceux qui ont accompagné mes premiers pas dans biologie, je pense notamment à Audrey et Céline. Un petit clin d'œil à Newton, parce qu'il m'aura bien fait rire depuis 2 ans ;-)

Mes derniers remerciements vont à Romain, ma moitié, qui m'a soutenu et aidé tout au long de ces années, qui a toujours cru en moi et avec qui j'espère pouvoir poursuivre encore longtemps notre bout de chemin, aussi loin que le vent nous portera.

Contents

I . Introduction

From an evolutionist's point of view, papillomaviruses are very successful infectious agents. They induce unapparent chronic infections that rarely kill the host and periodically shed infectious virions for their transmission. In humans, certain types of papillomaviruses (HPV) referred to as High-Risk HPV (HR-HPV), are predominant risk factors for the development of cervical cancers and other epithelial cancers as well. HPVs are the most frequent sexually transmitted agent causing pathologies ranging from asymptomatic infections to benign hyperplasia or cancer and are therefore considered as a major public health concern. Interest has grown over the last decade to better understand the biology of these small DNA viruses.

A - Historic

Genital warts have been known for many centuries but until the 19th century, they were generally considered to be a form of syphilis or gonorrhea. It is only in 1907 that Giuseppe Ciuffo demonstrated the viral nature of human warts after cell filtrates from lesions were shown to transmit the disease [1]. Details on when and how the first strain of HPV was discovered are incomplete, but the papillomavirus was first glimpsed as a disease by R. Shope of Rockefeller University in the 1930s [2]. A rabbit strain of papillomavirus (nowadays referred to as CRPV, Cottontail Rabbit Papillomavirus) often causes horn-like warts on infected rabbits. The cause of these warts was not known at the time, but Shope took samples of the warts and injected them back into healthy rabbits. The healthy rabbits soon developed the same warts. Shope did not identify the exact viral strain, but he correctly deduced that the warts were caused by a virus. In addition to causing benign papillomas, some warts induced by CRPV were observed to undergo malignant progression [3]. As such, the Shope virus became an important model to study viral tumorigenesis. It was not until the 1970s and the advent of molecular cloning that researchers have started to better understand the biology of papillomaviruses. The sequencing of the genome led to the identification of the open reading frames (ORF) as putative viral genes resulting in the characterization of their functions. German virologist Harald zur Hausen proposed in 1976 that HPV was the cause of cervical cancer, a theory that other scientists originally rejected [4]. In 1984, the team of zur Hausen discovered HPV DNA in cervical cancer tumors, proving his theory [5]. In 2008, he received the Nobel Prize for this research.

B - General description

The identification in the early 1980s of HPV16 and HPV18 by the German team provided the field with HPV types present in most cervical cancers. Subsequent studies have sought to understand the natural history of HPV infection, determine the biological

properties of different HPV types, elucidate the role of viruses in pathogenesis and identify non-viral factors that may influence the outcome of an infection.

a . Classification

Papillomaviruses (PV) have been isolated in many host species from humans to birds or even reptiles. Given the high species-specificity of PV, there are hundreds of PV types. The most extensively studied are the Human Papillomaviruses (HPVs) with more than 200 different genotypes identified, of which 118 are fully sequenced [6–8]. If originally classified according to their host species, DNA sequencing of the HPV genomes has led to their classification according to the comparison of the L1 ORF, which encodes the HPV major structural protein (**Fig I.1**).

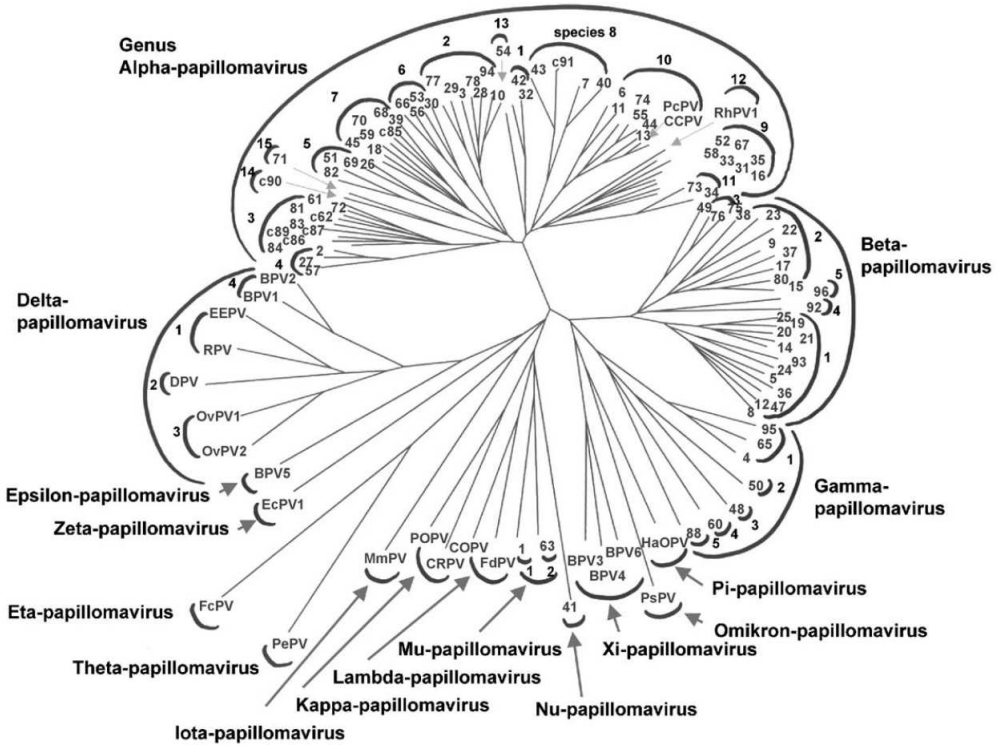

Figure I.1: **Phylogenetic tree of papillomaviruses**. Classification obtained by comparison of the L1 ORF sequences. The number at the end of each branch identifies a type of HPV, the other abbreviations refer to animal papillomavirus types. (From deVilliers *et al.* [7]).

The papillomaviruses are divided into several genera (designated with a greek letter) and further subdivided into genotypes (designated by a number for human papillomaviruses or a letter for animal papillomaviruses). HPVs are clustered into five genera: alpha, beta, gamma, mu and nu. This classification often correlates with the tropism and pathogenic potential of viruses, demonstrating that viruses and host have co-evolved. The most clinically relevant HPVs (those at the origin of genital cancers) are members of the alpha genus. Most of the human papillomaviruses of the alpha genus infect mucosal

epithelia while members of the other genera primarily infect the skin. The β HPVs include those associated with epidermodysplasia verruciformis (HPV5 and 8), a rare genetic predisposition to widespread non-genital HPV cancerous lesions.

b . HPV infection

HPVs are strictly epitheliotropic viruses that induce in the general population frequent asymptomatic infections. The HPV life cycle is tightly linked to the differentiation program of keratinocytes, their target cells. HPVs are implicated in the development of benign lesions of the skin (wart) or of mucous tissues (condylomas) as well as in malignant hyperplasia.

1 . Mucosal HPV

To establish infection, HPVs are believed to infect epithelial cells that possess proliferative capacities. As these cells are located at the basal layer of the epithelium, it is commonly accepted that it is microtrauma of the upper layers of the epithelium that provide access for the viruses to the basal cells. Such microtrauma occur frequently during intercourse and it is thus accepted that the standard transmission process is sexual intercourse. The prevalence of HPV infection is thus usually correlated with measures of sexual promiscuity like the number of lifetime sexual partners, recent change in sexual partners and history of other sexually transmitted infections. Genital HPV infection is the most common sexually transmitted viral infection [9, 10]. Infections by the genital HPVs are widespread and are associated with a broad range of clinical manifestations. HPV types like the HPV6 and 11 are the etiologic agents of benign hyperplasia such as genital warts or Condylomata acuminata. Although these lesions can resolve on their own, they are often recurrent and there is currently no long-term effective treatment. Other HPVs like HPV16, 18, 31, 33, are associated with the development of intra-epithelial neoplasia or pre-cancerous lesions. Most of the lesions regress spontaneously in less than a year [11, 12], however, in some cases, the lesions can persist, which represents a major risk for the development of cancer. HPV16 infections are more likely to persist than infection by other HPV types [13] and this particular HPV account by itself for more than 50% of cervical cancers. Cervical cancer is the second most common cancer in women worldwide, with more than 500,000 newly diagnosed cases each year [14], most of them occurring in developing countries. The disproportionately high burden of cervical cancers in developing countries and in medically underprivileged populations is mainly due to a lack of screening that allows detection of precancerous lesions at early stages [15]. Most cancers occur in the transformation zone of the cervix, at the junction between the endocervix and the stratified squamous epithelium of the exocervix. Histological classification of cervical intraepithelial neoplasia (CIN) grades the lesions from 1 to 3 based on the severity of the dysplasia: CIN1 corresponds to a mild dysplasia , CIN2 to moderate dysplasia, and CIN3 to severe dysplasia or carcinoma.

Infections of the oral mucosa by certain types of HPVs have also been reported for many years [16]. Most of Head and Neck cancers are caused by smoking and/or alcohol consumption but consistent data indicate that a subset of oral cancers, mainly in the upper oral area including tonsils, base of the tongue and soft palette, are attributable

to HPV infection, with HPV16 accounting for about 90% of the HPV-positive tumors [17–19].

2 . Cutaneous HPV

Skin warts are benign papilloma most frequently found on the hands and feet and lesions are usually small dome-shaped papules with a keratotic and verrucous surface [20]. They occur commonly in older children and young adults [21]. Since HPVs need to infect basal cells, maceration of the skin is believed to be a predisposition for infection. There is a good correlation between HPV types and clinical lesions with HPV1, 2, 4, 27, 57 and 65 usually associated with common and plantar warts, whereas HPV3, 10, 28 and 41 are mostly associated with flat warts. The lesions tend to be self resolving, but complete clearance may take up to several years. Regression of lesions is most likely due to an effective immune response and the low incidence of warts in older individuals might suggest that immune mechanisms have rendered them resistant to infection.

Patients with a rare genetic disorder, Epidermodysplasia verruciformis (EV), have a unique susceptibility to cutaneous HPV infections [22–24]. Starting during childhood, warts quickly spread over the body, tend to persist, and may even progress to squamous cell carcinomas. About a third of EV patients develop skin cancers and if most of the lesions remain local, distant metastases may appear. Patients with epidermodysplasia verruciformis are usually infected with multiple types of HPVs, including the common types that affect individuals of the general population. The development of tumorous lesions are primarily associated with HPV types referred to as EV-HPVs, mostly HPV5 and 8, which are present in 90% of cancers in EV patients. EV-HPVs were for long believed to be specific to the disease since they could not be detected in patients without EV, except in immunosuppressed populations. However, with the increasing sensitivity of detection methods, it was possible to show the presence of the EV-HPVs in the general population as well [25]. EV occurs as an inherited disorder with an autosomal recessive pattern. Two distinct chromosomal loci have been linked to the HPV predisposition in EV patients: EV1 and EV2. Two adjacent genes, EVER1 and EVER2, were identified within the major locus EV1, whose nonsense mutations are associated with the development of the disease [26, 27], suggesting a potential pivotal role in the control of the infection.

A high prevalence of HPV DNA in squamous cell skin carcinoma of immunosuppressed, but also of immunocompetent patients, has renewed the great interest in a possible etiologic role of HPVs in nonmelanoma skin cancer. Nonmelanoma skin cancers (NMSC), a frequent form of skin cancers, can be divided into two groups: basal cell carcinomas (BCC) and squamous cell carcinomas (SCC). Usually developing on skin areas exposed to sunlight, these tumors rarely metastasize. Consistent findings of certain HPV types in SCC makes HPV an attractive co-factor of UV radiation for cancer development [28]. Because of the low copy numbers of HPV DNA in skin cancers, probably not every tumor cell contains viral DNA, which is compatible with cutaneous HPVs being possibly important for tumor initiation, but not for maintenance of the malignant phenotype. This would suggest a "Hit-and-Run"mechanism, where the presence of viral DNA would be sufficient, in association with co-factors, to trigger malignant conversion, but the subsequent loss or dilution of viral DNA copies would not impede further cancer progression [29].

c . Genomic organization

Papillomaviruses are small, non-enveloped, DNA viruses infecting squamous epithelial cells. Viral particles are approximately 55 nanometers in diameter [30] containing a single double-stranded circular DNA molecule of about 8,000 base pairs within an icosahedral capsid of 72 capsomers (**Fig I.2**).

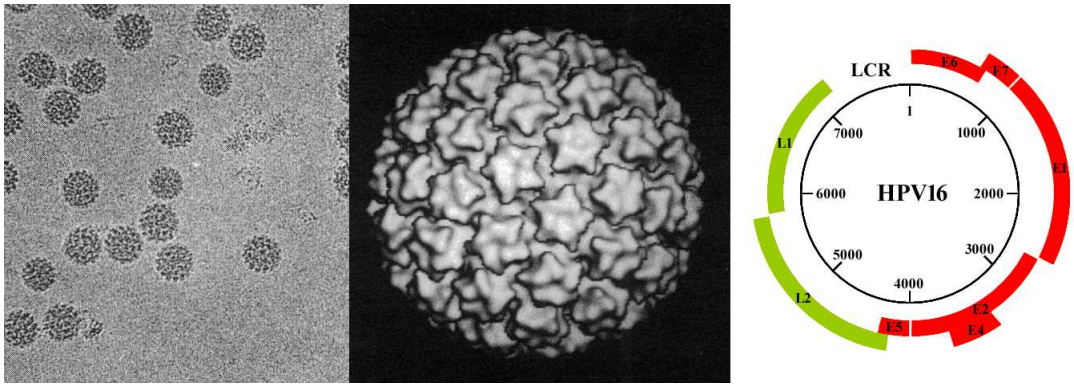

Figure I.2: **Papillomavirus particles and genome**. Left: Electron micrograph of Bovine papillomavirus 1 (BPV1). Middle: Three dimension image reconstitution of virion particle structures (Adapted from Baker *et al.* [30]). Right: HPV16 genomic map. The number in the circle indicate the nucleotide position. The ORF regions are represented either in red for the early proteins or in green for the late proteins. The long control region (LCR) is represented at the top.

The genomes of numerous human and animal papillomaviruses have been fully sequenced and it appears that the genomic organization of most PVs is similar (**Fig I.2**). All open reading frames are located on the same DNA strand, meaning that only one strand serves as a template for transcription. ORFs are classified either as early (E), encoding the regulatory proteins or late (L), encoding the structural proteins. A region called LCR (for Long control Region) of about 1kb devoid of ORF contains the viral origin of replication as well as important transcriptional control elements. Each of these regions are described in the following sections.

1 . LCR

The LCR of papillomaviruses contains sequence elements that are responsive to cellular factors, as well as virally encoded regulatory proteins (**Fig I.3**). Typically, the LCR includes a tissue-specific enhancer that plays an important role for viral gene expression in keratinocytes and may also be important for viral latency. The LCR contains numerous transcription binding sites such as AP1, SP1, Oct1, and YY1 among others [31–34]. No specific factors have been identified to be responsible for the keratinocyte-restricted activity of the enhancer. It seems that the specificity would rather be conferred by a complex interplay among these multiple ubiquitous transcription factors (see [35] for review).

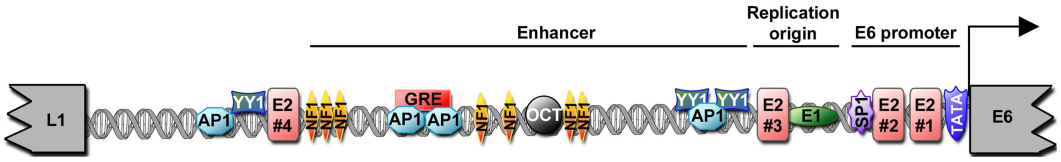

Figure I.3: **HPV16 Long Control Region**. Schematic map of the long control region of HPV16. The end of the L1 and the beginning of the E6 ORFs are indicated. The E2 binding sites (E2BS) are numbered from 1 to 4 and colored in red.

The LCR also harbors a glucocorticoid response element that is differentially regulated among the HPVs [36–38]. In addition to the binding sites for cellular transcription factors, the LCR contains recognition elements for the virally encoded E2 protein (referred to as E2 binding sites or E2BS) as well as the origin of viral DNA replication where the E1 helicase is loaded. At least four E2BS are present in the LCR of all α-HPVs and display a conserved relative position near the initiation start of the early promoter allowing the coordinated regulation of viral DNA transcription and replication. The LCR of the β-type HPVs is smaller and the regulatory elements are positioned differently than in the α-type HPVs [39, 40]. In particular, the E2 binding sites are located farther from the early promoter, which greatly influences the regulation of viral gene expression [41]. This will be developed later in this manuscript.

2 .Early region

The early region represents about half of the total genome and encodes proteins required during the early steps of the viral life cycle, before the onset of viral genome amplification. This region encodes four to seven proteins:

- E1 - E1 is the longest of all HPV proteins with 600 to 650 amino acids, but is also well conserved in sequence [42, 43]. E1 contains a DNA-binding domain as well as a helicase domain and an ATP fixation motif [44, 45]. E1 has DNA-dependent adenosine triphosphatase (ATPase) and DNA helicase activities [46, 47], which are pivotal for viral replication. E1 functions both as a DNA binding protein that binds with a weak affinity the viral origin of replication and as an helicase to unwind DNA ahead of the replication fork. The binding of E1 to the LCR is stabilized through its cooperative binding with the E2 protein [48, 49] (and [35] for review) which, in contrary to E1, associates with high affinity to its E2BS in the viral origin. Through interactions with E1 and E2, cellular replication factors such as DNA polymerase α-primase [50–52], topoisomerase I [53], and the single-stranded DNA-binding protein RPA [54] are recruited to the replication origin for assembly into an active replication complex. E1 then assembles in a hexameric complex that, in the presence of ATP, unwinds and bends the viral DNA and initiates replication [55]. In addition to being required for replication initiation, E1 is also important for DNA elongation and moves along the viral DNA template [49].

- E2 - E2 is a key protein for the viral life cycle and has a wide impact on the viral gene expression. The role of E2, being the focus of this thesis, will be detailed in a

specific section later in the introduction.

- E4 - The E4 ORF overlaps with that of E2 but is encoded in a different frame. Although E4 is classified as an early gene, it is expressed during the late stages of the viral life cycle, concomitantly with L1 and L2 [56–59]. The E4 protein has been shown to be synthesized from a spliced mRNA obtained from a donor site in E1 and an acceptor site in E2 at the beginning of the E4 ORF. It therefore results in the translation of the first five amino acids of E1 at the N-terminus of E4 and generates a protein often referred to as E1^E4. Recently, other transcripts were described and correspond to fusions between the N-terminal part of E2 and the complete ORF of E4 [60].

E1^E4 transcripts are produced throughout the HPV life cycle; however, the highest levels are found in the differentiated suprabasal layers [61]. E4 is primarily a cytoplasmic protein and colocalizes with the intermediate filament network. It has been shown to induce the reorganization and degradation of this keratin filament network to potentially favor viral particle release by weakening the upper layer of the epithelia [62, 63]. In addition, E4 was also shown to be localized at the mitochondria which leads to their detachment from the microtubules followed by a reduction of the mitochondria membrane potential and subsequent apoptosis induction. It was hypothesized that it could further facilitate the release of newly synthesized viral particles [64]. E4 is also able to arrest the cell cycle in G2, presumably to create a pseudo S-phase that optimizes conditions for viral DNA replication [65].

- E5 - The HPV E5 proteins are small hydrophobic proteins localized predominantly at the membranes of the endoplasmic reticulum, and occasionally at other cellular membranes [66, 67], whose role is still nebulous [68]. The E5 ORF is absent in the genome of many HPVs, such as beta-, gamma- and mu-HPVs, indicating that this protein is not essential for the life cycle of these viruses but rather can participate in infection and transformation. HPV16 E5 self-associates and this oligomerization is mainly mediated by hydrophobic interactions [69]. The E5 protein displays some transforming activities since it was shown to induce cell growth and tumorigenic transformation in transgenic mice in association with host factors [70–72]. It has been proposed that E5 can associate directly with the EGF (Epidermal Growth Factor) receptor, which may augment downstream receptor signaling and thus promote cell proliferation. In addition, E5-mediated effects on endosomal maturation and movement have been implicated in alteration of host cell antigen presentation [73, 74], therefore potentially favoring immune evasion.

- E6 & E7 - Overall, E6 and E7 disrupt or usurp multiple cellular signaling pathways to maintain infected cells in a proliferative state necessary for viral replication. However, in the case of HR-HPVs, it can lead to an increased genomic instability, and can result in full transformation. To understand the key role of E6 and E7 in HPV infection, it is important to keep in mind that the HPV life cycle is closely linked to the differentiation process of keratinocytes. Indeed, HPVs rely on host factors and in particular, on the host replication machinery to achieve their replication. However, the cells normally exit the cell cycle upon detachment from the epithelium to enter the differentiation process. One crucial aspect of HPV infection is therefore to uncouple the cellular proliferation and differentiation capacities, and this is mainly mediated through the activities of E6 and E7. The HPV E7 protein binds to Rb (Retinoblastoma) family members and targets them

for degradation [75]. This results in the release and activation of E2F transcription factors that drive expression of S-phase genes. The interaction between E7 and Rb is conserved for different types of HPVs but a much higher affinity is observed with High-risk HPVs [76, 77]. The binding of Rb by E7 and the subsequent forced S-phase gene expression would normally lead to cell growth inhibition and apoptosis through p53-dependent pathways. However, the E6 proteins from HR-HPVs have evolved to target the tumor suppressor p53 for degradation [78], thereby preventing cell growth inhibition and other p53-mediated responses. To interfere with p53 functions, E6 recruits the cellular E3 ubiquitin ligase E6-associated protein (E6AP), which leads to the ubiquitination and proteasomal degradation of p53 [79, 80]. Low-risk HPV E6 proteins can also complex with E6AP but this does not result in p53 degradation [81, 82]. The E6 protein of high-risk HPV types also plays a role in mediating cell proliferation independently of E7 through its C-terminal PDZ-binding domain, which can mediate suprabasal cell proliferation [83, 84]. These combined activities make high-risk HPV E6 and E7 proteins the primary transforming viral oncoproteins. The action of high-risk E6 and E7 proteins in targeting cell cycle regulators to maintain S-phase competence in differentiating cells results in perturbation of many cell cycle checkpoints. In long-term HPV-infected cells, this may lead to the accumulation of cellular mutations over a long period of time, which further promotes progression toward cancer [85].

Both E6 and E7 are small proteins, approximately 18 and 13 kDa in size, respectively. Despite this small size, E6 interacts with numerous cellular partners [86, 87] by, among others, recognizing Leucine-rich motifs containing the LxxLL consensus sequence or PDZ domains [83, 88–90]. Possessing nuclear localization signals, High-risk HPV E6 proteins are localized in the nucleus and in the cytoplasm while those of Low-risk HPVs are mainly cytoplasmic [91]. E7 accumulates both in the cytoplasm and in the nucleus and the presence of both nuclear localization and export signals in its C-terminal domain suggests that E7 is able to shuttle between the nucleus and the cytoplasm [92].

To conclude, the E6 and E7 proteins do not have an intrinsic enzymatic activity but rather act on the targeting of host cellular factors and the subsequent deregulation of cellular pathways to promote viral replication. The main oncogenic activities of HR-HPVs can be attributed to E6 and E7, but other early viral proteins also contribute to the overall carcinogenic conversion, like the E5 proteins and as we will discussed later, the E2 protein.

3 .Late region

The late region of the HPV genome encodes two structural proteins: the major capsid protein L1 and the minor capsid protein L2. Both proteins are essential for the assembly of the viral capsid but also for viral entry.

- L1 - L1 is the main factor for viral capsid oligomerization resulting in the formation of capsomers, the basic structural unit of the viral capsid.

L1 is synthesized during the late stages of infection, in the upper layer of infected epithelia, and is required for virion production and assembly. The viral capsid is composed of 360 L1 molecules, and up to 72 copies of the minor capsid protein, L2. L1 spontaneously oligomerises into pentamers, termed capsomers, which are the primary constituents of the

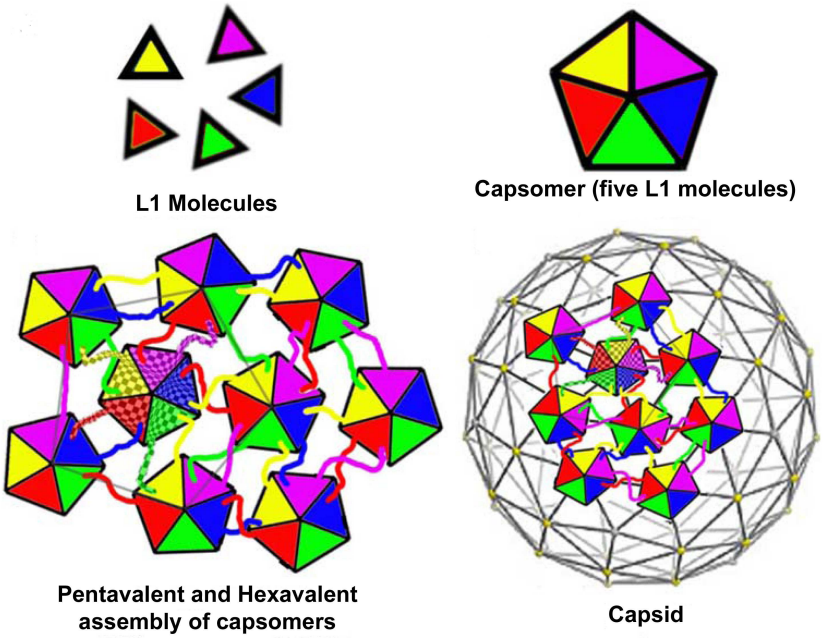

Figure I.4: **Interaction between capsomers**. L1 molecules associate to form pentameric capsomers. Pentamer contacts form pentavalent or hexavalent capsomers. Adapted from Modis *et al.* and Pereira *et al.* [93, 94].

outer capsid shell. Capsomers are linked together by the carboxy-terminal domains of L1 and are stabilized by intercapsomeric disulfide bonds between highly conserved cysteine residues (**Fig I.4**).

L1 mediates the primary attachment of viral particles to the cell surface or extracellular matrix (ECM) of target cells [95, 96]. Surface binding triggers changes in the conformation of the L1 protein resulting in the exposure of the N-terminal end of the minor capsid protein L2, necessary for further internalization steps. After viral uncoating, L1 segregates from the L2/viral genome complex in an endocytic compartment and is targeted to lysosomes for degradation [97].

The L1 protein contains all the intrinsic information required to form the capsid structure and is therefore sufficient to produce virus like particles (VLP) that mimic the native virion structure [98]. VLPs provide an efficient system to understand papillomavirus particle assembly, structure, and the binding of virus particles to the host cell.

Two vaccines against HPV infections have passed rigorous human trials and are currently commercialized: namely, Gardasil (Merck) and Cervarix (GSK). These vaccines exploit the fact that L1 expression alone leads to formation of VLPs, which have proven highly effective at producing a strong immune response that can protect against infection in humans [99, 100].

- L2 - L2 is a multifunctional structural protein naturally incorporated in the capsid. The L2 proteins are located beneath the L1 pentamers [101] meaning that there are 72 potential sites in the capsid in which L2 could be found (72 L1 capsomers in the capsid), although not all of them are believed to be occupied. L2 molecules seem to

interact with each other in an intercapsomeric-dependent manner, with the C-terminal region of one L2 molecule closely apposed to the N-terminal region of another [101]. Moreover, L2 is able to interact with L1 primarily through hydrophobic contacts [102]. L2 is critical for establishment of infection and evidence suggests an essential role for the L2 protein in many different steps of viral uptake by the infected cell [94, 103, 104]. L2 has been implicated in virion binding to the cell surface following its cleavage by furin [105], and is essential for the release of viral DNA from the endosomes [103], which seems facilitated by an L2 membrane-destabilizing motif [106] and to require L2 furin precleavage despite this step occurs at the cell surface [107]. L2 also helps in the transport and entry of the viral DNA to the nucleus [108]. However, the mechanisms involved in these processes are still evasive. Interactions of the L2 protein with microtubule motor proteins [109, 110] have been reported, suggesting that movement along microtubules might take part in the transport of HPV DNA to the nucleus. L2 proteins harbor two nuclear localization signals, suggesting that L2 could mediate nuclear import of viral genomes via nuclear pore complexes [111].

Interestingly, L2 also plays a pivotal role in virion production during viral life cycle, helping both in virion components gathering during the assembly of viral particles [112] and in the encapsidation of the viral genome through its interaction with the HPV DNA [104, 113, 114]. An interaction between L2 and E2 has been identified [115], another viral genome binding protein, which could further help the recruitment of HPV DNA to the site of virion assembly.

L2 is also necessary to generate virus-like particles. VLPs can form in the absence of L2 but L1-only VLPs have been noted to be more variable in size and shape than L1/L2 VLPs, leading to the speculation that L2 improves capsid formation. Similarly, L2 seems to increase VLP internalization when included in the particle [104, 116].

Immunological research on L2 has revealed promising cross-neutralisation potential [117, 118]. These characteristics are important in considering L2 or L2-containing constructs as vaccine candidates. Currently, phase-I trials of a vaccine utilizing an L2 construct are planned, and results of such trials are eagerly anticipated [119–121].

d . Viral cycle

HPVs are strictly epitheliotropic viruses whose life cycle is closely linked to the differentiation process of keratinocytes. Once the viral particle reaches the basal layer of the epithelium, presumably following microwounding, it enters the cell and establish infection. As the infected cell undergoes differentiation and moves toward the upper layer in the epithelium, the HPV proteins hijack the replication machinery to trigger genome amplification and later, viral production. Ultimately, virions are shed from the uppermost layer of the epithelium (**Fig I.5**).

1 . Non-productive cycle

In the undifferentiated cells comprising the proliferating part of the epithelium, the viral genome is maintained at low copy number as a stable nuclear episome. Only a subset of viral genes, the early genes, is selectively expressed in the basal compartment,

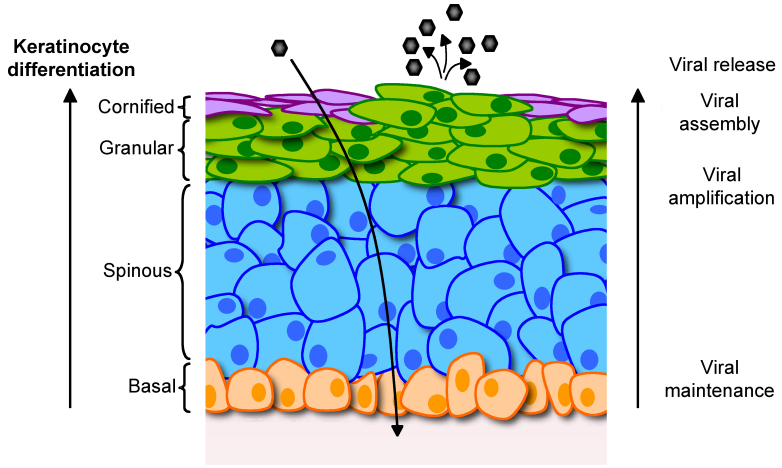

Figure I.5: **HPV life cycle**. Names of the epithelium layers are indicated on the left and the different viral steps on the right.

and therefore no new viruses are produced. Consequently, the first phase of the HPV life cycle in the epithelium basal layers is commonly referred to as the non-productive cycle.

Attachment and entry

For non-enveloped viruses such as HPVs, the protein coat covers and protects the viral DNA and provides the initial interaction site of the viral particle with the host cell [94]. After receptor engagement, the viral particle is internalized and its coat is disassembled to allow the encapsidated genome to access the cellular transcription and replication machinery in the nucleus (**Fig I.6**).

Initial binding is believed to occur at the basal membrane which underlies the epithelium. The viral particle reaches these regions only after their exposure by mechanical or chemical trauma [122, 123]. The initial interaction depends primarily on L1 [124–126]. Early work investigating host cell entry of HPVs showed that this process is initiated by binding of the virus particle to cell surface receptors, which are widely expressed and evolutionary conserved. Heparan sulfate proteoglycans (HSPG) are frequently found in the extracellular matrix and on the surface of most cells and were proposed as initial attachment receptors for HPV particles [95, 127]. HSPG function as more than simple attachment factors as they promote essential conformational changes in the viral capsid. However, HSPGs are clearly not the cell surface receptors that mediate virion internalization or later events in infection [105]. Accumulating evidence suggests that a secondary receptor or co-receptor is also involved in the infectious internalization of HPV subsequent to interaction with HSPG [128]. A role for L2 in facilitating infection via interaction with a secondary receptor has been suggested [108].

Initial capsid interaction with HSPG results in a conformational change which induces the exposure of a highly conserved furin cleavage motif in the N-terminal part of the L2 protein. After cleavage, an additional conformational change is likely to expose the binding site for the secondary cell receptor [105].

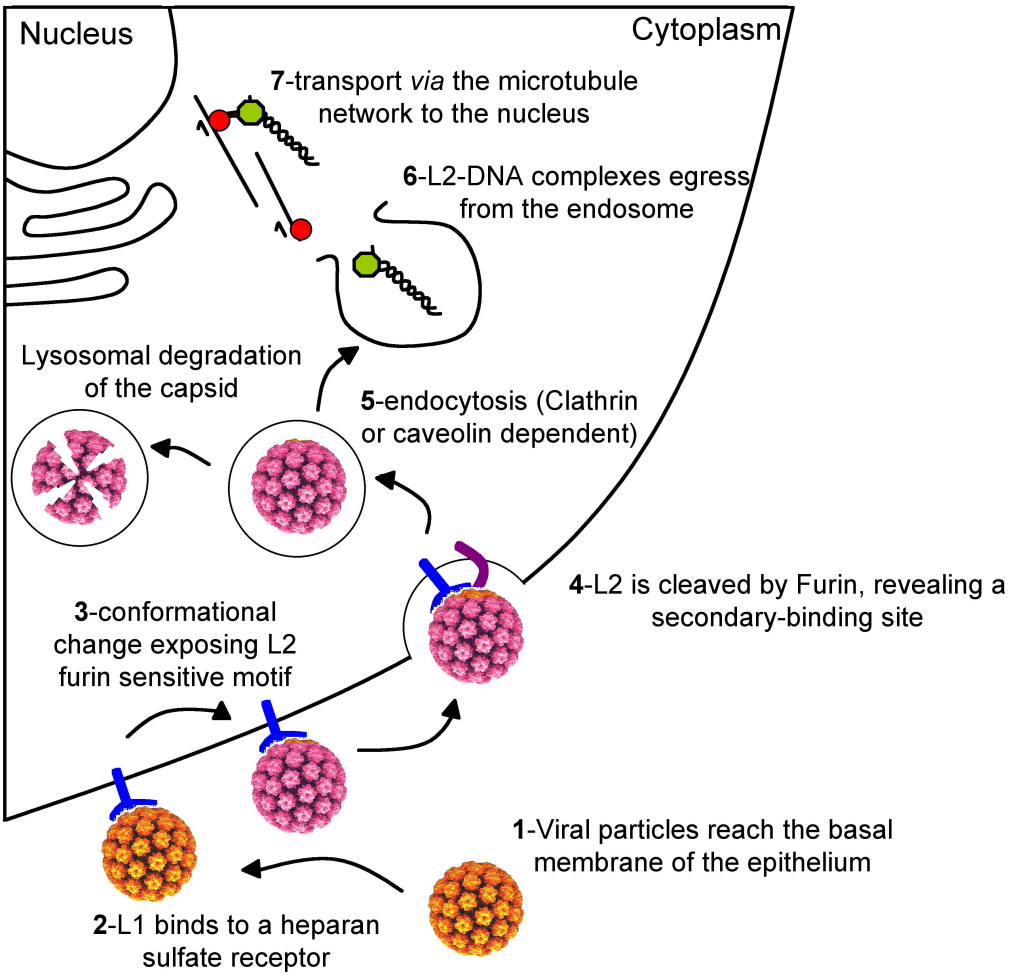

Figure I.6: **Model for HPV entry**. Schematized representation of viral entry processes. Modification in the virion color symbolizes the conformational change. Adapted from *Pereira et al.* [94].

Internalization and uncoating

After binding to cell surface receptors, HPV particles are internalized into the cell to establish the infection. To date, the dynamics of HPV interaction with the cell surface during the initial stages of infection are not completely understood and both the entry mechanisms and the molecules involved are still a matter of scientific debate [129]. Productive entry of HPVs involves internalization by endocytosis, a process that for HPVs occurs slowly and asynchronously over a period of several hours [130]. Several endocytic pathways have been described and clathrin- and caveolae-mediated are two main pathways presumably used by HPVs [131–135]. Both trafficking routes eventually converge to the endosomal pathway.

A carrier vesicle intermediate is then used to deliver virions into endosomes and lysosomes. Uncoating is not observed until approximately 8-12 hours after cell surface binding. Viruses have evolved several mechanisms to exit the endosomal compartment in order to

access the cytoplasm. For enveloped viruses, it usually requires fusion of membranes, either at the cell surface or after internalization [136]. Non-enveloped viruses can either lyse [137] or generate a pore in the limiting vesicular membrane [138] to allow escape of the viral genome into the cytosol. Pathogens that proceed through the endosomal pathway during trafficking in the host cell also typically take advantage of the pH acidification of the endosomal compartments [139–141]. Acidic pH acts as a trigger for many viruses to undergo conformational changes, leading to a number of events that facilitate endosomal escape of virion proteins and genomes. Such events may include modification of the viral-receptor interaction, exposure of protease digestion motifs or partial to complete uncoating of the viral genome. Although a C-terminal region of the HPV minor capsid protein L2 has been identified to display pH-dependent membrane-destabilizing activity [106], the exact mechanism by which this structure may assist in endosomal escape remains unclear. L2 was shown to interact with the sortin nexin 17 (SNX17) and this was demonstrated to mediate the retention of HPV virions in late endosomes, preventing their rapid lysosomal degradation and thereby favoring L2-DNA complexes to egress from the endosomes [103]. L2 was also recently demonstrated to contain transmembrane motifs which were hypothesized to be able to oligomerize, potentially forming a pore through the endosomal membrane and therefore important for the transfer of the L2-DNA complexes through the membrane [142]. All data combined point to L2 as a critical factor for HPV infection and more particularly as a key protein for viral genome egress from the endosome. Meanwhile, L1 does not appear to exit from the endosomal compartment but is ultimately destructed in lysosomes, confirming that L2 is the main factor allowing viral DNA egress from the endosome.

Nuclear import

Infection by DNA viruses replicating in the nucleus requires transport of the viral genome through the cytoplasm, a complex barrier due to its viscosity and the presence of a dense network of microtubules, actin, and intermediate filaments. A common strategy adopted by viruses to overcome this obstacle has been to use the cellular transport machinery to move along microtubules. In particular, the microtubule disrupting drug nocodazole inhibits HPV infection [131, 132] suggesting that the microtubule network integrity is essential for the early steps of HPV infection. Cytoplasmic transport along microtubules is mediated by motor protein complexes, and L2 has been found to interact with the microtubule network via motor proteins [109, 110] suggesting that the L2-DNA complexes reach the nucleus by moving along the microtubule network. Additional questions are being asked concerning the way the genome enters the nucleus but recent data suggest that it may follow nuclear membrane breakdown during mitosis [143] rather than through active transport of the L2-genome complex via karyopherins [111]. Once in the nucleus, the L2/DNA complexes predominantly localize in distinct punctuate nuclear domains designated ND10 bodies or promyelocytic leukaemia (PML) oncogenic domains (PODs), as determined by their co-localization with PML, the ND10-defining protein [144]. The subsequent steps of the viral life cycle will be initiated in the nucleus.

2 . Productive cycle

In a normal epithelium, basal cells proliferate and undergo cell division, but at some point, a daughter cell loses contact with the underlying basement membrane, migrates toward the upper layers of the epithelia and this serves as a signal to exit the cell cycle and initiate a terminal differentiation process. As cells move through the distinct epidermal layers, they acquire a more differentiated phenotype, ultimately resulting in cell cornification, cell death, and shedding into the environment. However, for a sustained viral replication and genome amplification, HPVs need the cell to maintain an active replication machinery within the suprabasal layers of the infected epithelium and have therefore evolved activities to counteract the natural cell cycle exit. Because of the complex interplay between the arrest of cell division and the onset of terminal differentiation, the ability of HPV proteins to reinitiate cellular replication implies an underlying capacity to alter the cellular differentiation program [145]. As discussed earlier, E6 and E7 are involved in the hijacking of numerous host cell pathways to maintain the proliferative state and hamper the normal differentiation of keratinocytes [146], which generates hyperproliferative lesions or hyperplasia, like warts, characteristics of HPV productive infections. Also, continuing cell multiplication increases the reservoir of cells that will ultimately produce high amounts of infectious virions. On the other hand, completion of the HPV productive cycle also requires cell progression toward its natural differentiation course, since both capsid genes transcription from the late promoter and high levels of E1 and E2 production required for viral genome amplification depend on cellular factors only present in these differentiated cells. Virion release also occurs via the natural shedding of the cornified cells, the last step of epithelial differentiation. Consequently, despite the function of E6 and E7 in inducing proliferation, hyperproliferative cells at one point commit to the differentiation program. This probably occurs through the combined action of viral proteins opposing E6/E7 as will be discussed in more details later for the E2 protein, and of cellular events naturally directing keratinocytes differentiation. The HPV productive cycle thus depends on a complex and timely manipulation of the balance between cell proliferation and differentiation, which constitutes a unique characteristic of HPV infection.

Replication and genome amplification

Historically, the dependence of the HPV life cycle on cellular differentiation has impeded the study of the viral late functions. Most cell lines used to study HPVs are derived from malignancies and contain integrated viral genomes with impaired late functions. Continuing advances in organotypic or raft tissue culture systems have permitted the growth of differentiated keratinocytes *in vitro* and provided a permissive experimental system for the complete HPV life cycle [147, 148].

Together with the cellular replication machinery, replication of HPV genomes requires the viral helicase E1 and the origin-binding protein E2. Upon entry of the viral genome into the nucleus, a first step of limited viral DNA amplification establishes the genome at low copy number per cell. The viral genomes are then replicated concomitantly with the host cell for an average of one time per cell cycle and are maintained as stable episomes to ensure persistence in the basal dividing cells of the epithelium. The E2 protein has

a central role in maintaining the viral DNA in these replicating cells both by activating replication in association with E1 and by tethering the viral episome to the host chromosome during mitosis. The precise mechanism will be discussed later in this manuscript. Upon differentiation of infected cells, a burst of viral DNA synthesis occurs in cells that express high levels of E1 and E2 corresponding to productive replication and giving rise to viral genome amplification to more than 1000 copies per cell [149].

Assembly and virion egress

Little is known about viral particle assembly and release. The assembly of viral particles together with the encapsidation of viral DNA has been proposed to occur in the PML bodies [144, 150] but the exact mechanisms and precise sequence of events that lead to virion formation remain largely unknown. Viral particles can be observed in the granular layer of the epithelium but not in the lower layers. HPV is not a lytic virus, release of viral particles thereby does not occur by cell disruption but rather benefits from the natural course of keratinocyte differentiation when the cornified cells are shed from the epithelium. Indeed, the last steps of keratinocyte maturation involve nucleus loss and a type of apoptosis, resulting in the natural release of the uppermost cell from the epithelium referred to as desquamation. Liberation of virions might also be favored by the action of E4 on the cytokeratin network and/or by the induction of apoptosis by E2 as we will discussed later.

Overall, HPV infection takes place in a continuously renewing keratinized epithelium, an environment normally tightly controlled. HPVs both depend on and benefit from the host differentiation and proliferation regulation, and require the complex hijacking of these processes to ensure proper spreading of progeny. Subsequent development of HPV-associated cancers can be envisioned as a by-product of the infection that could arise from the delicate manipulation of the regulation of host cell cycle and differentiation program. The following section describes the critical steps that can lead to the establishment of cancer.

C - Carcinogenesis

As mentioned above, certain papillomaviruses are capable of inducing cellular transformation. In assays employing primary rodent cells and primary human fibroblasts and keratinocyte cultures, high-risk HPVs (like HPV16 or HPV18) can induce transformation or immortalization, whereas low-risk HPVs (HPV6 or HPV11) cannot [151, 152]. In High grade dysplasia, only a restricted number of genes are expressed, primarily E6 and E7. During long-term infection, these two viral genes appear to be the main drivers of the progression to cancer by orchestrating a series of pathogenic changes.

a . Proliferation

The most fundamental trait of cancer cells involves their ability to sustain uncontrolled proliferation. Normal tissues carefully control the production and release of growth-promoting signals that instruct entry into and progression through the cell cycle, thereby ensuring a homeostasis of cell number and maintenance of normal tissue architecture and function. Cancerous cells circumvent the powerful programs that negatively regulate cell proliferation, mainly through the actions of tumor suppressor genes. Two prototypical tumor suppressors encoding Rb and p53 proteins operate as central control nodes within two key complementary cellular regulatory circuits that govern cell's decisions between proliferation or cell cycle arrest or possibly cell death. The Rb protein integrates signals from diverse extracellular and intracellular sources and, in response, determines whether or not a cell should proceed through its proliferation cycle. As such, Rb is considered as a critical gatekeeper of cell cycle progression. On the other hand, p53 receives inputs from stress and abnormality sensors that function within the cell, and if the inner cell conditions are not optimal, p53 can call a halt to further cell cycle progression until these conditions have been normalized.

As discussed earlier, the E7 protein of HR-HPVs binds to the Rb protein family and target their degradation. The viral oncoprotein E7 therefore primarily functions to modulate the proliferation status of infected cells. Normal cells use the pRb family members to regulate the G1/S transition by sequestering the E2F family of transcription factors. For normal S-phase entry, pRb hyperphosphorylation is induced by the cell cycle regulators cyclin-dependent kinases (Cdks). This leads to the release of E2F and subsequent transactivation of its target genes essential for the expression of many proteins functionally involved in S-phase, such as those needed for DNA replication. In HPV-infected cells however, E7's binding to Rb induces its destabilization thereby resulting in continuous E2F activation and to an unregulated S-phase entry.

Beyond its role in Rb's degradation, E7 was shown to bind and activate Cdk2, a cyclin-dependent kinase that regulates S-phase entry independently of Rb [153]. E7 was also demonstrated to counteract the effect of the Cdk inhibitors p21 and p27 by directly binding to them and inhibiting their functions [145, 154, 155]. Since p21 and p27 are target genes of p53, their expression is additionally reduced by the E6-mediated degradation of p53. Furthermore, E7 binds and inhibits histone deacetylases, whose activity normally restrain S-phase progression [156]. Indeed, the removal of acetyl groups on histones allows a tighter DNA wrapping and therefore impedes DNA replication.

In addition to E7, several groups have also implicated the E5 protein as an activator of cellular proliferation. Indeed, E5 was shown to interfere with the degradation and/or trafficking of the epidermal growth factor receptor, which leads to the sustained activation of epidermal growth factor signaling [157]. Since E5 expression is commonly lost after integration, its role may be important primarily in the early stages of carcinogenesis.

The unlimited proliferation capacity of HPV-infected cells are the result of E6. Indeed, normal proliferating cells can divide only a limited number of times, since their lifespan is restricted by telomere shortening occurring at each cell division. This erosion of telomeres ultimately leads to cell death by senescence and is thought to be an intrinsic cellular mechanism preventing unlimited cell growth. During carcinogenic conversion, the activation of the telomerase enzymatic component hTERT usually occurs to overcome this barrier, therefore resulting in the prolongation of the cell's lifespan and promoting immortalization. Activation of hTERT has been observed in HPV16 E6-expressing cells and is thought to be, at least in part, due to an increase in hTERT transcription induced by the cooperative binding of E6 and Myc at the hTERT promoter [159]. E6 activation of hTERT was also reported to depend on its association with the ubiquitin ligase E6-AP,

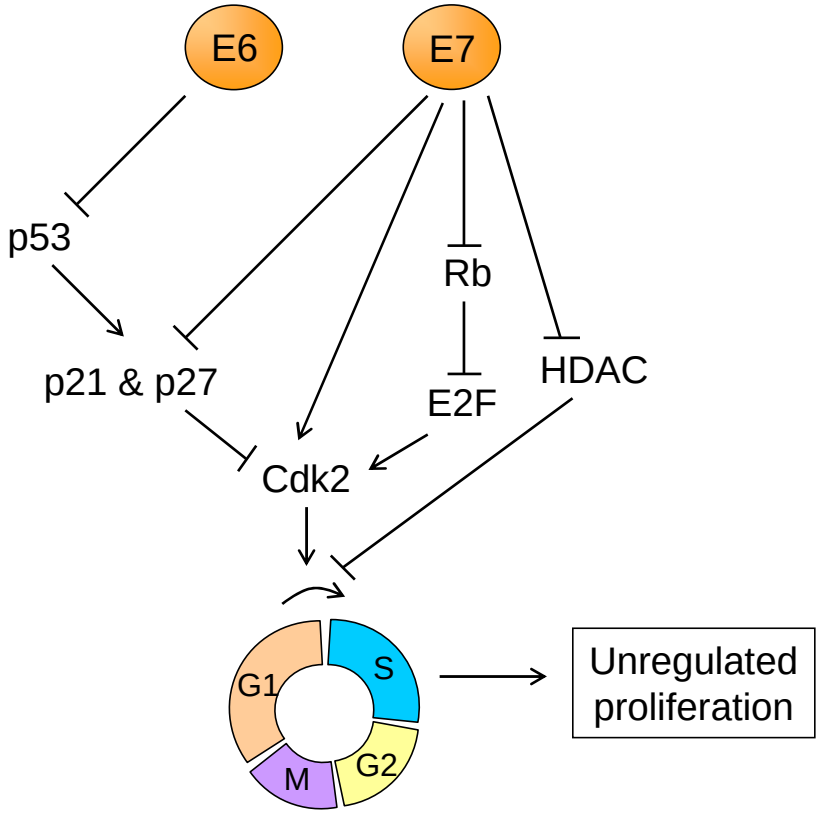

Figure I.7: **Molecular mechanisms of HPV-induced cellular proliferation.** Adapted from *Lehoux et al.* [158].

possibly implicating the degradation of an hTERT regulator [160].

The two oncoproteins E6 and E7 therefore cooperate to promote the deregulated and unlimited proliferation of HPV-infected cells, key events in the progression toward cancer development (**Fig I.7**).

b . Genome integration

Although the HPV genome is typically found in an episomal form in cervical lesions, viral integration has been reported to be associated with oncogenesis and is also correlated with more severe lesions [161, 162]. Viral DNA integration occurs at many sites throughout the host genome but it was shown to preferentially happen at fragile sites [163, 164], and in a given lesion, it usually involves only one locus [165]. It was shown that when more than one integration sites exist, expression of viral genes from a single site predominates by methylation of the others [166].

Integration is not a normal step of the HPV life cycle, but rather a random event that may confer a selective advantage to the host cell without any apparent advantage for the virus. Typically, viral DNA integration is associated with the loss of variably large fragments of the viral genome, and is characterized by the deletion of genes essential for the synthesis of an infectious virus, particularly the loss of the viral E2 gene (**Fig I.8**). In this integrated form, E6 and E7 remain intact and they are transcribed from the LCR.

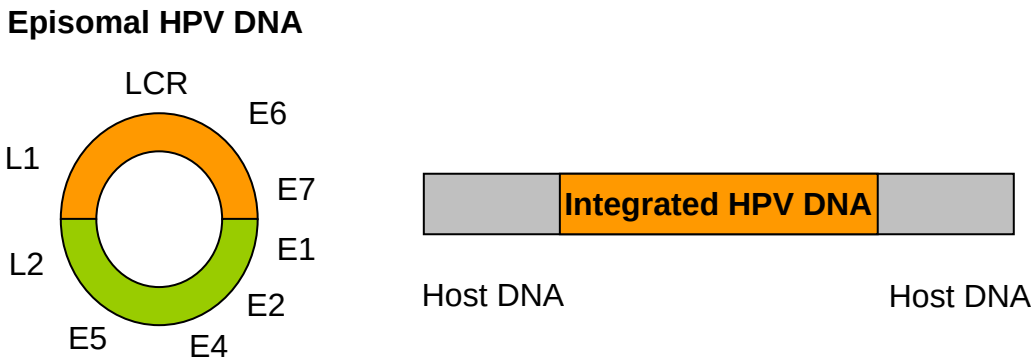

Figure I.8: **Integration of the HPV DNA**. In low grade lesions, the viral genome is maintained as an episome. Viral DNA integration into the host chromosome results in the loss of a long fragment of the HPV genome. Adapted from 5th edition of the book *Virology* [167].

Integration of HPV DNA represents a pivotal change that appears to stabilize E6 and E7 genes in a deregulated overexpression state. This is possibly due to the specific context of the host chromatin [168] and the concomitant loss of E2, which is the main repressor of their transcription as will be discussed later. Such E6 and E7 overexpression triggers cellular immortalization, deregulates proliferation, and enhances genomic instability, all cellular hallmarks that can contribute to development of the malignant phenotype [169].

The mechanism of HPV integration is not fully understood. Unlike retroviruses, HPVs

do not encode an integrase, and integration plays no role in the normal life cycle. Integration occurs *via* non-homologous recombination and must represent a chance occurrence, presumably favored by the generation of linear double strand replication intermediates that stimulate the repair machinery in the host. A number of studies have suggested that insertional mutagenesis may have a role in at least some cervical cancers. Cases have been reported in which viral genome integration has occurred within or adjacent to known oncogenes, most commonly within intronic sequences [164]. The most frequently observed integration sites, particularly in cervical cancers positive for HPV18, are in the region of the *MYC* gene [170, 171] and at the TERT locus [172]. However, very few studies have determined whether integration in these regions has an effect on expression of the candidate host oncogenes and consequently on the cell phenotype. In addition, given that transcription of the telomerase (hTERT) gene is inhibited by E2 [173] and activated by E6 [174], viral genome integration could result in telomerase activation and further promote immortalization of epithelial cells.

Overall, while HPV integration is a key event in the development of cervical cancers and is observed in about half of the HPV-associated cancer cells, it is a late event in the carcinogenic process.

c .Malignant transformation

Although infection with high-risk HPVs are necessary for the development of cancers attributable to HPV infection, it is clearly not sufficient. Cancers arise only after other factors have collaborated with infection. As discussed above, some changes are virus specific (like viral genome integration), but most changes associated with cancer progression actually result from genetic alterations of cellular genes or modifications of the host environment. They usually result from the long-term viral gene expression, include downregulation of tumor suppressor and proapoptotic genes, or upregulation of proto-oncogenes or antiapoptotic genes [175]. Cancerous lesions must accumulate an increasing number of mutations as they progress toward malignancy and invasion. Therefore, the identification of recurrent chromosomal alterations is of great importance for the understanding of the biology of these cancers. Based on the multigenic nature of cancer, multiple genetic alterations are likely to be found in a single tumor. Loss of heterozygosity is often observed in cervical cancerous lesions, generally thought of as an intermediate step in the inactivation of tumor suppressor genes, which requires the inactivation of both alleles in order to display the tumoral phenotype. There have also been several studies attempting to detect mutations in genes well known to have mutations in other tumors, including H-RAS, TP53, pRB, showing that HPV-associated cancerous lesions also display a pattern of point mutations similar to other cancers [176]. In addition, HPV-induced activation of telomerase ensures an unlimited number of cell divisions, therefore cells are more prone to accumulate mutations.

The development of solid tumors is a multistep process, which requires the combination of several key events. One critical feature that must be acquired by transformed cells is the ability to recruit a blood supply from existing vasculature, a process known as angiogenesis and associated with tumor invasion (**Fig I.9**). Angiogenesis involves the activation, migration, and proliferation of endothelial cells. This process enables a devel-

oping tumor to be fueled with nutrients and oxygen that sustain growth and expansion. Regulation of angiogenesis is tightly controlled [177]. The observation that tumors could be implanted into an avascular region suggested that tumors released diffusible activators of angiogenesis that could signal a quiescent vascular to begin capillary growth.

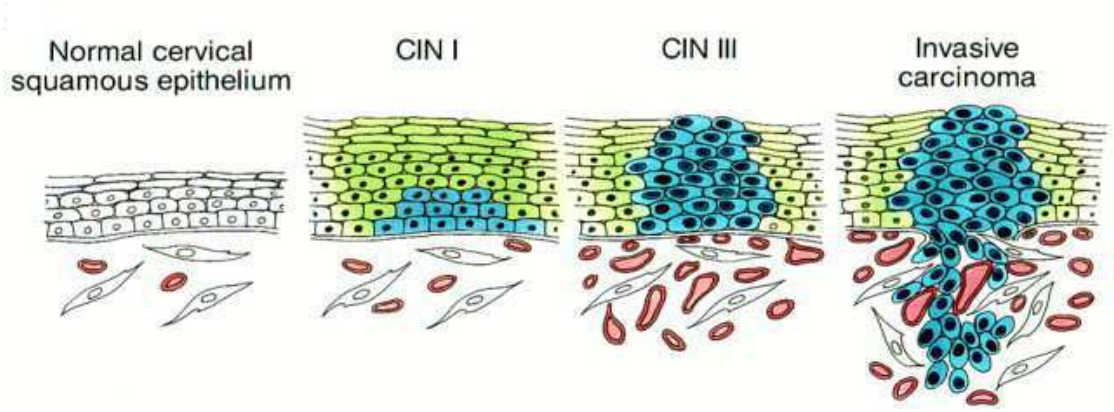

Figure I.9: **Cervical carcinoma invasion**. A modest increase in new vessel density is evident in CIN I lesions, while CIN III lesions show abundant new vessels, indicative of the angiogenic switch from vascular quiescence to sustained neovascularization. Red circles represent blood vessels and long cells represent fibroblasts. From Hanahan & Folkman [178].

There are two main classes of angiogenesis inducers: FGF (fibroblast growth factor) and VEGF (vascular endothelial growth factor). On the other hand, angiogenesis inhibitors, like interferon α and platelet factor 4, were shown to counteract inducing signals for new capillary growth. The angiogenic switch is therefore the result of the balance between inhibitors and inducers [178].

Previous studies have demonstrated that angiogenesis is an early event in cervical cancer that becomes apparent early in the carcinogenic conversion process. In both a transgenic mouse model expressing HPV16 early genes and in human cervical lesions ranging from low-grade dysplasia to invasive carcinoma, data indicate that angiogenesis occurs in a stepwise manner. Increasing vascular density and increasing expression of VEGF correlate with progression toward tumor development [175]. Expression of HPV16 E6 and E7 in primary foreskin keratinocytes is sufficient to alter the profile of expression of angiogenic factors [179, 180]. A correlation has been made between upregulation of HPV E6 and E7 gene expression and upregulation of VEGF expression, increase in microvascular density, and close apposition of neovasculature to the affected tissue.

In situ carcinoma arising from epithelial tissues progress to higher pathological grades of malignancy with local invasion followed by apparition of distant metastases (**Fig I.9**). Tumorigenic cancer cells are then typically associated with changes in their shapes and detachment from neighboring cells and from the underlying extracellular matrix. Invasion and metastasis are envisioned as a succession of biologic changes, beginning with local invasion, then intravasation into nearby blood and lymphatic vessels. Cells then transit through the lymphatic systems, and escape from the vessel lumina into the parenchyma

of distant tissues (extravasation). This process ends with the formation of small nodules of cancer cells (micrometastases), and finally the growth of micrometastatic lesions into macroscopic tumors, this last step being termed "colonization"[181].

Clinical reports on patients with HPV-positive tumors with distant metastasis in organs such as lung, liver and bone suggested an association between the presence of oncogenic human papillomaviral subtypes and the metastatic process. Several lines of evidence support a potential role of E6 and E7 in metastasis spreading. Stable expression of HPV16 E6 and E7 genes was shown to increase the pro-metastatic conversion of cell lines [182], and enhance migration of normal human keratinocytes [183]. E6 and E7 expression was also demonstrated to strengthen tumor cell capability to home into distant metastatic sites in conjunction with microenvironmental stimuli, including pro-migratory chemokine induction [184], which plays critical roles in determining organ-selective metastasis.

Malignant transformation thus appears as a complex process that requires both virus-independent and virus-induced modifications of the host cell. Given the aggressiveness of cervical cancers, metastases outbreak occurs rapidly during cancer progression. Cervical cancer is therefore very difficult to detect in time for efficient local treatment.

d .Immune evasion

Evidence suggests that cervical cancers have undergone changes enabling immune system evasion [185, 186]. HPVs must either avoid or negotiate the powerful host immune defense system. Host defense is a complex interplay between innate immunity (phagocytes, cytokines, complement...) and the adaptive immunity (antibody, effector cells...). The innate immunity is the first line of defense that detects pathogens and is believed to be able to clear up to 90% of microbial assaults. Although the innate immunity has no specific memory, it triggers the appropriate adaptive immune response using dendritic cells to activate naive T cells in the draining lymph node, which generates both highly specific lethal effector responses and long lasting memory cells. Adaptive response can therefore clear the host of viral infections and prevent re-infection.

Although a large fraction of the sexually active population has been infected with HPVs, progression from infection to cancer is a very rare event and prevents further completion of the viral life cycle. Thus, it is important to distinguish between the mechanisms used by the virus and those used by tumor cells to evade the immune attack.

The exclusively intra-epithelial life cycle of HPVs is central to understand the host response, since it has some key features that impact on the recognition and response of the host immune system to papillomaviruses. HPVs are not lytic viruses. Their life cycle is played out along the keratinocyte differentiation program, therefore no inflammation signals accompany viral infection and alert the innate immune sensors. In addition, keratinocytes are destined for death from natural causes. Virion release therefore occurs in the natural context of cell shedding further limiting inflammation. There are no or very few viruses in the bloodstream since the virus infects the host via microabrasions that leave the basal lamina intact, and is shed from mucosal or cutaneous surfaces far from vascular channels. Thus, there is poor access of the viral antigens to the draining lymph nodes where adaptive immune responses are initiated. In addition, HPVs encode only non-secreted proteins, expressed at low levels, reinforcing the difficulty for the host

immune system to detect the presence of an infection.

Despite these characteristics, an immune response is triggered during HPV infection. Indeed, spontaneous regression occurs for both cutaneous and ano-genital warts, while non-regressing lesions are characterized by a lack of immune cells. Studies of the natural infection history show that genital HPV infection is extremely common in sexually active young women. Most of these HPV infections "clear" *s*pontaneously, i.e. HPV DNA can no longer be detected in cervical tissue. The time taken to complete clearance varies greatly and is usually longer for the HR-HPV [13, 187]. However, if the immune response fails to clear or control the infection, then a persistent infection is established and generates an increased probability of progression to high grade cervical intra-epithelial neoplasia and invasive carcinoma.

The increased incidence of HPV-associated lesions in immunosuppressed individuals illustrates the critical importance of cell-mediated immune response in the resolution of HPV infections. Cell-mediated cytotoxicity is the most important effector mechanism for the control and clearance of viral infections and is implemented both by antigen-specific cytotoxic T cells and the so-called Natural Killers. HPV-specific cytotoxic T cells can be detected in patients with previous or ongoing HPV infection [188, 189]. The natural killer cells, on the other hand, are a subset of lymphocytes that kill virally infected or tumor cells lacking surface expression of MHC Class I molecules and there is evidence that they are important in HPV infections [190]. Why HPV infection remains ignored or undetected by the immune system during viral persistence is a central question. Indeed, HPV intrusion is detected by the professional Antigen Presenting Cells (APC) of the epithelia, the Langerhans cell (LC). The activated LC then migrates to the draining lymph node, process and present HPV antigens to naive T cells that then differentiate into activated effector cells and migrate back to the infected site to destroy the infected keratinocytes. However, during viral persistence, only a limited range of viral factors are expressed, only those required for maintaining the viral DNA. That results in the virus being practically invisible to the host, a viral strategy leading to persistent chronic infection. In addition, HPV-infected keratinocytes should induce a type 1 interferon responses - a powerful, generic, anti-viral, defense system. The type 1 interferons, IFN-α and IFN-β, have antiviral, antiproliferative, anti-angiogenic and immunostimulatory properties, acting as a bridge between innate and adaptive immunity by activating immature dendritic cells. Most DNA viruses have developed mechanisms for inhibiting interferon synthesis and signaling and in the case of papillomaviruses, it mostly relies on E6 and E7 (see [191] for a review).

Numerous serological studies using HPV virus-like particles (VLPs) have shown that infection with a genital HPV is followed by sero-conversion and type-specific antibodies to the major viral capsid protein L1. However, antibody concentrations are low even at the time of sero-conversion, which is not surprising given the poor access of infected cells to lymph nodes where immune responses are initiated.

Although these observations indicate that the immune system profoundly influences the outcome of an HPV infection, protective immunity is incompletely understood. The exact determinants that induce the immune responses responsible for regression as well as the balance between viral gene expression and alteration of normal immune response that leads to the establishment of a persistent infection remain largely unknown.

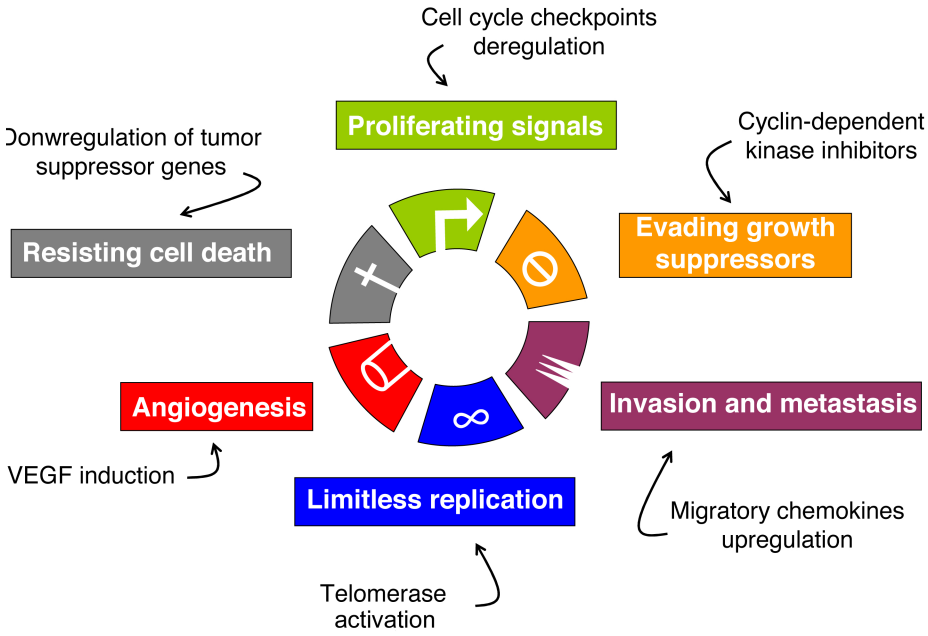

Figure I.10: **The multistep process of carcinogenic conversion.** Adapted from Hanahan & Weinberg [169].

Carcinogenesis can be seen as a Darwinian process involving sequential mutations providing the mutated cells with growth dominance over the normal neighboring cells, resulting in the increased representation of the mutated cells in the affected tissue. HPV carcinogenic conversion follows a cascade of events at the crossroad between cellular machinery hijacking to promote unlimited cell division, proliferation and migration, and hiding from the immune system to ensure a long-term persistence in the host cells (**Fig I.10**). Carcinogenic conversion is consistently associated with the loss of control over E6 and E7, which critically depends on the alteration of the normal function of a pivotal regulator of the HPV life cycle: the E2 protein.

D - The E2 proteins

E2 is a key protein in the HPV life cycle both for the productive cycle and persistent infection [192]. E2 is expressed from the early stages of infection and has been shown to participate in transcriptional regulation, viral DNA replication and mitotic segregation, all pivotal functions for the viral life cycle. Besides these historical roles, which rely on its aptitude to bind to specific DNA sequences in the viral genome, E2 has also been shown to modulate the host cells through direct protein-protein interactions independently of its binding to the viral DNA. In this section, E2's functions are extensively reviewed as well as their implications in the context of an HPV infection.

a .Structural properties

The E2 proteins are composed of 350 to 500 amino acids consisting of three distinct domains [193]: two conserved modular domains, a N-terminal domain of about 200 amino acids referred to as TAD (TransActivation Domain), and a 100 amino acid C-terminal DNA-Binding Domain (DBD), separated by an unstructured Hinge region, not conserved and of variable length. The crystal structure of the two conserved domains has been determined for several genital α-HPV E2 proteins [194–196], but no information is available concerning the structure of E2 from the cutaneous HPVs (**Fig I.11**). However, based on the high sequence homology, one can extrapolate that the E2 proteins from the cutaneous HPVs exhibit similar 3D structures as the α-HPV E2 proteins.

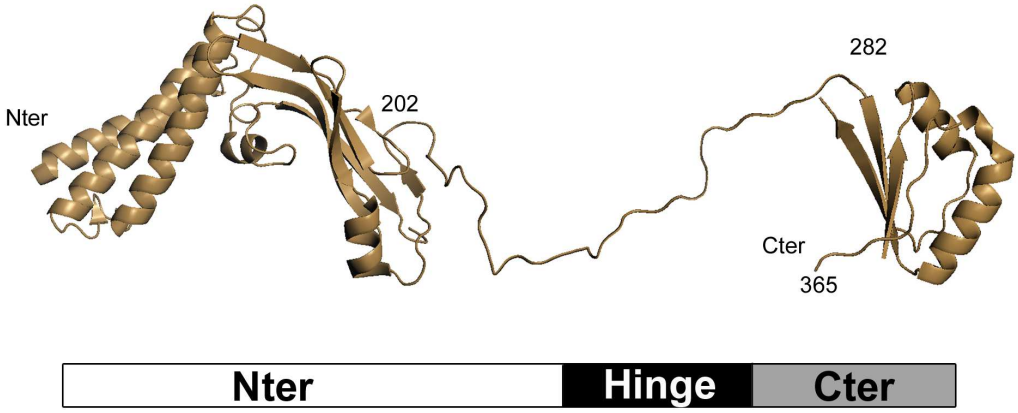

Figure I.11: **HPV16 E2 structure**. The N- and C-Terminal domains of E2 from HPV16 were obtained by crystallography and are represented here separated by a schematized hinge.

Transactivation domain

The transactivation domain of E2 is critical for E2's function in DNA replication, transcriptional regulation, apoptotic induction and a number of other functions.

It appears to be extremely sensitive to mutations, which often induce disruption of its conformation. The transactivation domain contains many residues that are invariant among all E2 sequences and are pivotal for its folding. The crystal structure was found to consist of a cashew-shaped domain made up of two regions [195]. The N-terminal half of the transactivation domain contains three antiparallel α-helices and the C-terminal half is almost entirely constituted by antiparallel β-sheet structures. The residues between these regions form a fulcrum that orientates the domain and closely groups each part against the others thereby forming a rather rigid structure. In addition, the E2 N-terminal domain of HPV16 was shown to be able to form stable dimers in solution [195], but this characteristic has not been observed with other HPV E2 proteins [197].

DNA-binding domain

The DNA-binding domain of E2 exhibits a typical three-dimensional structure, which has homology to other viral factors and is considered prototypic of some viral regulation factors [198]. This domain is composed of α-helices and β-sheets and contains both a strong dimerization interface and amino acids making contacts with DNA [194]. Although DNA binding is not required for the dimerization of E2's C-terminal domain, the binding to DNA sequences is only possible when E2 is dimerized [199, 200]. Dimerization of the C-terminal domain results in the formation of a β barrel by the association of the β-sheets of the two E2 proteins in dimer. The α-helices are located outside the barrel and interact with the major DNA groove. This is through this C-terminal domain that E2 interacts with its recognition sites in the viral genome to regulate viral genome transcription, replication and segregation.

Hinge

The hinge region of the E2 protein is a segment of variable length localized between the N- and C-terminal domains. It was predicted to adopt a random coil structure that could confer flexibility to E2 [201]. In contrast with the two well defined terminal domains, the function of the hinge region remains unclear. The hinge region sequence is not conserved but contain numerous binding motifs. Given that this domain is unfolded, these motifs are exposed and therefore extensively contribute to the interactions with the host cell proteins. The hinge domain was also shown to contain a Nuclear Localization Signal (NLS) which dictates E2's cellular localization [202].

b . Regulatory roles

E2 is considered as a central regulatory HPV protein, with roles spanning from regulation of viral gene expression to long-term maintenance of viral episomes. We can distinguish between two main ways for E2-mediated regulation of infection: E2 binding to E2-responsive elements in DNA sequences or autonomous activities independent of its binding to DNA sequences.

1 . Viral genome binding-associated activities

E2, through its C-terminal domain, binds as a dimer to its E2 binding sites, which are specific palindromic DNA sequences ACCG(N_4)CGGT present within the regulatory region (LCR) of HPVs. By binding to these sites, E2 recruits at the LCR a number of cellular proteins necessary to support transcription, replication and mitotic segregation of the viral genome. The stability of E2/DNA complexes differs according to the E2BS flanking sequences [203–205], and this has been proposed to modulate the effect of E2 on early promoter transcription [206].

Transcription

E2 is a transcriptional regulator and behaves as a transcriptional repressor in the context of the natural genital HPV E6/E7 promoters. Repression depends both on the interaction between E2 and its target DNA sequence and on the relative position of E2BS along the regulatory region. Within the genital HPV group, two E2BS overlap sequences of the early promoter (**Fig I.12**). By binding to these sites, E2 interferes with the binding of the cellular transcription factors TBP and SP1 by steric hindrance [203, 207, 208] which leads to repression of the early promoter transcription.

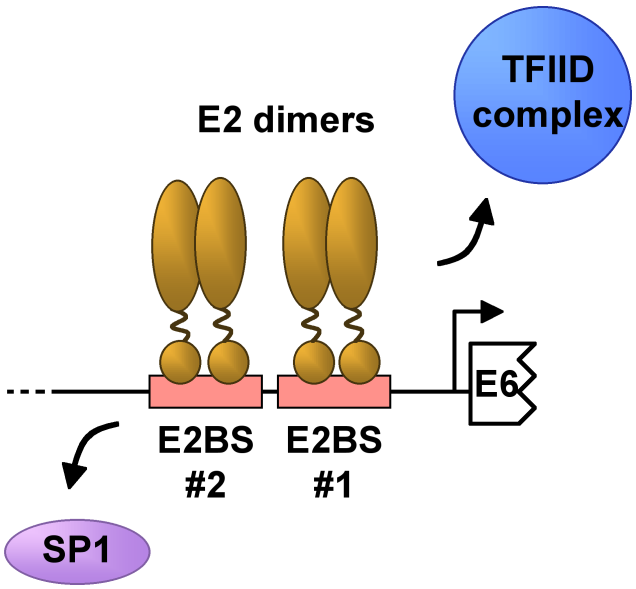

Figure I.12: **E2-mediated repression of the HPV early promoter**. Schematic representation of the molecular consequences of E2 binding to its binding sites most proximal to the early initiation start. Adapted from Desaintes & Demeret [35].

The repression of the early promoter is not only due to steric hindrance, but is also the outcome of E2 binding to cellular proteins involved in the recruitment of repressive complexes. For example, interaction of E2 with transcriptional regulatory proteins as the BRD4 protein, the histone demethylase SMCX, and the EP400 component of histone acetyltransferase complex have been shown to be required for efficient repression of the

HPV18 LCR [209]. In the LCR of cutaneous HPVs, the E2BS are located farther from the early transcription initiation start, and such a configuration leads to transcriptional activation by E2 [41].

Little is known about the regulation of the HPV late promoter. This promoter is localized within the E7 ORF [210, 211] and is only active during keratinocyte differentiation, which has impeded the study of its regulation in current experimental settings with cell lines. The late promoter is known to drive the transcription of the late genes L1 and L2. At this point, the role of E2 in regulating late gene expression is not clear but it is assumed that E2 activates the transcription of late genes from this promoter.

The transcriptional activation capacities of E2 were mostly studied using synthetic constructs containing multimerized E2BS upstream of a minimal promoter, or in *in vitro* transcription assays. It was determined that the HPV E2 proteins can activate promoters over large distances [212], and exhibit different transcriptional activation properties. The high risk E2 proteins were found to be intrinsically more potent transcriptional activators than the low risk E2 proteins [213]. These differences might result from variable interplay with the host transcriptional machinery. Indeed, interactions detected with E2 over a wide series of studies identified a large spectrum of transcriptional targets, ranging from basic factors to chromatin-related factors, suggesting that the E2 proteins are able to interfere at multiple levels of transcriptional processes as detailed below (reviewed in [192]).

The cellular transcription machinery is a complex assembly between basal and regulatory sequence-specific transcription factors. TFIID is a pivotal component of the cellular transcription initiation complex, and is composed of TBP and TBP-associated factors (TAF) [214]. Regulatory transcription factors bind promoter regions and allow the efficient recruitment of the transcription pre-initiation complex through co-activators, which mediates protein-protein interactions between regulatory transcription factors and the basal transcription machinery. E2 is known to interact with both basal (TAF, TBP, GTF2B [215–217]) or regulatory (SP1, HOXC9, NR4A1, C/EBP [218–221]) transcription factors, as well as with co-activators (TMF, [219]), indicating that its transcriptional properties rely, at least partly, on the modulation of the pre-initiation complex formation.

E2 is also able to act at the chromatin level to regulate transcription. Indeed, the best characterized partner necessary for E2-dependent transcription is BRD4, a member of the BET family of double bromodomain proteins that binds to acetylated tails of histones H3 and H4 and stimulates RNA polymerase II-dependent transcription by recruiting distinct transcriptional regulators [222]. Mutation of the amino acids responsible for the interaction with BRD4 (notably I73 and R37 for HPV16 E2) strongly affects the transcriptional capacities of E2, making BRD4 an essential mediator of E2's transcriptional properties [223, 224]. In addition, the E2 proteins are able to modulate transcription through modification of the histone code, by targeting histone-modification factors. In particular, several histone acetyltransferases PCAF, EP300, EP400 or CBP were shown to interact with E2 [225–228]. These factors function within macromolecular complexes, recruited at their target promoters by interacting with sequence-specific transcription factors, to loosen chromatin structure.

Lastly, the HPV E2 proteins interact with members of chromatin remodeling complexes involved in the deposition or displacement of nucleosomes. Two members of chromatin remodeling complexes are known E2 targets, NAP1L1 and hSNF5 [228, 229].

NAP1L1 is primarily involved in replication-coupled nucleosome assembly by mediating the incorporation of histones H2A-H2B dimers in nucleosomes. Its direct binding to E2 has been shown to enhance E2's transcriptional activation capacities, though this interaction might primarily impact on the replication activating functions of E2. On the other hand, the SWI/SNF chromatin remodeling complex was recently shown to enhance HPV18 E2-dependent transcription through direct binding of the core component hSNF5 to E2 [229], which might counteract the repressive effect of chromatin formation.

Therefore, E2 turns out to regulate viral transcription through a broad range of processes, and that this regulation might depend both on the cell context and on the HPV type. E2 is also known to have an impact on the transcriptional regulation of a number of cellular genes through interaction with various host transcription factors [218, 230–233]. This is believed to contribute to create permissive cell conditions for the successive stages of the viral life cycle.

Replication

By binding to the viral genome, E2 also activates the replication of viral DNA [234]. Indeed, the origin of replication lies within the promoter-proximal region of the LCR and contains E2BS as well as a binding site for the viral helicase E1 (reviewed in [35]). E2 both binds to E1 and to its cognate sites at high affinity, consequently targeting E1 at the origin of replication (**Fig I.13**).

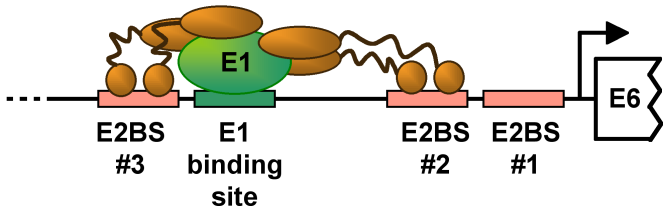

Figure I.13: **E2 helps the loading of the viral helicase E1 on the replication origin**. Schematic representation of the loading of E1 by E2 on the viral replication origin in the HPV LCR. Adapted from Desaintes & Demeret [35].

E2 also helps the loading of E1 by masking its non-specific binding activity [235] thereby increasing the specificity of E1 DNA binding. E2 additionally enhances replication by alleviating nucleosomes-mediated repression which allows efficient binding of the E1/E2 complex to the replication origin [236]. After loading, E2 is dissociated from the complex and E1 converts into a double hexameric ring helicase that encircles DNA [237, 238]. Lastly, E2 activates viral DNA replication by recruiting host cell replication factors like TOP1 or RPA through direct interaction [53, 239].

Mitotic segregation

The active mechanism of viral genome partitioning is not required when there are sufficient copies of the viral genome, since each daughter cell is likely to randomly contain

at least one genome. However, in the case of HPVs, only low copy number of the viral genome are present in the basal cells, and therefore a specific mechanism to ensure its segregation is needed. It was demonstrated that E2 is required for long-term viral episome maintenance within dividing cells [240] by mediating the tethering of viral genomes to mitotic chromosomes (**Fig I.14**). This ensures that the viral genomes are partitioned to daughter cells in approximately equal numbers, and guarantees that they are retained in the nucleus after cell division. The strategy of tethering viral genomes to host mitotic chromosomes using a viral DNA binding protein is common to many persistent DNA viruses, like in HHV-8 with the LANA protein or in EBV with EBNA-1 [241].

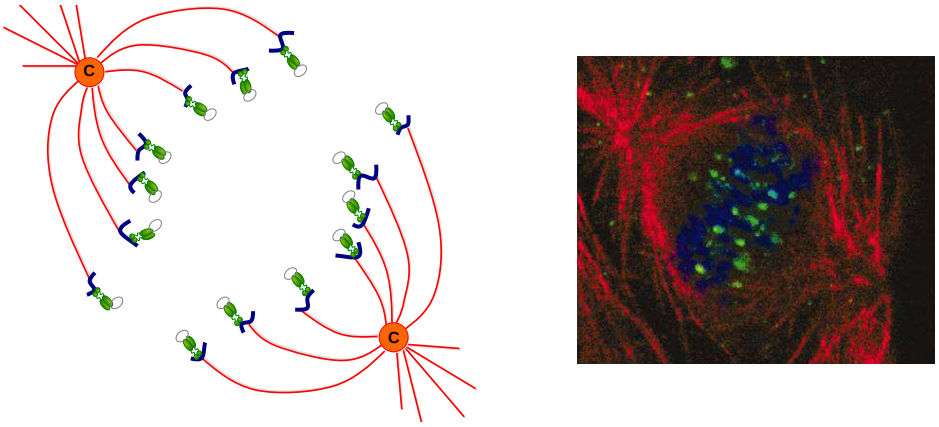

Figure I.14: **Mechanism of HPV genome partitioning**. On the left: schematic representation of the mechanism involved in segregation of viral episome. On the right immunofluorescence showing the pattern of genome partitioning. Microtubules are represented in red, Host DNA in blue, E2 dimers in green. HPV episomes are schematized as gray circle. The C letter stand for the microtubule Centrosome. Adapted from Oliveira et al. and Chapter 4 of the book Advances in Virus Research, 2008. [149, 242].

The first clue to the precise role of E2 in genome partitioning came from the observation that both the BPV-1 E2 protein and the viral genomes were observed as small speckles over the arms of all mitotic chromosomes [243]. It was later envisioned that E2 binds mitotic chromosomes through protein-protein interactions mediated by E2's transactivation domain. The DNA binding domain binds to multiple E2 binding sites in the viral genome and tethers it to the condensed chromosomes [240, 244]. Disruption of the E2 mediated interaction enabling the HPV genome tethering could have great therapeutic potential and so an important goal has been to identify and characterize the cellular chromosomal targets that mediate viral genome segregation.

The best characterized chromosomal target of E2, to date, is the cellular protein, Brd4 [245, 246]. Brd4 was shown to remain bound to mitotic chromatin [247]. Analysis of mutated E2 proteins showed that the mitotic chromosome binding activity of E2 correlates with Brd4 interaction [248]. However, while Brd4 is clearly an important component of the tethering complex for several papillomaviruses, it is not clear how it links E2/HPV genome to the cellular chromosomes. Indeed, Brd4 is not a stable tether, interacting only

with histones that are acetylated [247] but most of them become deacetylated during mitosis [249]. It was hypothesized that the interaction of Brd4 with E2 actively stabilizes Brd4 interaction with chromatin [246], suggesting that E2-HPV episome complexes do not passively hitchhike on mitotic chromosomes but instead strengthen the Brd4-chromatin interaction to ensure the stable transmission of the viral genome. In addition, it appears that Brd4 is not the only cellular factor relevant for viral segregation [250]. Other cellular proteins involved in papillomavirus genome partitioning were identified over the years, as ChLR1, a DNA helicase involved in sister chromatid cohesion [251] or Mlkp2, a microtubule motor kinesin [252].

E2 is thus essential both to partition viral episomes in basal and suprabasal proliferating cells and to maintain a reservoir of HPV genomes during long-term infections, and is thus a crucial determinant of viral persistence, a major risk factor for cancer development.

Splicing

In comparison with several well studied E2 gene products, the E2 proteins of the cutaneous HPVs are characterized by a relatively long hinge region rich in arginine, serine, and glycine residues. Arginine-serine (RS) dipeptide repeats are typical of the SR superfamily of proteins, which are primarily involved in the splicing of precursor mRNA [253–255]. SR proteins are crucial splicing factors and also direct alternative splicing by modulating splice site choice. RS domains have been shown to mediate protein-protein interaction between SR proteins [256].

Given the presence of RS-rich regions in their hinge, it was hypothesized that cutaneous E2 proteins may exert a function similar to that of cellular SR proteins in RNA splicing or interact with SR proteins. As predicted, the β-type HPV E2 proteins were shown to interact with a number of SR proteins as SFRS1 (best known as ASF/SF2) an essential sequence-specific splicing factor, as well as with SFRS2 (SC35) and SFRS7 (9G8) [257, 258].

In addition, functional evidence indicated that the HPV5 E2 protein can facilitate the splicing of transcripts synthesized from promoters activated by E2 itself, probably in a distance-dependent manner [257]. A plausible mechanism is that the RS-rich hinge of E2 recruits essential splicing factors, increasing their concentration around the HPV early promoter, and thereby making them available for the splicing of locally produced viral transcripts (**Fig I.15**).

The mucosal α-type HPV E2 proteins appear to modulate mRNA splicing as well, yet in another way since HPV16 E2 was shown to activate the expression of splicing factors such as ASF/SF2 [231]. This function seems critical given that viral genomes expressing an E2 protein defective for SFRS1 upregulation show reduce levels of viral RNA (reviewed in [259]). In addition, HPV16 E2 was recently shown to inhibit the polyadenylation of late viral mRNA [260]; further substantiating the participation of E2 to viral mRNA processing.

The E2 proteins therefore emerge as primary factors to orchestrate viral gene expression, both at the level of transcription and of mRNA splicing. Through its binding to the viral genome, the E2 protein ensures the coordinated regulation of viral DNA replication and transcription, regulate mRNA splicing and guarantees the segregation of the viral

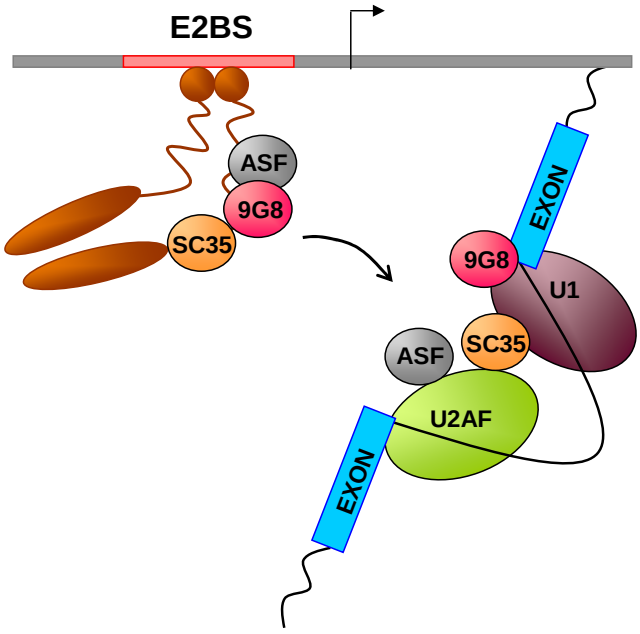

Figure I.15: **Model for the mechanistic action of E2 in assisting pre-mRNA splicing**. Schematic representation of E2 interacting with SR proteins increasing their concentration near the promoter. SR proteins are required for efficient splicing of introns. Adapted from Lai *et al.* [257].

episome throughout mitosis. All these activities make the E2 protein a pivotal factor for both productive viral life cycle and persistent infections.

2 .Host cell manipulation

For the last decade or so, several groups have demonstrated a number of additional E2 functions that are independent of its binding to the viral genome. These "autonomous" activities directly impact on the host cells and are likely required to provide appropriate cell conditions for the implementation of the productive viral life cycle within stratified epithelia. These functions are described in the following paragraphs.

Apoptosis

The pro-apoptotic activity of E2 is one of the first E2 functions described to be independent of its binding to the viral genome (reviewed in [261]). A striking aspect of this activity is that it was demonstrated to be specific for the HR-HPV E2 proteins [262, 263] first supporting the idea that E2 proteins might have developed specific activities that correlate with the HPV oncogenic power.

Initial demonstration of an induction of apoptosis by E2 in the absence of other vi-

ral sequences came from the observation that human foreskin keratinocytes infected by an adenovirus expressing HPV31 E2 underwent cell death [264] concomitantly with the demonstration that E2 could induce apoptosis in HeLa cells [265]. The role of E2 was later confirmed by Webster and colleagues [266]. Analysis of the different E2 domains showed that the amino-terminal domain alone was sufficient to induce apoptosis but independently of E2's transcriptional activity [267]. The two main pro-apoptotic pathways are dependent on caspases activation. On one hand, the intrinsic pathway can be induced by p53 and involves mitochondrial dysfunctions leading to the activation of caspase 9. On the other hand, the extrinsic pathway is dependent on death receptors signaling and activation of caspase 8. Both pathways eventually merge in the activation of effector caspases, which leads to subsequent cell death. Several interactions between HR-HPV E2 and effectors of both the extrinsic and the intrinsic pathways have been identified over the years, in particular with p53, CASP8 and C-FLIP [266, 268, 269] leading to the notion that regulation of cell death and survival by the HR-HPV E2 proteins is both redundant and complex (**Fig I.16**).

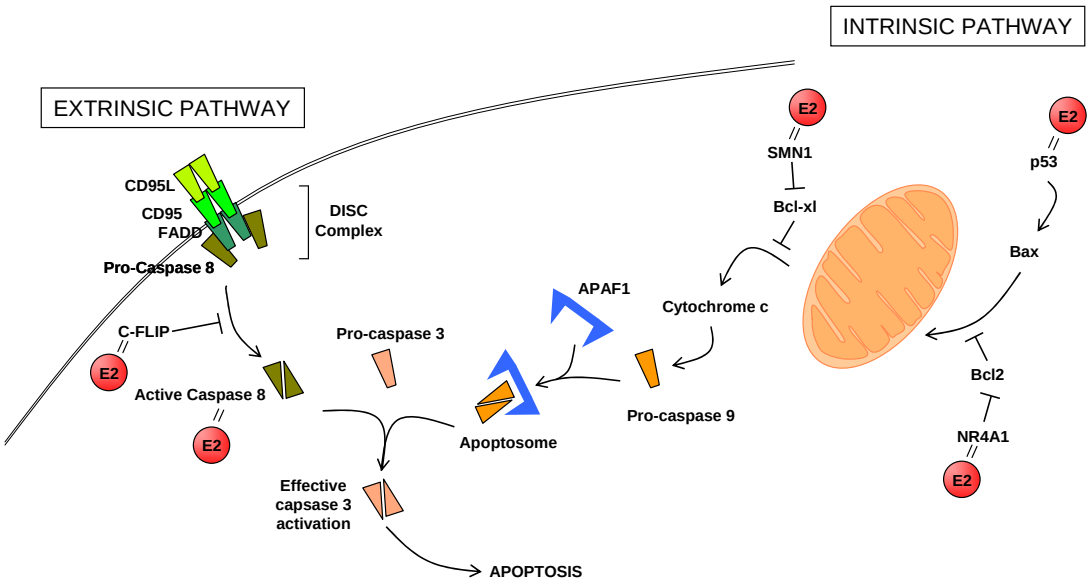

Figure I.16: **Main apoptotic pathways**. Schematic representation of E2 interacting with key factors of the intrinsic and extrinsic apoptosis pathways. Adapted from Blachon & Demeret [261].

Other known interactions may have implications for E2-induced apoptosis. For example, SMN1 was shown by Strasswimmer and colleagues to interact with E2 [270]. Recent evidence demonstrated that knock down of SMN1 increases neuronal cell death [271]. It is thus possible that the binding of SMN1 to E2 favors apoptosis. A yeast-two hybrid screen also identified NR4A1 as an interacting partner of HPV16 E2 [220]. This protein

is a member of the nuclear-receptor superfamily and plays a role in regulating cell growth and apoptosis [272]. NR4A1 triggers cytochrome c release from the mitochondria [273] both in the context of cancer cells and of virus-induced apoptosis [274, 275] by interacting with Bcl2 [276]. Therefore, by interacting with NR4A1, E2 could impact on the intrinsic apoptotic pathway. Lastly, E2 could have an impact on the last steps of apoptosis through its interaction with TOP1 [53], which has recently been shown to promote DNA fragmentation [277] characteristic of apoptotic cell death. These observations reinforce the notion that E2 uses various redundant pathways to regulate cell death and favor apoptosis completion by impacting on multiple steps of apoptosis.

Such extended regulation of apoptosis indicates that the E2 proteins actively take part to the manipulation of cell death or survival pathways, tightly orchestrated by the virus throughout infection.

Ubiquitin-Proteasome System targeting - Cell cycle

The first link between E2 and the cell cycle has been the observation of the induction of a G1 growth arrest due to repression of E6/E7 transcription and subsequent reactivation of the p53 and pRB pathways [278]. Despite this was not an intrinsic function of E2, it raised the idea that E2 could have an impact on cell cycle regulation. An E6/E7-independent role of E2 in regulating the cell cycle was later acknowledged when E2 was associated with the abrogation of mitotic checkpoints [264]. Later, Bellanger *et al.* described a G2/M arrest induced by the HR-HPV16 and 18 E2 proteins in HPV-negative cells [279]. The authors could show that E2-expressing cells were delayed in prophase and arrested in metaphase characterized by elevated levels of cyclin B/Cdk1 activity and extensive histone H3 phosphorylation. The mitotic arrest is dependent on the interaction of E2 with Cdc20 and/or Cdh1, two subunits of the Anaphase Promoting Complex/Cyclosome (APC/C), the ubiquitin-ligase responsible for metaphase-to-anaphase transition. Through these interactions, E2 inhibits the APC/C, leading to an overall stabilization of its substrates, in particular cyclin B. Lack of cyclin B degradation interferes with the normal progression through mitosis. Another APC/C substrate is Skp2 [280], a substrate recognition subunit of the ubiquitin ligase complex SCF. Stabilization by E2 of Skp2 activates the SCF complex leading to an increased degradation of the cell cycle regulators p21 and p27, and thereby an enhanced G1 to S transition [281]. Therefore it appears that interactions between the HR-HPV E2 proteins and the APC/C ubiquitin ligase complex strongly impacts on the host cell cycle. Moreover, the complex SCF^{skp2} has been shown to degrade the HPV18 E2 protein itself, suggesting a complex and controlled feedback mechanism [282].

In addition, induction of mitotic arrest leads to mitotic abnormalities and genomic instabilities due to cells encountering "mitotic slippage" or completing mitosis with abnormal cytokinesis. Over the past few years, such events have emerged as crucial precursors of carcinogenic conversion [181]. Indeed, abnormal mitoses, especially abnormal chromosome segregation following anaphase, leads to aneuploidy or DNA breaks. Given the interesting observation that only E2 proteins from high-risk HPVs could induce abnormal mitotic phenotypes [279], it was hypothesized that the high-risk E2 proteins could have a role of in premalignant stages of HPV-associated carcinogenic conversion [279].

There exist two types of E3 ubiquitin ligase complexes, the HECT (Homologous to E6AP C-terminus) proteins with an intrinsic catalytic activity, and the Cullin-based complexes composed of a cullin, a RING finger protein, and one or more substrate adapters [283–285]. Several interactions with proteins involved in ubiquitination mechanisms have also been shown to regulate E2 stability itself. The E2 proteins have short half-lives, and this rapid turnover has been shown to result from the ubiquitination of its N-terminal domain and followed by its proteasomal degradation [286–288]. Identification of the ubiquitin ligases responsible for the degradation of some E2 emerged only recently: as mentioned above, HPV18 E2 can be degraded by the ubiquitin ligase complex SCF^{skp2}, and this degradation is mediated by direct interaction with the adaptor Skp2 [282]. However, for E2 of HPV16, it was demonstrated that a cullin 3-based complex (BRC3) mediates ubiquitination and subsequent degradation [288]. It was hypothesized that 16E2 is recruited to this complex through BTB domain proteins, which are substrate adaptors of cullin 3-based complexes [288, 289]. In line with this hypothesis, two BTB domain-containing proteins, BTBD1 and BTBD2, were identified as partners of HPV16E2 protein in a yeast two-hybrid screen [219]. The process of E2 protein degradation is even more complex, since several cellular proteins have been shown to modulate E2's stability as TAX1BP1 or BRD4 [287, 290, 291]. Moreover, since HPVs infect different tissue niches, it is unlikely that all E2 proteins are degraded in the same way, and to our knowledge, there are no data available on the stability of E2 proteins from cutaneous β-HPV types.

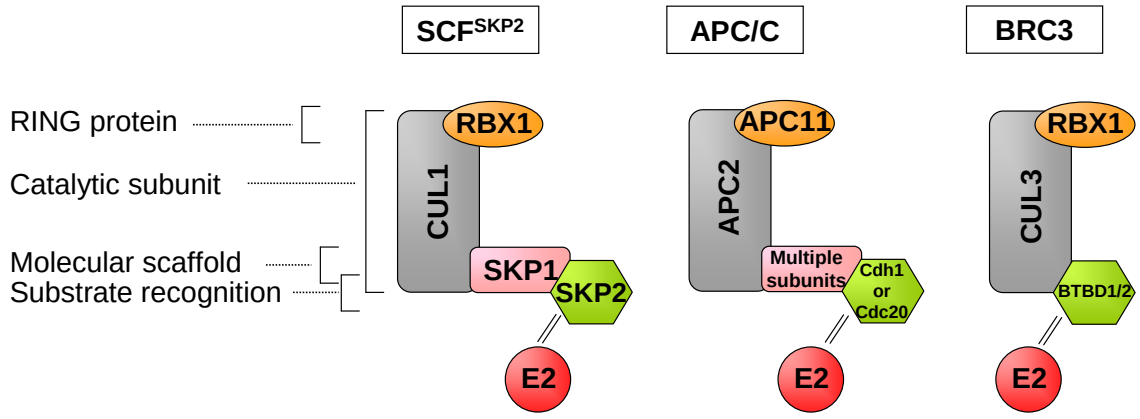

Figure I.17: **Targeting by E2 of ubiquitin ligase complexes**. Schematic representation of three E3 ubiquitin ligase complexes that have been shown to be regulated by the HPV E2 proteins, in particular by the targeting of the substrate recognition subunit.

Reports on the targeting of the ubiquitin-proteasome system by E2 seem to indicate a preferential targeting of ubiquitin ligases of cullin-based complexes, primarily through interactions with various substrate recognition subunits, the factors responsible for the

target specificity (**Fig I.17**). It could suggest that by interacting with the recognition component of these complexes, E2 could influence the choice between multiple proteins to degrade and therefore have a widespread effect on numerous cell pathways.

Cell differentiation and migration

Several recent lines of evidence point to a modulation by E2 of cellular genes involved in keratinocyte migration and differentiation. Indeed, E2 was shown to repress the transcription of ITGB4, which is a major integrin for keratinocyte attachment to the matrix and this leads to the detachment of keratinocytes from the underlying matrix [230]. Integrins play a key role in the development and maintenance of epidermal structures by linking keratinocytes to the underlying basement membrane. The anchorage of keratinocytes to the basal membrane mediated by β-integrins is a negative regulator of terminal differentiation, and cell detachment serves as a signal to commit to the differentiation program [292]. It thus appears that ligation of β-integrins to the basal membrane of the epithelia is a negative regulator of terminal differentiation. By repressing ITGB4, E2 would therefore "push"the cell to enter the differentiation program. E2 also activates the transcription of MMP9, a gene encoding for a protein involved in cell migration [232]. In addition, E2 was shown to bind to and cooperate with C/EBP factors for transcriptional activation [221]. From various studies, evidence is accumulating that C/EBP factors play a regulatory role in keratinocyte differentiation. In particular, forced expression of C/EBPβ in murine keratinocytes results in growth inhibition, a more highly differentiated phenotype, and up-regulation of keratin 1 and keratin 10, two early markers of keratinocyte differentiation [293]. Therefore E2 may contribute to enhanced activation of differentiation-regulated genes through binding C/EBP factors. Accordingly, in a microarray study, E2 was shown to modify the expression profile of cellular genes, among which many were involved in cell differentiation [294]. It was also observed that E2 induces phenotypical changes typical of terminally differentiated cells, with increased expression of differentiation markers [295]. Hence, E2 appears to promote differentiation by acting as a transcription factor which modulates the expression of a set of cellular genes.

The regulation of cell differentiation by E2 might also be mediated through the regulation of different signaling pathways. Indeed, it was recently published that E2 stimulates TNF-induced NFκB activation through direct interaction with TRAF5, an intermediate of the NFκB signaling pathway [296]. This could be involved in cell commitment to differentiation, since activation of NFκB by TNF induces differentiation [297]. Other known cellular proteins interacting with E2 have been implicated in cell differentiation or migration, such as CCHCR1, which affects the balance between proliferation and differentiation in keratinocytes [298, 299], or GNB2L1 reported to be involved in migration of carcinoma cells [300]. The interaction of these proteins with E2 in the context of infection might impact on their stability, localization and function, and thus promote the detachment of keratinocytes from the basal lamina and further induction of differentiation.

The consequences of such targeting are likely to be important for the global viral life cycle. Indeed, as discussed previously in this manuscript, the regulation of the balance between keratinocyte proliferation and differentiation is crucial during HPV infection, and is carefully tuned to allow the stepwise progression through the viral life cycle. By

promoting the natural differentiation program of keratinocytes, E2 actively takes part in this regulation, in particular by antagonizing the effects of E6 and E7.

Overall, E2 appears as a central regulatory HPV protein, involved in various mechanisms as wide as the control of viral genes expression or the extensive rewiring of host cell processes to ensure the proper viral expansion. Given its broad impact on the viral life cycle, its relatively well conserved sequence among the HPV family and the fact that this viral protein is expressed early during the infection, E2 is considered as a good candidate for the development of an anti-HPV therapeutic drug [301].

As described in this introductory chapter, HPV interplay with the host cell is both complex and extensive. Although vaccines are available to prevent the infection of the most common HPVs, there are still no treatment for already infected women. Therefore, there is a current need for the development of anti-viral drugs targeting HPV infection, the ultimate goal being to find a pan-HPV drug. E2 is acknowledged as a reasonably good target for the development of a therapeutic drug. However, more information are needed to better understand its impact on the host cell, and its potential implication in viral carcinogenesis. The focus of this thesis has been to study this key regulatory viral protein to advance on the understanding on how E2 can participate in the infection by hijacking the host cellular machinery. We chose to study E2 through the analysis of its interplay with the host cell proteome, provided that protein-protein interactions are an efficient way for viral proteins to hijack the host. To get a comprehensive overview of E2's interaction and of its involvement in pathogenesis, this study was carried out by comparing multiple HPV genotypes.

The second part of this work was dedicated to the study of one particular interaction identified during the first phase of this study. This interaction particularly drew our attention since it involved the E2 protein from HPV16, the most prevalent HPV in cancers, and a cellular protein, CCHCR1, involved in the regulation of keratinocyte proliferation, an uttermost important mechanism in development of HPV-associated cancer.

The following chapter describes the results obtained during this thesis as well as the potential outcome of such a large-scale approach to better understand HPV's pathogenesis and its possible use to identify specific biomarker of oncogenic HPVs.

II . Results

A - Viral interactomic

a .Context

Viruses are evolutionary constrain to keep their genomes as discrete as possible to expand throughout their host. To enable the expression of the viral genome, some viruses are equipped with the necessary tools for their development such as polymerases, helicases or proteases but they still rely on host additional factors and they must interfere with the host cell regulation in order to provide a proper environment for their spreading. The extensive range of host pathways and cellular functions targeted reflects how viral proteins have a widespread effect on the host cell physiology [302, 303]. To do so, viral proteins are platforms of interactions that are known to orchestrate host proteins localization and degradation, control their expression levels and post-translational modifications and rewire signal transduction in the infected cell. Therefore, finding interaction partners for a viral protein can reveal its function [304]. Significant efforts have been spent for the generation of comprehensive protein-protein interaction network maps in order to greatly improve the understanding of virus-host relationship and to determine the best candidates for assessing disease risk and targeting for therapies [305]. Mapping protein-protein interactions is crucial for unravelling the dynamic aspect of cellular networks, including when, where and for what purpose protein interactions occur [306]. The notion of interactome, defined as the complete map of physical interactions mediated by the proteins of an organism, reflects a drastic change in the way biological questions are addressed nowadays: taking as a whole all the events in the cell at once [307]. An additional step should be taken toward the development of integrative comparative interactomic approaches in order to decipher the pathogen-host interplays by comparing a range of different pathogenic potential, keeping in mind that differences in phenotypes might actually be the result of the interaction network properties [308]. Looking through a comparative approach can help answer the question of which mechanisms or pathways are redundantly triggered by multiple pathogens. However, a major obstacle to such cross-pathogen studies has been the shortage in High Throughput (HT) systems to map and characterize a large number of protein-protein interactions [309].

Yeast two-hybrid screens (Y2H) have been one of the most successful experimental system to identify binary interactions. This technique is based on the functional reconstitution of the GAL4 transcription factor by the interaction between two proteins. Screening cDNA libraries using Y2H is a powerful technique to identify new protein-protein interactions since it relies on genetic screenings and as such is an efficient and sensitive tool. However, this technique has been criticized for having several inherent limitations as for example the fact that protein pairs are artificially targeted to the nucleus which could favor false positive interactions. In addition, the coverage of interactions detected by yeast

two-hybrid is estimated to be around 20% of total protein-protein interactions, reflecting a high false-negative rate. Therefore interactions detected by yeast two-hybrid are not complete and should be further validated.

It was suggested that combining orthogonal methodologies substantially increases the reliability and the robustness of the protein-protein interaction datasets. Methods are considered orthogonal if they are independent, non-redundant and non-overlapping. To circumvent the intrinsic limitations of yeast-based detection methods, it is customary to use mammalian cell-based techniques.

Affinity purification followed by mass spectrometry (AP-MS) is a widely used technique based on the purification of protein complexes from cell extracts. The protein of interest is purified using an epitope tag and mass spectrometry is then used to identify bound proteins. But tagging and expressing each protein proves to be a laborious task and is relatively low-throughput compared to other available techniques. Additionally, transient interactions might be difficult to detect given that they have to go through many purification steps. Several other approaches are available such as LUMIER (LUminescence-based Mammalian IntERactome mapping) based on co-purification of two tagged proteins, one fused to a luciferase and the other with a tag allowing purification, or MAPPIT (Mammalian Protein-Protein Interaction Trap Assay) based on the reconstitution of the receptor of type I cytokines by the interaction between two proteins fused to fragments of the receptor [310]. But these techniques are limited both by the high rate of false positive and false negative interactions and by the fact that it is relatively difficult to apply them to a high throughput format.

Another technique that has proven very efficient to detect protein-protein interactions is based on protein complementation assays (PCA). In this strategy, two proteins of interest are fused to complementary fragments of a reporter protein. If the two fragments

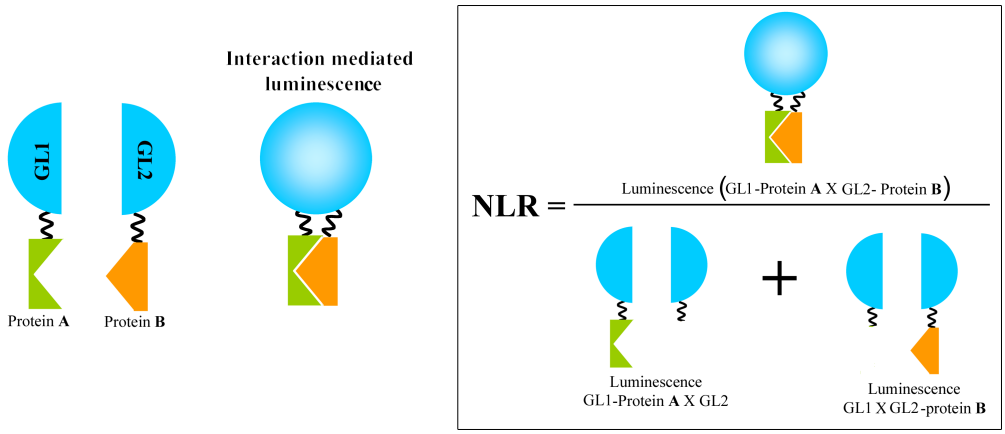

Figure II.1: **Scheme of the HT-GPCA**. Protein A and B are fused to a half of the *Gaussia princeps* luciferase. Upon interaction, the luciferase is reconstituted and luminescence can be measured. The interaction intensity is estimated from the ratio between the luminescence in the presence of both interacting partners divided by the sum of the luminescence in control wells where each partner is expressed alone against the complementary unfused fragment.

are brought in close proximity by an interaction, the reporter is reconstituted into its native structure and it is then possible to measure its activity. Initially developed with fluorescent proteins [311], it had several flaws as the fact that the transfected proteins had to be expressed at high levels to be distinguished from the normal cellular background fluorescence. The sensitivity was then later drastically improved by using reconstitution of luciferase proteins instead of fluorescent proteins [312]. In particular, fragments were engineered using the *Gaussia princeps* luciferase, a small 185 amino acids enzyme that catalyzes the oxidation of coelenterazine, a substrate that can freely enter the cells and generate luminescence 100-fold higher than Firefly or Renilla luciferases. This technique has been adapted in our laboratory by Yves Jacob to be used in a high throughput format and was therefore re-named HT-GPCA (High Throughput *Gaussia princeps* Complementation Assay, [313]). In this method, bait and prey proteins are expressed in fusion with two complementary inactive fragments of the *Gaussia princeps* luciferase. The interaction-mediated reconstitution of the luciferase is expressed as a Normalized Luminescence Ratio (NLR) calculated as shown in **Figure II.1**. Contrary to yeast two-hybrid approaches, a critical aspect of this assay is that it is implemented in mammalian cells, thereby ensuring proper post-translational modifications of proteins. More importantly, the proteins are expressed in HT-GPCA at their natural subcellular localizations, therefore discarding most of the false positive interactions observed in Y2H generated by the forced expression of proteins in the nucleus. HT-GPCA thus appears as a technique of choice to study interactions, and seems particularly appropriate to validate large-scale Y2H datasets

The sensitivity and the reliability of this assay has been benchmarked against two sets of proteins: an *a priori* negative interaction matrix composed of randomly selected cellular proteins and a positive interaction matrix composed of proteins whose interactions are supported in the literature by at least three independent experimental methods [313]. Using HT-GPCA, a clear segregation of the NLR values was observed between these two sets and setting the cut-off for positivity at an NLR of 3.5 discriminated positive interactions with less than 2.5% false positive interactions and with 70% recovery of true interactions (**Figure II.2**). These data place the HT-GPCA at the top of all existing high-throughput methods in terms of robustness and sensitivity to detect pair-wise interactions.

In line with the development of this HT-GPCA, our lab has embarked on the comparative interactomics study of the early proteins of papillomaviruses. In the case of HPVs, great divergence exists between infections with the different viral strains, ranging from variations in the tropism (oral, mucosal, cutaneous epithelia) to variations in the pathogenicity and more particularly, the oncogenic potential. The variability of HPV-associated lesions indicates that the interplay among viral and host proteins is likely to differ greatly from one strain to the other. However, only few studies actually compare a wide range of HPV genotypes, and concerning E2, it was still missing until this study.

b . The HPV E2 proteins interaction network

Given the major public health concern caused by genital cancers, the activities of viral early proteins have been far more extensively studied for the mucosal HR-HPV, and particularly HPV16 and HPV18. Consistent with the idea that comparative interactomics could give insights into the genotype specificities, we decided to map the virus-host

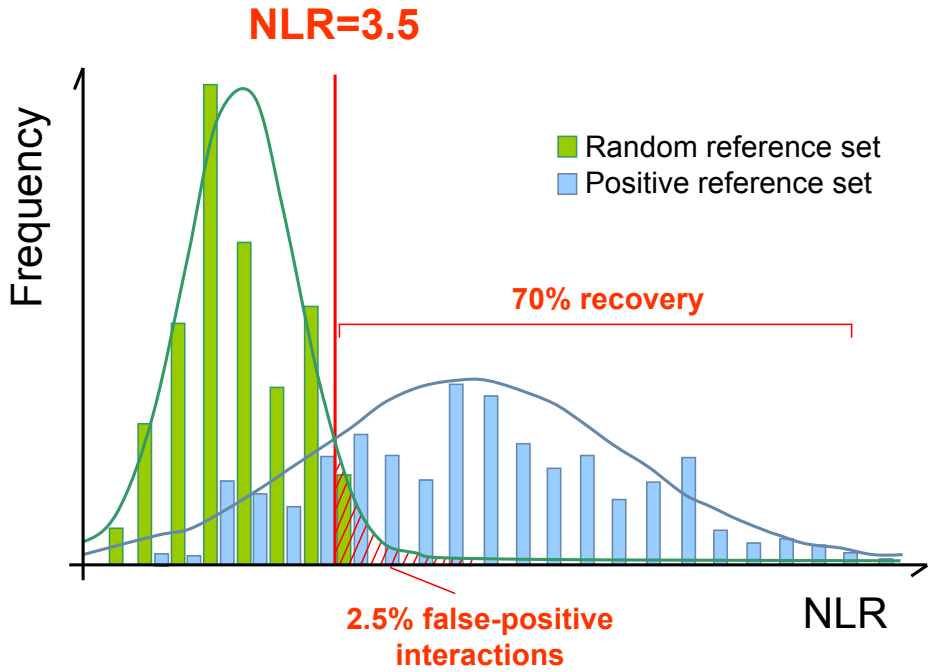

Figure II.2: **Determination of the cut-off for positive interactions**. Frequency distribution of the NLR values for the proteins of the positive interaction matrix (blue) and the random reference set of proteins (green) with their corresponding fitted Gaussian curves. The percentages of false-positive interactions and recovery determined with a threshold for positive interactions at 3.5 are represented in red. Adapted from Cassonnet *et al.* [313]

protein-protein interactions of E2 proteins from 12 HPV genotypes. In order to identify specific features that distinguish different pathological traits, we selected HPV strains representative of the viral natural diversity with differences in tropism (cutaneous HPV1, 3, 5, 8, 9 or mucosal HPV6, 11, 16, 18, 32, 33, 39) and in the oncogenic potential (LR-HPV 1, 3, 6, 9, 11, 32 or HR-HPV 5, 8, 16, 18, 33, 39).

1 .E2-host protein-protein interaction mapping

To get an overview of the cellular partners targeted by the HPV E2 proteins, we first conducted yeast two-hybrid (Y2H) screenings. Each of the 12 selected E2 proteins was used as bait to screen a Human keratinocyte (HaCaT) cDNA library. In total, we collected 251 interactions involving 202 distinct cellular proteins. Five cellular proteins corresponded to already known E2 interactors but were not always found with the same genotype as previously published. As mentioned before, the interactions detected by Y2H had to be validated with an orthogonal method. We thus decided to challenge a subset of cellular proteins selected from the initial Y2H screenings for pair-wise interaction with the whole panel of E2 proteins. The validation was conducted on the 12 E2 proteins selected for this study because, given the acknowledged high false-negative rate of Y2H, we suspected that interactions detected with a subset of the E2 proteins might have

escaped detection with the others.

A total of 102 cellular proteins identified in the Y2H screens were selected on several criteria: high confidence scoring (identified more than three times in the Y2H screens), functional relevance (proteins involved in replication, transcription...), proteins belonging to the same family (Zinc-finger proteins...). We also added positive controls (designated Gold Standards), which consisted in proteins known from the literature to interact with E2. The complete list of gold standards constituted 21 proteins including the five known E2 partners identified in the Y2H screens. A total of 121 cellular proteins were therefore assessed for pair-wide interactions against the 12 E2 proteins (1,452 interactions).

For this validation step, we used the HT-GPCA technique. Indeed, as discussed above, this method is adapted for high-throughput studies and has proven very reliable to detect protein-protein interactions. However, the ability of a given protein to generate false-positive or false-negative results might greatly differ according to its intrinsic biophysical and biochemical properties (folding, stability, presence of disordered regions...). We thus first wished to calibrate the specificity and sensitivity of HT-GPCA when applied to the study of E2. To that aim, we carried out a pilot experiment using the 12 E2 proteins against the 21 gold standards (**Fig II.3A**). Results are shown as a heatmap where the interaction intensities are reflected by a color gradient from black (no interaction) to light blue (strong interactions). In this representation, the E2 proteins are represented by columns and the cellular proteins by rows. Although not all interactions were recovered due to the use of different techniques with inherent false-negative interactions rate, 72% of previously published interactions between the E2 proteins and the gold standards were detected by HT-GPCA. This recovery rate is in good agreement with a previous report using this technique [313] and reflects the high sensitivity of HT-GPCA to detect true interactions.

We also estimated the false-positive rate associated with the E2 proteins by testing their interactions with a set of *a priori* non-interacting proteins randomly selected from the Human ORFeome. This rate was determined to be around 5.8% (**Fig II.3B**), which is a bit above the 2.5% determined previously and probably indicates that E2 is more prone to non-specific interactions than an average protein. As explain in the introduction of this chapter, this false-positive rate is the best of all techniques used to detect protein-protein interactions, which makes us very confident about the reliability of HT-GPCA. We next tested the specificity of the HT-GPCA using point mutants of the E2 proteins and taking advantage of the knowledge about the interaction between E2 and the cellular protein BRD4. Indeed, this interaction has been extensively studied [222, 245] and it has been repeatedly shown that introducing a single point mutation in E2 to change the Isoleucine in position 73 (for HPV16 E2) or 77 (for HPV18 E2) into an alanine strongly decreases BRD4 binding. This is consistent with the observation that this amino acid is part of the interaction interface, as determined by the crystal structure of E2 in complex with BRD4 C-terminal domain [223]. The mutation of this key amino acid in HPV16 and 18 E2 led to a 5-fold decrease in the NLR values obtained by HT-GPCA, reflecting a reduced interaction with BRD4 (**Fig II.3C**) thereby underlining a good accuracy of this technique to detect the loss of interactions. The residual interactions with the mutant E2 proteins corroborates the fact that the interaction between E2 and BRD4 is mediated by more than one amino acid and attests that HT-GPCA is also performant to detect even

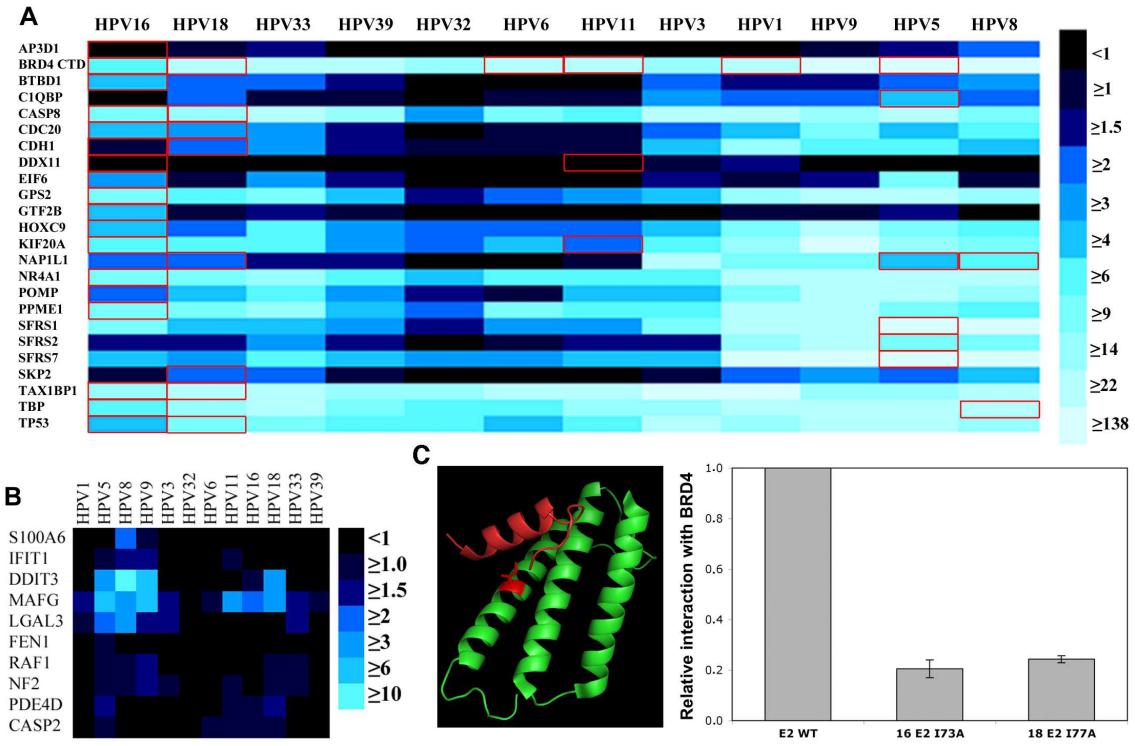

Figure II.3: **Sensitivity of the HT-GPCA using the E2 proteins**. (A) Interactions between the 12 E2 proteins (columns) and the Gold Standards (rows). The color scale is based on the NLR values. The red rectangles indicate the interactions previously published. (B) Interactions between the E2 proteins and a set of randomly selected cellular proteins. (C) Interactions between BRD4 CTD and the E2 proteins, either wild type or mutated (16E2I73A or 18E2I77A). On the left, a schematic 3D representation of the interaction between BRD4 CTD and the HPV16 E2 N-terminal domain, with a particular emphasis on the amino acid I73, essential for the interaction; and on the right, the result of the HT-GPCA. The results are represented as relative to the interaction with the Wild type E2 protein.

weak interactions. Overall, these results demonstrate the robustness and reliability of the HT-GPCA for detecting interactions involving the E2 proteins.

We then proceeded to study the interaction between the 12 E2 proteins and the complete set of cellular proteins selected from the Y2H screens. The raw NLR values obtained are shown in **Table 1** and the interaction profiles as a heatmap in **Figure II.4**. The first observation is that the interaction profiles on the left, corresponding to the E2 proteins from the β and μ-type HPVs, are brighter than the rest of the profiles associated with the α-type HPV E2 proteins. This reflects both a higher number of positive interactions and higher NLR values.

We verified that it was not due to a difference in their basal expression levels (**Fig II.5**). To that aim, E2's relative accumulation levels were deduced from the luciferase activity of the E2 proteins fused to the Firefly luciferase protein, as previously reported [287]. Despite a degree of heterogeneity, the basal expression levels of the E2 proteins did not correlate with the differences observed in the overall NLR levels between β/μ-

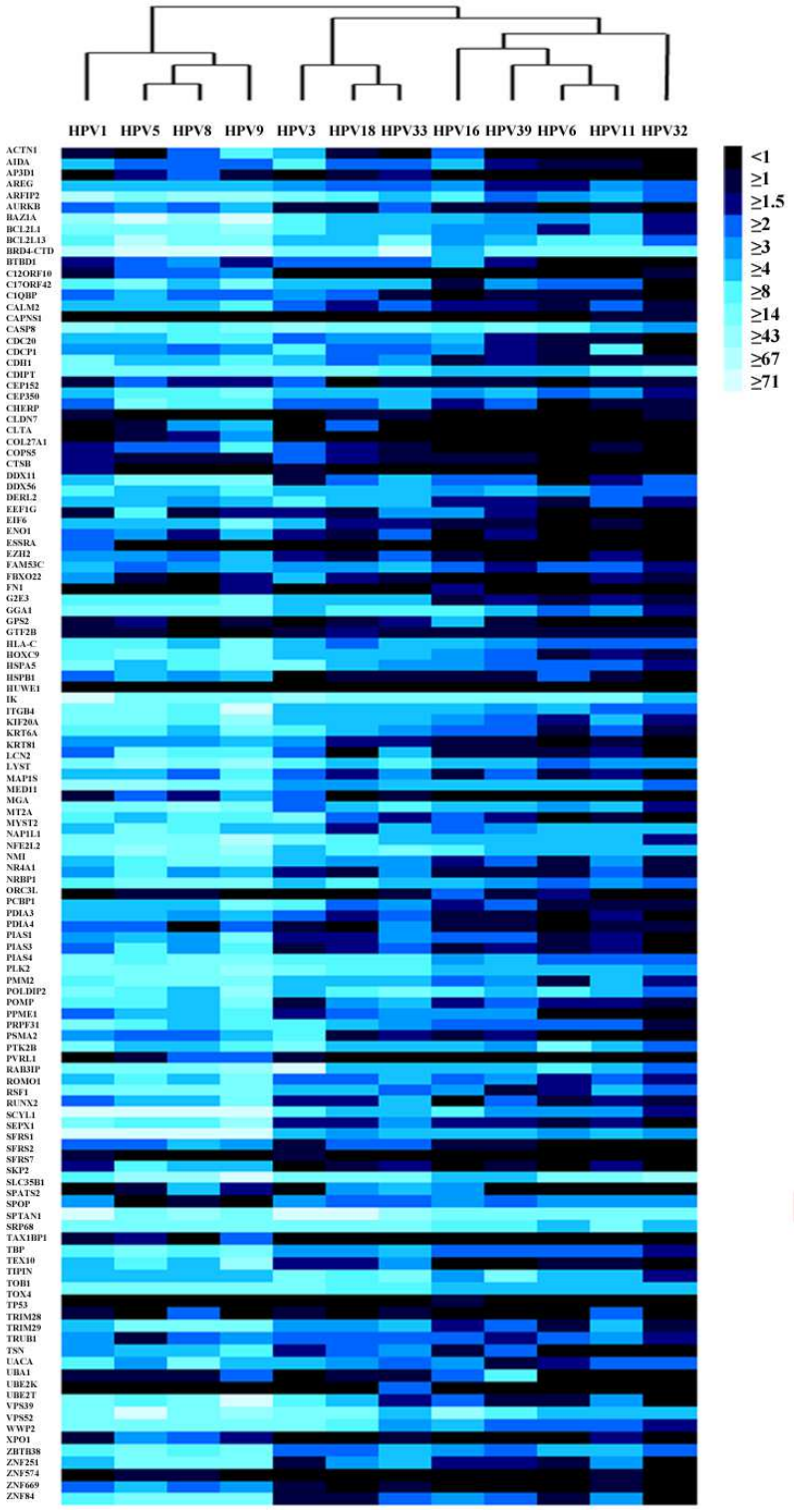

Figure II.4: **Complete interaction dataset between 12 E2 proteins and 121 cellular proteins by HT-GPCA**. Heat Map representing the complete interaction matrix obtained by HT-GPCA between the 12 E2 proteins (columns) and the 121 cellular proteins (rows). The intensity of interaction is represented by a color gradient, from black (no interaction) to light blue (strong interaction) based on the Normalized Luminescence Ratio (NLR). The interaction profiles of the E2 proteins were clustered according to their similarities by hierarchical clustering (tree above the heat map).

Table 1: Interaction levels between the 12 E2 proteins and the cellular proteins

Symbol	HPV1	HPV5	HPV8	HPV9	HPV3	HPV18	HPV33	HPV16	HPV39	HPV6	HPV11	HPV32
ACTN1	1.39	0.74	2.09	8.57	4.31	1.04	0.80	2.03	0.84	0.56	0.65	0.50
AIDA	5.87	2.45	2.72	2.99	9.35	2.43	2.39	4.39	1.75	1.13	1.12	0.68
AP3D1	0.79	1.53	2.21	1.18	0.87	1.00	1.78	0.71	0.70	0.82	0.86	0.81
AREG	5.75	5.40	4.68	5.89	3.89	2.89	2.91	3.10	1.98	1.58	3.27	2.16
ARFIP2	69.84	22.45	48.53	45.90	16.68	13.63	4.74	12.99	2.24	3.99	4.29	2.25
AURKB	2.11	3.91	2.84	4.56	1.07	1.15	2.82	1.08	1.12	0.50	1.03	0.73
BAZ1A	62.95	107.11	43.33	91.73	8.14	7.37	6.58	5.52	3.73	3.68	6.87	1.93
BCL2L1	25.68	29.83	25.52	50.97	9.30	6.15	5.38	3.19	3.23	1.58	7.66	1.66
BCL2L13	9.73	67.06	31.86	33.57	7.20	5.75	22.14	3.69	7.85	8.83	13.54	2.34
BRD4	67.20	200.56	200.90	153.11	14.25	26.40	71.59	6.89	22.97	27.15	35.94	21.65
BTBD1	1.95	2.69	3.00	1.92	2.15	2.46	2.39	4.49	1.50	0.48	0.82	0.61
C12ORF10	1.48	2.74	2.40	3.27	0.98	0.57	0.94	0.89	0.74	0.40	0.61	1.06
C17ORF42	10.47	16.97	5.71	15.25	4.52	4.39	7.03	1.33	3.53	2.64	2.92	0.75
C1QBP	2.30	4.26	2.83	2.88	3.25	2.09	1.04	0.97	1.13	1.08	1.04	0.72
CALM2	4.09	7.24	4.87	12.60	2.04	1.77	2.12	1.58	1.52	1.00	2.83	1.47
CAPNS1	0.79	0.34	0.48	0.34	0.61	0.88	0.67	0.93	0.80	0.59	1.15	1.13
CASP8	50.62	27.29	11.46	14.60	53.49	21.64	23.76	11.38	16.02	11.69	6.99	3.37
CCHCR1	67.86	18.65	27.94	31.98	15.59	7.59	3.07	271.62	1.48	3.17	2.99	1.02
CDC20	5.52	5.38	8.80	9.85	2.04	3.44	3.83	5.41	1.58	1.41	1.24	0.96
CDCP1	3.32	3.76	2.77	3.58	8.75	2.17	2.07	3.64	1.55	1.06	8.69	0.93
CDH1	15.42	7.67	5.70	8.33	4.21	2.50	3.26	1.44	1.55	1.21	1.24	1.04
CDIPT	42.50	17.49	18.20	31.41	39.49	18.85	12.46	6.66	7.84	6.10	8.42	18.84
CEP152	1.15	2.53	1.97	1.69	2.18	0.86	1.06	1.13	1.19	0.69	1.20	1.18
CEP350	7.01	10.13	11.41	14.11	4.43	5.09	6.41	3.79	4.00	2.44	3.57	1.63
CHERP	2.94	20.92	12.24	12.47	2.85	2.99	7.35	1.84	2.11	0.95	1.37	1.15
CLDN7	1.20	0.67	0.64	0.43	0.81	1.01	1.04	0.81	0.96	0.54	0.68	1.07
CLTA	0.94	1.32	3.08	7.09	0.34	2.74	0.97	0.39	0.31	0.18	0.34	0.29
COL27A1	0.76	1.10	1.89	3.12	0.73	0.55	0.51	0.28	0.28	0.26	0.21	0.27
COPS5	1.67	2.48	2.87	11.07	2.60	1.50	1.32	0.72	0.76	0.63	1.11	0.70
CTSB	1.87	1.21	1.32	1.17	2.84	1.53	1.39	1.01	1.22	0.90	0.97	0.66
DDX11	1.90	0.96	0.83	0.56	1.35	0.71	0.76	0.99	0.54	0.49	0.49	0.28
DDX56	7.89	27.49	18.04	31.23	1.19	2.16	5.69	2.40	2.39	0.87	1.72	2.14
DERL2	8.03	7.77	6.74	11.07	7.24	7.32	7.62	3.05	4.89	3.37	2.21	2.22
EEF1G	4.14	5.32	3.02	5.86	9.66	4.31	4.20	1.79	1.93	1.43	2.69	1.57
EIF6	1.23	10.53	1.15	1.70	1.54	1.35	3.56	3.88	1.70	0.46	0.73	0.52
ENO1	4.21	7.03	4.18	18.80	4.10	1.67	1.81	1.09	1.02	0.71	1.08	0.70
ESSRA	2.27	3.93	1.70	5.22	1.74	1.27	2.18	0.76	1.50	0.65	0.87	0.86
EZH2	2.19	0.55	0.46	0.83	0.30	0.35	0.68	0.35	0.48	0.23	0.47	0.42
FAM53C	3.59	3.62	2.06	5.74	1.59	1.23	2.73	1.05	0.94	0.64	1.59	0.60
FBXO22	4.42	2.95	3.04	4.75	3.18	3.59	5.57	2.64	1.85	2.58	2.15	1.52
FN1	3.40	1.40	0.80	1.59	7.95	1.60	1.22	0.98	0.94	0.96	1.85	1.07
G2E3	0.82	0.88	0.72	1.54	0.54	0.74	0.71	1.68	0.57	0.34	0.38	0.31
GGA1	8.15	9.89	9.26	24.61	6.10	7.77	4.56	1.38	1.68	1.19	1.50	1.27
GPS2	19.55	37.52	14.67	38.77	4.03	8.16	9.44	10.08	4.24	2.12	3.22	1.99
GTF2B	1.00	1.58	0.95	1.09	0.66	1.04	1.86	4.17	1.34	0.75	0.83	0.74
HLA-C	1.48	1.00	0.56	0.56	1.43	1.52	1.15	1.22	1.34	1.03	1.23	1.04
HOXC9	10.04	10.88	7.01	41.60	4.26	2.62	7.50	5.88	3.90	2.84	2.82	2.01
HSPA5	13.25	15.56	11.65	19.41	4.67	4.79	5.42	3.54	2.22	1.40	1.58	1.49
HSPB1	14.10	7.06	8.98	27.85	15.75	7.16	3.44	3.71	2.90	2.60	2.48	1.85
HUWE1	2.04	5.80	3.61	4.63	0.78	1.00	1.27	1.42	1.23	2.64	1.09	0.92
IK	0.56	0.74	0.53	0.33	0.76	0.44	0.64	0.64	0.44	0.34	0.51	0.63
ITGB4	104.73	34.88	21.80	29.72	63.62	33.27	20.80	15.96	14.62	21.47	15.67	7.54
KIF20A	21.92	15.87	11.76	167.93	7.12	6.60	6.15	6.80	3.46	5.17	2.90	2.53
KRT6A	17.72	29.59	10.69	45.11	7.31	6.73	5.35	3.06	2.94	1.77	4.31	1.60
KRT81	10.62	8.13	5.71	15.17	13.19	5.27	3.07	2.27	2.20	1.45	2.07	1.18
LCN2	3.69	3.81	3.92	6.00	3.64	1.54	1.76	1.38	1.40	0.92	1.07	0.88
LYST	2.26	15.49	8.72	9.39	2.41	0.80	6.20	1.11	1.14	1.03	1.51	0.68
MAP1S	30.89	62.59	26.34	55.74	11.58	7.74	8.19	4.05	5.38	2.33	3.69	3.09
MED11	5.37	6.84	2.53	8.71	2.09	1.62	3.28	1.30	2.37	1.25	1.62	0.55
MGA	44.53	46.24	28.14	26.55	3.92	3.74	6.03	6.88	4.37	4.58	4.53	2.31
MT2A	1.01	2.42	1.62	4.95	2.13	0.72	1.39	0.64	0.63	0.25	0.46	0.56
MYST2	15.36	28.16	44.61	23.38	2.98	6.13	10.93	4.08	4.80	3.93	4.89	1.93
NAP1L1	10.39	4.04	8.41	10.02	22.46	2.86	1.55	2.86	1.94	0.77	1.06	0.45
NFE2L2	4.68	25.04	13.45	5.93	6.59	1.69	4.69	2.51	3.14	4.02	4.74	4.46
NMI	26.89	30.60	40.06	67.90	37.20	12.82	5.81	6.22	4.28	4.75	5.13	1.87
NR4A1	25.71	44.77	32.00	62.60	6.00	10.29	16.76	9.37	7.80	6.92	6.87	4.37
NRBP1	5.73	12.76	17.02	18.21	5.51	3.60	3.29	1.91	2.77	1.26	3.26	1.28
ORC3L	3.33	11.11	3.08	6.63	1.94	1.28	3.37	1.27	1.26	1.14	2.04	1.12
PCBP1	13.89	21.59	17.67	30.04	5.59	9.96	4.96	4.10	3.03	2.88	3.36	2.42
PDIA3	0.87	1.19	1.07	0.72	0.66	0.68	1.09	2.52	1.07	1.81	0.71	0.52
PDIA4	6.46	5.81	7.13	19.35	9.69	2.04	3.01	1.65	2.03	1.11	1.47	1.23
PIAS1	4.23	5.80	3.33	6.71	2.58	1.64	2.71	1.11	1.16	0.80	1.72	0.73
PIAS3	2.07	2.71	0.91	2.07	1.45	0.99	3.34	1.05	1.33	0.63	1.47	1.02
PIAS4	3.68	5.76	3.17	16.77	1.83	1.66	3.59	2.46	2.29	1.04	1.63	0.73
PLK2	2.80	9.98	3.92	9.39	1.27	1.81	2.79	1.13	1.61	1.09	1.63	0.65
PMM2	28.34	11.18	19.11	22.96	8.34	10.51	12.81	3.82	5.05	2.04	2.95	2.44
POLDIP2	27.79	36.74	25.79	54.98	16.18	12.95	11.24	6.69	7.47	5.87	7.96	3.40
POMP	12.52	34.41	16.58	40.76	5.08	4.77	7.09	2.36	3.17	1.46	5.34	1.55
PPME1	29.71	12.38	7.38	45.72	6.68	9.20	14.85	11.27	5.23	9.77	7.24	2.81
PRPF31	8.60	11.24	5.88	21.58	1.38	3.76	5.19	1.82	2.28	1.82	1.94	1.07
PSMA2	2.74	7.89	4.85	8.50	1.75	2.15	3.86	3.22	3.07	0.91	0.72	0.69
PTK2B	14.44	11.96	4.46	8.98	9.26	5.14	3.75	2.68	2.31	2.07	2.02	1.48
PVRL1	3.36	2.96	2.59	7.01	10.03	1.48	1.68	1.43	1.69	0.83	0.89	0.89
RAB3IP	20.61	6.46	6.36	11.28	32.06	6.29	5.21	4.38	3.09	18.37	4.60	2.00

Note: Values represent the NLR obtained by HT-GPCA

Symbol	HPV1	HPV5	HPV8	HPV9	HPV3	HPV18	HPV33	HPV16	HPV39	HPV6	HPV11	HPV32
ROMO1	0.83	1.29	2.00	2.81	1.43	0.86	0.83	0.66	0.62	0.43	0.60	0.66
RSF1	35.57	34.15	34.67	61.10	89.78	6.85	4.92	6.25	7.99	8.15	4.88	2.12
RUNX2	5.75	8.46	5.58	17.59	2.50	2.47	5.20	2.74	3.90	1.71	2.55	1.50
SCYL1	21.49	14.50	16.29	22.37	4.35	7.94	2.97	3.62	1.41	1.94	5.75	2.30
SEPX1	2.61	6.96	5.96	19.75	1.53	1.71	4.11	0.96	2.14	1.05	1.89	1.00
SFRS1	82.18	167.51	183.54	127.99	13.21	4.01	4.17	10.66	3.30	3.85	3.82	1.57
SFRS2	20.50	13.78	10.00	45.77	1.99	1.76	3.06	1.64	1.81	1.04	1.79	0.80
SFRS7	168.71	280.03	782.45	227.78	5.90	3.05	7.68	4.51	5.14	3.84	4.76	3.58
SKP2	2.31	2.76	4.27	3.70	1.08	2.43	2.23	1.15	1.15	0.70	0.69	0.75
SLC35B1	1.23	0.68	0.57	0.50	1.29	0.83	0.78	0.80	0.83	0.53	0.71	0.50
SPATS2	1.55	9.99	6.92	5.61	0.62	1.15	1.70	0.88	1.05	0.52	1.59	0.71
SPOP	9.53	59.11	47.34	83.17	37.18	16.28	24.57	5.69	4.08	26.13	19.38	57.28
SPTAN1	0.64	1.06	7.16	1.59	0.68	3.20	4.15	3.27	0.39	0.92	0.80	0.37
SRP68	3.43	0.68	1.43	0.73	3.19	2.53	2.03	3.56	2.39	3.30	3.72	3.37
TAX1BP1	138.42	18.36	48.03	19.19	226.78	71.30	46.04	21.17	18.67	22.17	23.23	20.80
TBP	25.73	38.14	27.17	25.48	18.22	19.51	36.99	8.14	12.94	6.84	16.73	7.16
TEX10	1.00	1.52	0.97	2.57	0.42	0.38	0.68	0.35	0.34	0.24	0.34	0.18
TIPIN	12.26	14.13	10.57	24.35	3.52	3.86	7.29	2.43	2.05	2.15	2.44	1.58
TOB1	4.74	8.81	4.74	44.46	1.80	1.98	3.09	0.94	0.91	1.14	1.18	0.68
TOX4	6.08	5.80	5.28	7.29	25.11	13.41	16.64	3.23	18.94	7.28	5.94	1.68
TP53	14.23	26.43	17.11	31.42	23.81	11.38	12.83	5.68	6.61	4.71	7.89	6.58
TRIM28	0.71	0.94	0.65	0.42	0.45	0.53	0.78	1.45	0.68	0.39	0.35	0.28
TRIM29	1.18	0.94	2.17	0.90	1.11	0.82	1.10	0.48	0.66	0.37	2.17	0.41
TRUB1	6.36	17.88	14.62	19.39	3.51	3.94	6.49	1.86	2.28	1.12	6.34	1.20
TSN	3.43	1.20	2.34	3.14	2.64	2.23	2.14	2.84	1.55	2.54	3.47	1.96
UACA	3.54	6.59	5.31	10.98	1.79	2.25	3.13	1.01	2.38	0.88	0.87	0.53
UBA1	8.47	3.94	14.23	4.43	4.73	3.12	2.32	3.51	1.41	1.71	2.58	2.24
UBE2K	1.19	1.47	1.30	2.96	0.56	1.15	1.46	2.66	13.78	0.45	0.69	0.52
UBE2T	0.14	0.42	0.12	0.66	0.17	0.24	2.39	0.96	0.41	0.18	0.22	0.16
VPS39	25.08	11.58	20.04	114.54	10.33	6.13	1.73	2.36	1.20	1.05	3.80	0.83
VPS52	40.22	84.82	36.25	57.93	22.01	19.05	6.18	56.03	8.64	7.85	5.49	4.28
WWP2	25.60	18.21	19.10	37.36	16.12	9.35	3.57	4.06	2.69	2.35	2.21	1.51
XPO1	1.15	3.30	2.11	1.72	0.60	0.95	1.24	0.58	0.79	0.27	1.51	0.48
ZBTB38	10.34	16.94	10.24	38.02	2.94	2.24	4.99	3.22	2.54	4.13	4.23	2.03
ZNF251	6.91	32.82	38.13	32.14	1.34	3.13	4.04	1.76	1.97	1.26	3.55	0.80
ZNF574	0.70	1.04	1.05	0.84	0.56	0.54	0.88	0.43	0.79	0.30	1.18	0.57
ZNF669	2.11	7.27	2.51	3.04	1.04	0.54	1.16	0.55	0.60	0.41	1.11	0.34
ZNF84	9.58	27.72	34.62	24.95	1.33	1.41	2.94	3.43	2.10	1.38	3.11	0.76

Note: Values represent the NLR obtained by HT-GPCA

type HPV E2 proteins and the α-type HPV E2 proteins. The most striking examples illustrating this point are that of HPV33 E2, which exhibits the highest rate of interaction among the genital α-type HPV E2 proteins while accumulating at levels similar to that of HPV16 or 39 E2; or of HPV9 E2 having the strongest interaction rate in HT-GPCA while accumulating to modest levels compared to other β and μ E2 proteins (**Fig II.5**). Therefore variability in the expression levels of the E2 proteins could not explain the differences in the interaction rate. However, it is important to notice that HPV32 E2 yielded the fewest number of interactions and had a very low expression level, potentially suggesting that HPV32 E2 had an intrinsic expression problem in HT-GPCA.

One particularity of the β and μ-types HPV E2 proteins that could explain this higher interaction rate is that they contain longer hinge regions than the α-type E2 proteins. This region is an intrinsically disordered segment of the protein, which would be consistent with the observation that intrinsic disorder is a common feature of highly connected proteins [314]. Disordered regions offer many benefits for interactivity since structural plasticity allows the protein to adapt to many ligands, to switch rapidly from one partner to another and to bind several proteins at the same time even if the binding interfaces are adjacent.

A major strength of such large-scale validation is that it enabled us to compare E2's interactions between different HPV strains against a common set of cellular factors using the same detection assay, thereby providing a rigorous comparative mapping. It is therefore possible to compare the interaction profiles of the E2 proteins and therefore to extract general characteristics of the virus-host interplay. The interaction profiles were treated by hierarchical clustering and dendrograms were generated based on their similar-

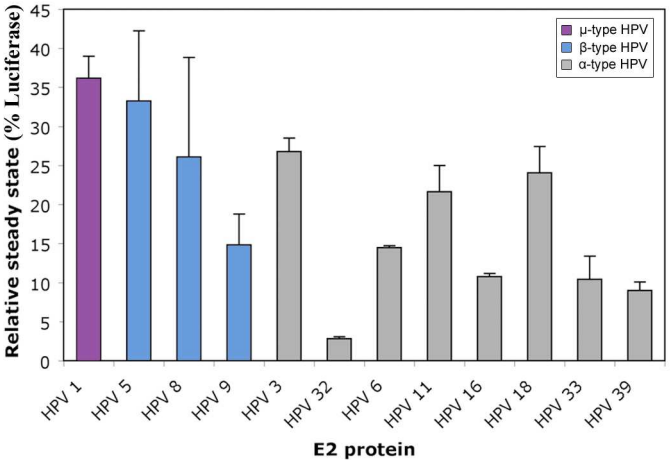

Figure II.5: **Expression levels of the 12 E2 proteins**. E2-Firefly luciferase fusion proteins were expressed in 293T cells and the firefly luciferase activity was determined 24h post-transfection. The results are expressed as a percentage of the activity obtained with the firefly luciferase only.

ities. The algorithm used takes into account all interaction values (positive and negative interactions) for each of the E2 proteins, then the distance between the profiles are compared two by two and classified according to their maximal proximities by agglomerative hierarchical clustering to obtain the interaction-based dendrogram. Concurrently, a dendrogram was generated using E2 polypeptidic sequences to create a phylogenetic tree. In the phylogenetic tree, we can observe that there are two degrees of segregation: there is a first dichotomy between the α-type E2 proteins on one branch and the other E2 on another branch which mostly reflects a dichotomy based on the tropism (the cutaneous HPV localized on one branch and the mucosal on the other). Second, the HR-HPV E2 segregates in separate branches than the LR-HPV E2. Both the dendrograms based on E2's interaction profiles and the phylogenetic tree were compared and a correlation coefficient was calculated to be over 0.9, demonstrating a high congruence (**Fig II.6**). Indeed, E2's interaction profiles can also be separated into two main groups: the interaction profiles of the β and μ E2 proteins and those of the α E2, both dendrograms therefore showing the same first branching dichotomy. Within each group, interaction profiles further clustered according to the HPV pathogenic potential (High-risk vs. Low-risk) even though it does not strictly follows the phylogeny. This result demonstrates that E2 engages different interaction patterns depending on HPV tropism and oncogenicity. Using a similar approach, a strong correlation had also been detected between the interaction profiles of the oncoproteins E6 and E7 using the same panel of HPV genotypes [315]. In the case of E6 and E7, it was somewhat expected since numerous activities of these oncoproteins have been for long associated with the oncogenic traits of HPVs and shown to be specific of the HR-HPV. Therefore, it was reasonable to hypothesize that, by looking only at the interactions of E6 and E7, it could be possible to distinguish interaction profiles of the HR-HPV and those of the LR-HPV. In contrast, E2 is principally envisioned as a basic factor essential for all HPVs through its regulatory roles in viral DNA transcription,

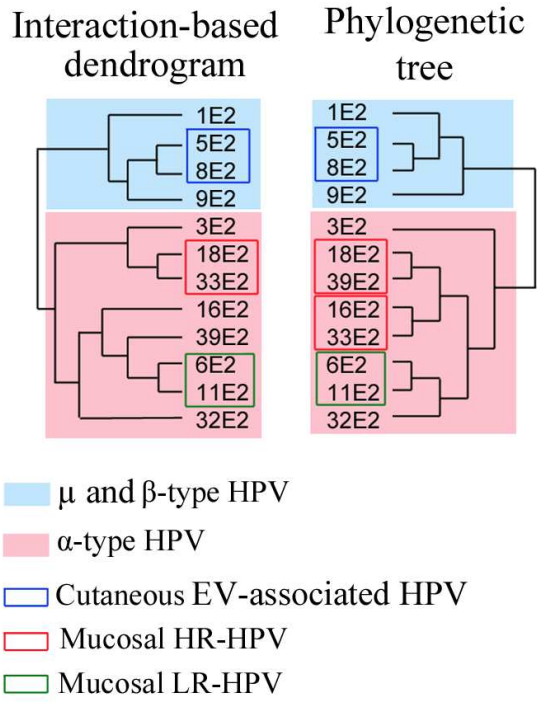

Figure II.6: **Comparison of interaction-based dendrogram and phylogenetic tree**. Left: The distance between the interaction profiles were calculated from the NLR data and a dendrogram was generated by agglomerative hierarchical clustering using the UPGMA method (Unweighted Pair Group Method with Arithmetic Mean). Right: E2 protein sequences were gathered to calculate the distance and to build the phylogenetic tree. The correlation coefficient was calculated with a Pearson test using the cophenetic distances between both the interaction and phylogenetic dendrograms and between randomly generated dendrograms and the phylogenetic tree. The congruence was estimated at 0.9 with a p-value $< 10^{-10}$. Both dendrograms were generated using JavaTreeView.

replication and segregation. Only few studies have actually identified specific activities associated with E2 and correlating with the HPV oncogenic potential. Here, we show that E2 differentially targets cellular partners according to both the tropism and HR or LR trait of HPVs. It indicates that E2 must, to some extent, contribute to the phenotypic characteristics of the virus and therefore take part to the pathogenesis of HPVs. Such comparative interaction mapping thus greatly improves the understanding of E2-mediated cell alterations.

2 .E2-host interaction network

Topology - Topological analysis of viral interaction networks can provide information about the global impact of viral proteins on the host network, as well as about the dynamics of the pathogenesis. We therefore proceeded to the topological study of E2's interacting partners. The "degree"parameter reflects the number of interactions engaged by a protein, reflecting its connectivity, while the degree distribution gives a measure of the network local dynamics. The degree distribution of the E2-host interaction network was compared to that of a Human interactome constructed from a human protein-protein interaction database (HPRD 2010, release 9) (**Fig II.7**).

We calculated the mean degree of this human interactome to be around eight, with 75% of the proteins having a degree lower than this value. This means that most of the proteins of the human interactome interact with less than eight other proteins, while few proteins are extremely connected. By contrast, in the E2 interactome, only 25% of the cellular proteins have a degree under eight (**Fig II.7**A), indicating that E2 preferentially targets highly connected cellular proteins also called Hub proteins. The study of the distribution of degree probability further substantiates the overrepresentation of highly connected proteins in the E2 interactome (**Fig II.7**B). If the distribution of degree probability can be fitted by a power law, it means that the network is scale-free [316], which seems to be

the case for the human interactome. In networks of this topology, most proteins (nodes) have one or few interactions (edges) and few proteins, the so-called hubs, have a very large number of interaction partners, which is consistent with the cumulative plot representing the protein degrees in the human interactome (**Fig II.7**A). Networks with this topology are also found in other complex networks such as the internet or social networks, and are resistant for attack of random nodes, by mutation for example, but sensitive to targeted attack of the hubs [317, 318]. Hubs are central in the cellular network and targeting these proteins can result in drastic changes in the network, forcing the rewiring of the main pathways or even leading to its fragmentation. The preferential targeting of Hub proteins was previously observed with several other viral factors [303, 319, 320] and could be a hallmark of viral proteins to hijack the cellular interactome. Indeed, it allows viral proteins to have a broad impact on the host cell by interacting with key proteins of the host signaling pathways. It thus likely maximizes the effect of viral proteins, in our case, E2, by allowing them to control a wide range of cellular functions *via* a minimum of interactions.

Functional enrichment - The study of virus-host interaction networks can also provide an overview of the targeted cellular functions. We thus next wanted to analyze the E2 interactome from a functional point of view to get insights into the roles of E2 that could emerge.

E2-targeted cellular proteins were gathered with the DAVID Bioinformatics database into functional families based on their Gene Ontology (GO) classification (**Fig II.8**). This analysis revealed that the cellular proteins targeted by E2 are distributed into five main functional families: regulation of transcription, regulation of apoptosis, RNA processing,

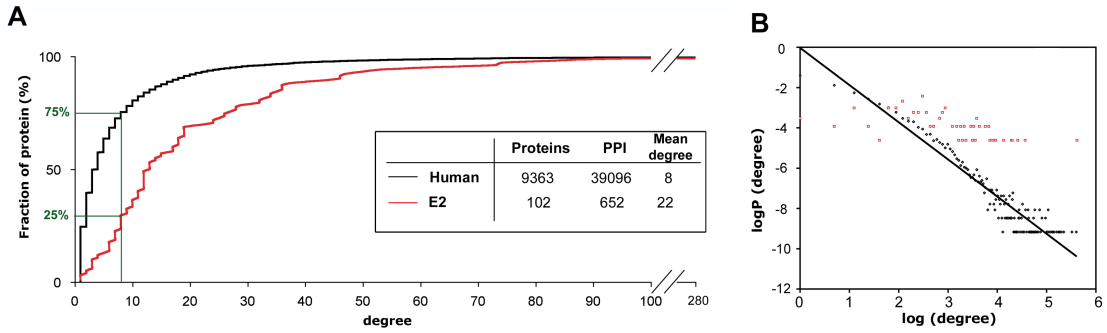

Figure II.7: **Topological analysis of the E2 interactome**. (A) Cumulative distribution of the node degree of the Human interactome (black) and the E2 interactome (Red). The average degree of the human interactome is estimated to be around 8 and is represented in green as well as the fraction of protein under the mean for each interactome. The inset summarizes the characteristics of each interactome. (B) Distribution of degree probability of the human (black) and the E2 interactome (red). P(degree) is the probability to connect K other proteins in the network. For the human interactome, the straight line represents the linear regression fit of the data (with a correlation coefficient R^2=0.91). For the E2 interactome, the data did not fit a linear regression.

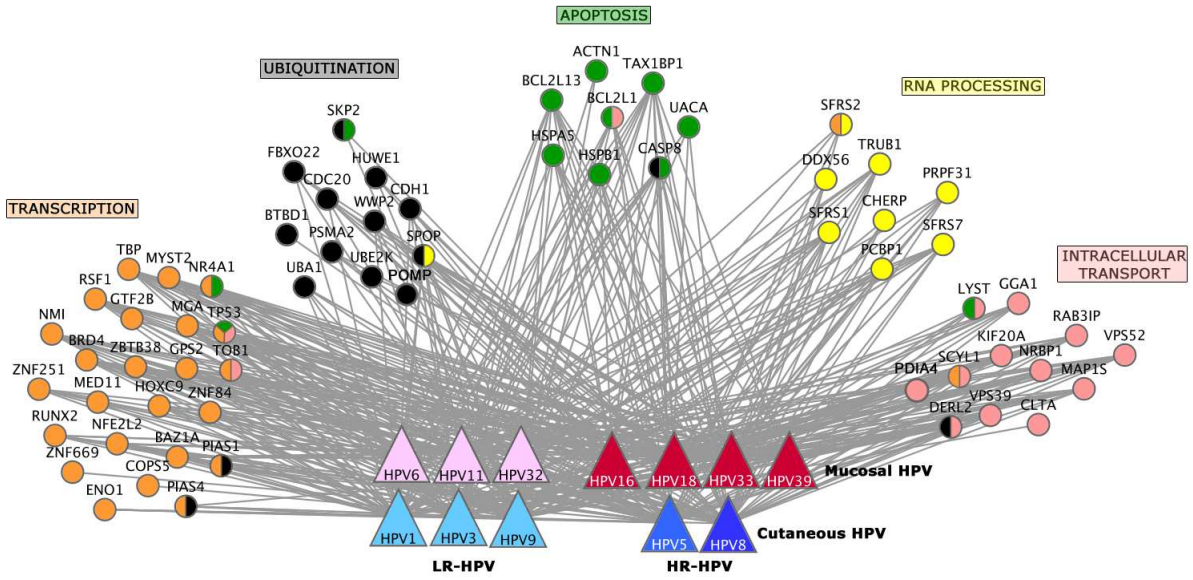

Figure II.8: **E2-targeted functional families**. Cellular proteins targeted by the E2 proteins classified into functional family based on their Gene Ontology classification. Each functional families is displayed by a specific color, proteins belonging to several families are multi-colored. Cellular proteins are represented by circular shapes and viral proteins by triangular shapes. The network representation was generated by Cytoscape.

ubiquitination processes and intracellular transport. The following section describes these functional categories and gives an overview of E2's biological functions across multiple HPV genotypes.

◇ Transcription regulation. The most represented and significant category of E2 partners emerging from the GO analysis consists of proteins involved in transcriptional regulation. This is consistent with the prominent role of E2 as a transcription factor and thus reinforces the pertinence of our interactomic approach. E2's interplay with the transcription machinery was detected across all genotypes tested and was quite heterogeneous (**Table 2**). Therefore, targeting of a large spectrum of transcription processes is likely to be a common trait of all E2 proteins. While E2 is able to interact with either transcriptional activators and repressors, activators are predominants. In addition, the interaction with transcriptional repressors seemed more pronounced for the β-type HPV E2 proteins. Our study has identified BRD4 as one of the strongest interactor of most of the E2 proteins, with a large range of interaction intensities. Of note, the interaction of BRD4 with HPV16 E2, while consistently detected, was the weakest compared to all of the E2 proteins tested. Others have shown that the E2 proteins interact with BRD4 with different affinities, with a less efficient binding for the α-type HPV E2 proteins. [250]. Yet, HPV16 E2 is among the most efficient transcriptional activator when tested on an E2-responsive promoter. This suggests that BRD4 has no significant contribution in the strength of transcriptional activity of HPV16 E2.

Table 2: Interaction levels between the 12 E2 proteins and cellular proteins involved in the regulation of transcription

	Symbol	HPV1	HPV5	HPV8	HPV9	HPV3	HPV18	HPV33	HPV16	HPV39	HPV6	HPV11	HPV32
Transcription activation	**BRD4**	+++	+++	+++	+++	+	++	+++	+	++	++	++	++
	SCYL1	++	+	+	++	+	+	+	+	-	-	+	-
	COPS5	-	-	-	+	-	-	-	-	-	-	-	-
	NMI	++	++	++	+++	++	+	+	+	+	+	+	-
	GTF2B	-	-	-	-	-	-	-	+	-	-	-	-
	HOXC9	+	+	+	++	+	-	+	+	+	-	-	-
	MED11	+	+	-	+	-	-	+	-	-	-	-	-
	NFE2L2	+	++	+	+	+	-	+	-	+	+	+	+
	NR4A1	++	++	++	+++	+	+	+	+	+	+	+	+
	RUNX2	+	+	+	+	-	-	+	-	+	-	-	-
	TBP	++	++	++	++	+	+	++	+	+	+	+	+
	TP53	+	++	+	++	++	+	+	+	+	+	+	+
	ZBTB38	+	+	+	++	-	-	+	+	-	+	+	-
Transcription activation & repression	PIAS1	+	+	+	+	-	-	-	-	-	-	-	-
	MGA	++	++	++	++	+	+	+	+	+	+	+	-
	MYST2	+	++	++	++	-	+	+	+	+	+	+	-
	RSF1	++	++	++	+++	+++	+	+	+	+	+	+	-
Transcription repression	**BAZ1A**	+++	+++	++	+++	+	+	+	+	+	+	+	+
	ENO1	+	+	+	+	+	-	-	-	-	-	-	-
	SFRS2	++	+	+	++	-	-	-	-	-	-	-	-
	GPS2	+	++	+	++	+	+	+	+	+	-	+	-
	TOB1	+	+	+	++	-	-	-	-	-	-	-	-
	PIAS4	+	+	+	+	-	-	+	-	-	-	-	-
	ZNF251	+	++	++	++	-	-	+	-	-	-	+	-
	ZNF669	-	+	-	-	-	-	-	-	-	-	-	-
	ZNF84	+	++	++	++	-	-	-	+	-	-	+	-

Note: The symbol - stands for a lack of interaction and from + to +++, increasing NLR values.
Please refer to table II.1 for details on the raw NLR values.
In bold are highlighted the members of the transcription core.

Also, the weaker interaction of BRD4 with HPV16 E2 could impact on other E2 functions as the segregation of viral genomes through mitosis, which could rely on the interaction with alternative factors in the case of HPV16, or on E2 expression levels since BRD4 has been shown to stabilize E2.

Among the list of transcriptional regulators targeted by E2, 10 are common targets of all E2 proteins and could thus constitute a "transcription core", essential for the basic transcriptional properties of E2. In this core, only TBP is known to directly bind the HPV LCR, the other sequence-specific transcription factors being rather involved in the regulation of cellular promoters. It could thus indicate that part of E2 transcriptional functions are devoted to the regulation of the host gene expression. This corroborates the observation that E2 is able to affect the expression of a number of cellular genes [218, 230–233] to contribute to appropriate cell conditions for successive stages of the viral life cycle. Other E2 targets act at the chromatin level and are components of chromatin-remodeling complexes. These cellular proteins have been shown to affect transcription either positively or negatively and could thus be instrumental for the timely regulation of viral gene expression throughout the viral life cycle.

Non-shared interactors would rather reflect intrinsic differences in the transcriptional properties of the E2 proteins and highlight mechanistic variability in E2 transcriptional functions. For example, data from the HT-GPCA shows GTF2B as an interactor specific to HPV16 E2. To provide a functional validation of this HT-GPCA result, we wanted to determine if this specific binding could be reflected functionally. To that aim, we used the pTK6E2BS-luc plasmid, a luciferase reporter plasmid containing six E2 binding sites (E2BS) upstream of the minimal TK promoter.

Co-expression of GTF2B with HPV16 E2 increased 16E2's transcriptional activation of this E2-responsive promoter by a 2.6 factor, while the effect on 18E2-mediated transactivation was minor (1.7 fold, **Fig II.9A**). Accordingly, siRNA-mediated silencing of GTF2B impaired the activation of transcription by 16E2 but not by 18E2 (**Fig II.9B**). These results substantiate both the functional relevance and the specificity of 16E2/GTF2B interaction. Hou and colleagues [206] have demonstrated that 16E2 is an especially potent transcriptional activator in a reconstituted E2-dependent *in vitro* test and hypothesized that it might result from a more efficient recruitment of general transcription regulators to promoter sequences. The specific targeting of GTF2B by HPV16 E2 substantiates this hypothesis.

Overall, mapping of E2-host interactions revealed a large variety of transcriptional regulators targeted by E2, and thus provides an experimental appraisal of the complex interplay between E2 proteins and the host cell transcriptional machinery.

◇ RNA processes. In addition to transcription, it was for long suspected that E2 regulates viral genes expression by the regulation of mRNA splicing. In particular, the targeting of regulators of RNA processing like proteins of the SR family was expected for the β-type HPV E2 proteins from previous report demonstrating their interactions. However, the identification of this functional family among all the E2 proteins tested provides evidence that this targeting is not restricted to the β-type E2 but is rather conserved among all HPVs. The α-type HPV E2 proteins yet exhibited greatly reduced interaction levels with the SR proteins of this family, which could be the result of the

presence of only short R-alternative sequences in their hinge regions instead of the long SR repeats in those of the β-type HPV E2 proteins. This could explain why these interactions have not been identified so far. As initially proposed for the E2 protein of HPV5 [257], interaction with SR-proteins might increase the local concentration of splicing factors near the newly synthesized viral transcripts, thereby facilitating the timely regulation of viral mRNA production. Overall, this targeting confirms the predominant effect of the E2 proteins from the β-type HPVs on regulation of mRNA but also expands this to the α-type HPV E2 proteins and therefore indicates that regulation of gene expression at the level of RNA is conserved among all HPVs.

$\diamond$ Regulation of apoptosis. Regulation of apoptosis emerged as a functional family targeted by E2, which indicates that the E2 proteins have an intrinsic capacity to interact with apoptosis regulators. This targeting is a common characteristic of viral proteins since manipulation of cell death and survival is essential during viral infections [321]. Among the targeted proteins of this family, three were bound by all E2 proteins (CASP8, TAX1BP1 and TP53). For TP53 and CASP8, the binding of E2 may not have similar functional consequences according to the HPV genotype. Indeed, the detection of an interaction between p53 and the E2 proteins of all genotypes was unexpected since previous studies described p53 binding as a characteristic of the HR-mucosal E2 proteins and demonstrated that the induction of p53-dependent apoptosis by E2 was specific to HR-HPVs [263]. We found here by HT-GPCA that the binding of p53 is actually conserved among HR and LR HPV genotypes, but it is possible that the interaction with the LR-HPV E2

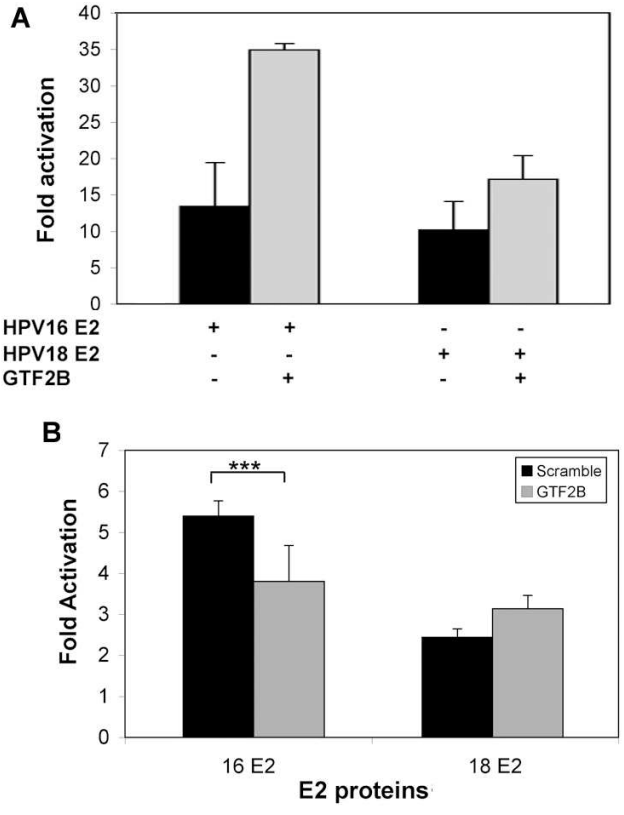

Figure II.9: **Functional impact of the specific interaction between GTF2B and HPV16 E2**. (A) HeLa cells were transfected by pTK6E2BS-Luc reporter and HPV16 or HPV18 E2 expression plasmids. Where indicated, GTF2B was added. Fold activation is given relative to TK6E2BS-Luc in the absence of E2. (B) HeLa cells were transfected with a pool of four siRNA targeting GTF2B or control siRNA (Scramble). 48 h post silencing, pTK6E2BS reporter plasmid was transfected along with E2 expression plasmids. Results are given as a fold activation relative to TK6E2BS basal activity in the presence of the same siRNA. Experiments were performed in triplicate with each bar representing the mean +/- SD. The stars (***) indicate a statistically significant difference between fold activation by 16E2 with a scramble siRNA or a GTF2B-directed siRNA (p-value $< 0{,}001$).

might alter a different p53 activity such as the induction of a cell cycle arrest. This notion is further supported by the detection of interactions between Caspase 8 and all the E2 proteins tested. Yet, it was previously shown that among mucosal HPVs, only HR-HPV E2 proteins are able to induce apoptosis through direct binding to caspase 8 [262, 268]. Here again, the interaction between CASP8 and the LR-HPV E2 proteins could be involved in a different mechanism and in particular in keratinocyte differentiation, since a role of caspase 8 in skin homeostasis has been detected [322]. Another hypothesis is that this interaction takes place at different times of the HPV life cycle. Indeed, E2-mediated caspase 8 activation is linked to the accumulation of E2 in the cytoplasm due to nucleo-cytoplasmic shuttling, which does not occur with LR-mucosal HPV E2 [262]. It can be envisioned that LR-HPV E2 proteins induce apoptosis only at late stages of the viral life cycle in the upper layers of infected epithelia where nuclear membrane breakdown naturally occurs as part of the keratinocyte differentiation program. This would allow the accumulation of E2-Caspase 8 complexes in the cytoplasm, enabling caspase activation and triggering cell death. Such delayed apoptotic induction could be involved in common viral processes such as viral particle release but could have escaped detection in cell culture conditions. For both p53 and caspase 8, the functional impact of interactions with the non-apoptotic E2 proteins could also be the opposite of interactions with the proapoptotic E2 proteins, *i.e.* interfering with death induction.

Overall, the functional targeting of apoptosis-regulatory factors shows that the E2 proteins actively take part in the manipulation of cell death and/or survival pathways, tightly orchestrated by the virus throughout infection.

◇ Ubiquitination. Controlling or rewiring the ubiquitin-proteasome pathway by viral proteins is likely to have a great impact on the host cell. The consequences could be multiple: degradation of antiviral molecules, diverting the substrate specificity of ubiquitin ligase... Previous studies have identified a targeting by E2 of several proteins involved in the regulation of ubiquitination processes, in particular of cullin-based E3 ligase complexes. As discussed in the introduction chapter, this targeting is primarily mediated by interactions with substrate adaptors. The comparative interaction mapping conducted here reinforces the notion of preferred targeting of E3 ubiquitin ligases by the E2 proteins. Indeed, out of the 16 factors of the ubiquitin-proteasome pathway found to interact with E2, eight are part of E3 ubiquitin ligase complexes and in particular substrate adaptors. Indeed BTBD1 and SPOP are BTB-proteins forming complexes with cullin3; FBX022, is a F-box protein subunit of cullin1-based complexes; and CDC20 and CDH1 are part of the APC/C. Binding of these two APC/C subunits was described as specific to HR-HPV E2 proteins [279], which was recovered for CDC20 among the mucosal HPVs. This binding was however also detected with cutaneous E2 proteins, suggesting that the involvement of the interaction with CDC20 in HPV pathogenesis could differ according to HPV tropism. Of note, the SPOP adapter binds all E2 proteins, but with a reduced efficiency for HPV16 E2. Conversely, the HPV16 E2 protein was the only one to bind another adapter of cullin3-complexes, BTBD1, suggesting that the targeting of E2 to cullin3-based complex is conserved but could be mediated through diverse interactions with adaptors. In addition, two HECT domain family ubiquitin ligases, HUWE1 and WWP2, were identified as novel E2 partners, further broadening the potential impact of E2 through interaction with ubiquitin ligases. This is compatible with the notion that

E2 could not only modify the action of ubiquitin ligases toward their natural substrates, as was shown for cyclin B or SKP2, but also could divert their specificity toward new substrates, the E2 proteins then acting as novel substrate adaptors. In that case, the targeting of ubiquitin ligase complexes by E2 would have a similar outcome than the binding of E6 to E6AP, which induces the hijacking of E6AP activity toward ubiquitination of p53 [323].

◇ Intracellular trafficking. The most unexpected functional family that has emerged from this analysis is composed of proteins involved in vesicle-mediated transport, affecting the dynamics and maintenance of intracellular membranous organelles such as endosomal and lysosomal transport, vesicle traficking to and from the Golgi, (**Fig II.10**).

Only one factor, VPS52, was bound by all the E2 proteins. Conversely, most of the E2 proteins interacted with several factors of this family, which highlights both a conserved and diversified targeting of intracellular trafficking factors by the E2 proteins. This brings the notion that such targeting is likely to underlie novel activities associated with E2. Given that E2 is primarily a nuclear transcription/replication factor, the targeting of cellular proteins involved in intracellular trafficking is a surprising aspect of our results. We therefore wished to validate a subset of the identified interactions by colocalization studies. We first focused on the shared target of all E2 proteins, VPS52 (Vacuolar Protein Sorting 52), a protein involved in vesicle trafficking from endosomes to the trans-Golgi network. When ectopically expressed in HaCat cells, VPS52 is distributed in vesicles as previously described [324]. Upon co-expression with 16, 18 or 39 E2, a consistent colocalization of VPS52 and E2 could be observed reflecting the conserved interaction detected by HT-GPCA (**Fig II.11**). Moreover, VPS52 vesicles clustered in a perinuclear region specifically in the presence of 16E2, which in turn massively redistributed in these vesicles. Such a drastic effect is in good agreement with the HT-GPCA interaction data where the interaction between VPS52 and HPV16 E2 is the highest of all interactions detected with this cellular protein.

We also studied the colocalization of CLTA (clathrin light chain) with HPV9 E2, the only E2 protein found by HT-GPCA to interact with this cellular protein. In the presence of HPV9 E2, CLTA is found not only in the cytoplasm but also in the same

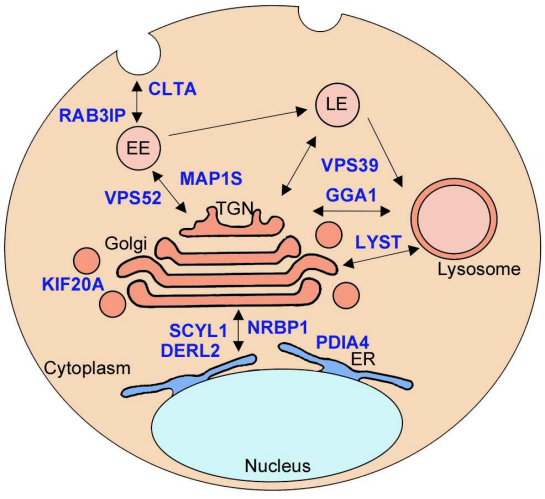

Figure II.10: **Proteins identified in the intracellular transport family**. Schematic representation of the function of the E2-targeted cellular proteins. Out of the 15 proteins constituting this functional category, most of them have a role in vesicular transport and are schematically represented between different cell compartments. The protein names are represented in blue. EE (early endosome), LE (late endosome), TGN (trans-golgi network), ER (endoplasmic reticulum).

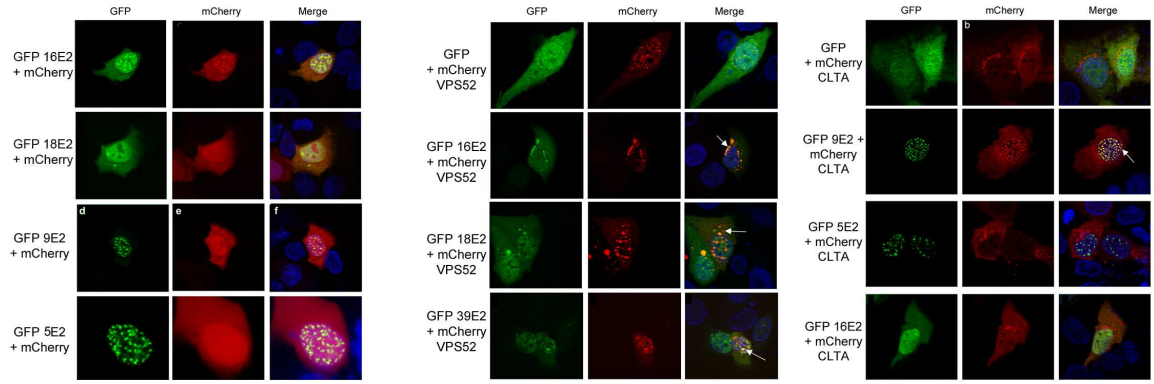

Figure II.11: **Colocalization between several E2 proteins and VPS52 and CLTA, two proteins of the intracellular trafficking family**. Left panel: HaCaT cells were cotransfected with expression plasmids for the indicated GFP-E2 proteins with an empty mCherry plasmid to assess their natural localization. Middle panel: HPV16, 18 and 39 E2 proteins in fusion with GFP were co-expressed with mCherry-VPS52. Right panel: HPV9, 5 and 16 E2 proteins in fusion with GFP were co-expressed with mCherry-CLTA. After fixation, the cells were subjected to fluorescence microscopy after counterstaining of the nucleus with DAPI. White arrows point to strong colocalization signals.

nuclear structures as E2. Neither the closely related E2 protein from HPV5, nor the E2 protein from HPV16 had the same effect on CLTA distribution, corroborating the interaction data obtained by HT-GPCA.

The colocalization studies therefore substantiate the targeting by E2 of proteins involved in cellular trafficking processes. In addition, they uncover several effects of E2 binding on the distribution of targeted factors, as observed with the relocalization of CLTA into the nucleus when interacting with HPV9 E2. It suggests that E2 proteins may have a strong impact on the overall intracellular trafficking mechanisms, whose biological significance will clearly require further investigation. Several hypotheses can be formulated concerning the implications of this targeting by the E2 proteins. First, proteins of this family are more concentrated around the Golgi apparatus, and the Golgi is known to be central in the translocation of processed viral antigens to and from the cytoplasmic membrane. It might thus suggest that, through this targeting, E2 could alter antigen presentation in infected keratinocytes. However the most striking feature of the proteins of this functional family is that they overlap with the pathway followed by viral particles to enter the cell and be transported through the cytoplasm. Our hypothesis is that targeting of this family of proteins reflects an involvement of E2 in the very early stages of infection and in the translocation of the viral genome to the nucleus. Indeed, E2 is able to link the viral DNA, but also the minor capsid protein L2 [115], which is pivotal for the trafficking of viral genomes from the capsid to the nucleus [325]. In addition, recent papers showed that viruses are transported by the retrograde pathway to the Golgi complex and the Trans-Golgi Network during HPV entry, further substantiating that the targeting by E2 of proteins involved in this mechanism could play a role in viral entry [326, 327]. A role of E2 in the early steps of infection would thus imply that the E2 protein

is encapsidated in the virion bound to the viral genome. Pseudovirion systems are for long used to understand viral entry and they indicate that E2 is dispensable for both the pseudovirion production and for the proper delivery of transducted DNA materials to the nucleus, in contrast to L2 [328]. However, only sparse information is available on the impact of E2 in pseudovirion's infectivity, and these studies were addressed only with BPV1 pseudovirions. In the BPV1 system, a study reported that E2 enhanced encapsidation of full-length viral DNA and may be packaged within the pseudovirion [329]. These findings however, were not corroborated in a more recent study [328]. In light with the emergence of the functional targeting of intracellular transport pathways by a number of different E2 proteins outlined in this work, we feel that this issue should be re-evaluated with HPV pseudovirions. We wonder whether the use of over-optimized systems for pseudovirion production (codon optimized ORF for example) could have hidden a potential involvement of E2 under normal conditions. Also, it should be taken into consideration that the pseudovirion experimental systems allow the efficient delivery only of DNA molecules smaller than the normal length of viral episomes. A potential role of E2 could favor the encapsidation of full-length viral DNA which could be uncovered using reporters of 8 kb containing E2-binding sites thereby mimicking authentic HPV genome.

Overall, these observations raise an intriguing possibility that E2 could take an active part in the very early steps of HPV infection, which we feel is worth investigating.

Overall our comparative approach allowed the identification of numerous new cellular partners of multiple HPV E2 proteins, and gave an overview of the potential functions targeted in the infected cell. Our strategy provided an unbiased mapping of E2-host protein-protein interactions through the identification of binding partners of the E2 proteins from a large panel of HPV genotypes including those that represent a lower public health concern. This study offers a unique opportunity to compare E2-host interaction profiles according to the HPV tropism and pathogenic potential. E2 interaction dendrograms correlated with HPV phylogeny, providing evidence of the contribution of E2 in HPV pathogenesis. Such contribution, however, does not rely on the targeting by E2 of different cell functions but our study rather highlights that functional targeting by E2 is achieved through a complex and variable interplay with the host cell proteins, and could potentially influence the HPV pathogenic power. Activities emerging from the E2-host interaction network corroborated the essential role of E2 in the control of gene expression, primarily through regulation of transcription but also through regulation of RNA processing. In addition, it emerged that E2 broadly impacts on the cell physiology through targeting of apoptosis, ubiquination and intracellular trafficking. These E2 activities could be essential to make the epithelial cells conductive for the productive viral cycle. In conclusion, this study constitutes a framework for future functional investigations and provides a solid basis to understand the implication of E2 in HPV pathogenesis.

B - Specific HPV16E2-CCHCR1 interaction

a . Context

One particular question that is recurrently asked in the field of HPV clinical studies is how to distinguish between genital lesions that will progress toward cancer from those that won't. There is a critical need to identify predictive markers in order to anticipate which women have a significant risk to develop high grade lesions and would therefore greatly benefit from early treatments. In particular, HPV16 is the most prevalent HR-HPV, accounting for over 50% of cancerous lesions. Understanding the differences between HPV16 pathogenesis and other HPVs could be the key to understand the particular properties that make it such a public health concern.

Contrary to most studies, the comparative approach developed during this thesis allows the assessment of interactions across multiple HPV genotypes against a common set of cellular proteins. This is thus a strategy of choice to identify specific interactions and indeed we were able to extract from the comparative interaction datasets several cellular proteins differentially bound by the mucosal HR-HPV and LR-HPV E2 proteins, which could be potentially interesting to use as biomarkers of the oncogenic HPVs. In particular, we have identified an interaction between the E2 protein of HPV16 and a cellular protein called CCHCR1 (Coiled-Coil alpha HeliCal Rod protein 1). The CCHCR1 gene is located in the major psoriasis susceptibility locus [330] and encodes a 782 amino acid long protein suggested to play a role in the balance between keratinocyte proliferation and differentiation [299, 331]. In particular, CCHCR1 has been associated with the complex regulation of basal keratinocyte proliferation, either as a negative regulator in mouse models [331] or rather activating proliferation in the context of cancerous cell lines [299]. It thus appears that the regulation of proliferation by CCHCR1 depends on the cell context and would impact on the early switch between differentiation and proliferation of epithelial cells. The last part of my PhD work has been devoted to further characterize the interaction between CCHCR1 and E2 from HPV16.

b . Characterization

CCHCR1 was isolated 13 times with HPV16 E2 in the initial yeast two-hybrid screenings described in the previous section. The cDNA sequences collected in the yeast clones corresponded to a unique sequence spanning from amino acid 137 of the full-length sequence (at the beginning of the first coiled-coil) to the stop codon. Since the CCHCR1 expression profile is highly complex, with a plethora of differently spliced mRNA, this isoform must be the one naturally expressed in HaCaT cells. An interaction between HPV16 E2 and CCHCR1 had already been identified in a yeast two-hybrid study in 2008 by Olejnik-Schmidt and colleagues [219], but this study was only conducted with HPV16 E2 so it was not informative regarding the specificity toward HPV16. We therefore decided to test CCHCR1 in HT-GPCA against the 12 E2 proteins included in our comparative study. This first allowed the validation of its interaction with HPV16 E2, but also enabled us to assess the specificity of this interaction across various HPV types (**Fig II.12A**). The profile of interaction shows that except for HPV16 E2, none of the mucosal HPV

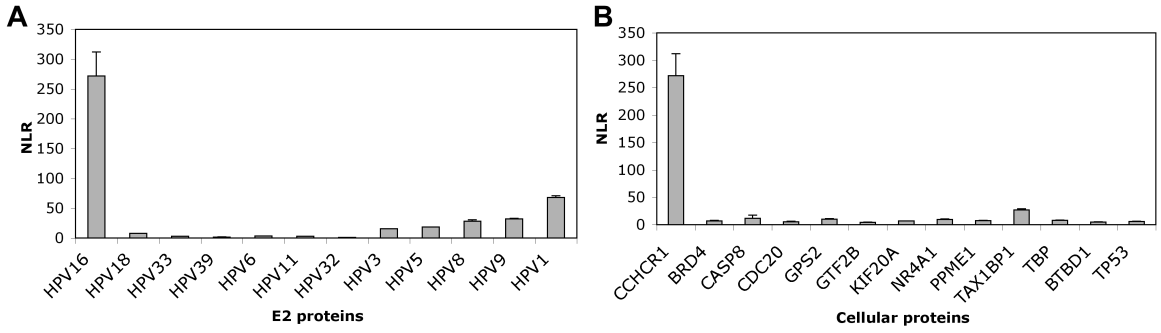

Figure II.12: **Specific HPV16E2-CCHCR1 interaction**. (A) Interactions between CCHCR1 and 12 E2 proteins from a panel of HPVs selected to be representative of their diversity. (B) Interaction between HPV16 E2 and a panel of cellular proteins known from the literature to interact with HPV16E2. Interactions were measured by HT-GPCA and errors bars represent Standards Errors to the Mean (SEM).

E2 protein interacted with CCHCR1, while E2 from the cutaneous HPV exhibited some level of binding, which however appears only marginal when compared to the interaction with HPV16 E2. We thus show that the interaction between CCHCR1 and E2 is specific to the HPV16 genotype.

In addition, the interaction between HPV16 E2 and CCHCR1 generated by far the strongest interaction level in the entire interaction profile obtained with this E2 protein by HT-GPCA (NLR above 250). This is exemplified in **Fig II.12B**, where the interaction of HPV16 E2 with a panel of Gold Standards is presented and shows that all interactions tested are weaker than the level observed with CCHCR1, even though they appear positive (NLR above 3.5). These results thus indicate that the interaction between HPV16 E2 and CCHCR1 is both highly specific to this HPV type and the strongest interaction identified so far in our conditions for this E2 protein.

We next constructed mutants of the HPV16 E2 protein to further characterize the interaction interface (**Fig II.13**). Serial deletions of HPV16 E2 N-terminal helices indicated that as soon as the first helix was deleted, the interaction was completely annihilated (**Fig II.14**, left panel). This is similar to the interaction between E2 and BRD4, which is known to span over the three α-helices in E2's N-terminal domain. In contrast, the deletion of all three helices did not substantially impact on the binding to TAX1BP1, thereby confirming the integrity of the expressed mutants.

Point mutants were next generated to determine on which side of the HPV16 E2 α-helices the interaction with CCHCR1 occurs (**Fig II.14**, right panel). We studied by HT-GPCA the interaction of CCHCR1 with HPV16 E2-R37A and HPV16 E2-I73A mutated at amino acids located on one side of the surface formed by the N-terminal helices and which are pivotal amino acids for BRD4 binding; and HPV16 E2-E39A where the mutated amino acid is exposed at the opposite helice surface, and is essential for the binding of the viral helicase, E1. The mutation of E39 had no effect on HPV16 E2 binding to CCHCR1. In contrast, mutations R37A and I73A drastically inhibited the interaction with CCHCR1, as well as with BRD4 as expected. TBP, used here as a control since

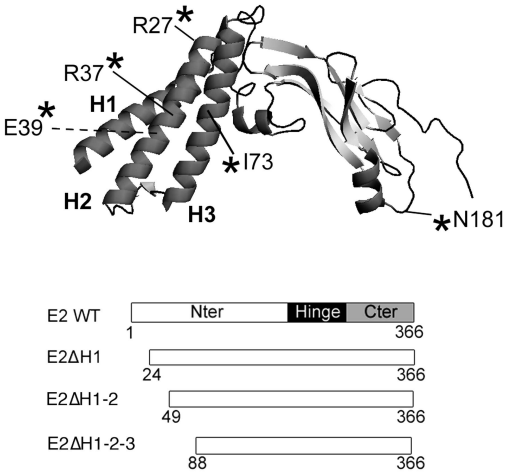

Figure II.13: **Mutation of the HPV16 E2 protein**. Schematic representation of the different mutations and deletions introduced into the HPV16 E2 protein. Top: 3D representation of HPV16 E2 N-terminal domain. The five point mutations are represented by a star. The amino acid E39 faces the other side of the helix and is therefore represented with a dashed line. H1, H2 and H3 designate the three α helices. Bottom: Representation of the serial deletion of HPV16 E2's α-helices.

it interacts with the C-terminal domain of E2, consistently bound all mutated HPV16 E2 proteins. These mutants revealed that CCHCR1 interaction interface in HPV16 E2 N-terminal domain is localized on a surface overlapping that of BRD4 binding. Keeping in mind that this interaction is specific for HPV16, we also mutated the amino acid R27 which is an arginine only in HPV16 E2 and is exposed on the same side as R37 and I73. However, this mutation had only little effect on the interaction between HPV16 E2 and CCHCR1. Also, we noticed that HPV16 E2 exhibited an asparagine at position 181 whereas a threonine was conserved among all other E2 proteins at the equivalent position. We reasoned that the binding interface of CCHCR1 could extend over a large area of HPV16 E2 N-terminal domain, and we therefore tested whether this amino acid could contribute to the interaction despite being localized outside the helical part of E2's

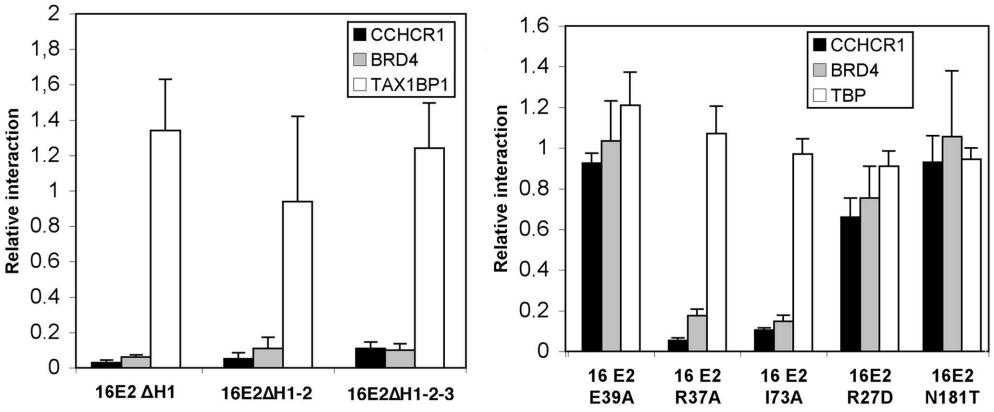

Figure II.14: **HPV16E2-CCHCR1 interaction interface mapping**. Left: Interactions between cellular proteins and HPV16 E2 N-terminal alpha helices deletion mutants. H1, H2, H3 are respectively the first three alpha helices and the symbol Δ symbolizes a deletion. BRD4 and TAX1BP1 are used as a controls. Right: Interactions between cellular proteins and HPV16 E2 N-terminal point mutants. BRD4 and TBP are used as a controls. Interaction are measured by HT-GPCA and errors bars represent Standards Errors to the Mean (SEM).

N-terminal domain. Mutating this asparagine into a threonine to mimic other E2 proteins did not affect the binding of CCHCR1 to HPV16 E2, therefore amino acids R27 and N181 are unlikely to participate in the specific interaction between CCHCR1 and HPV16 E2. One aspect, however, remains unclear: the interaction between BRD4 and E2 is conserved among all HPVs whereas the interaction with CCHCR1 is specific to E2 from HPV16. It is thus paradoxical that both factors share a common binding interface on HPV16 E2. The N-terminal domain is well conserved among the E2 proteins and, so far, we failed to identify a particular amino acid responsible for this specificity. One possibility is that this interaction requires the dimerization of E2 N-terminal domain, which could support the specificity toward HPV16. Indeed, in 2000, Antson et al. [195] resolved the structure of the N-terminal domain of HPV16, and showed that this domain by itself associates into dimers in solution, which they claim could be important for interactions with viral and cellular proteins. They further showed that substitution of a number of amino acids leads to dimer disruption, among which amino acids R37 and I73, that we demonstrate here as being crucial for the interaction with CCHCR1. By contrast, the crystal structure the N-terminal domains of E2 from HPV11 and HPV18 was determined to be fully monomeric, with no dimers formed in solution [332, 333]. Therefore, only the N-terminal domain of HPV16 E2 seems to be able to dimerize, even though it requires conserved amino acids. It is thus conceivable that the binding specificity of CCHCR1 could be brought by a particular surface on the dimerized HPV16 E2 N-terminal domain, and by specific amino acids motifs scattered over each monomer.

Given that we demonstrated that the binding surface of CCHCR1 on HPV16 E2 overlaps with that of BRD4, we wanted to determine if they could interfere with each

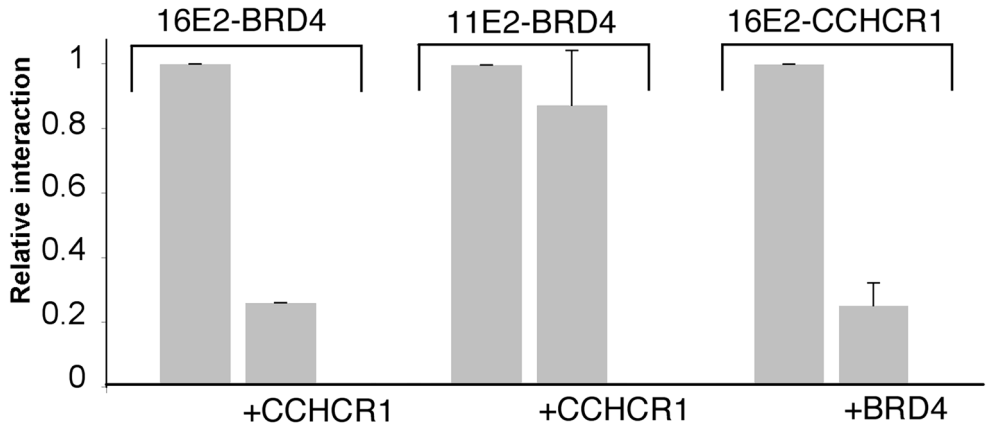

Figure II.15: **HPV16E2-CCHCR1 interaction competition**. Interaction between HPV16 E2 and BRD4 (left), HPV11 E2 and BRD4 (middle) or HPV16 E2 and CCHCR1 (right) was tested by HT-GPCA in the presence or not of a challenging protein: CCHCR1 (left and middle panels) or BRD4 (right). Results are reported to the NLR value in cells transfected by an empty plasmid instead of the challenging protein. Errors bars represent Standards Errors to the Mean (SEM).

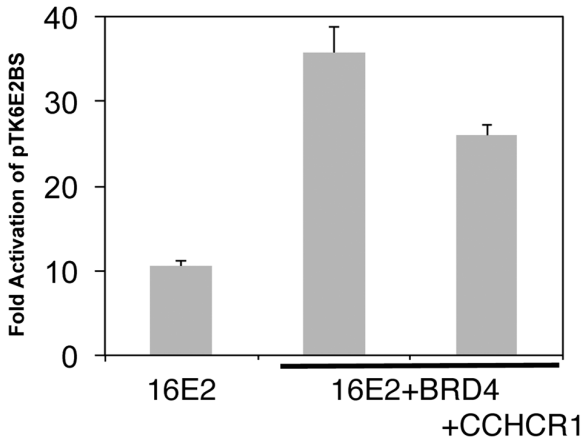

Figure II.16: **Effect of CCHCR1 on BRD4-mediated enhancement of E2 transcriptional activation**. HaCaT cells were transfected with a reporter plasmid (pTK6E2BS) containing six E2 binding sites along with HPV16 E2 and BRD4 and with or without CCHCR1. Luciferase signals were measured 30h later. Fold activation are relative to the experiment without E2. Errors bars represent SEM (Standards Errors to the Mean).

other. We adapted the HT-GPCA technique used to detect pairwise interactions by adding a third partner as a "challenger". If the third partner competes with the interaction assessed by HT-GPCA, it should decrease the interaction ratio. We started by looking at the interaction between HPV16 E2 and BRD4 and added flag-tagged CCHCR1. As shown in **Figure II.15** (left), addition of CCHCR1 induces a 5-fold decrease of the interaction between BRD4 and HPV16 E2. When the same experiment was performed with HPV11 E2 (**Figure II.15**, middle), known to interact with BRD4 but not with CCHCR1, the addition of CCHCR1 had only little effect on the BRD4-HPV11 E2 interaction, proving that it is indeed the interaction of CCHCR1 with HPV16 E2 that interferes with BRD4 binding. The reciprocal experiment consisted of adding flag-tagged BRD4 to the HPV16 E2-CCHCR1 interaction mix (**Figure II.15**, right) revealed that BRD4 interferes with this interaction as well. Taken together, these results indicate that there is a competitive binding between BRD4 and CCHCR1 for interaction with HPV16 E2.

BRD4 has a pivotal role in E2's transcriptional function, we therefore hypothesized that the competition with CCHCR1 could affect this function. The effect of CCHCR1 on E2-dependent transcription was assessed using the luciferase reporter system described in the previous section, pTK6E2BS-luc (**Fig II.16**). E2 alone activates transcription of this synthetic promoter of about 10 fold, which is within the range previously reported [213, 242, 290]. In the presence of BRD4, E2 transactivation was enhanced by three fold, in good agreement with previous reports [334]. However, when co-expressed with CCHCR1, the effect of BRD4 on E2 transcriptional activity was reduced. These results show that CCHCR1 binding to E2 functionally interferes with BRD4-mediated enhancement of E2 transactivation. Such functional interference probably results from the binding competition between BRD4 and CCHCR1 on HPV16 E2. It also indicates that, even though CCHCR1 associates with E2 in a similar manner as BRD4, it cannot restore an equivalent transcriptional function when in complex with HPV16 E2.

c .Consequences on E2's cellular localization

To further characterize at the cellular level the CCHCR1-HPV16 E2 interaction, we conducted fluorescent studies using GFP-HPV16 E2 fusions and CCHCR1 fused to a monomeric cherry fluorescent tag (**Fig II.17**). CCHCR1 is expressed as cytoplasmic dot-

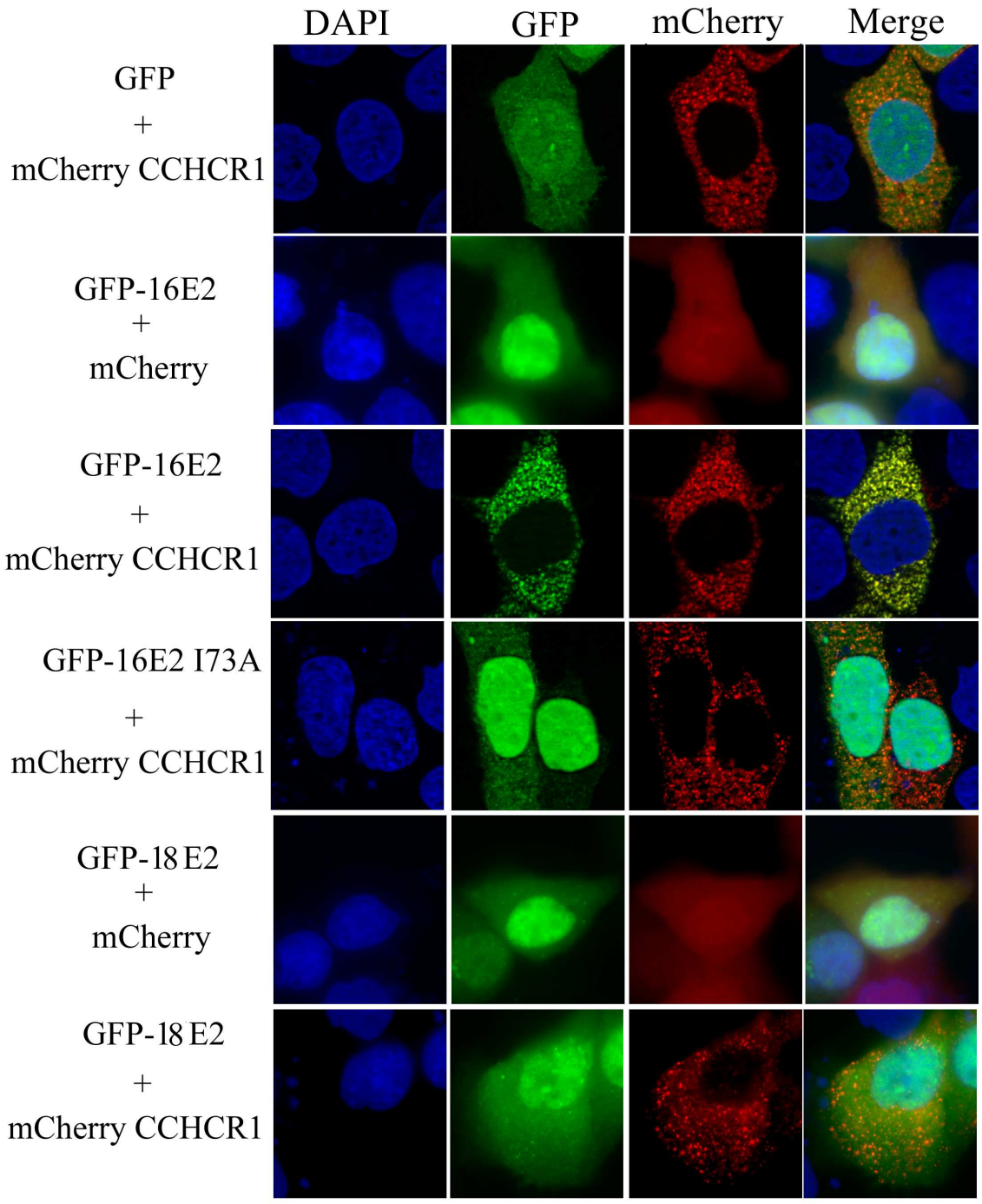

Figure II.17: **The effect of the HPV16E2-CCHCR1 interaction on E2's cellular localization**. From top to bottom : in HaCaT cells, ectopic expression of CCHCR1 shows punctuate staining (red) in the cytoplasm while HPV16 E2 displays a diffuse pattern both in the nucleus and in the cytoplasm. When co-expressed, the two protein signals overlap showing a strong delocalization of HPV16 E2 in the cytoplasmic dots typical of CCHCR1 expression. Neither the non-interacting E2 proteins 16E2 I73A nor HPV18 E2 show a delocalization from the nucleus upon CCHCR1 expression.

like structures corroborating other studies [298]. HPV16 E2 displays a diffuse pattern of expression primarily concentrated in the nucleus, which also confirms previous work [262]. When the two proteins are co-expressed, we observe a drastic change in the expression pattern of E2. Indeed, HPV16 E2 is now located in the same cytoplasmic dot-like structures as CCHCR1 where the two proteins perfectly colocalize. Concomitantly, a result of E2 being trapped in the cytoplasm is that it is massively delocalized from the nucleus which could therefore potentially affect its nuclear functions.

None of the non-interacting E2 proteins from HPV18 and HPV16 I73A mutant were delocalized from the nucleus upon CCHCR1 co-expression, indicating that the redistribution of HPV16 E2 is strictly correlated with its ability to bind to CCHCR1 (**Fig II.17**). By docking most of HPV16 E2 protein into the cytoplasm instead of the nucleus, this relocalization potentially affects the nuclear functions of E2. It therefore reinforces the negative effect of CCHCR1 on E2's transcriptional properties. At this point, however, we do not know if relocalization of E2 to the cytoplasm by CCHCR1 could promote E2's cytoplasmic function.

As discussed above, there is a competition between BRD4 and CCHCR1 for interaction with HPV16 E2. The interaction between E2 and BRD4 occurs in the nucleus and consequently does not affect the natural localization of E2 in contrast with the interaction between CCHCR1 and HPV16 E2 (**Fig II.18**). The delocalization of E2 into

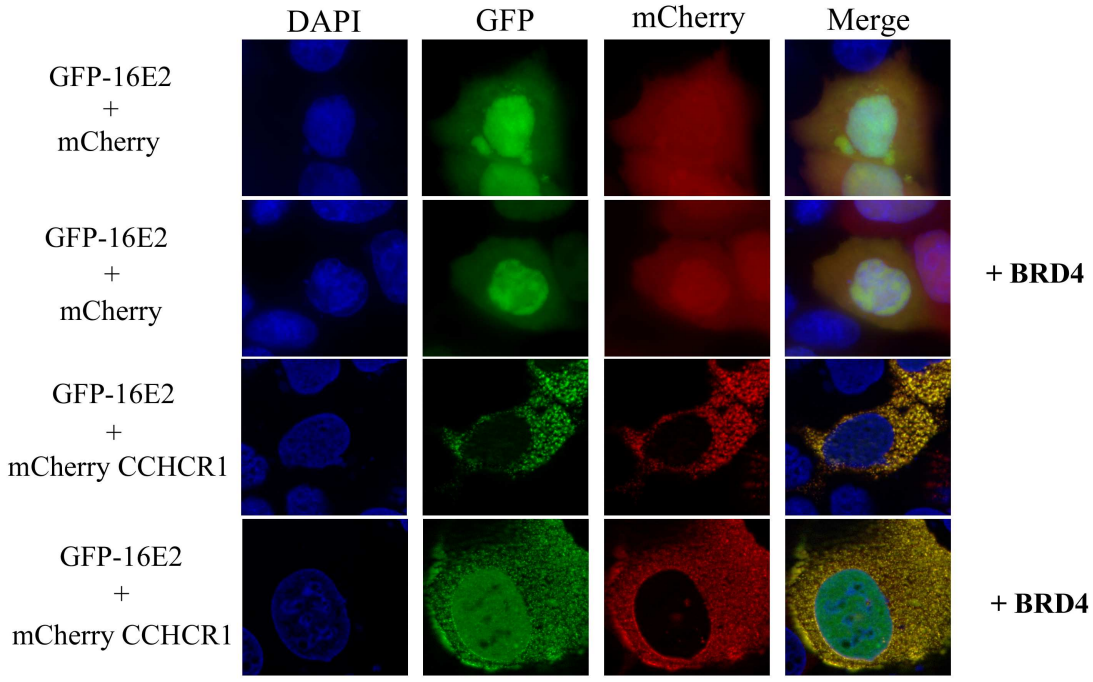

Figure II.18: **The BRD4-CCHCR1 competition**. HaCaT cells were transfected with GFP-HPV16 E2 and Ch-CCHCR1 or 3xF-BRD4 or all together. CCHCR1-induced total delocalization of E2 was less drastic in presence of BRD4 where E2 shows a pattern of expression both in the nucleus and in the cytoplasmic dot where it colocalizes with CCHCR1 showing that both 16E2-BRD4 and 16E2-CCHCR1 interactions co-exist in the cell.

the cytoplasm by CCHCR1 might therefore render E2 less accessible to interact with BRD4. Actually, when the three proteins are co-expressed (16 E2, BRD4 and CCHCR1; **Fig II.18**), E2 displays an intermediate pattern of expression: its natural pattern with a nuclear staining but also within the cytoplasmic dots colocalizing with CCHCR1.

This observation probably indicates that the interference between BRD4 and CCHCR1 for interaction with HPV16 E2 is mediated both by competitive binding to the same surface and by taking place in different cell compartments. It also suggests that both interactions - E2-BRD4 and E2-CCHCR1 - can occur at the same time in the cell.

d .Functional impact on keratinocyte differentiation

To get insights into the consequences of this interaction on the HPV16 life cycle, we focused our attention on the differentiation process of keratinocytes. Keratinocyte differentiation is a multi-step process that relies on the tightly regulated and sequential expression of a number of genes such as keratins, transglutaminase 1 or involucrin, which can be used as differentiation markers.

CCHCR1 is a factor that has been linked to the regulation of the switch between proliferation and differentiation. Given that this switch constitutes a key process hijacked by HPVs, it prompted us to pursue the study of a potential impact of this interaction on keratinocyte differentiation. To that aim, we looked by qRT-PCR at the mRNA expression levels of several differentiation markers chosen to be hallmarks of the differentiation process of keratinocytes: Keratin 14 (K14) as a marker of proliferation, Keratin 10 (K10) for early differentiation and Transglutaminase 1 (TGM1) for late differentiation (**Fig II.19**).

HaCaT cells were cultured in low-calcium containing media, a culture condition that allows the cells to be able to trigger differentiation while remaining proliferative [336, 337].

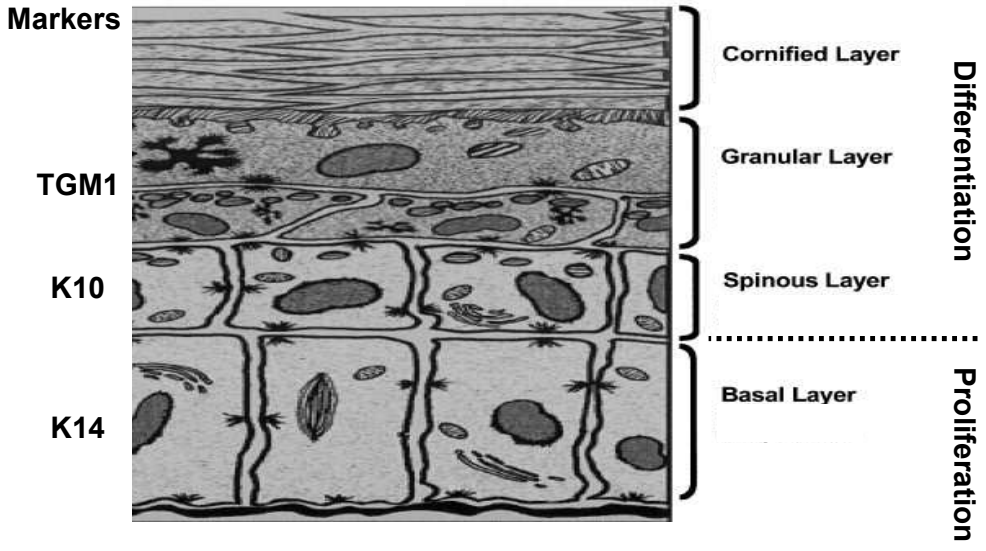

Figure II.19: **The epithelial different layers** Schematic representation of layers of the epithelium. The differentiation markers used in this study are pictured in their corresponding layers. Adapted from Bikle *et al.* [335].

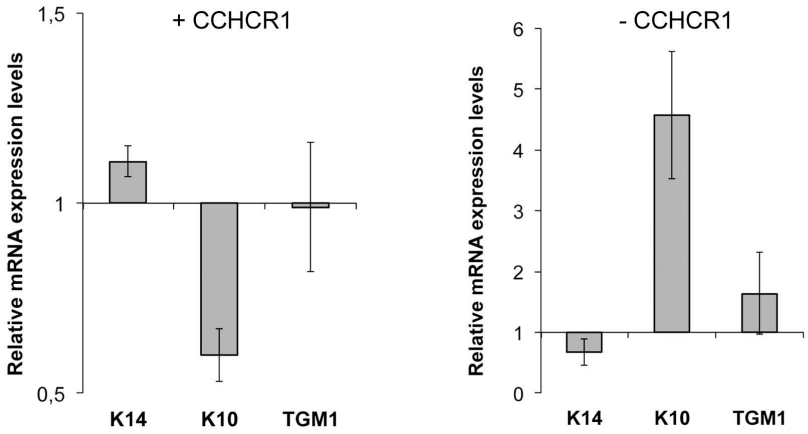

Figure II.20: **Effect of CCHCR1 on differentiation marker expression**. CCHCR1 was either overexpressed (left panel) or inhibited by siRNA (right panel) and the subsequent effect on the mRNA levels of three differentiation markers was monitored by qRT-PCR. Results are represented as relative to mock transfected cells (empty plasmid for overexpression or siSCRAMBLE for siRNA treatment). Error bars represent Standards Errors to the Mean (SEM).

First we wished to understand the effect of CCHCR1 in our conditions. As shown in **Figure II.20**, left panel, ectopic expression of CCHCR1 leads to a weak activation of K14 and to a repression of K10, while showing no effect on TGM1 expression. This is in good agreement with a previous report stating that CCHCR1 mainly stimulates proliferation of keratinocytes in the context of cell lines [299]. These results were confirmed by the reciprocal experiment using a siRNA to knock down CCHCR1, where K14 expression is repressed and K10 is activated (**Fig II.20**, right panel).

To characterize the effect of E2 on differentiation in our experimental system, HaCaT cells were infected by recombinant adenoviruses expressing GFP-E2 or GFP only. HPV16 E2 strikingly enhanced the expression of the early differentiation marker K10 and induced a drastic increase of its mRNA levels of more than 35 times (**Fig II.21A**). In these conditions, HPV16 E2 had either no or only a weak effect on the expression of K14 and TGM1, leading to the conclusion that HPV16 E2 has a strong impact on early differentiation. In comparison, the E2 protein from HPV18 had only a minor effect on K10 expression (**Fig II.21B**). Since the strong activation of K10 expression is specific to the E2 protein of HPV16, we wonder wether it might have a link with its interaction with CCHCR1. We thus conducted co-transfection experiments to assess the effect of the 16E2-CCHCR1 interaction on K10 expression (**Fig II.21C**). Expressing HPV16 E2 using mammalian plasmids instead of recombinant adenoviruses reduced the transfection efficiency, but a significant activation of K10 can still be observed in these conditions.

When HPV16 E2 was co-expressed with CCHCR1, the effect of E2 on K10 expression was decreased 2.5 times, lowering the level of K10 activation to less than 2 fold. We first figured that it might be caused by the repressive effect of CCHCR1 on K10, but the simple combination of CCHCR1 negative and E2 positive effects on K10 mRNA levels would only result in a minor decrease of E2-induced K10 activation (dotted line), while the effect of the co-transfection is much more drastic. This thus suggests that CCHCR1

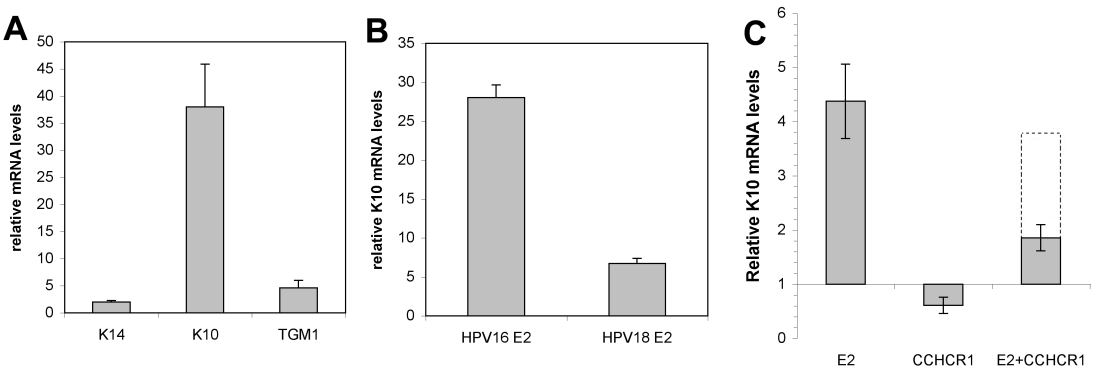

Figure II.21: **E2 and CCHCR1 effect on differentiation**. (A) HaCaT cells were infected by Ad-GFP16E2 and the subsequent effect on the mRNA levels of three differentiation markers was monitored by qRT-PCR. Results are represented as relative to experiments where cells are mock transfected (infection with Ad-GFP). (B) HaCaT cells were infected by Ad-GFP16E2 or Ad-GFP18E2 and the subsequent effect of the mRNA levels of K10 was monitored by qRT-PCR. Results are represented as relative to experiments where cells are mock transfected (infection with Ad-GFP). (C) HaCaT cells were transfected by HPV16 E2 and CCHCR1, either alone or together and the subsequent effect on the mRNA levels of K10 was monitored by qRT-PCR. Results are represented as relative to experiments where cells are mock transfected. The dotted line represents the expected effect on E2-mediated activation of K10 if we only take into account the direct repressive effect of CCHCR1 on K10. Errors bars represent Standards Errors to the Mean (SEM).

has a direct negative effect on E2-mediated activation of the expression of K10.

These results support a role of E2 in inducing differentiation. More importantly, they highlight that CCHCR1 negatively interferes with E2 functions on the activation of keratinocyte differentiation. The activation of K10 by E2 is likely to be transcriptional, and thus its inhibition by CCHCR1 would be related to the negative effect on E2's transactivation that we have demonstrated previously. In addition, we demonstrated that CCHCR1 induces a relocalization of E2 into the cytoplasm which could reinforce the repressive effect on K10 activation.

Overall our results indicate that CCHCR1 has a negative effect on HPV16E2 function by both physically interfering with the interaction of E2's major interactor BRD4 and by relocalizing E2 into the cytoplasm. They also suggest that, in the case of HPV16, both the transcriptional functions and the subcellular distribution of E2 might depend on the proportion of BRD4 and CCHCR1 present in keratinocytes along the differentiating epithelium, which may vary during the viral life cycle. This study also confirms a role of E2 in promoting keratinocyte differentiation, a mechanism otherwise impeded by the action of E6 and E7, therefore reinforcing the idea that E2 is an opposing force of the HPV oncoproteins. Given the repressive effect of CCHCR1 on HPV16 E2-mediated activation of K10 uncovered in this study, it can be envisioned that the specific interaction with CCHCR1 interferes with the induction of differentiation specifically for HPV16 E2 which

could therefore have a strong impact on HPV16's life cycle and/or pathogenesis.

III . Discussion

The HPV family is very wide and each member of this family has its own particularities: pathogenic and oncogenic potential, tropism specificity... To understand the mechanisms on which these particularities rely, it is important to switch from punctate studies to large-scale comparative approaches. The main focus of this thesis was the study of the viral regulatory protein E2. Provided that protein-protein interactions are the primary means for viral proteins to hijack and affect the host cellular signaling pathways, we decided to study E2 through its interaction network. Unlike most studies focusing on HPV16 and HPV18, the most clinically relevant HPVs, we decided to take advantage of the great diversity of the HPV family by carrying out a comparative study. To that aim, 12 different genotypes of HPVs covering their natural diversity were selected. In addition, most studies on E2's interactions focus on the role of E2 in transcriptional regulation. In our case, we did not want to focus our attention on one particular function of E2 by testing, for example, all the transcription factors of the cell to see whether they can interact with E2, but we rather wanted to objectively identify interactions engaged by E2. The best compromise for us to recover interacting partners of 12 E2 proteins in an unbiased approach was to perform yeast two-hybrid screenings. Genetic screens are an efficient way to probe at once potential interactions involving proteins expressed from cDNA libraries. In our case, we chose to use a cDNA library from HaCaT keratinocytes to be coherent with the natural HPV target cells during infections. We identified more than 200 cellular proteins interacting with at least one E2 protein. Only a low overlap of the interaction data was observed, meaning that most cellular proteins had been identified with only one of the E2 proteins included in the study. We therefore wondered if interactions identified with a subset of E2 proteins had not escaped detection with the others, given the acknowledged high false negative rate associated with yeast two-hybrid systems. This prompted us to go through a second-step validation of the results and to challenge cellular proteins recovered by yeast-two hybrid for interactions with the whole set of E2 proteins.

We decided to use the HT-GPCA as a validation technique. This assay is performed in mammalian cells, which seemed more physiologically relevant for the study of E2's interactions. It is based on the detection of luciferase signals and is therefore very sensitive and it is measured directly on lysed cells, without prior purification steps, so improving the detection of weak or transient interactions. Another main advantage is that it is adequate for a high-throughput format, which was required for this validation step including more than 1,400 interactions to test. However, a potential caveat of HT-GPCA could be that it involves the fusion of Gaussia fragments to the protein of interest, which could alter its natural folding, localization or function. We addressed this issue by verifying that all the E2 proteins fused to the Gaussia luciferase fragment were functional in a transactivation assay, thereby showing that they were properly folded and localized [338]. However, given the scale of the study, folding and function could not be verified for all the cellular

proteins and this constitutes a potential source of false negative results. Also, we noticed differences in the accumulation levels of the various E2 proteins [338], and this is likely to be true for the cell partners as well. Such disparity could obviously introduce a bias in the comparative interaction data. However, when addressing this important point, the results obtained showed that the expression levels are not directly correlated with the interaction rates between the E2 proteins as discussed earlier. In addition, in this comparative approach, assessing the interactions between a common set of viral and cellular proteins with the same methodology is considered, to some extent, to be free from such variability. Indeed, each interaction is estimated in the context of two interaction profiles: the interaction profile of a given cellular protein and that of a given viral protein, which helps assess the degree of specificity. Using interaction datasets as such a double-entry matrix greatly improves the ability to discriminate significant from non-specific interactions. For example, in the HT-GPCA assay, HPV32 E2 had an alarmingly low number of positive interactions, while in the yeast-two hybrid screens, this E2 protein yielded many of the new potential interactors. This clearly suggests that HPV32 E2 had an intrinsic difficulty in the HT-GPCA assay and that probably, the threshold for positive interaction specifically associated with this viral protein should be scaled down. It thus indicates that the number of binding partners for 32E2 detected here is likely underestimated. If we had studied interactions of HPV32 E2 in a more punctuate study rather than in the context of the identification of its interaction profile, we would have probably concluded that HPV32 E2 does not engage many interactions. Keeping the threshold for positive interactions at 3.5 allowed the consistent comparison of all E2 proteins, but probably resulted in the loss of certain interactions. If we were continuing the study of HPV32 E2, we would certainly take into consideration interactions detected with a NLR below 3.5 if they seemed functionally meaningful. This example illustrates the different levels of interpretation that can be relevant in such large-scale datasets. In my opinion, comparative approaches represent a major advance in interactomic studies.

Such large-scale studies are also efficient to broaden the scope of known interactions. For example, the study by HT-GPCA of the interactions between E2 proteins and the gold standards has greatly improved our knowledge on these interactions. Indeed, most of the known interactions of E2 involved HPV16, the most clinically relevant HPV from a public health point of view, while here, we tested these interactions with a large panel of E2 proteins. It allowed us to highlight that many of the interactions previously identified with one particular E2 protein are actually shared by several E2 proteins and it suggests that different E2 proteins often have similar targets. For example, some interactions, as with TP53 or caspase 8, had been detected in connection with E2 functions specific to a subset of genotypes, but were identified here as conserved across all HPV genotypes. An explanation for these results is that the HT-GPCA technique is more sensitive to detect weak interactions. A plausible hypothesis would be that, despite the conserved interaction, the functional impact of these interactions may be different according to the HPV genotype. These observations underline the importance of combining functional studies and interaction mapping to decipher E2 activities in relation with HPV pathogenesis. On the contrary, interaction data cannot always be generalized from functional studies addressing only a subset of E2 proteins. For example, the binding of GTF2B had been detected for both BPV1 E2 and HPV16 E2, and was assumed to be conserved and involved

in the activity of all E2 proteins [339]. However, binding to this factor turned out to be specific to HPV16 E2 and we showed that it accounts for some particular aspect of its transcriptional function. These results demonstrate the significant contribution of combining several methodologies to understand the panel of interactions mediated by viral proteins and the strength of our comparative approach to interpret interaction specificities.

In addition, another great advantage of such large-scale studies is that by comparing interaction profiles, global information about viral proteins can be extracted. An example of that is the observation that the interaction profiles of the β-type HPV E2 proteins are brighter than those of the α-type HPV E2. Brighter profiles reflect more interactions and stronger interactions. This led us to the conclusion that this difference in the interaction rate could be mediated by the β HPV E2 longer unfolded hinge regions since disordered regions were shown to be enriched in interaction motifs [314]. Therefore, this comparative approach provided us with an intrinsic difference between β and α E2 proteins that could result in drastic changes in the regulation of their respective viral life cycle.

However, to get more details from these types of large-scale datasets, it is required to go through more statistical studies. We therefore continued the analysis using hierarchical clustering. One striking feature of this analysis is that it is possible to distinguish, to some extent, between HR and LR HPVs by only looking at E2's interactions. Two hypotheses can be raised: infections take place in different cellular niches where the host proteome is different. The distinction between the interaction profiles of the E2 proteins would thus reflect the difference in the infection site by either a HR or a LR HPV. Or, the difference in the outcome of the HPV infection, either HR or LR, is the consequence of E2 interaction patterns. This would mean that E2 from the HR-HPVs has evolved a different set of interactions than the LR-HPV E2 proteins and this participates in the particular traits of a HR infection. We believe that the second hypothesis is correct since E2 has been shown to have autonomous functions that promote carcinogenic conversion, such as induction of genomic instability. We therefore think that the differences in E2's interaction patterns from an HPV to the other somehow influence the outcome of the infection.

The study of E2's interaction network by itself has also greatly enhanced the global understanding of E2's interplay with the host cell. The prominent targeting of proteins involved in regulation of transcription was comforting given the acknowledged primary role of E2 in these processes. Indeed, we started this study with an unbiased approach as described earlier, and still identified this preferential targeting, reinforcing both the robustness of the strategy and the reliability of the obtained interaction dataset. Once again, the comparative aspect of this analysis offers a way to extract particular features of different E2 proteins in transcriptional regulation and confirmed that E2's impact on transcription is both complex and diversified among different HPV genotypes.

We did not identify particular targeting of cellular proteins involved in replication, which probably indicates that E2's participation to this process mainly occurs through its interaction with the viral helicase E1.

The targeting of proteins involved in ubiquitination, RNA processing or apoptosis was already known for some of the E2 tested. However, for these families, we have demonstrated that the targeting was conserved for the entire panel of E2 proteins. This provided an experimental appraisal of the complex interplay engaged by the various E2 proteins.

Interestingly, we have identified a targeting of cellular proteins involved in the regulation of intracellular trafficking. The emergence of this unexpected functional family raises a provocative hypothesis that E2 may be involved in early steps of viral infection, which will need to be further studied. The first step to challenge this theory would be to show the presence of E2 in the viral particle. This might prove difficult if only four E2 proteins (one on each E2 binding sites in the LCR) are included in the viral particle. Also, if E2 is indeed involved in the translocation steps of the viral genome to the nucleus in the early stages of infection, E2 must somehow participate in the overall viral infectivity. Literature is not clear on the subject as described in the result section, and in light of our present data, it seems important to re-address this question in order to better understand the HPV entry steps.

A constant targeting of positive and negative regulators of the same cellular processes emerges, highlighting a duality in E2 functions, which potentially drive opposing activities during different steps of the viral life cycle. We did not detect any specific functional targeting according to tropism or pathogenic power, suggesting that modulation of the identified cellular processes by E2 takes part in the general regulation of the viral life cycle of all HPVs. However, this common targeting proved to be achieved through diversified patterns of interactions, which probably drive the contribution of E2 into HPV pathogenesis.

The same approach was applied to study E6 and E7 of the same panel of HPV genotypes used here [87]. In contrast to the E2 interactome, numerous interactions were found specific to all the HR-HPV types or to all the LR-HPV types. This observation is not so surprising given that HPV oncogenic potential is known to mainly rely on the functions of these two proteins, and therefore E6 and E7 have evolved to target particular host factors according to the HR or LR trait of HPVs. Comparing the interactomes of E2, E6 and E7 led to the identification of some cellular targets shared by the different viral proteins. For example, Intergrin $\beta 4$ (ITGB4), appears as a strong interactor of both E2 and E7, or POMP, a proteasome subunit, is targeted by both E2 and E6. Given the strict evolutionary constrain of viruses, redundant targeting of the same proteins most probably reveals cellular mechanisms that are particularly important for HPV's infection. It should be worth looking in more details into these interactions, which could give insights into key processes for the viral life cycle, and could constitute particularly good targets for anti viral drugs. More generally, conducting systematic interactomics strategies on different proteins of a common set of multiple viral genotypes can point to cellular signaling pathways that are redundantly targeted by viral proteins and, as such, are likely to be central for the infection.

Such large-scale interactomic studies are also a prime choice to identify specific biomarkers of the most clinically relevant strains of HPVs. Indeed, we identified the human protein CCHCR1 as a specific interactor of the E2 protein from HPV16, the most prevalent HPV in cervical cancers. Given that CCHCR1 is not able to interact with the other closely related E2 proteins, its interaction with the HPV16 E2 likely contributes specifically to the HPV16 life cycle. Notably, it is tempting to speculate that this interaction could be involved in a pathogenic trait of this highly prevalent HPV and could be important to understand specific features of HPV16-induced carcinogenesis.

Actually, a link between CCHCR1 and cervical cancer was reported in 2005. Indeed,

in the microarray study of Santin and colleagues, CCHCR1 (herein named C6orf18) is identified among many other genes to be overexpressed in cervical cancer cultures when compared to normal cervical keratinocytes [340]. The authors stated that these genes could be triggered by HPV transformation and that understanding the molecular basis of HPV-mediated carcinogenesis could potentially help to identify more effective treatments. The second part of my thesis work, in lines with this hypothesis, shows that CCHCR1 might impact the pathogenesis of HPV16 by modifying the function of the regulatory E2 protein.

Current evidence indicates that CCHCR1 is associated with Psoriasis, a skin disease where overproliferative lesions of the skin can be observed. We were surprised that the E2 protein that specifically interacted with CCHCR1 was from a mucosal HPV. This probably means that CCHCR1 has a role in hyperproliferative lesions not only in skin keratinocytes. In addition, HPV16, besides its role in the genital mucosa, has been associated with the development of HPV-mediated oropharyngeal cancers and is thus also present in the oral mucosa. Therefore, the interaction between HPV16 E2 and CCHCR1 could have an outcome both in genital lesions and in oropharynx lesions. By contrast, E2 from the cutaneous HPV types only marginally interacted with CCHCR1, suggesting that CCHCR1 might not have a major role in HPV-associated skin warts or that it relies on additional factors not present in our assays. This study thus also enhances our knowledge about the involvement of CCHCR1 in pathologies of keratinized epithelia.

Mapping of the interaction domain shows that the binding interface of CCHCR1 on HPV16 E2 overlaps that of BRD4, the well-known interactor of E2. This leads to a competition between these two cellular proteins for the interaction with HPV16 E2. In the cell, there must be an equilibrium between E2 proteins bound to BRD4 and those bound to CCHCR1. As described earlier, in the interactomic study, we observed that the interaction of BRD4 to E2 was the weakest with HPV16, which potentially favors the competitive binding of CCHCR1 to HPV16 E2. The consequences of such a competitive binding can be multiple since BRD4 is associated with many functions of E2, and, in particular, we showed that it impacts on E2's transcriptional properties. The interaction with CCHCR1 also affects drastically the subcellular localization of HPV16 E2 by inducing its cytoplasmic retention. This could reinforce the deregulation of E2's nuclear functions and further participate in the alteration of E2's functional properties. It is therefore possible that in the case of HPV16, the transcriptional functions and the subcellular localization of E2 depend on the balance between its interaction with BRD4 and its interaction with CCHCR1, which can vary along the epithelium. This could correspond to a specific feature of HPV16.

Clues to the impact of the HPV16 E2-CCHCR1 interaction *in vivo* came from the study of keratinocyte differentiation. CCHCR1 has been associated with psoriatic lesions and was previously linked to proliferation of keratinocytes, in line with the results obtained here in HaCaT. Concerning E2, information about its role in keratinocyte differentiation is sparser but begins to emerge over the past few years (see introduction): induction of changes typical of differentiated cells, modification of expression profiles of genes involved in cell differentiation, interaction with proteins regulating differentiation. Notably, Burns *et al.* showed that the E2 protein from HPV16 was able to stimulate epithelial differentiation in HaCaT [295] in good agreement with the effect that we ob-

served here. In our conditions, HPV16 E2 drastically activates the expression of the early differentiation marker K10, which is not the case for HPV18 E2. The effect of E2 on K10 is severely affected by the binding of CCHCR1. This effect could be the result of both CCHCR1-dependent negative regulation on HPV16 E2's transcriptional properties and CCHCR1-mediated relocalization to the cytoplasm of this E2 protein.

In the context of the HPV life cycle, the virus requires that the host cell undertakes differentiation for genome amplification and viral egress. While the two HPV oncoproteins E6 and E7 promote cell proliferation, E2 is known to counter-balance their effects by direct repression of their expression. E2 also counteracts E6 and E7 by favoring cell differentiation. In the case of HPV16, the influence of CCHCR1 on E2 would interfere with the induction of the differentiation program, leaving a window of opportunity for an exaggerated stimulation of proliferation by E6 and E7. The fact that this interaction with CCHCR1 is specific to the E2 protein from HPV16, the most prevalent HPV in cancers, might account for the better capacities of HPV16 to generate uncontrolled hyperplasia with a high risk to undergo malignant conversion.

The functional repercussions of the interaction between CCHCR1 and E2 from HPV16 are multiple, spanning from interfering with essential interaction, defect in transcriptional activation to keratinocyte differentiation reprogramming. The interaction with CCHCR1 could somehow underlies the higher incidence of carcinogenic conversion of HPV16-associated lesions and could be envisioned as a unique biomarker.

In summary, this thesis establishes a comprehensive mapping of interactions between HPV E2 proteins and host cellular proteins using a robust comparative approach. We have provided a general overview of the landscape of human proteins interacting with E2 and the other HPV's early proteins, E6 and E7. Including in this study HPV strains that are often left aside yielded a number of new interactions that could have been missed by using only the most common HPV genotypes, and which turned out to be conserved for several E2 proteins. We therefore think that comparing such a relatively large spectra of HPV genotypes helped us to broaden the scope of interactions identified and led to the generation of reliable interaction datasets. This approach enabled us to further define the prominent mechanisms by which E2 hijacks the cellular machinery presumably for the benefit of the virus. Moreover, comparative studies can shed some light on how molecular differences among the virus types lead to varied pathological consequences. In particular, the identification of a unique HPV16 E2 interactor has directed us toward a better understanding of the specific pathogenic potential of HPV16, including determining the effects of this interaction on the differentiation process (**Fig III.1**).

The recent advent of large-scale proteomic techniques has boosted the progress in the study of papillomavirus-host cell interactions. The field has progressed from the initial discoveries of single protein-protein interactions, to later studies identifying several host proteins in complex with a single HPV strain, and now to studies examining multiple partners from many HPV types. HT-GPCA has allowed the detection of more and more interactors with higher confidence than ever before. Upcoming developments are expected to improve the visualization of complex-mediated luminescence in living cells. In that

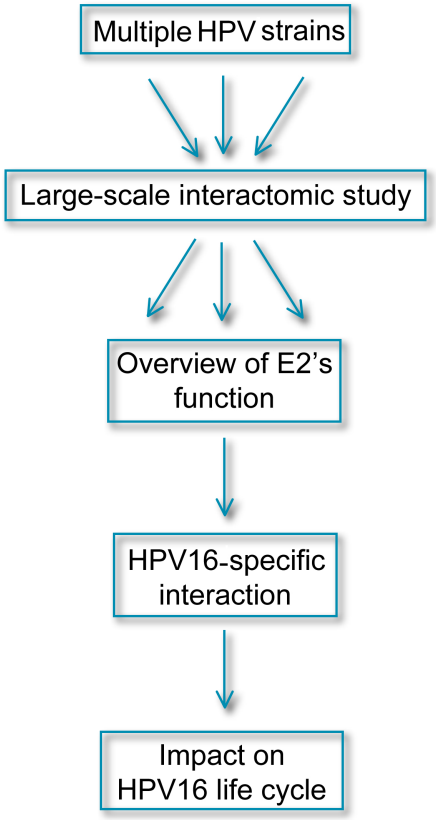

Figure III.1: **Schematic overview of this study**.

case, the HT-GPCA could be used to detect dynamic interactions in the context of the addition of drugs or complex inhibitors, thus allowing live monitoring of the disruption of an interaction. By using ordered arrays of ORFs taken from the increasing collection of the Human ORFeome, we foresee a point in the near future where HT-GPCA could be also used directly as a screening method. This should first drastically improve screening coverage since each protein pairs would be tested, and second, it should facilitate the automation of the technique. However, a more drastic high-throughput expansion of this assay should be brought up by the development of an *in vitro* HT-GPCA assay using *in vitro* expression systems based on human cell extracts. This should enable us to monitor protein-protein interactions in a controlled biochemical environment.

Studying virus-host interactions has become an important challenge since targeting protein-protein interactions is now considered as an attractive approach for the development of new therapeutic strategies. The use of small peptide-based strategies to counteract specific interactions has been studied for long but the ultimate goal to develop an efficient treatment would be to identify a pan-HPV inhibitor. To that end, efforts should be put into the expansion of systematic and comparative approaches to identify protein-protein interactions.

IV . Materials & Methods

1 . Cell culture and transfections

HEK-293T, HeLa and HaCaT cells were routinely maintained in DMEM supplemented with 10% fetal bovine serum at 37 °C in a 5% CO_2 incubator. Cells were transfected 24h after plating by linear PEI (polyethylenimine, Polysciences Inc). For siRNA studies, cells were reverse transfected by INTERFERin (Polyplus-Transfection). Scrambled siRNA were used as negative controls. 48 to 72h later, cells are collected and subjected to further analyses. The sequences of the siCCHCR1 are as follows: 5'-GAACUUGGAAGAGGGGAGGCA dTdT-3' and 5'-UGCCUCCCCUCUUCCAAGUUC dTdT-3'.

2 . Plasmids

The ORFs encoding for the E2 proteins were amplified from viral genomic DNA corresponding to the different HPV genotypes, cloned by the gateway recombinational cloning system (Invitrogen) into the entry vector pDON207 (Invitrogen), and were listed in the ViralORFeome database [341]. The E2 ORFs were then transferred into gateway-compatible destination vectors: pGBKT7-gw to generate E2-GAL4 DNA-binding domain fusion proteins for the yeast two-hybrid screen; pSPICA-N2-gw to generate proteins with amino acids 110 to 185 of the humanized *Gaussia princeps* luciferase in fusion with the N-terminus of E2 (GL2-E2 fusion proteins) for the High-Throughput Gaussia princeps Luciferase-based Complementation Assay (HT-GPCA); pCI-Neo-FLuc-gw to generate Firefly luciferase- E2 fusions proteins for steady state levels measurement; pCiNeo-3XFlag-gw for binding competition assay; peGFP-gw for fluorescence studies. The HPV16 E2 proteins with point mutations (N181T, R27D, R37A, I73A and E39A) or deleted of the N-terminal helices were obtained by PCR-directed mutagenesis.

Entry gateway plasmids for cellular partners were obtained either by PCR amplification from clones recovered by yeast two-hybrid originating from a HaCaT cDNA library (Clontech), or from the human ORFeome resource (hORFeome v3.1). The cellular ORF were transferred into gateway-compatible destination vectors pSPICA-N1-gw to generate proteins fused at the N-terminus with the amino acids 18 to 109 of humanized Gaussia luciferase (GL1-fusion proteins). A schematic representation of both E2 and cellular protein expression constructs can be found in **Figure IV.1**.

The CCHCR1 ORF was obtained by PCR amplification of CCHCR1 cDNA clones extracted from the yeast two-hybrid screen, originating from a HaCaT cDNA library (Clontech). Plasmids encoding BRD4 were kindly provided by Cheng-Ming Chiang and BRD4-CTD-NLS by J. Archambault. All ORFs were transferred into various gateway compatible destination vectors: pCherry-gw for fluorescence assay, pCiNeo-3XFlag-gw to generate Flag-tagged fusion proteins, pSPICA-N1-gw for interaction assay, or pCI-Neo-gw to express untagged proteins.

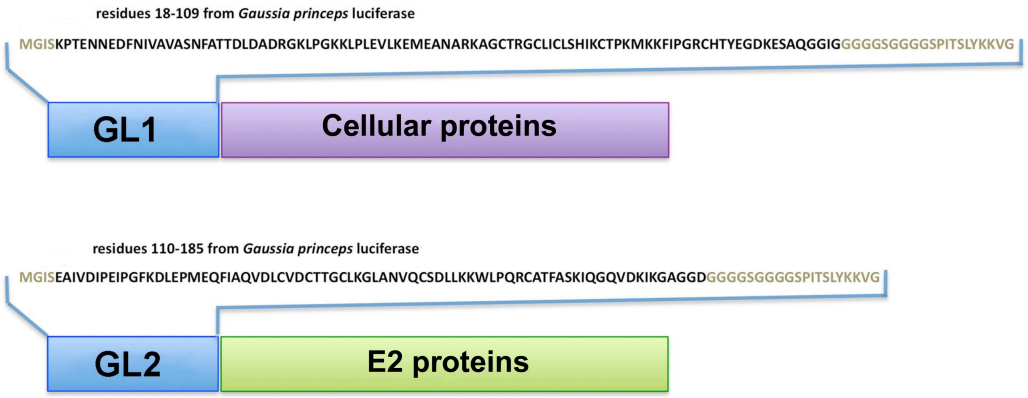

Figure IV.1: **GL1 and GL2 fusion proteins**. Schematic representation of the constructs used to express E2 and the cellular proteins in HT-GPCA. The amino acid sequences of the tags (black), as well as the flexible linkers (grey) are represented. Adapted from *Cassonnet et al.* [313].

The luciferase reporter (pTK6E2BS) driven by a E2-responsive promoter contained six E2-binding sites upstream the minimal TK promoter. E2BS sequences are as follows: (aACCGTTTTCGGTtaaACCGTTTTCGGTt)X3, designed after the study of Sanchez *et al.* [204] to be optimal for the binding of a large panel of E2 proteins. The RNA polymerase III-directed Renilla Luciferase plasmid (polIII-Ren) used as an internal control of transfection contained a 100-mer nucleotide encompassing the human Histone H1 promoter upstream of the Renilla ORF (hRluc).

Ad-GFPE2 and Ad-GFP constructs were obtained by bacterial recombination prior to this thesis and are described elsewhere [286].

3 .HT-GPCA

HEK-293T cells were seeded at 35,000 cells per well in 96-well plates. After 24h, cells were transfected by linear PEI (polyethylenimine) with pSPICA-N2-E2 and pSPICA-N1-cellular protein constructs (100 ng each), for expression of the GL2-E2 and GL1-fusion proteins, where GL1 and GL2 are two inactive fragments of the Gaussia princeps luciferase. 10 ng of a CMV-firefly luciferase reporter plasmid was added to normalize for transfection efficiency. Cells were lysed 24h post-transfection in 40 μL of Renilla luciferase lysis buffer (Promega) for 30 minutes. The *Gaussia princeps* luciferase activity was measured on 30 μL of total cell lysate by a luminometer Berthold Centro XS LB960 after injection of 100 μL of the Renilla luciferase substrate (Promega). Firefly luciferase was measured on the remaining 10 μL lysate with Firefly luciferase substrate. Gaussia Luciferase activity was divided by Firefly luciferase activity for each sample, giving a normalized Gaussia luminescence. Each normalized Gaussia luciferase activity was calculated from the mean of triplicate samples. For a given pair of proteins (A and B), the normalized Gaussia luminescence of cells coexpressing GL1-A+GL2-B proteins was divided by the sum of normalized Gaussia luminescence of each partner coexpressed with the cor-

responding empty plasmid: GL1-A+GL2-B/((GL1-A +GL2) + (GL1 + GL2-B)). This gave a Normalized Luminescence Ratio (NLR) corresponding to the reconstituted Gaussia luciferase activity (see **Figure II.1**), thus reflecting the level of interaction between protein pairs.

4 .Hierarchical clustering and topology

Interaction data analyses were performed using the R statistics package. Raw NLR interaction data were separated into categories in order to minimize the dispersion of NLR values. Cut-off thresholds of each category were determined with the goal of maintaining the same frequency distribution across all categories. An Euclidian distance matrix was calculated from the data categories using R. The interaction dendrogram was calculated using the complete (UPGMA) linkage method. E2 protein sequences were clustered using the phylip package. Protein distances were calculated with the prodist program, using default parameters. The phylogenetic dendrogram was generated with the neighbor program using the UPGMA method and default parameters. Both interaction and phylogenetic dendrograms were generated using JavaTreeView. A Pearson correlation coefficient was calculated with R using the cophenetic distances between both interaction and phylogenetic dendrogram to determine the closeness of the two dendrograms. The label order for the intensity data was then randomly changed to generate 100,000 random dendrograms. The cophenetic distance matrix for these randomized dendrograms was compared to the cophenetic distance matrix from the phylogenetic dendrogram with a Pearson correlation (cor) function from R. The number of standard deviations between this correlation and the mean of the distribution of the correlation between the random and the phylogenetic dendrogram was used to calculate the p-value. A Cumulative Density Function of the randomized dataset was compared to a normal distribution generated by the R function "rnorm"using the same mean and standard deviation from the randomized dataset to check the normality of the data.

The E2 interaction networks were generated with the cytoscape software with interactions scoring positive in HT-GPCA (NLR above 3.5). The degree of each cellular protein in both E2 and HPRD-based human interactomes were extracted from cytoscape and were collected to calculate the degree distribution. To determine the overrepresented GO (Gene Ontology) terms in the interaction dataset and to evaluate the gathering of E2 targets by functional categories, we used the DAVID bioinformatic database [342].

5 .Fluorescence

HaCaT cells grown in coverslip in 6-well plates and were co-transfected by linear PEI with expression plasmids for GFP-fused E2 proteins (3 μg) and Cherry-fused CCHCR1 (1 μg). 24h post transfection, cells were fixed with 4% paraformaldehyde for 30 min at 4 $^\circ$ C, washed in PBS, permeablized with 0.1%Triton 100X-containing PBS and stained for 30 min with DAPI. Cells were mounted with CitiFluor. Fluorescent Images were acquired using a ZEISS Apotome microscope.

6 . Transactivation

5x10^5 HaCaT plated cells were transfected 24h later with 100ng of the E2-dependent reporter plasmid pTK6E2BS. Other expression vectors included 25 ng of a Renilla luciferase-expressing plasmid (polIII-Ren)) as an internal control for normalization purpose, 100ng of pCINeo-driven HPV16 E2, and 0.8 μg of 3XFLAG-tagged challenging proteins BRD4 and CCHCR1. Cells were harvested 30h post-transfection, lysed in Passive lysis buffer according to manufacturer's instructions and luciferase activities were measured with Dual Glo substrates (Promega).

7 . Differentiation

For the differentiation assay, HaCaT cells were maintained in calcium-free DMEM (Invitrogen) in order to keep HaCaT cells in an undifferentiated state. Infections with recombinant adenoviruses expressing the GFP-E2 fusion proteins or GFP only were done at a multiplicity of infection of 250, in 1mL of DMEM complemented with 4μM polybrene for 1h at 37 $^\circ$C. The medium was then replaced by fresh medium with 10% fetal bovine serum. Cells were collected 24h or 48h later and subjected to RNA extraction.

8 . RNA extraction and RT-qPCR

Total RNA was isolated by TRIzol Reagent (Invitrogen) according to the manufacturer's instructions, and used for cDNA synthesis by Superscript II (Invitrogen). The cDNAs were used as templates for quantitative PCR using SYBR Green PCR master mix (Applied Biosystems). Primers used for RT-PCR are listed below. None of the primer sequences showed genomic cross-reactivity with other genes.

Table 1: List of primers used for qPCR analyses

CCHCR1(Forward)	5'-GCCAAGGCCGTGGTCTCCTTG-3'
CCHCR1(Reverse)	5'-CCCTCTCTAGCTCCTGCAAGCG-3'
K14 (Forward)	5'-GCGGATGACTTCCGCACCAAGTATGAG-3'
K14 (Reverse)	5'-CCTTCAGGCTCTCAATCTGCATCTCC-3'
K10 (Forward)	5'-GATGTGAATGTGGAAATGAATGCTGCCC-3
K10 (Reverse)	5'-GTTCCTTGCTCTTTTCATTGAACCAGGC-3'
TGM1 (Forward)	5'-CAGTGCTGCGCTGCCTGGGTC-3'
TGM1 (Reverse)	5'-CCGGCCTCTTCATCCAGCAGTC -3'

The $\Delta\Delta$Ct method was used to calculate fold changes. The cycle threshold (Ct) values of the gene of interest are correlated to the mRNA amount. For normalization, the GAPDH housekeeping gene was used as control and amplified in the same assay. The Ct value of the housekeeping gene was subtracted from the Ct value of the gene of interest, which gives a ΔCt. This represents the relative amount of the gene of interest transcripts. The fold increase induced by the E2 proteins or CCHCR1 is calculated by comparing these ΔCt values with the ΔCt values extracted from mock-transfected cells ($\Delta\Delta$Ct). The data are presented as the fold change in gene expression normalized to

an endogenous reference gene and relative to the untreated control, in our case, addition of E2 or CCHCR1 using the $2^{\Delta\Delta ct}$ method as follows: for the untreated control sample, $\Delta\Delta$Ct equals zero and 2^0 equals one, so that the fold change in gene expression relative to the untreated control equals one, by definition. Therefore, if the results gives a $2^{\Delta\Delta ct}$ of 0.5, it means that there is a 2-fold decrease of the gene of interest transcripts in the treated conditions, *i.e.* in the presence of E2 or CCHCR1.

Bibliography

[1] **Ciuffo, G.** (1907). *Innesto positivo con filtrato di verruca volgare.* Giorn Ital Mal Venereol **48(1)**, 12–17.

[2] **Shope, R., Hurst, E.W.** (1933). *Infectious papillomatosis of rabbits: With a note on the histopathology.* The Journal of Experimental Medicine **58(5)**, 607–624.

[3] **Rous, P., Beard, J.W.** (1935). *The progression to carcinoma of virus-induced rabbit papillomas (shope).* The Journal of Experimental Medicine **62(4)**, 523–548.

[4] **Zur Hausen, H.** (1976). *Condylomata acuminata and human genital cancer.* Cancer Research **36(1)**, 794.

[5] **Boshart, M., Gissmann, L., Ikenberg, F., Kleinheinz, A., Scheurlen, F., zur Hausen, H.** (1984). *A new type of papillomavirus DNA, its presence in genital cancer biopsies and in cell lines derived from cervical cancer.* EMBO Journal **3(5)**, 1151–1157.

[6] **Bravo, I., Alonso, Á.** (2007). *Phylogeny and evolution of papillomaviruses based on the E1 and E2 proteins.* Virus Genes **34(1)**, 249–262.

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

[8] **Schiffman, M., Rodriguez, A., Chen, Z., Wacholder, S., Herrero, R., Hildesheim, A., Desalle, R., Befano, B., Yu, K., Safaeian, M., Sherman, M., Morales, J., Guillen, D., Alfaro, M., Hutchinson, M., Solomon, D., Castle, P., Burk, R.** (2010). *A population-based prospective study of carcinogenic human papillomavirus variant lineages, viral persistence, and cervical neoplasia.* Cancer Research **70(8)**, 3159–3169.

[9] **Baseman, J., Koutsky, L.** (2005). *The epidemiology of human papillomavirus infections.* J Clin Virol **32(1)**, 16–24.

[10] **Schiffman, M., Kjaer, S.** (2003). *Chapter 2: Natural history of anogenital human papillomavirus infection and neoplasia.* J Natl Cancer Inst Monogr **31(1)**, 14–19.

[11] **Giuliano, A., Harris, R., Sedjo, R., Baldwin, S., Roe, D., Papenfuss, M., Abrahamsen, M., Inserra, P., Olvera, S., Hatch, K.** (2002). *Incidence, prevalence, and clearance of type-specific human papillomavirus infections: The young women's health study.* J Infect Dis **186(4)**, 462–9.

[12] **Franco, E., Villa, L., Sobrinho, J., Prado, J., Rousseau, M., Desy, M., Rohan, T.** (1999). *Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer.* J Infect Dis **180(5)**, 1415–23.

[13] **Liaw, K., Hildesheim, A., Burk, R., Gravitt, P., Wacholder, S., Manos, M.M., Scott, D., Sherman, M., Kurman, R., Glass, A., Anderson, S., Schiffman, M.** (2001). *A prospective study of human papillomavirus (HPV) type 16 DNA detection by polymerase chain reaction and its association with acquisition and persistence of other HPV types.* Journal of Infectious Diseases **183(1)**, 8–15.

[14] **Jemal, A., Bray, F., Center, M., Ferlay, J., Ward, E., Forman, D.** (2011). *Global cancer statistics.* CA: A Cancer Journal for Clinicians **61(2)**, 69–90.

[15] **Mathew, A., George, P.** (2009). *Trends in incidence and mortality rates of squamous cell carcinoma and adenocarcinoma of cervix–worldwide.* Asian Pacific Journal of Cancer Prevention **10(4)**, 645–50.

[16] **Praetorius, F.** (1997). *HPV-associated diseases of oral mucosa.* Clinics in Dermatology **15(3)**, 399–413.

[17] **Gillison, M.** (2004). *Human papillomavirus-associated head and neck cancer is a distinct epidemiologic, clinical, and molecular entity.* Semin Oncol **31(6)**, 744–754.

[18] **Smith, E., Ritchie, J., Summersgill, K., Klussmann, J., Lee, J., Wang, D., Haugen, T., Turek, L.** (2004). *Age, sexual behavior and human papillomavirus infection in oral cavity and oropharyngeal cancers.* International Journal of Cancer **108(5)**, 766–772.

[19] **Begum, S., Cao, D., Gillison, M., Zahurak, M., Westra, W.** (2005). *Tissue distribution of human papillomavirus 16 DNA integration in patients with tonsillar carcinoma.* Clinical Cancer Research **11(16)**, 5694–5699.

[20] **Cardoso, J., Calonje, E.** (2011). *Cutaneous manifestations of human papillomaviruses: a review.* Acta Dermatovenerol Alp Panonica Adriat. **20(3)**, 145–54.

[21] **Rübben, A., Kalka, K., Spelten, B., Großendorf-Conen, E.** (1997). *Clinical features and age distribution of patients with HPV 2/27/57-induced common warts.* Archives of Dermatological Research **289(1)**, 337–340.

[22] **Gewirtzman, A.and Bartlett, B., Tyring, S.** (2008). *Epidermodysplasia verruciformis and human papilloma virus.* Current Opinion in Infectious Disease **21(2)**, 141–6.

[23] **Patel, T., Morrison, L.K., Rady, P., Tyring, S.** (2010). *Epidermodysplasia verruciformis and susceptibility to HPV.* Disease Markers **29(3)**, 199–206.

[24] **Majewski, S., Jabłońska, S., Orth, G.** (1997). *Epidermodysplasia verruciformis. immunological and nonimmunological surveillance mechanisms: Role in tumor progression.* Clinics in Dermatology **15(3)**, 321–334.

[25] **Boxman, I.L.A., Berkhout, R.J.M., Mulder, L.H.C., Wolkers, M.C., Bavinck, J.N.B., Vermeer, B.J., ter Schegget, J.** (1997). *Detection of human papillomavirus DNA in plucked hairs from renal transplant recipients and healthy volunteers.* Journal of Investigative Dermatology **108(1)**, 712–715.

[26] **Ramoz, N., Rueda, L., Bouadjar, B., Montoya, L., Orth, G., Favre, M.** (2002). *Mutations in two adjacent novel genes are associated with epidermodysplasia verruciformis.* Nat Genet **32(4)**, 579–581.

[27] **Ramoz, N., Taieb, A., Rueda, L., Montoya, L., Bouadjar, B., Favre, M., Orth, G.** (2000). *Evidence for a nonallelic heterogeneity of epidermodysplasia verruciformis with two susceptibility loci mapped to chromosome regions 2p21-p24 and 17q25.* J Invest Dermatol **114(6)**, 1148–53.

[28] **Biliris, K.A., Koumantakis, E., Dokianakis, D.N., Sourvinos, G., Spandidos, D.A.** (2000). *Human papillomavirus infection of non-melanoma skin cancers in immunocompetent hosts.* Cancer Lett **161(1)**, 83–8.

[29] **Pfister, H.** (2003). *Chapter 8: Human papillomavirus and skin cancer.* JNCI Monographs **2003(31)**, 52–56.

[30] **Baker, T., Newcomb, W., Olson, N., Cowsert, L., Olson, C., Brown, J.** (1991). *Structures of bovine and human papillomaviruses. Analysis by cryoelectron microscopy and three-dimensional image reconstruction.* Biophys J **60(6)**, 1445–1456.

[31] **Bednarek, P., Lee, B., Gandhi, S., Lee, E., Phillips, B.** (1998). *Novel binding sites for regulatory factors in the human papillomavirus type 18 enhancer and promoter identified by in vivo footprinting.* Journal of Virology **72(1)**, 708–716.

[32] **Tan, S., Baker, C., Stünkel, W., Bernard, H.** (2003). *A transcriptional initiator overlaps with a conserved YY1 binding site in the long control region of human papillomavirus type 16.* Virology **305(2)**, 486–501.

[33] **Gloss, B., Yeo-Gloss, M., Meisterernst, M., Rogge, L., Winnacker, E., Bernard, H.** (1989). *Clusters of nuclear factor I binding sites identify enhancers of several papillomaviruses but alone are not sufficient for enhancer function.* Nucleic Acids Research **17(9)**, 3519–3533.

[34] **Bernard, H., Apt, D.** (1994). *Transcriptional control and cell type specificity of HPV gene expression.* Archives of Dermatology **130(2)**, 210–215.

[35] **Desaintes, C., Demeret, C.** (1996). *Control of papillomavirus DNA replication and transcription.* Semin Cancer Biol **7(6)**, 339–47.

[36] **Gloss, B., Bernard, H., Seedorf, K., Klock, G.** (1987). *The upstream regulatory region of the human papilloma virus-16 contains an E2 protein-independent enhancer which is specific for cervical carcinoma cells and regulated by glucocorticoid hormones.* EMBO Journal **6(12)**, 3735–43.

[37] **Bromberg-White, J., Meyers, C.** (2002). *The upstream regulatory region of human papillomavirus type 31 is insensitive to glucocorticoid induction.* Journal of Virology **76(19)**, 9702–9715.

[38] **Chen, Y., Huang, L., Chen, T.** (1996). *Differential effects of progestins and estrogens on long control regions of human papillomavirus types 16 and 18.* Biochemical and Biophysical Research Communications **224(3)**, 651–659.

[39] **Reh, H., Pfister, H.** (1990). *Human papillomavirus type 8 contains cis-active positive and negative transcriptional control sequences.* Journal of General Virology **71(10)**, 2457–2462.

[40] **Ensser, A., Pfister, H.** (1990). *Epidermodysplasia verruciformis associated human papilomaviruses present a subgenus-specific organization of the regulatory genome region.* Nucleic Acids Research **18(13)**, 3919–3922.

[41] **Guido, M.C., Zamorano, R., Garrido-Guerrero, E., Gariglio, P., Garcia-Carranca, A.** (1992). *Early promoters of genital and cutaneous human papillomaviruses are differentially regulated by the bovine papillomavirus type 1 E2 gene product.* J Gen Virol **73(1)**, 1395–400.

[42] **Thorner, L., Lim, D., Botchan, M.** (1993). *DNA-binding domain of bovine papillomavirus type 1 E1 helicase: structural and functional aspects.* Journal of Virology **67(10)**, 6000–6014.

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

[44] **White, P., Pelletier, A., Brault, K., Titolo, S., Welchner, E., Thauvette, L., Fazekas, M., Cordingley, M., Archambault, J.** (2001). *Characterization of recombinant HPV6 and 11 E1 helicases: Effect of ATP on the interaction of E1 with E2 and mapping of a minimal helicase domain.* Journal of Biological Chemistry **276(25)**, 22426–22438.

[45] **Sarafi, T., McBride, A.** (1995). *Domains of the BPV-1 E1 replication protein required for origin-specific DNA binding and interaction with the E2 transactivator.* Virology **211(2)**, 385 – 396.

[46] **Bream, G., Ohmstede, C., Phelps, W.** (1993). *Characterization of human papillomavirus type 11 E1 and E2 proteins expressed in insect cells.* Journal of Virology **67(5)**, 2655–2663.

[47] **Seo, Y., Müller, F., Lusky, M., Hurwitz, J.** (1993). *Bovine papilloma virus (BPV)-encoded E1 protein contains multiple activities required for BPV DNA replication.* Proceedings of the National Academy of Sciences **90(2)**, 702–706.

[48] **Titolo, S., Pelletier, A., Sauvé, F., Brault, K., Wardrop, E., White, P., Amin, A., Cordingley, M., Archambault, J.** (1999). *Role of the atp-binding domain of the human papillomavirus type 11 e1 helicase in E2-dependent binding to the origin.* Journal of Virology **73(7)**, 5282–5293.

[49] **Seo, Y., Müller, F., Lusky, M., Gibbs, E., Kim, H., Phillips, B., Hurwitz, J.** (1993). *Bovine papilloma virus (BPV)-encoded E2 protein enhances binding of E1 protein to the BPV replication origin.* Proceedings of the National Academy of Sciences **90(7)**, 2865–2869.

[50] **Conger, K.L., Liu, J.S., Kuo, S.R., Chow, L.T., Wang, T.S.F.** (1999). *Human papillomavirus DNA replication: Interactions between the viral E1 protein and two subunits of human DNA polymerase α-primase.* Journal of Biological Chemistry **274(5)**, 2696–2705.

[51] **Masterson, P.J., Stanley, M.A., Lewis, A.P., Romanos, M.A.** (1998). *A c-terminal helicase domain of the human papillomavirus E1 protein binds E2 and the DNA polymerase α-primase p68 subunit.* Journal of Virology **72(9)**, 7407–7419.

[52] **Park, P., Copeland, W., Yang, L., Wang, T., Botchan, M.R., Mohr, I.J.** (1994). *The cellular DNA polymerase alpha-primase is required for papillomavirus DNA replication and associates with the viral E1 helicase.* Proceedings of the National Academy of Sciences **91(18)**, 8700–8704.

[53] **Clower, R.V., Hu, Y., Melendy, T.** (2006). *Papillomavirus E2 protein interacts with and stimulates human topoisomerase I.* Virology **348(1)**, 13–8.

[54] **Loo, Y.M., Melendy, T.** (2004). *Recruitment of replication protein a by the papillomavirus E1 protein and modulation by single-stranded DNA.* Journal of Virology **78(4)**, 1605–1615.

[55] **Sedman, J., Stenlund, A.** (1998). *The papillomavirus E1 protein forms a DNA-Dependent hexameric complex with ATPase and DNA helicase activities.* Journal of Virology **72(8)**, 6893–6897.

[56] **Brown, D.R., Bryan, J.T., Rodriguez, M., Katz, B.P.** (1992). *Factors associated with detection of human papillomavirus E4 and L1 proteins in condylomata acuminata.* Journal of Infectious Diseases **166(3)**, 512–517.

[57] **Brown, D.R., Fan, L., Jones, J., Bryan, J.** (1994). *Colocalization of human papillomavirus type 11 E1ˆE4 and L1 proteins in human foreskin implants grown in athymic mice.* Virology **201(1)**, 46 – 54.

[58] **Doorbar, J., Parton, A., Hartley, K., Banks, L., Crook, T., Stanley, M., Crawford, L.** (1990). *Detection of novel splicing patterns in a HPV16-containing keratinocyte cell line.* Virology **178(1)**, 254–262.

[59] **Doorbar, J., Campbell, D., Grand, R.J., Gallimore, P.H.** (1986). *Identification of the human papilloma virus-1a E4 gene products.* EMBO J **5(2)**, 355–362.

[60] **Tan, C.L., Gunaratne, J., Lai, D., Carthagena, L., Wang, Q., Xue, Y.Z., Quek, L.S., Doorbar, J., Bachelerie, F., Thierry, F., Bellanger, S.** (2012). *HPV-18 E2 ˆE4 chimera: 2 new spliced transcripts and proteins induced by keratinocyte differentiation.* Virology **429(1)**, 47– 56.

[61] **Doorbar, J., Foo, C., Coleman, N., Medcalf, L., Hartley, O., Prospero, T., Napthine, S., Sterling, J., Winter, G., Griffin, H.** (1997). *Characterization of events during the late stages of HPV16 infection in vivo using high-affinity synthetic fabs to E4.* Virology **238(1)**, 40 – 52.

[62] **McIntosh, P.B., Laskey, P., Sullivan, K., Davy, C., Wang, Q., Jackson, D.J., Griffin, H.M., Doorbar, J.** (2010). *E1ˆE4-mediated keratin phosphorylation and ubiquitylation: a mechanism for keratin depletion in HPV16-infected epithelium.* Journal of Cell Science **123(16)**, 2810–2822.

[63] **Doorbar, J., Ely, S., Sterling, J., McLean, C., Crawford, L.** (1991). *Specific interaction between HPV-16 E1ˆE4 and cytokeratins results in collapse of the epithelial cell intermediate filament network.* Nature **352(6338)**, 824–827.

[64] **Raj, K., Berguerand, S., Southern, S., Doorbar, J., Beard, P.** (2004). *E1ˆE4 protein of human papillomavirus type 16 associates with mitochondria.* Journal of Virology **78(13)**, 7199–7207.

[65] **Davy, C.E., Jackson, D.J., Wang, Q., Raj, K., Masterson, P.J., Fenner, N.F., Southern, S., Cuthill, S., Millar, J.B.A., Doorbar, J.** (2002). *Identification of a G2 arrest domain in the E1ˆE4 protein of human papillomavirus type 16.* Journal of Virology **76(19)**, 9806–9818.

[66] **Conrad, M., Bubb, V.J., Schlegel, R.** (1993). *The human papillomavirus type 6 and 16 E5 proteins are membrane-associated proteins which associate with the 16-kilodalton pore-forming protein.* Journal of Virology **67(10)**, 6170–6178.

[67] **Hwang, E.S., Nottoli, T., DiMaio, D.** (1995). *The HPV16 E5 protein: Expression, detection, and stable complex formation with transmembrane proteins in COS cells.* Virology **211(1)**, 227–233.

[68] **Venuti, A., Paolini, F., Nasir, L., Corteggio, A., Roperto, S., Campo, M., Borzacchiello, G.** (2011). *Papillomavirus E5: the smallest oncoprotein with many functions.* Molecular Cancer **10(1)**, 140–158.

[69] **Yang, D.H., Wildeman, A.G., Sharom, F.J.** (2003). *Overexpression, purification, and structural analysis of the hydrophobic E5 protein from human papillomavirus type 16.* Protein Expression and Purification **30(1)**, 1–10.

[70] **Leptak, C., Ramon y Cajal, S., Kulke, R., Horwitz, B.H., Riese, D.J., Dotto, G.P., DiMaio, D.** (1991). *Tumorigenic transformation of murine keratinocytes by the E5 genes of bovine papillomavirus type 1 and human papillomavirus type 16.* Journal of Virology **65(12)**, 7078–7083.

[71] **Chen, S.L., Mounts, P.** (1990). *Transforming activity of E5a protein of human papillomavirus type 6 in NIH 3T3 and C127 cells.* Journal of Virology **64(7)**, 3226–3233.

[72] **Genther Williams, S.M., Disbrow, G.L., Schlegel, R., Lee, D., Threadgill, D.W., Lambert, P.F.** (2005). *Requirement of epidermal growth factor receptor for hyperplasia induced by E5, a high-risk human papillomavirus oncogene.* Cancer Research **65(15)**, 6534–6542.

[73] **Ashrafi, G.H., Haghshenas, M.R., Marchetti, B., O'Brien, P.M., Campo, M.S.** (2005). *E5 protein of human papillomavirus type 16 selectively downregulates surface HLA class I.* International Journal of Cancer **113(2)**, 276–283.

[74] **Zhang, B., Li, P., Wang, E., Brahmi, Z., Dunn, K.W., Blum, J.S., Roman, A.** (2003). *The E5 protein of human papillomavirus type 16 perturbs MHC class II antigen maturation in human foreskin keratinocytes treated with interferon.* Virology **310(1)**, 100–108.

[75] **Dyson, N., Howley, P.M., Munger, K., Harlow, E.** (1989). *The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product.* Science **243(4893)**, 934–7.

[76] **Gage, J.R., Meyers, C., Wettstein, F.O.** (1990). *The E7 proteins of the nononcogenic human papillomavirus type 6b (HPV-6b) and of the oncogenic HPV-16 differ in retinoblastoma protein binding and other properties.* Journal of Virology **64(2)**, 723–730.

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

[78] **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.

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

[80] **Scheffner, M., Huibregtse, J.M., Vierstra, R.D., Howley, P.M.** (1993). *The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53.* Cell **75(3)**, 495–505.

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

[82] **Li, X., Coffino, P.** (1996). *High-risk human papillomavirus E6 protein has two distinct binding sites within p53, of which only one determines degradation.* Journal of Virology **70(7)**, 4509–16.

[83] **Lee, S.S., Weiss, R.S., Javier, R.T.** (1997). *Binding of human virus oncoproteins to hDlg/SAP97, a mammalian homolog of the drosophila discs large tumor suppressor protein.* Proceedings of the National Academy of Sciences **94(13)**, 6670–6675.

[84] **Nguyen, M.L., Nguyen, M.M., Lee, D., Griep, 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.

[85] **Duensing, S., Munger, K.** (2004). *Mechanisms of genomic instability in human cancer: Insights from studies with human papillomavirus oncoproteins.* International Journal of Cancer **109(2)**, 157–162.

[86] **Tungteakkhun, S., Duerksen-Hughes, P.** (2008). *Cellular binding partners of the human papillomavirus E6 protein.* Archives of Virology **153(1)**, 397–408.

[87] **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 gaussia princeps luciferase.* Methods **58(4)**, 349–359.

[88] **Chen, J.J., Hong, Y., Rustamzadeh, E., Baleja, J.D., Androphy, E.J.** (1998). *Identification of an alpha helical motif sufficient for association with papillomavirus E6.* Journal of Biological Chemistry **273(22)**, 13537–13544.

[89] **Huibregtse, J.M., Scheffner, M., Howley, P.M.** (1993). *Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins.* Molecular and Cellular Biology **13(8)**, 4918–4927.

[90] **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., Muller, I., Stote, R.H., Cavarelli, J., Vande Pol, S., Travé, G.** (2013). *Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins.* Science **339(6120)**, 694–698.

[91] **Masson, M., Hindelang, C., Sibler, A.P., Schwalbach, G., Travé, G., Weiss, E.** (2003). *Preferential nuclear localization of the human papillomavirus type 16 E6 oncoprotein in cervical carcinoma cells.* Journal of General Virology **84(8)**, 2099–2104.

[92] **Knapp, A.A., McManus, P.M., Bockstall, K., Moroianu, J.** (2009). *Identification of the nuclear localization and export signals of high risk HPV16 E7 oncoprotein.* Virology **383(1)**, 60–68.

[93] **Modis, Y., Trus, B.L., Harrison, S.C.** (2002). *Atomic model of the papillomavirus capsid.* EMBO Journal **2(18)**, 4754–4762.

[94] **Pereira, R., Hitzeroth, I., Rybicki, E.** (2009). *Insights into the role and function of L2, the minor capsid protein of papillomaviruses.* Archives of Virology **154(2)**, 187–197.

[95] **Giroglou, T., Florin, L., Schafer, F., Streeck, R.E., Sapp, M.** (2001). *Human papillomavirus infection requires cell surface heparan sulfate.* Journal of Virology **75(3)**, 1565–1570.

[96] **Culp, T.D., Budgeon, L.R., Christensen, N.D.** (2006). *Human papillomaviruses bind a basal extracellular matrix component secreted by keratinocytes which is distinct from a membrane-associated receptor.* Virology **347(1)**, 147–159.

[97] **Bienkowska-Haba, M., Williams, C., Kim, S.M., Garcea, R.L., Sapp, M.** (2012). *Cyclophilins facilitate dissociation of the human papillomavirus type 16 capsid protein L1 from the L2/DNA complex following virus entry.* Journal of Virology **86(18)**, 9875–9887.

[98] **Casini, G.L., Graham, D., Heine, D., Garcea, R.L., Wu, D.T.** (2004). *In vitro papillomavirus capsid assembly analyzed by light scattering.* Virology **325(2)**, 320–327.

[99] **Villa, L.L., Ault, K.A., Giuliano, A.R., Costa, R.L., Petta, C.A., Andrade, R.P., Brown, D.R., Ferenczy, A., Harper, D.M., Koutsky, L.A., Kurman, R.J., Lehtinen, M., Malm, C., Olsson, S.E., Ronnett, B.M., Skjeldestad, F.E., Steinwall, M., Stoler, M.H., Wheeler, C.M., Taddeo, F.J., Yu, J., Lupinacci, L., Railkar, R., Marchese, R., Esser, M.T., Bryan, J., Jansen, K.U., Sings, H.L., Tamms, G.M., Saah, A.J., Barr, E.** (2006). *Immunologic responses following administration of a vaccine targeting human papillomavirus types 6, 11, 16, and 18.* Vaccine **24(27–28)**, 557–5583.

[100] **Brown, D.R., Fife, K.H., Wheeler, C.M., Koutsky, L.A., Lupinacci, L.M., Railkar, R., Suhr, G., Barr, E., Dicello, A., Li, W., Smith, J.F., Tadesse, A., Jansen, K.U.** (2004). *Early assessment of the efficacy of a human papillomavirus type 16 L1 virus-like particle vaccine.* Vaccine **22(21–22)**, 2936–2942.

[101] **Buck, C.B., Cheng, N., Thompson, C.D., Lowy, D.R., Steven, A.C., Schiller, J.T., Trus, B.L.** (2008). *Arrangement of L2 within the papillomavirus capsid.* Journal of Virology **82(11)**, 5190–5197.

[102] **Finnen, R.L., Erickson, K.D., Chen, X.S., Garcea, R.L.** (2003). *Interactions between papillomavirus L1 and L2 capsid proteins.* Journal of Virology **77(8)**, 4818–4826.

[103] **Bergant Marsic, M., Ozbun, M.A., Campos, S.K., Myers, M.P., Banks, L.** (2012). *Human papillomavirus L2 facilitates viral escape from late endosomes via sorting nexin 17.* Traffic **13(3)**, 455–467.

[104] **Holmgren, S.C., Patterson, N.A., Ozbun, M.A., Lambert, P.F.** (2005). *The minor capsid protein L2 contributes to two steps in the human papillomavirus type 31 life cycle.* Journal of Virology **79(7)**, 3938–3948.

[105] **Day, P.M., Lowy, D.R., Schiller, J.T.** (2008). *Heparan sulfate-independent cell binding and infection with furin-precleaved papillomavirus capsids.* Journal of Virology **82(24)**, 12565–12568.

[106] **Kamper, N., Day, P.M., Nowak, T., Selinka, H.C., Florin, L., Bolscher, J., Hilbig, L., Schiller, J.T., Sapp, M.** (2006). *A membrane-destabilizing peptide in capsid protein L2 is required for egress of papillomavirus genomes from endosomes.* Journal of Virology **80(2)**, 759–768.

[107] **Richards, R.M., Lowy, D.R., Schiller, J.T., Day, P.M.** (2006). *Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection.* Proceedings of the National Academy of Sciences **103(5)**, 1522–1527.

[108] **Yang, R., Day, P.M., Yutzy, W.H., Lin, K.Y., Hung, C.F., Roden, R.B.S.** (2003). *Cell surface-binding motifs of L2 that facilitate papillomavirus infection.* Journal of Virology **77(6)**, 3531–3541.

[109] **Florin, L., Becker, K.A., Lambert, C., Nowak, T., Sapp, C., Strand, D., Streeck, R.E., Sapp, M.** (2006). *Identification of a dynein interacting domain in the papillomavirus minor capsid protein L2.* Journal of Virology **80(13)**, 6691–6696.

[110] **Bossis, I., Roden, R.B.S., Gambhira, R., Yang, R., Tagaya, M., Howley, P.M., Meneses, P.I.** (2005). *Interaction of tSNARE syntaxin 18 with the papillomavirus minor capsid protein mediates infection.* Journal of Virology **79(11)**, 6723–6731.

[111] **Darshan, M.S., Lucchi, J., Harding, E., Moroianu, J.** (2004). *The L2 minor capsid protein of human papillomavirus type 16 interacts with a network of nuclear import receptors.* Journal of Virology **78(22)**, 12179–12188.

[112] **Day, P.M., Roden, R.B.S., Lowy, D.R., Schiller, J.T.** (1998). *The papillomavirus minor capsid protein, L2, induces localization of the major capsid protein, L1, and the viral transcription/replication protein, E2, to PML oncogenic domains.* Journal of Virology **72(1)**, 142–150.

[113] **Okun, M.M., Day, P.M., Greenstone, H.L., Booy, F.P., Lowy, D.R., Schiller, J.T., Roden, R.B.S.** (2001). *L1 interaction domains of papillomavirus L2 necessary for viral genome encapsidation.* Journal of Virology **75(9)**, 4332–4342.

[114] **Zhou, J., Sun, X.Y., Louis, K., Frazer, I.H.** (1994). *Interaction of human papillomavirus (HPV) type 16 capsid proteins with HPV DNA requires an intact L2 N-terminal sequence.* Journal of Virology **68(2)**, 619–625.

[115] **Heino, P., Zhou, J., Lambert, P.F.** (2000). *Interaction of the papillomavirus transcription/replication factor, E2, and the viral capsid protein, L2.* Virology **276(2)**, 304–14.

[116] **Zhou, J., Sun, X.Y., Stenzel, D.J., Frazer, I.H.** (1991). *Expression of vaccinia recombinant HPV 16 L1 and L2 ORF proteins in epithelial cells is sufficient for assembly of HPV virion-like particles.* Virology **185(1)**, 251–257.

[117] **Gambhira, R., Karanam, B., Jagu, S., Roberts, J.N., Buck, C.B., Bossis, I., Alphs, H., Culp, T., Christensen, N.D., Roden, R.B.S.** (2007). *A protective and broadly cross-neutralizing epitope of human papillomavirus L2.* Journal of Virology **81(24)**, 13927–13931.

[118] **Pastrana, D.V., Gambhira, R., Buck, C.B., Pang, Y.Y.S., Thompson, C.D., Culp, T.D., Christensen, N.D., Lowy, D.R., Schiller, J.T., Roden, R.B.** (2005). *Cross-neutralization of cutaneous and mucosal papillomavirus types with anti-sera to the amino terminus of L2.* Virology **337(2)**, 365–372.

[119] **Huh, W.K., Roden, R.B.** (2008). *The future of vaccines for cervical cancer.* Gynecologic Oncology **109(2)**, 48–56.

[120] **Ma, B., Maraj, B., Tran, N.P., Knoff, J., Chen, A., Alvarez, R.D., Hung, C.F., Wu, T.C.** (2012). *Emerging human papillomavirus vaccines.* Expert Opinion on Emerging Drugs **17(4)**, 469–492.

[121] **Gersch, E.D., Gissmann, L., Garcea, R.L.** (2012). *New approaches to prophylactic human papillomavirus vaccines for cervical cancer prevention.* Antivir Ther **17(3)**, 425–434.

[122] **Schelhaas, M., Ewers, H., Rajamäki, M.L., Day, P.M., Schiller, J.T., Helenius, A.** (2008). *Human papillomavirus type 16 entry: Retrograde cell surface transport along actin-rich protrusions.* PLoS Pathog **4(9)**, e1000148.

[123] **Roberts, J.N., Buck, C.B., Thompson, C.D., Kines, R., Bernardo, M., Choyke, P.L., Lowy, D.R., Schiller, J.T.** (2007). *Genital transmission of HPV in a mouse model is potentiated by nonoxynol-9 and inhibited by carrageenan.* Nat Med **13(7)**, 857–861.

[124] **Roden, R.B., Kirnbauer, R., Jenson, A.B., Lowy, D.R., Schiller, J.T.** (1994). *Interaction of papillomaviruses with the cell surface.* Journal of Virology **68(11)**, 7260–7266.

[125] **Muller, M., Gissmann, L., Cristiano, R.J., Sun, X.Y., Frazer, I.H., Jenson, A.B., Alonso, A., Zentgraf, H., Zhou, J.** (1995). *Papillomavirus capsid binding and uptake by cells from different tissues and species.* Journal of Virology **69(2)**, 948–54.

[126] **Sapp, M., Day, P.M.** (2009). *Structure, attachment and entry of polyoma- and papillomaviruses.* Virology **384(2)**, 400–409.

[127] **Combita, A.L., Touze, A., Bousarghin, L., Sizaret, P.Y., Munoz, N., Coursaget, P.** (2001). *Gene transfer using human papillomavirus pseudovirions varies according to virus genotype and requires cell surface heparan sulfate.* FEMS Microbiol Lett **204(1)**, 183–188.

[128] **Selinka, H.C., Florin, L., Patel, H.D., Freitag, K., Schmidtke, M., Makarov, V.A., Sapp, M.** (2007). *Inhibition of transfer to secondary receptors by heparan sulfate-binding drug or antibody induces noninfectious uptake of human papillomavirus.* Journal of Virology **81(20)**, 10970–10980.

[129] **Horvath, C., Boulet, G., Renoux, V., Delvenne, P., Bogers, J.P.** (2010). *Mechanisms of cell entry by human papillomaviruses: an overview.* Virology Journal **7(1)**, 11.

[130] **Buck, C.B., Thompson, C.D., Pang, Y.Y.S., Lowy, D.R., Schiller, J.T.** (2005). *Maturation of papillomavirus capsids.* Journal of Virology **79(5)**, 2839–2846.

[131] **Selinka, H.C., Giroglou, T., Sapp, M.** (2002). *Analysis of the infectious entry pathway of human papillomavirus type 33 pseudovirions.* Virology **299(2)**, 279–287.

[132] **Day, P.M., Lowy, D.R., Schiller, J.T.** (2003). *Papillomaviruses infect cells via a clathrin-dependent pathway.* Virology **307(1)**, 1–11.

[133] **Spoden, G., Freitag, K., Husmann, M., Boller, K., Sapp, M., Lambert, C., Florin, L.** (2008). *Clathrin- and caveolin-independent entry of human papillomavirus type 16: Involvement of tetraspanin-enriched microdomains (TEMs).* PLoS ONE **3(10)**, e3313.

[134] **Bousarghin, L., Touzé, A., Sizaret, P.Y., Coursaget, P.** (2003). *Human papillomavirus types 16, 31, and 58 use different endocytosis pathways to enter cells.* Journal of Virology **77(6)**, 3846–3850.

[135] **Smith, J.L., Campos, S.K., Ozbun, M.A.** (2007). *Human papillomavirus type 31 uses a caveolin 1- and dynamin 2-mediated entry pathway for infection of human keratinocytes.* Journal of Virology **81(18)**, 9922–9931.

[136] **Earp, L., Delos, S., Park, H., White, J.** (2005). *The many mechanisms of viral membrane fusion proteins.* Curr Top Microbiol Immunol **285(1)**, 25–66.

[137] **Meier, O., Greber, U.F.** (2004). *Adenovirus endocytosis.* The Journal of Gene Medicine **6(S1)**, 152–163.

[138] **Hogle, J.M.** (2002). *Poliovirus cell entry: Common structural themes in viral cell entry pathways.* Annual Review of Microbiology **56(1)**, 677–702.

[139] **Tsai, B.** (2007). *Penetration of nonenveloped viruses into the cytoplasm.* Annual Review of Cell and Developmental Biology **23(1)**, 23–43.

[140] **Gruenberg, J., van der Goot, F.G.** (2006). *Mechanisms of pathogen entry through the endosomal compartments.* Nat Rev Mol Cell Biol **7(7)**, 495–504.

[141] **Smith, J.L., Campos, S.K., Wandinger-Ness, A., Ozbun, M.A.** (2008). *Caveolin-1-dependent infectious entry of human papillomavirus type 31 in human keratinocytes proceeds to the endosomal pathway for ph-dependent uncoating.* Journal of Virology **82(19)**, 9505–9512.

[142] **Bronnimann, M.P., Chapman, J.A., Park, C.K., Campos, S.K.** (2013). *A transmembrane domain and GxxxG motifs within L2 are essential for papillomavirus infection.* Journal of Virology **87(1)**, 464–473.

[143] **Pyeon, D., Pearce, S.M., Lank, S.M., Ahlquist, P., Lambert, P.F.** (2009). *Establishment of human papillomavirus infection requires cell cycle progression.* PLoS Pathog **5(2)**, e1000318.

[144] **Day, P.M., Baker, C.C., Lowy, D.R., Schiller, J.T.** (2004). *Establishment of papillomavirus infection is enhanced by promyelocytic leukemia protein (PML) expression.* Proceedings of the National Academy of Sciences **101(39)**, 14252–14257.

[145] **Jones, D.L., Alani, R.M., Munger, K.** (1997). *The human papillomavirus E7 oncoprotein can uncouple cellular differentiation and proliferation in human keratinocytes by abrogating p21cip1-mediated inhibition of cdk2.* Genes Dev **11(16)**, 2101–2111.

[146] **Sherman, L., Schlegel, R.** (1996). *Serum- and calcium-induced differentiation of human keratinocytes is inhibited by the E6 oncoprotein of human papillomavirus type 16.* J Virol **70(5)**, 3269–3279.

[147] **Frattini, M.G., Lim, H.B., Doorbar, J., Laimins, L.A.** (1997). *Induction of human papillomavirus type 18 late gene expression and genomic amplification in organotypic cultures from transfected DNA templates.* Journal of Virology **71(9)**, 7068–72.

[148] **Ozbun, M.A., Meyers, C.** (1997). *Characterization of late gene transcripts expressed during vegetative replication of human papillomavirus type 31b.* Journal of Virology **71(7)**, 5161–72.

[149] **McBride, A.A.** (2008). *Chapter 4 replication and partitioning of papillomavirus genomes.* In A.J.S. Karl Maramorosch, F.A. Murphy, eds., *Advances in Virus Research*, volume 72, 155 – 205. Academic Press.

[150] **Swindle, C.S., Zou, N., Van Tine, B.A., Shaw, G.M., Engler, J.A., Chow, L.T.** (1999). *Human papillomavirus DNA replication compartments in a transient DNA replication system.* Journal of Virology **73(2)**, 1001–1009.

[151] **Schlegel, R., Phelps, W., Zhang, Y., Barbosa, M.** (1988). *Quantitative keratinocyte assay detects two biological activities of human papillomavirus DNA and identifies viral types associated with cervical carcinoma.* EMBO Journal **7(10)**, 3181–7.

[152] **Storey, A., Pim, D., Murray, A., Osborn, K., Banks, L., Crawford, L.** (1988). *Comparison of the in vitro transforming activities of human papillomavirus types.* EMBO Journal **7(6)**, 1815–20.

[153] **He, W., Staples, D., Smith, C., Fisher, C.** (2003). *Direct activation of cyclin-dependent kinase 2 by human papillomavirus E7.* Journal of Virology **77(19)**, 10566–10574.

[154] **Zerfass-Thome, K., Zwerschke, W., Mannhardt, B., Tindle, R., Botz, J.W., Jansen-Durr, P.** (1996). *Inactivation of the cdk inhibitor p27KIP1 by the human papillomavirus type 16 E7 oncoprotein.* Oncogene **13(11)**, 2323–30.

[155] **Funk, J.O., Waga, S., Harry, J.B., Espling, E., Stillman, B., Galloway, D.A.** (1997). *Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein.* Genes & Development **11(16)**, 2090–2100.

[156] **Longworth, M.S., Laimins, L.A.** (2004). *The binding of histone deacetylases and the integrity of zinc finger-like motifs of the E7 protein are essential for the life cycle of human papillomavirus type 31.* Journal of Virology **78(7)**, 3533–3541.

[157] **DiMaio, D., Mattoon, D.** (2001). *Mechanisms of cell transformation by papillomavirus E5 proteins.* Oncogene **20(54)**, 7866–73.

[158] **Lehoux, M., D'Abramo, C.M., Archambault, J.** (2009). *Molecular mechanisms of human papillomavirus-induced carcinogenesis.* Public Health Genomics **12(5-6)**, 268–80.

[159] **Veldman, T., Liu, X., Yuan, H., Schlegel, R.** (2003). *Human papillomavirus E6 and Myc proteins associate in vivo and bind to and cooperatively activate the telomerase reverse transcriptase promoter.* Proceedings of the National Academy of Sciences **100(14)**, 8211–8216.

[160] **Galloway, D., Gewin, L., Myers, H., Luo, W., Grandori, C., Katzenellenbogen, R., McDougall, J.** (2005). *Regulation of telomerase by human papillomaviruses.* Cold Spring Harbor Symposia on Quantitative Biology **70(1)**, 209–215.

[161] **Klaes, R., Woerner, S.M., Ridder, R., Wentzensen, N., Duerst, M., Schneider, A., Lotz, B., Melsheimer, P., Doeberitz, M.v.K.** (1999). *Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes.* Cancer Research **59(24)**, 6132–6136.

[162] **Hopman, A.H., Smedts, F., Dignef, W., Ummelen, M., Sonke, G., Mravunac, M., Vooijs, G.P., Speel, E.J.M., Ramaekers, F.C.** (2004). *Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical chromosome abnormalities.* The Journal of Pathology **202(1)**, 23–33.

[163] **Thorland, E.C., Myers, S.L., Gostout, B.S., Smith, D.I.** (2003). *Common fragile sites are preferential targets for HPV16 integrations in cervical tumors.* Oncogene **22(8)**, 1225–1237.

[164] **Wentzensen, N., Vinokurova, S., Doeberitz, M.v.K.** (2004). *Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract.* Cancer Research **64(11)**, 3878–3884.

[165] **Ziegert, C., Wentzensen, N., Vinokurova, S., Kisseljov, F., Einenkel, J., Hoeckel, M., von Knebel Doeberitz, M.** (2003). *A comprehensive analysis of HPV integration loci in anogenital lesions combining transcript and genome-based amplification techniques.* Oncogene **22(25)**, 3977–3984.

[166] **Van Tine, B.A., Kappes, J.C., Banerjee, N.S., Knops, J., Lai, L., Steenbergen, R.D.M., Meijer, C.L.J.M., Snijders, P.J.F., Chatis, P., Broker, T.R., Moen, P.T., Chow, L.T.** (2004). *Clonal selection for transcriptionally active viral oncogenes during progression to cancer.* Journal of Virology **78(20)**, 11172–11186.

[167] **Fields, B., Knipe, D., Howley, P.** (2007). *Fields Virology , 5e.* Knipe, Fields Virology. Wolters Kluwer Health/Lippincott Williams & Wilkins.

[168] **Bouallaga, I., Massicard, S., Yaniv, M., Thierry, F.** (2000). *An enhanceosome containing the jun b/fra-2 heterodimer and the hmg-i(y) architectural protein controls hpv 18 transcription.* EMBO reports **1(5)**, 422–7.

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

[170] **Durst, M., Croce, C.M., Gissmann, L., Schwarz, E., Huebner, K.** (1987). *Papillomavirus sequences integrate near cellular oncogenes in some cervical carcinomas.* Proceedings of the National Academy of Sciences **84(4)**, 1070–1074.

[171] **Ferber, M.J., Thorland, E.C., Brink, A.A., Rapp, A.K., Phillips, L.A., McGovern, R., Gostout, B.S., Cheung, T.H., Chung, T.K.H., Fu, W.Y., Smith, D.I.** (2003). *Preferential integration of human papillomavirus type 18 near the c-myc locus in cervical carcinoma.* Oncogene **22(46)**, 7233–7242.

[172] **Ferber, M.J., Montoya, D.P., Yu, C., Aderca, I., McGee, A., Thorland, E.C., Nagorney, D.M., Gostout, B.S., Burgart, L.J., Boix, L., Bruix, J., McMahon, B.J., Cheung, T.H., Chung, T.K.H., Wong, Y.F., Smith, D.I., Roberts, L.R.** (2003). *Integrations of the hepatitis b virus (HBV) and human papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers.* Oncogene **22(24)**, 3813–3820.

[173] **Lee, D., Kim, H.Z., Jeong, K.W., Shim, Y.S., Horikawa, I., Barrett, J.C., Choe, J.** (2002). *Human papillomavirus E2 down-regulates the human telomerase reverse transcriptase promoter.* Journal of Biological Chemistry **277(31)**, 27748–27756.

[174] **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.

[175] **Lazo, P.A.** (1999). *The molecular genetics of cervical carcinoma.* Br J Cancer **80(12)**, 2008–2018.

[176] **Dokianakis, D.N., Sourvinos, G., Sakkas, S., Athanasiadou, E., Spandidos, D.A.** (1998). *Detection of HPV and ras gene mutations in cervical smears from female genital lesions.* Oncol Rep **5(5)**, 1195–1198.

[177] **Xi, L., Wang, S., Wang, C., Xu, Q., Li, P., Tian, X., Wu, P., Wang, W., Deng, D., Zhou, J., Ma, D.** (2009). *The pro-angiogenic factors stimulated by human papillomavirus type 16 E6 and E7 protein in C33a and human fibroblasts.* Oncology reports **21(1)**, 25–31.

[178] **Hanahan, D., Folkman, J.** (1996). *Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis.* Cell **86(3)**, 353–364.

[179] **Toussaint-Smith, E., Donner, D.B., Roman, A.** (2004). *Expression of human papillomavirus type 16 E6 and E7 oncoproteins in primary foreskin keratinocytes is sufficient to alter the expression of angiogenic factors.* Oncogene **23(17)**, 2988–2995.

[180] **Chen, W., Li, F., Mead, L., White, H., Walker, J., Ingram, D., Roman, A.** (2007). *Human papillomavirus causes an angiogenic switch in keratinocytes which is sufficient to alter endothelial cell behavior.* Virology **367(1)**, 168–174.

[181] **Hanahan, D., Weinberg, R.A.** (2011). *Hallmarks of cancer: The next generation.* Cell **144(5)**, 646–674.

[182] **Chen, L., Ashe, S., Singhal, M.C., Galloway, D.A., Hellstrom, I., Hellstrom, K.E.** (1993). *Metastatic conversion of cells by expression of human papillomavirus type 16 E6 and E7 genes.* Proceedings of the National Academy of Sciences **90(14)**, 6523–6527.

[183] **Charette, S.T., McCance, D.J.** (2007). *The E7 protein from human papillomavirus type 16 enhances keratinocyte migration in an Akt-dependent manner.* Oncogene **26(52)**, 7386–7390.

[184] **Amine, A., Rivera, S., Opolon, P., Dekkal, M., Biard, D.S.F., Bouamar, H., Louache, F., McKay, M.J., Bourhis, J., Deutsch, E., Vozenin-Brotons, M.C.** (2009). *Novel anti-metastatic action of cidofovir mediated by inhibition of E6/E7, CXCR4 and Rho/ROCK signaling in HPV+ tumor cells.* PLoS ONE **4(3)**, e5018.

[185] **Frazer, I., Thomas, R., Zhou, J., Leggatt, G., Dunn, L., McMillan, N., Tindle, R., Filgueira, L., Manders, P., Barnard, P., Sharkey, M.** (1999). *Potential strategies utilised by papillomavirus to evade host immunity.* Immunol Rev **168(1)**, 131–42.

[186] **Stern, P.L.** (2005). *Immune control of human papillomavirus (HPV) associated anogenital disease and potential for vaccination.* J Clin Virol **32(1)**, 72–81.

[187] **Franco, E.L., Villa, L.L., Sobrinho, J.P., Prado, J.M., Rousseau, M.C., Désy, M., Rohan, T.E.** (1999). *Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer.* Journal of Infectious Diseases **180(5)**, 1415–1423.

[188] **Nakagawa, M., Stites, D.P., Farhat, S., Sisler, J.R., Moss, B., Kong, F., Moscicki, A.B., Palefsky, J.M.** (1997). *Cytotoxic T lymphocyte responses to E6 and E7 proteins of human papillomavirus type 16: Relationship to cervical intraepithelial neoplasia.* Journal of Infectious Diseases **175(4)**, 927–931.

[189] **Evans, C., Bauer, S., Grubert, T., Brucker, Cand Baur, S., Heeg, K., Wagner, H., Lipford, G.B.** (1996). *HLA-A2-restricted peripheral blood cytolytic T lymphocyte response to HPV type 16 proteins E6 and E7 from patients with neoplastic cervical lesions.* Cancer Immunology, Immunotherapy **42(3)**, 151–60.

[190] **Albarran-Somoza, B., Franco-Topete, R., Delgado-Rizo, V., Cerda-Camacho, F., Acosta-Jimenez, L., Lopez-Botet, M., Daneri-Navarro, A.** (2006). *CEACAM1 in cervical cancer and precursor lesions: Association with human papillomavirus infection.* Journal of Histochemistry & Cytochemistry **54(12)**, 1393–1399.

[191] **Kanodia, S., Fahey, L., Kast, W.** (2007). *Mechanisms used by human papillomaviruses to escape the host immune response.* Curr Cancer Drug Targets **7(1)**, 79–89.

[192] **Muller, M., Demeret, C.** (2012). *The HPV E2-host protein-protein interactions: A complex hijacking of the cellular network.* Open Virol J **6(1)**, 173–89.

[193] **Giri, I., Yaniv, M.** (1988). *Structural and mutational analysis of E2 transactivating proteins of papillomaviruses reveals three distinct functional domains.* EMBO Journal **7(9)**, 2823–9.

[194] **Hegde, R.S., Androphy, E.J.** (1998). *Crystal structure of the E2 DNA-binding domain from human papillomavirus type 16: implications for its DNA binding-site selection mechanism.* Journal of Molecular Biology **284(5)**, 1479–1489.

[195] **Antson, A.A., Burns, J.E., Moroz, O.V., Scott, D.J., Sanders, C.M., Bronstein, I.B., Dodson, G.G., Wilson, K.S., Maitland, N.J.** (2000). *Structure of the intact transactivation domain of the human papillomavirus E2 protein.* Nature **403(6771)**, 805–809.

[196] **Harris, S., Botchan, M.** (1999). *Crystal structure of the human papillomavirus type 18 E2 activation domain.* Science **284(5420)**, 1673–7.

[197] **Wang, Y., Coulombe, R., Cameron, D.R., Thauvette, L., Massariol, M.J., Amon, L.M., Fink, D., Titolo, S., Welchner, E., Yoakim, C., Archambault, J., White, P.W.** (2004). *Crystal structure of the E2 transactivation domain of human papillomavirus type 11 bound to a protein interaction inhibitor.* Journal of Biological Chemistry **279(8)**, 6976–6985.

[198] **Bochkarev, A., Barwell, J.A., Pfuetzner, R.A., Bochkareva, E., Frappier, L., Edwards, A.M.** (1996). *Crystal structure of the DNA-binding domain of the epstein-barr virus origin-binding protein, EBNA1, bound to DNA.* Cell **84(5)**, 791–800.

[199] **Dostatni, N., Thierry, F., Yaniv, M.** (1988). *A dimer of BPV-1 E2 containing a protease resistant core interacts with its DNA target.* EMBO J **7(12)**, 3807–3816.

[200] **McBride, A.A., Byrne, J.C., Howley, P.M.** (1989). *E2 polypeptides encoded by bovine papillomavirus type 1 form dimers through the common carboxyl-terminal domain: transactivation is mediated by the conserved amino-terminal domain.* Proceedings of the National Academy of Sciences **86(2)**, 510–514.

[201] **Ham, J., Dostatni, N., Gauthier, J.M., Yaniv, M.** (1991). *The papillomavirus E2 protein: a factor with many talents.* Trends in Biochemical Sciences **16(0)**, 440 – 444.

[202] **Zou, N., Lin, B.Y., Duan, F., Lee, K.Y., Jin, G., Guan, R., Yao, G., Lefkowitz, E.J., Broker, T.R., Chow, L.T.** (2000). *The hinge of the human papillomavirus type 11 E2 protein contains major determinants for nuclear localization and nuclear matrix association.* Journal of Virology **74(8)**, 3761–3770.

[203] **Demeret, C., Desaintes, C., Yaniv, M., Thierry, F.** (1997). *Different mechanisms contribute to the E2-mediated transcriptional repression of human papillomavirus type 18 viral oncogenes.* J Virol **71(12)**, 9343–9.

[204] **Sanchez, I.E., Dellarole, M., Gaston, K., de Prat Gay, G.** (2008). *Comprehensive comparison of the interaction of the E2 master regulator with its cognate target DNA sites in 73 human papillomavirus types by sequence statistics.* Nucleic Acids Res **36(3)**, 756–69.

[205] **Stubenrauch, F., Pfister, H.** (1994). *Low-affinity E2-binding site mediates downmodulation of E2 transactivation of the human papillomavirus type 8 late promoter.* J Virol **68(11)**, 6959–66.

[206] **Hou, S.Y., Wu, S.Y., Chiang, C.M.** (2002). *Transcriptional activity among high and low risk human papillomavirus E2 proteins correlates with E2 DNA binding.* J Biol Chem **277(47)**, 45619–29.

[207] **Dong, G., Broker, T.R., Chow, L.T.** (1994). *Human papillomavirus type 11 E2 proteins repress the homologous E6 promoter by interfering with the binding of host transcription factors to adjacent elements.* J Virol **68(2)**, 1115–27.

[208] **Dostatni, N., Lambert, P.F., Sousa, R., Ham, J., Howley, P.M., Yaniv, M.** (1991). *The functional BPV-1 E2 trans-activating protein can act as a repressor by preventing formation of the initiation complex.* Genes Dev **5(9)**, 1657–71.

[209] **Smith, J.A., White, E.A., Sowa, M.E., Powell, M.L., Ottinger, M., Harper, J.W., Howley, P.M.** (2010). *Genome-wide siRNA screen identifies SMCX, EP400, and Brd4 as E2-dependent regulators of human papillomavirus oncogene expression.* Proc Natl Acad Sci U S A **107(8)**, 3752–7.

[210] **Grassmann, K., Rapp, B., Maschek, H., Petry, K.U., Iftner, T.** (1996). *Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA.* Journal of Virology **70(4)**, 2339–49.

[211] **Hummel, M., Hudson, J.B., Laimins, L.A.** (1992). *Differentiation-induced and constitutive transcription of human papillomavirus type 31b in cell lines containing viral episomes.* Journal of Virology **66(10)**, 6070–6080.

[212] **Thierry, F., Dostatni, N., Arnos, F., Yaniv, M.** (1990). *Cooperative activation of transcription by bovine papillomavirus type 1 E2 can occur over a large distance.* Mol Cell Biol **10(8)**, 4431–7.

[213] **Kovelman, R., Bilter, G.K., Glezer, E., Tsou, A.Y., Barbosa, M.S.** (1996). *Enhanced transcriptional activation by E2 proteins from the oncogenic human papillomaviruses.* J Virol **70(11)**, 7549–60.

[214] **Goodrich, J.A., Tjian, R.** (1994). *TBP-TAF complexes: selectivity factors for eukaryotic transcription.* Curr Opin Cell Biol **6(3)**, 403–9.

[215] **Enzenauer, C., Mengus, G., Lavigne, A., Davidson, I., Pfister, H., May, M.** (1998). *Interaction of human papillomavirus 8 regulatory proteins E2, E6 and E7 with components of the TFIID complex.* Intervirology **41(2-3)**, 80–90.

[216] **Carrillo, E., Garrido, E., Gariglio, P.** (2004). *Specific in vitro interaction between papillomavirus E2 proteins and TBP-associated factors.* Intervirology **47(6)**, 342–9.

[217] **Benson, J.D., Lawande, R., Howley, P.M.** (1997). *Conserved interaction of the papillomavirus E2 transcriptional activator proteins with human and yeast TFIIB proteins.* J Virol **71(10)**, 8041–7.

[218] **Steger, G., Schnabel, C., Schmidt, H.M.** (2002). *The hinge region of the human papillomavirus type 8 E2 protein activates the human p21(WAF1/CIP1) promoter via interaction with Sp1.* J Gen Virol **83(3)**, 503–10.

[219] **Olejnik-Schmidt, A.K., Schmidt, M.T., Kedzia, W., Gozdzicka-Jozefiak, A.** (2008). *Search for cellular partners of human papillomavirus type 16 E2 protein.* Arch Virol **153(5)**, 983–90.

[220] **Olejnik-Schmidt, A.K., Schmidt, M.T., Gozdzicka-Jozefiak, A.** (2006). *Orphan nuclear hormone receptor NR4A1 interacts with HPV16 E2 regulatory protein.* Cell Mol Biol Lett **11(1)**, 102–8.

[221] **Hadaschik, D., Hinterkeuser, K., Oldak, M., Pfister, H.J., Smola-Hess, S.** (2003). *The papillomavirus E2 protein binds to and synergizes with C/EBP factors involved in keratinocyte differentiation.* J Virol **77(9)**, 5253–65.

[222] **Wu, S.Y., Chiang, C.M.** (2007). *The double bromodomain-containing chromatin adaptor brd4 and transcriptional regulation.* J Biol Chem **282(18)**, 13141–5.

[223] **Abbate, E.A., Voitenleitner, C., Botchan, M.R.** (2006). *Structure of the papillomavirus DNA-tethering complex E2:brd4 and a peptide that ablates HPV chromosomal association.* Mol Cell **24(6)**, 877–89.

[224] **Senechal, H., Poirier, G.G., Coulombe, B., Laimins, L.A., Archambault, J.** (2007). *Amino acid substitutions that specifically impair the transcriptional activity of papillomavirus E2 affect binding to the long isoform of brd4.* Virology **358(1)**, 10–7.

[225] **Lee, D., Hwang, S.G., Kim, J., Choe, J.** (2002). *Functional interaction between p/CAF and human papillomavirus E2 protein.* J Biol Chem **277(8)**, 6483–9.

[226] **Lee, D., Lee, B., Kim, J., Kim, D.W., Choe, J.** (2000). *cAMP response element-binding protein-binding protein binds to human papillomavirus E2 protein and activates E2-dependent transcription.* J Biol Chem **275(10)**, 7045–51.

[227] **Muller, A., Ritzkowsky, A., Steger, G.** (2002). *Cooperative activation of human papillomavirus type 8 gene expression by the E2 protein and the cellular coactivator p300.* J Virol **76(21)**, 11042–53.

[228] **Cha, S., Seo, T.** (2011). *hSNF5 is required for human papillomavirus E2-driven transcriptional activation and DNA replication.* Intervirology **54(2)**, 66–77.

[229] **Rehtanz, M., Schmidt, H.M., Warthorst, U., Steger, G.** (2004). *Direct interaction between nucleosome assembly protein 1 and the papillomavirus E2 proteins involved in activation of transcription.* Mol Cell Biol **24(5)**, 2153–68.

[230] **Oldak, M., Smola, H., Aumailley, M., Rivero, F., Pfister, H., Smola-Hess, S.** (2004). *The human papillomavirus type 8 E2 protein suppresses beta4-integrin expression in primary human keratinocytes.* J Virol **78(19)**, 10738–46.

[231] **Mole, S., Milligan, S.G., Graham, S.V.** (2009). *Human papillomavirus type 16 E2 protein transcriptionally activates the promoter of a key cellular splicing factor, SF2/ASF.* J Virol **83(1)**, 357–67.

[232] **Akgul, B., Garcia-Escudero, R., Ekechi, C., Steger, G., Navsaria, H., Pfister, H., Storey, A.** (2011). *The E2 protein of human papillomavirus type 8 increases the expression of matrix metalloproteinase-9 in human keratinocytes and organotypic skin cultures.* Med Microbiol Immunol **200(2)**, 127–35.

[233] **Lee, D., Kim, H.Z., Jeong, K.W., Shim, Y.S., Horikawa, I., Barrett, J.C., Choe, J.** (2002). *Human papillomavirus E2 down-regulates the human telomerase reverse transcriptase promoter.* J Biol Chem **277(31)**, 27748–56.

[234] **Chiang, C., Ustav, M., Stenlund, A., Ho, T., Broker, T., Chow, L.** (1992). *Viral E1 and E2 proteins support replication of homologous and heterologous papillomaviral origins.* Proceedings of the National Academy of Sciences **89(13)**, 5799–5803.

[235] **Stenlund, A.** (2003). *E1 initiator DNA binding specificity is unmasked by selective inhibition of non-specific DNA binding.* EMBO J **22(4)**, 954–963.

[236] **Li, R., Botchan, M.R.** (1994). *Acidic transcription factors alleviate nucleosome-mediated repression of DNA replication of bovine papillomavirus type 1.* Proceedings of the National Academy of Sciences **91(15)**, 7051–7055.

[237] **Sedman, T., Sedman, J., Stenlund, A.** (1997). *Binding of the E1 and E2 proteins to the origin of replication of bovine papillomavirus.* Journal of Virology **71(4)**, 2887–96.

[238] **Schuck, S., Stenlund, A.** (2005). *Assembly of a double hexameric helicase.* Molecular Cell **20(3)**, 377–389.

[239] **Li, R., Botchan, M.R.** (1993). *The acidic transcriptional activation domains of VP16 and p53 bind the cellular replication protein A and stimulate in vitro BPV-1 DNA replication.* Cell **73(6)**, 1207–1221.

[240] **Ilves, I., Kivi, S., Ustav, M.** (1999). *Long-term episomal maintenance of bovine papillomavirus type 1 plasmids is determined by attachment to host chromosomes, which is mediated by the viral E2 protein and its binding sites.* J Virol **73(5)**, 4404–12.

[241] **McBride, A.A., Oliveira, J.G., McPhillips, M.G.** (2006). *Partitioning viral genomes in mitosis: same idea, different targets.* Cell Cycle **5(14)**, 1499–502.

[242] **Oliveira, J.G., Colf, L.A., McBride, A.A.** (2006). *Variations in the association of papillomavirus E2 proteins with mitotic chromosomes.* Proceedings of the National Academy of Sciences **103(4)**, 1047–52.

[243] **Skiadopoulos, M.H., McBride, A.A.** (1998). *Bovine papillomavirus type 1 genomes and the E2 transactivator protein are closely associated with mitotic chromatin.* Journal of Virology **72(3)**, 2079–2088.

[244] **Bastien, N., McBride, A.A.** (2000). *Interaction of the papillomavirus E2 protein with mitotic chromosomes.* Virology **270(1)**, 124–134.

[245] **You, J., Croyle, J.L., Nishimura, A., Ozato, K., Howley, P.M.** (2004). *Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes.* Cell **117(3)**, 349–60.

[246] **McPhillips, M.G., Ozato, K., McBride, A.A.** (2005). *Interaction of bovine papillomavirus E2 protein with Brd4 stabilizes its association with chromatin.* Journal of Virology **79(14)**, 8920–8932.

[247] **Dey, A., Chitsaz, F., Abbasi, A., Misteli, T., Ozato, K.** (2003). *The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis.* Proceedings of the National Academy of Sciences **100(15)**, 8758–8763.

[248] **Baxter, M.K., McPhillips, M.G., Ozato, K., McBride, A.A.** (2005). *The mitotic chromosome binding activity of the papillomavirus E2 protein correlates with interaction with the cellular chromosomal protein, Brd4.* Journal of Virology **79(8)**, 4806–4818.

[249] **Kruhlak, M.J., Hendzel, M.J., Fischle, W., Bertos, N.R., Hameed, S., Yang, X.J., Verdin, E., Bazett-Jones, D.P.** (2001). *Regulation of global acetylation in mitosis through loss of histone acetyltransferases and deacetylases from chromatin.* Journal of Biological Chemistry **276(41)**, 38307–38319.

[250] **McPhillips, M.G., Oliveira, J.G., Spindler, J.E., Mitra, R., McBride, A.A.** (2006). *Brd4 is required for E2-mediated transcriptional activation but not genome partitioning of all papillomaviruses.* J Virol **80(19)**, 9530–43.

[251] **Parish, J.L., Bean, A.M., Park, R.B., Androphy, E.J.** (2006). *ChlR1 is required for loading papillomavirus E2 onto mitotic chromosomes and viral genome maintenance.* Mol Cell **24(6)**, 867–76.

[252] **Yu, T., Peng, Y.C., Androphy, E.J.** (2007). *Mitotic kinesin-like protein 2 binds and colocalizes with papillomavirus E2 during mitosis.* J Virol **81(4)**, 1736–45.

[253] **Valcarcel, J., Green, M.R.** (1996). *The SR protein family: pleiotropic functions in pre-mRNA splicing.* Trends in Biochemical Sciences **21(8)**, 296 – 301.

[254] **Manley, J.L., Tacke, R.** (1996). *SR proteins and splicing control.* Genes & Development **10(13)**, 1569–1579.

[255] **Fu, X.D.** (1995). *The superfamily of arginine/serine-rich splicing factors.* RNA **1(7)**, 663–680.

[256] **Wu, J.Y., Maniatis, T.** (1993). *Specific interactions between proteins implicated in splice site selection and regulated alternative splicing.* Cell **75(6)**, 1061–1070.

[257] **Lai, M.C., Teh, B.H., Tarn, W.Y.** (1999). *A human papillomavirus E2 transcriptional activator. the interactions with cellular splicing factors and potential function in pre-mRNA processing.* J Biol Chem **274(17)**, 11832–41.

[258] **Lai, M.C., Lin, R.I., Huang, S.Y., Tsai, C.W., Tarn, W.Y.** (2000). *A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins.* J Biol Chem **275(11)**, 7950–7.

[259] **McFarlane, M., Graham, S.V.** (2010). *Human papillomavirus regulation of SR proteins.* Biochem Soc Trans **38(4)**, 1116–21.

[260] **Johansson, C., Somberg, M., Li, X., Backstrom Winquist, E., Fay, J., Ryan, F., Pim, D., Banks, L., Schwartz, S.** (2012). *HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation.* EMBO Journal **31(14)**, 3212–27.

[261] **Blachon, S., Demeret, C.** (2003). *The regulatory E2 proteins of human genital papillomaviruses are pro-apoptotic.* Biochimie **85(8)**, 813–9.

[262] **Blachon, S., Bellanger, S., Demeret, C., Thierry, F.** (2005). *Nucleo-cytoplasmic shuttling of high risk human papillomavirus E2 proteins induces apoptosis.* J Biol Chem **280(43)**, 36088–98.

[263] **Parish, J.L., Kowalczyk, A., Chen, H.T., Roeder, G.E., Sessions, R., Buckle, M., Gaston, K.** (2006). *E2 proteins from high- and low-risk human papillomavirus types differ in their ability to bind p53 and induce apoptotic cell death.* J Virol **80(9)**, 4580–90.

[264] **Frattini, M.G., Hurst, S.D., Lim, H.B., Swaminathan, S., Laimins, L.A.** (1997). *Abrogation of a mitotic checkpoint by E2 proteins from oncogenic human papillomaviruses correlates with increased turnover of the p53 tumor suppressor protein.* EMBO J **16(2)**, 318–331.

[265] **Desaintes, C., Demeret, C., Goyat, S., Yaniv, M., Thierry, F.** (1997). *Expression of the papillomavirus E2 protein in HeLa cells leads to apoptosis.* EMBO Journal **16(3)**, 504–14.

[266] **Webster, K., Parish, J., Pandya, M., Stern, P.L., Clarke, A.R., Gaston, K.** (2000). *The human papillomavirus (HPV) 16 E2 protein induces apoptosis in the absence of other HPV proteins and via a p53-dependent pathway.* J Biol Chem **275(1)**, 87–94.

[267] **Demeret, C., Garcia-Carranca, A., Thierry, F.** (2003). *Transcription-independent triggering of the extrinsic pathway of apoptosis by human papillomavirus 18 E2 protein.* Oncogene **22(2)**, 168–75.

[268] **Thierry, F., Demeret, C.** (2008). *Direct activation of caspase 8 by the proapoptotic E2 protein of HPV18 independent of adaptor proteins.* Cell Death Differ **15(9)**, 1356–63.

[269] **Wang, W., Fang, Y., Sima, N., Li, Y., Li, W., Li, L., Han, L., Liao, S., Han, Z., Gao, Q., Li, K., Deng, D., Meng, L., Zhou, J., Wang, S., Ma, D.** (2011). *Triggering of death receptor apoptotic signaling by human papillomavirus 16 E2 protein in cervical cancer cell lines is mediated by interaction with c-FLIP.* Apoptosis **16(1)**, 55–66.

[270] **Strasswimmer, J., Lorson, C.L., Breiding, D.E., Chen, J.J., Le, T., Burghes, A.H., Androphy, E.J.** (1999). *Identification of survival motor neuron as a transcriptional activator-binding protein.* Hum Mol Genet **8(7)**, 1219–26.

[271] **Garcera, A., Mincheva, S., Gou-Fabregas, M., Caraballo-Miralles, V., Llado, J., Comella, J.X., Soler, R.M.** (2011). *A new model to study spinal muscular atrophy: neurite degeneration and cell death is counteracted by BCL-X(L) overexpression in motoneurons.* Neurobiol Dis **42(3)**, 415–26.

[272] **Woronicz, J.D., Calnan, B., Ngo, V., Winoto, A.** (1994). *Requirement for the orphan steroid receptor Nur77 in apoptosis of T-cell hybridomas.* Nature **367(6460)**, 277–81.

[273] **Li, H., Kolluri, S.K., Gu, J., Dawson, M.I., Cao, X., Hobbs, P.D., Lin, B., Chen, G., Lu, J., Lin, F., Xie, Z., Fontana, J.A., Reed, J.C., Zhang, X.** (2000). *Cytochrome c release and apoptosis induced by mitochondrial targeting of nuclear orphan receptor TR3.* Science **289(5482)**, 1159–64.

[274] **Holmes, W.F., Soprano, D.R., Soprano, K.J.** (2003). *Comparison of the mechanism of induction of apoptosis in ovarian carcinoma cells by the conformationally restricted synthetic retinoids CD437 and 4-HPR.* J Cell Biochem **89(2)**, 262–78.

[275] **Lee, J.M., Lee, K.H., Weidner, M., Osborne, B.A., Hayward, S.D.** (2002). *Epstein-Barr virus EBNA2 blocks Nur77- mediated apoptosis.* Proceedings of the National Academy of Sciences **99(18)**, 11878–83.

[276] **Lin, B., Kolluri, S.K., Lin, F., Liu, W., Han, Y.H., Cao, X., Dawson, M.I., Reed, J.C., Zhang, X.K.** (2004). *Conversion of Bcl-2 from protector to killer by interaction with nuclear orphan receptor Nur77/TR3.* Cell **116(4)**, 527–40.

[277] **Sordet, O., Goldman, A., Pommier, Y.** (2006). *Topoisomerase II and tubulin inhibitors both induce the formation of apoptotic topoisomerase I cleavage complexes.* Mol Cancer Ther **5(12)**, 3139–44.

[278] **Dowhanick, J.J., McBride, A.A., Howley, P.M.** (1995). *Suppression of cellular proliferation by the papillomavirus E2 protein.* Journal of Virology **69(12)**, 7791–9.

[279] **Bellanger, S., Blachon, S., Mechali, F., Bonne-Andrea, C., Thierry, F.** (2005). *High-risk but not low-risk hpv E2 proteins bind to the apc activators cdh1 and cdc20 and cause genomic instability.* Cell Cycle **4(11)**, 1608–15.

[280] **Wei, W., Ayad, N.G., Wan, Y., Zhang, G.J., Kirschner, M.W., Kaelin, W. G., J.** (2004). *Degradation of the SCF component Skp2 in cell-cycle phase G1 by the anaphase-promoting complex.* Nature **428(6979)**, 194–8.

[281] **Bornstein, G., Bloom, J., Sitry-Shevah, D., Nakayama, K., Pagano, M., Hershko, A.** (2003). *Role of the SCFSkp2 ubiquitin ligase in the degradation of p21Cip1 in S phase.* J Biol Chem **278(28)**, 25752–7.

[282] **Bellanger, S., Tan, C.L., Nei, W., He, P.P., Thierry, F.** (2010). *The human papillomavirus type 18 E2 protein is a cell cycle-dependent target of the SCFSkp2 ubiquitin ligase.* J Virol **84(1)**, 437–44.

[283] **Bernassola, F., Karin, M., Ciechanover, A., Melino, G.** (2008). *The HECT family of E3 ubiquitin ligases: multiple players in cancer development.* Cancer Cell **14(1)**, 10–21.

[284] **Rotin, D., Kumar, S.** (2009). *Physiological functions of the HECT family of ubiquitin ligases.* Nat Rev Mol Cell Biol **10(6)**, 398–409.

[285] **Bosu, D.R., Kipreos, E.T.** (2008). *Cullin-RING ubiquitin ligases: global regulation and activation cycles.* Cell Div **3(1)**, 7–20.

[286] **Bellanger, S., Demeret, C., Goyat, S., Thierry, F.** (2001). *Stability of the human papillomavirus type 18 E2 protein is regulated by a proteasome degradation pathway through its amino-terminal transactivation domain.* J Virol **75(16)**, 7244–51.

[287] **Gagnon, D., Joubert, S., Senechal, H., Fradet-Turcotte, A., Torre, S., Archambault, J.** (2009). *Proteasomal degradation of the papillomavirus E2 protein is inhibited by overexpression of bromodomain-containing protein 4.* J Virol **83(9)**, 4127–39.

[288] **Zheng, G., Schweiger, M.R., Martinez-Noel, G., Zheng, L., Smith, J.A., Harper, J.W., Howley, P.M.** (2009). *Brd4 regulation of papillomavirus protein E2 stability.* J Virol **83(17)**, 8683–92.

[289] **Furukawa, M., He, Y.J., Borchers, C., Xiong, Y.** (2003). *Targeting of protein ubiquitination by BTB-Cullin 3-roc1 ubiquitin ligases.* Nat Cell Biol **5(11)**, 1001–7.

[290] **Wang, X., Naidu, S.R., Sverdrup, F., Androphy, E.J.** (2009). *Tax1BP1 interacts with papillomavirus E2 and regulates E2-dependent transcription and stability.* J Virol **83(5)**, 2274–84.

[291] **Lee, A.Y., Chiang, C.M.** (2009). *Chromatin adaptor Brd4 modulates E2 transcription activity and protein stability.* J Biol Chem **284(5)**, 2778–86.

[292] **Green, H.** (1977). *Terminal differentiation of cultured human epidermal cells.* Cell **11(2)**, 405–416.

[293] **Zhu, S., Oh, H.S., Shim, M., Sterneck, E., Johnson, P.F., Smart, R.C.** (1999). *C/EBPβ modulates the early events of keratinocyte differentiation involving growth arrest and keratin 1 and keratin 10 expression.* Molecular and Cellular Biology **19(10)**, 7181–7190.

[294] **Ramirez-Salazar, E., Centeno, F., Nieto, K., Valencia-Hernandez, A., Salcedo, M., Garrido, E.** (2011). *Hpv16 E2 could act as down-regulator in cellular genes implicated in apoptosis, proliferation and cell differentiation.* Virol J **8(1)**, 247.

[295] **Burns, J.E., Walker, H.F., Schmitz, C., Maitland, N.J.** (2010). *Phenotypic effects of HPV-16 E2 protein expression in human keratinocytes.* Virology **401(2)**, 314–21.

[296] **Boulabiar, M., Boubaker, S., Favre, M., Demeret, C.** (2011). *Keratinocyte sensitization to tumour necrosis factor-induced nuclear factor kappa B activation by the E2 regulatory protein of human papillomaviruses.* J Gen Virol **92(1)**, 2422–7.

[297] **Basile, J.R., Zacny, V., Munger, K.** (2001). *The cytokines tumor necrosis factor-alpha (TNF-α) and TNF-related apoptosis-inducing ligand differentially modulate proliferation and apoptotic pathways in human keratinocytes expressing the human papillomavirus-16 E7 oncoprotein.* J Biol Chem **276(25)**, 22522–8.

[298] **Tiala, I., Suomela, S., Huuhtanen, J., Wakkinen, J., Holtta-Vuori, M., Kainu, K., Ranta, S., Turpeinen, U., Hamalainen, E., Jiao, H., Karvonen, S.L., Ikonen, E., Kere, J., Saarialho-Kere, U., Elomaa, O.** (2007). *The CCHCR1 (HCR) gene is relevant for skin steroidogenesis and downregulated in cultured psoriatic keratinocytes.* J Mol Med (Berl) **85(6)**, 589–601.

[299] **Suomela, S., Elomaa, O., Skoog, T., Ala-aho, R., Jeskanen, L., Parssinen, J., Latonen, L., Grenman, R., Kere, J., Kahari, V.M., Saarialho-Kere, U.** (2009). *CCHCR1 is up-regulated in skin cancer and associated with EGFR expression.* PLoS One **4(6)**, e6030.

[300] **Cao, X.X., Xu, J.D., Xu, J.W., Liu, X.L., Cheng, Y.Y., Li, Q.Q., Xu, Z.D., Liu, X.P.** (2011). *RACK1 promotes breast carcinoma migration/metastasis via activation of the rhoA/rho kinase pathway.* Breast Cancer Res Treat **126(3)**, 555–63.

[301] **D'Abramo, C.M., Archambault, J.** (2011). *Small molecule inhibitors of human papillomavirus protein - protein interactions.* Open Virol J **5(1)**, 80–95.

[302] **Davey, N.E., Trave, G., Gibson, T.J.** (2011). *How viruses hijack cell regulation.* Trends Biochem Sci **36(3)**, 159–69.

[303] **Dyer, M.D., Murali, T.M., Sobral, B.W.** (2008). *The landscape of human proteins interacting with viruses and other pathogens.* PLoS Pathog **4(2)**, e32.

[304] **Uetz, P., Rajagopala, S., Dong, Y., Haas, J.** (2004). *From ORFeomes to protein interaction maps in viruses.* Genome Res **14(10)**, 2029–2033.

[305] **Bonetta, L.** (2010). *Protein-protein interactions: Interactome under construction.* Nature **468(7325)**, 851–854.

[306] **Vidal, M., Cusick, M.E., Barabasi, A.L.** (2011). *Interactome networks and human disease.* Cell **144(6)**, 986–98.

[307] **Carvunis, A., Roth, F., Calderwood, M., Cusick, M., Superti-Furga, G., Vidal, M.** (2013). *Chapter 3 - interactome networks.* In *Handbook of Systems Biology*, 45 – 63. Academic Press, San Diego.

[308] **Rozenblatt-Rosen, O., Deo, R., Padi, M., Adelmant, G., Calderwood, M., Rolland, T., Grace, M., Dricot, A., Askenazi, M., Tavares, M., Pevzner, S., Abderazzaq, F., Byrdsong, D., Carvunis, A., Chen, A., Cheng, J., Correll, M., Duarte, M., Fan, C., Feltkamp, M., Ficarro, S., Franchi, R., Garg, B., Gulbahce, N., Hao, T., Holthaus, A., James, R., Korkhin, A., Litovchick, L., Mar, J., Pak, T., Rabello, S., Rubio, R., Shen, Y., Singh, S., Spangle, J., Tasan, M., Wanamaker, S., Webber, J., Roecklein-Canfield, J., Johannsen, E., Barabasi, A.L., Beroukhim, R., Kieff, E., Cusick, M., Hill, D., Munger, K., Marto, J., Quackenbush, J., Roth, F., DeCaprio, J., Vidal, M.** (2012). *Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins.* Nature **487(7408)**, 491–495.

[309] **Venkatesan, K., Rual, J.F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X., Goh, K.I., Yildirim, M.A., Simonis, N., Heinzmann, K., Gebreab, F., Sahalie, J.M., Cevik, S., Simon, C., de Smet, A.S., Dann, E., Smolyar, A., Vinayagam, A., Yu, H., Szeto, D., Borick, H., Dricot, A., Klitgord, N., Murray, R.R., Lin, C., Lalowski, M., Timm, J., Rau, K., Boone, C., Braun, P., Cusick, M.E., Roth, F.P., Hill, D.E., Tavernier, J., Wanker, E.E., Barabasi, A.L., Vidal, M.** (2009). *An empirical framework for binary interactome mapping.* Nat Methods **6(1)**, 83–90.

[310] **Eyckerman, S., Verhee, A., der Heyden, J., Lemmens, I., Ostade, X., Vandekerckhove, J., Tavernier, J.** (2001). *Design and application of a cytokine-receptor-based interaction trap.* Nat Cell Biol **3(12)**, 1114–1119.

[311] **Magliery, T., Wilson, C., Pan, W., Mishler, D., Ghosh, I., Hamilton, A., Regan, L.** (2005). *Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism.* J Am Chem Soc **127(1)**, 146–157.

[312] **Remy, I., Michnick, S.W.** (2006). *A highly sensitive protein-protein interaction assay based on gaussia luciferase.* Nat Methods **3(12)**, 977–9.

[313] **Cassonnet, P., Rolloy, C., Neveu, G., Vidalain, P.O., Chantier, T., Pellet, J., Jones, L., Muller, M., Demeret, C., Gaud, G., Vuillier, F., Lotteau, V., Tangy, F., Favre, M., Jacob, Y.** (2011). *Benchmarking a luciferase complementation assay for detecting protein complexes.* Nat Methods **8(12)**, 990–2.

[314] **Haynes, C., Oldfield, C.J., Ji, F., Klitgord, N., Cusick, M.E., Radivojac, P., Uversky, V.N., Vidal, M., Iakoucheva, L.M.** (2006). *Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes.* PLoS Comput Biol **2(8)**, e100.

[315] **Neveu, G., Cassonnet, P., Vidalain, P., Rolloy, C., Mendoza, J., Jones, L., Tangy, F., Muller, M., Demeret, C., Tafforeau, L., Lotteau, V., Rabourdin-Combe, C., Trave, G., Dricot, A., Hill, D., Vidal, M., Favre, M., Jacob, Y.** (2012). *Comparative analysis of virus-host interactomes with a*

mammalian high-throughput protein complementation assay based on gaussia princeps luciferase. Methods **58(4)**, 349–359.

[316] **Barabasi, A.** (2009). *Scale-free networks: a decade and beyond.* Science **325(5939)**, 412–413.

[317] **Albert, Jeong, Barabasi, A.** (2000). *Error and attack tolerance of complex networks.* Nature **406(6794)**, 378–382.

[318] **Braun, P., Gingras, A.** (2012). *History of protein-protein interactions: from egg-white to complex networks.* Proteomics **12(10)**, 1478–1498.

[319] **Calderwood, M.A., Venkatesan, K., Xing, L., Chase, M.R., Vazquez, A., Holthaus, A.M., Ewence, A.E., Li, N., Hirozane-Kishikawa, T., Hill, D.E., Vidal, M., Kieff, E., Johannsen, E.** (2007). *Epstein-barr virus and virus human protein interaction maps.* Proc Natl Acad Sci U S A **104(18)**, 7606–11.

[320] **de Chassey, B., Navratil, V., Tafforeau, L., Hiet, M.S., Aublin-Gex, A., Agaugue, S., Meiffren, G., Pradezynski, F., Faria, B.F., Chantier, T., Le Breton, M., Pellet, J., Davoust, N., Mangeot, P.E., Chaboud, A., Penin, F., Jacob, Y., Vidalain, P.O., Vidal, M., Andre, P., Rabourdin-Combe, C., Lotteau, V.** (2008). *Hepatitis C virus infection protein network.* Mol Syst Biol **4(1)**, 230.

[321] **Lamkanfi, M., Dixit, V.M.** (2010). *Manipulation of host cell death pathways during microbial infections.* Cell Host Microbe **8(1)**, 44–54.

[322] **Lee, P., Lee, D.J., Chan, C., Chen, S.W., Ch'en, I., Jamora, C.** (2009). *Dynamic expression of epidermal caspase 8 simulates a wound healing response.* Nature **458(7237)**, 519–23.

[323] **Beaudenon, S., Huibregtse, J.M.** (2008). *HPV E6, E6AP and cervical cancer.* BMC Biochem **9(1)**, 1–4.

[324] **Perez-Victoria, F.J., Mardones, G.A., Bonifacino, J.S.** (2008). *Requirement of the human garp complex for mannose 6-phosphate-receptor-dependent sorting of cathepsin d to lysosomes.* Mol Biol Cell **19(6)**, 2350–62.

[325] **Pereira, R., Hitzeroth, I., Rybicki, E.P.** (2009). *Insights into the role and function of L2, the minor capsid protein of papillomaviruses.* Arch Virol **154(2)**, 187–97.

[326] **Lipovsky, A., Popa, A., Pimienta, G., Wyler, M., Bhan, A., Kuruvilla, L., Guie, M.A., Poffenberger, A.C., Nelson, C.D.S., Atwood, W.J., DiMaio, D.** (2013). *Genome-wide siRNA screen identifies the retromer as a cellular entry factor for human papillomavirus.* Proceedings of the National Academy of Sciences **110(18)**, 7452–7457.

[327] **Day, P.M., Thompson, C.D., Schowalter, R.M., Lowy, D.R., Schiller, J.T.** (2013). *Identification of a role for the trans-golgi network in human papillomavirus 16 pseudovirus infection.* Journal of Virology **87(7)**, 3862–3870.

[328] **Buck, C.B., Pastrana, D.V., Lowy, D.R., Schiller, J.T.** (2004). *Efficient intracellular assembly of papillomaviral vectors.* J Virol **78(2)**, 751–7.

[329] **Zhao, K.N., Hengst, K., Liu, W.J., Liu, Y.H., Liu, X.S., McMillan, N.A., Frazer, I.H.** (2000). *BPV1 E2 protein enhances packaging of full-length plasmid DNA in BPV1 pseudovirions.* Virology **272(2)**, 382–93.

[330] **Asumalahti, K., Laitinen, T., Itkonen-Vatjus, R., Lokki, M., Suomela, S., Snellman, E., Saarialho-Kere, U., Kere, J.** (2000). *A candidate gene for psoriasis near HLA-C, HCR (Pg8), is highly polymorphic with a disease-associated susceptibility allele.* Hum Mol Genet **9(10)**, 1533–1542.

[331] **Tiala, I., Wakkinen, J., Suomela, S., Puolakkainen, P., Tammi, R., Forsberg, S., Rollman, O., Kainu, K., Rozell, B., Kere, J., Saarialho-Kere, U., Elomaa, O.** (2008). *The PSORS1 locus gene CCHCR1 affects keratinocyte proliferation in transgenic mice.* Hum Mol Genet **17(7)**, 1043–1051.

[332] **Abbate, E., Berger, J., Botchan, M.** (2004). *The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2.* Genes & Development **18(16)**, 1981–1996.

[333] **Wang, Y., Coulombe, R., Cameron, D.R., Thauvette, L., Massariol, M.J., Amon, L.M., Fink, D., Titolo, S., Welchner, E., Yoakim, C., Archambault, J., White, P.W.** (2004). *Crystal structure of the E2 transactivation domain of human papillomavirus type 11 bound to a protein interaction inhibitor.* J Biol Chem **279(8)**, 6976–85.

[334] **Schweiger, M.R., You, J., Howley, P.M.** (2006). *Bromodomain protein 4 mediates the papillomavirus E2 transcriptional activation function.* J Virol **80(9)**, 4276–85.

[335] **Bikle, D., Oda, Y., Xie, Z.** (2004). *Calcium and 1,25(OH)2D: interacting drivers of epidermal differentiation.* The Journal of Steroid Biochemistry and Molecular Biology **89-90(1)**, 355 – 360.

[336] **Boukamp, P., Petrussevska, R.T., Breitkreutz, D., Hornung, J., Markham, A., Fusenig, N.E.** (1988). *Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line.* The Journal of Cell Biology **106(3)**, 761–771.

[337] **Micallef, L., Belaubre, F., Pinon, A., Jayat-Vignoles, C., Delage, C., Charveron, M., Simon, A.** (2009). *Effects of extracellular calcium on the growth-differentiation switch in immortalized keratinocyte HaCat cells compared with normal human keratinocytes.* Experimental Dermatology **18(2)**, 143–151.

[338] **Muller, M., Jacob, Y., Jones, L., Weiss, A., Brino, L., Chantier, T., Lotteau, V., Favre, M., Demeret, C.** (2012). *Large scale genotype comparison of human papillomavirus E2-host interaction networks provides new insights for E2 molecular functions.* PLoS Pathog **8(6)**, e1002761.

[339] **Benson, J.D., Lawande, R., Howley, P.M.** (1997). *Conserved interaction of the papillomavirus E2 transcriptional activator proteins with human and yeast TFIIB proteins.* Journal of Virology **71(10)**, 8041–7.

[340] **Santin, A.D., Zhan, F., Bignotti, E., Siegel, E.R., Cané, S., Bellone, S., Palmieri, M., Anfossi, S., Thomas, M., Burnett, A., Kay, H.H., Roman, J.J., O'Brien, T.J., Tian, E., Cannon, M.J., Jr., J.S., Pecorelli, S.** (2005). *Gene expression profiles of primary HPV16- and HPV18-infected early stage cervical cancers and normal cervical epithelium: identification of novel candidate molecular markers for cervical cancer diagnosis and therapy.* Virology **331(2)**, 269–291.

[341] **Pellet, J., Tafforeau, L., Lucas-Hourani, M., Navratil, V., Meyniel, L., Achaz, G., Guironnet-Paquet, A., Aublin-Gex, A., Caignard, G., Cassonnet, P., Chaboud, A., Chantier, T., Deloire, A., Demeret, C., Le Breton, M., Neveu, G., Jacotot, L., Vaglio, P., Delmotte, S., Gautier, C., Combet, C., Deleage, G., Favre, M., Tangy, F., Jacob, Y., Andre, P., Lotteau, V., Rabourdin-Combe, C., Vidalain, P.O.** (2010). *ViralORFeome: an integrated database to generate a versatile collection of viral ORFs.* Nucleic Acids Res **38(Database issue)**, D371–8.

[342] **Huang da, W., Sherman, B.T., Zheng, X., Yang, J., Imamichi, T., Stephens, R., Lempicki, R.A.** (2009). *Extracting biological meaning from large gene lists with DAVID.* Curr Protoc Bioinformatics **Chapter 13**, Unit 13 11.

Presentations

This work was presented in national and international meetings:

- Oral presentation at the 28th HPV conference, San Juan, Porto Rico, **USA, 2012**: *Characterization of a specific interaction between the HPV16 E2 protein and CCHCR1.* Muller M & Demeret C.

- Poster at the doctoral school days, Paris, **France, 2012**: *E2-host protein-protein interaction network from 12 HPV.* Muller M, Jacob Y, Favre M, Demeret C.

- Poster at the "virology days of the Pasteur Institute", Touquet, **France, 2011**: *Characterization of a HPV16 E2 specific biomarker: CCHCR1.* Muller M & Demeret C.

- Oral presentation at the HPV Association 4th meeting, Amboise, **France, 2011**: *Analyse fonctionnelle d'un marqueur spécifique de HPV16 E2.* Muller M & Demeret C.

- Oral presentation at the 27th HPV conference, Berlin, **Germany, 2011**: *E2 virus-host protein-protein interaction network from 12 HPV.* Muller M, Jacob Y, Favre M, Demeret C.

- Oral presentation at the DNA tumor virus meeting, Trieste, **Italy, 2011**: *E2-host protein-protein interaction network from 12 HPV.* Muller M, Jacob Y, Favre M, Demeret C.

- Poster at the "Young researcher in life science" congress, Paris, **France, 2011**: *Comparative analysis of the E2 interaction network discriminates cutaneous and mucosal HPV.* Muller M, Jacob Y, Favre M, Demeret C.

- Poster at the "virology days of the Pasteur Institute", St Malo, **France, 2010**: *Comparative analysis of the E2 interaction network discriminates cutaneous and mucosal HPV.* Muller M, Jacob Y, Favre M, Demeret C.

- Oral presentation at the HPV Association 3rd meeting, Sedan, **France, 2010**: *Analyse comparative des réseaux d'interactions cellulaires des protéines E2 de 12 HPV.* Muller M, Jacob Y, Favre M, Demeret C. Muller M, Jacob Y, Favre M, Demeret C.

Reference articles

This thesis summarizes the results published in the following articles:

• Cassonnet P, Rolloy C, Neveu G, Vidalain PO, Chantier T, Pellet J, Jones L, Muller M, Demeret C, Gaud G, Vuillier F, Lotteau V, Tangy F, Favre M, Jacob Y. Benchmarking of a luciferase-based assay for detecting protein-protein interactions. **Nature Methods 8 (990-992), 2011**.

• Muller M, Jacob Y, Jones L, Weiss A, Brino L, Chantier T , Lotteau V, Favre M, Demeret C. Large Scale Genotype Comparison of Human Papillomavirus E2-Host Interaction Networks Provides New Insights for E2 Molecular Functions. **PLoS Pathogens 8(6): e1002761, 2012**.

• Neveu G, Cassonnet P, Vidalain PO, Rolloy C, Mendoza J, Jones L, Tangy F, Muller M, Demeret C, Tafforeau L, Lotteau V, Rabourdin-Combe C, Trave G, Dricot A, Hill D, Vidal M, Favre M, Jacob Y. Comparative analysis of virus-host interactomes with HT-GPCA, a high-throughput mammalian assay to probe protein interactions. **Methods 58(1), 349-359, 2012**.

• Muller M, Cassonnet P, Jacob Y & Demeret C. A Comparative Approach to Characterize the Landscape of Host-Pathogen Protein-Protein Interactions. **J Vis Exp. In press. 2013**.

• Muller M & Demeret C. CCHCR1 interacts with the Human Papillomavirus type 16 E2 protein and disrupts the differentiation regulation of keratinocytes. **Submitted**.

A review has also been published in the course of this thesis:

• Muller M & Demeret C. The HPV E2-host protein-protein interactions: a complex cellular network hijacking. **Open J. Virol, 6(1), 173-189, 2012**.