Université Sorbonne Paris Cité

Thèse de doctorat

de l'Université Sorbonne Paris Cité

Préparée à l'Université Paris Diderot

Ecole doctorale Bio SPC – Bio Sorbonne Paris Cité – ED 562

Unité de Génétique Moléculaire des Virus à ARN – Institut Pasteur

Delineating the interplay between the PB2 protein of influenza A viruses and the host Ubiquitin Proteasome System.

Présentée par Elise Biquand

Thèse de doctorat de Microbiologie – Option Virologie

Dirigée par Caroline Demeret

Présentée et soutenue publiquement à Paris, le 31 octobre 2017

Président du jury : Pr. Pierre Emmanuel Ceccaldi – Université Paris Diderot – France

Rapporteur : Pr. Martin Schwemmle – University of Freiburg – Allemagne

Rapporteur : Pr. Lionel Tafforeau – Université de Mons – Belgique

Examinateur : Dr. Isabelle Jupin - Université Paris Diderot – France

Examinateur : Dr. Nathalie Arhel - Université de Montpellier – France

Directeur de thèse : Dr. Caroline Demeret - Université Paris Diderot – France

Aknowledgements

I learned recently that this section, and the abstract, might be the only parts of this manuscript you will read. So, I hope that discovering what all these great people did for this work will stimulate your curiosity, and entice you to read more.

I first want to deeply thank the members of my thesis jury. Thanks to Pr. Martin Schwemmle and Pr. Lionel Tafforeau who accepted to review this manuscript, to Dr. Isabelle Jupin and Dr. Nathalie Arhel who agreed to be examiners during my defense, and finally to Pr. Pierre-Emmanuel Ceccaldi who accepted to chair the jury.

Je remercie ensuite l'université Paris Diderot – Paris 7 qui a financé mon doctorat, et m'a donné la chance d'avoir une première expérience d'enseignement en tant que moniteur. Ainsi que le département de virologie de l'Institut Pasteur qui m'a permis de terminer sereinement mon travail grâce à une bourse de soudure.

J'adresse mes remerciements chaleureux au Pr. Sylvie van der Werf qui m'a permis de rejoindre ce laboratoire, pour sa bienveillance et ses encouragements constants durant mon parcours de recherche et d'enseignement.

Mes remerciements les plus sincères vont au Dr. Caroline Demeret pour son encadrement scientifique et pour tout ce qu'elle m'a transmis. Merci de m'avoir soutenue pour toutes mes activités extra-recherche et de m'avoir donnée l'opportunité de partir à Hong Kong. Merci aussi pour ton aide lors de la rédaction de ce manuscrit.

Un merci particulier à Marion Declercq, ces quelques années auraient été beaucoup moins plaisantes, et sûrement beaucoup moins productives sans une amie comme toi. Merci pour tout (le détail prendrait trop de place et risquerait de faire des jaloux ^^) ! Je te souhaite le meilleur pour la suite ;)

Merci à tous ceux avec qui j'ai eu la chance de travailler dans le cadre de mon projet de thèse, et plus particulièrement : Patricia Cassonnet, Yves Jacob, Marwah Karim, Nadia Naffakh, Sandie Munier, Cyril Barbezange et Bernadette Crescenzo. Vous êtes pour beaucoup dans la réussite de mon projet de thèse. Je remercie également Sandie pour son soutien durant mon monitorat, Cyril pour la relecture de ce manuscrit et Bernadette pour m'avoir aidé à garder un esprit (presque) sain dans un corps sain ! Je remercie également les deux stagiaires que j'ai eu la chance d'encadrer durant ce projet : Katarzyna Krzywicka et Léo Bertrand.

I also want to thank Dr Sumana Sanyal for welcoming me in her team at HKU Pasteur Research pole in Hong Kong for one month, and Pr. Roberto Bruzzone for making it possible. I'm grateful to the Institut Pasteur network for the grant "Calmette et Yersin" that helped funding my internship. A huge thanks to Akhee Jahan who taught me all she knows about

CrispR and showed me a little bit around, I really enjoyed our talks! Thanks also for the figures on OTUB1 that you allowed me to use in my manuscript (go to chapter 3 for those curious to see ☺).

Un grand merci à tous les membres passés et présents de mon laboratoire d'accueil : Mélanie Albert, Andréa Alexandru, Florian Andrieux, Mégane Babiak, Amrish Baidjoe, Marion Barbet, Sylvie Behilil, Mathilde Benassaya, Camille Berthaud, Solene Brun, Christine Detchepare, Cédric Diot, Mélanie Dos Santos, Vincent Enouf, Anne-Marie Filodeau, Guillaume Fournier, Lisa Fourniol, Matthieu Fritz, Jean-Michel Heraud, Catherine Isel, Louis Jones, Marie Lazzerini, Sonia Martine Keddani, Stéphane Léandri, Annette Martin, Isabelle Poste, Camille Roquin, Clio Socratous and Jingshu Zhang. J'espère n'avoir oublié personne ! Ce fut un plaisir de travailler à vos côtés. Je vous remercie d'avoir toujours très bien fait semblant de trouver délicieux mes gâteaux plus ou moins réussis ^^. Je vous souhaite le meilleur pour la suite. J'ai aussi une pensée particulière pour Noémie Ammeux qui nous a quittés trop tôt.

Thanks to my PhD colleagues: Marion Declercq, Marwah Karim, Kuang-Yu Chen, Nicolas Malausse and Usama Ashraf. I wish you all the best for the coming years, and later ;)

Je remercie également l'équipe de l'Ecole Supérieure de Biotechnologie de Strasbourg avec laquelle nous avons collaboré pour la construction de la banque UPS et son criblage: Murielle Masson, Yves Nominé, Juline Poirson et Gilles Travé.

I also want to thank my former supervisors Dr. Vincenzo di Bartolo, Dr. Urban Friberg and Dr. Shunichi Takahashi, for their support and goodwill, giving me the opportunity to setup my first experiments in very different fields.

On ne fait pas d'études aussi longues sans être bien entourée, alors un grand merci à tous mes amis que je n'ai pas encore cités et qui m'ont soutenue depuis des années. Mes amis d'il y a longtemps déjà : Marion, Charlotte, Amandine, Anne mais aussi Jennifer chez qui je peux toujours venir me reposer et manger des pralulines, et Anaïs qui m'a fait voyager presque au bout du monde ! Et mes amis d'il y a … presque aussi longtemps finalement ^^ : Alice, Camille, Catherine, Florian, Guillaume, Lucille et Margot et enfin Laura que je remercie pour ses beaux schémas (que vous trouverez en introduction dans les techniques d'interactomique !), mais surtout pour son soutien dans les moments difficiles et pour tous les bons moments que nous avons passés.

Mes derniers remerciements vont à ma famille sans qui je n'en serais pas là. Je remercie mes parents qui ont toujours mis toutes les chances de mon côté pour que je puisse faire ce que j'aime, et pour leur soutien sans faille depuis toutes ces années. Merci à mon père d'avoir relu un nombre incalculable de dossiers tout au long de mes études ! Je remercie aussi ma mère pour ses analyses statistiques, à l'heure où j'écris ses lignes il y a toujours des graphiques de distribution de NLR sur le tableau de la cuisine ^^. Je n'oublie pas ma petite sœur Ariane, qui a toujours voulu être plus grande, et qui sera finalement docteur avant moi, bien joué miss ;) Merci du fond du cœur !

Résumé détaillé - Français

Les virus influenza de type A, B, et C sont les agents étiologiques de la grippe. Parmi les virus de type A, les virus de sous-types $H1N1_{pdm09}$ et H3N2 sont responsables d'épidémies de grippe saisonnières, et causent de 250 000 à 500 000 décès par an dans le monde. D'autres virus influenza A représentent un risque d'émergence virale zoonotique, comme l'a démontré la pandémie de 2009 provoquée par un virus H1N1 d'origine porcine. De plus, on observe un nombre croissant d'infections humaines par des virus aviaires très pathogènes pour l'homme, comme les virus de sous-types H7N9 et H5N1, particulièrement en Chine et en Asie du Sud-Est, ce qui peut laisser craindre l'émergence d'une épidémie liée à ces virus. Les virus influenza A représentent donc sous différents aspects un problème de santé publique à l'échelle mondiale.

Les virus influenza sont des virus enveloppés à ARN négatif comportant huit segments, codant au moins douze protéines virales, dont trois protéines composent la polymérase : PA, PB1 et PB2. Chaque segment d'ARN est encapsidé sur toute sa longueur par des oligomères de protéines NP, et ses extrémités sont associées à un complexe polymérase, formant une ribonucléoparticule virale (RNPv). La polymérase assure la transcription et la réplication des segments d'ARN viraux dans le noyau des cellules infectées, et est à ce titre un élément essentiel du cycle viral. Les différentes activités de la polymérase sont probablement mises en place au cours du cycle viral à travers une variété d'interactions avec des protéines de la cellule hôte. La caractérisation de ces interactions représente un enjeu crucial pour mieux comprendre la biologie des virus influenza et leur adaptation à l'hôte.

Nous avons abordé cette caractérisation en nous focalisant sur le Système Ubiquitine Protéasome (SUP), en raison de son rôle majeur dans l'homéostasie cellulaire. L'ubiquitination des protéines correspond à une des principales modifications post-traductionnelles. Elle repose sur l'activité de trois groupes d'enzymes. Les enzymes E1 activent l'ubiquitine et la transfèrent sur les enzymes de conjugaison E2. Ces dernières vont ensuite se lier à des complexes ligases E3, qui reconnaissent le substrat et vont permettre sa liaison à l'ubiquitine, soit par transfert direct depuis l'enzyme E2 soit via une liaison

intermédiaire de l'ubiquitine sur la ligase E3. L'ubiquitination est un phénomène finement contrôlé par la cellule, et réversible grâce aux déubiquitinases qui clivent les liaisons entre ubiquitines et entre ubiquitine et substrat. Il existe plusieurs types d'ubiquitination en fonction du type de liaison entre les ubiquitines et de la longueur des chaines d'ubiquitine, on parle de "code ubiquitine". En effet, en fonction de son type, l'ubiquitination va contrôler la dégradation des protéines par le protéasome, leur localisation subcellulaire et/ou leur activité. Elle est impliquée dans tous les mécanismes cellulaires et nombre de virus interagissent avec les SUP lors de l'infection, avec un effet pro- ou antiviral.

De plus en plus d'études démontrent l'importance de facteurs du SUP dans l'infection par les virus influenza A. Il a notamment été montré que l'entrée virale dépend de l'E3-ligase ITCH qui permet au virus d'échapper à l'endosome tardif et de libérer les RNPv dans le cytoplasme. A contrario la déubiquitinase USP11 empêche la mono-ubiquitination de NP qui est essentielle à la réplication du virus. Il existe d'autres exemples de l'implication des facteurs du SUP dans l'infection par les virus influenza A, mais aucune étude jusqu'à présent ne s'est focalisée spécifiquement sur les interactions entre SUP et protéines virales.

Nous avons étudié les interactions entre le SUP et la polymérase virale au travers de la protéine PB2. La sous-unité PB2 de l'ARN polymérase virale a été choisie pour détecter ces interactions, puisqu'elle constitue l'interface principale entre la polymérase virale et le protéome cellulaire. En effet sa structure cristallographique laisse apparaître un domaine flexible et accessible pour les interactions avec les facteurs de la cellule hôte. Par ailleurs PB2 est connue pour être un élément clef de la pathogénèse virale. En conséquence, nous avons dressé les cartes d'interaction entre le SUP et cinq PB2 provenant de souches de virus influenza A de virulence différentes chez l'homme : deux souches saisonnières circulant actuellement dans la population humaine ($H1N1_{pdm09}$ et H3N2), deux souches hautement pathogènes chez l'homme (H7N9 et $H1N1_{1918}$) et une souche de laboratoire ($H1N1_{WSN}$).

Nous avons construit une banque contenant 570 facteurs du SUP, ce qui représente environ 60% des facteurs SUP humain connus. Puis nous avons mis au point une méthodologie permettant de réaliser un crible comparatif des interactions entre cette banque SUP et les cinq protéines PB2. Le criblage haut-débit d'interactions protéine-protéine a été réalisé en utilisant une méthode développée au laboratoire, la PCA (*Protein Complementation Assay*). Cette méthode est basée sur la complémentation de fragments de l'enzyme Gaussia luciférase,

permettant de reconstituer une activité enzymatique par interaction des protéines auxquelles ils sont fusionnés. Il permet donc de détecter des interactions directes binaires entre couples de protéines co-exprimées en cellules humaines. Cette méthode, utilisée à plusieurs reprises pour la validation d'interactions identifiées par double hybride, a été utilisée pour la première fois dans notre étude comme méthode de détection primaire d'interactions entre protéines PB2 et SUP, ce qui permet un criblage systématique de la banque du SUP. Nous avons optimisé les conditions d'un tel criblage et mis au point les méthodes d'analyse permettant l'identification des partenaires potentiels des protéines PB2. Une seconde étape de validation des interactions sélectionnées a ensuite été réalisée, en re-testant, par une méthode basée sur la PCA, tous les interacteurs SUP potentiels sélectionnés à l'issue du premier criblage permis tout d'abord de valider les interactions PB2/SUP du premier criblage pour identifier de manière robuste les interacteurs du SUP avec les protéines PB2. La forte proportion des interactions validées (52%), démontre la bonne qualité du criblage primaire en PCA. De plus, cette validation nous a permis d'obtenir des profils d'interaction comparatifs entre les différentes protéines PB2 et les facteurs du SUP identifiés.

L'analyse des similarités de profils d'interaction PB2/UPS des souches étudiées a permis de mettre en évidence une corrélation avec le temps de circulation de chaque souche dans la population humaine. Les profils sont séparés en deux groupes : d'un côté les souches d'origine aviaires récemment introduites chez l'homme et n'ayant pas circulé dans la population humaine (H7N9), ou pour une durée très courte (<1an pour $H1N1_{pdm09}$); de l'autre les souches adaptées à l'homme qui dérivent les unes des autres ($H1N1_{1918}$, $H1N1_{WSN}$, H3N2). Ce dernier groupe est par ailleurs organisé selon le temps de circulation dans la population humaine. Ces résultats suggèrent une implication du SUP dans l'adaptation à l'hôte, facteur crucial dans l'émergence d'une épidémie. Des données de comparaison du SUP humain et aviaire, qui ne sont pas encore disponibles à ce jour, permettraient d'étayer cette hypothèse.

Cette première phase de cartographie a permis de sélectionner 42 facteurs du SUP interagissant avec au moins une des protéines PB2 étudiées. Ces dernières interagissent préférentiellement avec des enzymes E3 ligases ainsi que des déubiquitinases. Ces deux familles du SUP permettent un contrôle fin de l'ubiquitination, et leur ciblage par les protéines PB2, retrouvé pour d'autres protéines virales, induit un impact profond sur la cellule infectée.

Nous avons ensuite caractérisé le rôle fonctionnel des partenaires de PB2 dans le cycle viral des virus saisonniers en comparaison avec la souche de laboratoire par des expériences de déplétion transitoire à l'aide d'ARN interférants (siARN). Les souches naturelles saisonnières $H1N1_{pdm09}$ et H3N2 infectant de manière très inefficace les lignées cellulaires humaines de laboratoire, il a été nécessaire de construire des souches saisonnières adaptées à ces lignées cellulaires. Nous avons obtenu et caractérisé des souches saisonnières adaptées, qui donnent la possibilité d'étudier in vitro leur cycle viral sur lignées de cellules humaines. Ces études fonctionnelles ont permis de mettre en évidence la participation de la plupart des facteurs sur SUP identifiés comme interacteurs de PB2 dans l'infection. De plus, nous avons détecté une implication différentielle de certains partenaires SUP de PB2 dans l'infection par les différentes souches virales étudiées, mettant en évidence la diversité fonctionnelle de l'interaction PB2/SUP selon les souches virales.

Nous avons ensuite étudié plus en détail le rôle de déubiquitinases dans le déroulement du cycle viral. Nous avons mis en évidence que l'une d'entre elle, PAN2 est impliquée dans les phases précoces de l'infection. Cette protéine présente un site catalytique de deubiquitination inactif et est connue pour son implication dans les processus de deadenylation des ARN messagers (ARNm). Nous avons donc supposé qu'elle pouvait agir sur la stabilité des ARNm viraux. Cependant nos résultats montrent qu'il n'en est rien. Nous supposons donc que c'est sa propriété de pseudo-déubiquitinase inactive qui est mise en jeu lors de l'infection. Elle pourrait se fixer sur une protéine et empêcher l'action d'une autre déubiquitinase. D'autres expériences, au niveau protéique cette fois, devraient permettre de confirmer cette hypothèse.
Nous avons mis en évidence que OTUD6A stimule l'activité d'une polymérase virale reconstituée hors contexte infectieux, mais n'avons pas encore déterminé comment cela impacte sur le cycle viral d'OTUD6A. Nous avons finalement montré que OTUB1 semble impliquée dans les phases tardives du cycle viral. Nos collaborateurs de l'Institut Pasteur de Hong-Kong ont montré que l'expression d'OTUB1 est induite lors de l'infection par les virus de sous-type H1N1, et est accompagnée par sa translocation dans le noyau. De plus, OTUB1 est nécessaire à l'infection par les virus de sous-type H1N1 mais pas pour les sous-types H3N2. Cette DUB semble jouer un rôle dans la production de cytokines par la cellule en réponse à l'infection. De plus des données préliminaires montrent qu'OTUB1 pourrait être impliquée dans l'assemblage des nouveaux virions.

En conclusion, la nouvelle méthodologie haut débit que nous avons mise en place pour la cartographie des interactions entre PB2 et les facteurs du SUP a permis de révéler une corrélation entre le profil d'interaction et l'adaptation à l'hôte. Les cartes d'interaction obtenues nous ont permis de montrer que les protéines PB2 des virus influenza A interagissent exclusivement avec les ubiquitines ligases et les déubiquitinases, qui permettent de moduler le "code ubiquitine". Nous avons également mis en évidence la participation de nombreux facteurs du SUP dans l'infection par IAV, et l'implication différentielle de certains facteurs du SUP selon les souches, la deubiquitinase OTUB1 en est un exemple. Les déubiquitinases semblent d'ailleurs impliquées dans toutes les phases de l'infection.
La nouvelle méthodologie que nous avons mise en place peut être appliquée à d'autres protéines virales, et devrait permettre d'incrémenter significativement nos connaissances des interactions virus/ SUP.

Mots clefs : Virus influenza de type A, polymérase, protéine PB2, Système Ubiquitine Protéasome, Déubiquitinases, interactomique comparative, complémentation protéique, interactions virus-hôte.

Abstract

An estimated 10%-20% of the world's population is affected each year by seasonal epidemic influenza, causing about 250,000 to 500,000 fatal cases. The pandemic risk reinforces the trait of influenza A virus (IAV) infection as a public health issue. The virus life cycle critically relies on its ability to manipulate the host proteome. Besides, the ubiquitin-proteasome system (UPS) is involved in many regulatory processes in mammalian cells by inducing protein degradation, mediating protein activation or shaping their sub-cellular localisation. Therefore, UPS is a prime target hijacked by viruses. Recent evidence indicates that an intricate regulatory network involving viral proteins and the cellular UPS is likely to contribute to viral replication and immune evasion of influenza A viruses. However, usurpation of the host UPS by IAV is far from being comprehensively deciphered.

To gain better understanding, we assessed the interplay between the human UPS and the PB2 subunit of the influenza A virus polymerase through a global proteomic profiling approach. For that purpose, an UPS-dedicated library of 590 human cDNAs, comprising 63% of the whole human UPS, was constituted and characterised. In an initial screening, UPS factors were challenged using a high-throughput split luciferase assay for interaction with the PB2 protein from 5 influenza A strains of different pathogenicity in human. A total of 80 UPS factors emerged as potential PB2 partners, of which 42 were validated as high-confidence PB2 partners for at least one of the strains. Further comparison of interaction profiles of the 5 PB2 with the UPS by hierarchical clustering revealed an interaction dendrogram fitting with the circulation time in the human population.

Functional importance of interactors was tested by siRNA-mediated knock down experiments using luciferase tagged recombinant IAV viruses. Depletion of 36 out of the 42 tested UPS factors showed an effect on the infection with all or a subset of IAV strains, underlying the strong functional output of the developed methodology. Among these factors three deubiquitinases (DUBs) were further studied to decipher their involvement in IAV viral cycle. We have shown that they are involved in early and late stages of the infection and began to draw their function in viral cycle. We demonstrated with our colleagues in Hong-Kong that OTUB1 is involved in the host cytokine response and most probably in virus assembly. OTUD6A was also shown to be implicated in late stages of the infection but we still don't know its exact role. Contrariwise, the inactive DUB PAN2, which is part of poly-adenylation complexes, is implicated in the early phase of IAV infection, but surprisingly apparently not through viral mRNA regulation. More work is on going to specify the mechanisms of these DUBs implication in IAV infection.

Keywords: Influenza A viruses, polymerase, PB2 protein, Ubiquitin Proteasome System, deubiquitinases, comparative interactomics, protein complementation assay, host-pathogen interactions.

Table of content

Table of Figures

Table of Tables

Abbreviations

AP-MS	**A**ffinity **P**urification coupled with **M**ass **S**pectrometry
CRL	**Cullin**-based **R**ing ubiquitin **L**igases
Cter	**C ter**minal
Gluc1	N-terminal part of the ***G****aussia princeps* **luc**iferase
Gluc2	C-terminal part of the ***G****aussia princeps* **luc**iferase
HT	**H**igh **T**hroughput
IAV	**I**nfluenza **A** **V**irus
DUB	**Deub**iquitinase
MOI	**M**ultiplicity **O**f **I**nfection
NLR	**N**ormalized **L**uminescence **R**atio
NP	**Nu**cleo**p**rotein
Nter	**N ter**minal
PA	**P**olymerase **A**cidic
PB1	**P**olymerase **B**asic **1**
PB2	**P**olymerase **B**asic **2**
PCA	**P**rotein **C**omplementation **A**ssay
pfu	**p**laque **f**orming **u**nit
PPI	**P**rotein-**P**rotein **I**nteraction
PRS	**P**ositive **R**eference **S**et
PT	**P**ositive **T**hreshold
RdRP	**R**NA **d**ependent **R**NA **P**olymerase
vRNP	**v**iral **R**ibo**N**ucleo**P**article
RRS	**R**andom **R**eference **S**et
siRNA	**s**mall **i**nterfering **RNA**
SRF	**S**ubstrate **R**ecognition **F**actor
Ub	**Ub**iquitin
UPS	**U**biquitin **P**roteasome **S**ystem
vRNA	**v**iral **RNA**
Y2H	**Y**east **Two**-**H**ybrid

Bibliographic review and problematic

UPS pathway and viruses

1. The ubiquitin Proteasome System (UPS)

The ubiquitin proteasome system (UPS) is a major component of cellular homeostasis that is well conserved in eukaryotic organisms. Ubiquitination is a main post-translational modification, which has been known for a long time to control protein degradation [1]. However, ubiquitination is also involved in other protein properties, such as sub-cellular localisation [2, 3], activation, or interactions [4]. Therefore, the UPS is determinant for the regulation of numerous cell functions, such as DNA repair [4, 5], protein activation [6], cell proliferation [7, 8] and cell death or immune response [9, 10].

Ubiquitin (Ub) is encoded by four genes (UBC, UBB, UBA52 and UBA80) and is transcribed and translated as a linear fusion consisting of multiple copies of ubiquitin, or of ubiquitin fused to the amino terminus of two ribosomal proteins, 60S ribosomal protein L40 and 40S ribosomal protein S27a. The ubiquitin is a 76-amino acid protein with seven Lys residues (Figure 1). Ubiquitin is covalently attached to target proteins via an isopeptide bond between its C-terminal glycine and a lysine residue of the acceptor substrate. Each ubiquitin can be subjected to further ubiquitin addition on one of the seven lysine residues, forming isopeptide-linked ubiquitin chains attached to the substrate protein. Another chain type, Met1-linked is generated when ubiquitin is linked to the N-terminus of another ubiquitin. A multitude of distinct complex chains can be generated, which act as signals with distinct cellular outcomes. Ubiquitination can act as a code, called the ‘ubiquitin code’, to store and transmit information [11] (See The ubiquitin code p16). Ubiquitin can also be linked to proteins of the ubiquitin-like family such as SUMO (small ubiquitin-like modifier), NEDD8 (neural precursor cell expressed, developmentally down-regulated 8), and ISG15 (interferon-stimulated gene 15) (Figure 3).

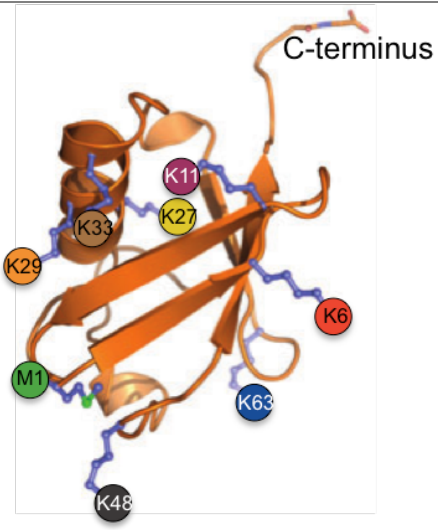


Figure 1: The ubiquitin protein**.** Structure of ubiquitin showing the seven lysine residues (K) and methionine 1 (N-terminal residue, M1). Spheres indicate amino acid groups used in ubiquitin chain formation. Adapted from [11].

Ubiquitination happens in three steps characterised by the serial transfer of the ubiquitin molecule to different enzymes (Figure 2). In a first step, the ubiquitin is activated by an E1 activating enzyme, which forms a thiol ester with the C-terminal Gly of the ubiquitin, thereby activating it for nucleophilic attack. The ubiquitin is then transferred to an E2 conjugating enzyme by trans-esterification before it is finally attached, usually through an isopeptide bond, to the substrate protein thanks to an E3 ligase. E3 ligases recognise protein substrates either through dedicated subunits named “substrate recognition factors” (SRF) for the Cullin-based E3 complexes, or by specific motifs for the HECT-E3 non-Cullin Ring-E3 ligases. In addition, deubiquitinases allow the reversibility of ubiquitination and thus contribute to the fine tuning of this central cellular pathway (Figure 2).

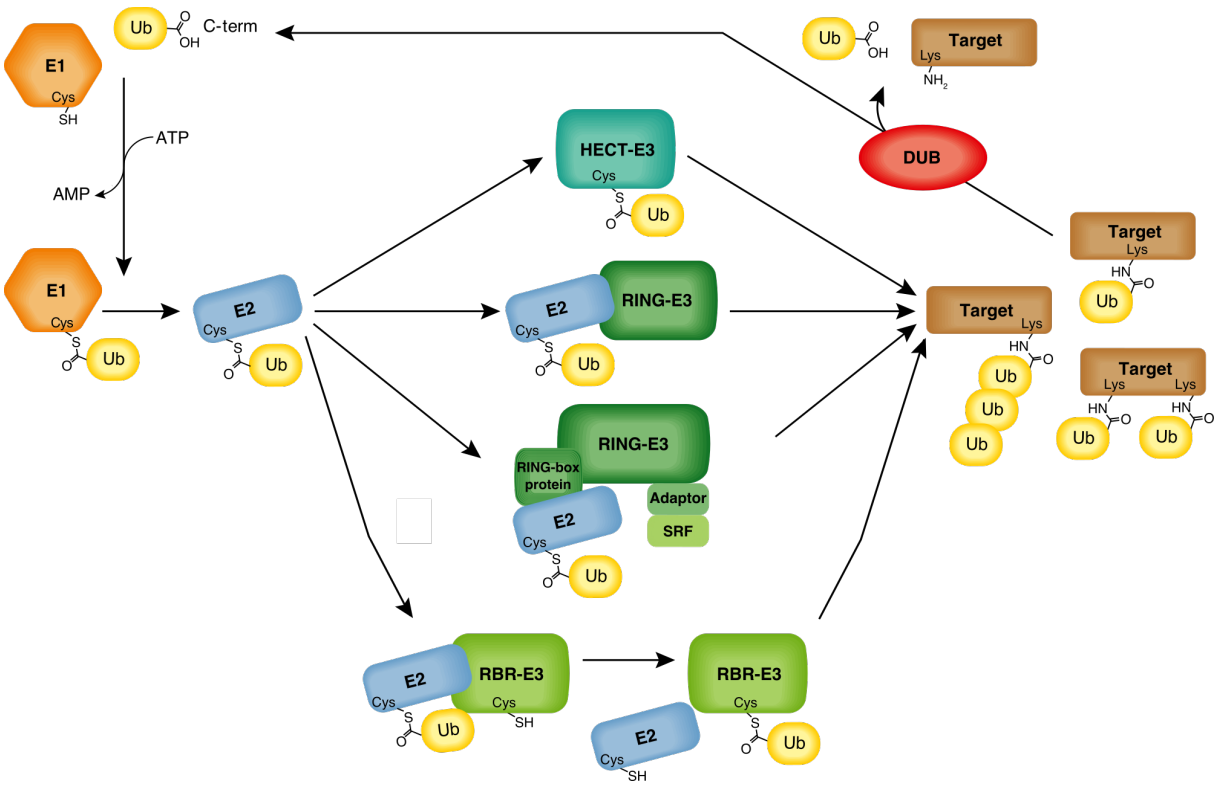


Figure 2: Ubiquitination and deubiquitination pathways. Substrate ubiquitination happens in three stages. 1- Activation of the ubiquitin in an ATP-dependent manner by an E1 activating enzyme. 2- Transfer of ubiquitin to an E2 conjugating enzyme to form a thioester intermediate. 3- The final transfer of ubiquitin onto the protein target, mediated by E3-ligases that either form a thioester intermediate with the ubiquitin (HECT E3-ligases), or mediate a direct transfer of the ubiquitin from the E2 onto the protein (monomeric RING E3-ligases and Cullin-based complexes Cul-E3), or function as RING/HECT-type hybrids (RBR E3-ligases). Deubiquitinases (DUBs) remove the ubiquitin from the protein substrates and allow the reversibility of ubiquitination. Adapted from [12].

2. The ubiquitin code

The diversity of the ubiquitin chains that can be formed on substrate proteins is referred as the ubiquitin code. Rather than being a simple code that assigns one specific output to each signal, ubiquitination has emerged as a cellular language in which information transfer depends both on the nature of the signal and on the context in which it is sent [13]. The ubiquitin code is based on two type of ubiquitin chains: homotypic and heterotypic chains that can integrate ubiquitin-like proteins or be modified by post-translational modifications (Figure 3). The simplest signal is given by a mono-ubiquitin. In homotypic chains, all ubiquitins of the chains are connected through the same lysine or methionine residue, leading to linear structures. Whereas heterotypic chains include different types of linkage and thus can develop branched structures. Moreover, the complexity of ubiquitin chains increases with the

addition of ubiquitin-like proteins such as SUMO, ISG15 or NEDD8, and other post-translational modifications, including acetylation and phosphorylation (Figure 3). The ubiquitin code has been studied a lot in the recent years, but some questions remain regarding the topology of the ubiquitin chains or their receptor for example, that new methodologies will most probably help to answer.

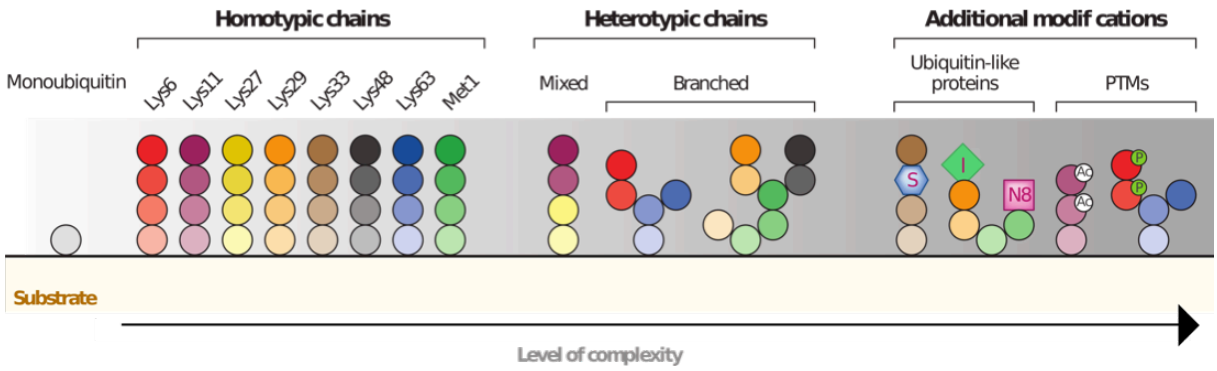


Figure 3 : The ubiquitin code. Schematic representation of the ubiquitin code. Mono-ubiquitin, can be extended on lysine residues (Lys) or on methionine1 (Met1) residue, forming homotypic polyubiquitin chains. More complex heterotypic branched polymers are formed when extension features more than one linkage type. The ubiquitin code can become even more complex with the addition of other signaling proteins, such as SUMO (S), ISG15 (I) or NEDD8 (N8), as well as with other post-translational modifications, including acetylation (Ac) and phosphorylation (P). Adapted from [14].

Substrate ubiquitination can target proteins to different fates, such as proteasomal degradation, interaction with other proteins, or sub-cellular localisation. The type of ubiquitination determines the effect on the substrate. It is well established that Lys48 chains can target proteins to degradation by the proteasome, and constitute the ubiquitin-canonical signals for proteasome-mediated degradation (Table 1). The 26S proteasome is a complex of 33 different subunits, which are arranged into an elongated structure composed of a central 20S core particle with one or two terminal 19S regulatory particle. The combination of multiple active sites with different selectivity contributes to rapid and processive degradation of substrates that have entered the catalytic chamber [15]. Lys48 linkages are the most abundant linkage in all organisms subjected to quantitative proteomic analysis, and their levels increase rapidly when the proteasome is inhibited [16]. However, other linkages, termed “non-canonical” signals, can also be recognised by the proteasome, such as Lys11 linkages that are catalysed by the APC/C complex during mitosis [17] or less frequently Lys29 linkages [18] and Lys63 homotypic and heterotypic chains [19]. Moreover, the Lys63 linkages are most commonly targeting plasma membrane proteins to degradation in

lysosomes [20]. Lys6 and Lys63 have been shown to play a role in mitophagy, a process leading to the removal of damaged mitochondria through autophagy [21, 22] (Table 1).

Ubiquitination is also able to regulate non-degradative processes. Mono-ubiquitination is especially involved in the targeting of proteins to membranar compartments like endosomes, through interactions with proteins at the surface of the cellular compartments [23]. Protein interactions can also be regulated by Lys63-linked chains, in particular during DNA damage response [24] (Table 1). Heterotypic chains of Lys11 and Lys63 but also homotypic Met1 chains are known to regulate the activation of innate immunity pathways through the induction of protein-protein interaction upon the ubiquitination of signaling intermediates, such as NF-κB [25, 26] or Interferon [27] (Table 1). Lys27 linkages have been linked in the same way to innate immunity but also to DNA damage response [28, 29]. Moreover, the Lys29, Lys48 and Lys63 linkages have been shown to be implicated in Wnt signaling pathway [30, 31]. Besides regulating protein activation, ubiquitination also plays a role in their localisation. Lys33 chains have been implicated in post-Golgi protein trafficking, and Lys63 in endocytic and secretory pathways in which they serve as sorting signals for membrane proteins [32] (Table 1). Finally, heterotypic chains and post-translational modifications of ubiquitin, such as phosphorylation, further extend the regulation possibilities. It has been shown for example that ubiquitin phosphorylation is involved in mitophagy pathways [33]. The complexity of the ubiquitin code and the plethora of outcomes on the substrate proteins make protein ubiquitination a central player of cell homeostasis.

Table 1: Overview of the cellular process in which the different ubiquitin chains are involved.

Lys6	Mitophagy, DNA damage response
Lys11	Mitophagy, Proteosomal degradation, Cell cycle regulation
Lys27	DNA damage response, Protein secretion
Lys29	Proteosomal degradation, Epigenetic regulation, Wnt signaling, Innate immunity
Lys33	Intracellular trafficking
Lys48	Proteosomal degradation, Cell cycle regulation
Lys63	Mitophagy, Innate immunity, Proteosomal degradation
Met1	Innate immunity

3. Ubiquitination components

Ubiquitination is driven by the activities of three types of enzymes: E1 activating enzyme, E2 conjugating enzymes and E3 ligases (Figure 2). To date around 10 E1s, 44 E2s and more than 1300 members of E3 families have been identified in human UPS [34, 35] (Digital supplementary data 1).

- *E1 activating enzymes*

The E1 activating enzymes are central to ubiquitin conjugation since they initiate the cascade of conjugation by activating their ubiquitin. E1 enzymes "activate" the ubiquitin molecule by catalysing the reaction of ATP and Ub, involving the active cysteine of the E1 enzyme, providing Ub C-terminus with a high-energy thiolester bond.

- *E2 conjugating enzymes*

This activated ubiquitin is then transferred to the E2 conjugating enzyme through a trans-esterification reaction, a process know as ubiquitin conjugation. The ubiquitin-conjugated E2 then associates with E3 ubiquitin ligase, itself complexed with a substrate protein, and facilitates the formation of an isopeptide bond between Ub and a lysine (K). Competitive binding experiments have shown that E2 must disengage from the E1 to be able to bind the E3 ligases, their binding being mutually exclusive [36] (Figure 4). The structure of E2 enzymes is quite conserved, with a core catalytic domain of around 150 amino acids called UBC domain (Figure 4). This domain adopts a typical α/β-fold, with four α-helices and a four-stranded β-sheet. The E3 binding site is composed of loops and many E2s have short N- and/or C-terminal extensions that can impart important functionality. Indeed some E2 can only mediate protein mono-ubiquitination, whereas others synthesise poly-ubiquitin chains of a specific linkage, or can both link ubiquitin to protein substrate and assemble chains of specified linkages [37]. This specificity depends on the ability of the E2 to bind the incoming ubiquitin in a defined orientation, which will determine chain linkage type and length. In fact, the combinatorial effect of the different E2s and E3s is responsible for the generation of different ubiquitinated protein structures. Ubiquitin chain synthesis is not well understood, but it most probably relies on the interaction of the chain-building machinery with the existing

ubiquitin chain [37]. Furthermore E2 enzymes are known to be regulated by E3 ligases binding, and recent structural evidences identified putative ubiquitin binding sites on the E2s that may enhance processivity and assembly of chains of a defined linkage type [38]. Spotlights are currently turning to these small proteins that are no more considered as simple links in the ubiquitination pathway, but as noteworthy players [37].

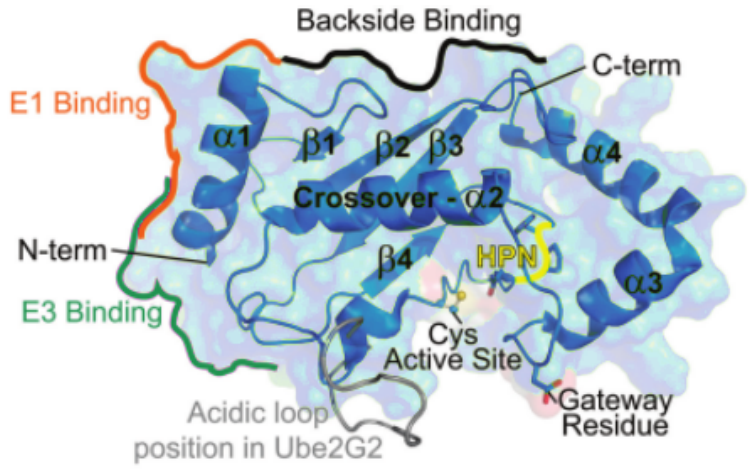


Figure 4: E2 structure. A representative UBC domain is pictured in green (Ube2D3; PDBID: 2FUH) with important structural features and common binding surfaces to the E1 ubiquitin activating enzyme and to E3 ubiquitin ligases. Since the acidic loop was not resolved in this structure, its position is represented (in grey) according to its position in Ube2G2 (PDBID: 2CYX) when the UBC domain is aligned with Ube2D3. [37]

- *E3 ligases, in charge of substrate specificity*

The most abundant and diverse proteins of the UPS are the E3 ligases, which is not surprising considering that they are driving the substrate specificity for ubiquitination. E3 ligases can be grouped in three families according to the mechanism of ubiquitin transfer and to the associated domains.

E3-RING ligases represent the largest family with around 400 members (Digital supplementary data 1). They have either a RING domain (Really Interesting New Gene) containing zinc, or a Ubox domain with the same finger fold without zinc. These domains bind to the ubiquitin thioester-linked E2 and position the substrate lysine, which is nucleophile, in proximity to the reactive E2-ubiquitin thioester bond, facilitating transfer of the ubiquitin to the substrate (Figure 5). The E3-RING can work as monomers, homo- or heterodimers, or be constituted as complexes. This is the case of the Cullin-RING E3 complexes (CLRs) assembled on a Cullin scaffold, which binds a RING-box protein at their N terminus and an adaptor protein or a substrate receptor at their C terminus. Another

important multi subunit E3 is the anaphase-promoting complex/cyclosome (APC/C), a large assembly of 19 subunits that includes a RING subunit (Apc11) and a Cullin-like subunit (Apc2) [39].

E3-HECT ligases display an HECT (homologous to the E6AP carboxyl terminus) domain that catalyses ubiquitin transfer to the substrate via an intermediate linkage of the ubiquitin to a catalytic cysteine on the ligase. The conserved HECT domain is located at the C-terminus of the proteins and is folded in two parts: a N-terminal lobe that interacts with the E2 and is separated by a flexible hinge from a C-terminal lobe, which contains the catalytic cysteine. The N-terminus of the ligase, which is less conserved recognises the protein substrate (Figure 5) [40].

The last family is an intermediate between the E3-HECT and E3-RING. E3-RBR (RING-BetweenRING-RING) ligases contain 2 RING domains (RING1 and RING2) separated by an in-between-RING domain (B) that folds like the RING2 domain while lacking the catalytic cysteine residue (Figure 5). Indeed, RING1 interacts with the E2, which ubiquitin is transferred to a catalytic cysteine on the RING2 and then to the substrate. This mechanism is similar to the mechanism of the E3-HECT ligases [12].

As every protein, E3 ligase activities are regulated by multiple mechanisms. E3 ligase interactors can be of different types: stimulators, inhibitors, adaptors and substrates. Interactions with E2 ligases and substrates are essential, but interaction with other proteins can also modulate E3 activity and/or their localisation, which will determine the availability of the substrate. For example, the HECT-type E3 ligase Smurf1 interacts with Smad7 in response to TGFß. This interaction induces the export of Smurf1/Smad7 complex to the cytoplasm, where Smurf1 ubiquitinates TGF-ßR and enhance its turnover [41], underlying the importance of interaction and localisation in the activity of E3 ligases. Another type of E3 ligase regulation is protein post-translational modifications. E3 ligases can be neddylated, phosphorylated and/or ubiquitinated. Cullin-based complexes comprising the RING finger containing factor RBX1 have been shown to be neddylated on the Cullin in close proximity to RBX1. This Nedd8 conjugation increases ubiquitin ligase activity by promoting the recruitment of the E2 through direct interaction between Nedd8 and the E2 [42]. These regulation mechanisms allow the specificity and regulation of ubiquitination in the dynamic environment of a living cell.

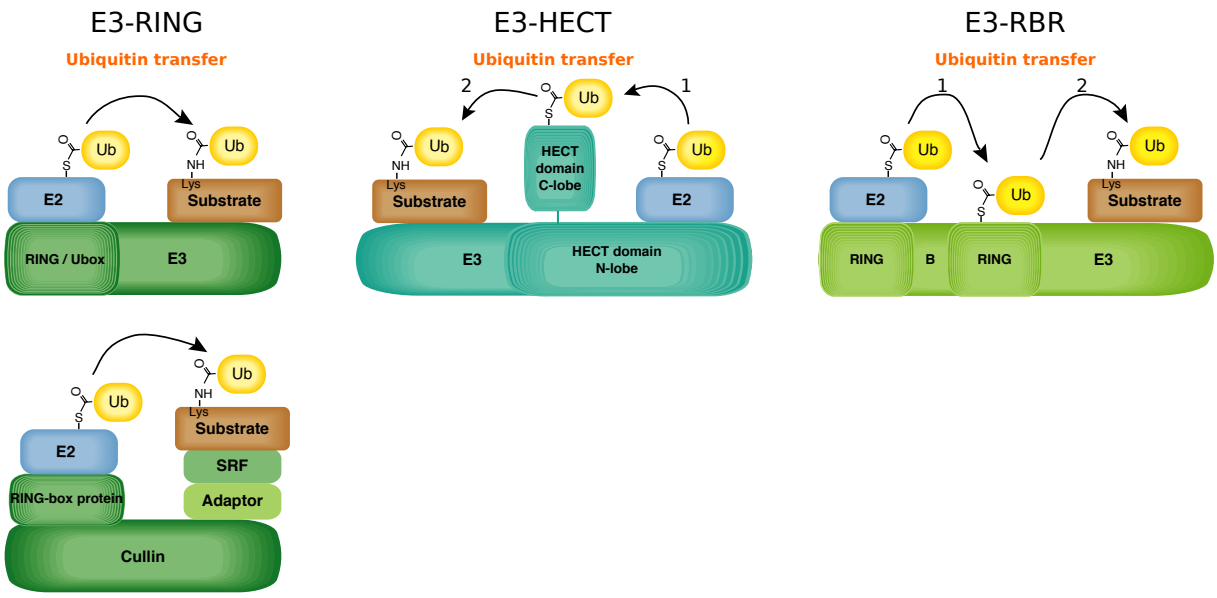


Figure 5: E3 ligases families. E3 ligases are grouped in three families according to the mechanism of transfer of the ubiquitin on the substrate and to the associated domains. E3-RING ligases display RING or Ubox domains and can be either monomeric or dimeric. The Cullin-based complexes include also an adaptor and a Substrate Recognition Factor (SRF). Thanks to the scaffold action of the Cullin, ubiquitin is transferred from the E2 to the substrate. E3-HECT ligases have HECT domains with two lobes that transiently link the ubiquitin on the C-lobe before it is transferred to the substrate. The RBR family is composed of E3 ligases with 2 RING domains flanking an In Between RING domain (B), and has similar mechanism of ubiquitin transfer than the HECT family. Adapted from [43].

- *Deubiquitination, a key element of the UPS*

Around 130 Deubiquitinases (DUBs) are known in human [34, 35] (Digital supplementary data 1), consisting of thiol proteases grouped in 5 families - the Ubiquitin-Specific Proteases (USPs), the Ovarian TUmor proteases (OTUs), the Ubiquitin C-terminal Hydrolases (UCHs), the Josephin family, the Motif Interacting with Ubiquitin (MIU) family [44] - and of zinc metalloproteases of the JAMM family [45].

DUBs have three major functions in the cell (Figure 6). They allow the generation of free ubiquitin from the transcribed fusion of multiple ubiquitin polypeptides, thus fuelling the free ubiquitin pool. DUBs can also remove ubiquitin or ubiquitin chains from proteins, with different possible outputs: either the reversal or the editing of protein ubiquitination [46]. Reversal of protein ubiquitination in turn can lead to protein stabilisation by rescue from degradation. Ubiquitin editing leads to the modification of the length or type of ubiquitin

chains linked to the substrate. To finish, DUBs are key players in ubiquitin homeostasis by recycling ubiquitins: free ubiquitin molecules are reintroduced in the free ubiquitin pool, either directly from the substrate or by cleaving ubiquitin chains that have been removed from proteins [47].

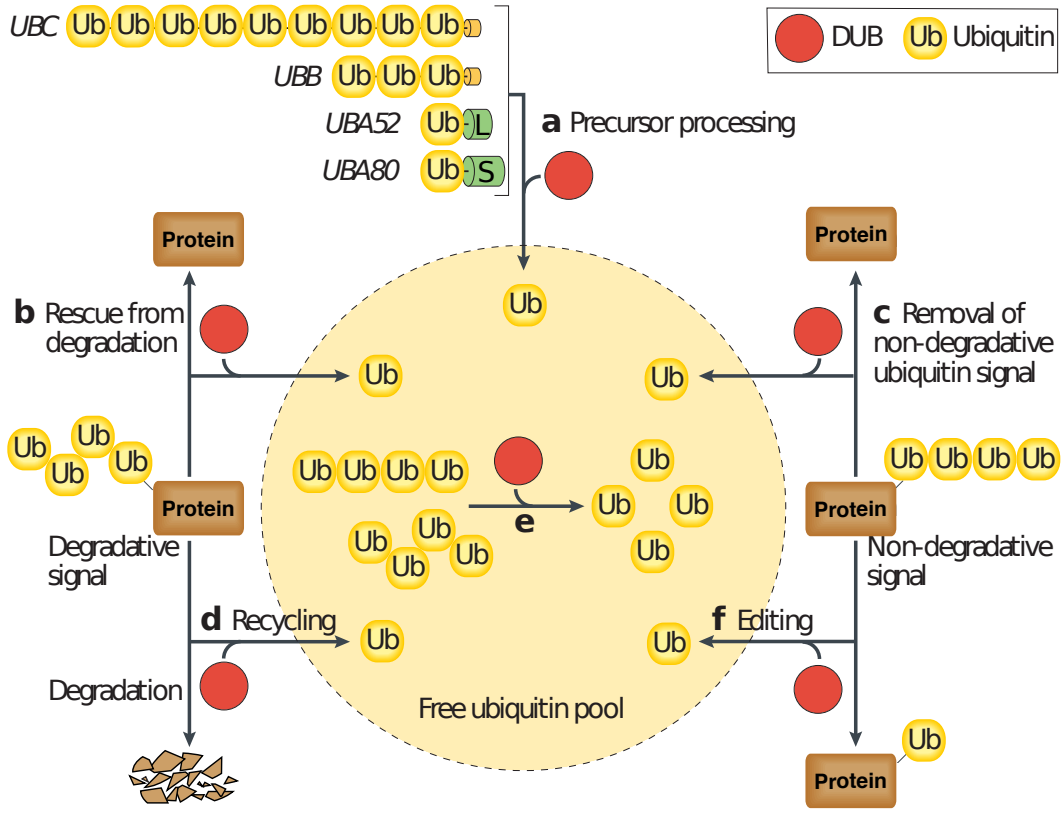


Figure 6: Functions of DUBs. DUBs cleave free ubiquitins from precursors. Ubiquitin is produced in precursors consisting of multiple copies of ubiquitin, or of ubiquitin fused to the amino terminus of two ribosomal proteins, 40S ribosomal protein L40 (L) and 60S ribosomal protein S27a (S) (**a**). DUBs remove ubiquitin from ubiquitinated proteins (b and c). Deubiquitination can rescue proteins from degradation (**b**) or remove a non-degradative ubiquitin signal (**c**). DUBs maintain ubiquitin homeostasis by recycling ubiquitin and preventing degradation of the ubiquitin and of substrates of the 26S proteasome and lysosomal pathways (**d**) and ensure ubiquitin return back to the free ubiquitin pool by disassembling ubiquitin chains removed from substrates (**e**). DUBs can participate to ubiquitin chains editing, corresponding to the exchange of one type of ubiquitin chain for another, by removing ubiquitin from the substrate. Adapted from [47].

Many layers of DUB specificity support all these activities. The first is the linkage specificity (Figure 7a), which is known for some of the DUBs in the OTU [48, 49], USP [50] or JAMM [51] families. Lys48 and Lys63 linkage-specific members have been described in theses families. Each lysine in ubiquitin has a unique sequence context, which might be used for specific recognition by DUBs. Moreover, studies have shown that the DUB structure can

explain its positioning on the ubiquitin chain [52]. Structure can also explain the second layer of specificity that is the cleavage site position: either internal to the ubiquitin chain (endo-deubiquitination) or at the extremity of the chain (exo-deubiquitination) (Figure 7b). Endo-DUBs must accommodate ubiquitin molecules on either side of the cleavage site, whereas exo-DUBs only need to bind to a single ubiquitin [50]. The last cleavage releases free ubiquitin molecules, whereas the endo-deubiquitination releases ubiquitin chains that can be further cleaved by DUBs, producing free ubiquitin (Figure 7e and Figure 6). The proximal ubiquitin, linked to the substrate, is less accessible and might be cleaved by substrate-specific DUBs (Figure 7c-d) that can accommodate a larger range of sequence in their proximal binding site [47]. Furthermore, a mono-ubiquitination can lead to the linkage of a new ubiquitin chain.

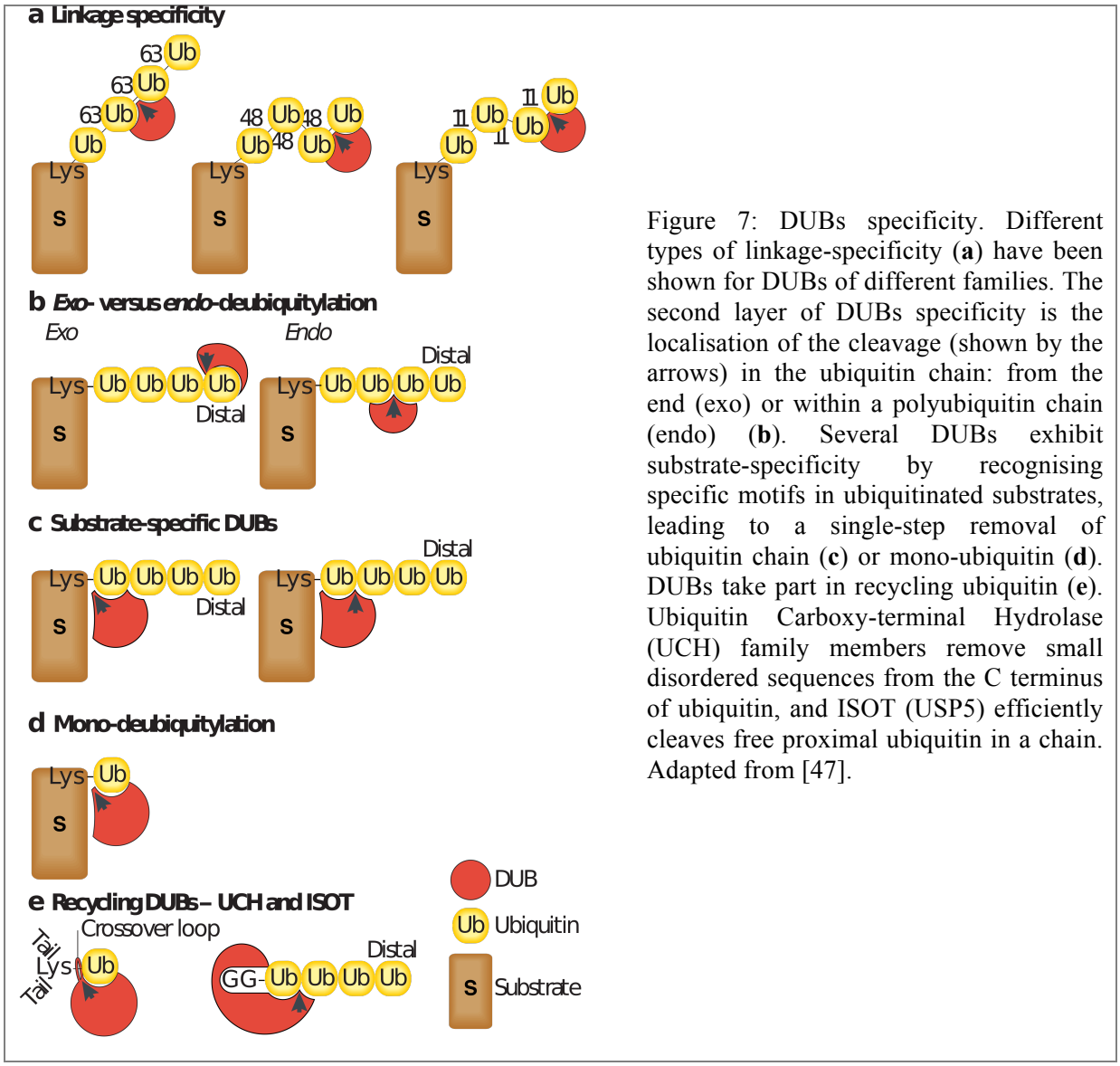


Figure 7: DUBs specificity. Different types of linkage-specificity (**a**) have been shown for DUBs of different families. The second layer of DUBs specificity is the localisation of the cleavage (shown by the arrows) in the ubiquitin chain: from the end (exo) or within a polyubiquitin chain (endo) (**b**). Several DUBs exhibit substrate-specificity by recognising specific motifs in ubiquitinated substrates, leading to a single-step removal of ubiquitin chain (**c**) or mono-ubiquitin (**d**). DUBs take part in recycling ubiquitin (**e**). Ubiquitin Carboxy-terminal Hydrolase (UCH) family members remove small disordered sequences from the C terminus of ubiquitin, and ISOT (USP5) efficiently cleaves free proximal ubiquitin in a chain. Adapted from [47].

DUBs activities can be regulated via three major mechanisms. The first one is post-translational modification. For example, the activity of CYCLD is known to be regulated by phosphorylation [53], and ubiquitination regulates ATXN3 [54]. Secondary, interaction with other proteins, among which other DUBS or E3 ligases, can inhibit or activate the DUBs, or can direct them to specific ubiquitinated substrates [55]. To finish, the localisation of the enzyme is crucial, as it determines the protein substrates exposed to the DUBs. USP8 for example is one of the many DUBS that have been shown to be recruited to EGFR-containing early endosomes after an EGF signal and to be involved in the EGFR recycling [56].

4. UPS and viral infection

Viruses of all families are known to manipulate the UPS, using ubiquitin at most stages of the infection. First, during the viral entry and the nucleocapsid transport, when inhibition of the proteasome in cell can modulate the nucleocapsid transport and/or the disassembly of several viruses. For example, the internal capsid protein VI of adenovirus 5 is ubiquitinated through the recruitment of Nedd4 E3 ligases, which is mandatory after virus release from the endosome for the microtubule-dependent nucleocapsid transport to the nucleus [57]. Conversely, a portion of the intracellular particles is ubiquitinated and degraded during Adeno-Associated Viruses (AAV) type 2 and 5 infections, so that proteasome inhibitors enhance AAV infection [58]. The transcription of the viral genome is also described, in many cases, as an UPS-sensitive stage. The transactivator EBNA1 of Epstein-Barr virus (EBV) has been shown to bind to DUB USP7, inducing the deubiquitination of histone 2A and increasing the binding of EBNA1 to the oriP site in the viral genomic DNA [59]. For many viruses, the last stages of the viral cycle, such as assembly, budding and release, also involve the UPS. Ebola virus matrix protein eVP40, for example, is ubiquitinated through its interaction with the E3-HECT ligase WWP1. This post-translational modification is thought to promote more efficient self-assembly, maturation, and scission of eVP40, enhancing the release of virus-like particles [60].

Besides completion of virus cycle *per se*, UPS is also involved in the cellular response to infection, and interactions between viral proteins and UPS are not always virus-driven. Immune responses are highly regulated by ubiquitination of signaling intermediates [61]. Interferon (IFN) production is the host's first line of defence against viral infection, and

distinct activation steps in NF-κB and IRF3 signaling, that lead to IFN production, involve K63-linked ubiquitination of signaling intermediates [62]. However, there are numerous examples in which viruses employ the UPS to block either the production of IFN, or its anti-viral effects. One way to block IFN response is, for example, to induce the degradation of interferon regulatory factors (IRF), as does the rotavirus NSP1 protein by inducing the degradation of IRF3, IRF5, and IRF7 [63]. NSP1 acts as an E3 ligase, interacting with components of CRL complexes, most notably the Cul3 and Cul1 complexes. Unexpectedly, IRF are not targeted to the proteasome via these interactions, and the mechanism inducing their degradation is still unknown [64].

The cell cycle can also be diverted by viruses through UPS signaling, as for example the E6 and E7 oncoproteins of High-Risk (HR) Human PapillomaViruses (HPV). The E7 protein targets pRB family members for degradation to trigger G1/S progression through the release the cell cycle factors such as E2F1 [65]. The E6 protein of HR-HPVs functions as an adaptor protein, which redirect the HECT ubiquitin ligase E6AP toward p53 for proteasome-dependent degradation [66, 67]. The coordinated UPS-mediated host factor degradation by HR-HPV E6 and E7 is required to maintain infected keratinocytes in a proliferative state, mandatory for viral replication.

Numerous other examples of UPS hijacking by viruses are reviewed in [10, 67, 68], showing how viruses generally subvert the host UPS. But some viruses also express proteins that function on their own as UPS factors, notably as E3 ligases and DUBs. Lists of viral proteins with ubiquitin ligase or deubiquitinase activity can be found in [69]. Not surprisingly, most of the targets of these viral “UPS related factors” are part of the anti-viral cell response. Recently, the HSV-1-encoded Ub-specific protease UL36 was shown to deubiquitinate TRAF3, thereby counteracting IFN-β signaling [70].

Overall, a complex and multi-layered set of interactions exists between viruses and the UPS. Viruses inhibit or induce ubiquitination of protein targets, either by diverting the host UPS proteins or by coding their own factors (Figure 8).

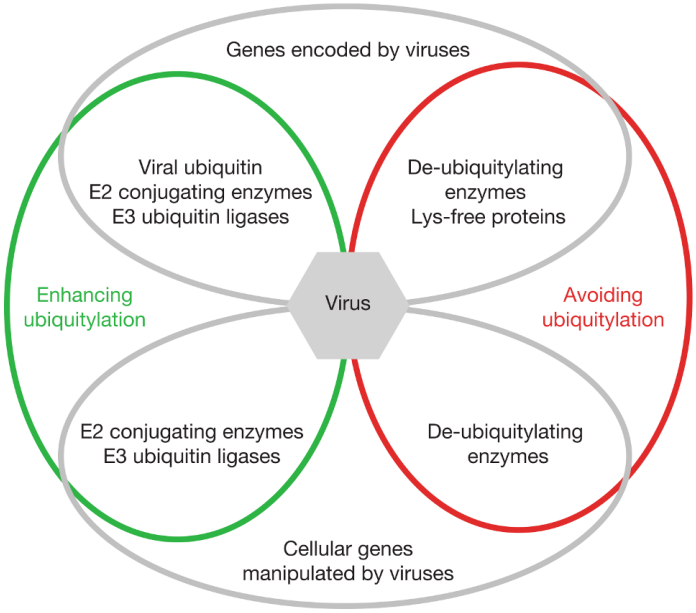


Figure 8: Virus and ubiquitin system manipulation. Viruses can divert the UPS by inducing or inhibiting ubiquitination, targeting either viral or cellular proteins, thanks to the manipulation of cellular factors or to the encoding their own ligases or DUBs. [10]

Influenza A viruses

Influenza are enveloped viruses of the *Orthomyxoviridae* family. Their genome is constituted of 7 or 8 single-stranded RNA segments of negative polarity. Four types of influenza (A, B, C and D) are described according to the antigenicity of the M1 and NP proteins [71]. Influenza B and C viruses are almost restricted to humans, while influenza D viruses have been reported to infect *bovidae* and swine species [72, 73]. Influenza A viruses (IAV) represent the main influenza associated-public health issue. IAVs are further classified into subtypes based on the antigenic properties of their two major surface glycoproteins, the hemagglutinin (HA, H1-H16) and the neuraminidase (NA, N1-N9). Two other IAV subtypes have been discovered in bats: H17N10 and H18N11 [74] (Figure 9). Viruses of the $H1N1_{pdm09}$ and H3N2 subtypes are responsible for current seasonal epidemics in humans. IAV also infect a wide range of birds and mammals. The highest diversity of IAVs and almost all the subtypes (H1-H16; N1-N9) have been isolated from wild aquatic birds of the Anseriformes (ducks, swans and gooses) and Charadriiformes (gulls) orders. They are considered as the natural reservoir of IAVs. In these species the infection is usually asymptomatic. From their natural reservoirs, IAVs can be transmitted to numerous species, including humans [75] (Figure 9), and thus represent a potential risk of zoonotic pandemics [76].

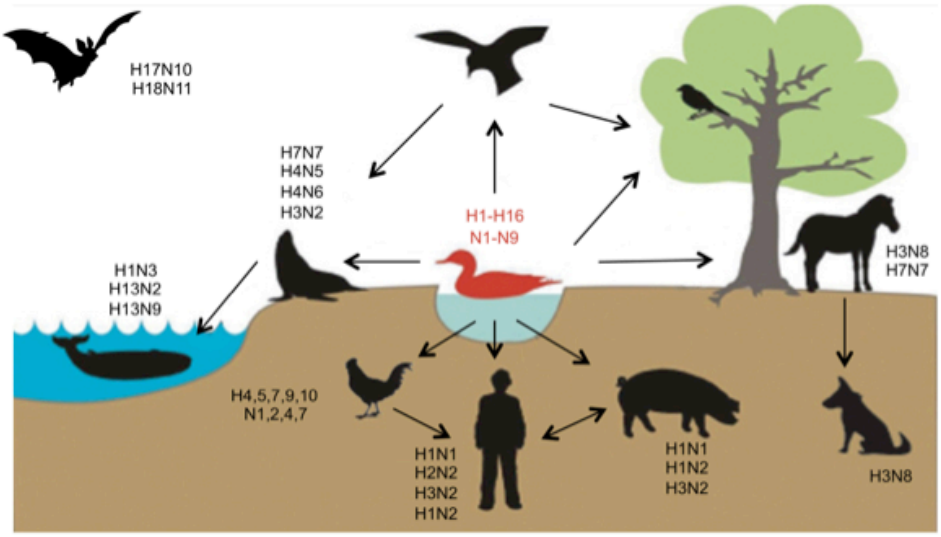


Figure 9: Influenza A reservoir is the wild aquatic birds population (in red), where the largest diversity of sub-types is found (H1-H16 ; N1-N9). From this reservoir IAV are able to infect a wide diversity of terrestrial and aquatic species (in black). The subtypes known to have circulated in given species are indicated with the corresponding image. Avian viruses can be transmitted directly to human or through an intermediate host, as swine or poultry. Adapted from [71].

1. Epidemiology and evolution

Since the 20th century, only H1N1, H2N2, H1N2 and H3N2 IAVs have circulated into the human population (Figure 9 and Figure 10). It is estimated that 5% to 10% of the worldwide population is infected yearly, IAV causing up to 500,000 deaths [77] and significant economical costs [78]. Viruses are transmitted by aerosols, droplets, or through contaminated skin/surfaces contacts. Typically, after a 1-2 days incubation period following infection, the onset of symptoms is abrupt. Influenza symptoms are not specific and include, among others, intense fever, headaches, myalgia, nasal congestion and cough. Rarely, IAV infection leads to severe, potentially fatal outcomes that depend largely on the patients' comorbidity factors and age (populations <2 and >65 years old are more likely to present a poor outcome) [77].

Two main mechanisms drive IAV evolution. The first one, called antigenic drift, is due to the low fidelity of the viral polymerase, with a mutation rate of 10^{-4}, which introduces around 1 mutation per genome per replication cycle [79]. Thus, the viral population in one individual is a complex composite of variants, named quasi-species, which ensure the maintenance of IAVs into the human population [80]. The second phenomenon is the reassortment that can

occur during co-infection, resulting in a genetically mixed virion, called reassortant. Reassortments between viruses of the same subtype or between subtypes play a key role in virus intra and inter-species evolution [81-83]. Occasionally, a new subtype of influenza virus is introduced in the human population, against which humans are immunogically naïve; this phenomenon is called antigenic shift. The novel virus type can rapidly spread, and potentially cause a pandemic.

Influenza outbreaks have apparently occurred since at least the Middle Ages, if not since ancient times [84]. Since the beginning of the 20th century, four distinct pandemics have occurred in the human population. Phylogenetic analyses revealed the zoonotic origin of pandemic IAVs [84]. The 1918, 1957, 1968 and 2009 pandemic viruses are thought to be reassortants harbouring a mix of viral genomic segments from human and animal IAVs (Figure 10), indicating the major importance of genetic reassortment in the emergence of novel IAVs into the human population. Indeed, the $H1N1_{pdm09}$ virus is derived from human, swine and avian viruses [85] (Figure 10). Viruses of the 1957 and 1968 pandemics derived from a reassortment between human viruses circulating at that time and a least one avian virus (Figure 10). The H1N1 pandemic of 1918, also called “Spanish flu”, killed 20 to 50 million of people, mostly because of subsequent bacterial infections. The pandemics of 1957 (H2N2), 1968 (H3N2), and 2009 ($H1N1_{pdm09}$) appeared less deadly [71].

Establishment of a new lineage in the human population requires: 1) transmission from animal to human; 2) an efficient replication in human; 3) a human-to-human transmission. Adaptation to an intermediate domestic species, as swine or poultry, allows the amplification of viruses in close contact to human. Human viruses preferentially bind to sialic acid attached to galactose by an α2,6 linkage, when avian and equine influenza viruses bind to α2,3Gal-linked sialic acids [86]. Swine displaying on all its respiratory tract sialic acids attached to galactose by an α2,6 and α2,3 linkage could constitute a “mixing vessel” favouring the emergence of reassortant with potential thread for human [87]. Pre-existing immunity can restrict the replication of viruses, and the vitality of the reassortant compared to circulating strains is key in its emergence. Thus, the majority of human infections by animal viruses are not followed by an emergence, and the massive virus propagation in the human population leading to a pandemic is a rare event. It remains that IAVs represent a risk of pandemics that is a continuous and unresolved threat for humans [76].

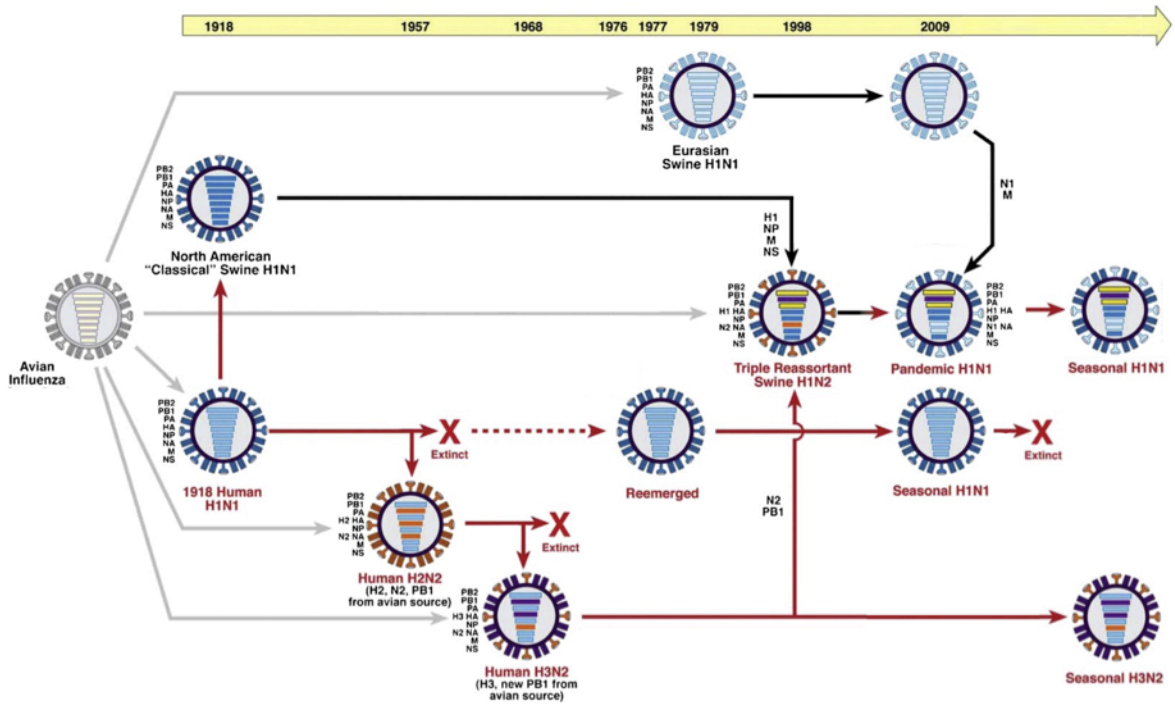


Figure 10: Circulation of human influenza A viruses since the 1918 pandemics. Grey arrows indicate avian origin of one or more segments; dashed red arrow indicates a period without circulation of H1N1 in humans; red arrows indicate circulation of human IAV lineages with potential swine transmission; black arrows indicate circulation of swine IAV lineages. Segment exchanges through reassortment are indicated. The 1918 pandemic virus was initially thought to derive from an avian virus, but the most recent phylogenetic studies suggest that it is a reassortant between a human and an animal (possibly a swine) virus. Adapted from [76, 88].

2. IAV particle and genome

Influenza viruses are spherical enveloped viruses (Figure 11a). The envelope is composed of a lipid bilayer, derived from the cellular membrane, in which are incorporated 3 viral proteins: the HA and NA spike glycoproteins and the M2 membrane channel protein. Additionally, virion envelope incorporates numerous host proteins whose function is largely unknown [89]. The homotrimers of HA are responsible for the virus attachment to the cellular receptors and for the fusion of the cellular and viral membranes. Conversely, the homotetramers of NA play a role in the release of newly formed virions, thanks to their sialidase activity. After virus entry, tetramers of M2 form ionic channels inducing an acidification of the viral particle required for the release of the viral segments from the endosome. The matrix protein M1 covers the inside part of the envelope and is involved in the assembly and budding of the newly formed virions. Although they are called non-structural proteins, low amounts of

NS2/NEP (Non Structural protein 2 / Nuclear Export Protein) and NS1 proteins are also present in the particle [90]. The viral genome is composed of 8 single-stranded RNA segments of negative polarity (vRNA). Each segment is encapsidated with oligomers of the nucleoprotein (NP) and associated to a viral polymerase, and thus form a viral ribonucleoparticles (vRNP). The viral RNA-dependent RNA polymerase consists of three subunits, polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2), and polymerase acidic protein (PA), which together are responsible for both transcription and replication of viral genome (Figure 11b). The genome of IAV encodes up to 18 distinct proteins, thanks to splicing and alternative translation, and their major functions are indicated in Table 2.

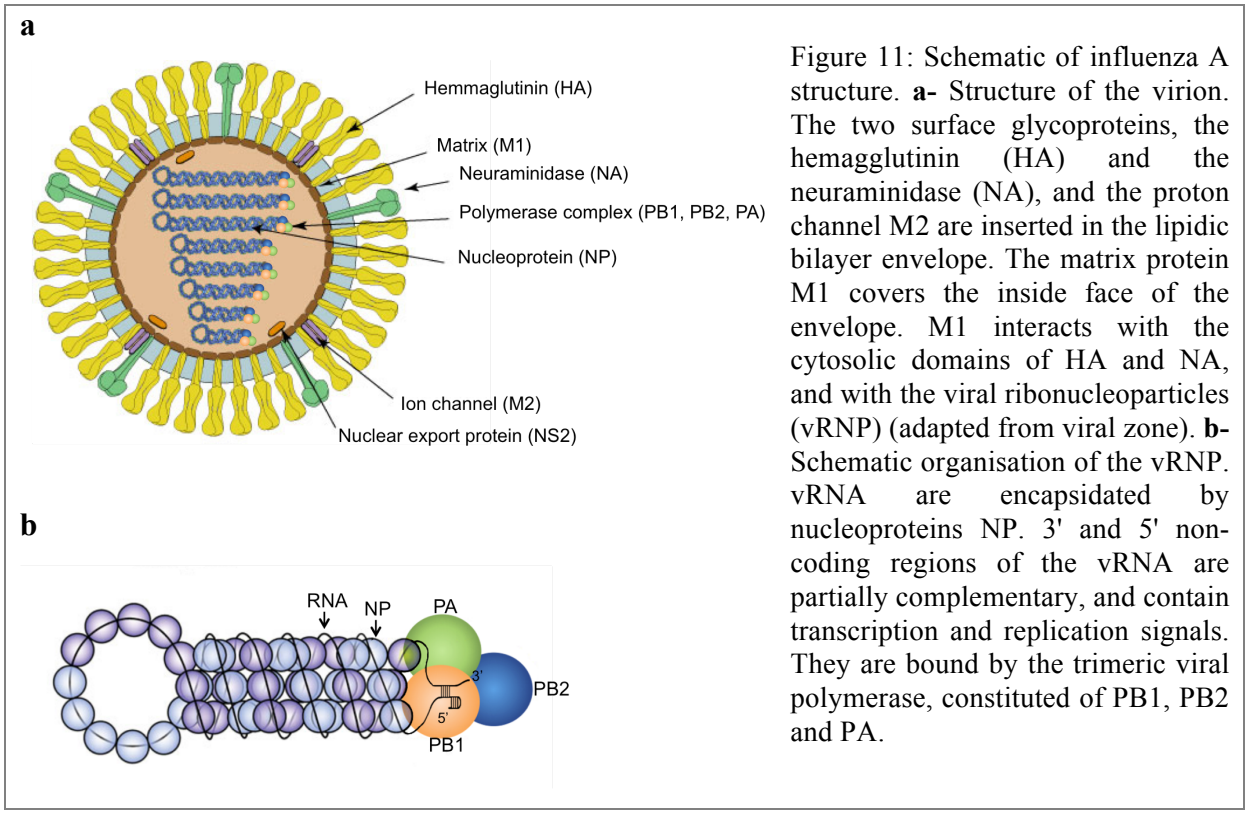
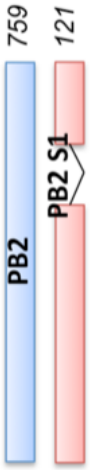
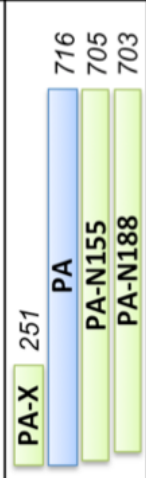
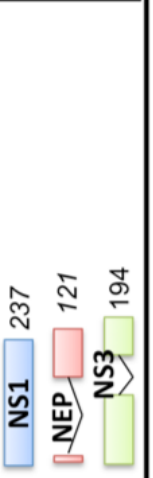


Figure 11: Schematic of influenza A structure. **a-** Structure of the virion. The two surface glycoproteins, the hemagglutinin (HA) and the neuraminidase (NA), and the proton channel M2 are inserted in the lipidic bilayer envelope. The matrix protein M1 covers the inside face of the envelope. M1 interacts with the cytosolic domains of HA and NA, and with the viral ribonucleoparticles (vRNP) (adapted from viral zone). **b-** Schematic organisation of the vRNP. vRNA are encapsidated by nucleoproteins NP. 3' and 5' non-coding regions of the vRNA are partially complementary, and contain transcription and replication signals. They are bound by the trimeric viral polymerase, constituted of PB1, PB2 and PA.

Table 2: ORFs encoded by IAV segments and functions of the related proteins. ORFs of the A/Puerto Rico/8/34 virus genome are schematically represented. The essential ORFs from intronless or unspliced mRNAs are boxed in blue, the essential ORFs from spliced mRNAs are boxed in red, and the remaining ORFs are boxed in green. The length of the proteins is indicated (in italic, amino acids number), as well as expected molecular size and main functions. Adapted by Cedric Diot from [71, 91].

Segment Length	Schematic view of the ORFs encoded by the viral segments	Theoric molecular sizes	Principal functions of the viral proteins
1 - PB2 2341 nt		PB2: 86 kDa PB2-S1: 55kDa	PB2: Component of the vRNP Cap binding of cellular mRNAs PB2-S1: Unkown
2 - PB1 2341 nt		PB1: 87 kDa PB1-F2: 10 kDa N40: 82 kDa	PB1: Component of the vRNP RNA-dependent RNA-Polymerase PB1-F2: Apoptosis, Regulation of polymerase activity N40: Unknown
3 - PA 2233 nt		PA: 84 kDa PA-X: 29kDa	PA: Component of vRNP Endonuclease (cap-snatching) PA-X: Host gene expression shut-off PA-N155 and PA-N182: Unknown
4 - HA 1778 nt		61 kDa	Surface glycoprotein Cell receptor binding, virus-endosome envelope fusion
5 - NP 1565 nt		56 kDa	Component of the vRNP Encapsidation of vRNA, nucleocytoplasmic trafficking of vRNPs
6 - NA 1413 nt		50 kDa	Surface glycoprotein Liberation of virus progeny
7 - M 1027 nt		M1: 28 kDa M2: 11 kDa M42: 11kDa	M1: Matrix protein Nuclear export of vRNPs and budding of neosynthesized virions M2: proton channel Release of vRNPs M42: M2 like function?
8 - NS 890 nt		NS1: 27 kDa NEP: 14 kDa NS3 : 20 kDa	NS1: Multifunctional protein Repression of the innate immune response, host gene expression shut-off, regulation of the polymerase activity, control of viral mRNAs fate NEP/NS2 : Nuclear export of neosynthesized vRNPs NS3 : Host adaptation

3. Viral cycle

The viral cycle begins by the binding of the HA to the cellular sialic acids that are linked to membrane glycolipids and glycoproteins. The attachment induces the clathrin-mediated endocytosis of the virus [92], although IAV can also enter via clathrin independent pathways [93, 94] (Figure 12A). Acidification of the endosome opens the M2 proton channel, which acidify the viral particle. The pH drop induces a transconformation of the HA, leading to exposure of a peptide fusion, which causes the fusion of viral and endosomal membranes [95]. Acidification of the particle destabilises the interactions between M1 and the vRNP, allowing vRNP release in the cytoplasm [96] (Figure 12C). vRNP are then translocated into the nucleus. The nuclear import of the vRNPs and NP relies mainly on a non-conventional N-terminus NLS on the NP that is recognised by α-importins of the nuclear pore complex [97]. Viral transcription and replication take place in the nucleus thanks to the viral RNA-dependant RNA polymerase (RdRP) [98].

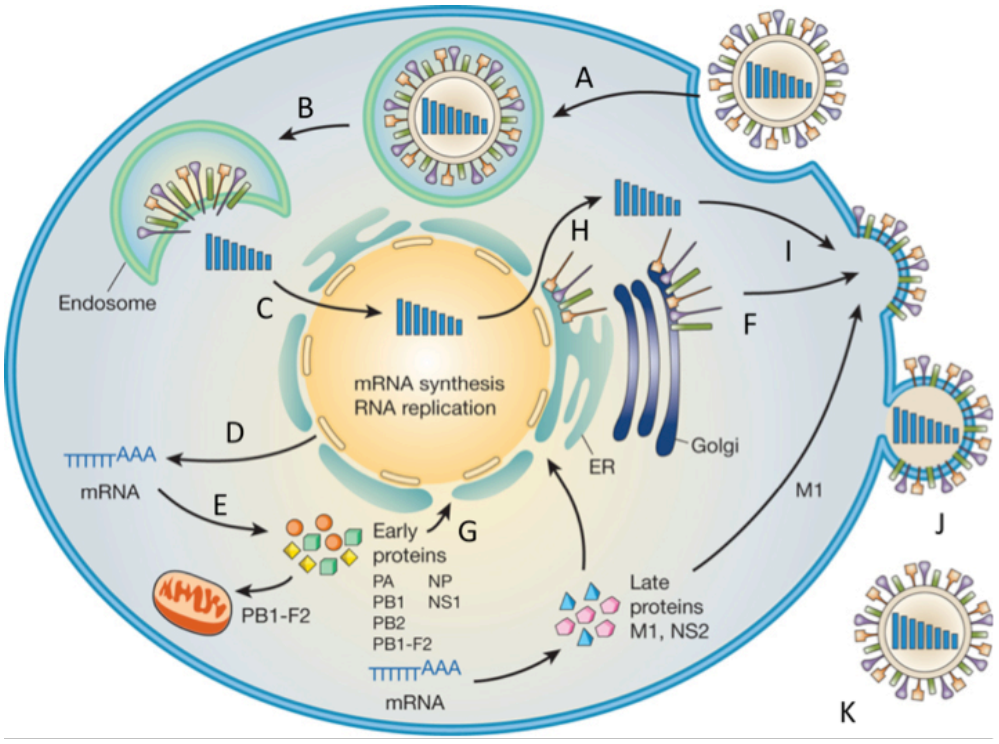


Figure 12: Influenza A virus life cycle. **A**- Attachment of the HA to the cellular receptor. **B**- Fusion of viral and endosomal envelopes, cytoplasmic release of vRNPs. **C**- Nuclear import of vRNPs. **D**- Nuclear export of viral mRNAs. **E**- Translation of viral proteins. **F**- Routing of envelope protein to the Trans-Golgi Network. **G**- Nuclear import of viral proteins responsible of the viral replication. **H**- Nuclear export of the neosynthesised vRNPs. **I**- Transport of vRNPs to the cellular envelope. **J**- Virus budding. **K**- Release of virus progeny. Adapted from [99].

Upon vRNPs entry into the nucleus, a first round of transcription occurs, where the polymerase transcribes vRNA into viral messenger RNAs (mRNA) [100]. The vRNA 5'NCR and 3' NCR constitute the promoter for transcription of each viral segment, which relies on the primed synthesis of a viral mRNA using host capped oligonucleotides by a process referred to as "cap-snatching". To initiate transcription, the viral polymerase first captures selected, capped, non-coding or pre-messenger host RNA associated with the transcribing RNA-polymerase II, via the Cap-binding domain of PB2. Then, PA cleaves it 8-14 nucleotides from the cap via its endonuclease activity. The resulting short-capped oligonucleotides serve as primers for the synthesis of viral mRNAs, which are last poly-adenylated by the stuttering of the polymerase on a poly-U sequence located near the 5'end of the RNA template (Figure 13).

The viral replication is taking place in two steps. The RNA segment of each vRNP is first copied in complementary cRNA strand of positive polarity, which forms cRNP after the loading of NP oligomers and the association of a polymerase complex (Figure 14). Each cRNP serves as template for the synthesis of vRNA segments, giving rise to progeny vRNPs (reviewed in [101]). For vRNA replication, the RdRp operates an unprimed RNA polymerisation, and this process is dependent upon newly synthesised PB1, PB2, PA and NP proteins. Newly formed polymerase act in trans during the replication, the 3′ end of the vRNA template needs to be released by the vRNP-associated polymerase to be copied by the trans-acting polymerase into the 5′ end of the new cRNA which then binds to the replicating trans-acting polymerase. After 5′ terminus binding, the first NP is added to the product cRNA adjacent to the replicating polymerase, initiating encapsidation of cRNA in a 5′ to 3′ direction. Thus, encapsidation of cRNA would be triggered by the sequence specific binding of the trans-acting RNA polymerase to the cRNA 5′ terminus [100]. The PB1 subunit performs the RNA polymerisation and harbours functional motifs common to all RNA-dependent RNA polymerases [102]. The influenza polymerase is devoid of proof reading activity (or 3'-5' exonuclease activity), accounting for the high mutation rate in the replicated viral segments. Newly synthesised vRNPs are either transcribed to provide high levels of viral protein expression (secondary transcription), or are exported from the nucleus to be incorporated into novel virions.

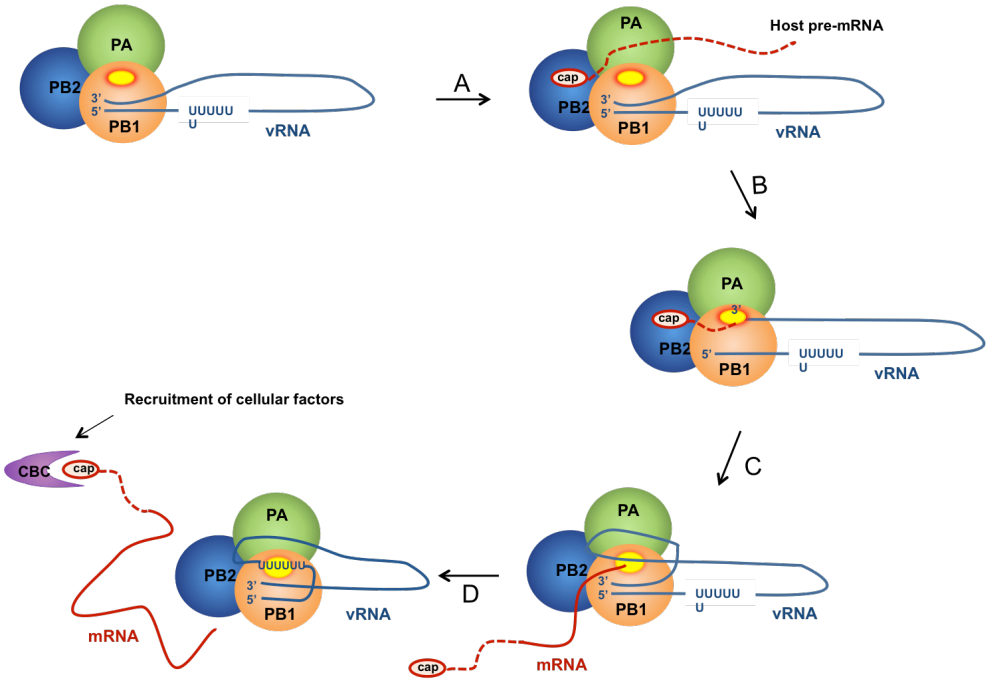


Figure 13: Model for the IAV RNA-dependant RNA polymerase transcription. The viral polymerase, consisting of the PB1, PB2 and PA subunits, is bound to the partially complementary 5´ and 3´ termini of the vRNA segment. PB2 binding to a host pre-mRNA initiates the transcription (**A**). The host pre-mRNA is then cleaved by the endonuclease activity of PA, the 3´ end of the capped RNA primer is positioned in the active site of PB1, along with the 3´ end of the vRNA template, allowing transcription initiation to take place (**B**). Once the 3´ end of vRNA has been copied it re-binds the polymerase, a process possibly facilitated by base-pairing between the vRNA termini (**C**). Viral mRNA are last poly-adenylated by the stuttering of the polymerase on a poly-U sequence located near the 5' end of the RNA template (**D**). The 5´ cap of mRNA is released from PB2 and is bound by the cellular nuclear cap-binding complex (CBC) triggering the recruitment of cellular factors for mRNP assembly. Adapted from [100].

The transcribed viral mRNA are exported into the cytoplasm and translated by the cellular machinery (Figure 12D-E). HA, NA and M2 proteins are routed to the trans-Golgi network (Figure 12F) and other viral proteins including the PB1, PB2, PA and NP mandatory for polymerase activity and vRNP assembly are imported into the nucleus (Figure 12G).

The Newly assembled vRNP are exported into the cytoplasm thanks to the NEP/NS2 protein (Figure 12H) and are, according to the generally-accepted model [103], specifically transported to the plasma membrane where sets of 8 distinct segments are incorporated in the viral particles (Figure 12I). The NA cleaves the sialic acids present at the surface of the cell and viral particles, thereby promoting the release of virus particles and preventing their aggregation [104] (Figure 12K). Following virions release, extracellular proteases harbouring trypsin-like activities and expressed in epithelia cleave the surface HA proteins into two

domains HA1 and HA2, which nevertheless remain covalently bound via a disulfide bound. This cleavage is essential for the fusion function of HA, hence for the infectivity of the produced viral particles [105]. The HA of highly pathogenic avian viruses of the H5 and H7 subtypes present a multibasic cleavage site, which can be cleaved in the TGN by the PC6 and/or furin proteases that are ubiquitous and therefore confer a wider tissue tropism to these viruses [104].

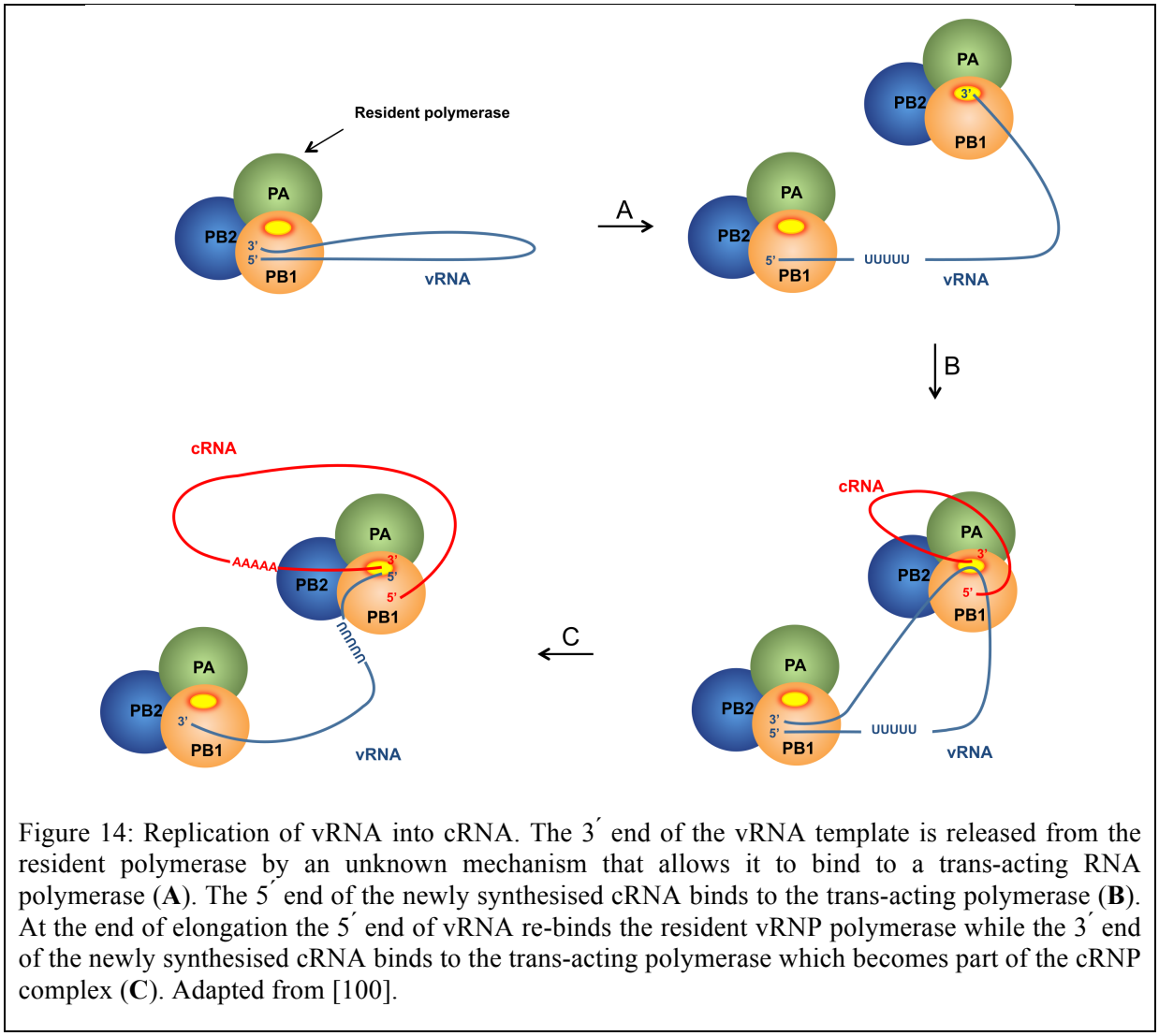


Figure 14: Replication of vRNA into cRNA. The 3´ end of the vRNA template is released from the resident polymerase by an unknown mechanism that allows it to bind to a trans-acting RNA polymerase (**A**). The 5´ end of the newly synthesised cRNA binds to the trans-acting polymerase (**B**). At the end of elongation the 5´ end of vRNA re-binds the resident vRNP polymerase while the 3´ end of the newly synthesised cRNA binds to the trans-acting polymerase which becomes part of the cRNP complex (**C**). Adapted from [100].

4. The viral polymerase

Insights regarding the trimeric RdRp emerged recently thanks to the near-concomitant structure resolution of RdRp complexes from influenza A, B, and C viruses [106-108] (Figure 15). At the structural level, the polymerases of each influenza virus type are strikingly similar.

They consist in a central invariant core made of the PB1 subunit, the N-terminal part of PB2 (PB2-Nter), and the C-terminal part of PA (PA-C). This core encompasses a central catalytic fold where the primed and unprimed synthesis of viral RNA occurs. The C-terminal two-third of PB2 (PB2-C) and the endonuclease domain of PA (PA-ENDO) are distinctly arranged around the polymerase core, leading to strikingly divergent polymerase configurations (Figure 15a). The Cap-binding domain, 627 domain and NLS domain of PB2 constitute such flexibly linked domains, and are accessible in the polymerase structure, as it is the case for PA-ENDO (Figure 15). The flexibly linked domains adopt stable and distinctive packing relative to the invariant core depending upon the binding of viral RNA ends, suggesting that different conformations are supporting the transcription and replication activities of the polymerase. A high flexibility of the polymerase complex has been detected in solution [109]. Indeed, a wide range of alternative polymerase conformation seems to exist, owing to many different relative conformations that the PB2-Cter domain seem to adopt [110]. One may speculate that the association of host factors could assist in the polymerase conformation remodelling and could contribute to stabilise alternative polymerase conformations.

5. The PB2 protein

The PB2 protein is acknowledged to be the main interactor with cellular factors within the viral polymerase. Moreover, most of the adaptive mutations in the polymerase complex concern the PB2 subunit (reviewed in [111]). For example, the position 627 is a well-known determinant of host range and pathogenicity [112]: PB2-627K is found in human strains whereas it is a PB2-627E in avian viruses. Recently, the cellular factor ANP32A has been shown to underlie the species-specificity of polymerase based on the replication efficiency of 627 residue of PB2 [113]. Other PB2 adaptive mutations that are located in the C-terminal, 627 and NLS domains, *i.e.* 569T, 588T, 591R, 661M, 682G, 701D, 702K are mostly lying at the surface [110], suggesting they concern amino acids required for optimal interactions with host cell factors, and for full polymerase functioning. In the N-terminal part of PB2 within the polymerase core, several adaptive mutations are lining the interaction interface with PB1, suggesting that these residues could be involved in interactions of the isolated PB2 subunit with host factors. A number of other host factors have been identified as acting on the viral life cycle via an interaction with the polymerase [114]. Examples of such an interaction are α importins with PB2 [115]. It has been suggested that the binding of PB2 to α-importins is

involved in the efficiency of the polymerase activity of 627K-human adapted polymerase, independently of its role in nuclear import of PB2 [116].

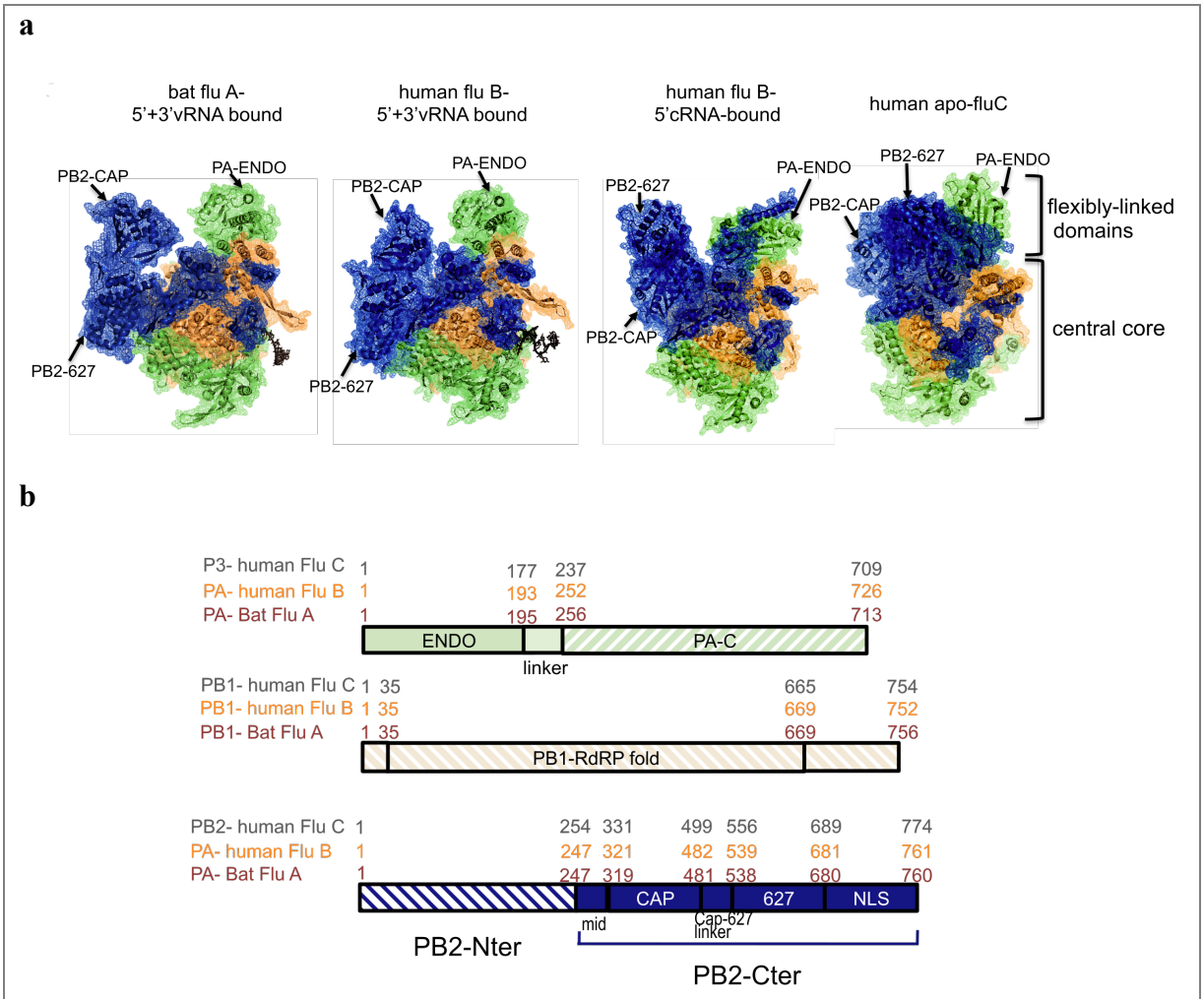


Figure 15: The influenza A polymerase. **a-** Overview of the different polymerase structures. The overall shape of the polymerase complexes in four resolved structures are shown with a fixed position of the central core. The PB1 subunit is coloured beige, PB2 blue, PA green. The pdb are as follows: bat FluA bound to 5'+3'vRNA pdb 4WSB [108]; human FluB bound to 5' + 3'vRNA pdb4WSA [107]; human FluB bound to 5'cRNA pdb 5EPI [109]; human apo-FluC pdb5d9 [106]. **b-** Schematics of the PA /P3, PB1 and PB2 domain organization. Numbering is given for human bat FluA, human FluB and human FluC as indicated. Hatched domains are those constituting the central core polymerase [110].

High-throughput interaction mappings led to the identification of numerous host factors interacting either with the RNP in an infectious context [117, 118], or with isolated polymerase subunits [119-121]. Despite limited overlap in the polymerase host partners identified, several cell functions emerged as being targeted by the viral polymerase, among which protein chaperoning, RNA metabolism, and nucleo-cytoplasmic transport. The

functional links between these interactions and the activities of the polymerase or of its isolated subunits are however, still largely undeciphered. According to the resolved structure, it can be hypothesised that a substantial part of these interactions are taking place through the PB2 protein, as it provides large, flexible and shape-changing interfaces accessible for interactions with host proteins.

6. Influenza A viruses and UPS

An increasing number of evidence point to the importance of the UPS in influenza A infection, in particular in innate immunity, viral entry and uncoating. The E3 ligase ITCH has been shown to promote the release of IAV vRNP from late endosomes via the ubiquitination of M1 [122]. Cullin3 emerged recently as a player in endosome maturation, and is also mandatory for viral uncoating in association with the SPOPL Substrate Recognition Factor [123-125]. Lastly, the viral capsids carry unanchored ubiquitin chains that activate, by mimicking misfolded protein aggregates, a histone deacetylase 6 (HDAC6)-dependent pathway that seems to be essential to IAV uncoating [126] (Figure 16).

After vRNA entry in the nucleus, the NP is mono-ubiquitinated, which seems important for virus replication. NP ubiquitination is counteracted by the DUB USP11, which depletion enhances virus production [127]. In addition, ubiquitination of PB1, PB2, and PA also enhances the viral polymerase activity [127-130], but the mechanism of their ubiquitination remains unknown. Conversely, the E3 ligase TRIM32 mediates K48-linked ubiquitination of PB1, inducing its degradation and reducing viral polymerase activity. The viral NS2 protein prevents the ubiquitination and degradation of the cellular factor AIMP2, which promotes viral replication by inducing the switch from ubiquitination to SUMOylation on M1 K242 residue [131].

The last aspect of IAV infection that is known to be linked to the UPS is the innate immunity. NS1 is the major suppressor of the IFN response during viral infection. It antagonises the E3 ligase TRIM25, which mediates RIG-I signal transduction by ubiquitinating its CARD domain [132, 133] (Figure 16).

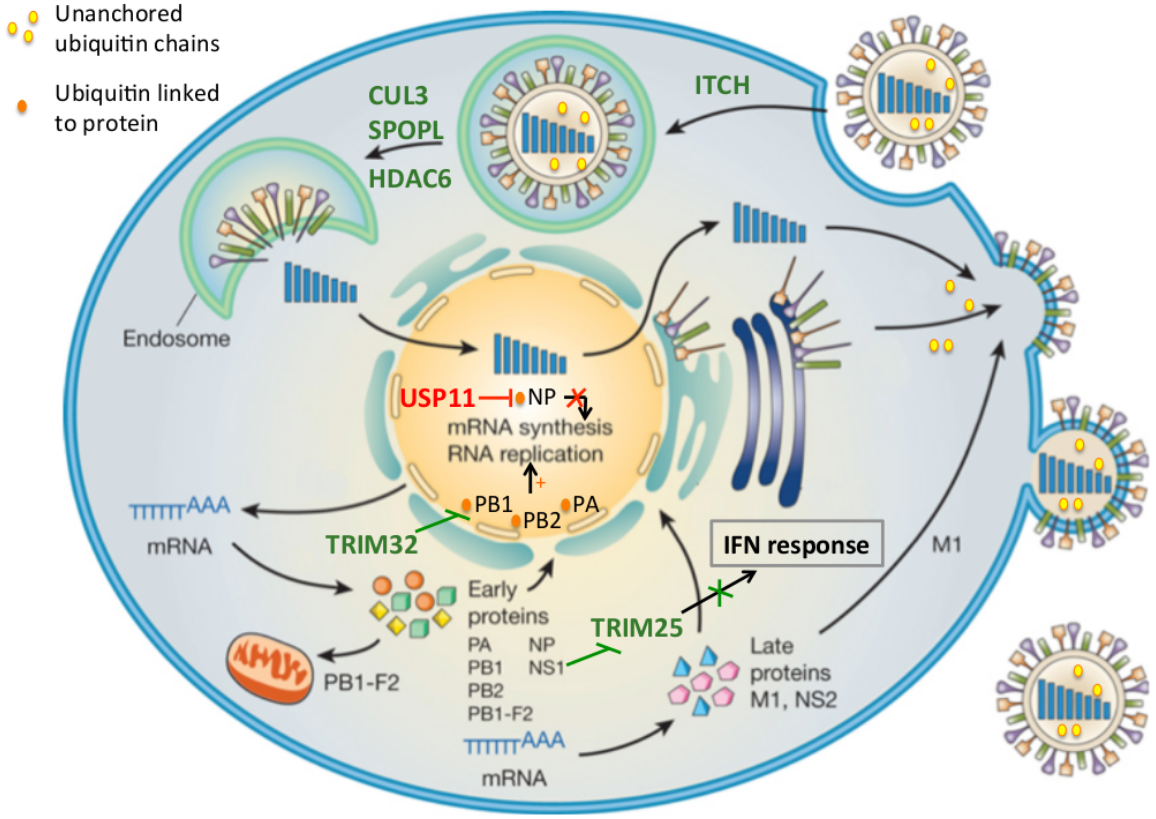


Figure 16: Interplay between influenza A virus and the cellular UPS. E3 ligases components (in green) ITCH, CUL3, SPOPL and HDAC6 are crucial for IAV entry and uncoating. The uncoating relies on the aggresome diversion by the virus thanks to unanchored ubiquitin chains contained into the virion (yellow). The DUB (in red) USP11 inhibits viral replication by counteracting NP mono-ubiquitination. The ubiquitination of viral polymerase subunits (orange) enhances the viral polymerase activity. But TRIM32 mediates K48-linked ubiquitination of PB1 and its proteosomal degradation. NS1 inhibits the interferon (IFN) immune response through the inhibition of the E3 ligase TRIM25.

Host-pathogen interactomic

Viruses are obligatory intracellular pathogens relying on host factors to complete their cycle. They must interfere with the host cell in order to provide a proper environment for their replication and spreading, as described previously (*UPS and viral infection* p25 and *Influenza A viruses and UPS* p39). Moreover, viruses have to counteract the cellular antiviral response to be produced and able to spread. Virus-host interactions have a wide effect on host physiology by modulating host proteins expression levels, post-translational modifications, localisations, degradations and interactions [67, 134]. As nicely summarises by Laura Bonetta [135]: "*Show me your friends, and I'll know who you are.*" also applies to proteins, finding

their interaction partners can reveal their functions. Discovering with which cellular proteins virus factors interact, and understanding their impact on the human Protein-Protein Interactions (PPI) network, is key to understand viral cycle, to find candidates for risk disease assessment and possible drug targets. Hence, the growing number of virus-host interactomes published in the recent years. But mapping interactions is not an easy task. Many of them are transient, and others occur only in specific cellular contexts. A battery of methods has been developed to detect virus-host protein-protein interaction at large scale.

1. Common bench assays to map Protein-Protein Interactions (PPI)

- *Yeast two-hybrid (Y2H)*

Yeast two-hybrid is a widely used technique to detect binary protein-protein interactions in a high throughput format. This system is based on the functional reconstitution of a yeast transcription factor by the interaction between two proteins, one fused to the DNA Binding Domain (DBD) and the other to the Activation Domain (AD) (Figure 17). Reconstitution of the transcription factor activates the transcription of a reporter gene, usually the HIS3 gene that allows yeast growth on selective medium lacking histidine [136]. The yeast two-hybrid is a highly efficient genetic screen for the positive selection of protein-protein interactions, which quickly became a leading detection method to map large-scale PPI between one factor (the bait) and a set of proteins expressed from a cDNA library (the preys). This efficient and user-friendly screening tool has however several inherent limitations. Since based on the nuclear localisation of a transcriptional reporter system, it is limited in the analysis of transcriptional activators as well as of proteins localised to membrane compartments. However, although the proteins to be tested are forced into the yeast nucleus for interaction, no bias between nuclear or non-nuclear proteins was observed [137]. Another reproach often made to the yeast two-hybrid system is the incomplete eukaryotic post-translational modifications of the expressed proteins, but it is nonetheless able to detect modification-dependent interactions up to a certain extent. Thus, the yeast cell is considered as offering an environment sufficiently natural for the analysis of protein interactions of other species [137]. However, when the same search of the human interactome was interrogated using two different yeast strains by Braun *et al.* [137], many interactions were uniquely detected in one screen. The authors concluded that different implementations of the same technology can lead

to different results, which may partially explain why datasets acquired by different groups exhibit limited overlap. Last, one main concern that has been raised for this technique is the risk of non-homogenous representation of the different cDNA preys in the ORF libraries, which strongly depends upon library coverage and quality. Indeed the Y2H system has a high false negative rate, being acknowledged to detect around 20% of existing PPI [137]. To overcome these limitations, others PPI detection methods have been developed.

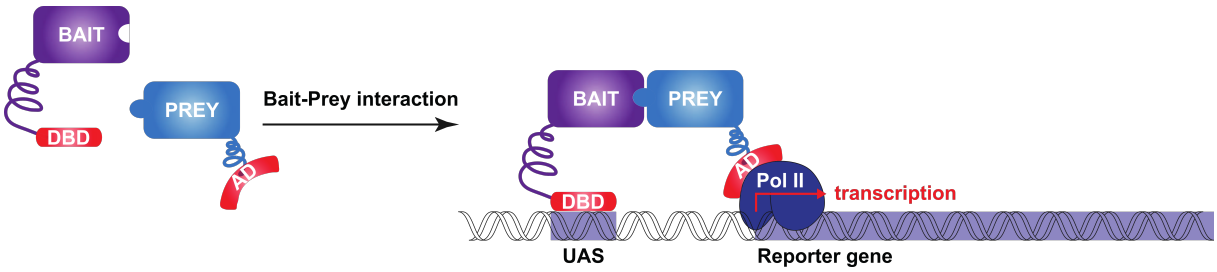


Figure 17: Yeast two-hybrid principle. Prey and bait are fused respectively to the DNA Binding Domain (DBD) and the Activation Domain (AD) of a yeast transcription factor. If the proteins of interest interact, the transcription factor is reconstituted and a reporter gene is expressed. Design by Laura Desban.

- *MAmmalian Protein-Protein Interaction Trap - MAPPIT*

MAPPIT is a two-hybrid technology based on functional complementation of a cytokine receptor signaling pathway. A bait protein is fused to a chimeric signaling-deficient hybrid cytokine receptor consisting in the extracellular part of a type I cytokine receptor and the transmembrane and intracellular domains of the murine leptin receptor deficient in STAT3 (Signal Transducer and Activator of Transcription 3) recruitment. When co-expressed with a prey that is fused to a receptor fragment containing functional STAT3 recruitment sites, the hybrid receptor is functionally complemented to efficiently recruit STAT3, leading to a restored signaling upon cytokine stimulation. STAT3 molecules are activated, migrate to the nucleus and induce transcription of a STAT3-responsive reporter gene [138, 139] (Figure 18). This assay as the advantage to take place in the cytoplasm which offers better interaction conditions for cytoplasmic proteins, but more importantly to take place in mammalian cells, which allow the expression of the full panel of post-translational modifications.

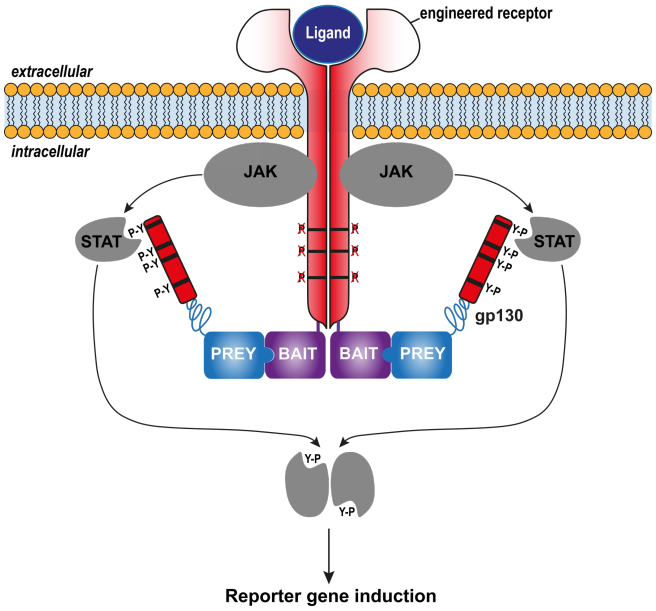


Figure 18: MAPPIT principle. The bait protein is fused to an engineered receptor with JAK kinase but lacking STAT activation sites. The prey is linked to an active STAT binding site. If the prey and the bait interact, then STAT activation will occur and activate a reporter gene system such as luciferase or puromycin resistance. Adapted from [138] by Laura Desban.

- *Protein-fragment Complementation Assays - PCA*

Protein-fragment complementation assays are based on the reconstitution of a reporter protein upon interaction of protein pairs expressed in fusion with autonomous fragments of the split reporter. These methods require that the two autonomous fragments of the split reporter do not have intrinsic activity (low background) and that the reconstituted activity is sufficiently sensitive. PCA methods were initially developed with fluorescent proteins [140], but the expression level of the tested proteins had to be high in order to distinguish the signal from the normal cellular background fluorescence. The sensitivity was later increased by the use of enzymes, in particular the *Gaussia princeps* luciferase, a small 185 amino acids enzyme that catalyses the oxidation of coelenterazine, to generate a luminescence signal 100-fold higher than Renilla or Firefly luciferases [141]. The High-Throughput *Gaussia princeps* Protein Complementation Assay (HT-GPCA), which has been developed in our laboratory [142], relies on the ability of interacting protein pairs expressed in fusion with the Gluc1 and Gluc2 complementary fragments of the *Gaussia princeps* luciferase to reconstitute a functional enzyme (Figure 19). This assay is performed by transfection of mammalian 293T cells, and allows the detection of binary protein-protein interactions in the context of mammalian cells. It has been used in the past to validate interactions detected by yeast two-hybrid and compare interaction patterns of protein from different genotypes of Human Papillomaviruses [143, 144].

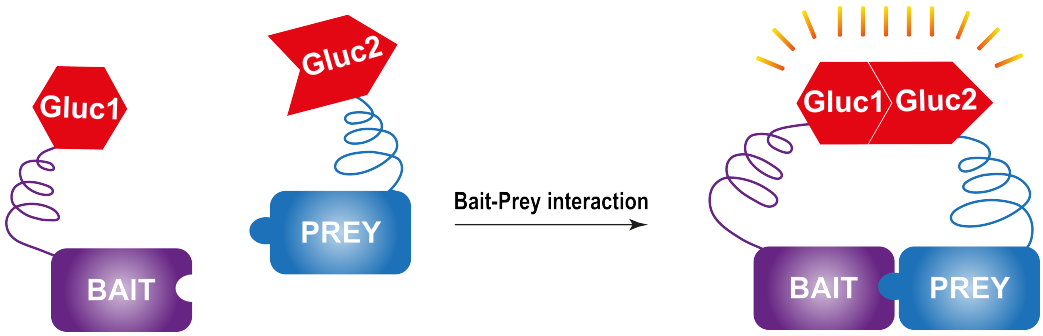


Figure 19: Protein Complementation Assays, the example of the *Gaussia princeps* Complementation Assay (GPCA). Bait and prey are fused to the Gluc1 and Gluc2 complementary fragments of the copepod *Gaussia princeps* luciferase. If the proteins interact the luciferase activity is restored and a luminescent signal can be detected upon addition of coelenterazine, the luciferase substrate. Design by Laura Desban.

Two by two co-transfection of a prey and a library of cDNA enable the systematic screening of interactions using split-reporter fusion expression plasmids. The main disadvantages of the PCA methods were lying in their more restricted format, both in terms of scale of the library and of experimental procedures. But nowadays the gateway cloning enables the high-throughput cloning of large cDNA collection representing an exhaustive set of human cDNA. The ORFeome is a human cDNA library comprising to date 12,000 cDNA of the 20,000 protein-coding human genes, and is still continuously implemented [145]. Automation of different screening steps with robots will enable in the near future to significantly enhance the experimental HT-GPCA format. For my PhD project, we developed an interactomics strategy using the HT-GPCA as a PPI screening method to provide a systematic screening in a comparative approach.

- *Affinity purification coupled with mass spectrometry (AP-MS)*

The methods described previously are all based on detection of binary PPI, whereas affinity purification followed by mass spectrometry (AP/MS) is a powerful approach to characterise multiprotein complexes. This approach, that has been use commonly on a large scale, is composed of two steps. First, an affinity reagent directed against a protein of interest is used to purify the complex. Purification is often realised thanks to TAP (Tandem Affinity Purification) tags which are dual-affinity tags allowing a sequential purification. Then the purified components are identified by mass spectrometry (Figure 20). The methodology

requires strong bioinformatics analysis to identify the proteins after mass spectrometry. AP-MS is a powerful tool to compare the composition of complexes according to conditions, but does not allow deciphering interactions inside the complex. This lack can nonetheless be compensated by the combination with a binary-interaction detection method.

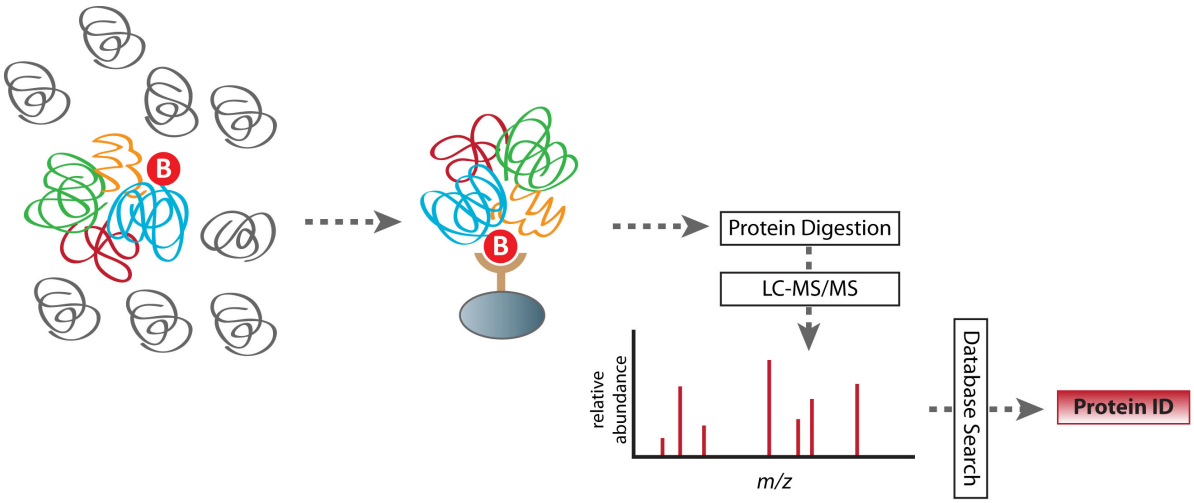


Figure 20: Affinity Purification - Mass Spectrometry. Protein complexes are purified from the protein lysates using an affinity reagent directed against a protein of interest. Complex components are then digested with trypsin, and subjected to LC-MS/MS analysis. Database searching is used for protein identification. Adapted from [146].

2. Improving interaction maps

- *Combining methods*

Information obtained from the binary interaction mapping studies allow the construction of interaction networks between proteins, called interactomes. The confidence of mapped interactions can be improved by validating them with an orthogonal method. As seen with the yeast two-hybrid, even with the same methodology different sets of experiments lead to small recovery rate [137]. Thus, interactions that have been detected by different assays can be considered as highly confident. Different methods have been developed to score binary interactions according to the type of screen that has been used and the overlap in different studies [137], or according to their network topology by clustering [147] for example.

Combining yeast two-hybrid system and AP-MS enables to obtain a more comprehensive view of an interaction network. First, both the composition of protein complexes and the interactions between individual protein pairs is elucidated, providing a more detailed view of the network. Second, interactions missed by one technology may be detected by the other, resulting in a more robust network (Figure 21). Finally, although many relevant interactions will only be detected by one approach, those that are confirmed by both Y2H and AP/MS represent a series of interactions of increased confidence.

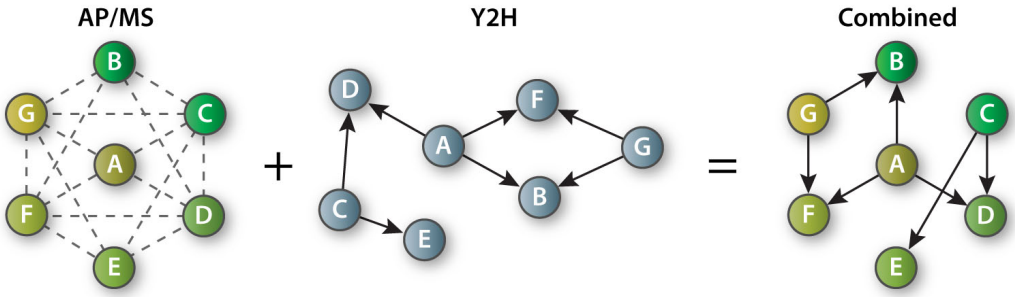


Figure 21: Combining AP/MS with Y2H presents a more detailed picture of interaction networks by identifying both protein complexes and binary interactions between proteins. [146]

- *Predicting interactions*

Different groups are working on improving interactomes by adding predicted interactions. The structure of the protein is compared to other structures in datasets and predictions are made on potential interactors according to known interactions between similarly structured proteins. This is based on the postulate that when two proteins are structurally homologous, they are more likely to have common interactors. When the structure is unknown, predictive structure can be calculated according to the sequence, by similarity with resolved structures [148]. Thanks to the structure, docking can also be done to test if two structures can interact and define the interacting domains [149]. Another tool that has been proposed to build and to analyse interactomes is cross-species comparisons. By comparing orthologues and their known interactions, it is possible to infer interactions in a species in which they have not been search for yet. Moreover, if a detected interaction in a species is also found in its orthologues, it increases the confidence of the interaction. Thus, conservation among species can be added as a criterion for confidence score calculation [150].

3. Influenza interactomics

- *Interactomes deduced from physical PPI map*

Interactomes have been constructed for influenza proteins mostly by affinity purification and mass spectrometry. They were conducted by transfecting the protein of interest or by infecting the cells, and thus explored one or several IAV protein interactions (Table 3). According to the methodology and the criteria chosen to select the interactions, the number of interactors found for the different proteins is quite heterogeneous. For example, 30 interactors were detected for PA in yeast two-hybrid by Tafforeau *et al.* [151], 134 with AP-MS by Wang *et al.* [152] and around 300 with AP-MS by Bradel-Tretheway *et al.* [121] and Watanabe *et al.* [120]. As described previously, it is to note that AP-MS allows the identification of complexes, and yeast two-hybrid binary interactions. This can explain the discrepancy between the studies of Tafforeau *et al*. [151] and Bradel-Tretheway *et al.* [121]. The higher number of interactors found by Watanabe *et al.* [120] compared to Wang *et al.* [152] could be due to differences in selection criteria, which highlights their crucial role when mapping interactions. Nevertheless, all these interactomes showed the targeting of common major cellular functions such as RNA processing, transcription, translation, transport and cell cycle.

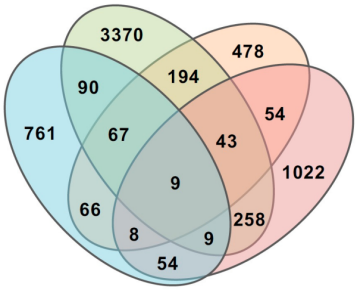
Even more than comparing the number of interactors found in each study, it is worth comparing the interactors they describe. Zao *et al.* [153] compared four IAV interactomics studies made by AP-MS [120, 153-155]. They found that the overlap of specific protein interactions identified by these studies is relatively low (Figure 22). They propose that reasons for the discrepancy could be the false positives and false negatives that are endemic to the methodology, or differences in the studied conditions such as virus strain, cell type, and statistical algorithm used for "cleaning" the dataset (Figure 22). Nonetheless, a total of 136 common protein interactions have been found in at least three studied approach, that can then be considered as high confidence interactions. These interactions concern mostly PB2 (132), PB1 (102), M2 (91) and NP (70). Of note, a lot of interactions are shared by multiple viral proteins. For example, PB1 and PB2 share 48 interactors, highlighting a strong redundancy for the targeting of cellular proteins. These latter might represent host factors which regulation is essential for implementation of the viral cycle.

Table 3: Summary of influenza interactomes studies. The viral proteins tested, the number of detected interactions, cellular functions impacted and the methodologies used are indicated. AP-MS: affinity purification coupled with mass spectrometry; Y2H: yeast two-hybrid.

Viral protein(s) tested	Number of detected interactions	Virus	Cellular functions impacted	Methodology	Reference
vRNP	45	H1N1 WSN A/WSN/33	Ubiquitination machinery Translation Transcription Cytoskeleton	Transfection AP-MS Peptide and Protein Prophet data analyses	Mayer *et al.* [156]
Polymerase	10		Cellular RNA synthesis, modification and transport	Transfection AP-MS	Jorba *et al.* [157]
PB1, PB2, PA, NP, NS1, NS2, HA, NA, M1, M2	87	H1N1 PR8 A/Puerto Rico/34	RNA binding Transcription Transport Signaling : NFκB, apoptosis, MAPK and WNT	Y2H and genome-wide expression profiling	Shapira *et al.* [119]
	66	H3N2 A/Udorn/72			
PA	306	H5N1 A/Vietnam/ 1203/04	Mitochondrial dysfunction Oxidative phosphorylation Aminoacyl-tRNA biosynthesis Ran signalling Caveolar mediated signaling	Adenovirus expression AP-MS Peptide and Protein Prophet data analyses	Bradel-Tretheway *et al.* [121]
PA-PB1	68		Translation Protein folding Proteolysis Regulation of transcription/RNA splicing PolII network Intracellular transport		
Polymerase	34				
PB1, PB2, PA, NP	PB1: 18 PB2: 59 PA: 30 NP: 5	H1N1 PR8 and WSN	Viral infectious cycle Cell cycle RNA processing Nucleocytoplasmic transport Transcription DNA repair Translation Response to unfolded protein DNA replication	Y2H	Tafforeau *et al.* [151]
PB1, PA		H5N1 A/VietNam/ 1194/04			
PA		H5N1 A/Turkey/ 651242/2006			
PB2		H3N2 A/Victoria/3/75			
M2-Cter	47	H5N1 A/Goose/Guangdong/1/96		Y2H	Ma *et al.* [158]
NS1	32	H1N1 A/New Caledonia/2006		Prediction and validation by Y2H and GST pull-down	de Chassey *et al.* [159]
NS1, NS2	NS1 only: 33 NS2:only: 28 Both: 18	H1N1 PR8 and WSN H5N1 A/chicken/Scotland/1959 A/chicken/Kurgan/3/2005 A/Vietnam/1194/2004 H7N7 A/chicken/Belgium/2003 A/equine/Prague/1/1956 H9N2	Transcription Translation Immune response signaling Vesicular transport Apoptosis Nuclear transport	Y2H Analysis with pISTil pipeline siRNA validation	de Chassey *et al.* [160]

		A/chicken/Guangdong/6/9 H15N8 A/duck/Australia/348/1983			
PA	6	H1N1 WSN A/WSN/33		Y2H, Co-IP GST-pulldown	Hsu *et al.* [161]
PB2	388	H1N1 WSN A/WSN/33	Transcription Translation Cell cycle mRNA splicing	Transfection AP-MS	Watanabe *et al.* [120]
PB1	322				
PA	304				
HA	351				
NP	574				
NA	675				
M1	659				
M2	531				
NS1	113				
NS2	42				
PB1-F2	81				
vRNP	171	H1N1 WSN A/WSN/33 PB2-strep	Transport Ubiquitination machinery Response to unfolded protein Transcription	Infection AP-MS	York *et al.* [117]
NS2	4	H1N1 WSN A/WSN/33		Y2H, Co-IP GST-pulldown	Gao *et al.* [131]
PB1, PB2, PA, NP, NS1, NS2, HA, NA, M1, M2, PB1-F2	849	H1N1 WSN A/WSN/33 PB2-strep	Amino acid metabolism Immune response signalling mRNA splicing Transport	Transfection AP-MS ComPASS and MiST algorithms	Tripathi *et al.* [154]
PB1, PB2, PA, HA, NA, NS1, NS2, M2	804	H1N1 PR8 A/Puerto Rico/ 1934	Amino acid metabolism Transport Transcription Translation Antigen processing and presentation	Infection AP-MS MiST algorithm	Heaton *et al.* [155]
NS1	64	H1N1 PR8 A/Puerto Rico/ 1934	Spliceosome Neurotrophin signalling Cell cycle	Transfection AP-MS Peptide and Protein Prophet data analyses	Kuo *et al.* [162]
PA-X	56	H5N1 A/Chicken/ShanXi/2/ 2006	Mitochondrial transport Lipid transport Nucleosome assembly RNA processing	Transfection AP-MS	Li *et al.* [163]
PA	134	H5N1 A/Chicken/ShanXi/2/ 2006	Translation Protein localisation DNA replication Protein transport Proteasome Cell cycle	Transfection AP-MS	Wang *et al.* [152]
PB1, PB2, PA, NP, NS1, NS2, HA, NA, M1, M2	357	H1N1 PR8 A/Puerto Rico/1934	mRNA splicing RNA transport Virus modulation	Transfection AP-MS SAINT algorithm	Wang *et al.* [164]
PB1, PB2, NP, NS1, M2	185 associated with at least 3 strains	H1N1 and H5N1 A/WSN/33, A/NewYork/ 1682/2009, A/Vietnam/ 1203/2004			
NS1, NP		H3N2 A/Aichi/2/1968			

The study from Lai *et al.* [165] used proteome-wide network approach to decipher topological and biological characteristics of the IAV interactome. They constructed an IAV–human PPI network considering two types of interactions: direct and indirect interactions with IAV proteins based on the VirHostNet database [166]. They found that IAV interacts with highly connected and central proteins of the human interactome, located in sub-networks showing few inter-connections. PB2 and NS1 are the two most connected viral proteins. Targeting highly connected hubs induce a strong impact on host though few interactions. Whereas interactions with less connected proteins allow a more fine tuning of the diversion.



References		Strain	Cell line	Viral infection	Statistical algorithm
Watanabe et al.	[120]	WSN	HEK293	No	N/A
Tripathi et al.	[154]	WSN	HEK293	No	ComPASS and MiST
Heaton et al.	[155]	PR8	A549	Yes	MiST
Wang et al.	[164]	PR88,WSN, Aichii, NY/2009, VN/2004	HEK293	Yes	SAINT

Figure 22: Venn diagram of the interactions found in the different AP-MS studied in the article of Zao *et al.* [153]. Conditions of these studies are indicated in front of the reference. Adapted from [167].

The description of these interactions maps can lead to direct applications, as described in the work of Watanabe *et al.* [120], which presented a screen for interaction by high-throughput AP-MS and its functional validation by siRNA. They argue that, in case of a pharmaceutical treatment, identified cellular factors essential for viral cycle will be less prone to mutation than viral proteins, and are thus good candidates for antiviral drug development. From the AP/MS screening, they found 1292 cellular factors interacting with one of the viral protein, 323 of them proved to be involved in viral cycle by siRNA experiments and 44 were sensitive to drug treatment. Finally, they suggest two drug targets: GBF1 and JAK1, which inhibition drastically reduced influenza virus titres in cultured cells with no appreciable effect on cell viability, demonstrating the direct use that can be made out of PPI maps construction.

- *Addition of PPI prediction*

Previously described approaches for interaction prediction have been used to complete the picture. De Chassey *et al.* [159] predicted NS1 potential interactors by first selecting

structural homologues of NS1, then identifying interactors of these homologues in databases. These proteins were considered as putative interactors of NS1, and were ranked according to a score that favours proteins independently identified from multiple structural homologues. Thirty-two potential interactions with NS1 have then been validated by yeast two-hybrid and GST pull-down. Y2H and pull-down confirmed respectively 81% and 72% of the predicted interactions. With a 100% recovery rate with proteins ranking high in their scoring system, this method allows the selection of very high confidence interactors. They then expanded this technique to the whole set of influenza A proteins, increasing the known network with 106 predicted interactions (Figure 23). It is noteworthy that few UPS factors have been identified by these approaches as interactors of PA, PB1, PB2, NP and NS1.

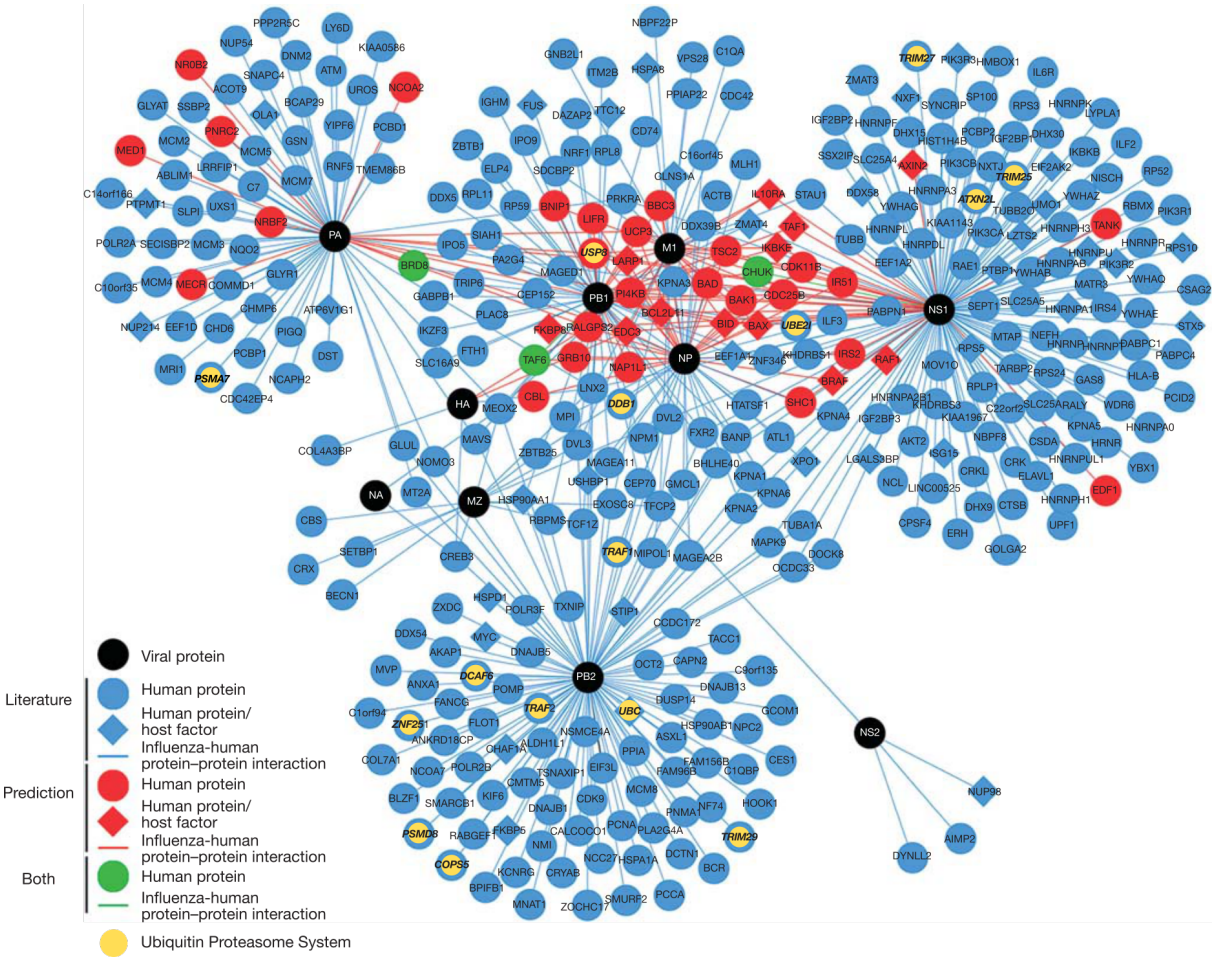


Figure 23: Influenza–human interactome combining predicted and literature data. Network composed of 515 interactions involving 10 influenza viral proteins and 364 human proteins. One hundred and six interactions were obtained from predictions, 407 from literature and 2 from both. Forty-seven of 364 human proteins have been identified to be host factors directly controlling viral replication. Members of the UPS are highlighted in yellow. Adapted from [159].

- *Comparative interactomics*

Comparison of interaction maps between strains enables to discriminate unvarying core pathogen-host PPIs from those involved in virulence or host specificity. Even though comparative interactomics were performed for other pathogens, as human papillomaviruses [143, 144], the recent study of Wang *et al.* [164] is the only one which takes advantage of a comparative interactomics for IAV viruses. In this work, Wang and collaborators screened by AP-MS for interactions of PB1, PB2, NP, NS1 and M2 proteins from five strains of IAV with the human proteome. They identified host factors interacting with at least three of these strains, pointing to proteins that might be crucial for IAV infection in a general manner. They focused on NS1 protein from H1N1, H3N2 and H5N1 viruses to compare their interactions with the human proteome, leading to the identification of new strain-specific NS1 interactors, including factors involved in the NF-κB signalling and in interleukin production. NS1 was already known to be a key player of the interplay with the host immune response, but these unique interactions suggested distinct perturbation strategies according to the virus strain, which might be key in *de novo* virulence and host determinants. Additional comparative interactomics with other IAV viral proteins will allow completing the map of conserved and specific interactions, and helping identify virulence factors and host determinants.

- *Interactomes focused on polymerase or vRNP complexes*

A number of interactomics studies focused on the viral polymerase or on the vRNP [117, 121, 151, 156, 157] (Table 3). They globally revealed enrichment in host factors implicated in processes directly linked to the polymerase activity, such as the RNA polymerase II-dependent transcription or the processing of mRNA. In addition, factors involved in other cellular mechanisms, like translation, cell cycle, nucleo-cytoplasmic transport, response to unfolded proteins, were also enriched in the vRNP/polymerase-associated factors. In the polymerase, many interfaces of the PB1, PB2 and PA subunits are engaged in inter-subunit interactions for the formation of the polymerase complex. However, as noted earlier, large interfaces for interaction are still exposed in the polymerase complex, notably provided by the PB2-Cter flexibly linked domains and by the PA endonuclease domain (Figure 15a). Moreover, interactions identified with isolated polymerase sub-units may also happen when the subunits are not in complex, beside the transcription or replication stages of the viral infection.

Some host cell partners have been identified as interactors of reconstituted polymerase complexes [157] or vRNPs [156], or with vRNP during infection [117], pointing to host factors associating with the transcription and/or replication-competent polymerase. Here again, a poor overlap exists between these polymerase/vRNP-directed interactomes regarding the individual identified factors. However, some common cellular functions are found to be targeted through interaction with the polymerase/vRNP, such as mRNA processing, nucleo-cytoplasmic transport or translation [117, 156, 157]. Interestingly, factors of the UPS have been described in several studies as interacting with the polymerase [117, 156, 159] (Figure 23 and Table 3). A few of the vRNP/UPS interactions have been explored more precisely, like USP11 and TRIM32 [127, 131] (see *Influenza A viruses and UPS* p39), but no study has been published yet to map the interactome between the UPS and IAV viruses.

To conclude, large-scale PPI maps that have been applied to IAV have provided a comprehensive view of IAV/host interplay [119, 120, 154, 155, 168], which, however, is still not exhaustive and warrants more validation to reach reliability given the small overlap observed between studies. The most suitable validation would be functional, but the relevance of all the identified factors cannot be assessed in such comprehensive interaction maps. Indeed, most of the IAV interaction maps have led to the functional characterisation of a single or small set of identified factors, creating a gap between the degree of knowledge reached through interactomics at the level of virus-host interaction and that reached regarding their biological relevance.

Aim of the work

The ubiquitin proteasome pathway is an essential cell regulator and a huge number of interactions are already known between the UPS and many viruses [69]. Growing evidence shows the implication of the UPS in influenza A virus infection at different stages of the viral cycle. Large-scale interactomics studies conducted on IAV proteins showed interactions with factors of the UPS. However, the knowledge of the role of the UPS in IAV infection is still incomplete. My PhD project has been devoted to provide a more comprehensive picture of the role of UPS in IAV infection, following several working steps.

I first undertook the systematic interaction mapping of the PB2 proteins of several IAV strains with the human UPS. We focused on the PB2 protein because it provides the viral polymerase with the main interface with the human proteome, as revealed by published host-IAV interactomes and by the resolved structure of the polymerase complex. Moreover, PB2 is a key player of IAV pathogenicity [111] and we were interested in comparing the interaction profiles of IAV strains according to their pathogenicity. I therefore developed a comparative interactomic strategy to assess the interplay of the human UPS with the PB2 of five IAV strains of various degrees of virulence. For that purpose, we used for the first time a high-throughput luciferase based complementation assay (HT-GPCA) as a screening tool for binary interactions between the PB2 protein and factors of the human UPS. We developed a straightforward interactomics screening strategy, providing both robust PB2/UPS interaction maps and enabling the rigorous comparison of PB2/UPS interaction profiles according to the strain.

I then drew a primary estimate of the biological relevance of the identified UPS partners of PB2. In these functional studies, I examined the productive infectious cycle of the $H1N1_{WSN}$ laboratory-adapted strain as well as of viruses derived from the $H1N1_{pdm09}$ and H3N2 seasonal strains upon siRNA-mediated depletion of the individual PB2 interactors. By such multi-strains functional assessment, we intended to assess the convergence and/or diversity of the functional outcome of the PB2/UPS interplay. I finally focused on a subset of PB2/UPS interactors: the DUBs, to study in more depth their role in the viral life cycle.

In all, the results obtained along my PhD project showed an extensive interplay between the viral polymerase PB2 protein and the human UPS, which recapitulated the IAV adaptation and evolution in humans. Most of the identified UPS factors proved relevant to IAV infection, in a strain-independent or strain-specific manner, revealing a diversified involvement of UPS/PB2 interplay in infection according to the strains. Functional studies on the DUBs pointed to different viral cycle steps impacted by these factors.

Comparative profiling of the Ubiquitin Proteasome System interplay with the influenza A virus PB2 polymerase protein

Benchmarking the HT-GPCA methodology for the detection of interactions with the PB2 proteins

To assess the interactions between the PB2 protein and the UPS we choose to use the high-throughput *Gaussia princeps* luciferase complementation assay (HT-GPCA). This assay based on the reconstitution of a luciferase activity allows testing binary interactions in mammalian cells (see *PCA* p43). Moreover the low false positive rate, around 5% [142-144], and the possibility to assess interactions in a high throughput format enhance the suitability of this technique for the comparative mapping of interactions between PB2 proteins and the UPS.

Since the library of host factors that we used for the screening is fused to the Gluc1 fragment, corresponding to the N-terminal part of the *Gaussia princeps* luciferase, PB2 has to be fused to the Gluc2 fragment (Figure 24a). In order to choose the position of the fusion on the PB2, we assessed the interactions between PB2 fused to Gluc2 in N- or C-terminus position, with a Positive Reference Set (PRS) of proteins known from the literature to interact with PB2, and a Random Reference Set (RRS) of proteins *a priori* not interacting with PB2 (Figure 24b). The PB2 proteins used for HT-GPCA evaluation were those chosen for the comparative screening with the human UPS and originated from five IAV strains of various degree of virulence in humans, *i.e.* the laboratory-adapted strain $H1N1_{WSN}$ (A/WSN/33), the low to mildly virulent human seasonal viruses $H1N1_{pmd09}$ (A/Bretagne/7608/2009) and H3N2 (A/Centre/1003/2012), and the highly virulent viruses $H1N1_{1918}$ (A/Brevig mission/1/1918) and H7N9 (A/Anhui/1/2013).

In both configurations consisting of the Gluc2 fragment fused to the N-terminus or the C-terminus of PB2 (Gluc2-PB2 and PB2-Gluc2 respectively), the PRS and RRS gave distinct range of signals (Figure 24b) allowing the proper discrimination between interacting and non-interacting proteins with PB2. Regarding the PRS, interactions with the nuclear import

factors, involving the extreme C-terminal Nuclear Localisation Signal motif of the PB2 proteins, generated lower luciferase signals in the Gluc2-PB2 than in the PB2-Gluc2 configuration. However, similar levels of PPI detection were obtained with both N- and C terminal positions of the Gluc2 tag, for the other PRS, as the DDX3X, RED, and eIF4e proteins. In the tridimensional structure of the viral polymerase [108], the C-terminal two third of the PB2 protein encompassing the Cap binding and the 627-NLS domains (Figure 15a), are distinctly arranged according to the type of vRNA extremity bound. In contrast the N-terminus of the PB2 protein takes part to the core polymerase forming a stable fold with the PB1 and the PA C-terminal domain in which the N-terminal helices of PB2 are exposed [108] (Figure 15a). According to these elements, we choose to use the N-terminal GLuc2 fusion of the PB2 proteins for UPS interaction screening in order to enhance the chance of detecting most possible interactions with the PB2-Cterminal domains. The HT-GPCA can tolerate a certain distance between the interacting proteins, as shown by the detection of triple complexes [142, 169], so that the Gluc2 fused to N-terminus of PB2 would enable the detection of interactions involving the C-terminal part of PB2. It is however assumed that, using a single configuration for HT-GPCA could be at the cost to loose some interactions strictly dependent upon the position of the Gluc2 tag. For more exhaustiveness, the PPI screening should be applied using 2 (N-ter, C-ter) or even 4 (Gluc1 in N- and C-ter, Gluc2 in N- and C-ter) possible configurations, which is out of scale for high throughput screening.

We noticed variations in the overall levels of luminescence generated by the PRS and RRS according to the PB2 protein, which were not related to differences in Gluc2-PB2 protein accumulation, as detected by western blot (Figure 25). They rather likely reflect differences in the intrinsic binding properties of the 5 PB2 proteins.

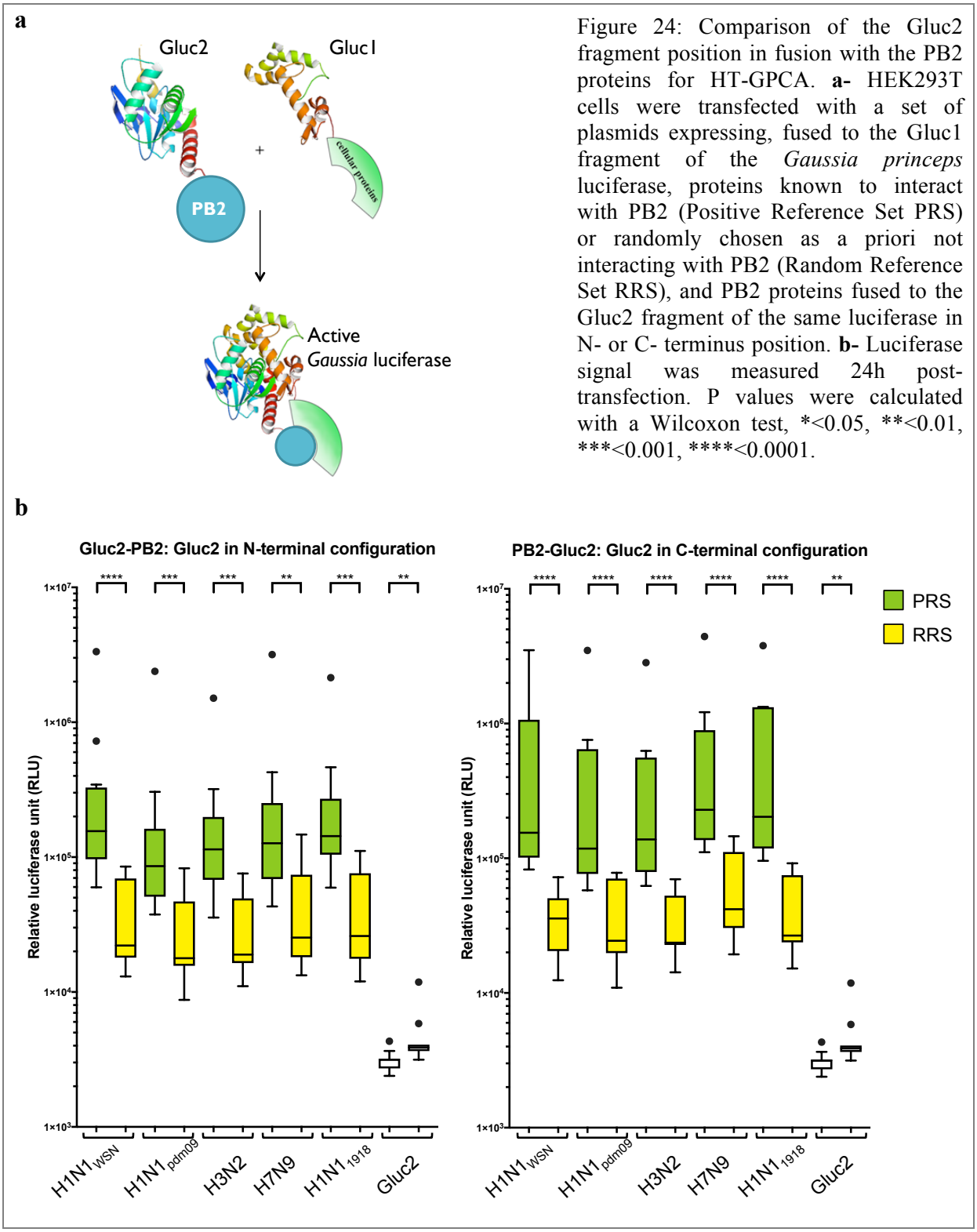


Figure 24: Comparison of the Gluc2 fragment position in fusion with the PB2 proteins for HT-GPCA. **a-** HEK293T cells were transfected with a set of plasmids expressing, fused to the Gluc1 fragment of the *Gaussia princeps* luciferase, proteins known to interact with PB2 (Positive Reference Set PRS) or randomly chosen as a priori not interacting with PB2 (Random Reference Set RRS), and PB2 proteins fused to the Gluc2 fragment of the same luciferase in N- or C- terminus position. **b-** Luciferase signal was measured 24h post-transfection. P values were calculated with a Wilcoxon test, *<0.05, **<0.01, ***<0.001, ****<0.0001.

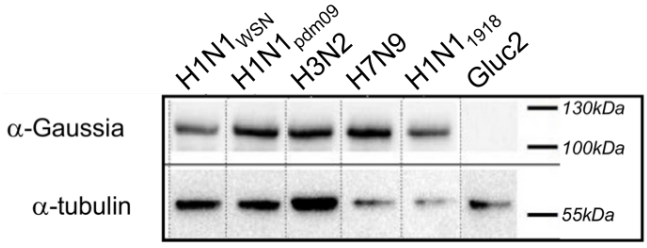


Figure 25: Expression levels of the Gluc2-PB2 proteins in HT-GPCA settings. The Gluc2-fused PB2 proteins in whole cell lysate of HT-GPCA samples were analysed by Western-blotting using rabbit anti-Gaussia antibodies, anti-tubulin was used for loading control.

The UPS-targeted library

A paramount issue for our systematic interactomics strategy was the quality and exhaustiveness of the library representing the screened subarray of the human proteome. We constructed an UPS-dedicated library suitable for PPI detection via the HT-GPCA method, by assembling 590 cDNAs encoding for 558 unique human UPS factors (Digital supplementary data 1), cloned in fusion with the N-terminal Gluc1 fragment of the *Gaussia princeps* luciferase [170]. The coverage of this library was estimated in comparison with an *in silico* human UPS database, designated "Human UPS", that contained a total of 1,277 factors obtained by combining the existing DuDe and UUCE ubiquitin-dedicated databases [34, 35] (Digital supplementary data 1). Our UPS-dedicated library covers 50% of the E1 activating enzymes, 64% of the E2 ubiquitin conjugating enzymes, 46% of the non-cullin E3 ubiquitin ligases, 28% of the cullin-based E3 ubiquitin ligases and 52% of the DUBs (Figure 26). With few exceptions (E3-DWD, DUB-ULP), the distribution of the different sub-categories of E3 ubiquitin ligases and of DUBs in the UPS-dedicated library is comparable to the database-derived human UPS composition (Figure 26). This was of key importance since E3 ligases and DUBs provide the human UPS its complexity and flexibility towards protein targeting, and are consequently the major targets of viruses [171]. The UPS-dedicated library thus offers an accurate representation of the global human UPS, ready to use for interaction mapping by HT-GPCA with any protein.

We then estimated the background HT-GPCA signal of the UPS library. For this purpose, the library was tested against the complementary Gluc2 fragment alone. The vast majority of constructs gave a weak HT-GPCA signal. Log10 transformation of the luminescence values resulted in their nearly normal distribution (Figure 27). The small shoulder on the right side of the distribution curve at 4.2 corresponds to 23 Gluc1-UPS factors/Gluc2 protein pairs, which give high luminescence values against the Gluc2 alone (Annexe 1, Digital supplementary data

2). These 23 Gluc1-UPS are likely pro-aggregating and should be considered with high precaution when tested with a Gluc2-fused bait.

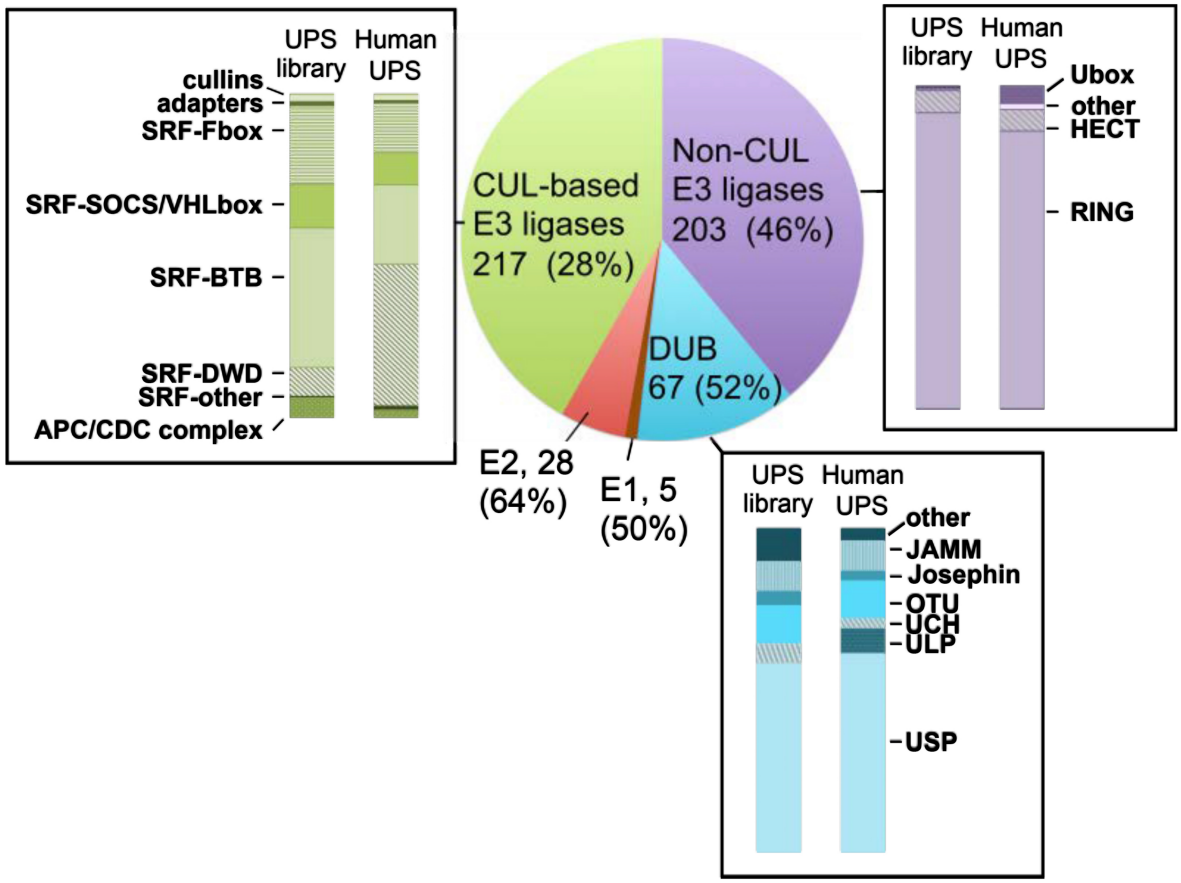


Figure 26: Pie-chart representation of the UPS-dedicated library composition. For each category of the UPS, the number of factors is indicated, with the percentage of coverage of the *in silico* Human UPS database in brackets. The representativeness of sub-categories for the DUB, cullin-based (CUL-based) and non-cullin (Non-CUL) E3 ligases are further detailed for the Human UPS and the UPS-dedicated library.

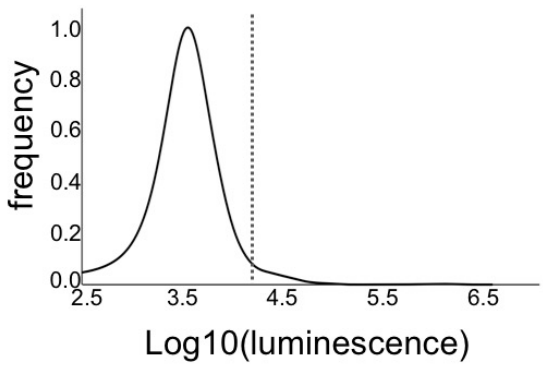


Figure 27: HT-GPCA screening of the UPS library with Gluc2. HEK-293T cells were co-transfected with plasmids encoding Gluc1-UPS and the complementary fragment Gluc2 alone. Twenty-four hours after transfection, the cells were lysed and the luciferase activity was measured. The log10 of the luminescence values were represented according to the frequency in the KDE-plot representation. On the curve, the dashed line indicates the position of the "shoulder". *This experiment was carried out by Juline Poirson at the Ecole Superieure de Biotechnologie de Strasbourg.*

HT-GPCA screening of PB2-UPS interactions

The screening of the 590 factors of the UPS-dedicated library was conducted via 12 independent HT-GPCA experiments, wherein a set of 25 to 72 exploratory Gluc1-UPS factors were screened for interactions with the five individual Gluc2-PB2 proteins, along with factors of the PRS and RRS (Digital supplementary data 3). For each experiment, a Positive Threshold (PT) was calculated per PB2 protein, based on the distribution of the luminescence values generated by the PB2-UPS pairs. The PT corresponded to the third quartile + 1.5 times interquartile space (PT = Q3 + 1.5(IQR)). UPS factors generating outlier luminescence values above the PT were selected as potential interacting partners (Figure 28a and Annexe 2). This selection was validated by the distribution of the PRS and RRS relatively to the PT of the corresponding PB2 protein (Figure 28a and Annexe 2). A fraction of the PRS consistently generated luminescence values under the PT (Digital supplementary data 3), indicating the high stringency of the selection of positive interactions using the defined PT. While the PT differed between the various PB2 proteins, their relative levels were preserved between experiments (Figure 28b), according to differences in PB2 intrinsic binding properties detected with the PRS and RRS sets of proteins (Figure 25). Such disparity did not alter the sensitivity of HT-GPCA to identify potential interacting partners, since the number of selected UPS factors was independent of the PT level of the PB2 proteins (Pearson's product-moment correlation test: R2=-0.2207, p-value= 0.09015) (Figure 28c). The initial PB2/UPS screening identified 91 UPS cDNA expressing putative interacting partners of at least one PB2 protein, corresponding to 80 unique UPS proteins (Figure 28d) (Digital supplementary data 3). Twenty-four factors were selected with all the 5 PB2 proteins, whereas the remaining 56 factors exhibited distinct interaction patterns between strains (Figure 28d).

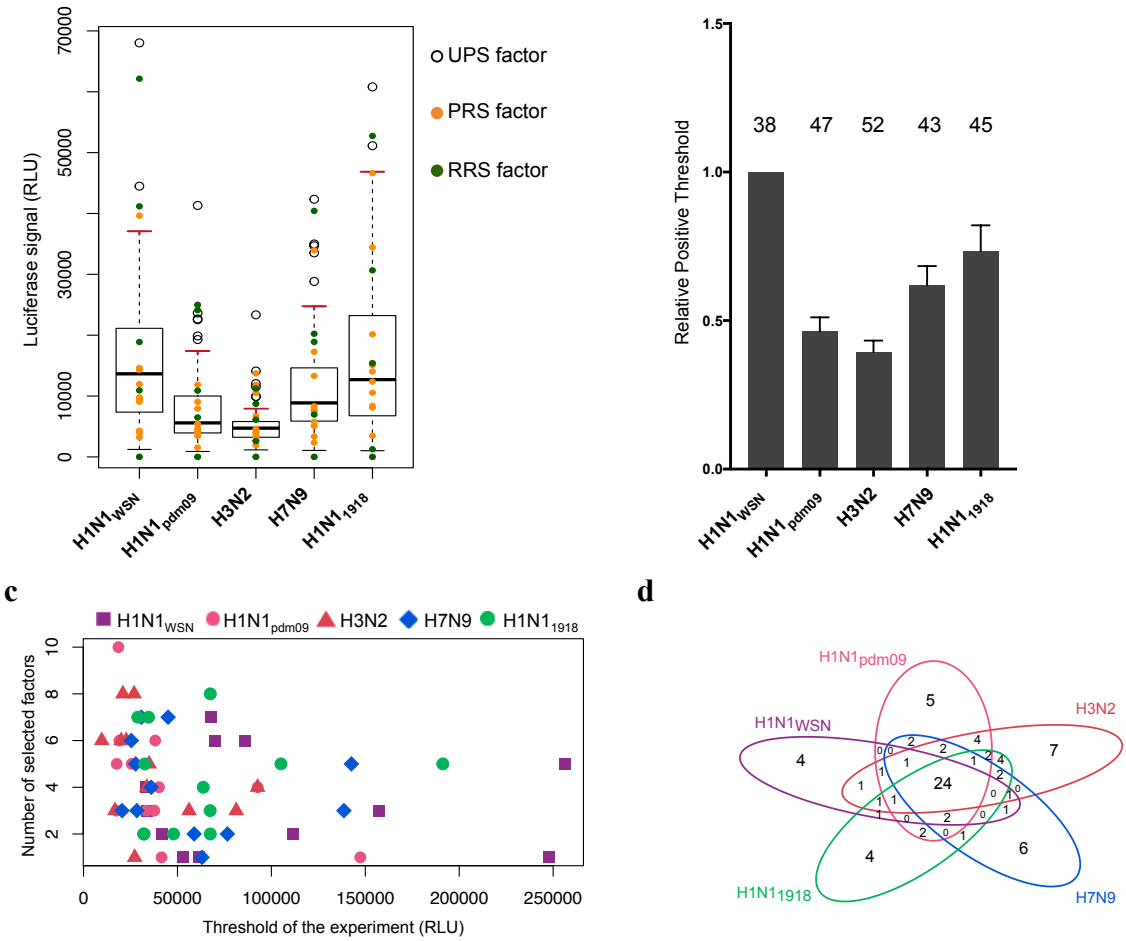


Figure 28: First screening of the UPS library with the PB2 proteins in HT-GPCA experiments. **a-** Example: screening results of the first experiment. For each PB2 protein, whisker-plots were generated from the luminescence values of the UPS-PB2 pairs. The Positive Threshold (PT = Q3 + 1.5(IQR)) is highlighted in red for each PB2. Outlier luminescence values of tested factors, represented in open circles, were selected as potential positive interactions. The RRS and PRS values (yellow and green dots respectively) were plotted over the whisker-plots afterwards, to evaluate the accuracy of the outlier-based selection. **b-** Positive Thresholds (PT) obtained in each individual HT-GPCA experiment were calculated relative to the PT of PB2 $H1N1_{WSN}$in the corresponding experiment. The total number of UPS factors selected is given at the top of each histogram for each of the PB2 protein. **c-** Number of selected factors according to the PT in each experiment for the 5 PB2 proteins. There is no correlation between the value of the PT and the number of selected factors (Pearson's product-moment correlation test: R^2=-0.2207, p-value= 0.09015). **d-** Venn diagram indicating the numbers of common UPS factors identified with the different PB2 proteins.

Post-screening retesting of the PB2-UPS interacting pairs and hierarchical clustering of the PB2-UPS interaction profiles

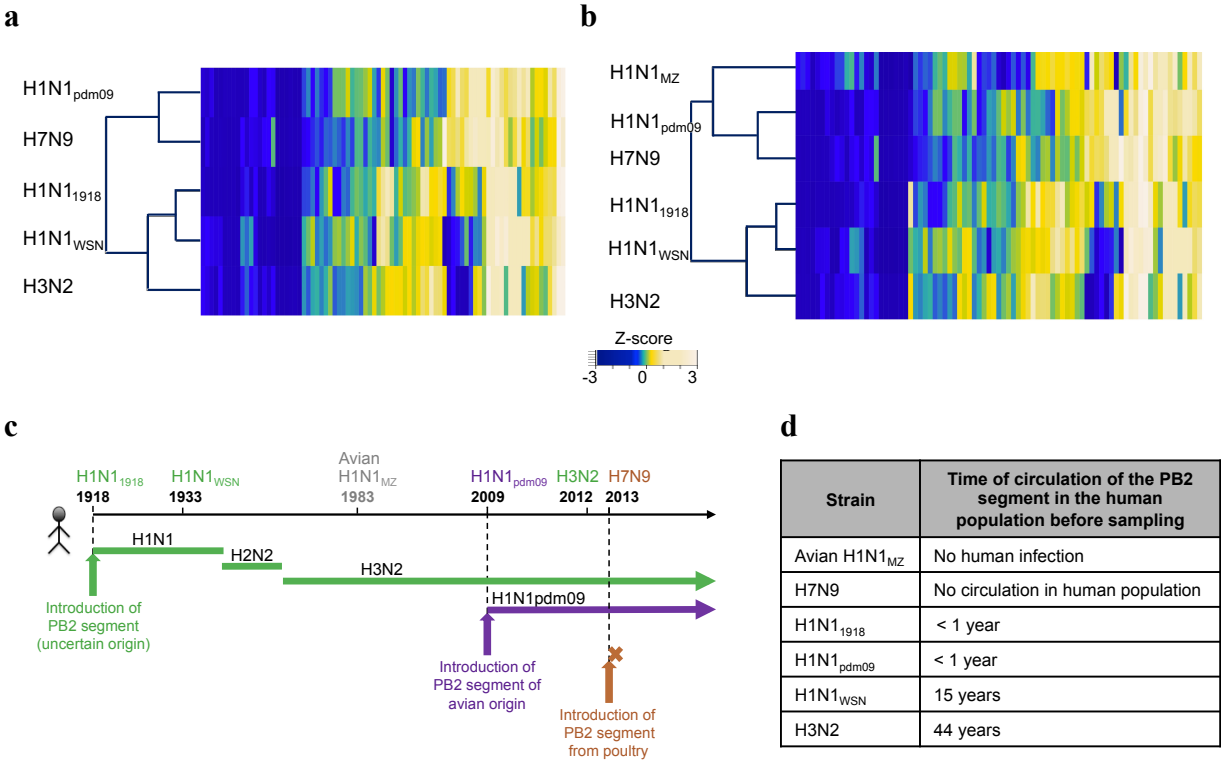
Validating interactions detected in the initial screening is necessary to enhance the robustness of the interaction data set. Therefore, the interactions selected from the initial screen were retested by applying the NLR (Normalized Luminescence Ratio) method to the HT-GPCA assay [142]. This method, which takes into account the background interaction level of the Gluc1- or Gluc2-fusion partners, has been shown to accurately discriminate PB2 interacting partners from non-interacting proteins in an infectious context [118]. To improve assay detectability and to allow a parallel comparison, each of the 80 selected UPS factors was retested against all five PB2 proteins, in three biological replicates, independently of their initial pattern of selection (Digital supplementary data 4). For each PB2 protein individually, the NLR values obtained with the 11 factors of the RRS were used to calculate a 99.73 % confidence interval, which corresponds in a normally distributed population to the mean +/- 3 standard deviations. The upper limit of this confidence interval was used as the threshold for the selection of positive interactions (Digital supplementary data 4). The interactions were considered as validated in the NLR retesting when they were three times out of three experiments above the positive threshold, which maintained the highest stringency for the identification of PB2/UPS interactions. With these criteria, 38 factors detected with the screening were not validated for interaction with PB2 with the NLR retesting. Forty-two UPS factors were validated for their interaction with at least one of the five PB2 proteins studied (Annexe 3). This represents a 52% validation rate of the initial screening. It is assumed that some positive interactions might have been overlooked by such stringent selection criteria, especially those positive 2 times in the NLR test. We nevertheless kept these criteria to obtain a more robust PB2/UPS interaction map and considered the 38 UPS factors selected as the high confidence set of UPS interactors with PB2.

The pattern of PPI validation according to the PB2 strain overlapped partly the pattern of the first screening. Ten UPS factors were validated with all the PB2 proteins identified in the first screening, 22 with only a subset of the PB2 proteins identified in the first screening, and 23 were validated with PB2 proteins that did not score positive in the first screening (Annexe 3). The main explanation of this partial overlap likely lies in the stringency in the positive selection with NRL retesting. Indeed, most of the interactions with discrepancies between the first screening and validation were scoring positive 2 times over 3 experiments in the NLR

retesting assays. Only six factors scored positive in NLR retesting 0 or 1 time over the 3 retesting experiments. For those factors, the initial enhanced luciferase activity detected in the first screening was likely due to a high interaction background. The NLR retesting enables to detect it against the empty Gluc2 fragment and to exclude the corresponding factors. Their reduced number indicates that interaction background-induced selection in the first screening occurred only for a minority of UPS factors, and highlights the suitability of the HT-GPCA assay for high-throughput screening of protein-protein interactions. Overall, the comparison between the initial HT-GPCA screen and the NLR retesting highlights the absolute necessity to proceed to a secondary PPI assay. It also indicates that the selection criteria chosen after the NLR retesting are probably over-stringent, and that a number of interactions between selected UPS factors and particular PB2 proteins might have been omitted. These criteria were nevertheless retained to select for a set of factors with the highest confidence interactions for further functional exploration. It implies however that the patterns of interactions considered as validated would not accurately reflect the strain specificity of PB2/UPS interplay.

The NLR retesting experiments performed with the five PB2 proteins and the same matrix of UPS factors enabled a rigorous comparative interaction mapping. As discussed above, the extremely strict criteria for selecting the highest-confidence UPS/PB2 interactions after NLR resting experiment precluded from deducing the UPS/PB2 strain specificity. We reasoned that the comparison of all NLR generated in a single retesting experiment, including validated and non-validated UPS factors, is more appropriate to compare the interaction profiles of the PB2 proteins of the different strains. The whole set of selected UPS factors but 4 were retested at once one time, in order to obtain an experimentally homogenous set of NLR for each PB2 protein. PB2 interaction profiles were derived from z-score calculation for each PB2 protein, and were compared by agglomerative hierarchical clustering between the different strains (see *Materials and Methods* p111). This comparison was used to calculate an interaction dendrogram (Figure 29a). PPI profile-based clustering separates viral strains in two groups: on one hand strains recently introduced in the human population from an avian origin ($H1N1_{pdm09}$ and H7N9) and associated with a short time (<1year) ($H1N1_{pdm09}$) or without any (H7N9) circulation in humans; on the other hand human adapted-strains that sequentially derive from each other ($H1N1_{1918}$, $H1N1_{WSN}$, H3N2) (Figure 29c). The PB2/UPS interaction profiles of this latter group further clustered according to the duration of circulation of the PB2 segment in the human population (Figure 29d). We next analysed the interaction profile

of PB2 from the purely avian strain A/Mallard/Marquenterre/Z2371/83 ($H1N1_{MZ}$) with the same set of UPS factors. Z-scores were calculated for the obtained NLR, independently from the other PB2, and used to compare the UPS/PB2 $H1N1_{MZ}$ profile to those of the five PB2 from human-infecting strains (Figure 29b) (Digital supplementary data 4). $H1N1_{MZ}$ PB2 gathered with the other avian-origin $H1N1_{pdm09}$ and H7N9 PB2 proteins, but the latter having recently gained human infection potential remained closer to each other (Figure 29b). It can thus be hypothesised from the PB2-UPS interaction dendrogram that the interplay between the PB2 polymerase protein and the human UPS may have evolved along with virus adaptation to humans, *i.e* acquisition of the capacity to infect human from animal reservoir (H7N9 *vs* $H1N1_{MZ}$), gaining human-to-human transmission capability ($H1N1_{pdm09}$ *vs.* H7N9), then duration of circulation in the human population ($H1N1_{1918}$ < $H1N1_{WSN}$ < H3N2).



Strain	Time of circulation of the PB2 segment in the human population before sampling
Avian $H1N1_{MZ}$	No human infection
H7N9	No circulation in human population
$H1N1_{1918}$	< 1 year
$H1N1_{pdm09}$	< 1 year
$H1N1_{WSN}$	15 years
H3N2	44 years

Figure 29: Interaction dendrograms deduced from the PB2/UPS interaction mapping. **a-** Hierarchical clustering of the 5 indicated PB2 proteins according to their interaction profiles with the human UPS. Heat maps represent the z-scores derived from an experimentally homogenous set of NLR obtained by the retesting of 75 UPS factors at once in NLR retesting (Digital supplementary data 4). **b-** Hierarchical clustering of PB2/UPS interaction profiles including PB2 $H1N1_{MZ}$. **c-** Schematics delineating, for the IAV strains used in this study: the sampling year; the duration (in years) of PB2 segment circulation in the human population (in brackets); the year of PB2 segment introduction in humans (vertical arrows, the grey arrow indicates the infection of a human patient without any spread in the human population); the course of the PB2 segment in the human population (plain lines), the dashed line represents the presence of the PB2 segment originating from the former H1N1 in the H3N2 strain (adapted from [172]).

The PB2 proteins of $H1N1_{pdm09}$ (which contains a PB2 segment of avian origin, the IAV strain has crossed the host barrier to humans, and gained the ability to spread in humans) and H7N9 (avian PB2 segment, the strain crossed the human host barrier species but is unable to sustain human to human transmission) gathered closer in the interaction dendrogram than the purely avian $H1N1_{MZ}$ PB2 (avian PB2 segment, did not cross the human barrier species). At the nucleotide sequence level in contrast, the PB2 segments of the strains that do not spread in humans (H7N9 and $H1N1_{MZ}$) are closer than the PB2 $H1N1_{pdm09}$ that gained the ability to sustains human to human transmission, even though the H7N9 crossed the human species barrier (Figure 30). The UPS/PB2 interaction profiles might thus help predicting the crossing potential of human species and transmission ability of IAV strains in the human population more accurately that vRNA segment sequence analysis. Such comparison-based analysis highlights how the systematic IAV-host comparative profiling can pinpoint to potentially significant biological information relative to IAV infection, barrier species crossing, and viral adaptation, which are to be further explored.

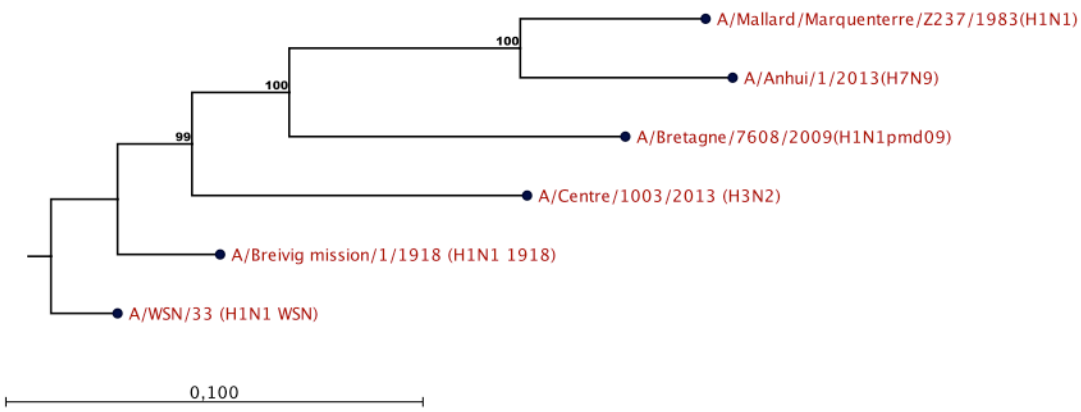


Figure 30: Phylogenetic tree calculated with the nucleotide sequences of the PB2 proteins using neighbour joining algorithm, Jukes-Cantor distance measurement and 1000 replicates for bootstrap calculation.

Combined UPS-library screening increases the robustness of interaction maps

In parallel to our study, the UPS-dedicated library was screened using the same strategy for interactions with human papillomavirus type 16 (HPV16), by our collaborators in Strasbourg. HPV16 is the most prevalent HPV type in HPV-associated cancers. Its oncogenic potential is based on the immortalising activities of the E6 and E7 oncoproteins, which are known to target respectively the tumour suppressors TP53 and pRB to degradation.

Remarkably, neither PB2 proteins nor E6 or E7 were found to target E2 enzymes (Figure 31a) (Annexes 4 and 5). Instead, viral baits bind specifically to E3 Ub ligases, suggesting that the viral proteins induce the UPS remodelling mainly through subversion of E3 ligases, in line with the UPS rewiring by other viruses [69]. Nevertheless, the interplays turn out to be specific to each of the viral protein, as observed from the comparison between HPV16 E6, E7 and the high confidence set of PB2 interactors (Figure 31b).

Only three UPS factors were found to interact with the three viral proteins, PB2, E6 and E7 (Figure 31b) (Annexe 5): TRAF5, TRAF3IP2 and UBB. The last two also exhibited strong GPCA signal with the Gluc2 fragment alone (Annexe 1), reflecting their pro-aggregating trait. The other selected UPS factors that are also GLuc2 outliers are binding to only one of the viral protein, indicating that despite an elevated level of aspecific binding, these factors still exhibit binding specificity toward viral proteins. It can also be noticed that the UPS factors generating outliers luciferase values against empty Gluc2, were not particularly recovered as viral protein interactors. Indeed, 9 of the 23 Gluc2 outliers are found among the viral proteins interactors, and they interact differently with the different proteins (Annexe 1). Therefore, UPS factors that are Gluc2 outliers do not necessarily generate false positive interactions, but their strong aspecific interaction properties have to be taken in consideration for further analysis.

Most of the UPS factors identified were specific to only one viral bait, with 7 UPS factors being exclusively targeted by E6, 26 UPS factors by E7 and 32 UPS factors by the PB2 proteins. It is important to stress that because specific sets of UPS factors emerge from the combined screening with three viral proteins, each exhibiting a high level of interaction with the host proteome, it strongly enhances the robustness of our interactomics data, and

demonstrates both the robustness and the stringency of the HT-GPCA based interactomic strategy.

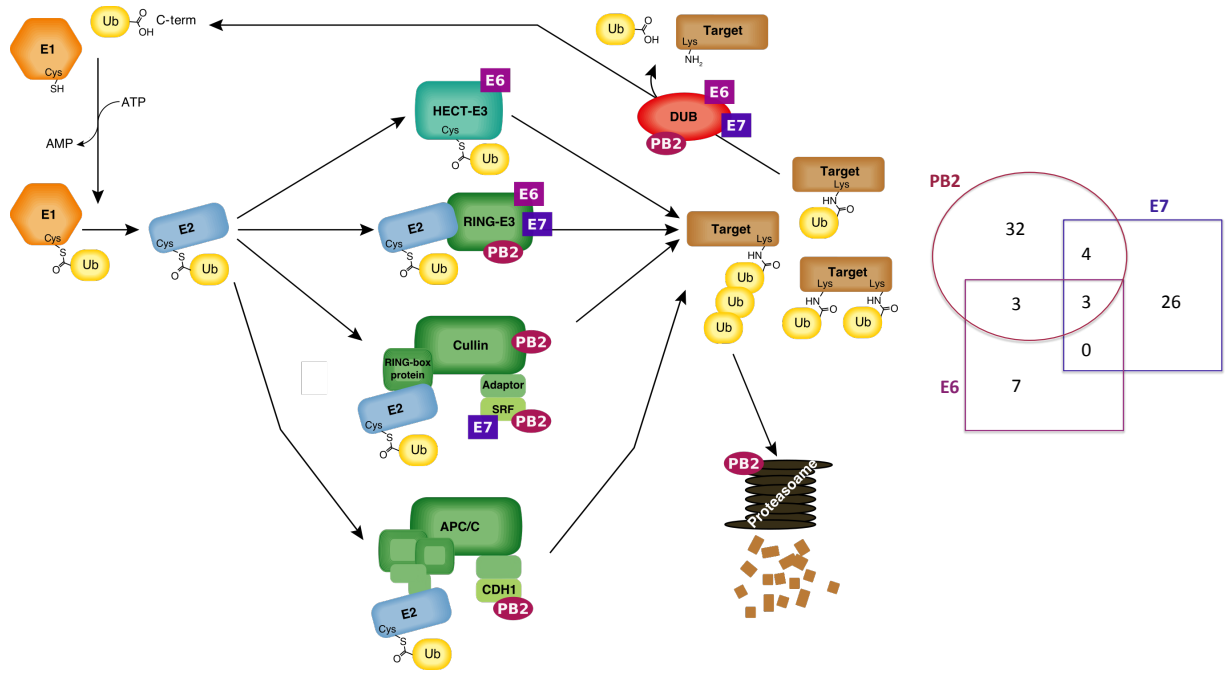


Figure 31: Comparison of the targeting of the ubiquitin proteasome pathway by the influenza A virus PB2 polymerase subunit and the human papillomaviruses HPV16 E6 and E7 oncoproteins. **a-** E6 and E7 interact with some E3 ubiquitination complexes, whereas the PB2 proteins have a larger spectrum of interactions. **b-** Venn diagram illustrating the number of UPS factors found to interact with IAV PB2 proteins, HPV16 E6 and E7 proteins.

Functional impact of PB2 interactors

The involvement of the UPS targets of PB2 in the production of infectious viral particle was estimated to assess the biological relevance of identified PB2/UPS interplays. For these initial, unrefined functional studies, only the top-confident set of UPS interactors were selected, *i.e.* those that were validated 3 times out of 3 experiments in NLR retesting. We examined the productive infectious cycle of the seasonal strains $H1N1_{pdm09}$ and H3N2 viruses and of the $H1N1_{WSN}$ laboratory-adapted strain upon siRNA-mediated depletion of the individual PB2 interactors. The seasonal strains $H1N1_{pdm09}$ and H3N2 grow poorly in the human pulmonary A549 cell line. We thus produced adapted $H1N1_{pdm09}$ and H3N2 viruses by reverse genetics, that harboured point mutations in the HA segment and an intact PB2 segment and efficiently infected A549 cells (see *How to finally tame the seasonal strains*

p73). Forty-one high-confidence PB2 interactors were assessed, while the ubiquitin (UBB) protein was excluded owing to its pleiotropic function and high aggregation properties. A549 cells transfected with siRNAs targeting the individual UPS factors, or transfected with a non-target siRNA, were infected for 24 hours with $H1N1_{WSN}$, adapted $H1N1_{pdm09}$ or H3N2 viruses.

The titre of infectious IAV particles produced was measured by plaque assays in the highly IAV sensitive MDCK cells. The effects of UPS factors depletion were calculated by comparing the virus titres produced with those obtained with non-target siRNA. Of the 41 PB2 interactors tested, 36 showed an effect on the infection with all or a subset of IAV strains, albeit at variable degrees. A statistically significant ($P<0.05$) reduction of viral production was observed with the depletion of respectively 13, 26 and 31 factors for H3N2, $H1N1_{pdm09}$ and $H1N1_{WSN}$. The depletion of most corresponding UPS factors resulted in a decrease in virus production, suggesting that they are beneficial for IAV viruses. The decrease ranged between 2.3 and 4.8 fold for H3N2, 1.7 and 7 fold for $H1N1_{pdm09}$ and 2.3 and 50 fold for $H1N1_{WSN}$ (Figure 32) (Annexe 6). The variable effects of the siRNAs suggest that the identified UPS factors may have critical roles or modulate IAV infection. Only one factor induced an enhanced viral production upon depletion in a statistically significant manner, FBXL19, increasing the viral production only for H3N2 infection.

We noticed a globally stronger functional dependency towards the UPS proteins interacting with PB2 of the H1N1 strains pdm09 and WSN, which may reflect that they more optimally exploit the UPS. It could alternatively result from a lessened efficiency in our experimental settings to detect the H3N2 strain sensitivity to siRNA-mediated cell factors depletion *in vitro*.

Knockdown efficiency and cell viability were assessed for the different siRNAs (Figure 33), no correlation was noticed between the slight loss of cell viability observed for several siRNA and the functional effect of these siRNA on the virus cycles of the three tested strains.

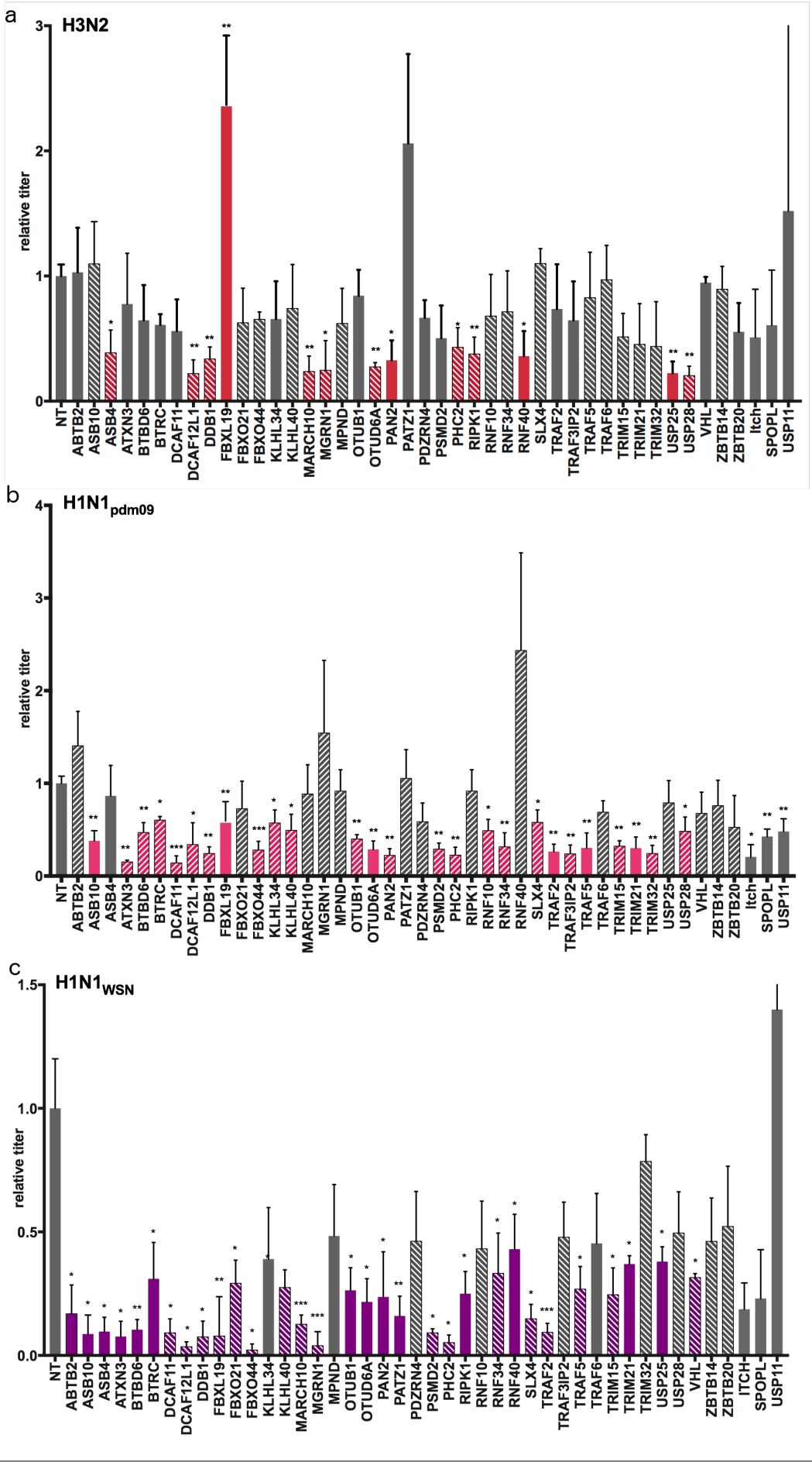

a
H3N2
relative titer
b
H1N1pdm09
relative titer
c
H1N1WSN
relative titer
NT
ABTB2
ASB10
ASB4
ATXN3
BTBD6
BTRC
DCAF11
DCAF12L1
DDB1
FBXL19
FBXO21
FBXO44
KLHL34
KLHL40
MARCH10
MGRN1
MPND
OTUB1
OTUD6A
PAN2
PATZ1
PDZRN4
PSMD2
PHC2
RIPK1
RNF10
RNF34
RNF40
SLX4
TRAF2
TRAF3IP2
TRAF5
TRAF6
TRIM15
TRIM21
TRIM32
USP25
USP28
VHL
ZBTB14
ZBTB20
Itch
SPOPL
USP11
ITCH

Figure 32: Functional evaluation of PB2 UPS partner involvement in the multiplication of H3N2, $H1N1_{pdm09}$ and $H1N1_{WSN}$. A549 cells transfected by Non-Target (NT) or UPS targeting siRNA for 48 hours were infected with $H1N1_{pdm09}$ (**a**) or H3N2 (**b**) at a MOI of 10^{-3} and with $H1N1_{WSN}$ (**c**) at a MOI of 10^{-4}. The titres of produced viruses were determined in the supernatants collected 24 hours post infection by plaque forming assay, and expressed as ratio relative to the titres obtained with non-target siRNA. UPS factors with statistically significant effects were labelled in red (H3N2), pink ($H1N1_{pdm09}$) or purple ($H1N1_{WSN}$), hatched bars indicate that PPI has been detected with the corresponding PB2 protein, taking into account positive scoring of at least two times in the NRL retesting experiments, plain bars that no interaction was detected (validated less than two times in NLR re-testing). Data represent means ±s.e.m. of three independent experiments, p values were calculated with a Wilcoxon test, *<0.05, **<0.01, ***<0.001.

Six UPS factors were involved in the infection of the 3 IAV strains tested (Figure 34), while the remaining 30 were differentially modulating infection according to IAV strains (Annexe 6), revealing the diversity of the functional outcome of the identified UPS factors. Most of the UPS factors that had an effect on viral production were also among those found to bind the PB2 protein of the corresponding strain. This correlation is enhanced when taking into account interactions positive twice in NLR retesting, again advocating for the overstatement of the 3 times positive scoring for considering validated interactions. In rare occasions however, an UPS factor found implicated in IAV infection by siRNA experiment did not consistently interact with the corresponding PB2 protein (scoring positive only once in NLR retesting). This was the case for ASB4 and ATXN3 for example, which affected the production of $H1N1_{WSN}$ viral particle upon depletion but were not validated as interacting partners with the PB2 $H1N1_{WSN}$ protein in NLR retesting. The same applied for TRIM21, which impacted $H1N1_{pmd09}$ and H3N2 infection but was validated for interaction only with only the H3N2 PB2 protein. These factors may be involved in the viral life cycle independently of the binding of the PB2 protein, maybe through a functional interaction with another component of the vRNP. In any case, the correlation between functional impact and PB2 interaction does not necessarily mean that the role of the UPS factor in IAV infection is mediated through binding to PB2. The demonstration of the direct role of the PB2/UPS interactions will require studying mutations that disrupt the PB2/UPS interactions.

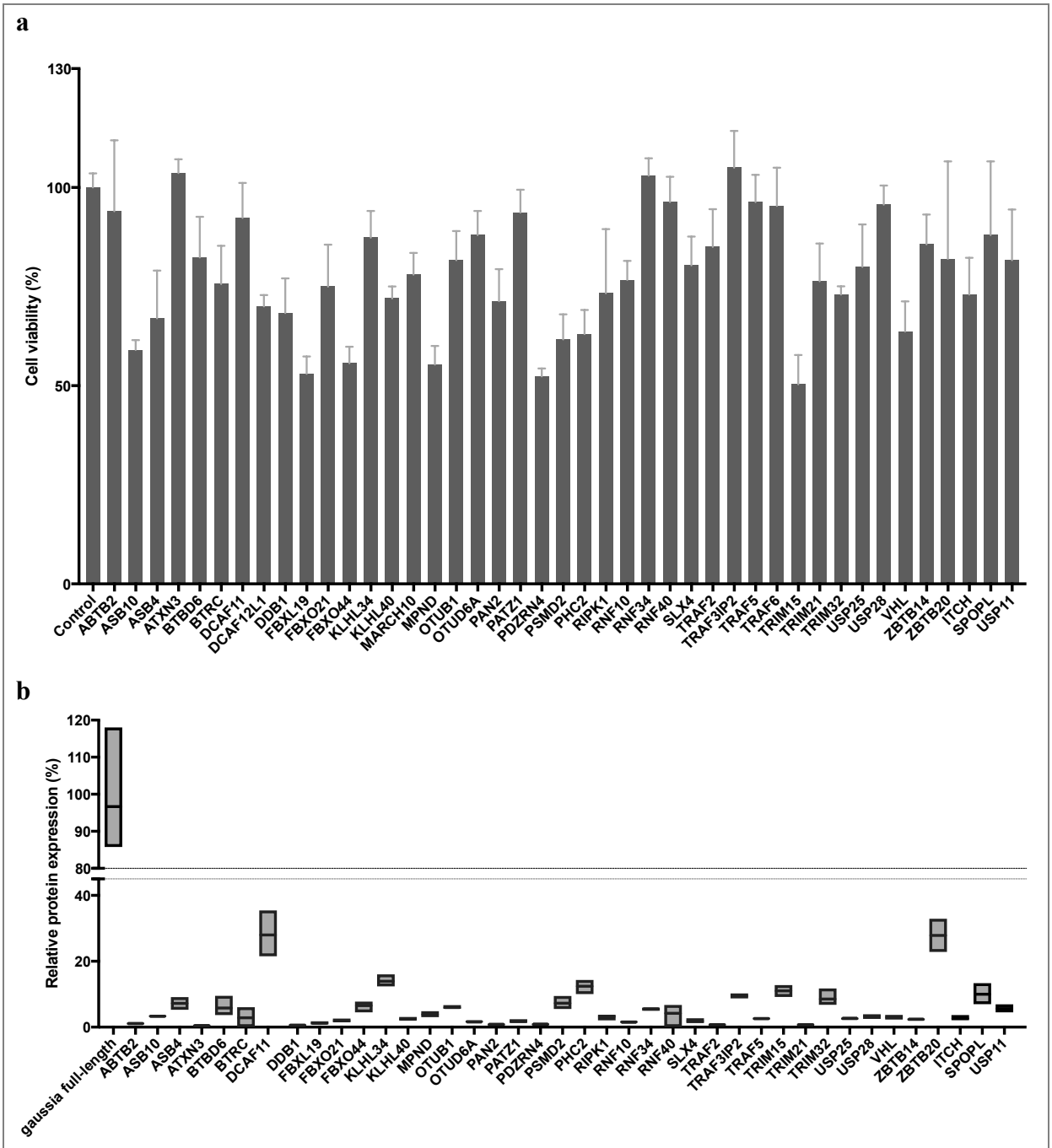


Figure 33: Toxicity and silencing efficiency of siRNAs. **a-** Toxicity of siRNA. A549 cells were transfected with 25 nM of siRNA and cell viability was determined at 72 hours post transfection using trypan blue. The results are expressed as the mean percentages ± SEM (n=3). **b-** Silencing efficiency of siRNA. A549 cells were transfected with 25 nM of control or UPS-targeting siRNAs and with plasmids encoding the corresponding UPS protein fused with the full-length *Gaussia princeps* luciferase (pGlucFL-UPS). Ratios of the luciferase activities obtained in cells transfected with the UPS-targeting siRNA to the one obtained in cells transfected with the control siRNA are shown. The results are represented as floating bars with a line at the mean.

The six UPS factors discovered to be essential for infection with the three viruses consist of members of the cullin-based E3 ligase complexes (DCAF12L, DDB1, FBXL19), a component of the H2A histone Ubiquitin ligase complex (PHC2, 24), and DUBs (OTUD6A, PAN2). These functional studies also identified three factors (TRIM32, TRAF3-IP2, KHLH34) that specifically modulate infection with the $H1N1_{pdm09}$ virus from the 2009 pandemic (Figure 34).

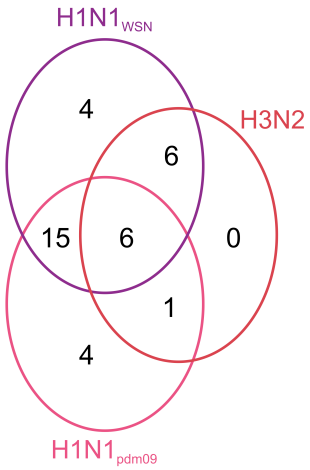


Figure 34: Venn diagram illustrating the number of UPS factors found to be involved in the multiplication of one or several IAV viruses.

The UPS factors that were found to mediate $H1N1_{WSN}$ infection were ectopically expressed fused to the mCherry protein, to evaluate their over-expression effect on infection with a recombinant $H1N1_{WSN}$ virus expressing the mCitrine fluorescent reporter (Figure 35). Twelve factors increased viral replication, as measured by the percentage of mCitrine-positive cells (Figure 35) in the cherry-UPS expressing cells, further validating their role in IAV infection. Similar analysis could not have been done with the adapted $H1N1_{pdm09}$ and H3N2 viruses, because the corresponding recombinant mCitrine reporter viruses could not be produced.

In all, the multi-strains functional validation highlighted an unexpected diversity in the functional consequences of the PB2-mediated UPS hijacking, in particular between the seasonal $H1N1_{pdm09}$ and H3N2 strains (Figure 34). Indeed, a number of identified UPS factors turned out to have strain-specific effects on IAV infection, while 6 UPS factors were shown to be involved in the life cycle of the three strains tested.

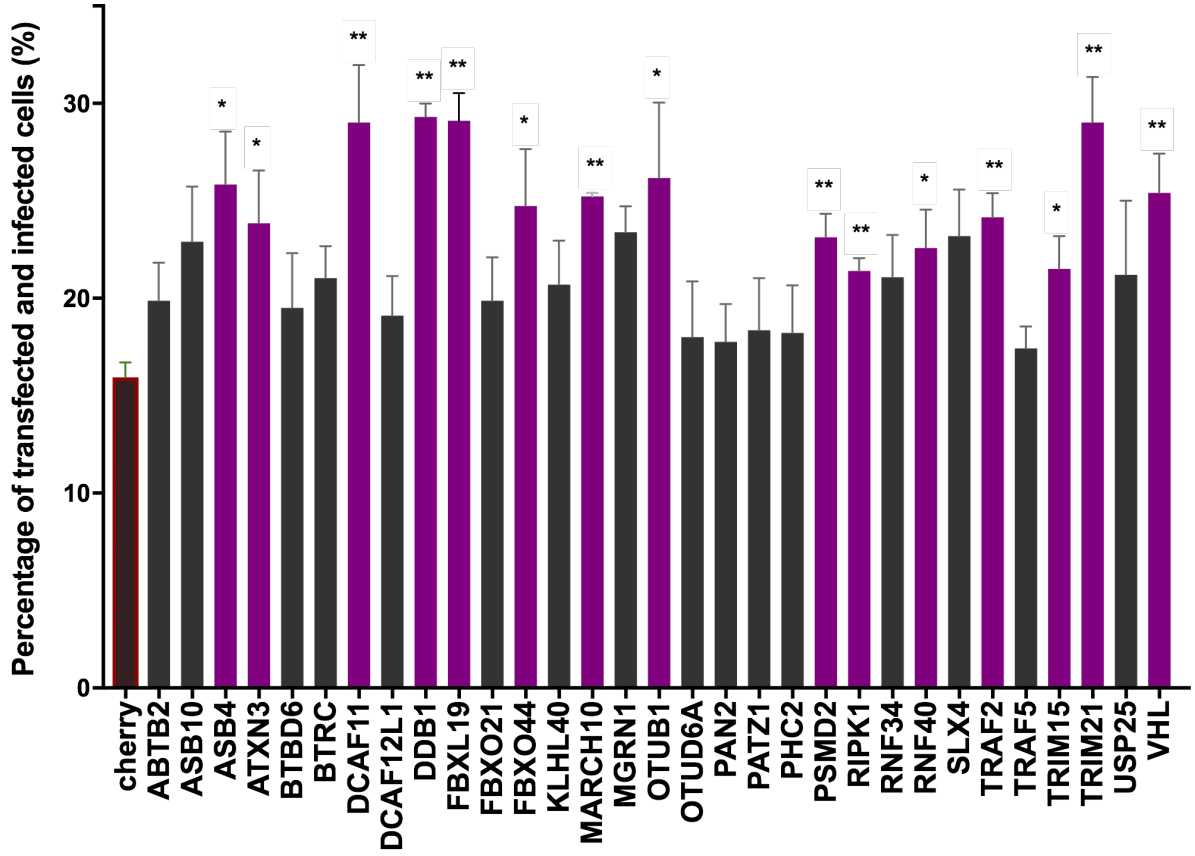


Figure 35: Effect of UPS factors over-expression. Plasmids expressing the UPS factors fused to mCherry were transfected in HEK-293T cells. Twenty-four hours after transfection, the cells were infected at a MOI of 1 with a reporter $H1N1_{WSN}$ virus expressing the mCitrine fluorescent protein for 18 hours. The percentage of mCitrine-positive infected cells in Cherry-UPS expressing cells was determined by FACS analysis. Data represent means ±s.e.m. of three independent experiments, p values were calculated with a Wilcoxon test, *<0.05, **<0.01, ***<0.001. Purple bars indicate factors with a statistical difference in the percentage of transfected and infected cells compare to the control.

How to finally tame the seasonal strains

Most of the experiments with IAV presented in the literature were performed with laboratory-adapted strains, as the seasonal strains are difficult to work with owing to their poor growth in cell lines. For my PhDs project, based on the comparison of different strains, it was decisive to conduct functional studies with the seasonal strains $H1N1_{pdm09}$ and H3N2 in parallel to the laboratory adapted $H1N1_{WSN}$ to assess the functional relevance of identified UPS/PB2 interplays at least for the three strains manipulated in BSL2+ facilities. The functional validation of highly pathogenic strains, requiring BL3 facilities will possibly be conducted later. Our initial infection experiments in the A549 human pulmonary cell line using $H1N1_{pdm09}$ and H3N2 viruses, obtained by reverse genetics and derived from biological

samples, showed a so poor infection rate that *in vitro* studies in A549 cells were not possible. Moreover, this was obligatory since human siRNA cannot be used in the only cell line where these seasonal viruses grow efficiently: MCDK, which are canine cells. We therefore took advantage of the work performed earlier in the laboratory by Cyril Barbezange, who obtained a 2 log increase in the viral titre of $H1N1_{pdm09}$ and H3N2 viruses originating from patient swabs, after 5 passages on A549 cells at a controlled MOI (Figure 36a-b). We sequenced the HA and NA segments of the $H1N1_{pdm09}$ and H3N2 viruses efficiently replicating on A549 (stock and 5th passage) and identified mutations in the HA segments, that we suspected to be responsible for viruses adaptation to A549 cells. Surprisingly, those mutations are not positioned in the receptor binding domain of the HA.

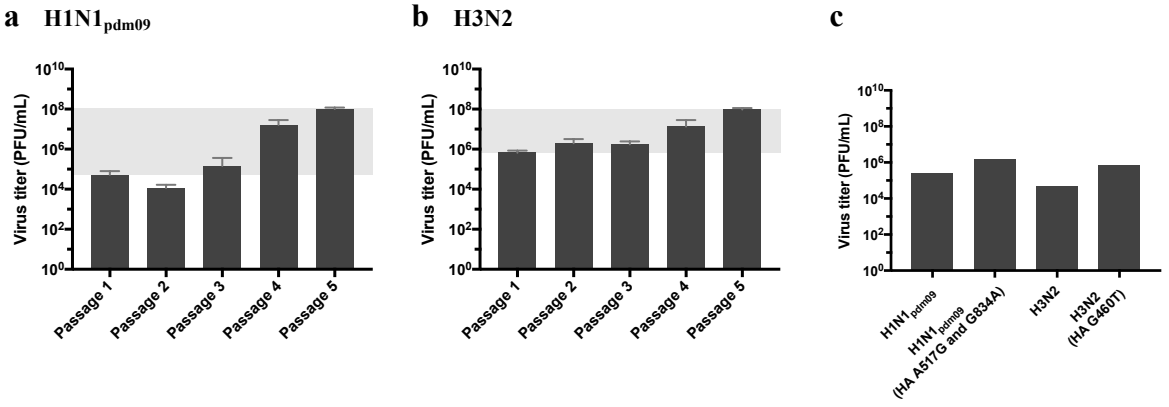


Figure 36: Adaptation of $H1N1_{pdm09}$ and H3N2 viruses to A549 pulmonary cells by serial passages. A549 cell were infected with swab-derived $H1N1_{pdm09}$ (**a**) at a MOI of 10^{-1} or H3N2 (**b**) at a MOI of 10^{-2}. The titres of produced viruses were determined in the supernatants collected 72 hours post infection by plaque forming assay. **c-** Identified mutations in the HA segment were introduced in the reverse genetics system and produced mutated viruses were compared to wild types for infection of A549 cells, at a MOI of 10^{-2}. The titres of produced viruses were determined in the supernatants collected 72 hours post infection by plaque forming assay.

We introduced these mutations in the respective reverse genetics systems for $H1N1_{pdm09}$ and H3N2, and amplified the recovered HA mutated viruses in A549 cells. The HA-mutated viruses gained more than one log titre when produced in A549 cells compared to the non-mutated reverse genetics-produced viruses (Figure 36c). The adapted $H1N1_{pdm09}$ and H3N2 obtained by reverse genetics could be produced at sufficiently high titres (around 10^6 pfu/mL) to be used in siRNA experiments in A549 cell lines, contrary to the non-mutated viruses. They were used for the functional validation of the UPS/PB2 interplay (see *Functional impact of PB2 interactors* p67).

Interestingly, the adapted $H1N1_{pdm09}$ and H3N2 viruses generated by reverse genetics never reached the titres of the adapted corresponding viruses obtained from biological samples (patients swab). This almost 2 log titre difference between adapted seasonal viruses originating from swabs and from reverse genetics system is likely due to different initial virus populations: the population for viruses originating from the reverse genetics are likely to be more homogenous, compared to the natural cloud of viral quasi-species from patient swab. The importance of quasi-species for IAV fitness has been documented, and we have here an illustration of this phenomenon. The under representation of quasi-species in the reverse genetics system could represent a limitation for IAV studies that are to be taken in consideration. However, reverse genetics allows the introduction of mutations and the production of reporter viruses, which both make this system compulsory for IAV infection studies [173].

Deubiquitinases and influenza A virus infection

The screening of the UPS-dedicated library identified interactors of the PB2 proteins belonging to different families of the ubiquitin pathway, mainly proteins of E3 ubiquitin ligases complexes but also 7 deubiquitinases (DUB) (Figure 37a). In parallel, our collaborators in Hong Kong found that six DUBs were activated during influenza A infection (unpublished data) using an activity-based ubiquitin-specific protease (USP) profiling [174]. Among these DUBs, two were part of the high-confidence set of PB2 interactors. We focused on the role of the 11 DUBs that were found to interact with PB2 and/or be activated during influenza A infection.

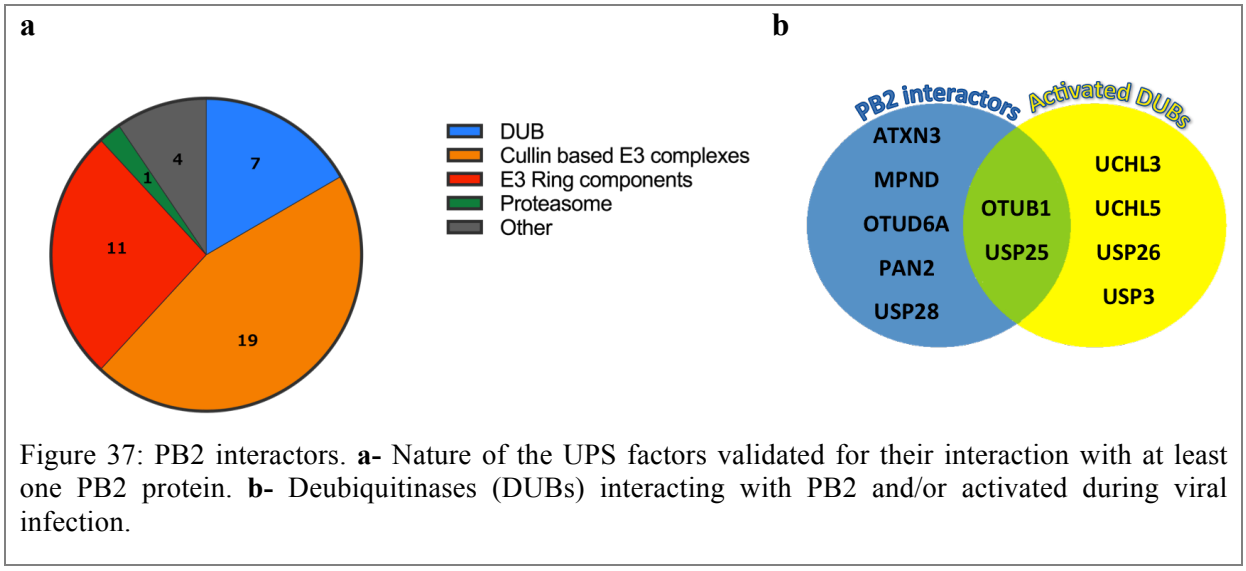


Figure 37: PB2 interactors. **a-** Nature of the UPS factors validated for their interaction with at least one PB2 protein. **b-** Deubiquitinases (DUBs) interacting with PB2 and/or activated during viral infection.

Functional exploration of the DUBs

In order to study the role of the DUBs in the IAV viral cycle, we conducted experiments where A549 cells were treated with DUB siRNA, and then infected with at a low MOI of a $H1N1_{WSN}$ Nano-luciferase reporter virus for 24 hours (multi-cycle settings) [175]. The reporter viruses contained a PB2 segment with the Nano-luciferase ORF preceded by a FMDV 2A cleavage site (Figure 38a). Translation of the mRNA produced from the modified PB2 segment leads to the production of the PB2 protein and of a Nano-luciferase released co-translationally from the PB2 polypeptide by cleavage of the FMDV A2 site. In this way, the

function of the PB2 protein is not affected and accordingly the $H1N1_{WSN}$ PB2-nanoluc recombinant viruses are only marginally attenuated compared to wild type $H1N1_{WSN}$ [175]. The production of viral mRNA from the PB2 segment is measured through the detection of the Nano-luciferase activity. Twenty-four hours post-infection with the $H1N1_{WSN}$-Nanoluc reporter virus, the production of viral particles was measured by taking the supernatant to infect non-treated A549 cells for 6 hours (single viral cycle) then reading the luciferase signal (Figure 38b). The detection of Nano-luciferase activity reflects the number of PB2-Nanoluc segments from infecting virus, *i.e.* the amount of infectious viral particles produced by the A549 cells treated or not with DUBs siRNA.

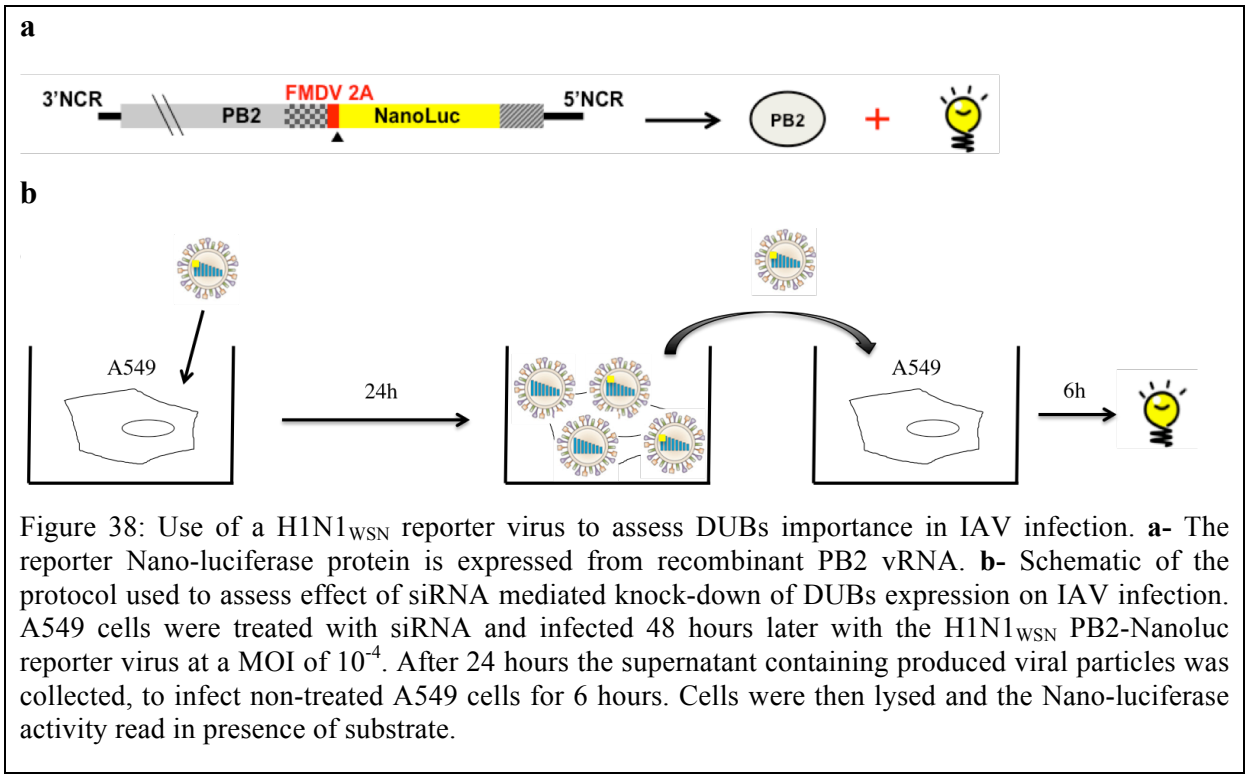


Figure 38: Use of a $H1N1_{WSN}$ reporter virus to assess DUBs importance in IAV infection. **a-** The reporter Nano-luciferase protein is expressed from recombinant PB2 vRNA. **b-** Schematic of the protocol used to assess effect of siRNA mediated knock-down of DUBs expression on IAV infection. A549 cells were treated with siRNA and infected 48 hours later with the $H1N1_{WSN}$ PB2-Nanoluc reporter virus at a MOI of 10^{-4}. After 24 hours the supernatant containing produced viral particles was collected, to infect non-treated A549 cells for 6 hours. Cells were then lysed and the Nano-luciferase activity read in presence of substrate.

A decrease in viral particles production could be detected upon the depletion of all of the DUB, albeit to different levels (Figure 39). This is not surprising considering that the selection of these DUBs was initially based on their interaction with PB2 and/or their activation during infection. Interestingly all DUBs with a functional impact seemed to favour infection, as their depletion decreased the viral detection. Three of them decreased the luminescence activity by more than 75%: OTUD6A and OTUB1, which are both active deubiquitinases; and PAN2, a pseudo-DUB with an inactive catalytic site, which is part of the poly(A)-nuclease (PAN) deadenylation complex. For simplification, we will still name it as a member of the DUB family in the following experiments. Of note, the depletion of two of

these DUBs, MPND and UPS28 led to a decrease in $H1N1_{WSN}$ production that was not statistically significant when measured by plaque forming assay (Figure 32), showing that the Nano-luciferase reporter system might be more sensitive.

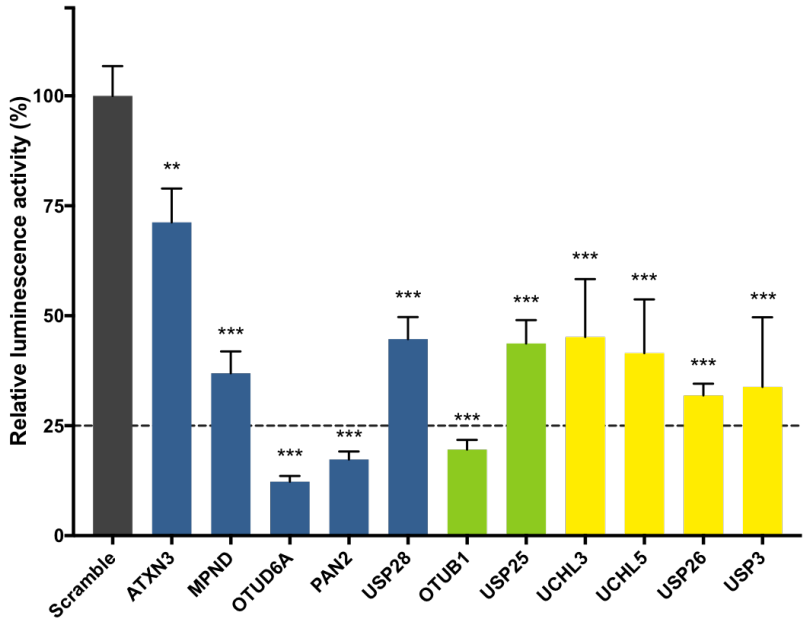


Figure 39: Effect of DUBs siRNA knock-down (KD) on $H1N1_{WSN}$ infection using a luminescent reporter virus. The expression of the Nano-luciferase reflecting the titre of produced viruses is measured 24 hours post infection in the different siRNA DUB KD A549 cells. DUBs found to interact with PB2 are coloured in blue, the DUBs found to interact with PB2 and activated during viral infection are in green, the DUBs only found to be activated during infection are in yellow. Data represent means ±s.e.m. of three independent experiments, p values were calculated with a Wilcoxon test, *<0.05, **<0.01, ***<0.001.

We then assessed the interaction map of the two PB2-interacting DUBs with the highest deletion impact on infection (OTUD6A and PAN2) (Figure 39) and of the DUBs activated during infection, with the major proteins of $H1N1_{WSN}$: PB1, PB2, PA, NP, M1, NS1 and NS2, HA and NA. The position of the Gluc fragment fusion on each viral protein was chosen from previous experiments performed in the lab to better conserve the viral polymerase activity in a mini-replicon assay (data not shown), or according to the described structure or known interactions (Table 4). Because for most of the viral proteins, fusion with the Gluc1 fragment was more appropriate, all DUBs were switched to a Gluc2 fusion, leading to an inverted configuration compared to the initial DUB/PB2 PPI screening.

Table 4: Fusion position of the Gluc1 tag on the viral proteins for the HT-GPCA screening for interactions with the host deubiquitinases.

Viral protein	PB1	PB2	PA	NP	M1	NS1	NS2	HA	NA
Gluc1 tag fusion position	C-terminal	C-terminal	C-terminal	C-terminal	N-terminal	C-terminal	C-terminal	C-terminal	C-terminal

For these experiment, a double threshold was used for PPI selection. Indeed, both the viral proteins and the DUBs were tested against the random reference set (RRS) to determine the threshold of positive interactions for each partner, as described in chapter 2 (upper limit of the confidence interval). In each experiment the interaction was selected if the NLR of the tested proteins was higher than the individual thresholds of each DUB and viral partner (Figure 40). This method allowed taking into account the stickiness of both the viral protein and of the tested DUBs, evaluated with the RRS, to define the interacting pairs. For the NA and HA capsid proteins, the RRS generated highest luciferase values than the DUBs in two independent experiments, leading to negative interactions for all tested DUBs. It is not clear at present whether these results reflect that the HT-GPCA assay used is not appropriate to detect PPI with these two membrane proteins, or whether it reflects real negative interactions. The HA and NA proteins were consequently removed from further PPI analysis.

The NLR experiment was repeated 3 times for the viral internal proteins, and an interaction was considered as validated if scoring 3 times positive with the double threshold method (Table 5). With these selection criteria, most of the interactions with PB2 were recovered apart for OTUB1, and PAN2 detected 2 times over 3. On the contrary USP26, which interact with PB2 in the present screening, had not been validated with any strain in the initial PB2/UPS screening (Annexe 3). We have shown that the nature of the tag (Gluc1 or Gluc2 in the HT-GPCA assay) and its position (N or C-terminal) can modify the detection and/or stability of an interaction (Figure 24 and data not shown). Therefore, the configuration change in the present HT-GPCA assay, where the PB2 proteins are fused in C-terminus with the Gluc1 fragment can explain why some interactions between the DUBs and PB2 have not been recovered.

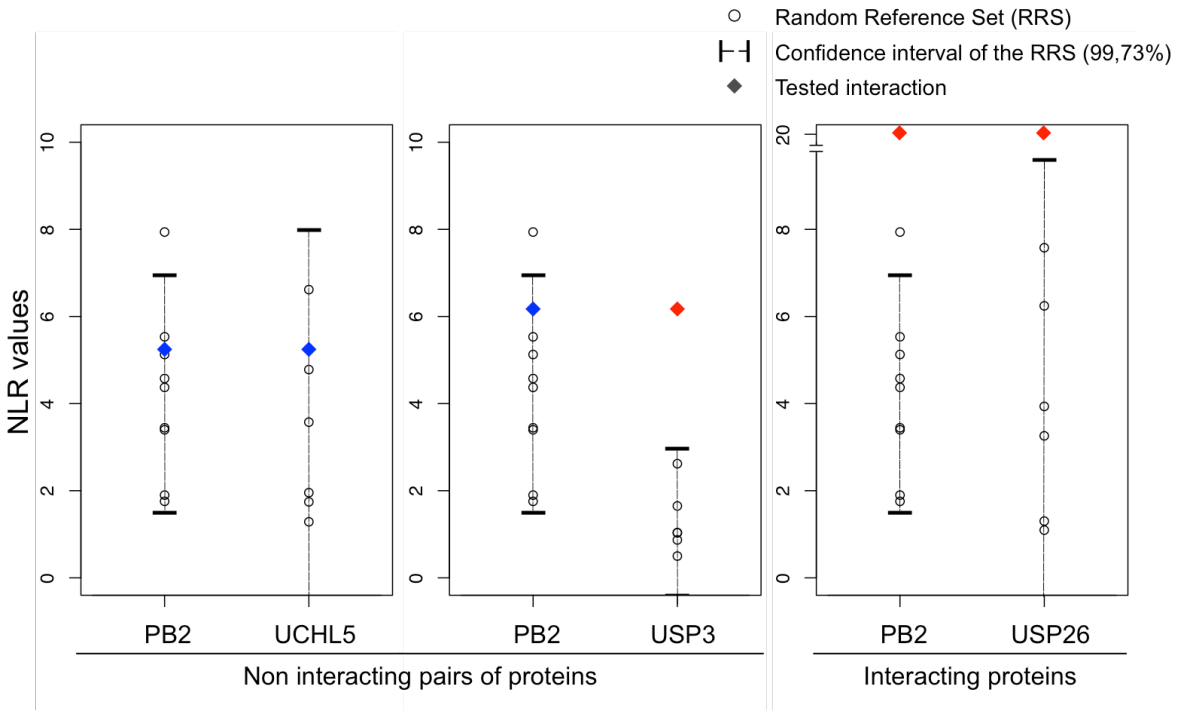


Figure 40: Principle of the double threshold applied for the interaction screening of the DUBs and the internal proteins of the IAV. Each protein of interest (DUBs and viral proteins) is tested against a random reference set and a threshold calculated according to the upper limit of à 99,73% confidence interval. The interaction was considered positive when the NLR of the protein pair was higher than the two individual thresholds. In this example, PB2 was considered interacting with USP26 but not with UCHL5 or USP3.

Table 5: Redundant targeting of studied DUBs by H1N1$_{WSN}$ viral factors. HEK-293T cells were co-transfected with plasmids encoding the Gluc2-DUB and Gluc1-viral protein. The NLR were calculated for each interaction and compared to the NLR given by the interaction of each partner with the RRS using 99,73% confidence interval. Numbers indicate how many times, out of the 3 experiments, the interaction was detected above the threshold of both partners (double threshold). DUBs found to interact with PB2 in the initial screen are coloured in blue, the DUBs found to interact with PB2 and activated during viral infection are in green, the DUBs only found to be activated during infection are in yellow.

	OTUD6A	PAN2	OTUB1	USP25	UCHL5	UCHL3	USP26	USP3
PB1				2	3			1
PB2	3	2		3			3	
PA					1	1	1	
NP					2	1	3	3
M1					2		1	3
NS1	3	2	3	3	3	3	3	3
NS2	3	1		3		1		

In addition, the interaction between PB2 and OTUB1 or PAN2 has been detected by a classical co-immunoprecipitation (Figure 41). This corroborates that an interaction exists between PB2 and these two DUBs, suggesting that detection of these PPI by HT-GPCA did not work in the present Gluc configuration.

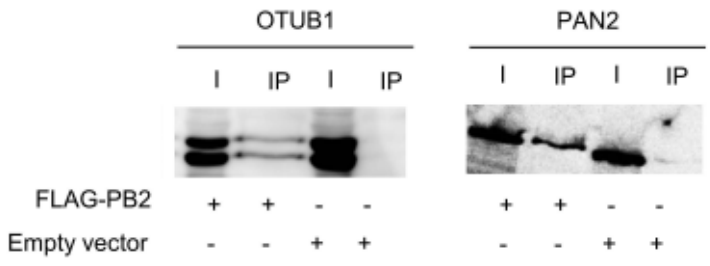


Figure 41: Co-immunoprecipitation of OTUB1 and PAN2 with PB2. HEK-293T cells were co-transfected with a plasmid encoding UPS protein fused to Gluc1 fragment and the plasmid pCineo3xFLAG-PB2 or the empty vector. The Gluc1-fused DUBs proteins immunoprecipitated on ANTI-FLAG magnetic beads (IP) and present in the input (I) were analysed by gel electrophoresis and Western-blotting using rabbit anti-Gaussia antibodies.

It is noticeable that the PA subunit of the viral polymerase did not interact with any of the DUB, while in the opposite the NS1 protein seems to interact with almost every DUBs. While such high degree of NS1 interaction questions whether NS1 does not engage aspecific interaction with many cell factors, the NLR obtained with the RRS are not particularly high with NS1 compared to the other viral proteins and significantly they are consistently below those obtained with the DUBs. We thus believe that the NS1/DUB PPI detected by the HT-GPCA assay are relevant and reflect an extended interplay between the NS1 protein and DUBs. NS1 thus seems to be pleiotropic regarding interactions with DUBs. To confirm the specificity of these interactions it would be necessary to assess the interaction of NS1 with proteins of other pathways.

On the DUBs side, some exhibit high aspecific binding, which was detected with the RRS and was taken into account thanks to the double threshold strategy. This observation underlies the need to perform RRS analysis for both partners when exploring in depth small-scale interaction maps, which is not feasible in a high throughput format.

A strikingly redundant targeting by viral proteins could be detected for the positive DUBs (Table 5), suggesting that these DUBs are regulated by the virus at multiple levels. We cannot rule out that some VP/DUBs interactions have been missed, which would have been detected

upon different configuration of the Gluc fusions for HT-GPCA assay. The redundant targeting of the DUB deduced from this interaction map may consequently be underestimated. It is interesting to note that among the DUBs bound by IAV proteins, only USP26 emerged as a positive interactor of an HPV16 protein, the E7 protein (Annexe 4), first corroborating the accuracy of the interactions detected with IAV proteins, and second indicating that the redundant targeting of this set of DUBs seems specific to IAV viruses.

OTUD6A can modulate viral polymerase activity in a mini-replicon assay context

We evaluated the effect of the three DUBs which depletion affected the most the viral production (OTUB1, OTUD6A and PAN2) on the viral polymerase activity in A549 cells either over-expressing or depleted for the DUB. In over-expression experiments, cells were transfected with the viral replication proteins PB1, PB2, PA and NP, a luciferase reporter gene under the polymerase control (mini-replicon), and a CMV-driven renilla expression plasmid as a transfection control. Twenty-four hours post-infection the luciferase and renilla activities were measured and the polymerase activity was calculated relatively to the transfection efficiency.

To assess the effect of DUBs over expression, strep-tagged DUBs expression vectors were added in the mini-replicon assay. A plasmid expressing the N-terminal part of the PB2 protein was taken as a control, as it was published to decrease the viral polymerase activity by competing with the full-length PB2 for interaction with PB1, thereby inhibiting the formation of an active polymerase [176]. The expression of strep-tagged RRS proteins was used to estimate replication activity without DUB expression. In such conditions, we could observe that ectopic expression of PAN2 did not alter the activity of the polymerase (Figure 42a). In contrast, the overexpression of OTUD6A has a strong interfering effect on the viral polymerase activity, and this latter is also reduced by OTUB1 expression to a lesser extent.

However, in mini-replicon assays performed in siRNA-treated A549 cells, only the depletion of OTUD6A led to a significant increase in polymerase activity (Figure 42b). We conclude that the OTUD6A DUB impacts on viral polymerase transcription and/or replication activities, whereas OTUB1 does not prove relevant to polymerase activities.

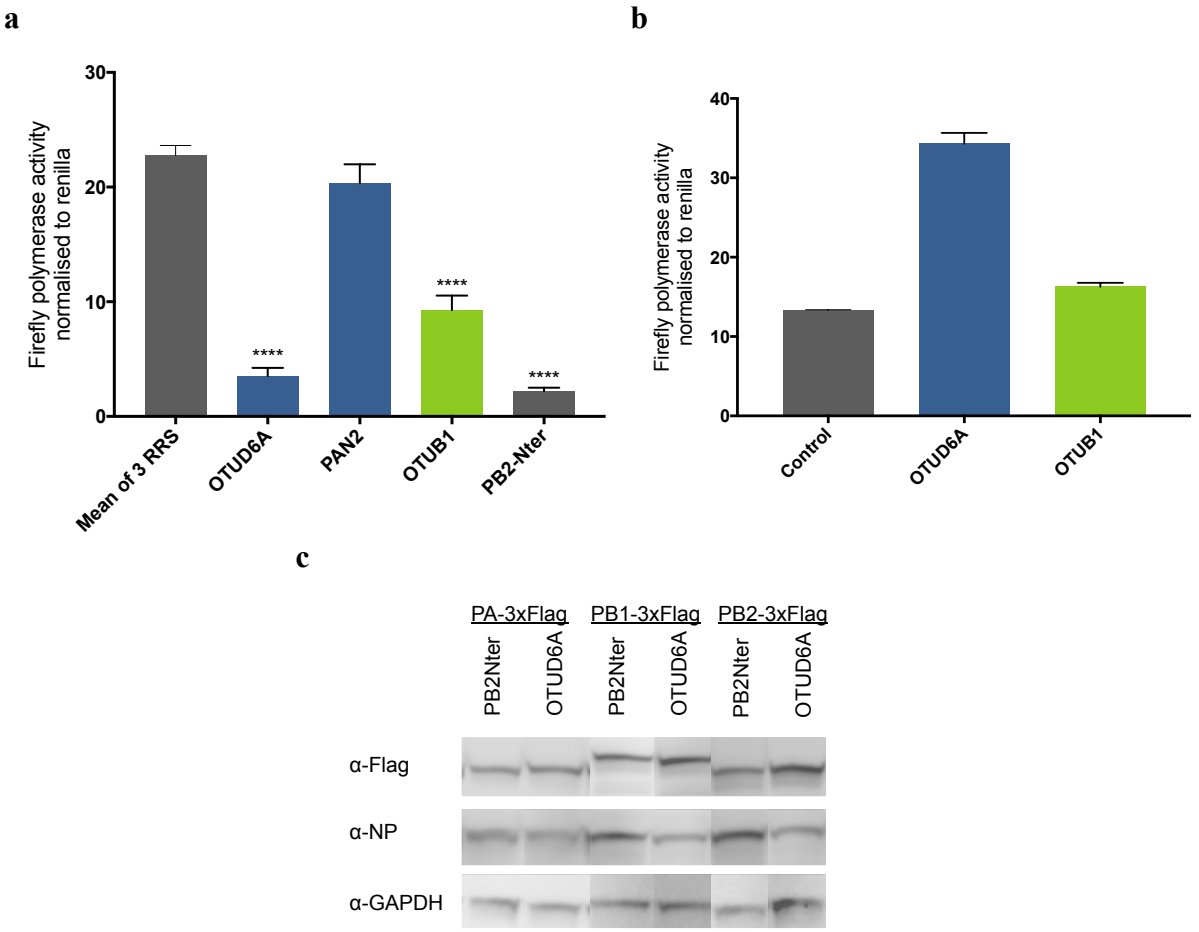


Figure 42: Effect of OTUB1, OTUD6A and PAN2 on viral polymerase activity. **a-** A549 cells were transfected with plasmids expressing the DUB of interest, the viral polymerase complex, NP, a luciferase reporter gene under the control of the viral polymerase and a renilla expressing plasmid as a transfection control. After 24 hours the luciferase and the renilla activity were measured, and the polymerase activity calculated according to the transfection efficiency. Data represent means ±s.e.m. of three independent experiments, p-values were calculated with a Wilcoxon test, *<0.05, **<0.01, ***<0.001, ****<0.0001. **b-** The same assay was performed in siRNA treated A549, in which the expression of the DUB is then depleted. Data represent means ±s.e.m. of three samples of the same experiment. **c-** Polymerase complex protein expression level was assessed in the same context as the first assay (**a**), but with co-transfecting a plasmid expressing one of the polymerase protein fused to a flag-tagged. 24 hours post-transfection the cells were lysed in laemmli buffer and the samples loaded for western blot anti-flag analysis.

We then assessed whether the effect of OTUD6A on polymerase activity was due to a modified expression of the viral replication proteins. We over-expressed OTUD6A as in the mini-replicon assay, including each time one of the replication protein expressed fused to a flag tag. The expression levels of the tagged proteins were assessed by western-blot (Figure 42c). OTUD6A expression did not significantly modify the expression levels of the replication proteins, leading us to conclude that the effect of OTUD6A on the polymerase activity was not a consequence of altered viral proteins stability.

PAN2 is involved in early stages of infection

Single cycle experiments were conducted to determine if the early stages of infection, preceding virus assembly, budding and release, were affected by the depletion of OTUD6A, OTUB1 and PAN2. SiRNA treated A549 cells were infected with $H1N1_{WSN}$ virus at high MOI of 3 for 3, 6 and 9 hours and western blot were performed to detect the expression level of NP, HA, M1 and PB2 proteins (Figure 43). The depletion of PAN2 decreased the expression of all these proteins, as detected at 6 and 9 hours post-infection, while OTUB1 or OTUD6A depletion did not affect any protein level.

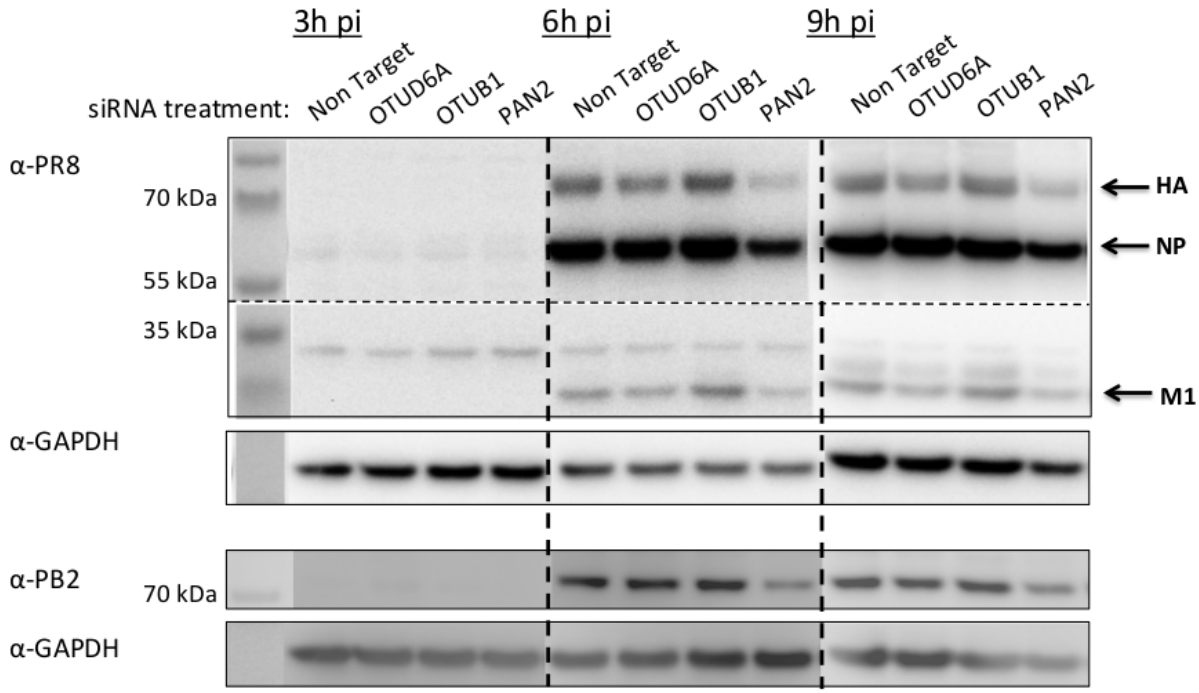


Figure 43: Expression levels of viral proteins in siRNA treated A549 cells at 3, 6 and 9 hours post-infection with $H1N1_{WSN}$. Western blot membranes were first incubated with antibodies against PR8 and GAPDH, then stripped and incubated with antibodies against PB2 and GAPDH.

The single cycle experiment indicated that PAN2 affects the sequential accumulation of viral proteins, but it does not interfere with viral polymerase activity (Figure 42). Owing to its poly-adenylase function, we explored whether PAN2 could affect the stability of the viral mRNA. To this end, we quantified the mRNA of the different segment (Figure 39), by specific qRT-PCR in A549 cells treated with PAN2 siRNA at 6 hours post-infection. The depletion of PAN2 did not alter the level of any mRNA, suggesting that the effect of PAN2 depletion is not due to mRNA stability (Figure 44a). In these experiments we detected an

increase of PAN2 mRNA during viral infection (Figure 44b). It has now to be studied whether the stability of other cellular protein mRNA is impacted and how it impact the viral replication.

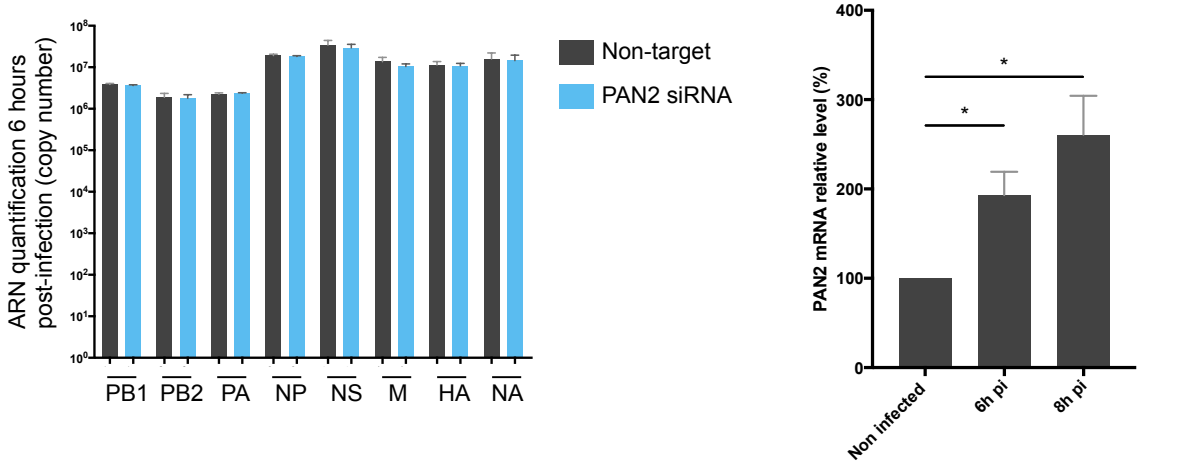


Figure 44: mRNA level of the diffent IAV segments and PAN2 during the course of infection. PAN2 or non-target siRNA treated A549 cells were infected with $H1N1_{WSN}$ virus at a MOI of 1. After 6 hours, mRNA were extracted and reverse-transcribed before qPCR with primers specific of the viral mRNA segment or PAN2. **a-** Viral mRNA level in A549 cells at 6 hours post-infection relatively to the non-target expression level. **b-** PAN2 mRNA level in A549 cells infected during 6 or 8 hours. Data represent means ±s.e.m. of three independent experiments, p values were calculated with a Wilcoxon test, 0.05>*>0.01.

OTUB1 is implicated in host cytokine response

Unlike PAN2 or OTUD6A, OTUB1 was found to be differentially involved in virus infection according to their subtype (Figure 32 and Figure 45). H1N1 subtypes seem dependent on OTUB1 expression, whereas no effect was detected in the different assay involving H3N2 viruses, this results being confirmed by our collaborators in Hong Kong (data not shown).

As OTUB1 is important for H1N1 viral infection, expression and localisation were assessed during infection with $H1N1_{pdm09}$ virus. Our collaborators in Hong Kong observed that OTUB1 expression was transitorily induced during the $H1N1_{pdm09}$ infection in A549 cells, with a maximum expression around 12 hours post-infection (Figure 46a). Moreover, OTUB1 localisation is shifted from cytoplasm to nucleus localisation, (Figure 46b), where it co-localises with the PB2 protein (Figure 47a,b).

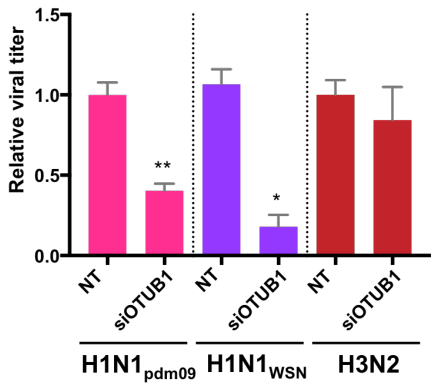


Figure 45: Differential involvement of OTUB1 in IAV infection according to subtype. A549 cells were treated with siRNA for 48 hours and infected with $H1N1_{pdm09}$, $H1N1_{WSN}$ or H3N2 during 24 hours at a MOI of 10^{-3}. The viral production was then assessed by plaque assay. Data represent means ±s.e.m. of three independent experiments with triplicates for $H1N1_{pdm09}$ and H3N2 and one experiment with independent triplicates for $H1N1_{WSN}$. P values were calculated with a two-tailed non-parametric Student's t-test, *<0.05, **<0.01, ***<0.001.

We performed a pull-down experiment in an infectious context, using a $H1N1_{WSN}$virus expressing a strep tagged PB2. HEK-293T cells ectopically expressing cherry-OTUB1 were infected at a high MOI with a tagged-PB2 expressing $H1N1_{pdm09}$ virus. After 24 hours cells were lysed and proteins precipitated with strepTactin beads. Pulled-down proteins were analysed by western-blotting using rabbit anti-cherry antibodies. The cherry-OTUB1 protein could be pulled-down along with the strep-tagged PB2 protein, indicating that the PB2/OTUB1 interaction occurs during $H1N1_{pdm09}$ infection, and suggesting that this interaction can be set up with PB2 as part of the polymerase complex (Figure 47c).

We have shown that OTUB1 was not implicated in early steps of viral infection (Figure 43) and did not modulate viral polymerase activity (Figure 42). Furthermore OTUB1 has been described as an enhancer of TGFβ cytokines signaling by inhibiting the degradation of the active complex SMAD2/3 [9], as well as an intermediate in IL1β–induced NF-κB–dependent gene expression [177]. Considering evidences of OTUB1 implication in host cytokine response, its expression was assessed after interferon β treatment in A549 cells. Similarly as upon IAV infection, OTUB1 was activated by an interferon β (IFN) treatment (Figure 48a), indicating that this factor is an IFN responsive gene. Moreover, in OTUB1 knockout (KO) cell lines, obtained by our collaborators using the CrispR-Cas9 methodology, the production of pro-inflammatory cytokines upon infection with $H1N1_{pdm09}$ was found to be lowered compared to the wild type cells (Figure 48b). Taken together, these data point to a plausible role of OTUB1 in virus-induced cell response.

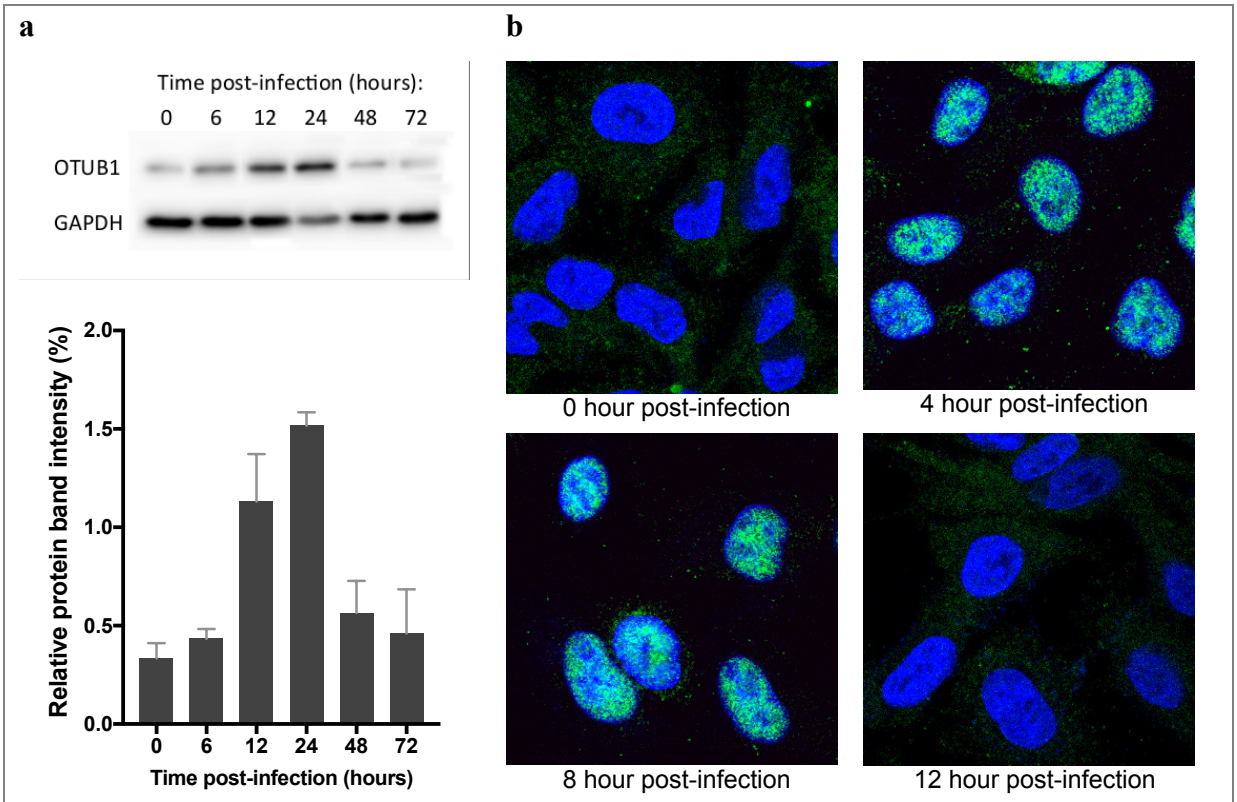


Figure 46: Up-regulation of OTUB1 and shift in sub-cellular localisation in A549 cells during IAV infection. **a-** A549 cells were infected with $H1N1_{pdm09}$ at a MOI of 0.1 and OTUB1 protein expression level was assessed in the whole cell lysates at different time points upon infection. The protein band intensity was measured using ImageJ software (NIH). Data represent the mean fold change of two independent experiments normalised with housekeeping GAPDH levels. **b-** OTUB1 shuttles from the cytosol into the nucleus during infection as early as 4 hours post infection. A549 cells were subjected to $H1N1_{pdm09}$ infection with a MOI of 2. Cells were then fixed at different time points, permeabilised, and stained with DAPI (dark blue) and Alexa fluor 488- OTUB1 antibody (green). Serial sections were collected with a 0.5 µm increment on the z- axis. *These experiments were carried out by Akhee Jahan at HKU-Pasteur Research Pole.*

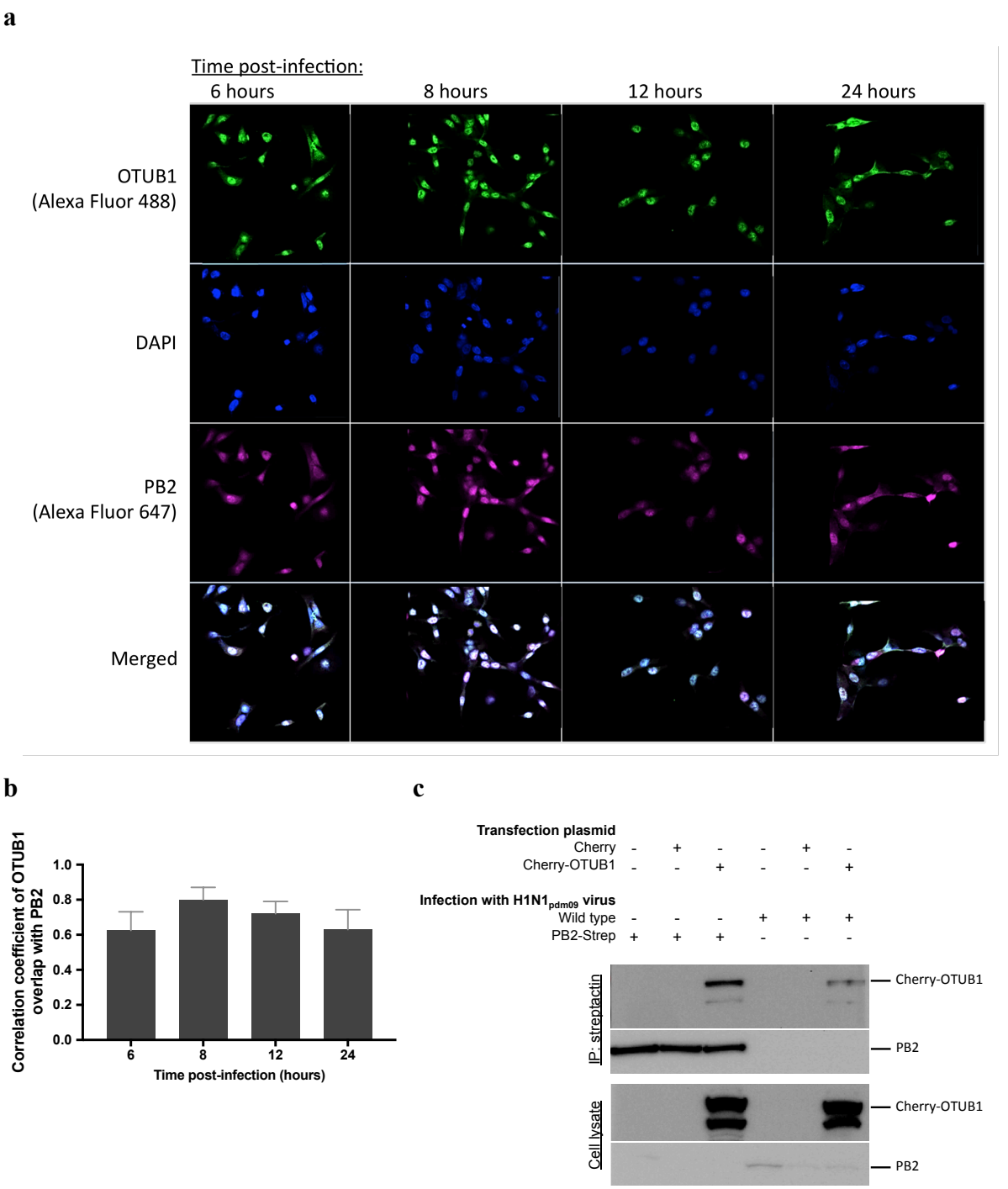


Figure 47: OTUB1 co-localises and interacts with PB2 during H1N1pdm09 infection. **a,b-** IAV viral subunit PB2 co-localises with OTUB1 starting 6hours post infection. A549 cells were infected with H1N1pdm09 at a MOI of 8. A549 cells were then fixed at different time points, permeabilised, and stained with DAPI (dark blue) and Alexa-fluor 488 anti-OTUB1 (green) and Alexa-fluor 645 anti-PB2 (pink). Serial sections were collected with a 0.5 µm increment on the z-axis. Z-orthogonal and 3D images were rendered and analysed using the Zeiss ZEN co-localisation module. Data represent the mean co-localisation coefficient ±s.d. of triplicates. *This experiment was carried out by Akhee Jahan at HKU-Pasteur Research Pole.* **c-** OTUB1 interacts with PB2 during infection. A549 cells were transfected with cherry or cherry-OTUB1 expressing plasmid and infected 24 hours later with wild type H1N1pdm09 or PB2-strep H1N1pdm09 at a MOI of 1. After 24 hours cells were lysate and incubated with streptactin beads. Pull down was performed, followed by western blot and membrane was incubated with anti-cherry and anti-PB2 antibodies.

In all, our data indicate a probable dual function of OTUB1 on the infection with H1N1 viruses: it is accumulated and targeted to the nucleus during infection, is required for efficient viral replication, and at the same time seems involved in the antiviral response by mediating cytokine production. More work is on going to understand the role of OTUB1 in H1N1 infection, as well as its strain-specificity. In collaboration with the team of Dr Sumana Sanyal in the HKU-Pasteur Research Pole of Hong Kong, the mechanisms of OTUB1 up-regulation during the course of infection will be further explored, as well as the balance existing between its pro-viral functions and its involvement in pro-inflammatory response that is expected to be anti-viral. Preliminary results obtained from immunofluorescence imaging led to the hypothesis that OTUB1 would be implicated in the late stages of the viral cycle for the assembly of the viral particle.

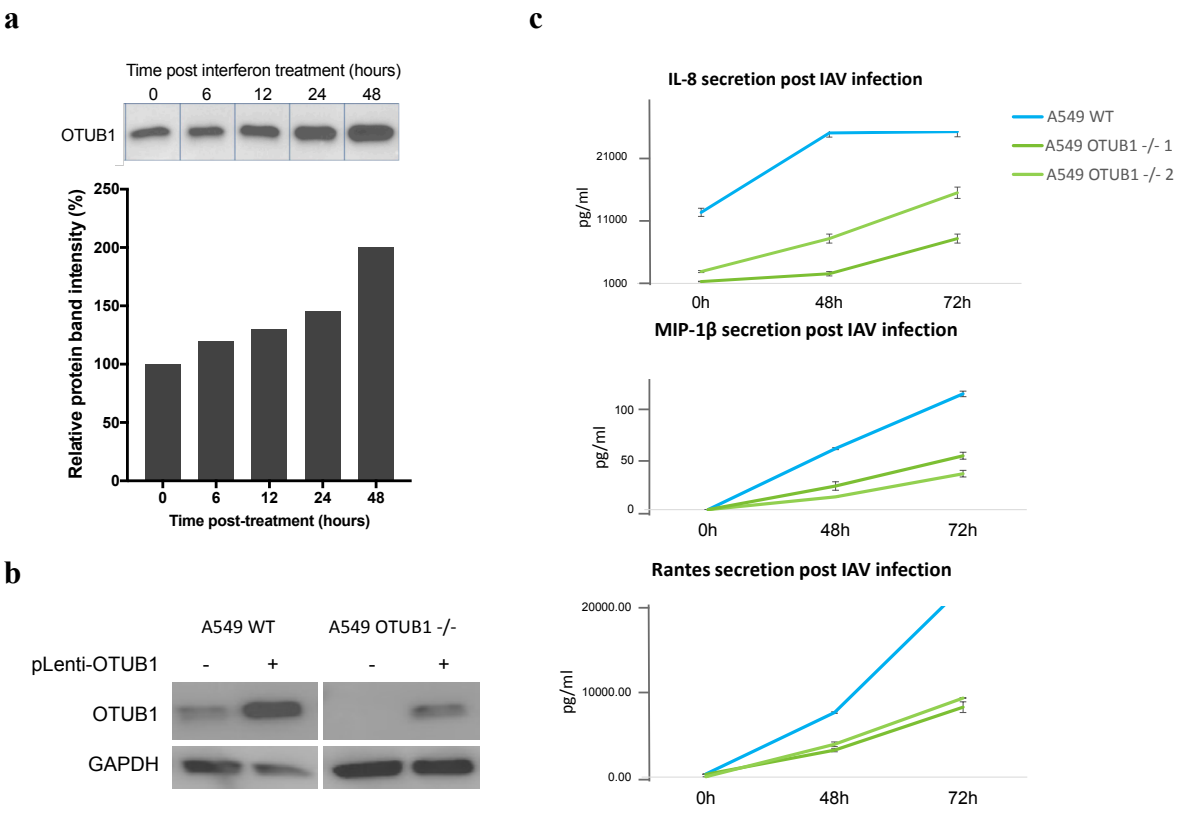


Figure 48: OTUB1 is part of the host cytokine response. **a-** OTUB1 expression induced with interferon β treatment in A549cells. Cells were subjected to treatment with IFN β (100U/ml) for 24 hours, and OTUB1 protein level was assessed in the whole cell lysates at different time points following treatment. The protein band intensity was measured using ImageJ software. Data represent the mean fold change of two independent experiments normalised with housekeeping GAPDH. **b-** OTUB1 knockout cells were produced by CrispR-Cas9. The knockout was confirmed by rescue of OTUB1 expression with a pLenti expressing OTUB1. OTUB1 expression was assessed by western blot. **c-** Two A549 cell lines depleted for OTUB1 expression by CrispR (OTUB1-/-) were infected with $H1N1_{pdm09}$ at a MOI of 0.1. Cytokines were measured using bead based FACS experiments, in supernatants that were collected at the indicated times post infection. Data represent means ±s.d. of three independent experiments. *These experiments were carried out by Akhee Jahan at HKU-Pasteur.*

Discussion

The UPS library

The usurpation of UPS is a recognised, significant process of viral infections, which has been shown to take part in viral pathogenesis [178]. As described earlier, the role of UPS in influenza virus infection is recognised, but a global picture of viral protein interplay with UPS is lacking. To fill this gap of knowledge, we focused our interactomics strategy on the human UPS.

To select human UPS factors from the ORFeome library, we took advantage of two *in silico* UPS-dedicated databases, the UUCD [35] and the DuDe [34] databases, where proteins of the human UPS are listed based on the genome-wide search of protein domain signatures for the various enzyme families (E1, E2, E3, DUB). These databases, while demonstrating a strong overlap, still present variations. We therefore combined them to obtain a more complete theoretical human UPS array, that we designated “human UPS”, thus enabling an exhaustive identification of all UPS-related factors available in the ORFeome collection. In the resulting *in silico* human UPS, lots of factors did not have recognised UPS-related function but were selected due to the presence of protein domains or motifs frequently found in UPS factors. Some domains/motifs are specifically allocated to factors of the UPS pathway, such as F-Boxes typically found in substrate factors of Cullin1-containing E3 ligases. In contrast, other domains/motifs frequently found in UPS factors are also regularly found in a variety of proteins of other functions, meaning they are not strictly dedicated to UPS-related factors. This is the case of the BTB, WD domains, or zinc fingers of the RING type. For those factors we performed a manual sorting according to their described function in the UniProt database [179], and factors identified as implicated in irrelevant cell functions were removed from the library. This was in particular the case of some BTB domain-containing proteins of the KCTD family related to K-channel tetramerisation. Of note, the substrate recognition factor of CUL4-based CRL are characterised by the presence of a DWD ("DDB1-binding WD40 protein") motif, a 16-residues motif in WD40 repeats. This motif is found in numerous WD containing proteins as a PPI domain. The E3-DWD category in the theoretical human UPS contains 271 genes, of which only 70 have a reviewed function in the UPS. This category is

one of the least represented in our UPS library, but only one of our 21 E3-DWD has no UPS-related function, meaning that, among the lacking E3-DWD, only 50 were reviewed as UPS factors. The under-representation of E3-DWD is thus over-evaluated regarding the UPS-related factors. We had in mind in the elaboration of the UPS-dedicated library to provide the higher reachable exhaustiveness with regard to human UPS. That is why we chose to retain all factors, for which it was not possible to rule out an involvement in the UPS. In all, the collection of UPS factors assembled in our UPS-dedicated library represents roughly 50% of the theoretical human UPS, which represents to our knowledge an unprecedented coverage. In addition, our strategy allowed us to screen UPS factors not yet or scarcely explored in the literature, thereby avoiding bias of selection based on functional knowledge.

Power and limits of the HT-GPCA based comparative interactomics strategy

The HT-GPCA method has been shown to enable the specific and sensitive detection of protein-protein interactions (PPI) [142-144]. It has been used previously as a secondary PPI detection method tool for the validation of large number of interactions engaged by viral proteins of human papillomaviruses (HPV) and identified in yeast two-hybrid screens, thereby delivering improved virus-host interaction maps [143, 144]. Detection of false positive interactions, as well as the false negative rates, are inherent to all large scale interaction maps or interactomics studies, but those are the only way to assess host/pathogen interplays at a global level. Previous studies showed that HT-GPCA has a particularly low false negative level for PPI detection, between 70% and 80% of known interactions (interactions with the Positive Reference Set) being recovered, while the false positive PPI detection was estimated to be less than 5% [142-144]. Despite such excellent performances, we could not exclude that a number of interactions may have not been detected by our HT-GPCA screening. Indeed, the PPI detection might be altered by the fusion configuration of both partners: the nature of the Gluc fragment, Gluc1 or Gluc2 trans-complementing fragments of the *Gaussia princeps* luciferase; and its location at the N- or C-terminal position of the protein. We experienced such PPI detection variability during HT-GPCA benchmarking with the PB2 proteins of the 5 strains to be studied. We compared the detected interactions between a set of PRS and RRS factors with the PB2 proteins fused to the Gluc2 fragment in N- or C-terminal position. We observed that some PPI are strongly dependent upon Gluc configuration. This is especially

true for the factors involved in the nuclear import of PB2, while several of them were not recovered for any PB2 in any configuration. It is assumed that using a single configuration for HT-GPCA would be at the cost to likely loose some interactions. For more exhaustiveness the PPI screening should be applied using 2 (N-ter, C-ter) or even 4 (Gluc1 in N- and C-ter, Gluc2 in N- and C-ter) possible configurations, which is out of scale for high throughput screening. The most effective PPI results were obtained with PB2 from the $H1N1_{WSN}$ strain, which is among the most studied strains and was used to identify the vast majority of the virus-host interactions previously published. It is therefore not surprising, that the PRS were mostly recovered with this strain, which however is maintained in cell lines for *in vitro* experiments since a long time. This questions the relevance of the host factors found to interact with factors of this strain in regard to interactions that happen with field strains.

One of the caveats that may arise from comparative interactomics studies is the difference in intrinsic, aspecific binding properties of the proteins to be compared. To overcome this issue, we proposed using a set of a priori non-interacting proteins, randomly pick up in the human ORFeome collection, to assess the aspecific binding of each viral factor. Determination of a threshold for positive PPI according to the signals measured with this RRS set enable to take into account the aspecific binding level of each viral prey. Importantly, the variability in intrinsic properties of the viral baits is not problematic for the comparative aspect of this interaction mapping, as demonstrated with the 5 PB2 proteins of different IAV strains. On the other hand however, the host factors inevitably exhibit a wide range of intrinsic binding level. The measurement of the host factors against the unfused reporter protein fragment (Gluc2 in our case) enables to partially control this point, by identifying host proteins of high aggregation properties that are to be taken cautiously for PPI detection (Annexe 1).

In this study, we pioneered the use of the split-*Gaussia princeps* luciferase complementation assay (GPCA) for systematic, unbiased screening for binary interactions, in order to challenge all the host proteins for interaction with the PB2 viral proteins in the same assay. Such strategy drastically improves the screening coverage of virus-host proteome interplay. However, approximately 20,000 protein-coding genes exist in humans [180, 181], corresponding to even more proteins when considering the multiple isoforms. Therefore, addressing all human proteins for potential PPI would require more than 200 millions of protein pairs to be tested [182]. While such scale is out of range by now, it is feasible by focusing on sub-arrays of the human proteome thanks to the increasing collection of human

cDNA provided by the human ORFeome v8.1 [145]. This continuously implemented collection contains products of more than 11,149 human cDNA, thereby covering more than half of the human coding genes. The human ORFeome should reach exhaustiveness in term of human proteins representation in the next future. Its current version already enables the near-systematic levels of PPI mapping, when considering a single protein isoform per protein coding gene. For now, systematic studies focused on specific cellular pathway can provide comprehensive maps associating PPI datasets to functional knowledge. Interaction proteomic is indeed a straightforward approach to assess viral-host interplay and to further comprehend its impact on the cell thank to network analysis. The evaluation of the same matrix of host factors for interactions with multiple viral proteins provides a rigorous comparative interaction mapping, and may shed light to differences in virus-driven targeting of the host proteome, or specific cellular response, potentially underlying some traits of viral pathogenic potential.

Combined screening strategy

The UPS-dedicated library screening strategy was realised in parallel with the E6 and E7 proteins of the Human Papillomavirus HPV16 by our colleagues in Strasbourg, through a close collaboration. The use of a common and constant set of RRS with all the viral proteins reinforced the characterisation of each RRS interaction level. We clearly detected that some of our RRS are pro-aggregative proteins, and we are confident that our RRS set is representative of different aspecific binding characteristics, which is essential for the screening strategy.

Different interplays with the UPS emerged with the PB2 proteins, HPV16 E6 and HPV16 E7. Contrariwise, the interplays between the UPS and the PB2 proteins from different strains were significantly related to each others, which certified the stringency of the viral protein-UPS interactions identified, and strongly supported the robustness of our interactomics strategy. In conclusion, our UPS-dedicated library and the HT-GPCA screening strategy have been used successfully to identify novel interactions between viral proteins and the ubiquitin proteasome system. Moreover the assessment of the UPS-dedicated library with empty Gluc2, and the combinatorial screening with three viral baits constituted valuable information about potential aspecific interactions that could be taken into account for further screening experiments. We thus think that the split-luciferase strategy we developed allows the straightforward delivery

of high-quality comparative mapping of virus interactions with the human UPS. Importantly, it can be directly applied to proteins from any other pathogens, to define the landscape of pathogen-UPS interplays.

The benefit of comparative interactomics: the hierarchical clustering

Influenza pandemics can occur when zoonotic influenza viruses adapt for efficient replication and transmission in humans. The natural incompatibility of avian IAV strains with the human host can be explained by the lack of factors supporting avian viral replication in human cells or the presence of restriction factors that cannot be counteracted by avian strains until adaptive mutations are selected [183]. We showed by agglomerative hierarchical clustering that the PB2/UPS interaction patterns are segregating according to the lineage of origin (avian or human) of the PB2 segment, as well as of its time of circulation in the human population.

Introduction of the purely avian PB2 in the interaction dendrogram suggests that the PB2/UPS interaction profiles may be relevant to the abilities of avian strains to infect human before they acquire human-to-human transmission potential. As such, the PB2/UPS interplay may help to prognosticate human infection capabilities of IAV strains emerging from avian reservoirs. Analysis of the PB2/UPS interaction for a larger number of human and avian strains collected at different time points will be needed to support this possibility.

No comparison has been made yet between the human and the avian UPS. Nevertheless, a study focused on RIG-I activation in duck and human showed that unanchored ubiquitin may play the main role in RIG-I signaling in duck, whereas it seemed to be only an accessory mechanism to covalent ubiquitination in human [184]. It would be interesting to compare the human and the avian UPS to know how they differ, since as shown with the previous example, adaptation of IAV strains to counteract different signalling pathway in avian and human hosts likely contributes to interspecies transmission and pandemic risk.

PB2/UPS interplay

We identified a rather extensive interplay between PB2 and the human UPS, and showed that PB2 proteins bind to E3 ubiquitin ligases or E3-related factors ("others" in our UPS-dedicated library), to DUBs, and to a proteasome subunit, but not to E1 or E2 enzymes. Moreover, the PB2 proteins interact with all RING E3 ligases and Cullin-based ubiquitin ligases (CRL), mainly through their Substrate Recognition Factors (SRF) (Figure 49, Digital supplementary data 5). This suggests that the PB2 proteins induce an UPS remodelling mainly through subversion of E3 ligases recognition, or that they are targeted by the host, thereby affecting host and/or viral protein ubiquitination, as it was often observed with viruses through UPS hijacking. Similar E3-enriched interplays were detected for the HPV16 E6 and E7 proteins. Intriguingly, the PB2 protein targets an extended pattern of E3 ligases, encompassing RING-type E3 ligases, and substrate recognition factors of all types of CRL, without any preference for binding to specific E3 ligases sub-categories (Figure 49). In contrast a strong enrichment for targeting SRF of CUL3-CRL has been observed for the E7 protein [170] (Annexe 4). For E6, no specific E3-ligase targeting enrichment could be observed, but it is the only viral protein found to target the HECT E3 ligase family (Figure 31, Annexe 4), the smallest E3 ligase family.

Of note, PB2 associates with three SRF of CRL complexes based on Cullin 4 (DCAF4L1, DCAF4L2, DCAF11) and with Cullin 4B itself. Interestingly, these factors were shown to belong to the same protein complex in the recently delivered Bioplex network, obtained through high-throughput affinity purification and mass spectrometry experiment [185]. This complex contains also the COP9 signalosome, involved in the regulation of CRLs through Cullin deneddylation. This suggests a possible impact of PB2 on COP9-mediated regulation of Cullin E3 Ub ligase activities. Interestingly, the COP9 signalosome complex recently emerged as a significant component of the Influenza virus and host interface, with an involvement in viral replication [154].

Three subunits of the proteasome, PMSB2, PMSB6 and PMSD8, were reported to interact with PB2, as detected in TAP/mass spectrometry, leading to the recognition of protein degradation as a cell function targeted by influenza virus [154]. None of these proteasome subunits have been addressed in the present study because they were not present in our UPS-library. However, of the 4 proteasome subunits present in the library, one, PSMD2, has been

identified as an interacting partner of PB2, substantiating the interplay between PB2 and the proteasome. In view of this observation, it would be worth to assess each individual proteasome subunit for interaction with PB2, and to explore the functional impact of such interplay on the regulation of protein degradation in the cell.

The interaction with several DUBs also suggests that PB2 might interfere with the versatility of ubiquitination process. A different subset of DUBs was targeted by the HPV16 E6 and E7 proteins (Annexes 4 and 5), suggesting that it might be a general mechanism during viral infection.

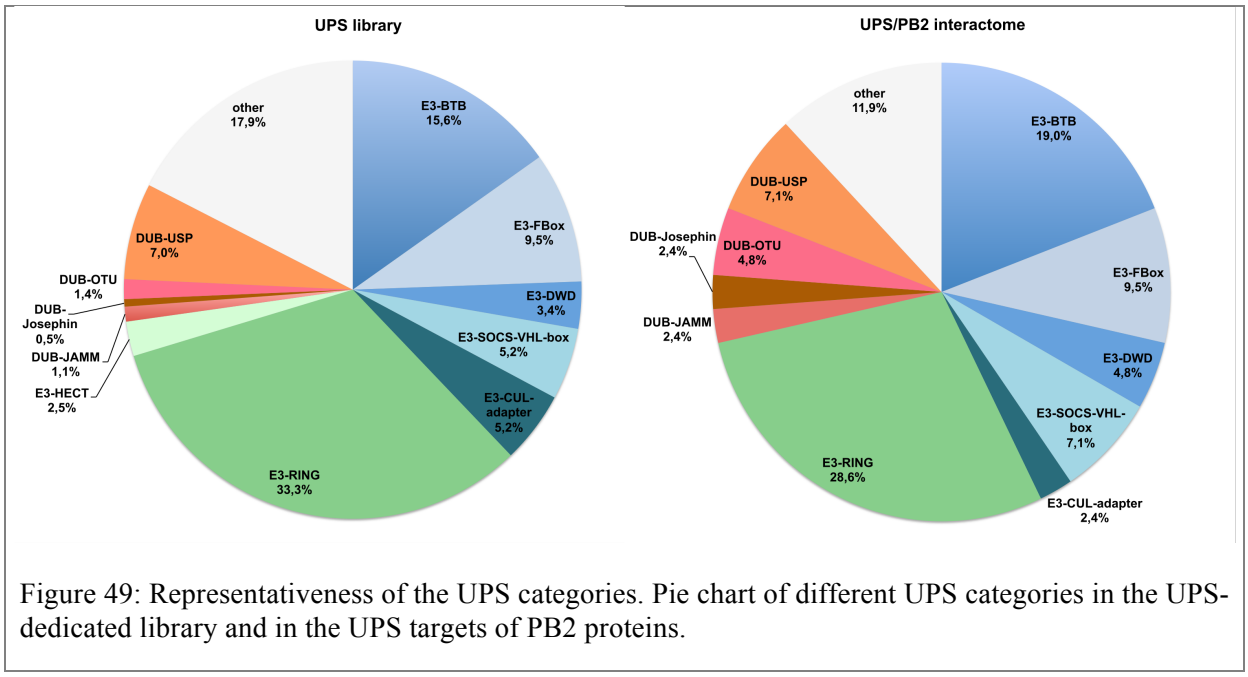


Figure 49: Representativeness of the UPS categories. Pie chart of different UPS categories in the UPS-dedicated library and in the UPS targets of PB2 proteins.

Comparison with other vRNP interactomes

Virus-host protein interaction networks can be deduced from the obtained binary PPI datasets. A PB2/UPS interaction map encompassing all the strains tested was plugged to the human interactome, recovered from the BioGRID PPI database [186]. Some of the PB2 targets are interacting with each other (Figure 50a), underlying an intricate interplay where PB2 binds to UPS proteins that already act in complexes. When all human factors interacting with the UPS partners of PB2 are considered, an elaborated PB2-UPS-Human Proteome interaction map emerges (Figure 50b), revealing the possible impact of PB2 on the human proteome through UPS targeting.

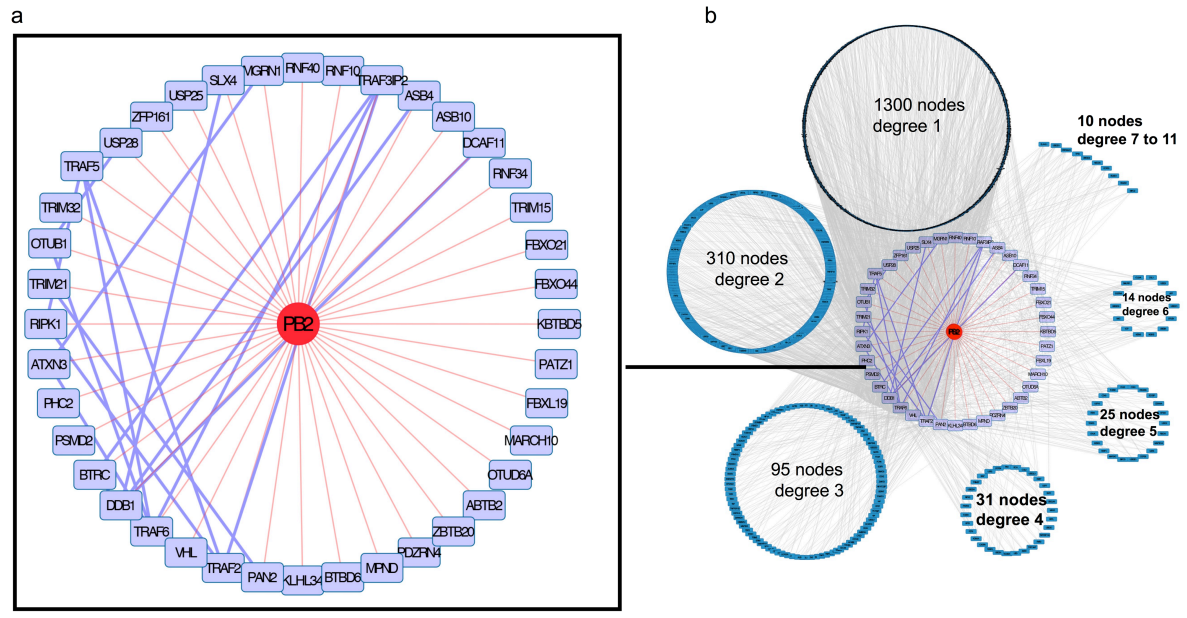


Figure 50: The PB2/UPS interactome. **a-** Map of the PB2-UPS interactions. The PB2-UPS direct interactions are labelled in red, and include interactions detected with all the PB2 proteins tested. Interactions existing between the UPS targets of PB2 are labelled in purple. **b-** Plugging of the PB2-UPS network onto the human interactome. The PB2 UPS targets are labelled in purple. The host proteins they connect are organised in circles, according to their degree of connection (node degree) with the UPS factors interacting with PB2. A node degree of 1 means an interaction with 1 UPS factors of the PB2/UPS interactome, of 2 with 2 UPS factors of the PB2/UPS interactome etc.

Previous mass-spectrometry-based (MS) proteomic studies identified 171 [117] and 45 [156] host proteins in complex with the vRNP proteins (Table 3). Also the combination of the results obtained in a number of High Throughput-MS-based proteomics with isolated replication proteins pointed to 21 high-confidence PB2-associating host factors [168]. These proteomics studies identified proteins in complex with PB2 or with the vRNP without giving access to the direct interactions underlying complexes formation. We searched for the vRNP/PB2 associated host proteins in the PB2-UPS-Human Proteome network, and found that 52 [117], 29 [156] and 10 [153] of the proteomics-identified vRNP/PB2 associated factors are binding partners of the UPS targets of PB2 (Figure 51). This suggests that these host factors could be recruited to the vRNP/PB2 through their binding to an UPS factors that directly binds to PB2. Of the 42 UPS factors that we detected as directly binding to PB2, 24 have a published interaction with at least one vRNP/PB2-associated host factor. In addition, some of the UPS partners of PB2 interact with a large proportion of vRNP-associated factors, such as PAN2 for which an interaction has been detected with 45 vRNP-associated host factors, meaning that PAN2 could act as a scaffold for the recruitment of numerous host

factors to the vRNP through PB2. The DDB1 protein was found associated with the vRNP in two independent studies [117] [156], and we showed that it is a direct partner of PB2, suggesting that PB2 is implicated in recruiting DDB1 to the vRNP. DDB1 has also been shown to interact with the vRNP in an infectious context, but it was not recovered with MS under expression of the viral polymerase alone [156], suggesting that the PB2 protein is interacting with DDB1 when expressed alone or in the vRNP context, but not in the context of an RNA-free, unbound viral polymerase. These differences might rely on the different conformations adopted by PB2 in the polymerase complex [110].

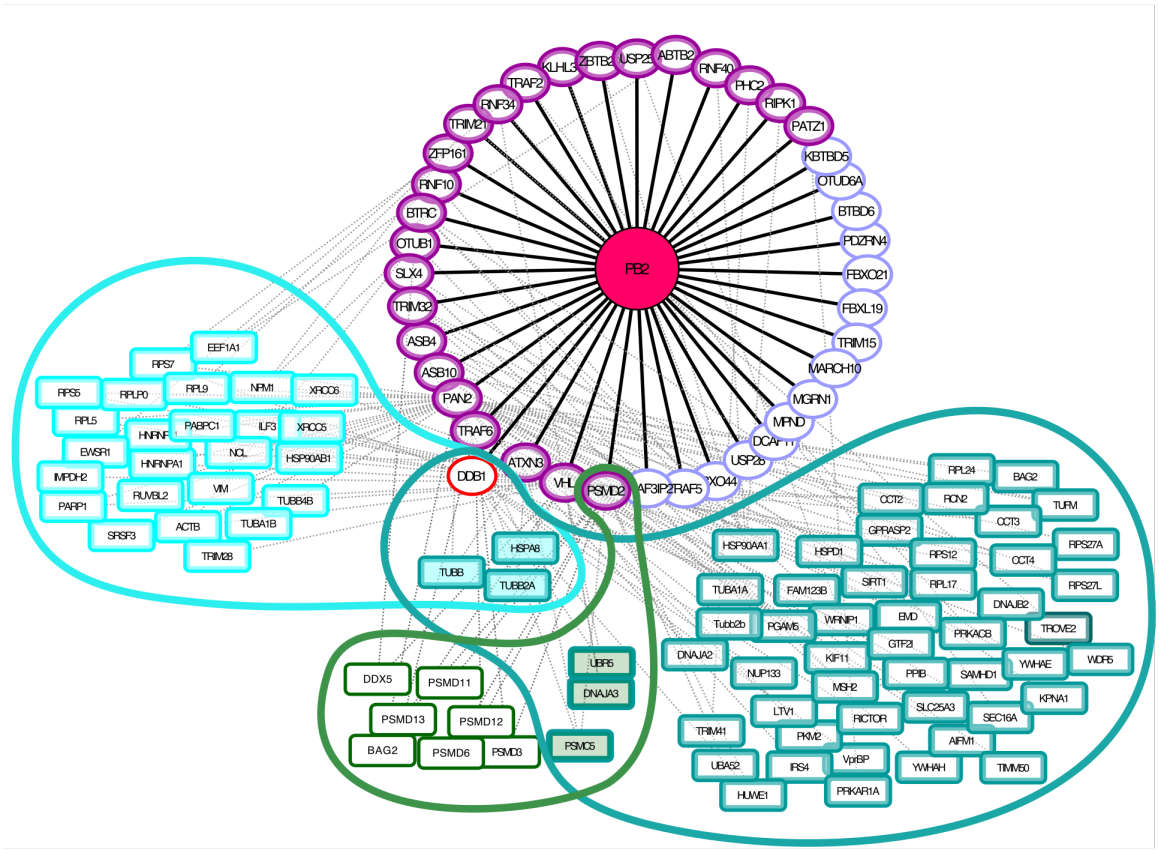


Figure 51: Merge of the PB2-UPS interactome and host partners in complex with the vRNP as detected by AP-MS. Several host factors identified as associated with the vRNP [117] [156], or found associated with the PB2 proteins in ≥ 3 independent studies in Zao *et al.* [153], are interacting partners of UPS factors binding to PB2. Interactors identified in York *et al.* [117] are coloured cyan, blue for those identified in Mayer *et al.* [156], and green for those listed in Zhao *et al.* [153]. The set of corresponding factors are circled in colours to highlight commonly identified vRNP/PB2 associated factors.

Several host factors have been identified in multiple MS studies (Table 3) and turn out to associate with factors binding to PB2, which altogether define these factors as high confidence vRNP-associated factors, most probably recruited via the PB2 protein. On the

other hand, several vRNP interactors detected by MS associated are UPS factors that were not identified in the present study as PB2 UPS partners. It suggests that these UPS factors were pulled down with the vRNPs via binding to NP, PB1 or PA, or through interactions with other factors of the vRNP interactome, but we cannot rule out that they may represent HT-GPCA false negatives as well.

Altogether, the combined analysis of MS-based and HT-GPCA based approaches to explore host factors associated with the vRNP or PB2 highlights the complementary nature of the two strategies, both providing valuable information regarding viral-host interplays that need to be assessed functionally to explore their relevance relative to IAV infection.

Functional exploration of the UPS targets of PB2

Of the 42 top-confident PB2 interactors identified in our screening, a total of 36 were shown by siRNA-based experiments to affect IAV replication, with the depletion of 13, 25 and 31 UPS factors decreasing the production of the H3N2, $H1N1_{pmd09}$ and $H1N1_{WSN}$ viruses, respectively. Such high functional output emphasises the strong power of this interactomic strategy to detect biologically relevant partners of PB2.

SiRNA experiments pointed to six factors involved in the virus cycle of all three low to mildly virulent IAV strains studied, and could represent a core set of UPS factors generally involved in the regulation of IAV infection. They consist on members of Cullin-based E3 complexes (DCAF12L, DDB1 and FBXL19), a component of the H2A histone Ubiquitin ligase complex (PHC2), and DUBs (OTUD6A and PAN2). DDB1 is an adapter of multiple Cullin4-based E3 ligases complexes and has pleiotropic effects, whereas the SRF DCAF12L1 integrated in some of these complexes has only one documented target: the SH2 domain-containing protein 2A. The SRF of Cullin1-based complex FBXL19 has been shown to target the receptor of IL-33 (ST2L) for degradation, thereby limiting the inflammatory effects of IL-33 [187]. Its association with the PB2 protein could alter inflammation in response to IAV infection. We suspect that the binding of PB2 to SRF could rewire the E3 ligases complex toward novel cellular or viral factors for ubiquitination. The PCH2 protein is a component of a complex required to maintain the transcriptionally repressive state of many cellular genes. Its interaction with PB2 might be involved in the colocalisation of the vRNPs to inactive

chromatin that has been detected at late time of IAV infection [188]. The OTUD6A deubiquitinase hydrolyzes all ubiquitin chains but the K48-linked chains, but has been barely described in a functional point of view. Contrariwise, the PAN2 protein contains an inactive catalytic deubiquitinase domain that could alter proteins ubiquitination through interaction with other DUBs or E3 ligases, but is, as of now, only known to form an mRNA deadenylation complex with PAN3 [189].

For some other UPS factors, a differential effect was observed from strain to strain, highlighting an unexpectedly diversified functional interplay between IAV and the UPS. Strain-specific disparities may provide clues on the molecular mechanisms underlying differences in pathogenicity, while conserved effects could serve as targets for broad-spectrum therapeutics.

These results underlie the importance of working on different circulating strains and not only on laboratory-adapted viruses such as $H1N1_{WSN}$. Moreover, as discussed previously, interactions that are described with $H1N1_{WSN}$ are not constantly detected in other strains. This questions the relevance of strains with a long time adaptation to cell culture conditions. However, tools and techniques are usually developed with these viruses that are much more easy to manipulate and grow well in laboratory conditions. Indeed the $H1N1_{pdm09}$ and H3N2 seasonal strains we used in this study were producing lower viral titres than $H1N1_{WSN}$, which lead us to generate viruses adapted to A549 cell line (see *How to finally tame the seasonal strains* p73). Their cycle however, is not well characterised and less is known from the literature, making these viruses more difficult to work with. In addition, they are not easily manipulable, and the introduction of tag for example is not always successful. Despite a lot of trials, I was not able to rescue $H1N1_{pdm09}$ and H3N1 viruses with a PB2 fused to a Gluc fragment, which was achieved for $H1N1_{WSN}$, but I nonetheless managed to rescue them with PB2 fused to a strep tag. I am convinced that it was worthwhile using seasonal-derived strains such as the $H1N1_{pdm09}$ and H3N2 in my project considering the unexpected diversified outcome of the UPS factors depletion that I could detect according to the strains. Especially, we showed that the natural strains are surprisingly less susceptible to UPS depletion, allowing us to point to the UPS factors that are most probably involved in infection with natural strains and not solely linked to over-optimisation of $H1N1_{WSN}$ in cell culture.

Surprisingly, apart from FBXL19 with H3N2, none of the factors we tested seemed to act as inhibitor of IAV infection (Figure 32). It should be stressed that a role in the life cycle could have been missed for several UPS PB2 interactors in our experimental settings. This is evidenced by the fact that UPS factors SPOPL and UPS11, shown to be involved in the IAV infection in previous studies [4, 14] did not exhibit any consistent effect in our experimental settings (Figure 32). In contrast, a 5-fold decrease in the progeny $H1N1_{WSN}$ and $H1N1_{pdm09}$ virus titres was recovered upon depletion of Itch, in line with published results [15]. A role of SPOPL in the endocytic-mediated cell entry of IAV has been demonstrated by means of NP detection at early time points post-infection in SPOPL-depleted cells. It is thus plausible that the multi-cycle settings used here did not enable to detect defects in such early stages of IAV life cycle. It could conceivably be caught up in the course of multi-cycle infection. The UPS11 DUB has been described as an inhibitor of IAV infection, its depletion increasing viral production [4]. The low MOI used in our study (10^{-3} and 10^{-4} pfu/cell for the seasonal strains and for $H1N1_{WSN}$, respectively), relative to the MOI of 1 used in the published study, may have disfavoured the detection of positive effects on infection.

Involvement of DUBs in IAV infection

Besides the interactions detected between PB2 and DUBs, we also showed that these DUBs interact with other viral proteins. On another side, all the viral internal proteins except PA engaged interaction with at least one of these DUBs, highlighting a strong degree of redundancy for DUB targeting by IAV proteins. We were surprised not to detect more redundancy linked to the function of the viral proteins, for example enriched interactions with replication proteins PB1/PB2/PA/NP, which would have reflected factors interacting with the vRNP and possibly affecting viral polymerase activity. The NS1 protein seems to interact with all the DUBs tested but one. Whether the NS1 protein engages particularly extended interactions specifically with the DUBs, the whole UPS, the full human proteome, or with specific subarrays, would be possible to assess using the UPS library as well as other pathway-dedicated libraries. Small-scale HT-GPCA libraries recapitulating the IFN, NF-κB and TNF signaling pathways are available and could be used for that purpose.

We have shown that several DUBs are implicated at diverse stages of the infection: in all tested strains for OTUD6A and PAN2, and specifically in H1N1 subtype strains for OTUB1.

OTUD6A and OTUB1 belong to the OTU family of deubiquitinases, which show linkage specificity. OTUB1 specifically cleaves Lys48 linkage [45], whereas OTUD6A is specific of Lys27, Lys29 or Lys33 linkage [49]. Interestingly, PB2 proteins that can interact with both OTUD6A and OTUB1 covers almost all the possible linkage types, which might be of importance in the fine tuning of the UPS interplay. Nevertheless, our collaborators and we also found that some UPS deubiquitinases that favour viral infection do not show linkage specificity. This could consequently enable a more global diversion of cellular pathways if redirected by viruses.

The results obtained in our study concerning OTUD6A were quite puzzling. Whereas its depletion decreased viral production in multi-cycle settings, it also induced an increased viral polymerase activity in the context of the mini-replicon assay. This effect was however not detected in single cycle setting experiments. It should be taken into account that the polymerase activity was assessed in mini-replicon assay in an artificial setting without all the other viral proteins and segments. Variation in OTUD6A expression may then modulate the viral polymerase activity in a non-infectious context, but this is probably not the mechanism that is involved during infection. The fact that OTUD6A is known to be over-expressed in the lung compared to other tissues [45] reinforces the arguments pointing to it being an interesting factor regarding IAV infection. More experiments have to be conducted to understand how this DUB, which does not seem to be implicated in early steps of the infection, is involved in the viral-host interplay during the viral cycle. If OTUD6A is not involved in early steps of infection, it might be involved in the assembly and/or budding of the virions, or in the host immune response (Figure 52). Immunofluorescence staining and imagery in the course of infection can help to investigate the first hypothesis, whereas study of the cytokine and immune pathways can address the second one, as for OTUB1.

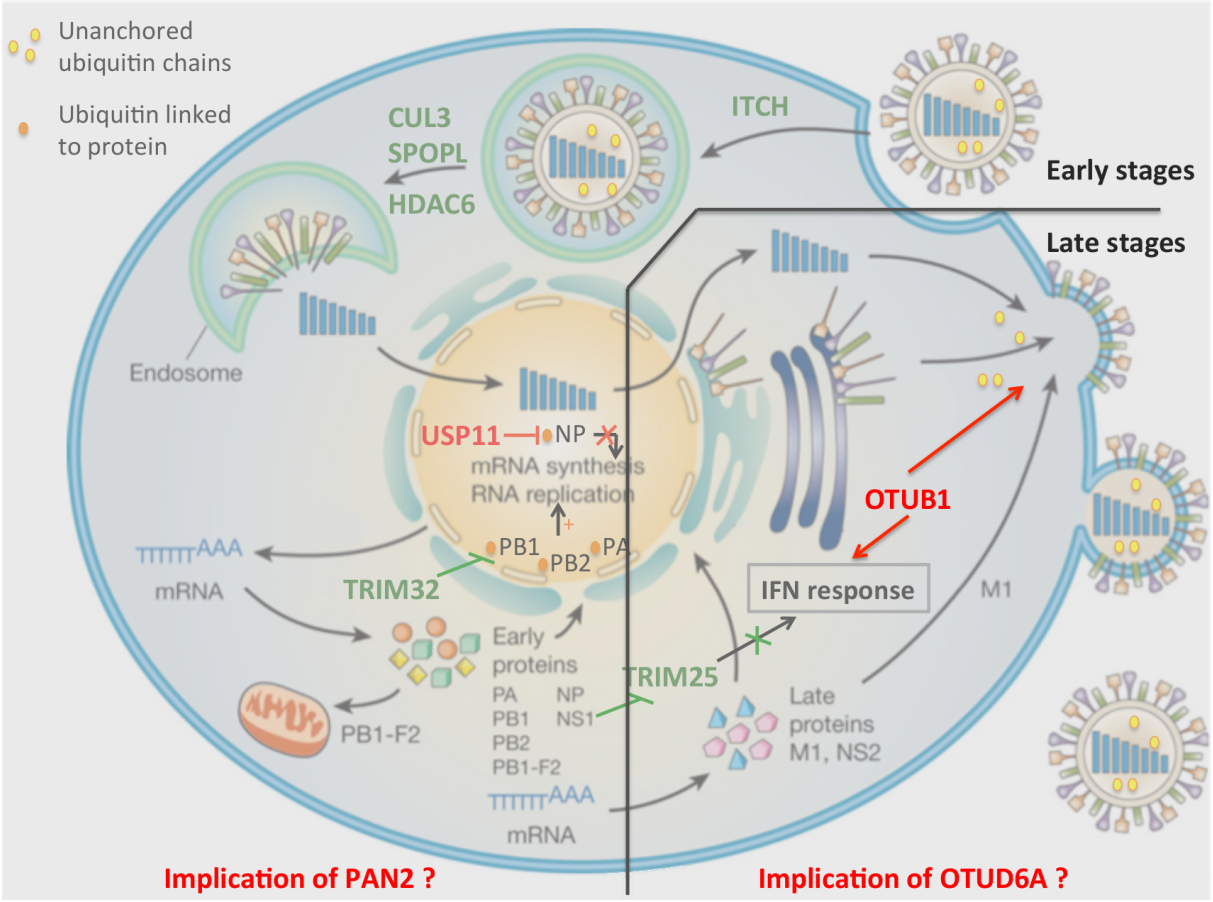


Figure 52: Implication of the studied DUBs in IAV life cycle. OTUB1 is implicated in interferon response of the cell and virus assembly. Whereas no clear roles have been elucidated for PAN2 and OTUD6A. However, PAN2 might be implicated in early stages and OTUD6A in late stages of the infection.

We already know a little more about OTUB1 role in the infection. We have shown with our collaborators that OTUB1 is involved specifically in late infection stages of H1N1 viruses, as well as in the host pro-inflammatory cytokine response (Figure 52). Preliminary results of our collaborators in Hong Kong led us to think that OTUB1 might also be implicated in the virus' assembly. Work is on going to decipher the dual function of OTUB1 in both promoting viral production and being involved in the production of pro-inflammatory cytokines, these two functions being opposite at the first glance. It would also be interesting to address the question of the strain specificity. As OTUB1 is involved in the viral production of H1N1 strains, how do H3N2 strains realise the same function without involving OTUB1?

The last studied DUB was PAN2, for which the role in infection was not yet deciphered. In another context, the cellular response to hypoxia, PAN2 is known to co-localise in the cytoplasmic P-bodies with the hypoxia-inducible factor 1α (HIF1A) mRNA, and to protect it

from degradation in a 3′-UTR-dependent but poly(A)-tail-length-independent manner [189]. Moreover, evidence points to P-bodies implication in influenza infection: NP has been shown to be targeted to P-bodies at early time point of infection, while NS1 counteracts this host restriction mechanism [190]. P-bodies have also been reported to serve as reservoir for cellular 5'mRNA caps for the viral cap-snatching mechanism of hantaviruses, which nucleocapsid protein accumulates in P-bodies, where it sequesters protected 5'mRNA caps [191]. Influenza virus transcription also relies on cap-snatching [192], but no link has been made so far with P-bodies. Lastly, the P-bodies formation relies on deadenylation, which implies PAN2, PAN3, CAF1 and CCR4, deadenylases [193]. Given these elements, we have hypothesised that PAN2 could affect protein expression level through different mechanisms: post-transcriptional modification of poly(A)-tail length, protection of mRNA or availability of 5'mRNA caps. All of these hypotheses would induce a decrease in viral mRNA in the absence of PAN2. But it appeared that the amount of viral mRNA is not affected by the depletion of PAN2. However, PAN2 could also regulate the mRNA level of a cellular protein important for the viral cycle. Another hypothesis is that PAN2 is acting through its pseudo-DUB characteristics. Indeed PAN2, also called USP52, is an USP with an inactive enzymatic site, and might be able to bind substrates as a DUB [189]. It is then possible that, by binding to some substrate, PAN2 would protect it from an active DUB or would play a scaffolding role for other PPI. Experiments will have to be conducted at the protein level to explore this hypothesis. Ubiquitination of the viral proteins in presence or absence of PAN2 would be interesting to address first.

Conclusion

The new systematic interaction mapping methodology that we developed enables to build interaction profiles in a comparative manner, based on a well characterised human UPS-dedicated library, ready to use for other screenings. The screening of viral proteins with the UPS library will benefit from the already mapped UPS/IAV and HPV proteins interactomes by pointing to specific UPS interactors or generally viral-targeted UPS factors. The UPS-dedicated comparative interactomic strategy constitutes a valuable resource, available to be applied to other pathogens to decipher UPS/pathogen proteins interplay that could also provide valuable insights into significant pathogen traits.

The comparative profiling of the IAV strains points to a correlation between PB2/UPS interplay and time of virus circulation in the human population, providing an experimental evidence for the role of host-viral interplay in the adaptation processes. Hence, studying the factors that interact with highly human-adapted strains but not with the avian strains, or strains that have shortly circulated in humans, could reveal new mechanisms of adaptation.

The UPS factors that were shown to interact with PB2 in this study implement the known interactome map of IAV proteins with the human proteome. PB2 interacts with different E3 ligases of the RING family and the Cullin family through their substrate recognition factors (SRF). These interactions can lead either to impaired ubiquitination of usual substrates, or to the E3 ligases rewiring to new substrates. This function was known for the E6 protein of oncogenic human papillomaviruses, that redirects E6AP to p53, but this observation is new for influenza viruses. The large panel of interactions between PB2 and E3 ligases suggests a lot of different functional outcomes depending on the implicated substrate recognition factor.

Besides, we have shown redundant interactions of a number of DUBs with different IAV proteins. Whether these interactions are induced by the virus during the viral cycle, or mediated by the cell immune response, this redundancy likely point to factors critical for the host-pathogen interplay.

The interactomic strategy identified six factors involved in the replication cycle of all studied strains. Because these factors should be involved in the regulation of all IAV, their study, combined with the resolution of the crystal structure of the polymerase [108] can be the starting point to broad-spectrum drug design based on the disruption of these critical interactions between the viral proteins and host factors.

However, the role of the UPS seems to be mainly strain specific since 30 of the UPS factors interacting with PB2 differentially affect the replication of the tested viruses. For example, we have shown that the deubiquitinase OTUB1 is crucial for the seasonal $H1N1_{pdm09}$ strain but not for the H3N2 strain we studied, whereas both of these strains circulate well in human. This shows that the infection process of IAV may differ between strains of similar pathogenicity. Going further into these topics is of prime interest for the development of drug targeting specific strains.

Materials and Methods

Plasmids

Gateway®-compatible Destination GPCA (pDEST) vectors pSPICA-N1 and pSPICA-N2 both derived from the pCiNeo mammalian expression vector, and expressed respectively the Gluc1 and 2 complementary fragments of the *Gaussia princeps* luciferase linked to the N-terminal ends of tested proteins after recombinatorial cloning (Gateway® system, Invitrogen). The ORFs encoding for the PB2 from influenza viruses A/WSN/33 ($H1N1_{WSN}$), A/Bretagne/7608/2009 ($H1N1_{pdm09}$), A/Centre/1003/2012 (H3N2), A/Anhui/1/2013 (H7N9), A/BrevigMission/1/1918 ($H1N1_{1918}$) and A/Mallard/Marquenterre/Z2371/83 ($H1N1_{MZ}$) were cloned into Gateway®-DONOR vector pDONR207, the resulting Entry clones were transferred into Gateway®-pDEST pSPICA-N2 to produce the Gluc2-PB2 expressing plasmids. The ORFs encoding the $H1N1_{WSN}$proteins PB1, PB2, PA, NP, M1, NS1 and NS2 were cloned into Gateway®-DONOR vector pDONR207, the resulting Entry clones were transferred into Gateway®-pDEST pSPICA-N1 to produce the Gluc1-viral protein expressing plasmids.
The ORFs encoding UPS factors were obtained as Gateway®-Entry plasmids from the human ORFeome [145] and transferred in Gateway®-pDEST pSPICA-N1 to obtain the Gluc1-UPS expressing plasmids and for some DUBs into Gateway®-pDEST pSPICA-N2 to produce the Gluc2-DUB expressing plasmids.
The RRS (Random Reference Set) contains human ORFs encoding proteins randomly picked from the human ORFeome and *a priori* not interacting with the viral protein PB2: LRCC28, NXP2, NFE2L1, GSTT1, GYPA, DPYSL2, UGT3A1, DBH, PLEKHA9, NXPH1, CNTN2, SLC7A13, CACNG7. The PRS (Positive Reference Set) corresponds to human ORFs, encoding NFX1, RUVBL2, NUP50, PTGES3, KPNA2 proteins pick up from the human ORFeome, shown to bind PB2 in the literature.

For the expression of proteins fused to mCherry, gaussia full-length or strep tags, the ORFs were introduces by Gateway® in a pDEST plasmid containing the fusion.

Cell lines

HEK-293T and A549 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal calf serum (FCS). MDCK-SIAT cells were grown in Modified Eagle's Medium (MEM) supplemented with 5% FCS.

OTUB1 knock-out cell line

A549 cell lines knockout for the expression of OTUB1 were obtained by CrispR-Cas9. The cells were transfected with a pSpCas9(BB)-2A-Puro (PX459) plasmid expressing the Cas9, the sgRNA and a puromycin resistance gene. The sgRNA sequence for OTUB1 are the following: TTGTGAATTCAGTGAAGCCC and GATTGCTGTGCAGAACCCTC. Twenty-four hours post-transfection cells were treated with puromycin antibiotic at 2µg/mL for 3 days. After puromycin treatment cells were isolated and mono-clonal cultures were tested for OTUB1 protein depletion by western blot.

Viruses

The A/WSN/33(H1N1) virus was produced by reverse genetics as described in [173]. The recombinant A/WSN/33 virus expressing mCitrine or the Nano-luciferase was produced by reverse genetics using a pPolI-PB2 plasmid with a sequence encoding the self-cleaving 2A peptide from porcine teschovirus followed by the mCitrine coding sequence, as described in [175].

Influenza viruses A/Centre/1003/2012(H3N2) and A/Bretagne/7608/2009($H1N1_{pdm09}$) were provided by the National Influenza Centre at the Institut Pasteur (Paris, France), and passaged 5 times on A549 cells at a MOI of 10^{-2}. After the 5th passage, the titres of adapted viruses produced on A549 had increased by more than 2 log. The hemagglutinin (HA) of the adapted viruses $H1N1_{pdm09}$ and H3N2 was sequenced. One mutation in the H3N2 HA (G460T) and two mutations in the $H1N1_{pdm09}$ HA (A517G and G834A) were identified. They were introduced into the HA-encoding plasmid in the A/Centre/1003/2012(H3N2) and H1N1A/Bretagne/7608/2009($H1N1_{pdm09}$) reverse genetics system, to produce A549-adapted H3N2 and $H1N1_{pmd09}$ viruses.

HT-GPCA and NLR retesting

HEK-293T cells were seeded in white 96-well plates at a concentration of $2,5x10^4$ cells per well. After 24 hours, cells were transfected using linear PEI with 300 ng of pSPICA-N2 and 100 ng of pSPICA-N1 expressing UPS factors. At 24 h post-transfection, cells were washed with 50 µl of PBS and lysed with 40 µl of Renilla Lysis Buffer (Promega, E2820) during 30 min. *Gaussia princeps* luciferase enzymatic activity was measured using a Berthold Centro LB960 luminometer by injecting 50 µl per well of luciferase substrate reagent (Promega, E2820) and counting luminescence during 10 seconds. Results were expressed as luciferase activity (RLU) or as a fold change normalised over the sum of controls, specified herein as Normalised Luminescence Ratio (NLR). For a given protein pair A/B, NLR = (Gluc1-A + Gluc2-B) / [(Gluc1-A + Gluc2) +(Gluc1 + Gluc2-B)].

Immunoprecipitation and western blot analysis

Immunoprecipitation

The PB2 coding sequence of $H1N1_{WSN}$ was cloned into the pCineo-3xFLAG vector allowing expression of the protein with a Nter-3xFlag tag. HEK-293T cells were plated in 6-well plates at $4x10^5$ cells per well the day before transfection with 0,5 µg of plasmid Flag-PB2 or empty vector (pCineo Flag) and 0,5µg of plasmid expressing Gluc1 tagged OTUB1 or PAN2. Cells were harvested after 24h and lysed by adding 100 µl of lysis buffer (50 mM Tris HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% TRITON X-100 and proteases inhibitors). After centrifugation at 12000 rpm for 20 min at 4°C, 80 µl of the cleared lysates were incubated overnight with 10 µl ANTI-FLAG M2® Magnetic Beads (Sigma) at 4°C. Beads were washed five times with lysis buffer and were resuspended with SDS sample buffer. Total proteins (input corresponding to 20%) and immunoprecipitated proteins on ANTI-FLAG magnetic beads were analysed by gel electrophoresis and Western-blotting. The PVDF membrane was incubated with rabbit anti-Gaussia luciferase (NEB, E8023S) as primary antibody and with goat HRP-conjugated anti-rabbit IgG as secondary antibody (Advansta). The Gluc1 tagged UPS proteins were revealed by chemioluminescence using WesternBrightTM SiriusTM (Advansta) and visualised by a LAS 4000 camera.

Western blots

Protein extracts were prepared in Laemmli buffer. Immunoblot membranes were incubated with primary antibodies directed against the *Gaussia* luciferase (E8023S, NEB), A/PR8/8/34 virions [194], GAPDH (Pierce) or OTUB1 (Abcam). They were revealed with secondary antibodies (GE Healthcare) or peroxydase-conjugated strepTactin (IBA) and with the ECL 2 substrate (Pierce). The chemiluminescence signals were acquired using the G-Box and the GeneSnap software (SynGene).

siRNA assays

Small interfering RNAs were purchased from Dharmacon (ON-TARGETplus SMARTpools and Non-targeting Control pool). A549 cells were transfected with 25 nM of siRNA using the DharmaFECT1 transfection reagent (Dharmacon). At 48 hours post transfection, cells were infected with the $H1N1_{WSN}$, (MOI of 10^{-4}) or adapted H3N2 or $H1N1_{pdm09}$ viruses (MOI of 10^{-3}) for 24 h. Plaque assays using MDCK-SIAT cells were performed as described in [195].

Efficiency of siRNA was controlled using expression plasmids for UPS factors fused to *Gaussia princeps* luciferase as described in [175], except for DCAF12L1 for which knock-down efficiency was controlled using qRT-PCR. For this factor cell lysates were harvested and subjected to the total RNA extraction using the RNeasy Mini Kit (Qiagen). RT-qPCR was conducted using the following primers: forward primer 5'-AATGCGCTCTACACCCACTG-3', reverse primer 5'- TTGACCAAAGGACCACTCACT-3' with the protocol of the Light Cycler RNA Amplification Kit SYBR Green 1 (Roche). The cellular GAPDH mRNA in the infected cells was also quantified as an internal control by using the above RT-qPCR. The DCAF12L1 siRNA decreased by 50 % the level of DCAF12L1 mRNA (not shown).

Cell viability was determined using trypan blue counting 48 hours post transfection with 25 nM of the siRNAs.

Mini-replicon assay

A549 cells were treated with siRNA. After 48 hours, cells were transfected using linear PEI with 150 ng of pCINeo plasmid expressing a DUB fused to strep, 25 ng of pCDNA3 plasmid expressing PB1, PB2, PA of $H1N1_{WSN}$, 50 ng of pCDNA3 plasmid expressing NP of $H1N1_{WSN}$ and 5 ng of pCINeo plasmid expressing a renilla protein and a pPR7 plasmid expressing a luciferase protein under the control of the viral polymerase. At 24 hours post-transfection, renilla and luciferase activity were measured using the Dual-Glo® Luciferase Assay System (Promega) as recommended.

Single cycle experiments

Small interfering RNAs were purchased from Dharmacon (ON-TARGETplus SMARTpools and Non-targeting Control pool). A549 cells were transfected with 25 nM of siRNA using the DharmaFECT1 transfection reagent (Dharmacon). At 48 hours post transfection, cells were infected with the $H1N1_{WSN}$at a MOI of 3 for 3, 6 and 9 hours. Western blots were performed as described above.

Over-expression-based experiments

Infection with mCitrine reporter virus

HEK-293T cells were seeded in white 96-well plates at a concentration of $3x10^4$ cells per well. After 24 hours, cells were transfected using linear PEI with 150 ng of pCINeo plasmid expressing an UPS factor fused to mCherry. At 24 h post-transfection, cells were infected with the WSN-PB2-2A-mCitrine virus at a MOI of 1 for 18h. Cells were then fixed with a 4% PFA solution and analysed by flow-cytometry (Attune NxT, Thermo Fisher Scientific).

Mini-replicon assay

A549 cells were seeded in white 96-well plates at a concentration of $3x10^4$ cells per well. After 24 hours, cells were transfected using linear PEI with 150 ng of pCINeo plasmid expressing a DUB fused to strep, 25 ng of pCDNA3 plasmid expressing PB1, PB2, PA of

$H1N1_{WSN}$, 50 ng of pCDNA3 plasmid expressing NP of $H1N1_{WSN}$ and 5 ng of pCINeo plasmid expressing a renilla protein and a pPR7 plasmid expressing a luciferase protein under the control of the viral polymerase. At 24 hours post-transfection, renilla and luciferase activity were measured using the Dual-Glo® Luciferase Assay System (Promega) as recommended.

mRNA quantification

To measure mRNA level, mRNA were first extracted with the RNeasy Mini Kit (Qiagen) according to recommendations. Then a reverse-transcription was performed with a SuperScript™ II Reverse Transcriptase (Invitrogen) and an oligo dT. Finally the cDNAs were used as templates for quantitative PCR (qPCR) using a LightCycler 480 (Roche) with the Luna® Universal qPCR Master Mix. Primers used for PB2 transcripts were described previously [169]. Plasmids containing the sequences of interest were used for calibration curves and absolute quantification. Primers for GAPDH mRNA can be provided upon request.

Immunofluorescence assays

A549 cells on coverslips were infected with the $H1N1_{pdm09}$ virus, at a MOI of 0.1, incubated with DMEM supplemented with 2% FCS, fixed with PBS- 4% paraformaldehyde for 20 min, and permeabilised with PBS-0.1% Triton X100 for 10 min. Cells were incubated with DAPI, anti-Otub1 (Abcam) and anti-PB2 primary antibodies and then Alexa-fluor 488 and 645 secondary antibodies. The samples were mounted in ProLong Gold mounting medium (Invitrogen) and analysed under a fluorescence microscope (Zeiss).

Statistics

Whisker-plot representation

The distributions of the relative luminescence values were represented by whisker-plot boxes. Whisker length corresponds to 1.5 times the interquartile range (IQR) that is equal to the

difference between the upper and lower quartiles (IQR=Q3-Q1). The upper whisker, defined by the third quartile (Q3) plus 1.5 times the interquartile range, *i.e.* Q3+1.5(IQR), was defined as the positive threshold (PT).

Determination of the confidence interval

To estimate the significance of a NLR value for a given protein-protein pair by comparison to the RRS sampling signal, a confidence interval was calculated for the RRS dataset, considering the estimated standard error SE and a confidence level of 99.73% (which corresponds in a normally distributed population to the mean +/- 3 standard deviations) by using the following expression: μ-t.SE + μ+t.SE where t is the critical value for α two-sided student test and for (n-1) degrees of freedom. We considered the NLR value of a new sample as statistically significantly different from the RRS if its value is larger than the upper confidence bound determined for the RRS dataset.

Hierarchical clustering

NLR values for each PB2 were normalised using Z-scores calculated as follows: Z-score = [NLR – mean (NLR for the PB2)] / standard deviation (NLR for the PB2). The clustering was made on Z-scores using Euclidian Distance parameter and word.D method with the R software [196].

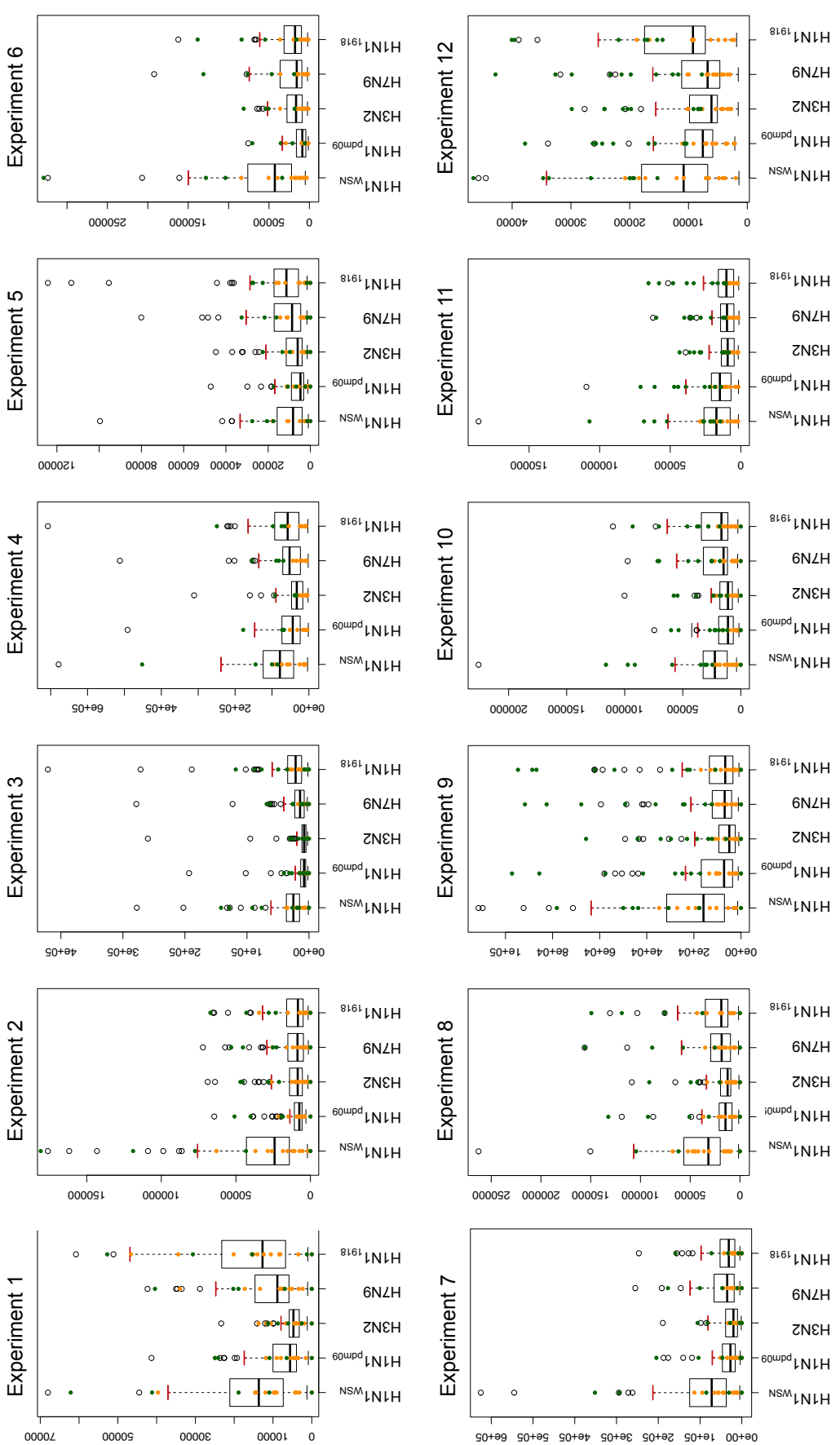
Annexes

Annexe 1

Screening of the UPS-dedicated library against the Gluc2 fragment of the *Gaussia princeps* luciferase using the HT-GPCA assay. The factors given the 23 highest values are presented. The validated interactions are indicated, after screening for interaction with influenza A PB2 proteins and human papillomavirus HPV16 E6 and E7 oncoproteins.

UPS factor	UPS clone	Isoform	Gluc2	Interaction validated with:		
				PB2	E6	E7
TRAF3IP2		2	1 312 329	X	X	X
TRAF3		1	80 430			X
UBB	stop at aa 78		68 874	X	X	X
KLHL25	Deletion 231-588		47 145			
NDFIP1		1	43 260			
TRIM22		1	39 917			X
RIPK1		1	34 723	X		
LRR1		1	33 945			
TRIM28		1	31 582			
TCEB2		1	31 148			
ZNF131	Deletion 1-80	2	28 795			
FBXO33			27 945			
SOCS2			26 090			
MPND		1	26 061	X		X
RNF26			22 124			
TRAF4		1	21 017			X
RNF175		1	20 544			
RAB40A			20 225			
BTRC		1	19 761	X		
XIAP			19 534			
UBE2A		1	19 437			X
CDC16		1	19 312			
FBXO7		1	18 940			

Annexe 2



UPS library screening. The 12 HT-GPCA experiments covering the screening of the whole UPS library are shown. In each experiment, whisker plots were generated from the luminescence values of the UPS-PB2 pairs. Outliers luminescence values, represented in circles, were selected as potential positive interactions. The RRS and PRS values (yellow and green dots respectively) were plotted in the whiskers plots afterwards to evaluate the accuracy of the outlier-based selection.

Annexe 3

Post-screening retesting of the PB2-UPS interacting pairs by NLR (Normalized Luminescence Ratio). Each interacting pair was compared to a threshold based on the NLR values of the RRS. Experiments were conducted 3 times and the number of positive scoring over the threshold is indicated. Interactions were validated if they scored positive 3 times over 3, for at least one PB2 protein.

	Number of positive scoring					UPS factors	
UPS	**H1N1$_{WSN}$**	**H1N1$_{pdm09}$**	**H3N2**	**H7N9**	**H1N1$_{1918}$**	**Validated**	**Not validated**
ABTB2	2	3	2	1	2	1	
ASB10	2	1	3	1	2	1	
ASB2	2	2	2	2	2		1
ASB4	1	1	3	1	3	1	
ASB6	2	1	2	1	1		1
ATXN3	1	3	2	2	2	1	
BAP1	2	2	2	1	1		1
BCL6B	2	2	2	1	2		1
BTBD12	3	3	3	3	3	1	
BTBD6	2	3	2	2	2	1	
BTRC	2	3	2	3	2	1	
C19orf40	0	0	0	0	0		1
CDC16	2	1	1	0	2		1
CDC20	2	2	1	2	1		1
CUL4B	2	2	1	2	1		1
DCAF11	3	3	2	2	2	1	
DCAF12L1	3	3	3	3	2	1	
DCAF4	2	1	1	1	2		1
DCAF4L1	2	2	2	1	2		1
DCAF4L2	2	2	2	1	2		1
DCAF5	2	1	1	1	1		1
DCAF7	1	2	1	0	1		1
DDB1	3	3	3	2	2	1	
FBXL12	2	2	2	1	2		1
FBXL18	2	1	2	1	1		1
FBXL19	3	2	2	2	1	1	
FBXL3	2	1	1	1	1		1
FBXO2	2	2	2	1	2		1
FBXO21	3	3	3	2	2	1	
FBXO44	3	3	3	2	2	1	
FBXW11	2	2	2	2	2		1
GLMN	1	2	2	2	2		1
GNB2	1	1	0	1	1		1
KBTBD5	3	3	3	3	3	1	
KCTD13	2	2	1	2	2		1
KLHL34	1	3	2	2	2	1	
LZTR1	2	1	2	2	2		1
MARCH10	3	3	3	3	2	1	

MGRN1	2	3	3	3	3	1	
MPND	2	3	2	1	2	1	
NEURL	1	2	2	0	2		1
OTUB1	1	3	2	2	2	1	
OTUD6A	2	2	3	3	2	1	
PAN2	2	3	2	3	2	1	
PATZ1	2	3	2	2	2	1	
PDZRN4	3	3	2	3	3	1	
PHC2	3	3	3	3	3	1	
PML	1	1	1	1	1		1
PSMD2	3	3	2	3	2	1	
RAB40B	2	1	1	1	1		1
RIPK1	2	3	3	2	3	1	
RNF10	3	3	3	3	2	1	
RNF111	0	1	1	2	1		1
RNF25	2	2	2	1	2		1
RNF34	3	3	3	3	3	1	
RNF40	2	3	2	1	2	1	
TNFAIP3	2	1	1	1	1		1
TRAF2	3	1	2	2	2	1	
TRAF3IP2	3	3	1	3	2	1	
TRAF5	3	2	3	2	3	1	
TRAF6	2	3	3	3	3	1	
TRIM15	3	3	3	3	2	1	
TRIM21	2	1	3	2	2	1	
TRIM3	1	2	2	1	2		1
TRIM32	3	3	3	2	2	1	
TRIM4	2	0	2	1	2		1
TRIM52	2	2	2	2	2		1
TRIM69	1	0	0	0	0		1
UBB	2	3	3	3	3	1	
UFD1L	0	1	0	0	0		1
USP25	2	3	2	2	2	1	
USP26	1	2	1	0	2		1
USP28	3	3	3	3	3	1	
VHL	3	3	2	3	2	1	
ZBTB10	1	0	1	0	1		1
ZBTB14	3	3	3	3	2	1	
ZBTB20	3	3	2	2	2	1	
ZBTB24	1	1	2	2	1		1
ZBTB43	2	1	1	1	1		1
ZBTB48	1	0	2	1	2		1
					Total	**42**	**38**

Annexe 4

UPS factors interacting with human papillomavirus HPV16 E6 and E7 oncoproteins. UPS families and domains are indicated.

SYMB	HPV16 E6	HPV16 E7	UPS Class & Family	Domain/motif
TRAF3IP2	■	■	other	TRAF binding domain
UBB	■	■	UB	ubiquitin
TRAF2		■	E3_RING	RING type zinc finger
TRAF3		■	E3_RING	RING type zinc finger
TRAF4		■	E3_RING	RING type zinc finger
TRAF5	■	■	E3_RING	RING type zinc finger
TRAF6	■		E3_RING	RING type zinc finger
MGRN1	■		E3_RING	RING type zinc finger
ZBTB20		■	E3_Cul3-CRL	BTB
ZBTB48		■	E3_Cul3-CRL	BTB
MPND		■	DUB	JAMM
ITCH	■		E3_HECT	HECT
UBE3A	■		E3_HECT	HECT
LNX4	■		E3_RING	RING type zinc finger PDZ domain
RNF25	■		E3_RING	RING type zinc finger
RNF40	■		E3_RING	RING type zinc finger
AIMP2	■		other	
TAX1BP1	■		other	
USP15	■		DUB	USP
UBE2A		■	E2	UBC
NEURL1		■	E3_RING	RING type zinc finger
RNF135		■	E3_RING	RING type zinc finger
SH3RF1		■	E3_RING	RING type zinc finger
TRIM22		■	E3_RING	RING type zinc finger
TRIM32		■	E3_RING	RING type zinc finger
RNF30		■	E3_RING	RING type zinc finger
TRIM72		■	E3_RING	RING type zinc finger
TRIM9		■	E3_RING	RING type zinc finger
ZNF219		■	E3_RING	RING type zinc finger
ZNF598		■	E3_RING	RING type zinc finger
DCAF15		■	E3_Cul4-CRL	WD repeats
BTBD15		■	E3_Cul3-CRL	BTB
KCTD13		■	E3_Cul3-CRL	BTB
NACC1		■	E3_Cul3-CRL	BTB
ZBTB32		■	E3_Cul3-CRL	BTB
ZBTB42		■	E3_Cul3-CRL	BTB
ZBTB43		■	E3_Cul3-CRL	BTB
ZBTB9		■	E3_Cul3-CRL	BTB
SHKBP1		■	E3_Cul3-CRL	BTB
TNFAIP1		■	E3_Cul3-CRL	BTB
USP26		■	DUB	USP
USP29		■	DUB	USP
USP33		■	DUB	USP

Annexe 5

UPS factors interacting with at least one PB2 protein. Interactions with human papillomavirus HPV16 E6 and E7 oncoproteins, UPS families and domains are indicated.

UPS	HPV16 interactor ?	UPS Class & Family	Domain/motif
ABTB2		E3-BTB	BTB
ASB10		E3_Cu2/5-CRL	Socs/VHL
ASB4		E3_Cu2/5-CRL	Socs/VHL
ATXN3		DUB	Josephin
BTBD6		E3-BTB	BTB
BTRC		E3_Cul1-CRL	F-box
DCAF11		E3_Cul4-CRL	WD repeats
DCAF12L1		E3_Cul4-CRL	WD repeats
DDB1		E3_Cul4-CRL	adapter
FBXL19		E3_Cul1-CRL	F-box
FBXO21		E3_Cul1-CRL	F-box
FBXO44		E3_Cul1-CRL	F-box
KBTBD5		E3_Cul3-CRL	BTB
KLHL34		E3_Cul3-CRL	BTB
MARCH10		E3_RING	RING type zinc finger
MGRN1	E6	E3_RING	RING type zinc finger
MPND	E7	DUB	JAMM
OTUB1		DUB	OTU
OTUD6A		DUB	OTU
PAN2		DUB	USP and WD repeats
PATZ1		E3_Cul3-CRL	BTB
PDZRN4		E3-BTB	BTB
PHC2		other: component of the PRC1-like complex	
PSMD2		proteasome	
RIPK1		other	kinase
RNF10		E3_RING	RING type zinc finger
RNF34		E3_RING	RING type zinc finger
RNF40	E6	E3_RING	RING type zinc finger
SLX4		E3-BTB	BTB
TRAF2	E7	E3_RING	RING type zinc finger
TRAF3IP2	E6 and E7	E3_RING	RING type zinc finger
TRAF5	E6 and E7	other	TRAF binding domain
TRAF6	E6	E3_RING	RING type zinc finger
TRIM15		E3_RING	RING type zinc finger
TRIM21		E3_RING	RING type zinc finger
TRIM32	E7	E3_RING	RING type zinc finger
UBB	E6 and E7	UB	ubiquitin
USP25		DUB	USP
USP28		DUB	USP
VHL		E3_Cul2-CRL	VHL
ZBTB14		E3-BTB	BTB
ZBTB20	E7	E3-BTB	BTB

Annexe 6

Relative viral titres produced upon siRNA mediated depletion of the UPS factors. Viral titres are calculated relative to titre obtained with non-target siRNA. P-values calculated from non parametric two-tailed student's-t test are as follows: 0.05 > * > 0.033 > ** > 0.002 > ***.

	$H1N1_{WSN}$	$H1N1_{pdm09}$	H3N2
ABTB2	0,17 *	1,41	1,03
ASB10	0,09 *	0,38 **	1,10
ASB4	0,1 *	0,86	0,39 *
ATXN3	0,08 *	0,15 **	0,78
BTBD6	0,1 **	0,47 **	0,65
BTRC	0,31 *	0,60 *	0,61
DCAF11	0,09 *	0,14 ***	0,56
DCAF12L1	0,04 *	0,34 *	0,22 **
DDB1	0,08 *	0,25**	0,34 **
FBXL19	0,08 **	0,6 **	2,37 **
FBXO21	0,29 *	0,73	0,63
FBXO44	0,02 *	0,29 ***	0,66
KLHL40	0,28 *	0,57 *	0,74
KLHL34	0,39	0,49 *	0,66
MARCH10	0,13 *	0,89	0,24 **
MGRN1	0,04 ***	1,54	0,25 **
MPND	0,48	0,92	0,62
OTUB1	0,26 *	0,40 **	0,84
OTUD6A	0,22 *	0,29 **	0,28 **
PAN2	0,24 *	0,23 **	0,33 **
PATZ1	0,16 **	1,06	2,06
PDZRN4	0,46	0,59	0,67
PSMD2	0,09 *	0,29 **	0,50
PHC2	0,05 *	0,23 **	0,43 *
RIPK1	0,25 *	0,92	0,38 **
RNF10	0,43	0,49 *	0,68
RNF34	0,33 *	0,41**	0,72
RNF40	0,43 *	2,44	0,36 *
SLX4	0,15 *	0,58 *	1,10
TRAF2	0,1 ***	0,26 **	0,74
TRAF3IP2	0,48	0,24 **	0,65
TRAF5	0,27 *	0,30 **	0,83
TRAF6	0,45	0,69	0,97
TRIM15	0,25 *	0,33 **	0,52
TRIM21	0,37 *	0,3 **	0,46
TRIM32	0,79	0,25 **	0,44
UBB	ND	ND	ND
USP25	0,38 *	0,79	0,22 **
USP28	0,50	0,49 *	0,21 **
VHL	0,32 *	0,68	0,95
ZBTB14	0,46	0,76	0,90
ZBTB20	0,52	0,53	0,55

Summary of the satatistically significant effects per strains

	$H1N1_{WSN}$
statistically significant UPS factors :	32
effect ≥ 10	13
10>effect≥4	8
effect < 4	11
max sat. significant decrease	50
min sat. significant decrease	2,33

	$H1N1_{pdm09}$	H3N2
statistically significant UPS factors :	26	13
effect ≥ 4	7	5
4>effect≥2	15	8
effect < 2	4	
max sat. significant decrease	7,14	4,76
min sat. significant decrease	1,74	2,33

Bibliography

1. Hershko A, Ciechanover A. The ubiquitin system. Annual review of biochemistry. 1998;67:425-79. Epub 1998/10/06. doi: 10.1146/annurev.biochem.67.1.425. PubMed PMID: 9759494.

2. Taniguchi T, Garcia-Higuera I, Andreassen PR, Gregory RC, Grompe M, D'Andrea AD. S-phase-specific interaction of the Fanconi anemia protein, FANCD2, with BRCA1 and RAD51. Blood. 2002;100(7):2414-20. Epub 2002/09/20. doi: 10.1182/blood-2002-01-0278. PubMed PMID: 12239151.

3. Li M, Brooks CL, Wu-Baer F, Chen D, Baer R, Gu W. Mono- versus polyubiquitination: differential control of p53 fate by Mdm2. Science. 2003;302(5652):1972-5. Epub 2003/12/13. doi: 10.1126/science.1091362. PubMed PMID: 14671306.

4. Hoege C, Pfander B, Moldovan GL, Pyrowolakis G, Jentsch S. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature. 2002;419(6903):135-41. Epub 2002/09/13. doi: 10.1038/nature00991. PubMed PMID: 12226657.

5. Nijman SM, Huang TT, Dirac AM, Brummelkamp TR, Kerkhoven RM, D'Andrea AD, et al. The deubiquitinating enzyme USP1 regulates the Fanconi anemia pathway. Mol Cell. 2005;17(3):331-9. Epub 2005/02/08. doi: 10.1016/j.molcel.2005.01.008. PubMed PMID: 15694335.

6. Palombella VJ, Rando OJ, Goldberg AL, Maniatis T. The ubiquitin-proteasome pathway is required for processing the NF-kappa B1 precursor protein and the activation of NF-kappa B. Cell. 1994;78(5):773-85. Epub 1994/09/09. PubMed PMID: 8087845.

7. Stegmeier F, Rape M, Draviam VM, Nalepa G, Sowa ME, Ang XL, et al. Anaphase initiation is regulated by antagonistic ubiquitination and deubiquitination activities. Nature. 2007;446(7138):876-81. Epub 2007/04/20. doi: 10.1038/nature05694. PubMed PMID: 17443180.

8. Gilberto S, Peter M. Dynamic ubiquitin signaling in cell cycle regulation. J Cell Biol. 2017. doi: 10.1083/jcb.201703170. PubMed PMID: 28684425.

9. Herhaus L, Al-Salihi M, Macartney T, Weidlich S, Sapkota GP. OTUB1 enhances TGFbeta signalling by inhibiting the ubiquitylation and degradation of active SMAD2/3. Nat Commun. 2013;4:2519. Epub 2013/09/28. doi: 10.1038/ncomms3519. PubMed PMID: 24071738; PubMed Central PMCID: PMCPMC3791481.

10. Randow F, Lehner PJ. Viral avoidance and exploitation of the ubiquitin system. Nat Cell Biol. 2009;11(5):527-U35. doi: 10.1038/ncb0509-527. PubMed PMID: WOS:000265640000005.

11. Komander D, Rape M. The ubiquitin code. Annual review of biochemistry. 2012;81:203-29. Epub 2012/04/25. doi: 10.1146/annurev-biochem-060310-170328. PubMed PMID: 22524316.

12. Smit JJ, Sixma TK. RBR E3-ligases at work. EMBO Rep. 2014;15(2):142-54. Epub 2014/01/29. doi: 10.1002/embr.201338166. PubMed PMID: 24469331; PubMed Central PMCID: PMCPMC3989860.

13. Yau R, Rape M. The increasing complexity of the ubiquitin code. Nat Cell Biol. 2016;18(6):579-86. Epub 2016/05/28. doi: 10.1038/ncb3358. PubMed PMID: 27230526.

14. Mevissen TET, Komander D. Mechanisms of Deubiquitinase Specificity and Regulation. Annual review of biochemistry. 2017;86:159-92. Epub 2017/05/13. doi: 10.1146/annurev-biochem-061516-044916. PubMed PMID: 28498721.

15. Saeki Y. Ubiquitin recognition by the proteasome. J Biochem. 2017. doi: 10.1093/jb/mvw091. PubMed PMID: 28069863.

16. Xu P, Duong DM, Seyfried NT, Cheng D, Xie Y, Robert J, et al. Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. Cell. 2009;137(1):133-45. Epub 2009/04/07. doi: 10.1016/j.cell.2009.01.041. PubMed PMID: 19345192; PubMed Central PMCID: PMCPMC2668214.

17. Jin L, Williamson A, Banerjee S, Philipp I, Rape M. Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. Cell. 2008;133(4):653-65. Epub 2008/05/20. doi: 10.1016/j.cell.2008.04.012. PubMed PMID: 18485873; PubMed Central PMCID: PMCPMC2696189.

18. Johnson ES, Ma PC, Ota IM, Varshavsky A. A proteolytic pathway that recognizes ubiquitin as a degradation signal. The Journal of biological chemistry. 1995;270(29):17442-56. Epub 1995/07/21. PubMed PMID: 7615550.

19. Saeki Y, Kudo T, Sone T, Kikuchi Y, Yokosawa H, Toh-e A, et al. Lysine 63-linked polyubiquitin chain may serve as a targeting signal for the 26S proteasome. The EMBO journal. 2009;28(4):359-71. Epub 2009/01/21. doi: 10.1038/emboj.2008.305. PubMed PMID: 19153599; PubMed Central PMCID: PMCPMC2646160.

20. Mukhopadhyay D, Riezman H. Proteasome-independent functions of ubiquitin in endocytosis and signaling. Science. 2007;315(5809):201-5. Epub 2007/01/16. doi: 10.1126/science.1127085. PubMed PMID: 17218518.

21. Ordureau A, Heo JM, Duda DM, Paulo JA, Olszewski JL, Yanishevski D, et al. Defining roles of PARKIN and ubiquitin phosphorylation by PINK1 in mitochondrial quality control using a ubiquitin replacement strategy. Proc Natl Acad Sci U S A. 2015;112(21):6637-42. Epub 2015/05/15. doi: 10.1073/pnas.1506593112. PubMed PMID: 25969509; PubMed Central PMCID: PMCPMC4450373.

22. Wang Y, Serricchio M, Jauregui M, Shanbhag R, Stoltz T, Di Paolo CT, et al. Deubiquitinating enzymes regulate PARK2-mediated mitophagy. Autophagy. 2015;11(4):595-606. Epub 2015/04/29. doi: 10.1080/15548627.2015.1034408. PubMed PMID: 25915564; PubMed Central PMCID: PMCPMC4502823.

23. d'Azzo A, Bongiovanni A, Nastasi T. E3 ubiquitin ligases as regulators of membrane protein trafficking and degradation. Traffic. 2005;6(6):429-41. doi: 10.1111/j.1600-0854.2005.00294.x. PubMed PMID: WOS:000229137300001.

24. Al-Hakim A, Escribano-Diaz C, Landry MC, O'Donnell L, Panier S, Szilard RK, et al. The ubiquitous role of ubiquitin in the DNA damage response. DNA repair. 2010;9(12):1229-40. Epub 2010/11/09. doi: 10.1016/j.dnarep.2010.09.011. PubMed PMID: 21056014.

25. Dynek JN, Goncharov T, Dueber EC, Fedorova AV, Izrael-Tomasevic A, Phu L, et al. c-IAP1 and UbcH5 promote K11-linked polyubiquitination of RIP1 in TNF signalling. The EMBO journal. 2010;29(24):4198-209. Epub 2010/11/30. doi: 10.1038/emboj.2010.300. PubMed PMID: 21113135; PubMed Central PMCID: PMCPMC3018797.

26. Hymowitz SG, Wertz IE. A20: from ubiquitin editing to tumour suppression. Nature reviews Cancer. 2010;10(5):332-41. Epub 2010/04/13. doi: 10.1038/nrc2775. PubMed PMID: 20383180.

27. Kayagaki N, Phung Q, Chan S, Chaudhari R, Quan C, O'Rourke KM, et al. DUBA: a deubiquitinase that regulates type I interferon production. Science. 2007;318(5856):1628-32. Epub 2007/11/10. doi: 10.1126/science.1145918. PubMed PMID: 17991829.

28. Gatti M, Pinato S, Maiolica A, Rocchio F, Prato MG, Aebersold R, et al. RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage. Cell Rep. 2015;10(2):226-38. Epub 2015/01/13. doi: 10.1016/j.celrep.2014.12.021. PubMed PMID: 25578731.

29. Wang Q, Liu X, Cui Y, Tang Y, Chen W, Li S, et al. The E3 ubiquitin ligase AMFR and INSIG1 bridge the activation of TBK1 kinase by modifying the adaptor STING. Immunity. 2014;41(6):919-33. Epub 2014/12/20. doi: 10.1016/j.immuni.2014.11.011. PubMed PMID: 25526307.

30. Tauriello DV, Maurice MM. The various roles of ubiquitin in Wnt pathway regulation. Cell cycle (Georgetown, Tex). 2010;9(18):3700-9. Epub 2010/10/12. doi: 10.4161/cc.9.18.13204. PubMed PMID: 20930545; PubMed Central PMCID: PMCPMC3047798.

31. Fei C, Li Z, Li C, Chen Y, Chen Z, He X, et al. Smurf1-mediated Lys29-linked nonproteolytic polyubiquitination of axin negatively regulates Wnt/beta-catenin signaling. Molecular and cellular biology. 2013;33(20):4095-105. Epub 2013/08/21. doi: 10.1128/mcb.00418-13. PubMed PMID: 23959799; PubMed Central PMCID: PMCPMC3811687.

32. Yuan WC, Lee YR, Lin SY, Chang LY, Tan YP, Hung CC, et al. K33-Linked Polyubiquitination of Coronin 7 by Cul3-KLHL20 Ubiquitin E3 Ligase Regulates Protein Trafficking. Mol Cell. 2014;54(4):586-600. Epub 2014/04/29. doi: 10.1016/j.molcel.2014.03.035. PubMed PMID: 24768539.

33. Herhaus L, Dikic I. Expanding the ubiquitin code through post-translational modification. EMBO Rep. 2015;16(9):1071-83. Epub 2015/08/14. doi: 10.15252/embr.201540891. PubMed PMID: 26268526; PubMed Central PMCID: PMCPMC4576978.

34. Hutchins AP, Liu S, Diez D, Miranda-Saavedra D. The repertoires of ubiquitinating and deubiquitinating enzymes in eukaryotic genomes. Molecular biology and evolution. 2013;30(5):1172-87. Epub 2013/02/09. doi: 10.1093/molbev/mst022. PubMed PMID: 23393154; PubMed Central PMCID: PMCPMC3670738.

35. Gao T, Liu Z, Wang Y, Cheng H, Yang Q, Guo A, et al. UUCD: a family-based database of ubiquitin and ubiquitin-like conjugation. Nucleic acids research. 2013;41(Database issue):D445-51. Epub 2012/11/23. doi: 10.1093/nar/gks1103. PubMed PMID: 23172288; PubMed Central PMCID: PMCPMC3531133.

36. Eletr ZM, Huang DT, Duda DM, Schulman BA, Kuhlman B. E2 conjugating enzymes must disengage from their E1 enzymes before E3-dependent ubiquitin and ubiquitin-like transfer. Nat Struct Mol Biol. 2005;12(10):933-4. Epub 2005/09/06. doi: 10.1038/nsmb984. PubMed PMID: 16142244.

37. Stewart MD, Ritterhoff T, Klevit RE, Brzovic PS. E2 enzymes: more than just middle men. Cell research. 2016;26(4):423-40. Epub 2016/03/24. doi: 10.1038/cr.2016.35. PubMed PMID: 27002219; PubMed Central PMCID: PMCPMC4822130.

38. Middleton AJ, Wright JD, Day CL. Regulation of E2s: A Role for Additional Ubiquitin Binding Sites? Journal of molecular biology. 2017. Epub 2017/06/20. doi: 10.1016/j.jmb.2017.06.008. PubMed PMID: 28625848.

39. Deshaies RJ, Joazeiro CA. RING domain E3 ubiquitin ligases. Annual review of biochemistry. 2009;78:399-434. Epub 2009/06/06. doi: 10.1146/annurev.biochem.78.101807.093809. PubMed PMID: 19489725.

40. Zheng N, Shabek N. Ubiquitin Ligases: Structure, Function, and Regulation. Annual review of biochemistry. 2017;86:129-57. Epub 2017/04/05. doi: 10.1146/annurev-biochem-060815-014922. PubMed PMID: 28375744.

41. Ebisawa T, Fukuchi M, Murakami G, Chiba T, Tanaka K, Imamura T, et al. Smurf1 interacts with transforming growth factor-beta type I receptor through Smad7 and induces receptor degradation. The Journal of biological chemistry. 2001;276(16):12477-80. Epub 2001/03/30. doi: 10.1074/jbc.C100008200. PubMed PMID: 11278251.

42. Sakata E, Yamaguchi Y, Miyauchi Y, Iwai K, Chiba T, Saeki Y, et al. Direct interactions between NEDD8 and ubiquitin E2 conjugating enzymes upregulate cullin-based E3 ligase activity. Nat Struct Mol Biol. 2007;14(2):167-8. Epub 2007/01/09. doi: 10.1038/nsmb1191. PubMed PMID: 17206147.

43. Morreale FE, Walden H. Types of Ubiquitin Ligases. Cell. 2016;165(1):248-.e1. Epub 2016/03/26. doi: 10.1016/j.cell.2016.03.003. PubMed PMID: 27015313.

44. Abdul Rehman SA, Kristariyanto YA, Choi SY, Nkosi PJ, Weidlich S, Labib K, et al. MINDY-1 Is a Member of an Evolutionarily Conserved and Structurally Distinct New Family of Deubiquitinating Enzymes. Mol Cell. 2016;63(1):146-55. Epub 2016/06/14. doi: 10.1016/j.molcel.2016.05.009. PubMed PMID: 27292798; PubMed Central PMCID: PMCPMC4942677.

45. Clague MJ, Barsukov I, Coulson JM, Liu H, Rigden DJ, Urbe S. Deubiquitylases from genes to organism. Physiol Rev. 2013;93(3):1289-315. doi: 10.1152/physrev.00002.2013. PubMed PMID: 23899565.

46. Wertz IE, O'Rourke KM, Zhou H, Eby M, Aravind L, Seshagiri S, et al. De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling. Nature. 2004;430(7000):694-9. Epub 2004/07/20. doi: 10.1038/nature02794. PubMed PMID: 15258597.

47. Komander D, Clague MJ, Urbe S. Breaking the chains: structure and function of the deubiquitinases. Nature reviews Molecular cell biology. 2009;10(8):550-63. Epub 2009/07/25. doi: 10.1038/nrm2731. PubMed PMID: 19626045.

48. Edelmann MJ, Iphofer A, Akutsu M, Altun M, di Gleria K, Kramer HB, et al. Structural basis and specificity of human otubain 1-mediated deubiquitination. Biochem J. 2009;418(2):379-90. Epub 2008/10/29. doi: 10.1042/bj20081318. PubMed PMID: 18954305.

49. Mevissen TE, Hospenthal MK, Geurink PP, Elliott PR, Akutsu M, Arnaudo N, et al. OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. Cell. 2013;154(1):169-84. Epub 2013/07/06. doi: 10.1016/j.cell.2013.05.046. PubMed PMID: 23827681; PubMed Central PMCID: PMCPMC3705208.

50. Komander D, Lord CJ, Scheel H, Swift S, Hofmann K, Ashworth A, et al. The structure of the CYLD USP domain explains its specificity for Lys63-linked polyubiquitin and reveals a B box module. Mol Cell. 2008;29(4):451-64. Epub 2008/03/04. doi: 10.1016/j.molcel.2007.12.018. PubMed PMID: 18313383.

51. Cooper EM, Cutcliffe C, Kristiansen TZ, Pandey A, Pickart CM, Cohen RE. K63-specific deubiquitination by two JAMM/MPN+ complexes: BRISC-associated Brcc36 and proteasomal Poh1. The EMBO journal. 2009;28(6):621-31. Epub 2009/02/14. doi: 10.1038/emboj.2009.27. PubMed PMID: 19214193; PubMed Central PMCID: PMCPMC2666030.

52. Sato Y, Yoshikawa A, Yamagata A, Mimura H, Yamashita M, Ookata K, et al. Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. Nature. 2008;455(7211):358-62. Epub 2008/09/02. doi: 10.1038/nature07254. PubMed PMID: 18758443.

53. Reiley W, Zhang M, Wu X, Granger E, Sun SC. Regulation of the deubiquitinating enzyme CYLD by IkappaB kinase gamma-dependent phosphorylation. Molecular and cellular biology. 2005;25(10):3886-95. Epub 2005/05/05. doi: 10.1128/mcb.25.10.3886-3895.2005. PubMed PMID: 15870263; PubMed Central PMCID: PMCPMC1087725.

54. Todi SV, Winborn BJ, Scaglione KM, Blount JR, Travis SM, Paulson HL. Ubiquitination directly enhances activity of the deubiquitinating enzyme ataxin-3. The EMBO journal. 2009;28(4):372-82. Epub 2009/01/21. doi: 10.1038/emboj.2008.289. PubMed PMID: 19153604; PubMed Central PMCID: PMCPMC2646149.

55. Yao T, Song L, Jin J, Cai Y, Takahashi H, Swanson SK, et al. Distinct modes of regulation of the Uch37 deubiquitinating enzyme in the proteasome and in the Ino80 chromatin-remodeling complex. Mol Cell. 2008;31(6):909-17. Epub 2008/10/17. doi: 10.1016/j.molcel.2008.08.027. PubMed PMID: 18922472; PubMed Central PMCID: PMCPMC2577292.

56. Row PE, Prior IA, McCullough J, Clague MJ, Urbe S. The ubiquitin isopeptidase UBPY regulates endosomal ubiquitin dynamics and is essential for receptor down-regulation. The Journal of biological chemistry. 2006;281(18):12618-24. Epub 2006/03/08. doi: 10.1074/jbc.M512615200. PubMed PMID: 16520378.

57. Wodrich H, Henaff D, Jammart B, Segura-Morales C, Seelmeir S, Coux O, et al. A capsid-encoded PPxY-motif facilitates adenovirus entry. PLoS pathogens. 2010;6(3):e1000808. Epub 2010/03/25. doi: 10.1371/journal.ppat.1000808. PubMed PMID: 20333243; PubMed Central PMCID: PMCPMC2841620.

58. Yan Z, Zak R, Luxton GW, Ritchie TC, Bantel-Schaal U, Engelhardt JF. Ubiquitination of both adeno-associated virus type 2 and 5 capsid proteins affects the transduction efficiency of recombinant vectors. Journal of virology. 2002;76(5):2043-53. Epub 2002/02/12. PubMed PMID: 11836382; PubMed Central PMCID: PMCPMC135943.

59. Sarkari F, Sanchez-Alcaraz T, Wang S, Holowaty MN, Sheng Y, Frappier L. EBNA1-mediated recruitment of a histone H2B deubiquitylating complex to the Epstein-Barr virus latent origin of DNA replication. PLoS pathogens. 2009;5(10):e1000624. Epub 2009/10/17. doi: 10.1371/journal.ppat.1000624. PubMed PMID: 19834552; PubMed Central PMCID: PMCPMC2757719.

60. Han Z, Sagum CA, Takizawa F, Ruthel G, Berry CT, Kong J, et al. Ubiquitin Ligase WWP1 Interacts with Ebola Virus VP40 to Regulate Egress. Journal of virology. 2017. Epub 2017/08/05. doi: 10.1128/jvi.00812-17. PubMed PMID: 28768865.

61. Davis ME, Gack MU. Ubiquitination in the antiviral immune response. Virology. 2015;479-480:52-65. Epub 2015/03/11. doi: 10.1016/j.virol.2015.02.033. PubMed PMID: 25753787; PubMed Central PMCID: PMCPMC4774549.

62. Malynn BA, Ma A. Ubiquitin makes its mark on immune regulation. Immunity. 2010;33(6):843-52. Epub 2010/12/21. doi: 10.1016/j.immuni.2010.12.007. PubMed PMID: 21168777; PubMed Central PMCID: PMCPMC3030984.

63. Graff JW, Ettayebi K, Hardy ME. Rotavirus NSP1 inhibits NFkappaB activation by inducing proteasome-dependent degradation of beta-TrCP: a novel mechanism of IFN antagonism. PLoS pathogens. 2009;5(1):e1000280. Epub 2009/01/31. doi: 10.1371/journal.ppat.1000280. PubMed PMID: 19180189; PubMed Central PMCID: PMCPMC2627925.

64. Lutz LM, Pace CR, Arnold MM. Rotavirus NSP1 Associates with Components of the Cullin RING Ligase Family of E3 Ubiquitin Ligases. Journal of virology. 2016;90(13):6036-48. Epub 2016/04/22. doi: 10.1128/jvi.00704-16. PubMed PMID: 27099313; PubMed Central PMCID: PMCPMC4907217.

65. McLaughlin-Drubin ME, Munger K. The human papillomavirus E7 oncoprotein. Virology. 2009;384(2):335-44. Epub 2008/11/15. doi: 10.1016/j.virol.2008.10.006. PubMed PMID: 19007963; PubMed Central PMCID: PMCPMC2661820.

66. Martinez-Zapien D, Ruiz FX, Poirson J, Mitschler A, Ramirez J, Forster A, et al. Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. Nature. 2016;529(7587):541-5. Epub 2016/01/21. doi: 10.1038/nature16481. PubMed PMID: 26789255; PubMed Central PMCID: PMCPMC4853763.

67. Gustin JK, Moses AV, Fruh K, Douglas JL. Viral takeover of the host ubiquitin system. Front Microbiol. 2011;2:161. Epub 2011/08/19. doi: 10.3389/fmicb.2011.00161. PubMed PMID: 21847386; PubMed Central PMCID: PMCPMC3147166.

68. Barry M, Fruh K. Viral modulators of cullin RING ubiquitin ligases: culling the host defense. Science's STKE : signal transduction knowledge environment. 2006;2006(335):pe21. doi: 10.1126/stke.3352006pe21. PubMed PMID: MEDLINE:16705129.

69. Calistri A, Munegato D, Carli I, Parolin C, Palu G. The ubiquitin-conjugating system: multiple roles in viral replication and infection. Cells. 2014;3(2):386-417. Epub 2014/05/09. doi: 10.3390/cells3020386. PubMed PMID: 24805990; PubMed Central PMCID: PMCPMC4092849.

70. Wang S, Wang K, Li J, Zheng C. Herpes simplex virus 1 ubiquitin-specific protease UL36 inhibits beta interferon production by deubiquitinating TRAF3. Journal of virology. 2013;87(21):11851-60. Epub 2013/08/30. doi: 10.1128/jvi.01211-13. PubMed PMID: 23986588; PubMed Central PMCID: PMCPMC3807349.

71. Knipe D, Howley P. Fields Virology2013.

72. Hause BM, Ducatez M, Collin EA, Ran Z, Liu R, Sheng Z, et al. Isolation of a novel swine influenza virus from Oklahoma in 2011 which is distantly related to human influenza C viruses. PLoS pathogens. 2013;9(2):e1003176. Epub 2013/02/15. doi: 10.1371/journal.ppat.1003176. PubMed PMID: 23408893; PubMed Central PMCID: PMCPMC3567177.

73. Ducatez MF, Pelletier C, Meyer G. Influenza D virus in cattle, France, 2011-2014. Emerging infectious diseases. 2015;21(2):368-71. Epub 2015/01/30. doi: 10.3201/eid2102.141449. PubMed PMID: 25628038; PubMed Central PMCID: PMCPMC4313661.

74. Mehle A. Unusual influenza A viruses in bats. Viruses. 2014;6(9):3438-49. Epub 2014/09/27. doi: 10.3390/v6093438. PubMed PMID: 25256392; PubMed Central PMCID: PMCPMC4189031.

75. Yoon SW, Webby RJ, Webster RG. Evolution and ecology of influenza A viruses. Curr Top Microbiol Immunol. 2014;385:359-75. Epub 2014/07/06. doi: 10.1007/82_2014_396. PubMed PMID: 24990620.

76. Taubenberger JK, Kash JC. Influenza virus evolution, host adaptation, and pandemic formation. Cell Host Microbe. 2010;7(6):440-51. Epub 2010/06/15. doi: 10.1016/j.chom.2010.05.009. PubMed PMID: 20542248; PubMed Central PMCID: PMCPMC2892379.

77. Organization WH. http://www.who.int/topics/influenza/en/.

78. Prager F, Wei D, Rose A. Total Economic Consequences of an Influenza Outbreak in the United States. Risk analysis : an official publication of the Society for Risk Analysis. 2017;37(1):4-19. Epub 2016/05/24. doi: 10.1111/risa.12625. PubMed PMID: 27214756.

79. Drake JW. Rates of spontaneous mutation among RNA viruses. Proc Natl Acad Sci U S A. 1993;90(9):4171-5. Epub 1993/05/01. PubMed PMID: 8387212; PubMed Central PMCID: PMCPMC46468.

80. Holmes EC. RNA virus genomics: a world of possibilities. The Journal of clinical investigation. 2009;119(9):2488-95. Epub 2009/09/05. doi: 10.1172/jci38050. PubMed PMID: 19729846; PubMed Central PMCID: PMCPMC2735898.

81. Dugan VG, Chen R, Spiro DJ, Sengamalay N, Zaborsky J, Ghedin E, et al. The evolutionary genetics and emergence of avian influenza viruses in wild birds. PLoS pathogens. 2008;4(5):e1000076. Epub 2008/06/03. doi: 10.1371/journal.ppat.1000076. PubMed PMID: 18516303; PubMed Central PMCID: PMCPMC2387073.

82. Holmes EC, Ghedin E, Miller N, Taylor J, Bao Y, St George K, et al. Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. PLoS biology. 2005;3(9):e300. Epub 2005/07/20. doi: 10.1371/journal.pbio.0030300. PubMed PMID: 16026181; PubMed Central PMCID: PMCPMC1180517.

83. Nelson MI, Viboud C, Simonsen L, Bennett RT, Griesemer SB, St George K, et al. Multiple reassortment events in the evolutionary history of H1N1 influenza A virus since 1918. PLoS pathogens. 2008;4(2):e1000012. Epub 2008/05/09. doi: 10.1371/journal.ppat.1000012. PubMed PMID: 18463694; PubMed Central PMCID: PMCPMC2262849.

84. Taubenberger JK, Morens DM. Pandemic influenza--including a risk assessment of H5N1. Revue scientifique et technique (International Office of Epizootics). 2009;28(1):187-202. Epub 2009/07/22. PubMed PMID: 19618626; PubMed Central PMCID: PMCPMC2720801.

85. Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature. 2009;459(7250):1122-5. Epub 2009/06/12. doi: 10.1038/nature08182. PubMed PMID: 19516283.

86. Connor RJ, Kawaoka Y, Webster RG, Paulson JC. Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. Virology. 1994;205(1):17-23. Epub 1994/11/15. doi: 10.1006/viro.1994.1615. PubMed PMID: 7975212.

87. Ito T, Couceiro JN, Kelm S, Baum LG, Krauss S, Castrucci MR, et al. Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. Journal of virology. 1998;72(9):7367-73. Epub 1998/08/08. PubMed PMID: 9696833; PubMed Central PMCID: PMCPMC109961.

88. Morens DM, Taubenberger JK, Fauci AS. The persistent legacy of the 1918 influenza virus. The New England journal of medicine. 2009;361(3):225-9. Epub 2009/07/01. doi: 10.1056/NEJMp0904819. PubMed PMID: 19564629; PubMed Central PMCID: PMCPMC2749954.

89. Shaw ML, Stone KL, Colangelo CM, Gulcicek EE, Palese P. Cellular proteins in influenza virus particles. PLoS pathogens. 2008;4(6):e1000085. Epub 2008/06/07. doi: 10.1371/journal.ppat.1000085. PubMed PMID: 18535660; PubMed Central PMCID: PMCPMC2390764.

90. Hutchinson EC, Denham EM, Thomas B, Trudgian DC, Hester SS, Ridlova G, et al. Mapping the phosphoproteome of influenza A and B viruses by mass spectrometry. PLoS pathogens. 2012;8(11):e1002993. Epub 2012/11/13. doi: 10.1371/journal.ppat.1002993. PubMed PMID: 23144613; PubMed Central PMCID: PMCPMC3493474.

91. Yamayoshi S, Watanabe M, Goto H, Kawaoka Y. Identification of a Novel Viral Protein Expressed from the PB2 Segment of Influenza A Virus. Journal of virology. 2015;90(1):444-56. Epub 2015/10/23. doi: 10.1128/jvi.02175-15. PubMed PMID: 26491155; PubMed Central PMCID: PMCPMC4702538.

92. Matlin KS, Reggio H, Helenius A, Simons K. Infectious entry pathway of influenza virus in a canine kidney cell line. J Cell Biol. 1981;91(3 Pt 1):601-13. Epub 1981/12/01. PubMed PMID: 7328111; PubMed Central PMCID: PMCPMC2112819.

93. de Vries E, Tscherne DM, Wienholts MJ, Cobos-Jimenez V, Scholte F, Garcia-Sastre A, et al. Dissection of the influenza A virus endocytic routes reveals macropinocytosis as an alternative entry pathway. PLoS pathogens. 2011;7(3):e1001329. Epub 2011/04/13. doi: 10.1371/journal.ppat.1001329. PubMed PMID: 21483486; PubMed Central PMCID: PMCPMC3068995.

94. Sieczkarski SB, Whittaker GR. Influenza virus can enter and infect cells in the absence of clathrin-mediated endocytosis. Journal of virology. 2002;76(20):10455-64. Epub 2002/09/20. PubMed PMID: 12239322; PubMed Central PMCID: PMCPMC136567.

95. Fontana J, Steven AC. At low pH, influenza virus matrix protein M1 undergoes a conformational change prior to dissociating from the membrane. Journal of virology. 2013;87(10):5621-8. Epub 2013/03/08. doi: 10.1128/jvi.00276-13. PubMed PMID: 23468509; PubMed Central PMCID: PMCPMC3648175.

96. Helenius A. Unpacking the incoming influenza virus. Cell. 1992;69(4):577-8. Epub 1992/05/15. PubMed PMID: 1375129.

97. Hutchinson EC, Fodor E. Nuclear import of the influenza A virus transcriptional machinery. Vaccine. 2012;30(51):7353-8. Epub 2012/06/02. doi: 10.1016/j.vaccine.2012.04.085. PubMed PMID: 22652398.

98. Martin-Benito J, Ortin J. Influenza virus transcription and replication. Advances in virus research. 2013;87:113-37. Epub 2013/07/03. doi: 10.1016/b978-0-12-407698-3.00004-1. PubMed PMID: 23809922.

99. Neumann G, Noda T, Kawaoka Y. Emergence and pandemic potential of swine-origin H1N1 influenza virus. Nature. 2009;459(7249):931-9. Epub 2009/06/16. doi: 10.1038/nature08157. PubMed PMID: 19525932; PubMed Central PMCID: PMCPMC2873852.

100. Fodor E. The RNA polymerase of influenza a virus: mechanisms of viral transcription and replication. Acta virologica. 2013;57(2):113-22. Epub 2013/04/23. PubMed PMID: 23600869.

101. Resa-Infante P, Jorba N, Coloma R, Ortin J. The influenza virus RNA synthesis machine: advances in its structure and function. RNA biology. 2011;8(2):207-15. Epub 2011/03/02. PubMed PMID: 21358279; PubMed Central PMCID: PMCPMC3127100.

102. Biswas SK, Nayak DP. Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1. Journal of virology. 1994;68(3):1819-26. Epub 1994/03/01. PubMed PMID: 8107244; PubMed Central PMCID: PMCPMC236644.

103. Gerber M, Isel C, Moules V, Marquet R. Selective packaging of the influenza A genome and consequences for genetic reassortment. Trends in microbiology. 2014;22(8):446-55. Epub 2014/05/07. doi: 10.1016/j.tim.2014.04.001. PubMed PMID: 24798745.

104. Air GM. Influenza neuraminidase. Influenza and other respiratory viruses. 2012;6(4):245-56. Epub 2011/11/17. doi: 10.1111/j.1750-2659.2011.00304.x. PubMed PMID: 22085243; PubMed Central PMCID: PMCPMC3290697.

105. Garten W, Braden C, Arendt A, Peitsch C, Baron J, Lu Y, et al. Influenza virus activating host proteases: Identification, localization and inhibitors as potential therapeutics. European journal of cell biology. 2015;94(7-9):375-83. Epub 2015/06/23. doi: 10.1016/j.ejcb.2015.05.013. PubMed PMID: 26095298.

106. Hengrung N, El Omari K, Serna Martin I, Vreede FT, Cusack S, Rambo RP, et al. Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. Nature. 2015;527(7576):114-7. Epub 2015/10/28. doi: 10.1038/nature15525. PubMed PMID: 26503046; PubMed Central PMCID: PMCPMC4783868.

107. Reich S, Guilligay D, Pflug A, Malet H, Berger I, Crepin T, et al. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature. 2014;516(7531):361-6. Epub 2014/11/20. doi: 10.1038/nature14009. PubMed PMID: 25409151.

108. Pflug A, Guilligay D, Reich S, Cusack S. Structure of influenza A polymerase bound to the viral RNA promoter. Nature. 2014;516(7531):355-60. Epub 2014/11/20. doi: 10.1038/nature14008. PubMed PMID: 25409142.

109. Thierry E, Guilligay D, Kosinski J, Bock T, Gaudon S, Round A, et al. Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. Mol Cell. 2015. Epub 2015/12/30. doi: 10.1016/j.molcel.2015.11.016. PubMed PMID: 26711008.

110. Biquand E, Demeret C. Structure resolution of the trimeric RNA-dependent RNA polymerase of influenza viruses: impact on our understanding of polymerase interactions with host and viral factors. Virologie. 2016;20(6):32-48. doi: 10.1684/vir.2016.0671.

111. Boivin S, Cusack S, Ruigrok RW, Hart DJ. Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. The Journal of biological chemistry. 2010;285(37):28411-7. Epub 2010/06/12. doi: 10.1074/jbc.R110.117531. PubMed PMID: 20538599; PubMed Central PMCID: PMCPMC2937865.

112. Subbarao K, Klimov A, Katz J, Regnery H, Lim W, Hall H, et al. Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. Science. 1998;279(5349):393-6. Epub 1998/02/07. PubMed PMID: 9430591.

113. Long JS, Giotis ES, Moncorge O, Frise R, Mistry B, James J, et al. Species difference in ANP32A underlies influenza A virus polymerase host restriction. Nature. 2016;529(7584):101-4. Epub 2016/01/08. doi: 10.1038/nature16474. PubMed PMID: 26738596; PubMed Central PMCID: PMCPMC4710677.

114. Eisfeld AJ, Neumann G, Kawaoka Y. At the centre: influenza A virus ribonucleoproteins. Nature reviews Microbiology. 2015;13(1):28-41. Epub 2014/11/25. doi: 10.1038/nrmicro3367. PubMed PMID: 25417656.

115. Resa-Infante P, Jorba N, Zamarreno N, Fernandez Y, Juarez S, Ortin J. The host-dependent interaction of alpha-importins with influenza PB2 polymerase subunit is required for virus RNA replication. PLoS One. 2008;3(12):e3904. doi: 10.1371/journal.pone.0003904. PubMed PMID: 19066626; PubMed Central PMCID: PMCPMC2588535.

116. Hudjetz B, Gabriel G. Human-like PB2 627K influenza virus polymerase activity is regulated by importin-alpha1 and -alpha7. PLoS pathogens. 2012;8(1):e1002488. Epub 2012/01/26. doi: 10.1371/journal.ppat.1002488. PubMed PMID: 22275867; PubMed Central PMCID: PMCPMC3262014.

117. York A, Hutchinson EC, Fodor E. Interactome analysis of the influenza A virus transcription/replication machinery identifies protein phosphatase 6 as a cellular factor required for efficient virus replication. Journal of virology. 2014;88(22):13284-99. Epub 2014/09/05. doi: 10.1128/jvi.01813-14. PubMed PMID: 25187537; PubMed Central PMCID: PMCPMC4249064.

118. Munier S, Rolland T, Diot C, Jacob Y, Naffakh N. Exploration of Binary Virus-Host Interactions Using an Infectious Protein Complementation Assay. Mol Cell Proteomics. 2013;12(10):2845-55. doi: 10.1074/mcp.M113.028688. PubMed PMID: WOS:000330537000014.

119. Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM, Wu L, et al. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. Cell. 2009;139(7):1255-67. Epub 2010/01/13. doi: 10.1016/j.cell.2009.12.018. PubMed PMID: 20064372; PubMed Central PMCID: PMCPMC2892837.

120. Watanabe T, Kawakami E, Shoemaker JE, Lopes TJ, Matsuoka Y, Tomita Y, et al. Influenza virus-host interactome screen as a platform for antiviral drug development. Cell Host Microbe. 2014;16(6):795-805. Epub 2014/12/04. doi: 10.1016/j.chom.2014.11.002. PubMed PMID: 25464832; PubMed Central PMCID: PMCPMC4451456.

121. Bradel-Tretheway BG, Mattiacio JL, Krasnoselsky A, Stevenson C, Purdy D, Dewhurst S, et al. Comprehensive proteomic analysis of influenza virus polymerase complex reveals a novel association with mitochondrial proteins and RNA polymerase accessory factors. Journal of virology. 2011;85(17):8569-81. Epub 2011/07/01. doi: 10.1128/jvi.00496-11. PubMed PMID: 21715506; PubMed Central PMCID: PMCPMC3165779.

122. Su WC, Chen YC, Tseng CH, Hsu PW, Tung KF, Jeng KS, et al. Pooled RNAi screen identifies ubiquitin ligase Itch as crucial for influenza A virus release from the endosome during virus entry. Proc Natl Acad Sci U S A. 2013;110(43):17516-21. Epub 2013/10/09. doi: 10.1073/pnas.1312374110. PubMed PMID: 24101521; PubMed Central PMCID: PMCPMC3808593.

123. Hubner M, Peter M. Cullin-3 and the endocytic system: New functions of ubiquitination for endosome maturation. Cellular logistics. 2012;2(3):166-8. Epub 2013/06/07. doi: 10.4161/cl.20372. PubMed PMID: 23739092; PubMed Central PMCID: PMCPMC3498076.

124. Huotari J, Meyer-Schaller N, Hubner M, Stauffer S, Katheder N, Horvath P, et al. Cullin-3 regulates late endosome maturation. Proc Natl Acad Sci U S A. 2012;109(3):823-8. Epub 2012/01/06. doi: 10.1073/pnas.1118744109. PubMed PMID: 22219362; PubMed Central PMCID: PMCPMC3271891.

125. Gschweitl M, Ulbricht A, Barnes CA, Enchev RI, Stoffel-Studer I, Meyer-Schaller N, et al. A SPOPL/Cullin-3 ubiquitin ligase complex regulates endocytic trafficking by targeting EPS15 at endosomes. Elife. 2016;5:e13841. doi: 10.7554/eLife.13841. PubMed PMID: 27008177; PubMed Central PMCID: PMCPMC4846373.

126. Banerjee I, Miyake Y, Nobs SP, Schneider C, Horvath P, Kopf M, et al. Influenza A virus uses the aggresome processing machinery for host cell entry. Science. 2014;346(6208):473-7. Epub 2014/10/25. doi: 10.1126/science.1257037. PubMed PMID: 25342804.

127. Liao TL, Wu CY, Su WC, Jeng KS, Lai MM. Ubiquitination and deubiquitination of NP protein regulates influenza A virus RNA replication. The EMBO journal. 2010;29(22):3879-90. Epub 2010/10/07. doi: 10.1038/emboj.2010.250. PubMed PMID: 20924359; PubMed Central PMCID: PMCPMC2989104.

128. Kirui J, Mondal A, Mehle A. Ubiquitination up-regulates influenza virus polymerase function. Journal of virology. 2016. Epub 2016/09/30. doi: 10.1128/jvi.01829-16. PubMed PMID: 27681127; PubMed Central PMCID: PMCPMC5110180.

129. Fu B, Wang L, Ding H, Schwamborn JC, Li S, Dorf ME. TRIM32 Senses and Restricts Influenza A Virus by Ubiquitination of PB1 Polymerase. PLoS pathogens. 2015;11(6):e1004960. Epub 2015/06/10. doi: 10.1371/journal.ppat.1004960. PubMed PMID: 26057645; PubMed Central PMCID: PMCPMC4461266.

130. Di Pietro A, Kajaste-Rudnitski A, Oteiza A, Nicora L, Towers GJ, Mechti N, et al. TRIM22 inhibits influenza A virus infection by targeting the viral nucleoprotein for degradation. Journal of virology. 2013;87(8):4523-33. Epub 2013/02/15. doi: 10.1128/jvi.02548-12. PubMed PMID: 23408607; PubMed Central PMCID: PMCPMC3624352.

131. Gao S, Wu J, Liu RY, Li J, Song L, Teng Y, et al. Interaction of NS2 with AIMP2 facilitates the switch from ubiquitination to SUMOylation of M1 in influenza A virus-infected cells. Journal of virology. 2015;89(1):300-11. Epub 2014/10/17. doi: 10.1128/jvi.02170-14. PubMed PMID: 25320310; PubMed Central PMCID: PMCPMC4301113.

132. Gack MU, Albrecht RA, Urano T, Inn KS, Huang IC, Carnero E, et al. Influenza A virus NS1 targets the ubiquitin ligase TRIM25 to evade recognition by the host viral RNA sensor RIG-I. Cell Host Microbe. 2009;5(5):439-49. Epub 2009/05/21. doi: 10.1016/j.chom.2009.04.006. PubMed PMID: 19454348; PubMed Central PMCID: PMCPMC2737813.

133. Rajsbaum R, Albrecht RA, Wang MK, Maharaj NP, Versteeg GA, Nistal-Villan E, et al. Species-Specific Inhibition of RIG-I Ubiquitination and IFN Induction by the Influenza A Virus NS1 Protein. PLoS pathogens. 2012;8(11):16. doi: 10.1371/journal.ppat.1003059. PubMed PMID: WOS:000311997100063.

134. Dyer MD, Murali TM, Sobral BW. The landscape of human proteins interacting with viruses and other pathogens. PLoS pathogens. 2008;4(2):e32. Epub 2008/02/20. doi: 10.1371/journal.ppat.0040032. PubMed PMID: 18282095; PubMed Central PMCID: PMCPMC2242834.

135. Bonetta L. Protein-protein interactions: Interactome under construction. Nature. 2010;468(7325):851-4. Epub 2010/12/15. doi: 10.1038/468851a. PubMed PMID: 21150998.

136. Fields C, Song O. A novel genetic system to detect protein-protein interactions. Nature. 1989;340:245-6.

137. Braun P, Tasan M, Dreze M, Barrios-Rodiles M, Lemmens I, Yu H, et al. An experimentally derived confidence score for binary protein-protein interactions. Nature methods. 2009;6(1):91-7. Epub 2008/12/09. doi: 10.1038/nmeth.1281. PubMed PMID: 19060903; PubMed Central PMCID: PMCPMC2976677.

138. Tavernier J, Eyckerman S, Lemmens I, Van der Heyden J, Vandekerckhove J, Van Ostade X. MAPPIT: a cytokine receptor-based two-hybrid method in mammalian cells. Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology. 2002;32(10):1397-404. Epub 2002/10/10. PubMed PMID: 12372116.

139. Lievens S, Lemmens I, Montoye T, Eyckerman S, Tavernier J. Two-hybrid and its recent adaptations. Drug discovery today Technologies. 2006;3(3):317-24. Epub 2006/10/01. doi: 10.1016/j.ddtec.2006.09.006. PubMed PMID: 24980535.

140. Magliery TJ, Wilson CG, Pan W, Mishler D, Ghosh I, Hamilton AD, et al. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. Journal of the American Chemical Society. 2005;127(1):146-57. Epub 2005/01/06. doi: 10.1021/ja046699g. PubMed PMID: 15631464.

141. Remy I, Michnick SW. A highly sensitive protein-protein interaction assay based on Gaussia luciferase. Nature methods. 2006;3(12):977-9. Epub 2006/11/14. doi: 10.1038/nmeth979. PubMed PMID: 17099704.

142. Cassonnet P, Rolloy C, Neveu G, Vidalain PO, Chantier T, Pellet J, et al. Benchmarking a luciferase complementation assay for detecting protein complexes. Nature methods. 2011;8(12):990-2. doi: 10.1038/nmeth.1773. PubMed PMID: 22127214.

143. Muller M, Jacob Y, Jones L, Weiss A, Brino L, Chantier T, et al. Large scale genotype comparison of human papillomavirus E2-host interaction networks provides new insights for e2 molecular functions. PLoS pathogens. 2012;8(6):e1002761. doi: 10.1371/journal.ppat.1002761. PubMed PMID: 22761572; PubMed Central PMCID: PMC3386243.

144. Neveu G, Cassonnet P, Vidalain PO, Rolloy C, Mendoza J, Jones L, et al. Comparative analysis of virus-host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase. Methods. 2012;58(4):349-59. Epub 2012/08/18. doi: 10.1016/j.ymeth.2012.07.029. PubMed PMID: 22898364; PubMed Central PMCID: PMCPMC3546263.

145. The ORFeome Collaboration: a genome-scale human ORF-clone resource. Nature methods. 2016;13(3):191-2. Epub 2016/02/26. doi: 10.1038/nmeth.3776. PubMed PMID: 26914201.

146. Utrecht U. Protein interaction mapping 2017 [cited 2017 August 2017]. Available from: https://web.science.uu.nl/developmentalbiology/boxem/interaction_mapping.html.

147. Kamburov A, Grossmann A, Herwig R, Stelzl U. Cluster-based assessment of protein-protein interaction confidence. BMC bioinformatics. 2012;13:262. Epub 2012/10/12. doi: 10.1186/1471-2105-13-262. PubMed PMID: 23050565; PubMed Central PMCID: PMCPMC3532186.

148. Mahajan G, Mande SC. Using structural knowledge in the protein data bank to inform the search for potential host-microbe protein interactions in sequence space: application to Mycobacterium tuberculosis. BMC bioinformatics. 2017;18(1):201. Epub 2017/04/06. doi: 10.1186/s12859-017-1550-y. PubMed PMID: 28376709; PubMed Central PMCID: PMCPMC5379762.

149. Ohue M, Matsuzaki Y, Uchikoga N, Ishida T, Akiyama Y. MEGADOCK: an all-to-all protein-protein interaction prediction system using tertiary structure data. Protein and peptide letters. 2014;21(8):766-78. Epub 2013/07/17. PubMed PMID: 23855673; PubMed Central PMCID: PMCPMC4443796.

150. Wiles AM, Doderer M, Ruan J, Gu TT, Ravi D, Blackman B, et al. Building and analyzing protein interactome networks by cross-species comparisons. BMC systems biology. 2010;4:36. Epub 2010/04/01. doi: 10.1186/1752-0509-4-36. PubMed PMID: 20353594; PubMed Central PMCID: PMCPMC2859380.

151. Tafforeau L, Chantier T, Pradezynski F, Pellet J, Mangeot PE, Vidalain PO, et al. Generation and comprehensive analysis of an influenza virus polymerase cellular interaction network. Journal of virology. 2011;85(24):13010-8. doi: 10.1128/JVI.02651-10. PubMed PMID: 21994455; PubMed Central PMCID: PMCPMC3233135.

152. Wang Q, Li Q, Liu R, Zheng M, Wen J, Zhao G. Host cell interactome of PA protein of H5N1 influenza A virus in chicken cells. Journal of proteomics. 2016;136:48-54. Epub 2016/02/02. doi: 10.1016/j.jprot.2016.01.018. PubMed PMID: 26828018.

153. Zhao M, Wang L, Li S. Influenza A Virus-Host Protein Interactions Control Viral Pathogenesis. Int J Mol Sci. 2017;18(8). Epub 2017/08/02. doi: 10.3390/ijms18081673. PubMed PMID: 28763020.

154. Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, et al. Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding. Cell Host Microbe. 2015;18(6):723-35. doi: 10.1016/j.chom.2015.11.002. PubMed PMID: 26651948.

155. Heaton NS, Moshkina N, Fenouil R, Gardner TJ, Aguirre S, Shah PS, et al. Targeting Viral Proteostasis Limits Influenza Virus, HIV, and Dengue Virus Infection. Immunity. 2016;44(1):46-58. Epub 2016/01/21. doi: 10.1016/j.immuni.2015.12.017. PubMed PMID: 26789921; PubMed Central PMCID: PMCPMC4878455.

156. Mayer D, Molawi K, Martinez-Sobrido L, Ghanem A, Thomas S, Baginsky S, et al. Identification of cellular interaction partners of the influenza virus ribonucleoprotein complex and polymerase complex using proteomic-based approaches. Journal of proteome research. 2007;6(2):672-82. Epub 2007/02/03. doi: 10.1021/pr060432u. PubMed PMID: 17269724; PubMed Central PMCID: PMCPMC2577182.

157. Jorba N, Juarez S, Torreira E, Gastaminza P, Zamarreno N, Albar JP, et al. Analysis of the interaction of influenza virus polymerase complex with human cell factors. Proteomics. 2008;8(10):2077-88. Epub 2008/05/21. doi: 10.1002/pmic.200700508. PubMed PMID: 18491320.

158. Ma H, Kien F, Maniere M, Zhang Y, Lagarde N, Tse KS, et al. Human annexin A6 interacts with influenza a virus protein M2 and negatively modulates infection. Journal of virology. 2012;86(3):1789-801. Epub 2011/11/25. doi: 10.1128/jvi.06003-11. PubMed PMID: 22114333; PubMed Central PMCID: PMCPMC3264383.

159. de Chassey B, Meyniel-Schicklin L, Aublin-Gex A, Navratil V, Chantier T, Andre P, et al. Structure homology and interaction redundancy for discovering virus-host protein interactions. EMBO Rep. 2013;14(10):938-44. Epub 2013/09/07. doi: 10.1038/embor.2013.130. PubMed PMID: 24008843; PubMed Central PMCID: PMCPMC3807215.

160. de Chassey B, Aublin-Gex A, Ruggieri A, Meyniel-Schicklin L, Pradezynski F, Davoust N, et al. The interactomes of influenza virus NS1 and NS2 proteins identify new host factors and provide insights for ADAR1 playing a supportive role in virus replication. PLoS pathogens. 2013;9(7):e1003440. Epub 2013/07/16. doi: 10.1371/journal.ppat.1003440. PubMed PMID: 23853584; PubMed Central PMCID: PMCPMC3701712.

161. Hsu WB, Shih JL, Shih JR, Du JL, Teng SC, Huang LM, et al. Cellular protein HAX1 interacts with the influenza A virus PA polymerase subunit and impedes its nuclear translocation. Journal of virology. 2013;87(1):110-23. Epub 2012/10/12. doi: 10.1128/jvi.00939-12. PubMed PMID: 23055567; PubMed Central PMCID: PMCPMC3536397.

162. Kuo RL, Li ZH, Li LH, Lee KM, Tam EH, Liu HM, et al. Interactome Analysis of the NS1 Protein Encoded by Influenza A H1N1 Virus Reveals a Positive Regulatory Role of Host Protein PRP19 in Viral Replication. Journal of proteome research. 2016;15(5):1639-48. Epub 2016/04/21. doi: 10.1021/acs.jproteome.6b00103. PubMed PMID: 27096427.

163. Li Q, Yuan X, Wang Q, Chang G, Wang F, Liu R, et al. Interactomic landscape of PA-X-chicken protein complexes of H5N1 influenza A virus. Journal of proteomics. 2016;148:20-5. Epub 2016/07/17. doi: 10.1016/j.jprot.2016.07.009. PubMed PMID: 27422376.

164. Wang L, Fu B, Li W, Patil G, Liu L, Dorf ME, et al. Comparative influenza protein interactomes identify the role of plakophilin 2 in virus restriction. Nat Commun. 2017;8:13876. doi: 10.1038/ncomms13876. PubMed PMID: 28169297; PubMed Central PMCID: PMCPMC5309701.

165. Lai YH, Li ZC, Chen LL, Dai Z, Zou XY. Identification of potential host proteins for influenza A virus based on topological and biological characteristics by proteome-wide network approach. Journal of proteomics. 2012;75(8):2500-13. Epub 2012/03/15. doi: 10.1016/j.jprot.2012.02.034. PubMed PMID: 22415277.

166. Guirimand T, Delmotte S, Navratil V. VirHostNet 2.0: surfing on the web of virus/host molecular interactions data. Nucleic acids research. 2015;43(Database issue):D583-7. Epub 2014/11/14. doi: 10.1093/nar/gku1121. PubMed PMID: 25392406; PubMed Central PMCID: PMCPMC4383936.

167. Gao Z, Hu J, Liang Y, Yang Q, Yan K, Liu D, et al. Generation and Comprehensive Analysis of Host Cell Interactome of the PA Protein of the Highly Pathogenic H5N1 Avian Influenza Virus in Mammalian Cells. Front Microbiol. 2017;8:739. doi: 10.3389/fmicb.2017.00739. PubMed PMID: 28503168; PubMed Central PMCID: PMCPMC5408021.

168. Shaw ML. The host interactome of influenza virus presents new potential targets for antiviral drugs. Reviews in medical virology. 2011;21(6):358-69. Epub 2011/08/09. doi: 10.1002/rmv.703. PubMed PMID: 21823192; PubMed Central PMCID: PMCPMC3207218.

169. Fournier G, Chiang C, Munier S, Tomoiu A, Demeret C, Vidalain PO, et al. Recruitment of RED-SMU1 complex by Influenza A Virus RNA polymerase to control Viral mRNA splicing. PLoS pathogens. 2014;10(6):e1004164. Epub 2014/06/20. doi: 10.1371/journal.ppat.1004164. PubMed PMID: 24945353; PubMed Central PMCID: PMCPMC4055741.

170. Poirson J, Biquand E, Straub ML, Cassonnet P, Nomine Y, Jones L, et al. Mapping the interactome of HPV E6 and E7 oncoproteins with the ubiquitin-proteasome system. The FEBS journal. 2017. Epub 2017/08/09. doi: 10.1111/febs.14193. PubMed PMID: 28786561.

171. Mahon C, Krogan NJ, Craik CS, Pick E. Cullin E3 ligases and their rewiring by viral factors. Biomolecules. 2014;4(4):897-930. Epub 2014/10/15. doi: 10.3390/biom4040897. PubMed PMID: 25314029; PubMed Central PMCID: PMCPMC4279162.

172. Wendel I, Matrosovich M, Klenk HD. SnapShot: Evolution of human influenza A viruses. Cell Host Microbe. 2015;17(3):416.e1. Epub 2015/03/15. doi: 10.1016/j.chom.2015.02.001. PubMed PMID: 25766297.

173. Fodor E, Devenish L, Engelhardt OG, Palese P, Brownlee GG, Garcia-Sastre A. Rescue of influenza A virus from recombinant DNA. Journal of virology. 1999;73(11):9679-82. Epub 1999/10/09. PubMed PMID: 10516084; PubMed Central PMCID: PMCPMC113010.

174. Jahan AS, Lestra M, Swee LK, Fan Y, Lamers MM, Tafesse FG, et al. Usp12 stabilizes the T-cell receptor complex at the cell surface during signaling. Proc Natl Acad Sci U S A. 2016;113(6):E705-14. Epub 2016/01/27. doi: 10.1073/pnas.1521763113. PubMed PMID: 26811477; PubMed Central PMCID: PMCPMC4760780.

175. Diot C, Fournier G, Dos Santos M, Magnus J, Komarova A, van der Werf S, et al. Influenza A Virus Polymerase Recruits the RNA Helicase DDX19 to Promote the Nuclear Export of Viral mRNAs. Sci Rep. 2016;6:33763. Epub 2016/09/23. doi: 10.1038/srep33763. PubMed PMID: 27653209; PubMed Central PMCID: PMCPMC5037575.

176. Kashiwagi T, Hara K, Nakazono Y, Uemura Y, Imamura Y, Hamada N, et al. The N-Terminal Fragment of a PB2 Subunit from the Influenza A Virus (A/Hong Kong/156/1997 H5N1) Effectively Inhibits RNP Activity and Viral Replication. PLoS One. 2014;9(12):17. doi: 10.1371/journal.pone.0114502. PubMed PMID: WOS:000345869700140.

177. Scholz CC, Cavadas MA, Tambuwala MM, Hams E, Rodriguez J, von Kriegsheim A, et al. Regulation of IL-1beta-induced NF-kappaB by hydroxylases links key hypoxic and inflammatory signaling pathways. Proc Natl Acad Sci U S A. 2013;110(46):18490-5. doi: 10.1073/pnas.1309718110. PubMed PMID: 24145445; PubMed Central PMCID: PMCPMC3832034.

178. Luo H. Interplay between the virus and the ubiquitin-proteasome system: molecular mechanism of viral pathogenesis. Current opinion in virology. 2016;17:1-10. Epub 2015/10/02. doi: 10.1016/j.coviro.2015.09.005. PubMed PMID: 26426962.

179. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, et al. UniProt: the Universal Protein knowledgebase. Nucleic acids research. 2004;32(Database issue):D115-9. Epub 2003/12/19. doi: 10.1093/nar/gkh131. PubMed PMID: 14681372; PubMed Central PMCID: PMCPMC308865.

180. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, et al. A draft map of the human proteome. Nature. 2014;509(7502):575-81. Epub 2014/05/30. doi: 10.1038/nature13302. PubMed PMID: 24870542; PubMed Central PMCID: PMCPMC4403737.

181. Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, et al. Mass-spectrometry-based draft of the human proteome. Nature. 2014;509(7502):582-7. Epub 2014/05/30. doi: 10.1038/nature13319. PubMed PMID: 24870543.

182. Rolland T, Tasan M, Charloteaux B, Pevzner SJ, Zhong Q, Sahni N, et al. A proteome-scale map of the human interactome network. Cell. 2014;159(5):1212-26. Epub 2014/11/25. doi: 10.1016/j.cell.2014.10.050. PubMed PMID: 25416956; PubMed Central PMCID: PMCPMC4266588.

183. Manz B, Schwemmle M, Brunotte L. Adaptation of avian influenza A virus polymerase in mammals to overcome the host species barrier. Journal of virology. 2013;87(13):7200-9. Epub 2013/04/26. doi: 10.1128/jvi.00980-13. PubMed PMID: 23616660; PubMed Central PMCID: PMCPMC3700283.

184. Zeng W, Sun L, Jiang X, Chen X, Hou F, Adhikari A, et al. Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. Cell. 2010;141(2):315-30. Epub 2010/04/21. doi: 10.1016/j.cell.2010.03.029. PubMed PMID: 20403326; PubMed Central PMCID: PMCPMC2919214.

185. Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, Szpyt J, et al. The BioPlex Network: A Systematic Exploration of the Human Interactome. Cell. 2015;162(2):425-40. Epub 2015/07/18. doi: 10.1016/j.cell.2015.06.043. PubMed PMID: 26186194; PubMed Central PMCID: PMCPMC4617211.

186. Chatr-Aryamontri A, Oughtred R, Boucher L, Rust J, Chang C, Kolas NK, et al. The BioGRID interaction database: 2017 update. Nucleic acids research. 2017;45(D1):D369-d79. Epub 2016/12/17. doi: 10.1093/nar/gkw1102. PubMed PMID: 27980099; PubMed Central PMCID: PMCPMC5210573.

187. Zhao J, Wei J, Mialki RK, Mallampalli DF, Chen BB, Coon T, et al. F-box protein FBXL19-mediated ubiquitination and degradation of the receptor for IL-33 limits pulmonary inflammation. Nature immunology. 2012;13(7):651-8. Epub 2012/06/05. doi: 10.1038/ni.2341. PubMed PMID: 22660580; PubMed Central PMCID: PMCPMC3643313.

188. Alfonso R, Lutz T, Rodriguez A, Chavez JP, Rodriguez P, Gutierrez S, et al. CHD6 chromatin remodeler is a negative modulator of influenza virus replication that relocates to inactive chromatin upon infection. Cell Microbiol. 2011;13(12):1894-906. doi: 10.1111/j.1462-5822.2011.01679.x. PubMed PMID: 21899694.

189. Bett JS, Ibrahim AF, Garg AK, Kelly V, Pedrioli P, Rocha S, et al. The P-body component USP52/PAN2 is a novel regulator of HIF1A mRNA stability. Biochem J. 2013;451(2):185-94. doi: 10.1042/BJ20130026. PubMed PMID: 23398456; PubMed Central PMCID: PMCPMC3632086.

190. Mok BW, Song W, Wang P, Tai H, Chen Y, Zheng M, et al. The NS1 protein of influenza A virus interacts with cellular processing bodies and stress granules through RNA-associated protein 55 (RAP55) during virus infection. Journal of virology. 2012;86(23):12695-707. doi: 10.1128/JVI.00647-12. PubMed PMID: 22973032; PubMed Central PMCID: PMCPMC3497642.

191. Mir MA, Duran WA, Hjelle BL, Ye C, Panganiban AT. Storage of cellular 5' mRNA caps in P bodies for viral cap-snatching. Proc Natl Acad Sci U S A. 2008;105(49):19294-9. Epub 2008/12/03. doi: 10.1073/pnas.0807211105. PubMed PMID: 19047634; PubMed Central PMCID: PMCPMC2614755.

192. Sikora D, Rocheleau L, Brown EG, Pelchat M. Influenza A virus cap-snatches host RNAs based on their abundance early after infection. Virology. 2017;509:167-77. Epub 2017/06/25. doi: 10.1016/j.virol.2017.06.020. PubMed PMID: 28646652.

193. Zheng D, Ezzeddine N, Chen CY, Zhu W, He X, Shyu AB. Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. J Cell Biol. 2008;182(1):89-101. doi: 10.1083/jcb.200801196. PubMed PMID: 18625844; PubMed Central PMCID: PMCPMC2447901.

194. Vignuzzi M, Gerbaud S, van der Werf S, Escriou N. Naked RNA immunization with replicons derived from poliovirus and Semliki Forest virus genomes for the generation of a cytotoxic T cell response against the influenza A virus nucleoprotein. The Journal of general virology. 2001;82(Pt 7):1737-47. Epub 2001/06/20. doi: 10.1099/0022-1317-82-7-1737. PubMed PMID: 11413386.

195. Matrosovich M, Matrosovich T, Garten W, Klenk HD. New low-viscosity overlay medium for viral plaque assays. Virology journal. 2006;3:63. Epub 2006/09/02. doi: 10.1186/1743-422x-3-63. PubMed PMID: 16945126; PubMed Central PMCID: PMCPMC1564390.

196. R Development Core Team. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing. Vienna, Austria.; 2016.

Articles arising from this work

- E. Biquand*, J. Poirson* et al. A comparative profiling of the Ubiquitin Proteasome System interplay with the influenza A virus PB2 polymerase protein recapitulating virus evolution in humans. Under revisions in mSphere, July 2017.

- J. Poirson*, E. Biquand* et al. Screening interactions between HPV E6 and E7 proteins and the Ubiquitin Proteasome System. Accepted in FEBS, July 2017.

- E. Biquand, C. Demeret. Structure resolution of the trimeric RNA-dependent RNA polymerase of influenza viruses: impact on our understanding of polymerase interactions with host and viral factors. Virologie, 2016, 20(6): 32-48. Review.

A comparative profiling of the Ubiquitin Proteasome System interplay with the influenza A virus PB2 polymerase protein recapitulating virus evolution in humans.

Elise Biquand*, Juline Poirson*, Marwah Karim, Marion Declercq, Nicolas Malausse, Patricia Cassonnet, Cyril Barbezange, Marie-Laure Straub, Louis Jones, Sandie Munier, Nadia Naffakh, Sylvie van der Werf, Yves Jacob, Murielle Masson[#], Caroline Demeret[#]

*, [#] these authors contributed equally to this work. Correspondance and requests for materials should be addressed to C.D. (mail caroline.demeret@pasteur.fr).

Abbreviations IAV influenza A virus; PPI protein-protein interaction; PRS positive reference set; RRS random reference set; CRL Cullin-based RING finger E3 Ub ligases; DUB de-ubiquitinase; NLR Normalized Luminescence Ratio.

ABSTRACT

Hijacking the host cell proteome is one cornerstone of a successful viral infection. Differential mapping of virus-host protein-protein interactions (PPI) based on comparative interactomics between multiple strains is an effective strategy to highlight correlations between network properties and biological or pathogenic traits. Herein, we developed an interactomic pipeline to deliver high-confidence comparative PPI maps between a given pathogen and the human Ubiquitin Proteasome System (UPS). This subarray of the human proteome represents a range of essential cellular functions and promiscuous targets for many viruses. The screening pipeline was applied to the influenza A virus (IAV) PB2 polymerase protein of five strains representing different levels of virulence in humans. An extensive PB2/UPS interplay has been detected, that recapitulates the evolution of IAV viruses in humans. UPS factors. Functional validation with several IAV strains, including seasonal $H1N1_{pdm09}$ and H3N2 viruses, confirmed the biological relevance of most identified UPS factors, and revealed strain-independent and strain-specific effects on IAV infection of UPS factors invalidation. Such strategy is applicable to proteins from any other virus or pathogen, providing a valuable resource to explore the UPS-pathogen interplay and its relationship with pathogenicity.

The typical virus life cycle sequentially involves cell entry, viral genome uncoating, transcription, replication, protein expression, particle assembly and egress, along with immune evasion. These processes require a large subset of the host cell machinery. Accordingly, viruses' capacity to infect, replicate, and propagate in a given host depends on their ability to appropriately subvert cell pathways and processes. Interaction proteomics, so-called interactomics, is a straightforward approach to assess virus hijacking of the cellular proteome, and decipher its impact on cellular functions. The human ORFeome v8.1 (CCSB), a cDNA collection assembling over 11,149 of the estimated 20,000 protein-coding genes existing in the human genome enables mapping virus-host PPI focused on subarrays of the human proteome [1, 2, 3].

In this study, we focused on the ubiquitin-proteasome system (UPS), a major pathway of protein post-translational modifications. Protein ubiquitination involves the step-wise action of three types of enzymes: the ubiquitin activating enzymes E1, the ubiquitin conjugating enzymes E2, and the ubiquitin ligases E3[4]. The E3 ligases perform the ultimate transfer of ubiquitin molecules to substrate proteins, and represent the most extended UPS families, consisting of about 1000 factors [5]. There are two distinct classes of E3 ligases, *i.e.* the non-cullin E3 ligases that carry out substrate recognition and ubiquitin ligase activity in a single protein, and the cullin-based Ring-E3 ligases (CRL), which act as protein complexes. In addition, protein ubiquitination can be reversed by de-ubiquitinases (DUBs). Ubiquitination determines stability, activity, or sub-cellular localization of targeted proteins [6, 7]. It is a highly versatile post-translational modification and is tightly regulated. Hijacking the UPS seems inherent to the replication cycles of viruses and other intracellular pathogens. Viruses manipulate protein ubiquitination to their advantage in various ways, to overcome host cell defense mechanisms, cell cycle control or cell death [8, 9]. The involvement of UPS in influenza virus infection is supported by an increasing number of studies, primarily focused on viral entry/uncoating and escape from cellular antiviral responses [10, 11, 12]. A functional interplay between the human UPS and the viral replication proteins PB1, PB2, PA and NP has also been reported, however, only a small number of underlying PPI has been detected [13, 14, 15]. The viral replication proteins have also been shown to be ubiquitinated during IAV infection, resulting in an increased polymerase activity, but the mechanisms of their ubiquitination remain uncharacterized [16, 17].

The PB2 protein is a subunit of the trimeric influenza virus polymerase, and a key player of IAV pathogenicity [18]. From the crystal structure of the viral polymerase complex, PB2 appears to provide large, flexible and shape-changing interfaces accessible for interactions with host proteins [19, 20]. We describe here a comparative interactomic strategy developed to explore the interplay between PB2 of 5 IAV strains of various degrees of virulence and the human UPS. A split-luciferase complementation Assay (HT-GPCA) [21] was

used as a systematic PPI screening method against a targeted library covering about half of the whole human UPS system. A high-confidence set of PB2/UPS interactions has been identified. The interaction profile-based clustering segregated according to the level of IAV strain adaptation to humans. Functional validation with three IAV strains indicated the biological relevance of identified UPS factors for IAV infection, in a common or a strain-specific manner.

RESULTS

Implementation of HT-GPCA for detecting UPS-PB2 PPIs

The HT-GPCA method relies on the ability of interacting protein pairs expressed in fusion with the Glc1 and Glc2 complementary fragments of the *Gaussia princeps* luciferase to reconstitute an active enzyme. We first evaluated whether this assay could successfully distinguish between interacting and non-interacting pairs, when the PB2 proteins from influenza A viruses (IAV) are expressed in fusion with the Glc2 fragment, corresponding to the C-terminal fragment of the luciferase. The PB2 proteins used for HT-GPCA evaluation were the ones chosen for the comparative screening with the human UPS, originating from 5 IAV strains of various degree of virulence in human, *i.e.* the laboratory-adapted strain $H1N1_{WSN}$ (A/WSN/33), the low to mildly virulent human seasonal viruses $H1N1_{pmd09}$ (A/Bretagne/7608/2009) and H3N2 (A/Centre/1003/2012), and the highly virulent viruses $H1N1_{1918}$ (A/Brevig mission/1/1918) and H7N9 (A/Anhui/1/2013).

The Glc2-PB2 proteins were co-expressed either with a set of nine previously reported PB2 interacting partners (Positive Reference Set, PRS) or with 11 *a priori* non-interacting proteins (Random Reference Set, RRS), fused to the N-terminal fragment of the *Gaussia princeps* luciferase (Glc1 fragment). HT-GPCA clearly discriminated between the PRS and RRS sets (supplementary figure 1). The observed variations in the overall levels of luminescence were not related with differences in Glc2-PB2 proteins accumulation, but likely reflect differences in the intrinsic binding properties of the 5 PB2 proteins (supplementary figure 1).

The UPS-targeted library

A UPS library suitable for PPI detection via the HT-GPCA method was first constructed by assembling 590 cDNAs encoding for 558 unique human UPS factors (supplementary data 1) cloned in fusion with the N-terminal Glc1 fragment of the *Gaussia princeps* luciferase (Poirson et al, submitted). The coverage of this library was estimated in comparison with an *in silico* human UPS database, designated “Human UPS”, which was

obtained by combining the existing DuDe and UUCE ubiquitin-dedicated databases [5, 22] (supplementary data 1). Our UPS library covers 50% of the human E1 activating enzymes, 64% of the E2 ubiquitin conjugating enzymes, 46% of the non-cullin E3 ubiquitin ligases, 28% of the cullin-based E3 ubiquitin ligases and 52% of the DUBs (figure 1). With few exceptions (E3-DWD, DUB-ULP), the distribution of the different sub-categories of E3 ubiquitin ligases and of DUBs in the UPS library is comparable to the Human UPS (figure 1). This is an important feature, since the E3 and DUBs are the major regulators of protein ubiquitination/de-ubiquitination, and thus prime targets of viruses.

HT-GPCA screening of PB2-UPS interactions

The entire UPS library screening was conducted via 12 independent HT-GPCA experiments, wherein a set of 25 to 72 exploratory Glc1-UPS factors were screened for interactions with the five individual Glc2-PB2 proteins along with the PRS and RRS (supplementary data 2). In each experiment, a Positive Threshold (PT) was calculated per PB2 protein, based on the distribution of the luminescence values generated by the PB2-UPS pairs. The PT corresponded to the third quartile + 1.5 times interquartile space (PT=Q3+1.5(IQR)). UPS factors generating outlier luminescence values above the PT were selected as potential interacting partners (figure 2-a and supplementary figure 2). This selection was validated by the distribution of the PRS and RRS relatively to the PT (figure 2-a and supplementary figure 2). A fraction of the PRS consistently generated luminescence values under the PT (supplementary data 2), indicating the high stringency of the selection of positive interactions from this initial screening using the defined PT.

While the PT differed between the various PB2 proteins, their relative levels were preserved between experiments (figure 2-b), according to differences in PB2 intrinsic binding properties detected with the PRS and RRS sets of proteins (supplementary figure 1). Such disparity did not alter the sensitivity of GPCA to identify potential interacting partners, since the number of selected UPS factors was independent of the PT level of the PB2 proteins (figure 2-c). The initial PB2/UPS screening identified 91 UPS clones, corresponding to 80 unique UPS proteins, as putative interacting partners of at least one PB2 protein (supplementary data 2). Twenty-four factors were selected with all the 5 PB2 proteins, whereas the remaining 56 factors exhibited distinct interaction patterns between strains (figure 2-d).

Post-screening retesting of the PB2-UPS interacting pairs

The interactions selected from the initial screen were retested by applying the NLR (Normalized Luminescence Ratio) method to the HT-GPCA assay [21]. This method, taking into account the background interaction level of the Glc1- or Glc2-fusion partners, has been

shown to accurately discriminate PB2 interacting partners from non-interacting proteins in an infectious context [23]. To improve assay detectability and to allow a parallel comparison, each of the 80 selected UPS factors was retested against all 5 PB2 proteins, in three biological replicates (supplementary data 3). For each PB2 protein individually, the NLR values obtained with the set of 11 RRS were used to calculate a 99.73 % confidence interval. The upper limit of this confidence interval was used as the threshold for the selection of positive interactions (supplementary data 3). With this selection strategy, 42 UPS factors scored 3 times over 3 with at least one PB2, and were considered as "high confidence" PB2 interactors.

Hierarchical clustering of the PB2-UPS interaction profiles

The PB2/UPS interaction profiles were determined using z-score transformation of the NLR. They were compared by agglomerative hierarchical clustering between the different strains (see material and methods), which in turn was used to calculate an interaction dendrogram (figure 3-a).

PPI profile-based clustering separates viral strains in two groups: on the one hand strains recently introduced in the human population from an avian origin ($H1N1_{pdm09}$ and H7N9), on the other hand human adapted-strains that sequentially derive from each other ($H1N1_{1918}$, $H1N1_{WSN}$, H3N2). The PB2/UPS interaction profiles of this latter group further clustered according to the duration of circulation of the PB2 segment in the human population (figure 3-c). We next analyzed the interaction profile of PB2 from the avian strain A/Mallard/Marquenterre/Z2371/83 ($H1N1_{MZ}$) with this set of UPS factors. z-scores were calculated for the obtained NLR, and used to compare the UPS/PB2 $H1N1_{MZ}$ profile to those of the 5 PB2 from human-infecting strains (supplementary data 3, figure 3-b). $H1N1_{MZ}$ PB2 gathered with the other avian-origin $H1N1_{pdm09}$ and H7N9 PB2 proteins, but the latter having recently gained human infection potential remained closer to each other (figure 3-b). It can thus be hypothesized from the PB2-UPS interaction dendrogram that the interplay between the PB2 polymerase protein and the human UPS may have evolved along with virus adaptation to humans*, i.e* acquisition of the capacity to infect human from animal reservoir (H7N9 vs $H1N1_{Mz}$), gaining human-to-human transmission capability ($H1N1_{pdm09}$ vs H7N9), then duration of circulation in the human population ($H1N1_{1918}$ < $H1N1_{WSN}$ < H3N2).

Functional impact of PB2 interactors

To demonstrate the biological relevance of identified PB2/UPS PPIs, we studied the involvement of the UPS targets of PB2 in the production of infectious viral particles. For that purpose, we examined the productive infectious cycle of the seasonal strains $H1N1_{pdm09}$ and H3N2 viruses and of the $H1N1_{WSN}$ laboratory-adapted strain upon siRNA-mediated depletion

of the individual PB2 interactors. The seasonal strains $H1N1_{pdm09}$ and H3N2 grow poorly in the human pulmonary A549 cell line. We thus produced by reverse genetics adapted $H1N1_{pdm09}$ and H3N2 viruses harboring an intact PB2 segment and point mutations in the HA segment, which efficiently infected A549 cells (see material and methods).

Forty-one high-confidence PB2 interactors were assessed, while the ubiquitin (UBB) protein was excluded owing to its pleiotropic function. A549 cells transfected with siRNAs targeting the individual UPS factors, or transfected with a non-target siRNA, were infected for 24 hours with $H1N1_{WSN}$, adapted $H1N1_{pdm09}$ or H3N2 viruses. Knockdown efficiency and cell viability were assessed for the different siRNAs (supplementary figure 3). The titer of infectious IAV particles produced was measured by plaque assays. The effects of UPS factors depletion were calculated by comparing the virus titers produced with those obtained with non-target siRNA. Of the 41 PB2 interactors tested, 36 showed an effect on the infection with all or a subset of IAV strains, albeit at variable degrees. A statistically significant ($P<0.05$) reduction of viral production was observed with the depletion of 13 (H3N2), 26 ($H1N1_{pdm09}$) and 31 ($H1N1_{WSN}$) factors. The decrease ranged between 2.3 and 4.8 fold for H3N2, 1.7 and 7 fold for $H1N1_{pdm09}$ and 2.3 and 50 fold for $H1N1_{WSN}$ (figures 4 and 5-a, supplementary data 4). The variable effects of the siRNAs suggest that the identified UPS factors may have critical roles or modulate IAV infection. Most of the UPS factors that had an effect on viral production were also among those found to bind the PB2 protein of the corresponding strain. The demonstration of the direct role of the PB2/UPS interactions will nonetheless require to study mutations that disrupt the PB2/UPS interactions.

The six UPS factors discovered to be essential for infection with the three viruses consist of members of the cullin-based E3 ligase complexes (DCAF12L, DDB1, FBXL19), a component of the H2A histone Ubiquitin ligase complex (PHC2, [24]), and DUBs (OTUD6A, PAN2) (figure 5-c, supplementary data 4). These functional studies also identified three factors (TRIM32, TRAF3-IP2, KHLH34) that specifically modulate infection with the $H1N1_{pdm09}$ virus from the 2009 pandemic.

The UPS factors that were found to mediate $H1N1_{WSN}$ infection were ectopically expressed fused to the mCherry protein, to evaluate their overexpression effect on infection with recombinant $H1N1_{WSN}$ virus expressing the mCitirne fluorescent reporter (figure 5-b). Twelve factors increased viral replication, as mesuread by the percentage of mCitrine-positive cells (figure 5-b), further validating their role in IAV infection.

The functional evaluation revealed that UPS factors that may have strain-independent or strain-specific effects on IAV infection, highlighting a differential involvement of the UPS in IAV infection (figure 5-c).

DISCUSSION

The split-luciferase strategy developed here for the comparative mapping of interactions between the IAV PB2 protein and factors of the human UPS shows a number of noticeable benefits allowing the straightforward delivery of high-quality, systematic interaction mapping. It takes advantage of excellent performances the high-throughput protein-protein interaction assay HT-GPCA for specific and sensitive detection of protein-protein interactions (PPI) [21, 25, 26]. Evaluation of the same matrix of host factors for interactions with multiple pathogen proteins provides a rigorous comparative interaction mapping, as applied here to the human Ubiquitin Proteasome System (UPS) and the PB2 protein of influenza A viruses. Importantly, the variable intrinsic properties of pathogens proteins are not problematic for the comparative aspect of this interaction mapping as demonstrated in this study with the 5 IAV PB2 proteins of different strains.

A paramount issue for such systematic interactomics strategy is the quality and exhaustiveness of the library representing a subarray of the human proteome. The UPS-dedicated library that was assembled and carefully characterized contained 558 unique UPS factors out of the 1277 described in databases (Poirson et al, submitted), covering one to two third of the different UPS categories (E1, E2, various E3 ligases classes). The relative UPS library composition parallels the database-derived human UPS composition, notably regarding the E3 ligases or their substrate recognitions factors. This is of key importance since E3 ligases provide the human UPS its complexity and flexibility towards protein targeting, and are consequently the major targets of viruses [27]. The UPS dedicated library thus offers an accurate representation of the global human UPS, ready to use for interaction mapping by HT-GPCA with any protein. The same strategy can be implemented with other sub-arrays of the proteome provided that HT-GPCA compatible libraries are available. Smaller libraries dedicated to the Interferon α/β, the NFκB and the TGFβ pathways actually exist, and have been used to benchmark the HT-GPCA using a human-human protein interaction matrix [21].

The usurpation of UPS is a recognized significant process of viral infections, which may underlie essential traits of the pathogenic potential [8, 9]. An interplay between the IAV replication proteins and the UPS has been detected [13, 15, 16, 17], which is barely characterized. The comparative interactomics strategy applied in this study addresses this knowledge gap. Indeed, we identified a rather extensive interplay between PB2 and the human UPS, and showed that PB2 binds to E3 ubiquitin ligases or E3-related factors (“others” in our UPS library), to DUBs, and to a proteasome sub-unit, but not to E1 or E2 enzymes. Moreover, the PB2 proteins interact with all categories of E3 ligases; RING ligases as well as Cullin-based RING E3 ligases complexes (CRL) (supplementary data 5). The targeted proteins in those

CRL complexes are mainly substrate recognition factors, in line with the dysregulation of cellular proteins ubiquitination, often observed with viruses through UPS hijacking. The targeting of several DUBs also suggests that PB2 might interfere with the versatility of ubiquitination process. Finally, PB2 appears to broadly target UPS proteins without sub-category preference (supplementary figure 4), possibly reflecting a wide manipulation of the UPS during IAV infection.

Virus-host protein interaction networks can be deduced from the obtained binary PPI datasets. A PB2/UPS interaction maps encompassing all the strains tested was plugged to the human interactome, recovered from the BioGRID PPI database. Some of the PB2 targets are interacting with each others (figure 6-a), underlying an intricate interplay where PB2 binds to UPS proteins that already act in complexes. When all human factors interacting with the UPS partners of PB2 are considered, an elaborated interaction map emerges (figure 6-b), revealing the possible impact of PB2 on the human proteome through UPS targeting. An affinity purification coupled with mass spectrometry was recently used to determine the comparative interactomes of the proteins of several influenza virus strains [28]. Interestingly, among the 132 host factors found in association with PB2 in this study, 29 are binding to UPS targets of PB2 detected in our study. These observations highlight the complementary nature of the two interactomics strategies: while mass spectrometry approaches characterize proteins in complexes, the systematic binary PPI screenings point to direct targets within the host proteome. Both of these strategies contributed to the identification of host factors functionally relevant to IAV infection [28].

Indeed, of the 41 top-confident PB2 interactors identified in our screening, a total of 36 were shown by siRNA-based experiments to affect IAV replication, with the depletion of 13, 25 and 31 UPS factors decreasing the production of the H3N2, $H1N1_{pmd09}$ and $H1N1_{WSN}$ viruses, respectively. Such high functional output emphasizes the strong power of this interactomic strategy to detect biologically relevant partners of PB2. Six UPS factors are involved in the replication cycle of all three low to mildly virulent IAV strains studied, and could represent a core set of UPS factors generally involved in the regulation of IAV infection. For some other UPS factors, a differential effect was observed from strain to strain, highlighting an unexpectedly diversified functional interplay between IAV and the UPS. Strain-specific disparities may provide clues on the molecular mechanisms underlying differences in pathogenicity, while conserved effects could serve as targets for broad-spectrum therapeutics.

The strain specificity of the virus-host interplay can be determined by comparing the interaction profiles of viral factors against the same set of host proteins [25, 26]. We showed by agglomerative hierarchical clustering that the PB2/UPS interaction patterns are segregating according to the lineage of origin of the PB2 segment as well as to the duration of its

circulation in the human population. The position of the purely avian PB2 in the interaction dendrogram suggests that the PB2/UPS interaction profiles might be relevant to detect the ability of avian strains to infect humans before they acquire human-to-human transmission potential. As such, the PB2/UPS interplay may help predict the potential of IAV strains emerging from avian reservoirs to cause human infections. Analysis of the PB2/UPS interaction for a larger number of human and avian strains collected at different time points will be needed to support this possibility.

The UPS-dedicated comparative interactomic strategy thus constitutes a valuable resource, available to be applied to other pathogens to decipher UPS/pathogen proteins interplay that could also provide valuable insights into significant pathogen traits.

METHODS

Plasmids

Gateway®-compatible Destination GPCA (pDEST) vectors pSPICA-N1 and pSPICA-N2 both derived from the pCiNeo mammalian expression vector, and expressed respectively the Glc1 and 2 complementary fragments of the *Gaussia princeps* luciferase linked to the N-terminal ends of tested proteins after recombinatorial cloning (Gateway® system, Invitrogen). The ORFs encoding for the PB2 from influenza viruses A/WSN/33 ($H1N1_{WSN}$), A/Bretagne/7608/2009 ($H1N1_{pdm09}$), A/Centre/1003/2012 (H3N2), A/Anhui/1/2013 (H7N9), A/BrevigMission/1/1918 ($H1N1_{1918}$) and A/Mallard/Marquenterre/Z2371/83 ($H1N1_{MZ}$) were cloned into Gateway®-DONOR vector pDONR207, the resulting Entry clones were transferred into Gateway®-pDEST pSPICA-N2 to produce the Glc2-PB2 expressing plasmids. The ORFs encoding UPS factors were obtained as Gateway®-Entry plasmids from the human ORFeome (the CCSB human ORFeome collection) and transferred in Gateway®-pDEST pSPICA-N1 to obtain the Glc1-UPS expressing plasmids. The RRS (Random Reference Set) contains human ORFs encoding proteins randomly picked from the human ORFeome and *a priori* not interacting with the viral protein PB2: LRCC28, NXP2, NFE2L1, GSTT1, GYPA, DPYSL2, UGT3A1, DBH, PLEKHA9, NXPH1, CNTN2, SLC7A13, CACNG7. The PRS (Positive Reference Set) corresponds to human ORFs, encoding NFX1, RUVBL2, NUP50, PTGES3, KPNA2 proteins pick up from the human ORFeome, shown to bind PB2 in the literature.

Cell lines

HEK-293T and A549 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS). MDCK-SIAT cells were grown in Modified Eagle's Medium (MEM) supplemented with 5% FCS.

HT-GPCA

HEK-293T cells were seeded in white 96-well plates at a concentration of $3x10^4$ cells per well. After 24 h, cells were transfected using linear PEI (polyethylenimine) with 300 ng of Glc2-PB2 expressing plasmid and 100 ng of Glc1-UPS expressing plasmid. At 24 h post-transfection, cells were washed with 100 µl of PBS and lysed with 40 µl of Renilla Lysis Buffer (Promega, E2820) for 1h. *Gaussia princeps* luciferase enzymatic activity was measured using a Berthold Centro LB960 luminometer by injecting 50 µl per well of luciferase substrate reagent (Promega, E2820) and counting luminescence during 10 seconds. Results were expressed as Relative Luminescence Unit (RLU).

NLR retesting

For the NLR method, the Glc2-PB2/Glc1-UPS pairs were tested in HT-GPCA along with controls consisting in 300 ng Glc2-PB2 + 100 ng Glc1 and 300 ng Glc2 + 100 ng Glc1-UPS. The NLR was calculated as the fold change normalized over the sum of controls. For a given protein pair A/B, NLR = (Glc1-A + Glc2-B) / [(Glc1-A + Glc2) +(Glc1 + Glc2-B)]. Retesting experiments were conducted three times for each UPS factors.

Viruses

The A/WSN/33(H1N1) virus was produced by reverse genetics as described in [29]. The recombinant A/WSN/33 virus expressing mCitrine was produced by reverse genetics using a pPolI-PB2 plasmid with a sequence encoding the self-cleaving 2A peptide from porcine teschovirus followed by the mCitrine coding sequence, as described in [30] for the Nanoluciferase ORF. Influenza viruses A/Centre/1003/2012(H3N2) and A/Bretagne/7608/2009($H1N1_{pdm09}$) were provided by the National Influenza Center at the Institut Pasteur (Paris, France), and passaged 5 times on A549 cells at m.o.i. 0.01 pfu/cell. After the 5^{th} passage, the titers of adapted viruses produced on A549 had increased by more than 2 log. The hemagglutinin (HA) of the adapted viruses H1N1pdm09 and H3N2 was sequenced. One mutation in the H3N2 HA (G460T) and two mutations in the $H1N1_{pdm09}$ HA (A517G and G834A) were identified. They were introduced into the HA-encoding plasmid in the A/Centre/1003/2012(H3N2) and H1N1A/Bretagne/7608/2009($H1N1_{pdm09}$) reverse genetics system, to produce A549-adapted H3N2 and $H1N1_{pmd09}$ viruses.

siRNA assays

Small interfering RNAs were purchased from Dharmacon (ON-TARGETplus SMARTpools and Non-targeting Control pool). A549 cells were transfected with 25 nM of siRNA using the DharmaFECT1 transfection reagent (Dharmacon). At 48 hours post transfection, cells were infected with the $H1N1_{WSN}$, (m.o.i. of 0,0001 pfu/cell) or adapted

H3N2 or $H1N1_{pdm09}$ viruses (m.o.i 0,001 pfu/cell) for 24 h. Plaque assays using MDCK-SIAT cells were performed as described in [31].

Efficiency of siRNA was controlled using expression plasmids for UPS factors fused to *Gaussia princeps* luciferase as described in [32], except for DCAF12L1 for which knock-down efficiency was controlled using qRT-PCR. For this factor cell lysates were harvested and subjected to the total RNA extraction using the RNeasy Mini Kit (Qiagen). RT-qPCR was conducted using the following primers : forward primer 5'-AATGCGCTCTACACCCACTG-3', reverse primer 5'-TTGACCAAAGGACCACTCACT-3') with the protocol of the Light Cycler RNA Amplification Kit SYBR Green 1 (Roche). The cellular GAPDH mRNA in the infected cells was also quantified as an internal control by using the above RT-qPCR. The DCAF12L1 siRNA decreased by 50 % the level of DCAF12L1 mRNA (not shown). Cell viability was determined using trypan blue counting 48 hrs post transfection with 25 nM of the siRNAs.

Overexpression-based experiments

HEK-293T cells were seeded in white 96-well plates at a concentration of $3x10^4$ cells per well. After 24 h, cells were transfected using linear PEI with 150 ng of pCINeo plasmid expressing an UPS factor fused to mCherry. At 24 h post-transfection, cells were infected with the WSN-PB2-2A-mCitrine virus at a m.o.i. of 1 pfu/cell for 18h. Cells were then fixed with a 4% PFA solution and analysed by flow-cytometry (Attune NxT, Thermo Fisher Scientific).

STATISTICS

Whisker-plot representation

The distributions of the relative luminescence values were represented by whisker-plot boxes. Whisker length corresponds to 1.5 times the interquartile range (IQR) that is equal to the difference between the upper and lower quartiles (IQR=Q3-Q1). The upper whisker, defined by the third quartile (Q3) plus 1.5 time the interquartile range, *i.e* Q3+1.5(IQR), was defined as the positive threshold (PT).

Determination of the confidence interval

To estimate the significance of a NLR value for a given protein-protein pair by comparison to the RRS sampling signal, a confidence interval was calculated for the RRS dataset, considering the estimated standard error SE and a confidence level of 99.73% (i.e. a risk α = 0.27%) by using the following expression: μ-t.SE + μ+t.SE where t is the critical value for α two-sided student test and for (n-1) degrees of freedom. We considered the NLR

value of a new sample as statistically significantly different from the RRS if its value is larger than the upper confidence bound determined for the RRS dataset.

Hierarchical clustering

NLR values for each PB2 were normalized using Z-scores calculated as follows: Z-score = [NLR – mean (NLR for the PB2)] / standard deviation. The clustering was made on Z-scores using Euclidian Distance parameter and word.D method with the R software [33].

ACKNOWLEDGMENTS

This work was supported by the Institut Pasteur in the frame of PTR call 2015 (PTR n° 546), EU FP7 PREDEMICS (Grant Agreement No. 278433), Labex IBEID. EB, MD were recipients of a MENRT PhD fellowship, JP was a recipient of a PhD fellowship of the Ligue Nationale Contre le Cancer and Alsace Contre le Cancer. MK has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 665850. We thank Marc Vidal and David Hill at the Center for Cancer Systems Biology (CCSB) at the Dana-Farber Cancer Institute (Boston, MA) for providing us with the human ORFeome collection. We are grateful to Véronique Guyot-Biquand (Université de Rennes 1, UMR CNRS 6552 Ethologie Animale et Humaine, Rennes, France) for her advices concerning statistical analyses of the data and to Shirit Einav for her suggestions and critical reading of the manuscript.

AUTHORS CONTRIBUTION

EB, JP, MK, MD, CD performed the experimental work; SM, NM constructed and characterized the mCitrine H1N1$_{WSN}$ reporter; EB, JP, PC, MLS contributed to the UPS library assembly and characterization; EB, MD, CB contributed to the adaption of H1N1pdm and H3N2 viruses; EB, LJ performed bioinformatics analyzes; EB, SM, NN, SW, YJ, MM, CD contributed to experimental design and analyses; CD wrote the manuscript.

LEGENDS TO FIGURES

Figure 1. Pie-chart representation of the UPS library. For each category of the UPS, the number of factors is indicated, as well as the coverage of the *in silico* Human UPS database (in brackets). The representativeness of sub-categories for the DUB, cullin-based (CUL-based) and non-Cullin (Non-CUL) E3-ligases are shown for the Human UPS and the UPS library.

Figure 2. First screening of the UPS library with the PB2 proteins **a-** Example of UPS factors screening in the first HT-GPCA experiment. For each PB2 protein, whisker-plots were generated from the luminescence values of the UPS-PB2 pairs. The Positive Threshold is highlighted in red for each PB2. Outliers luminescence values, represented in open circles, were selected as potential positive interactions. The RRS and PRS values (yellow and green dots respectively) were plotted over the whisker-plots afterwards, to evaluate the accuracy of the outlier-based selection. **b-** Positive Thresholds (PT) obtained in each individual HT-GPCA experiment were calculated relative to the PT of PB2 $H1N1_{WSN}$. The total number of UPS factors selected in the first screening is given at the top of each histogram for each of the PB2 protein. **c-** Number of selected factors according to the PT in each experiment for the 5 PB2 proteins. There is no correlation between the value of the PT and the number of selected factors (Pearson's product-moment correlation test: R^2=-0.2207, p-value= 0.09015). **d-** Venn diagram indicating the numbers of common UPS factors identified with the different PB2 proteins.

Figure 3. Interaction dendrograms deduced from the PB2/UPS interaction mapping. **a-** Hierarchical clustering of the 5 indicated PB2 proteins according to their interaction profiles with the human UPS. Heat maps represent the z-scores derived from an experimentally homogenous set of NLR obtained by the retesting of 75 UPS factors at once in the NLR retesting-2 experiment (supplementary data 3). **b-** Hierarchical clustering of PB2/UPS interaction profiles including PB2 $H1N1_{MZ}$. **c-** Schematics delineating, for the IAV strains used in this study: the sampling year; the duration (in years) of PB2 segment circulation in the human population (in brackets); the year of PB2 segment introduction in humans (vertical arrows), the grey arrow indicates the infection of a human patient without any spread in the human population; the course of the PB2 segment in the human population (plain lines), the dashed line represents the presence of the PB2 segment originating from the former H1N1 in the H3N2 strain (adapted from [34]).

Figure 4. Functional evaluation of the involvement of the UPS PB2 partners in the multiplication of H3N2 and $H1N1_{pdm09}$. A549 cell transfected by Non-Target (NT) or UPS targeting siRNA for 48 hours were infected by $H1N1_{pdm09}$ or H3N2 at a m.o.i of 0.001 pfu/cell. The titers of produced viruses were determined in the supernatants collected 24 hours post infection by plaque forming assay, and expressed as ratio relative to the titers obtained with non-target siRNA. UPS factors with statistically significant effects were labelled in red (H3N2) or pink ($H1N1_{pdm09}$), hatched bars indicate that PPI has been detected with the corresponding PB2 protein, taking into account positive scoring of at least two times in the NRL retesting experiments, plain bars that no interaction was detected (validated less than two times in

NLR restesting). Data represent means ±s.e.m of three independent experiments, p values were calculated with a two-tailed non-parametric Student's *t*-test, *<0.05, **<0.001, ***<0.001.

Figure 5. Functional evaluation of the involvement of the UPS PB2 partners in the multiplication of $H1N1_{WSN}$ virus. **a-** Effect of siRNA-mediated depletion**.** A549 cell transfected by Non-Target (NT) or UPS targeting siRNA for 48 hours were infected by $H1N1_{WSN}$ at a m.o.i of 0.0001 pfu/cell for 24 hours. Viral titers were determined by plaque forming assay, and expressed as ratio relative to the titers obtained with non-target siRNAs**.** UPS factors with statistically significant effects were labelled in purple, hatched bars indicate that PPI has been detected with the $H1N1_{WSN}$ PB2, plain bars that no interaction was detected. **b-** Effect of UPS factors overexpression. Plasmids expressing the UPS factors fused to mCherry were transfected in HEK293T cells. Twenty-four hours after transfection, the cells were infected at a m.o.i of 1 pfu/cell with a reporter $H1N1_{WSN}$ virus expressing the mCitrine fluorescent protein for 18 hours. The percentage of mCitrine-positive infected cells in Cherry-UPS expressing cells was determined by FACS analysis, and expressed as a ratio relative to the mCitrine-positive cells expressing the unfused mCherry. Data represent means ±s.e.m of three independent experiments, p values were calculated with a two-tailed non-parametric Student's *t*-test, *<0.05, **<0.001, ***<0.001 **c-** Venn diagram illustrating the number of UPS factors found to be involved in the multiplication of one or several IAV viruses.

Figure 6. The PB2/UPS interactome. **a-** Map of the PB2-UPS interactions. The PB2-UPS direct interactions are labeled in red, and comprise interactions detected with all the PB2 proteins tested. Interactions existing between the UPS targets of PB2 are labeled in violet. **b-** Plugging of the PB2-UPS network onto the human interactome. The PB2 UPS targets are labeled in violet, the host proteins they connect are organized in circles, according to their degree of connection (node degree) with the UPS factors interacting with PB2. A node degree of 1 means an interaction with 1 UPS factors of the PB2/UPS interactome, of 2 with 2 UPS factors of the PB2/UPS interactome etc...

Supplementary figure 1. a- Discrimination of PRS and RRS by HT-GPCA for the 5 PB2 proteins. Plasmids encoding the Glc1-RRS and Glc1-PRS were co-transfected with Glc2-PB2 proteins in HEK293T cells. The distributions of the luminescence values for protein pairs were represented with a Whisker-plot, where dots represent outliers values. Discrimination between PRS and RRS were validated with a Wilcoxon rank sum test, calculated p-values for the 5 PB2 proteins are: $H1N1_{WSN}$ $p=8.5x10^{-5}$, $H1N1_{pdm09}$ $p=1.13x10^{-5}$, H3N2 $p=2.75x10^{-4}$, H7N9 $p=1.13x10^{-5}$ and $H1N1_{1918}$ $p=8.50x10^{-5}$. **b-** Expression levels of the Glc2-PB2 proteins

in HT-GPCA settings. The Glc2-fused PB2 proteins in whole cell lysate of HT-GPCA samples were analyzed by Western-blotting using rabbit anti-Gaussia antibodies, anti-tubulin was used for loading control.

Supplementary figure 2. UPS library screening. The 12 HT-GPCA experiments covering the screening of the whole UPS library are shown. In each experiment, whisker plots were generated from the luminescence values of the UPS-PB2 pairs. Outliers luminescence values, represented in circles, were selected as potential positive interactions. The RRS and PRS values (yellow and green dots respectively) were plotted in the whiskers plots afterwards to evaluate the accuracy of the outlier-based selection.

Supplementary figure 3. Toxicity and silencing efficiency of siRNAs. **a-** Toxicity of siRNA. A549 cells were transfected with 25 nM of siRNA and cell viability was determined at 72 hours post transfection using trypan blue. The results are expressed as the mean percentages ± SEM (n=3). There is no correlation between the slight loss of cell viability observed for several siRNA and the functional effect of these siRNA on the virus cycles of the three tested strains. **b-** Silencing efficiency of siRNA. A549 cells were transfected with 25 nM of control or UPS-targeting siRNAs and with plasmids encoding the corresponding UPS protein fused with the full-length *Gaussia princeps* luciferase (pGlucFL-UPS). Ratios of the luciferase activities obtained in cells transfected with the UPS-targeting siRNA to the one obtained in cells transfected with the control siRNA are shown. The results are represented as floating bars with a line at the mean.

Supplementary figure 4. Representativeness of the UPS categories. Pie chart of different UPS categories in the UPS library and in the UPS targets of PB2.

REFERENCES

1. Clamp M*, et al.* Distinguishing protein-coding and noncoding genes in the human genome. *Proc Natl Acad Sci U S A* **104**, 19428-19433 (2007).

2. Ezkurdia I*, et al.* Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. *Hum Mol Genet* **23**, 5866-5878 (2014).

3. Collaboration OR. The ORFeome Collaboration: a genome-scale human ORF-clone resource. *Nat Methods* **13**, 191-192 (2016).

4. Hershko A, Ciechanover A. The ubiquitin system. *Annu Rev Biochem* **67**, 425-479 (1998).

5. Gao T*, et al.* UUCD: a family-based database of ubiquitin and ubiquitin-like conjugation. *Nucleic Acids Res* **41**, D445-451 (2013).

6. Chen ZJ, Sun LJ. Nonproteolytic functions of ubiquitin in cell signaling. *Mol Cell* **33**, 275-286 (2009).

7. Livneh I, Cohen-Kaplan V, Cohen-Rosenzweig C, Avni N, Ciechanover A. The life cycle of the 26S proteasome: from birth, through regulation and function, and onto its death. *Cell research* **26**, 869-885 (2016).

8. Luo H. Interplay between the virus and the ubiquitin-proteasome system: molecular mechanism of viral pathogenesis. *Curr Opin Virol* **17**, 1-10 (2016).

9. van de Weijer ML, Luteijn RD, Wiertz EJ. Viral immune evasion: Lessons in MHC class I antigen presentation. *Seminars in immunology* **27**, 125-137 (2015).

10. Gschweitl M*, et al.* A SPOPL/Cullin-3 ubiquitin ligase complex regulates endocytic trafficking by targeting EPS15 at endosomes. *Elife* **5**, e13841 (2016).

11. Rudnicka A, Yamauchi Y. Ubiquitin in Influenza Virus Entry and Innate Immunity. *Viruses* **8**, (2016).

12. Su WC*, et al.* Pooled RNAi screen identifies ubiquitin ligase Itch as crucial for influenza A virus release from the endosome during virus entry. *Proc Natl Acad Sci U S A* **110**, 17516-17521. doi: 17510.11073/pnas.1312374110. Epub 1312372013 Oct 1312374117. (2013).

13. Di Pietro A*, et al.* TRIM22 inhibits influenza A virus infection by targeting the viral nucleoprotein for degradation. *J Virol* **87**, 4523-4533 (2013).

14. Nailwal H, Sharma S, Mayank AK, Lal SK. The nucleoprotein of influenza A virus induces p53 signaling and apoptosis via attenuation of host ubiquitin ligase RNF43. *Cell death & disease* **6**, e1768 (2015).

15. Fu B, Wang L, Ding H, Schwamborn JC, Li S, Dorf ME. TRIM32 Senses and Restricts Influenza A Virus by Ubiquitination of PB1 Polymerase. *PLoS Pathog* **11**, e1004960 (2015).

16. Kirui J, Mondal A, Mehle A. Ubiquitination up-regulates influenza virus polymerase function. *J Virol*, (2016).

17. Liao TL, Wu CY, Su WC, Jeng KS, Lai MM. Ubiquitination and deubiquitination of NP protein regulates influenza A virus RNA replication. *EMBO J* **29**, 3879-3890 (2010).

18. Boivin S, Cusack S, Ruigrok RW, Hart DJ. Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. *J Biol Chem* **285**, 28411-28417. doi: 28410.21074/jbc.R28110.117531. Epub 112010 Jun 117510. (2010).

19. Pflug A, Guilligay D, Reich S, Cusack S. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* **516**, 355-360 (2014).

20. Thierry E*, et al.* Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. *Mol Cell* **61**, 125-137 (2016).

21. Cassonnet P*, et al.* Benchmarking a luciferase complementation assay for detecting protein complexes. *Nat Methods* **8**, 990-992 (2011).

22. Hutchins AP, Liu S, Diez D, Miranda-Saavedra D. The repertoires of ubiquitinating and deubiquitinating enzymes in eukaryotic genomes. *Mol Biol Evol* **30**, 1172-1187. doi: 1110.1093/molbev/mst1022. Epub 2013 Feb 1177. (2013).

23. Munier S, Rolland T, Diot C, Jacob Y, Naffakh N. Exploration of Binary Virus-Host Interactions Using an Infectious Protein Complementation Assay. *Mol Cell Proteomics* **12**, 2845-2855 (2013).

24. Awad S*, et al.* Mutation in PHC1 implicates chromatin remodeling in primary microcephaly pathogenesis. *Hum Mol Genet* **22**, 2200-2213 (2013).

25. Muller M*, et al.* Large scale genotype comparison of human papillomavirus E2-host interaction networks provides new insights for e2 molecular functions. *PLoS Pathog* **8**, e1002761 (2012).

26. Neveu G*, et al.* Comparative analysis of virus-host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase. *Methods* **58**, 349-359 (2012).

27. Mahon C, Krogan NJ, Craik CS, Pick E. Cullin E3 ligases and their rewiring by viral factors. *Biomolecules* **4**, 897-930 (2014).

28. Wang L*, et al.* Comparative influenza protein interactomes identify the role of plakophilin 2 in virus restriction. *Nat Commun* **8**, 13876 (2017).

29. Fodor E, Devenish L, Engelhardt OG, Palese P, Brownlee GG, Garcia-Sastre A. Rescue of influenza A virus from recombinant DNA. *J Virol* **73**, 9679-9682 (1999).

30. Tran V, Moser LA, Poole DS, Mehle A. Highly sensitive real-time in vivo imaging of an influenza reporter virus reveals dynamics of replication and spread. *J Virol* **87**, 13321-13329 (2013).

31. Matrosovich M, Matrosovich T, Garten W, Klenk HD. New low-viscosity overlay medium for viral plaque assays. *Virology journal* **3**, 63 (2006).

32. Diot C*, et al.* Influenza A Virus Polymerase Recruits the RNA Helicase DDX19 to Promote the Nuclear Export of Viral mRNAs. *Scientific reports* **6**, 33763 (2016).

33. R Development Core Team. **R: A language and environment for statistical computing**. (ed^(eds). R Foundation for Statistical Computing. Vienna, Austria. (2016).

34. Wendel I, Matrosovich M, Klenk HD. SnapShot: Evolution of human influenza A viruses. *Cell Host Microbe* **17**, 416.e411 (2015).

Figure 1

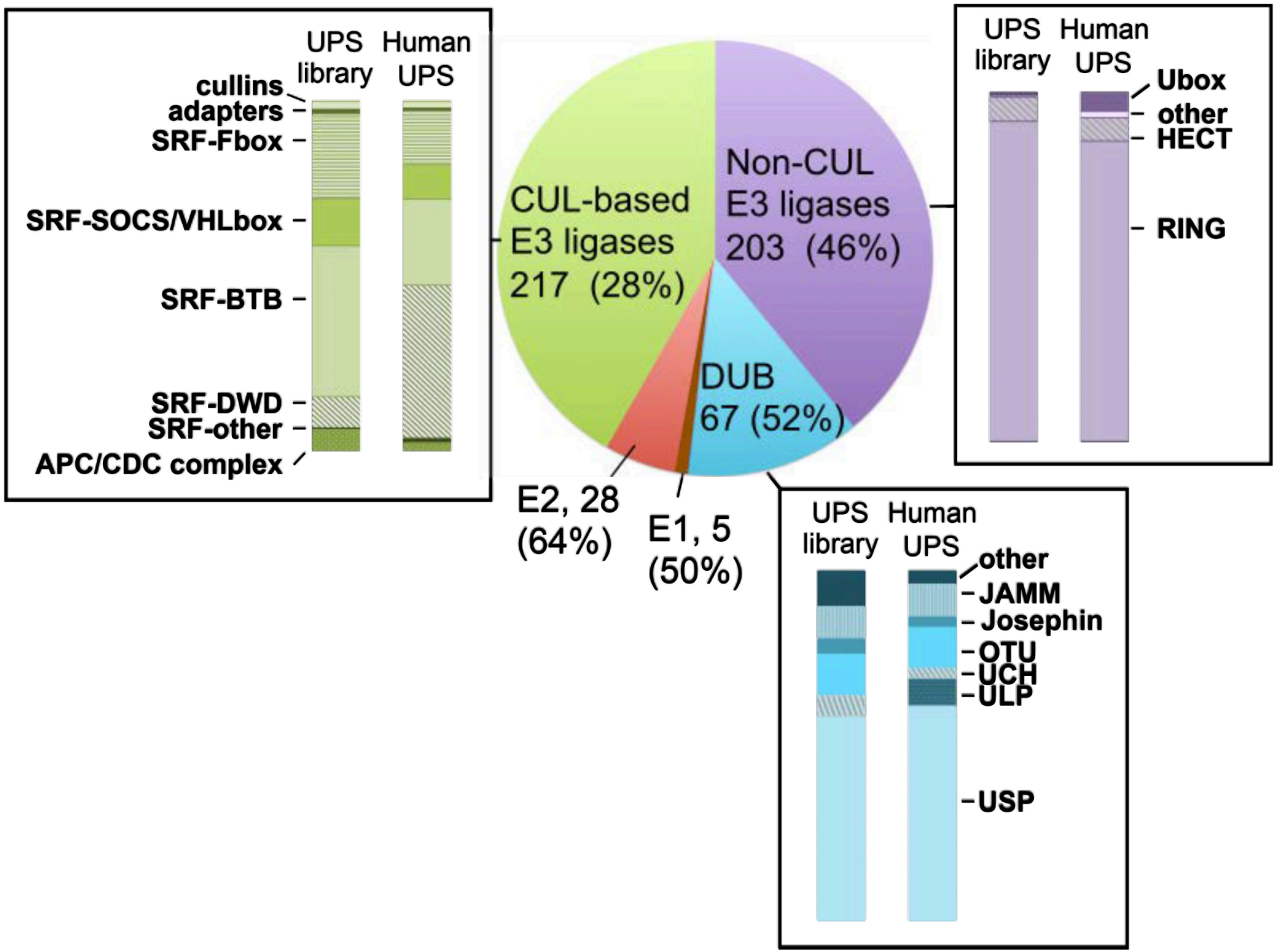

UPS library
Human UPS
cullins
adapters
SRF-Fbox
SRF-SOCS/VHLbox
SRF-BTB
SRF-DWD
SRF-other
APC/CDC complex
CUL-based E3 ligases 217 (28%)
Non-CUL E3 ligases 203 (46%)
DUB 67 (52%)
E2, 28 (64%)
E1, 5 (50%)
Ubox
other
HECT
RING
JAMM
Josephin
OTU
UCH
ULP
USP

Figure 2

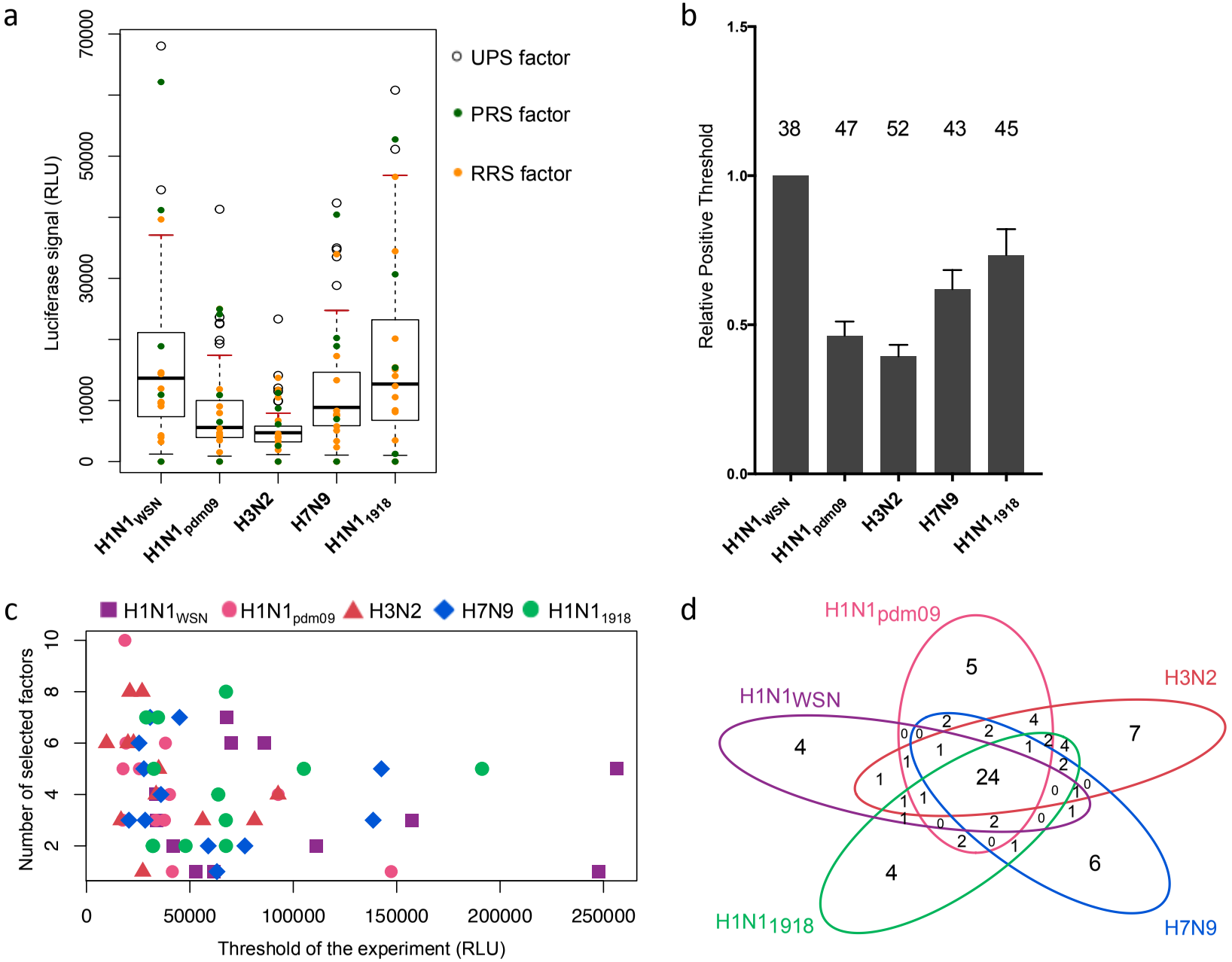

a
Luciferase signal (RLU)
0
10000
30000
50000
70000
UPS factor
PRS factor
RRS factor
H1N1WSN
H1N1pdm09
H3N2
H7N9
H1N11918
b
Relative Positive Threshold
0.0
0.5
1.0
1.5
38
47
52
43
45
H1N1WSN
H1N1pdm09
H3N2
H7N9
H1N11918
c
H1N1WSN
H1N1pdm09
H3N2
H7N9
H1N11918
Number of selected factors
2
4
6
8
10
0
50000
100000
150000
200000
250000
Threshold of the experiment (RLU)
d
H1N1pdm09
H3N2
H1N1WSN
H7N9
H1N11918
5
4
7
24
4
6

Figure 3

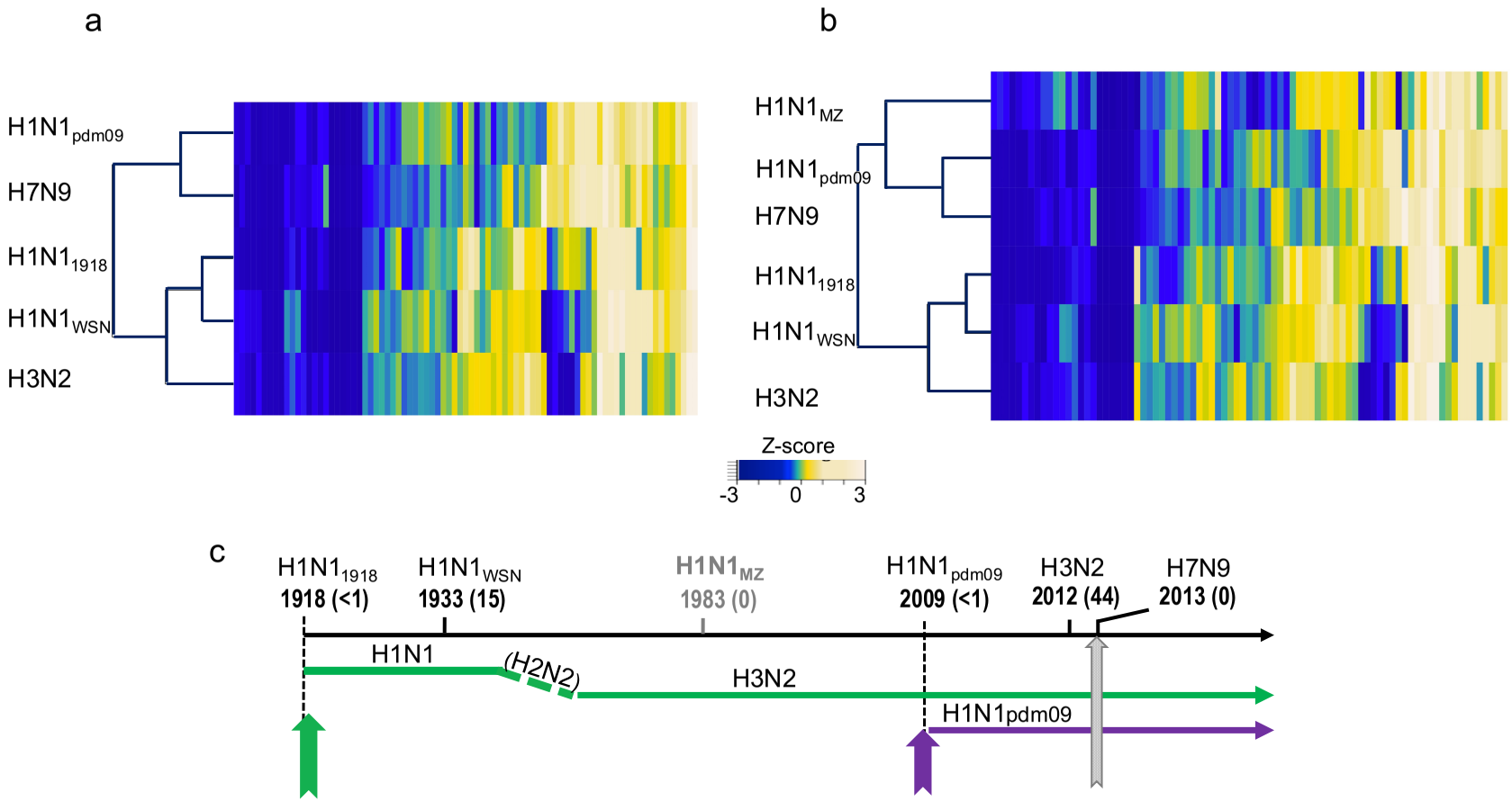

a
H1N1pdm09
H7N9
H1N1 1918
H1N1WSN
H3N2
b
H1N1MZ
H1N1pdm09
H7N9
H1N1 1918
H1N1WSN
H3N2
Z-score
-3
0
3
c
H1N1 1918
1918 (<1)
H1N1WSN
1933 (15)
H1N1MZ
1983 (0)
H1N1pdm09
2009 (<1)
H3N2
2012 (44)
H7N9
2013 (0)
H1N1
(H2N2)
H3N2
H1N1pdm09

Figure 4

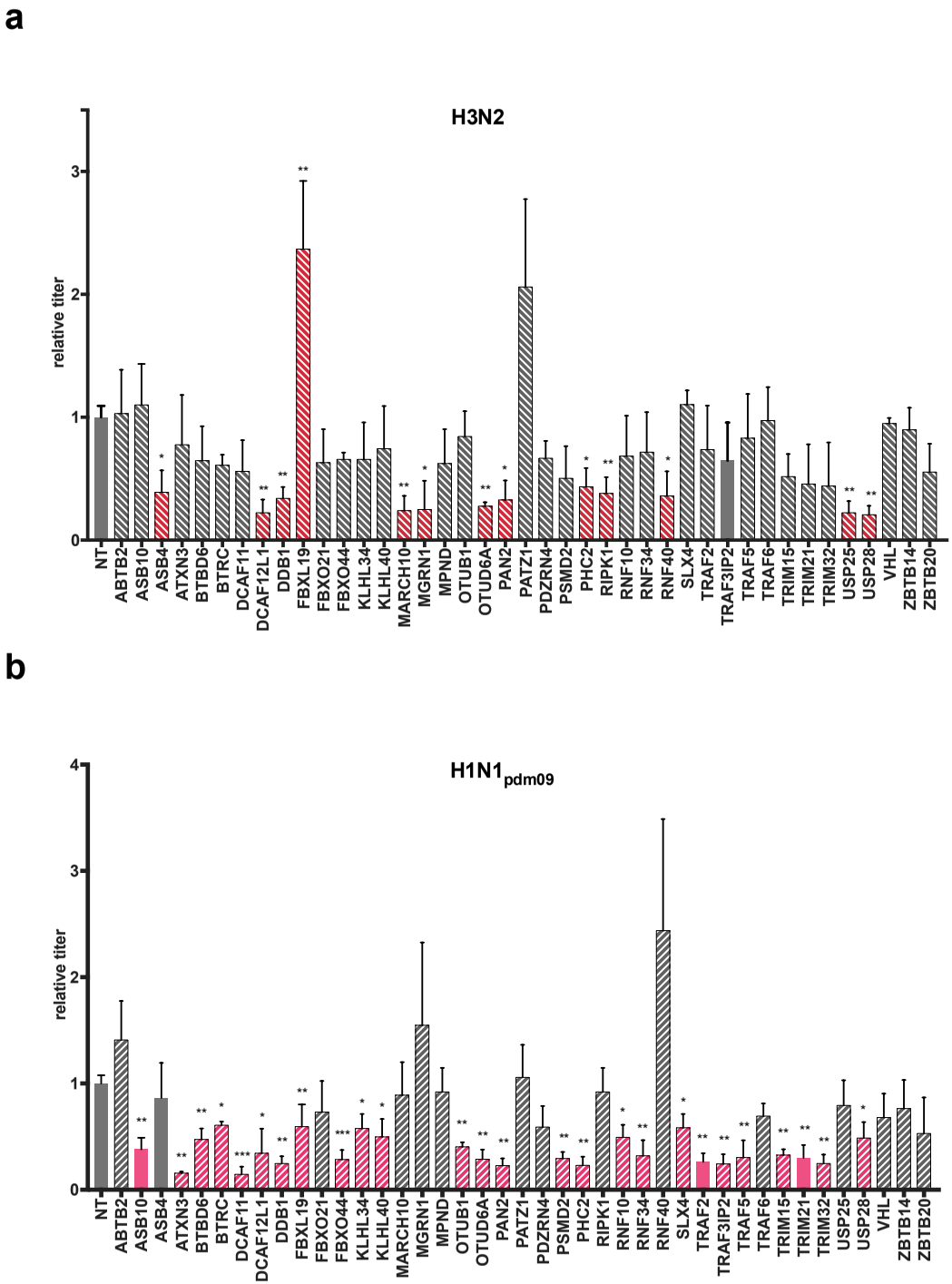

a
H3N2
relative titer
0
1
2
3
NT
ABTB2
ASB10
ASB4
ATXN3
BTBD6
BTRC
DCAF11
DCAF12L1
DDB1
FBXL19
FBXO21
FBXO44
KLHL34
KLHL40
MARCH10
MGRN1
MPND
OTUB1
OTUD6A
PAN2
PATZ1
PDZRN4
PSMD2
PHC2
RIPK1
RNF10
RNF34
RNF40
SLX4
TRAF2
TRAF3IP2
TRAF5
TRAF6
TRIM15
TRIM21
TRIM32
USP25
USP28
VHL
ZBTB14
ZBTB20
b
H1N1pdm09
relative titer
0
1
2
3
4
NT
ABTB2
ASB10
ASB4
ATXN3
BTBD6
BTRC
DCAF11
DCAF12L1
DDB1
FBXL19
FBXO21
FBXO44
KLHL34
KLHL40
MARCH10
MGRN1
MPND
OTUB1
OTUD6A
PAN2
PATZ1
PDZRN4
PSMD2
PHC2
RIPK1
RNF10
RNF34
RNF40
SLX4
TRAF2
TRAF3IP2
TRAF5
TRAF6
TRIM15
TRIM21
TRIM32
USP25
USP28
VHL
ZBTB14
ZBTB20

Figure 5

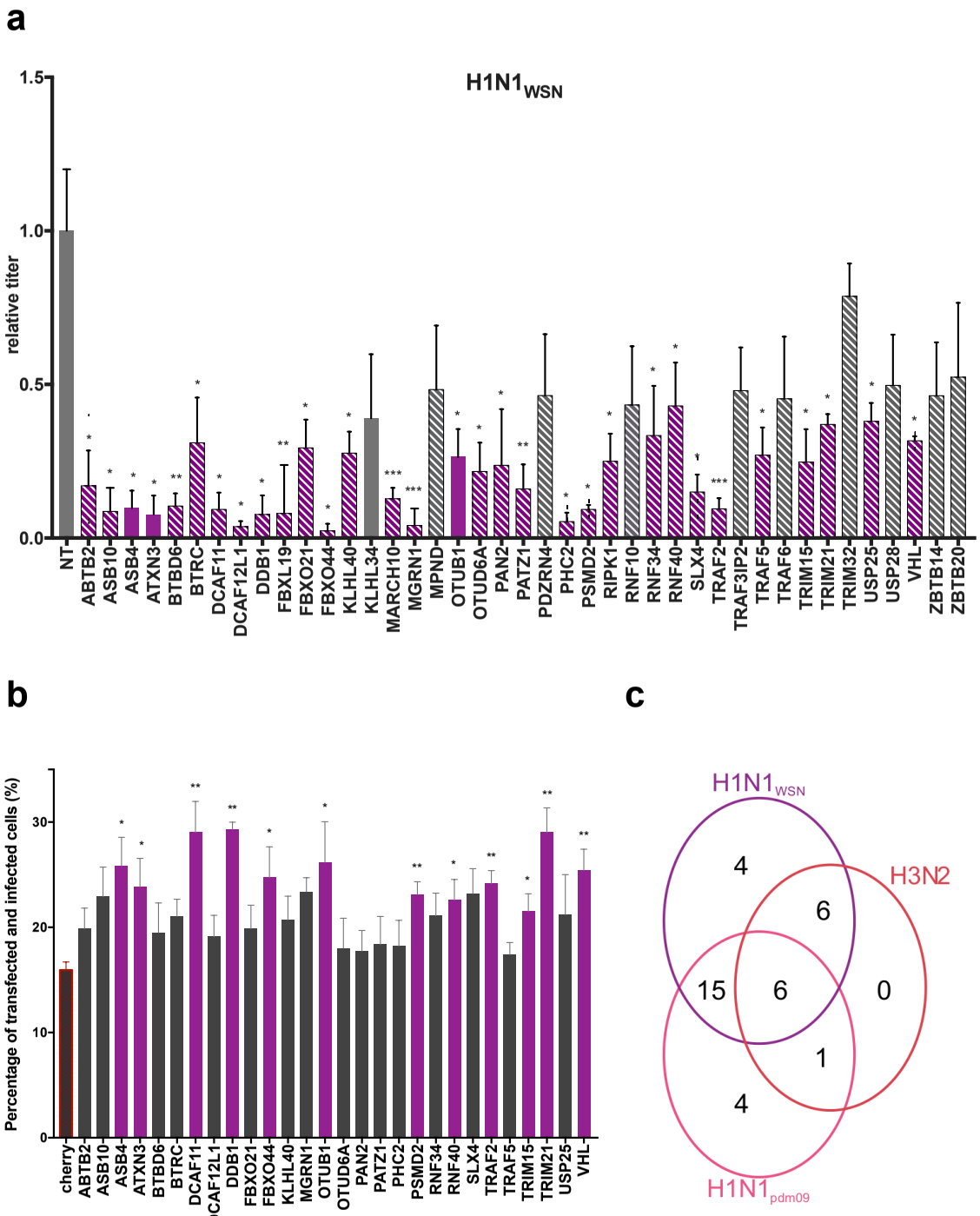
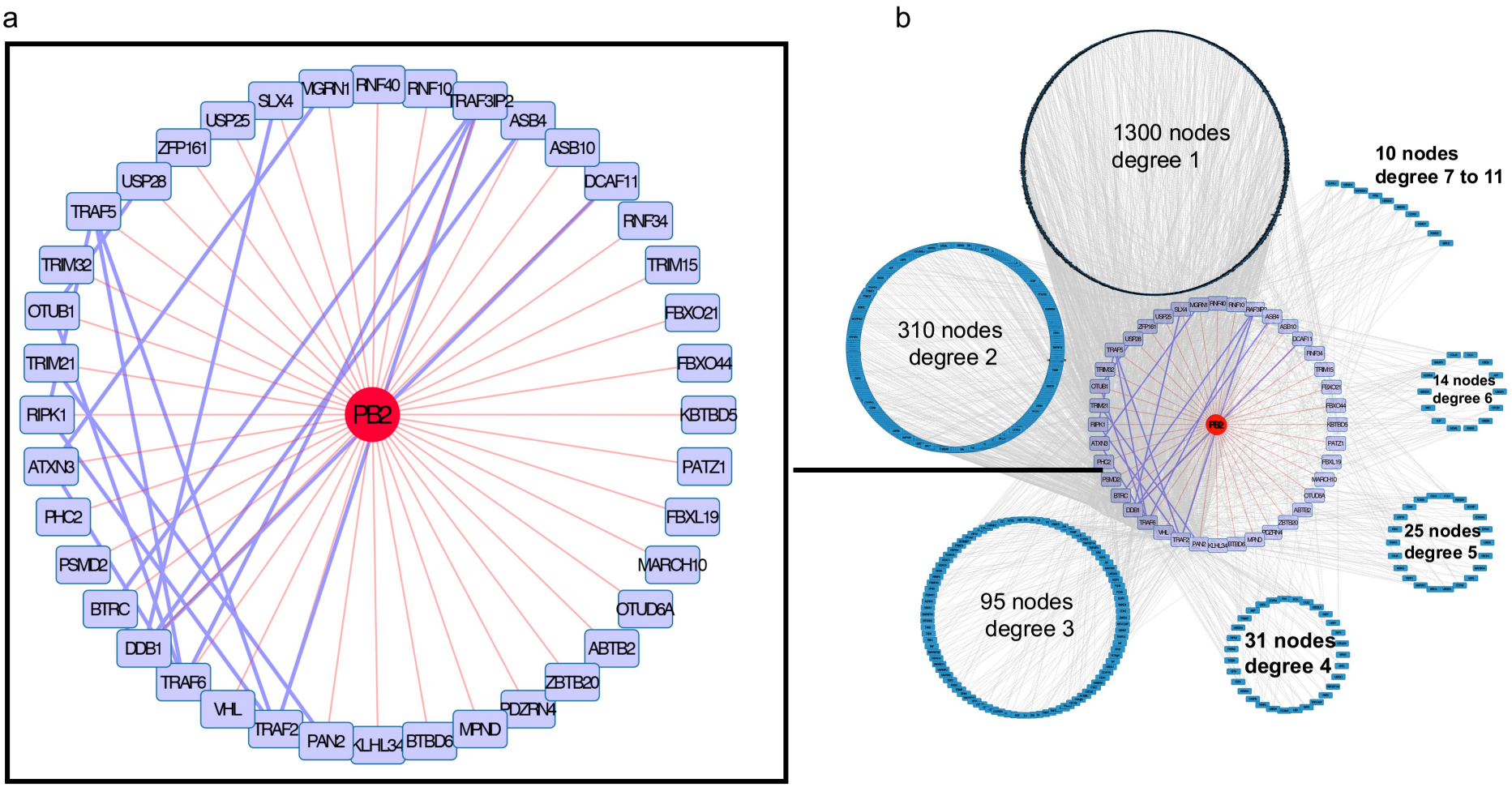
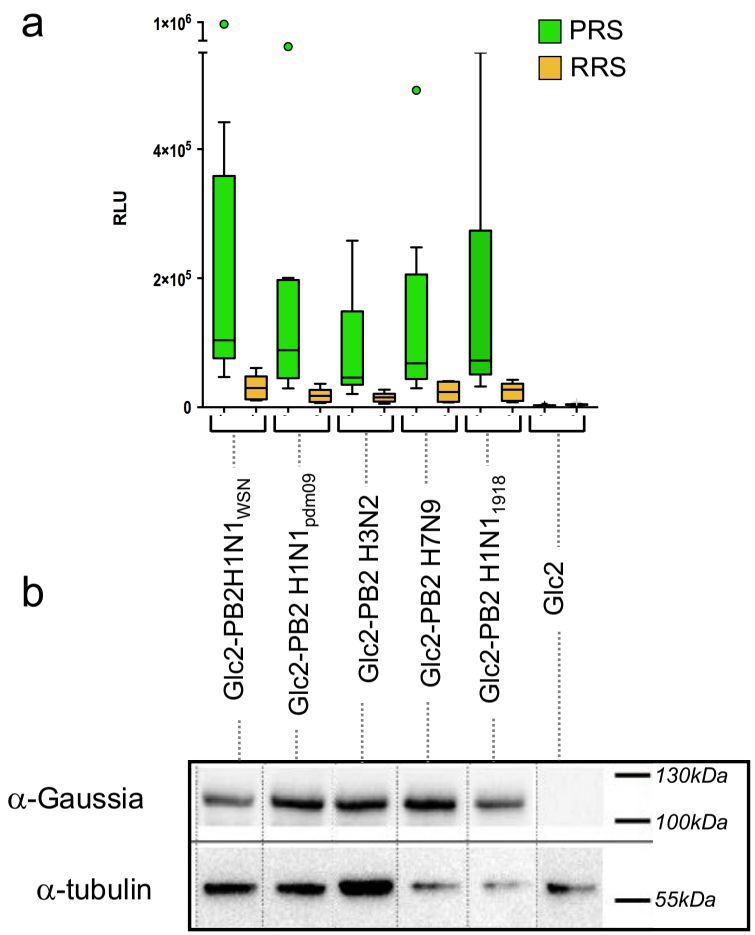

a
H1N1WSN
relative titer
1.5
1.0
0.5
0.0
NT
ABTB2
ASB10
ASB4
ATXN3
BTBD6
BTRC
DCAF11
DCAF12L1
DDB1
FBXL19
FBXO21
FBXO44
KLHL40
KLHL34
MARCH10
MGRN1
MPND
OTUB1
OTUD6A
PAN2
PATZ1
PDZRN4
PHC2
PSMD2
RIPK1
RNF10
RNF34
RNF40
SLX4
TRAF2
TRAF3IP2
TRAF5
TRAF6
TRIM15
TRIM21
TRIM32
USP25
USP28
VHL
ZBTB14
ZBTB20
b
Percentage of transfected and infected cells (%)
30
20
10
0
cherry
ABTB2
ASB10
ASB4
ATXN3
BTBD6
BTRC
DCAF11
DCAF12L1
DDB1
FBXO21
FBXO44
KLHL40
MGRN1
OTUB1
OTUD6A
PAN2
PATZ1
PHC2
PSMD2
RNF34
RNF40
SLX4
TRAF2
TRAF5
TRIM15
TRIM21
USP25
VHL
c
H1N1WSN
H3N2
H1N1pdm09
4
6
15
6
0
1
4

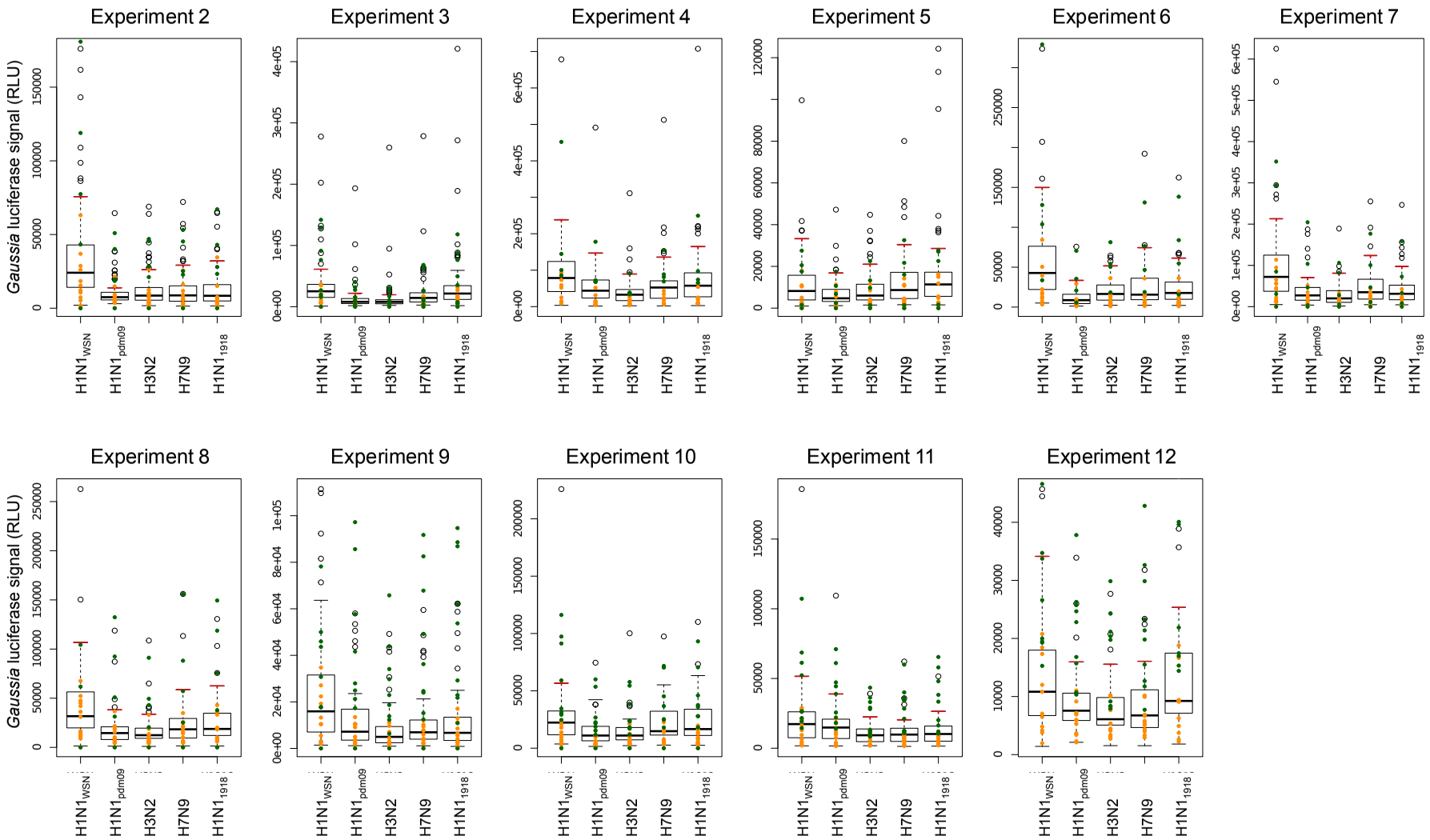
Figure 6


a
b
PB2
MGRN1
RNF40
RNF10
TRAF3IP2
ASB4
ASB10
DCAF11
RNF34
TRIM15
FBXO21
FBXO44
KBTBD5
PATZ1
FBXL19
MARCH10
OTUD6A
ABTB2
ZBTB20
PDZRN4
MPND
BTBD6
KLHL34
PAN2
TRAF2
VHL
TRAF6
DDB1
BTRC
PSMD2
PHC2
ATXN3
RIPK1
TRIM21
OTUB1
TRIM32
TRAF5
USP28
ZFP161
USP25
SLX4
1300 nodes
degree 1
10 nodes
degree 7 to 11
310 nodes
degree 2
14 nodes
degree 6
25 nodes
degree 5
95 nodes
degree 3
31 nodes
degree 4

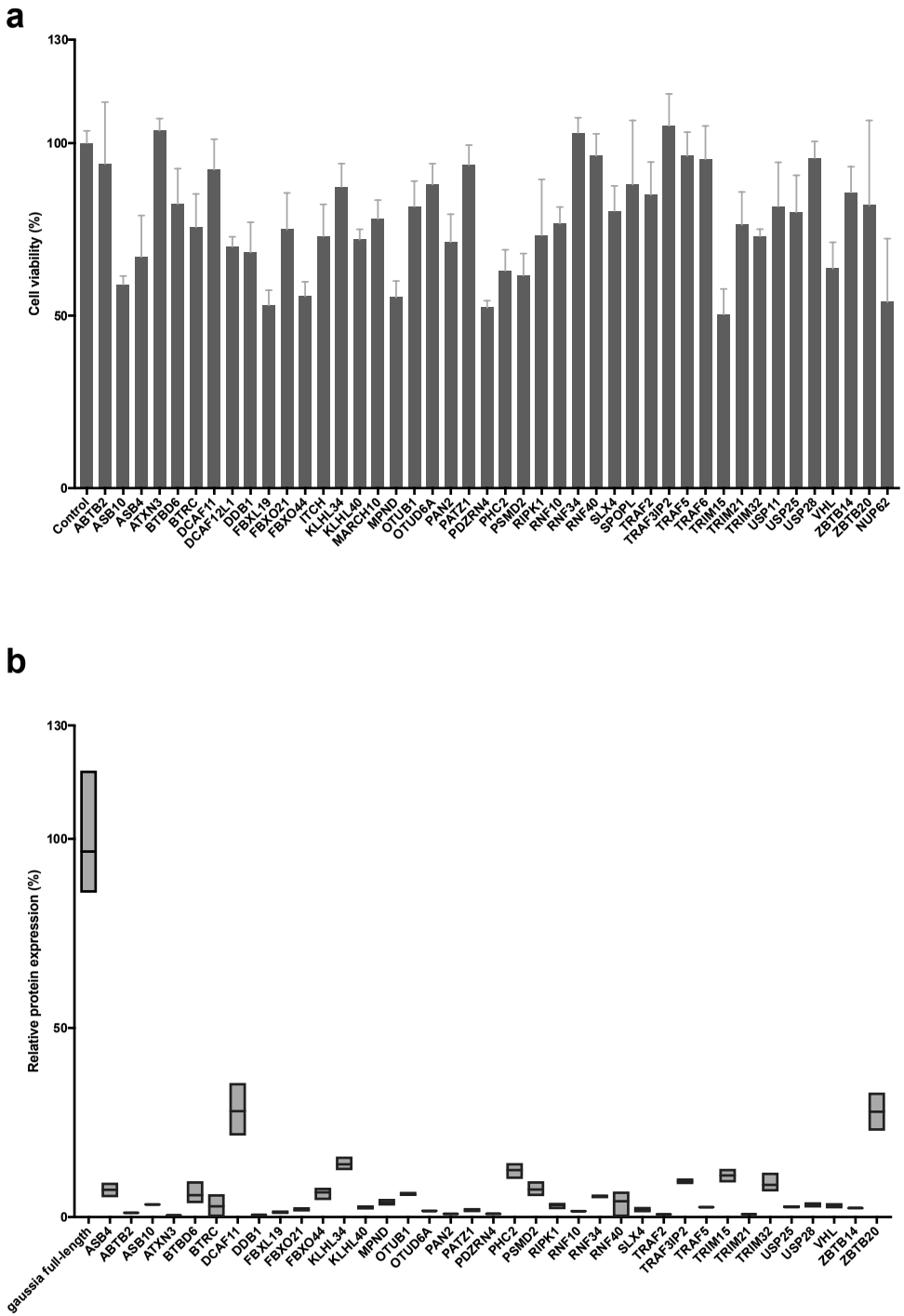
Supp Figure 1



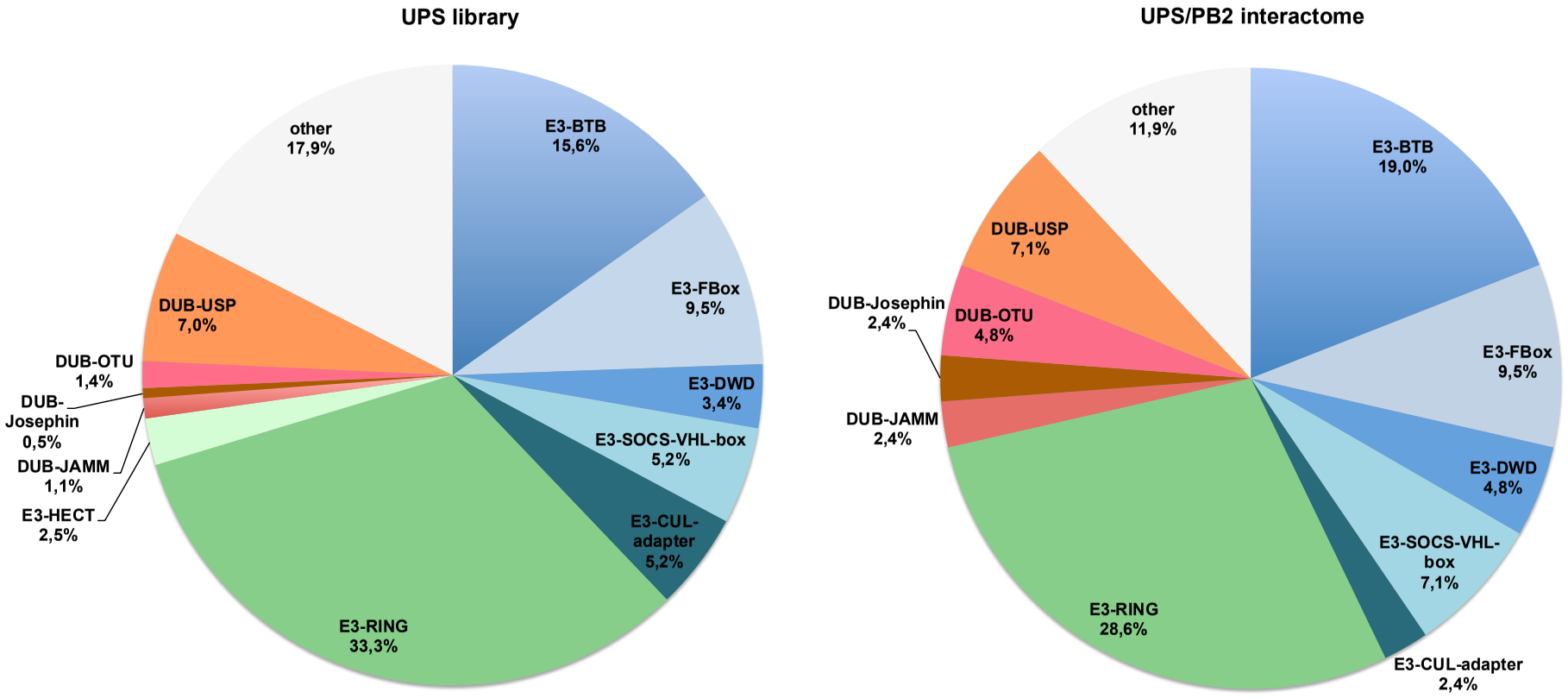
Supp Figure 2



Supp Figure 3



Supp Figure 4



The FEBS Journal

FEBS PRESS science publishing by scientists

Mapping the interactome of HPV E6 and E7 oncoproteins with the ubiquitin-proteasome system

Juline Poirson[1,*], Elise Biquand[2,*], Marie-Laure Straub[1], Patricia Cassonnet[2], Yves Nominé[3], Louis Jones[4], Sylvie van der Werf[2], Gilles Travé[3], Katia Zanier[1], Yves Jacob[2], Caroline Demeret[2,#] and Murielle Masson[1,#]

1 Ecole Supérieure de Biotechnologie Strasbourg, UMR-7242, CNRS, Université de Strasbourg, Illkirch, France
2 UMR 3569, CNRS, Unité de Génétique Moléculaire des Virus à ARN, Institut Pasteur, Université Paris Diderot, Paris, France
3 UMR 7104-Inserm U964, CNRS, IGBMC-CBI, Equipe labellisée Ligue 2015, Illkirch, France
4 Biostatistiques et biologie intégrative (C3BI), Institut Pasteur, Centre de bioinformatique, Paris, France

Keywords
HPV; interactomic; protein complementation assay; ubiquitination

Correspondence
M. Masson, Ecole Supérieure de Biotechnologie Strasbourg, UMR-7242, CNRS, Université de Strasbourg, 300, Boulevard S. Brant, 67412 Illkirch, France
Fax: +33 (0)368854683
Tel: +33 (0)368854763
E-mail: murielle.masson@unistra.fr

*Co-first authors.
#Co-last authors.

(Received 21 March 2017, revised 27 July 2017, accepted 3 August 2017)

doi:10.1111/febs.14193

Protein ubiquitination and its reverse reaction, deubiquitination, regulate protein stability, protein binding activity, and their subcellular localization. These reactions are catalyzed by the enzymes E1, E2, and E3 ubiquitin (Ub) ligases and deubiquitinases (DUBs). The Ub-proteasome system (UPS) is targeted by viruses for the sake of their replication and to escape host immune response. To identify novel partners of human papillomavirus 16 (HPV16) E6 and E7 proteins, we assembled and screened a library of 590 cDNAs related to the UPS by using the *Gaussia princeps* luciferase protein complementation assay. HPV16 E6 was found to bind to the homology to E6AP C terminus-type Ub ligase (E6AP), three really interesting new gene (RING)-type Ub ligases (MGRN1, LNX3, LNX4), and the DUB Ub-specific protease 15 (USP15). Except for E6AP, the binding of UPS factors did not require the LxxLL-binding pocket of HPV16 E6. LNX3 bound preferentially to all high-risk mucosal HPV E6 tested, whereas LNX4 bound specifically to HPV16 E6. HPV16 E7 was found to bind to several broad-complex tramtrack and bric-a-brac domain-containing proteins (such as TNFAIP1/KCTD13) that are potential substrate adaptors of Cullin 3-RING Ub ligases, to RING-type Ub ligases implicated in innate immunity (RNF135, TRIM32, TRAF2, TRAF5), to the substrate adaptor DCAF15 of Cullin 4-RING Ub ligase and to some DUBs (USP29, USP33). The binding to UPS factors did not require the LxCxE motif but rather the C-terminal region of HPV16 E7 protein. The identified UPS factors interacted with most of E7 proteins across different HPV types. This study establishes a strategy for the rapid identification of interactions between host or pathogen proteins and the human ubiquitination system.

Abbreviations
APC/C, anaphase-promoting complex/cyclosome; BTB, broad-complex tramtrack and bric-a-brac; CRL, Cullin-ring ubiquitin ligase; DUB, deubiquitinase; DWD, DBD1-binding W40 protein; Gluc1, Gaussia luciferase fragment 1; Gluc2, Gaussia luciferase fragment 2; Gluc, Gaussia luciferase; GPCA, Gaussia protein complementation assay; HECT, homology to E6AP C terminus; HPV, human papillomavirus; IQR, interquartile range; JAMM, Jab1/mov34/Mpr1 Pad1 N-terminal+; MJD, Machado/Josephin domain; NLR, normalized luminescence ratio; OTU, ovarian tumor protease; PBM, PDZ-binding motif; PPI, protein–protein interaction; PRS, positive reference set; RING, really interesting new gene; RLU, relative luciferase unit; RRS, random reference set; SOCS, suppressor of cytokine signaling; SRF, substrate recognition factor; Ubl, ubiquitin-like; Ub, ubiquitin; UCH, ubiquitin C-terminal hydrolase; UPS, ubiquitin-proteasome system; USP, ubiquitin-specific protease.

Introduction

The ubiquitin (Ub)-proteasome system (UPS) is the major pathway of intracellular protein degradation in eukaryotes [1]. Conjugation of Ub to cellular proteins is involved in mechanisms as diverse as control of cell proliferation, DNA repair, cell death, transcription, membrane trafficking, and immune response. The deregulation of ubiquitination has been implicated in human pathologies such as cancer and neurodegenerative and inflammatory diseases. Ubiquitination involves an enzymatic reaction cascade (E1 Ub-activating, E2 Ub-conjugating enzyme, and E3 Ub-ligating enzyme). Ub is covalently conjugated to a lysine residue or (less frequently) at the N-terminal α amino group of the substrate protein. Ubiquitination can be reversed by deubiquitinating proteases (DUB) [2]. The human genome encodes for two E1, 40 Ub-conjugating enzymes (E2), more than 800 E3 Ub ligase components (E3), and around 100 DUB. The E3 Ub ligases are largely responsible for substrate specificity and can be grouped into two main families, homology to E6AP C terminus (HECT)-type and the largest family of really interesting new gene (RING) finger-type Ub ligases [3]. As obligate intracellular parasites, viruses have evolved the capacity to exploit the UPS during their life cycle. Some viruses can encode their own ubiquitination enzymes. Most frequently, viruses divert cellular ubiquitination enzymes to escape host cell defense and to ensure virus propagation, by cell cycle or cell death control [4].

More than 200 different human papillomavirus (HPV) types have been characterized so far. Based on the sequence of the capsid L1 gene, HPV can be grouped into genera (such as alpha and beta). Most alpha HPVs exhibit a mucosal tropism, whereas most beta HPVs exhibit a cutaneous one. Persistent infection with high-risk mucosal HPV (such as HPV16, HPV18, HPV33) has been implicated in the carcinogenesis of cervical cancers [5], oropharyngeal cancers [6], and anal cancers [7], HPV16 type being the most frequent high-risk mucosal HPV. Beta HPV types (such as HPV5, 8) have been initially identified in skin lesions of patients with epidermodysplasia verruciformis (EV). Upon sun exposure, these lesions can evolve into squamous cell carcinoma. Recently, some evidences have suggested that some beta HPV types (HPV5, HPV8, HPV38) could promote nonmelanoma skin carcinogenesis of non-EV patients [8].

The oncogenic potential of HPV is based on the immortalizing activities of the E6 and E7 oncoproteins, which target for degradation of the tumor suppressors P53 and RB1, respectively. E6 functions as an adaptor protein to redirect the E6AP/UBE3A HECT Ub ligase to target some cellular proteins such as P53 for proteasome-dependent degradation [9–11]. On the other hand, one major activity of E7 is the binding to the members of the RB1 family (RB1, RBL1, RBL2) to trigger G1/S progression. HPV16 E7 has been detected in a complex with RB1 and a Cullin 2-based E3 Ub ligase [12] and has been shown to interact with the substrate recognition factor (SRF) Zer1 protein for RB1 degradation [13]. The half-life of HPV E6 and E7 proteins is also highly regulated, as both proteins are polyubiquitinated and degraded by the proteasome [14,15,16], while the DUBs Ub-specific protease 15 (USP15) and USP11 have been shown to stabilize E6 and E7, respectively [17] and [18].

Here, we have used the split *Gaussia princeps* luciferase protein complementation assay (GPCA) [19] to detect direct binary complexes between HPV16 E6 or E7 proteins and the UPS. For this purpose, we have assembled a library covering 50% of the entire human ubiquitination system. The library was screened against HPV16 E6 and E7 proteins and allowed us to identify novel UPS partners. We have also identified the motif/domain within HPV16 E6 and E7 proteins required for the UPS interactions. Finally, we compared the binding to the UPS factors between E6 and E7 proteins across six different HPV types.

Results

Assembling the UPS library

First, a library of cDNAs encoding for UPS factors was assembled from the ORFs encoding UPS factors present in the human ORFeome collection (CCSB, Dana-Farber Cancer Institute). We selected cDNAs encoding UPS factors (E1, E2, E3, DUB), to which we added cDNAs encoding for proteins of the Fanconi anemia complex (FA complex), the anaphase-promoting complex/cyclosome (APC/C), and some UPS-related proteins. In total, 590 cDNAs encoding for 575 unique protein entries were cloned in the pSPICA-N1 vector by Gateway cloning (Table 1). All cDNAs are flanked by attB1 recombination sites and are then easily transferred to any other compatible expression vectors. The pSPICA-N1 vector allows the expression of the UPS factor in fusion with a fragment of the *G. princeps* luciferase that can be used for the split *G. princeps* protein complementation assay (GPCA) [19]. All cDNA constructs of the library were fully sequenced and annotated. The resulting UPS library comprises five E1 or E1-like proteins, 28 E2-conjugating enzymes, 432 E3 Ub ligase components, three Ub

or Ub-like proteins, 63 DUB, 13 proteins of the APC/CDC complex, and 31 UPS-related proteins (Fig. 1). The largest protein class corresponds to E3 Ub ligases, with 14 HECT Ub ligases, 189 RING Ub ligases, and 230 potential components of Cullin-based RING Ub ligases (CRL). This last group includes (i) five Cullin proteins (Cullin 2, 3, 4A, 4B, and 5), (ii) 107 broad-complex tramtrack and bric-a-brac (BTB) domain proteins that can be SRF of Cullin 3-based Ub ligases (CRL3), (iii) 55 F-box proteins that are part of Cullin 1-based Ub ligases (CRL1), (iv) 27 suppressor of cytokine signaling (SOCS) box proteins that interact with Elongin B and C and are part of Cullin 5-based Ub ligases (CRL5), (v) 28 DBD1-binding W40 protein (DWD)-containing proteins that bind DDB1 and act as SRFs of Cullin 4-based Ub ligases (CRL4). The DUB family is represented by 39 USP-, eight ovarian tumor protease (OTU)-, eight Jab1/mov34/Mpr1 Pad1 N-terminal+ (JAMM)/MPN-, four Ub C-terminal hydrolase (UCH)-, three Machado/Josephin domain (MJD)-containing DUBs. The Ub and ubiquitin-like (Ubl) Conjugation Database (www.uucd.biocuckoo.org) has recently annotated and classified human UPS effectors according to the presence of motif/domain signature [20]. By comparison to this database, the current version of our UPS library contains 55% of the E1 and E1-like proteins, 73% of the E2-conjugating enzymes, 100% of the APC/CDC complex, 48% of the HECT Ub ligases, 42% of the RING finger Ub ligases, 44% of the Cullin-based Ub ligases (75% of BTB domain proteins, 93% of F-box proteins, 66% of SOCS proteins, and 10% of DWD domain proteins), and 51% of DUB. The current UPS library covers around 50% of the entire human ubiquitination system.

Screening of the UPS library against HPV16 E6 and E7 proteins

We used GPCA to perform a systematic unbiased screening for binary interactions between viral proteins and the UPS. The UPS library was screened against HPV16 E6 and E7 proteins. In a different work, the UPS library was also screened against the PB2 proteins of human influenza viruses (Biquand E. *et al.* manuscript in preparation). Briefly in GPCA, the luciferase activity is restored leading to luminescence signal when the two fragments of the Gaussia luciferase are in close contact resulting in the interaction between the prey and the bait fused to Gaussia luciferase fragments 1 and 2 (Gluc1 and Gluc2), respectively. To verify that the GPCA discriminates between positive and negative interactions, HPV16 E6 and E7 proteins were screened against (a) a random reference set (RRS) containing nine proteins supposedly nonbinder proteins and (b) a positive reference set (PRS) containing well-established binding partners (including E6AP [9], hADA3 [21], MAGI-1 [22], SCRIB [23], DLG3 [24], an homolog of DLG1 [25], IRF3 [26] for E6 and RB1 [27], RBL1 [28], EEF1A1 [29], PKM2 [29], PTPN14 [30] for E7). The GPCA signals obtained with the PRS were significantly higher than the ones obtained with RRS (Fig. 2, Table S1).

Then, we estimated the 'background' GPCA signal of the UPS library. For this purpose, the UPS library containing 590 Gluc1-fused UPS proteins was tested against the complementary Gluc2 fragment alone. The vast majority of constructs gave a weak GPCA signal. Representation of frequency as a function of the log10 of the luminescence values using a Kernel density (KDE) plot function results in a nearly normal distribution (Fig. 3A). The small shoulder on the right side of the distribution curve at 4.2 corresponds to 26 Gluc1-UPS factors/Gluc2 protein pairs including Gluc1-TRAF3IP2, which gives an exceptionally high luminescence value (Table S2). Given their high background signals, these 26 Gluc1-UPS proteins should be considered with high precaution when tested with a Gluc2-fused bait.

Each screen of the UPS library with either HPV16 E6 or E7 has been performed in a single transfection experiment, thus allowing for the assessment of the entire UPS library at once. A small fraction of UPS factors generated 'outlier' luminescence values (see the shoulders on the right of the curve) by comparison to noninteracting pairs (which is the case for the majority of the UPS factors; Fig. 3B,C). Thus, here we considered as potential interacting pairs, those with a GPCA signal above a standard threshold (T) on a Whisker plot representation. In this way, we found 81 and 65 outlier luminescence values for the E6 and E7 proteins, respectively (Fig. 3D–E and Table S3).

Retesting protein pairs by GPCA

To confirm the protein pairs selected from the initial screen, we retested them individually by GPCA and calculated the normalized luminescence ratio (NLR), as defined in [19]. The NLR takes into account the background GPCA signal, which was measured for each interaction partner as described above. In addition, the luminescence values obtained with the RRS and E6 or E7 protein were used to calculate a confidence interval. We thus considered a protein pair as interacting if the NLR was above the previously defined threshold of 3.5 [19] and above the upper limit

Table 1. Composition of the UPS library (symbol, Gene ID, Uniprot number, protein length, isoform), classification (E1, E2, E3, DUB, other), and annotation of domain/motif.

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
ABTB1	80325	Q969K4	478		2	E3	E3 CRL adaptor	BTB (POZ) domain
ABTB2	25841	Q8N961	839		2	E3	E3 CRL adaptor	BTB (POZ) domain
AIMP2	7965	Q13155	320			Other		
ANAPC1	64682	Q9H1A4	1944	Deletion 329–1944		Other	APC/CDC complex	
ANAPC10	10393	Q9UM13	185			Other	APC/CDC complex	
ANAPC11	51529	Q9NYG5	196		2	E3	E3-RING APC/CDC complex	RING-type zinc finger
ANAPC13	25847	Q9BS18	74			Other	APC/CDC complex	
ANAPC16	119504	Q96DE5	110			Other	APC/CDC complex	
ANAPC5	51433	Q9UJX4	755		1	Other	APC/CDC complex	
ANAPC7	51434	Q9UJX3	599	Deletion 1–34	1	Other	APC/CDC complex	
ARIH1	25820	Q9Y4X5	557			E3	E3-RING	RING- and IBR-type zinc fingers
ARIH2	10425	O95376	493			E3	E3-RING	RING- and IBR-type zinc fingers
ARMC5	79798	Q96C12	771		2	E3	E3 CRL adaptor	BTB (POZ) domain
ASB10	136371	Q8WXI3	467		1	E3	E3 CRL adaptor	SOCS box domain
ASB11	140456	Q8WXH4	323		1	E3	E3 CRL adaptor	SOCS box domain
ASB13	79754	Q8WXK3	173		2	E3	E3 CRL adaptor	SOCS box domain
ASB14	142686	A6NK59	302		2	E3	E3 CRL adaptor	SOCS box domain
ASB16	92591	Q96NS5	453			E3	E3 CRL adaptor	SOCS box domain
ASB17	127247	Q8WXJ9	295			E3	E3 CRL adaptor	SOCS box domain
ASB2	51676	Q96Q27	587		1	E3	E3 CRL adaptor	SOCS box domain
ASB3	51130	Q9Y575	518		1	E3	E3 CRL adaptor	SOCS box domain
ASB4	51666	Q9Y574	349		2	E3	E3 CRL adaptor	SOCS box domain
ASB5	140458	Q8WWX0	329		1	E3	E3 CRL adaptor	SOCS box domain
ASB6	140459	Q9NWX5	197		2	E3	E3 CRL adaptor	SOCS box domain
ASB6	140459	Q9NWX5	421		1	E3	E3 CRL adaptor	SOCS box domain
ASB7	140460	Q9H672	318		2	E3	E3 CRL adaptor	SOCS box domain
ASB8	140461	Q9H765	288			E3	E3 CRL adaptor	SOCS box domain
ASB9	140462	Q96DX5	252		3	E3	E3 CRL adaptor	SOCS box domain
ASB9	140462	Q96DX5	294		1	E3	E3 CRL adaptor	SOCS box domain
ATG7	74244	O95352	676		2	E1	E1-like	
ATXN3	4287	P54252	361		2	DUB	DUB	Josephin
ATXN7L3	56970	Q14CW9	347		1	DUB	SAGA complex	
BABAM1	29086	Q14CW9	329		1	DUB	BRCA1-A complex	
BACH1	571	O14867	736			E3	E3 CRL adaptor	BTB (POZ) domain
BAP1	8314	Q92560	729			DUB	DUB	UCH
BCL6	604	P41182	706		1	E3	E3 CRL adaptor	BTB (POZ) domain
BCL6B	255877	Q8N143	479			E3	E3 CRL adaptor	BTB (POZ) domain
BFAR	51283	Q9NZS9	450		1	E3	E3-RING	RING-type zinc finger
BIRC2	329	Q13490	618		1	E3	E3-RING	RING-type zinc finger
BIRC3	330	Q13489	604			E3	E3-RING	RING-type zinc finger
BIRC7	79444	Q96CA5	298		2	E3	E3-RING	RING-type zinc finger
BIRC8	112401	Q96P09	236			E3	E3-RING	RING-type zinc finger
BMI1	648	P35226	326			E3	E3-RING	RING-type zinc finger
BRCA1	672	P38398	1863	Deletion 1–530	1	E3	E3-RING	RING-type zinc finger
BRCA1	672	P38398	699		6	E3	E3-RING	RING-type zinc finger
BRCC3	79184	P46736	291	Deletion 7–12	2	DUB	DUB	JAMM/MPN
BRE	9577	Q9NXR7	383		1	E3/DUB	BRCA1-A complex	
BRE1A	56254	Q5VTR2	975			E3	E3-RING	RING-type zinc finger
BTBD1	53339	Q9H0C5	482		1	E3	E3 CRL adaptor	BTB (POZ) domain

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
BTBD10	84280	Q9BSF8	475	Deletion 1–291		E3	E3 CRL adaptor	BTB (POZ) domain
BTBD15	29068	Q8NCP5	570		1	E3	E3 CRL adaptor	BTB (POZ) domain
BTBD15	29068	Q8NCP5	453		3	E3	E3 CRL adaptor	BTB (POZ) domain
BTBD16	118663	Q32M84	507		2	E3	E3 CRL adaptor	BTB (POZ) domain
BTBD18	643376	B2RXH4	712			E3	E3 CRL adaptor	BTB (POZ) domain
BTBD6	90135	Q96KE9	410		2	E3	E3 CRL adaptor	BTB (POZ) domain
BTBD8	284697	Q5XKL5	378		1	E3	E3 CRL adaptor	BTB (POZ) domain
BTBD9	114781	Q96Q07	612			E3	E3 CRL adaptor	BTB/POZ domain
BTRC	8945	Q9Y297	605		1	E3	E3 CRL adaptor	F-box domain, WD repeats
C19orf40	91442	Q9BTP7	215			Other	FA complex	
CDC16	8881	Q13042	620		1	Other	APC/CDC complex	
CDC20	991	Q12834	499			Other	APC/CDC complex	WD repeats
CDC23	8697	Q9UJX2	597	Deletion 1–7	1	Other	APC/CDC complex	
CDC26	246184	Q8NHZ8	85			Other	APC/CDC complex	
CDC27	996	P30260	824		1,2	Other	APC/CDC complex	
CDH1	51343	Q9UM11	493		2	Other	APC/CDC complex	WD repeats
CGRRF1	10668	Q99675	332			E3	E3-RING	RING-type zinc finger
CNOT4	4850	O95628	572		2	E3	E3-RING	RING-type zinc finger
COPS5	10987	Q92905	334			DUB	DUB/COP9 signalosome	JAMM/MPN
COPS7A	50813	Q9UBW8	275			Other	COP9 signalosome subunit	
CUL2	8453	Q13617	745		1	E3	E3-CRL component	Cullin
CUL3	8452	Q13618	768		1	E3	E3-CRL component	Cullin
CUL4A	8451	Q13619	759		2	E3	E3-CRL component	Cullin
CUL4B	8450	Q13620	913		1	E3	E3-CRL component	Cullin
CUL5	8565	Q93034	780			E3	E3-CRL component	Cullin
CYHR1	50626	Q6ZMK1	192		3	E3	E3-RING	RING-type zinc finger
DCAF10	79269	Q5QP82	522		2	E3	E3 CRL adaptor	WD repeats
DCAF10	79269	Q5QP82	559	Deletion 1–301	1	E3	E3 CRL adaptor	WD repeats
DCAF11	80344	Q8TEB1	546		1	E3	E3 CRL adaptor	WD repeats
DCAF12	25853	Q5T6F0	453			E3	E3 CRL adaptor	WD repeats
DCAF12L1	139170	Q5VU92	463			E3	E3 CRL adaptor	WD repeats
DCAF15	90379	Q66K64	600			E3	E3 CRL adaptor	WD repeats
DCAF16	54876	Q9NXF7	216			E3	E3 CRL adaptor	WD repeats
DCAF4	26094	Q8WV16	495		1	E3	E3 CRL adaptor	WD repeats
DCAF4	26094	Q8WV16	395		4	E3	E3 CRL adaptor	WD repeats
DCAF4L1	285429	Q3SXM0	396			E3	E3 CRL adaptor	WD repeats
DCAF4L2	138009	Q8NA75	395			E3	E3 CRL adaptor	WD repeats
DCAF5	8816	Q96JK2	942	Deletion 294–942	1	E3	E3 CRL adaptor	WD repeats
DCAF7	10238	P61962	342		1	E3	E3 CRL adaptor	WD repeats
DCAF8	50717	Q5TAQ9	597		1	E3	E3 CRL adaptor	WD repeats
DDB1	1642	Q16531	1140		1	E3	E3 CRL adaptor	WD repeat beta propeller
DDB2	1643	Q92466	427		1	E3	E3 CRL adaptor	WD repeat beta propeller
DTL	51514	Q9NZJ0	730		1	E3	E3 CRL adaptor	WD repeats
DTX1	1840	Q86Y01	620			E3	E3-RING	RING-type zinc finger
DTX3	196403	Q8N9I9	347		1	E3	E3-RING	RING-type zinc finger

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
DTX3L	151636	Q8TDB6	740		1	E3	E3-RING	RING-type zinc finger
DTX4	23220	Q9Y2E6	513		2	E3	E3-RING	RING-type zinc finger
DZIP3	9666	Q86Y13	1208		1	E3	E3-RING	RING-type zinc finger
E6AP/ UBE3A	7337	Q05086	875	Deletion 1–290		E3	E3-HECT	HECT
ECT2L	345930	Q008S8	904			E3	E3 CRL adaptor	F-box domain
EED	267	O75530	466		2	E3	E3 CRL adaptor	WD repeats
ENC1	8507	O14682	489		1	E3	E3 CRL adaptor	BTB (POZ) domain
ERCC8	1161	Q13216	205		2	E3	E3 CRL adaptor	WD repeats
FAM175A	84142	Q6UWZ7	409	Deletion de 1–109		E3/DUB	BRCA1-A complex	MPN like
FAM175B	23172	Q15018	415	Deletion de 1–104		DUB	BRISC complex	MPN like
FANCA	2175	Q0VAP4	265		2	Other	FA complex	
FANCB	2187	Q8NB91	859			Other	FA complex	
FANCC	2176	Q9Y2Y4	487			Other	FA complex	
FANCD2	2177	Q9BXW9	241		4	Other	FA complex	
FANCE	2178	Q9HB96	536			Other	FA complex	
FANCF	2188	Q9NPI8	374			Other	FA complex	
FANCG	2189	O15287	622			Other	FA complex	
FANCJ	83990	Q9BX63	1249		1	Other	FA complex	
FANCL	55120	Q9WN38	375		1	E3	E3-RING FA complex	RING-type zinc finger
FANCM	57697	Q8IYD8	669		2	Other	FA complex	
FBXL12	54850	Q9NXK8	326		1	E3	E3 CRL adaptor	F-box domain
FBXL13	222235	Q8NEE6	735		1	E3	E3 CRL adaptor	F-box domain
FBXL14	144699	Q8N1E6	418			E3	E3 CRL adaptor	F-box domain
FBXL15	79176	Q9H469	300			E3	E3 CRL adaptor	F-box domain
FBXL16	146330	Q8N461	479		1	E3	E3 CRL adaptor	F-box domain
FBXL17	64839	Q9UF56	303		2	E3	E3 CRL adaptor	F-box domain
FBXL18	80028	Q96ME1	365		3	E3	E3 CRL adaptor	F-box domain
FBXL18	80028	Q96D16	259			E3	E3 CRL adaptor	F-box domain
FBXL19	54620	Q6PCT2	694		1,2	E3	E3 CRL adaptor	F-box domain
FBXL2	25827	Q9UKC9	423		1	E3	E3 CRL adaptor	F-box domain
FBXL20	84961	Q96IG2	436		1	E3	E3 CRL adaptor	F-box domain
FBXL3	26224	Q9UKT7	428			E3	E3 CRL adaptor	F-box domain
FBXL4	26235	Q9UKA2	621			E3	E3 CRL adaptor	F-box domain
FBXL5	26234	Q9UKA1	691		1	E3	E3 CRL adaptor	F-box domain
FBXL6	26233	Q8N531	539		1	E3	E3 CRL adaptor	F-box domain
FBXL7	23194	Q9UJT9	491		1	E3	E3 CRL adaptor	F-box domain
FBXL8	55336	Q96CD0	374			E3	E3 CRL adaptor	F-box domain
FBXO10	26267	Q9UK96	481		2	E3	E3 CRL adaptor	F-box domain
FBXO11	80204	Q86XK2	843		6	E3	E3 CRL adaptor	F-box domain
FBXO15	201456	Q8NCQ5	434		2	E3	E3 CRL adaptor	F-box domain
FBXO16	1575 74	Q8IX29	292		1	E3	E3 CRL adaptor	F-box domain
FBXO2	26232	Q9UK22	296			E3	E3 CRL adaptor	F-box domain
FBXO21	23014	Q8IUQ5	477		CRA_b	E3	E3 CRL adaptor	F-box domain
FBXO22	26263	Q8NEZ5	403		1	E3	E3 CRL adaptor	F-box domain
FBXO24	26261	O75426	580		1	E3	E3 CRL adaptor	F-box domain
FBXO25	26260	Q8TCJ0	291		3	E3	E3 CRL adaptor	F-box domain
FBXO27	126433	Q8NI29	283			E3	E3 CRL adaptor	F-box domain
FBXO28	23219	Q9NVF7	368		1	E3	E3 CRL adaptor	F-box domain
FBXO3	26273	Q9UK99	471		1	E3	E3 CRL adaptor	F-box domain
FBXO30	84085	Q8TB52	745	Deletion 1–70		E3	E3 CRL adaptor	F-box domain

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
FBXO32	114907	Q0VAQ6	210			E3	E3 CRL adaptor	F-box domain
FBXO33	254170	Q7Z6M2	555			E3	E3 CRL adaptor	F-box domain
FBXO34	55030	Q9NWN3	711			E3	E3 CRL adaptor	F-box domain
FBXO38	81545	Q6PIJ6	1117		2	E3	E3 CRL adaptor	F-box domain
FBXO38	81545	Q6PIJ6	943		3	E3	E3 CRL adaptor	F-box domain
FBXO39	162517	Q8N4B4	442	Deletion 177–442		E3	E3 CRL adaptor	F-box domain
FBXO4	26272	Q9UKT5	387		1	E3	E3 CRL adaptor	F-box domain
FBXO40	51725	Q9UH90	709			E3	E3 CRL adaptor	F-box domain
FBXO42	54455	Q6P3S6	717			E3	E3 CRL adaptor	F-box domain
FBXO44	93611	Q9H4M3	224		2	E3	E3 CRL adaptor	F-box domain
FBXO48	554251	Q5FWF7	155			E3	E3 CRL adaptor	F-box domain
FBXO5	26271	Q9UKT4	447		1	E3	E3 CRL adaptor	F-box domain
FBXO6	26270	Q9NRD1	293			E3	E3 CRL adaptor	F-box domain
FBXO7	25793	Q9Y3I1	522		1	E3	E3 CRL adaptor	F-box domain
FBXO9	26268	Q9UK97	403		3	E3	E3 CRL adaptor	F-box domain
FBXW10	10517	Q5XX13	1052		2	E3	E3 CRL adaptor	F-box domain, WD repeats
FBXW11	23291	Q9UKB1	529		B	E3	E3 CRL adaptor	F-box domain, WD repeats
FBXW12	285231	Q6X9E4	394		2	E3	E3 CRL adaptor	F-box domain, WD repeats
FBXW12	285231	Q494Z1	262			E3	E3 CRL adaptor	F-box domain, WD repeats
FBXW4	6468	P57775	412			E3	E3 CRL adaptor	F-box domain, WD repeats
FBXW7	55294	Q969H0	707		1	E3	E3 CRL adaptor	F-box domain, WD repeats
FEM1C	56929	Q96JP0	617			Other		
G2E3	55632	Q7L622	706			E3	E3-HECT	HECT
GLMN	11146	Q92990	594		1	E3	E3 CRL adaptor	Coiled coil domain
GMCL1	64395	Q96IK5	515			E3	E3 CRL adaptor	BTB (POZ) domain
GNB2	2783	P62879	340			E3	E3 CRL adaptor	WD repeat
HECTD3	79654	Q5T447	861	Deletion 1–284	1	E3	E3-HECT	HECT
HECW1	23072	Q76N89	1606		1,2	E3	E3-HECT	HECT
HERC3	8916	Q15034	368		2	E3	E3-HECT	HECT
HERC4	26091	Q5GLZ8	1057		1,2,	E3	E3-HECT	HECT
HERC6	55008	Q8IVUE3	1022	Deletion 1–643		E3	E3-HECT	HECT
HERPUD1	9709	Q15011	391			DUB	DUB	ERAD Ub-like domain
HSP90B1	7184	Q96GW1	315			Other		
IQUB	154865	Q8NA54	791	Deletion 374–791		Other		Ub-like domain
ITCH	83732	Q96J02	862	Deletion 1–16	2	E3	E3-HECT	HECT
JOSD1	9929	Q15040	202			DUB	DUB	Josephin
JOSD2	126119	Q8TAC2	188		1	DUB	DUB	Josephin
KBTBD2	25948	Q8IY47	623			E3	E3 CRL adaptor	BTB (POZ) domain
KBTBD3	143879	Q8NAB2	608			E3	E3 CRL adaptor	BTB (POZ) domain
KBTBD4	55709	Q9NVX7	518		1	E3	E3 CRL adaptor	BTB (POZ) domain
KBTBD5	131377	Q2TBA0	621		1	E3	E3 CRL adaptor	BTB (POZ) domain
KBTBD7	84078	Q8WVZ9	684			E3	E3 CRL adaptor	BTB (POZ) domain
KCNRG	283518	Q8N5I3	272			E3	E3 CRL adaptor	BTB/POZ domain
KCTD1	284252	Q719H9	257			E3	E3 CRL adaptor	BTB (POZ) domain
KCTD13	253980	Q8WZ19	329			E3	E3 CRL adaptor	BTB (POZ) domain
KCTD14	65987	Q9BQ13	255		2	E3	E3 CRL adaptor	BTB (POZ) domain
KCTD15	79047	Q96SI1	283		2	E3	E3 CRL adaptor	BTB (POZ) domain

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
KCTD16	57528	Q68DU8	428			E3	E3 CRL adaptor	BTB (POZ) domain
KCTD17	79734	Q8N5Z5	321		1	E3	E3 CRL adaptor	BTB (POZ) domain
KCTD17	79734	Q8N5Z5	297	Deletion 1–7	2	E3	E3 CRL adaptor	BTB (POZ) domain
KCTD19	146212	Q17RG1	926			E3	E3 CRL adaptor	BTB (POZ) domain
KCTD21	283219	Q4G0X4	260			E3	E3 CRL adaptor	BTB (POZ) domain
KCTD3	51133	Q9Y597	815		1	E3	E3 CRL adaptor	BTB (POZ) domain
KCTD3	51133	Q9Y597	813		2	E3	E3 CRL adaptor	BTB (POZ) domain
KCTD4	386618	Q8WVF5	259			E3	E3 CRL adaptor	BTB (POZ) domain
KCTD5	54442	Q9NXV2	234			E3	E3 CRL adaptor	BTB (POZ) domain
KCTD6	200845	Q8NC69	237			E3	E3 CRL adaptor	BTB (POZ) domain
KCTD7	154881	Q96MP8	289		1,2	E3	E3 CRL adaptor	BTB (POZ) domain
KCTD8	386617	Q6ZWB6	473			E3	E3 CRL adaptor	BTB (POZ) domain
KCTD9	54793	Q7L273	389			E3	E3 CRL adaptor	BTB (POZ) domain
KEAP1	9817	Q14145	624			E3	E3 CRL adaptor	BTB (POZ) domain
KIAA0317	9870	O15033	789		2	E3	E3-HECT	HECT
KIAA0368	23392	Q5VYK3	1845			Other		
KLHL1	57626	Q9NR64	748			E3	E3 CRL adaptor	BTB (POZ) domain
KLHL10	317719	Q6JEL2	608			E3	E3 CRL adaptor	BTB/POZ domain
KLHL11	55175	Q9NVR0	708			E3	E3 CRL adaptor	BTB (POZ) domain
KLHL12	59349	Q53G59	568		1	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL13	90293	Q9P2N7	655		1	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL14	57565	Q9P2G3	388		2	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL15	80311	Q96M94	604			E3	E3 CRL adaptor	BTB (POZ) domain
KLHL16	8139	Q9H2C0	597			E3	E3 CRL adaptor	BTB (POZ) domain
KLHL17	339451	Q86XA8	187			E3	E3 CRL adaptor	BTB (POZ) domain
KLHL18	23276	O94889	509		2	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL2	11275	O95198	593		1	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL20	27252	Q9BS75	230			E3	E3 CRL adaptor	BTB (POZ) domain
KLHL21	9903	Q9UJP4	597			E3	E3 CRL adaptor	BTB (POZ) domain.
KLHL22	84861	Q53GT1	634		1	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL23	151230	Q8NBE8	558			E3	E3 CRL adaptor	BTB (POZ) domain
KLHL25	64410	Q9H0H3	589	Deletion 231–588		E3	E3 CRL adaptor	BTB (POZ) domain
KLHL26	55295	Q53HC5	615			E3	E3 CRL adaptor	BTB (POZ) domain
KLHL28	54813	Q9NXS3	571		1	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL29	114818	Q96CT2	702	Deletion 1–200	2	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL3	26249	Q9UH77	587	Deletion 302–587	A	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL32	114792	Q96NJ5	620	Deletion 453–620	1	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL34	257240	Q8N239	644			E3	E3 CRL adaptor	BTB (POZ) domain
KLHL35	283212	Q6PF15	583	Deletion 1–220	1	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL4	56062	Q9C0H6	718		1	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL6	89857	Q8WZ60	621	Deletion 1–11	1	E3	E3 CRL adaptor	BTB (POZ) domain
KLHL9	55958	Q9P2J3	617			E3	E3 CRL adaptor	BTB (POZ) domain
LGALS3BP	3959	Q08380	585			E3	E3 CRL adaptor	BTB (POZ) domain
LNX1	84708	Q8TBB1	632		2	E3	E3-RING	RING-type zinc finger, PDZ
LNX2	222484	Q8N448	690			E3	E3-RING	RING-type zinc finger
LNX3	23024	Q9UPQ7	1066			E3	E3-RING	RING-type zinc finger
LNX4	29951	Q6ZMN7	778		2	E3	E3-RING	RING-type zinc finger
LONRF1	91694	Q17RB8	773			E3	E3-RING	RING-type zinc finger
LONRF2	164832	Q1L5Z9	511			E3	E3-RING	RING-type zinc finger
LONRF3	79836	Q496Y0	759		1,2	E3	E3-RING	RING-type zinc finger
LRR1	122769	Q96L50	414			E3	E3 CRL adaptor	
LRRC29	26231	Q8WV35	223			E3	E3 CRL adaptor	F-box domain
LRSAM1	90678	Q6UWE0	723			E3	E3-RING	RING-type zinc finger

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
LZTR1	8216	Q8N653	840			E3	E3 CRL adaptor	BTB (POZ) domain
MARCH 2	51257	Q9P0N8	246			E3	E3-RING	RING-type zinc finger
MARCH10	162333	Q8NA82	808	Deletion 657–808		E3	E3-RING	RING-type zinc finger
MARCH3	115123	Q86UD3	253			E3	E3-RING	RING-type zinc finger
MARCH5	54708	Q9NX47	278			E3	E3-RING	RING-type zinc finger
MARCH8	220972	Q5T0T0	291			E3	E3-RING	RING-type zinc finger
MDM2	4193	Q00987	491			E3	E3-RING	RING-type zinc finger
MDM4	4194	O15151	490			E3	E3-RING	RING-type zinc finger
MED8	112950	Q96G25	268			E3	E3 CRL adaptor	BC-box domain
MGRN1	23295	O60291	576			E3	E3-RING	RING-type zinc finger
MID1	4281	O15344	667			E3	E3-RING	B-box and RING-type zinc fingers
MKRN2	23609	Q9H000	416			E3	E3-RING	RING-type zinc finger
MKRN3	7681	Q13064	507			E3	E3-RING	RING- type zinc finger
MNAT1	4331	P51948	309			E3	E3-RING	RING-type zinc finger
MPND	84954	Q8N594	471			DUB	DUB	JAMM/MPN
MSL2	55167	Q9HCI7	577			E3	E3-RING	RING-type zinc finger
MYLIP	29116	Q8WY64	445			E3	E3-RING	RING-type zinc finger
MYNN	55892	Q9NPC7	610			E3	E3 CRL adaptor	BTB (POZ) domain
NACC1	112939	Q96RE7	527			E3	E3 CRL adaptor	BTB (POZ) domain
NAE1	8883	Q13564	534			E1	E1 (for NEDD8)	
NDFIP1	80762	Q9BT76	221			E3	E3-HECT activator	
NDFIP2	54602	Q9NV92	336	Deletion 1–94		E3	E3-HECT activator	
NEDD4L	23327	Q96PU5	975			E3	E3-HECT	HECT
NEURL1	9148	O76050	574		1	E3	E3-RING	RING-type zinc finger
NHLRC1	378884	Q6VVB1	395			E3	E3-RING	RING-type zinc finger
NSMCE1	1973 70	B2RDU2	256			E3	E3-RING	RING-type zinc finger
OTUB1	5561 1	Q96FW1	271		1	DUB	DUB	OTU
OTUB2	78990	Q96DC9	73		2	DUB	DUB	OTU
OTUD2	55432	Q5VVQ6	348		1	DUB	DUB	OTU
OTUD3	23252	Q5T2D3	398	Deletion 1–221		DUB	DUB	OTU
OTUD4	54726	Q01804	1049		3	DUB	DUB	OTU
OTUD6A	139562	Q7L8S5	288			DUB	DUB	OTU
OTUD6B	51633	Q8N6M0	293		1	DUB	DUB	OTU
PAN2	9924	Q504Q3	1202			DUB	DUB	USP
PATZ1	23598	Q9HBE1	537		4	E3	E3 CRL adaptor	BTB (POZ) domain
PCBP2	5094	Q15366	365		3,6	Other		
PCGF1	84759	Q9BSM1	259	Deletion 1–12	1	E3	E3-RING	RING-type zinc finger
PCGF5	84333	Q86SE9	256		1	E3	E3-RING	RING-type zinc finger
PCGF6	84108	Q9BYE7	275		3	E3	E3-RING	RING-type zinc finger
PHC2	1912	Q8IXK0	972	Deletion 1–535		Other		
PHF7	51533	Q9BWX1	381		1	E3	E3-RING	RING-type zinc finger
PJA1	64219	Q8NG27	643		1	E3	E3-RING	RING-type zinc finger
PJA2	9867	O43164	708		1	E3	E3-RING	RING-type zinc finger
PML	5371	P29590 3	781		13	E3	E3-RING	B-box and RING-type zinc fingers
POMP	51371	Q9Y244	141	Insertion of 28 aa upstream Met		Other	Proteasome chaperon	
PSMA8	143471	Q8TAA3	256	Deletion 243–256	1	Other		
PSMD14	10213	O00487	310			DUB	DUB	JAMM/MPN
PSMD2	5708	Q13200	908		1	Other	Proteasome subunit	
RAB40A	142684	Q8WXH6	277			E3	E3 CRL adaptor	SOCS box domain
RAB40AL	282808	P0C0E4	278			E3	E3 CRL adaptor	SOCS box domain
RAB40B	10966	Q12829	278			E3	E3 CRL adaptor	SOCS box domain

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
RAB40C	57799	Q96S21	281		1	E3	E3 CRL adaptor	SOCS box domain
RABGEF1	27342	Q9UJ41	491		2	E3	E3	A20 type zinc finger
RAG1	5896	P15918	931		2	E3	E3-RING	RING-type zinc finger
RBBP6	5930	Q7Z6E9	118		3	E3	E3-RING	RING-type zinc finger
RBX1	9978	P62877	108			E3	E3 CRL component	RING-type zinc finger
RC3H2	54542	Q9HBD1	506		3	E3	E3-RING	RING-type zinc finger
RCBTB1	55213	Q8NDN9	355		2	E3	E3 CRL adaptor	BTB (POZ) domain
RCBTB2	1102	O95199	551		1	E3	E3 CRL adaptor	BTB (POZ) domain
RCHY1	25898	Q96PM5	261		1	E3	E3-RING	RING-type zinc finger
RCHY1	25898	Q96PM5	179	Deletion 171–180	3	E3	E3-RING	RING-type zinc finger
RFFL	117584	Q8WZ73	327		3	E3	E3-RING	RING-type zinc finger
RFPL3	10738	O75679	288		2	E3	E3-RING	RING-type zinc finger
RFWD2	64326	Q8NHY2	731			E3	E3-RING	RING-type zinc finger
RFWD3	55159	Q6PCD5	774	Deletion 1–26		E3	E3-RING	RING-type zinc finger
RHOBTB1	9886	O94844	696			E3	E3 CRL adaptor	BTB (POZ) domain
RHOBTB2	23221	Q9BYZ6	727		1	E3	E3 CRL adaptor	BTB (POZ) domain
RHOBTB3	22836	O94955	611			E3	E3 CRL adaptor	BTB (POZ) domain
RING1	6015	Q06587	406		1	E3	E3-RING	RING-type zinc finger
RIPK1	8737	Q13546	671		1	Other	Kinase	
RLIM	51132	Q9NVW2	624		1	E3	E3-RING	RING-type zinc finger
RNF10	9921	Q8N5U6	811		1,2	E3	E3-RING	RING-type zinc finger
RNF103	7844	O00237	685			E3	E3-RING	RING-type zinc finger
RNF11	26994	Q9Y3C5	154			E3	E3-RING	RING-type zinc finger
RNF111	54778	Q6ZNA4	986		2	E3	E3-RING	RING-type zinc finger
RNF113A	7737	O15541	343			E3	E3-RING	RING-type zinc finger
RNF114	55905	Q9Y508	228		1	E3	E3-RING	RING-type zinc finger
RNF115	27246	Q9Y4L5	304			E3	E3-RING	RING-type zinc finger
RNF125	54941	Q96EQ8	232			E3	E3-RING	RING-type zinc finger
RNF126	55658	Q9BV68	326		1	E3	E3-RING	RING-type zinc finger
RNF126	55658	Q9BV68	311		2	E3	E3-RING	RING-type zinc finger
RNF13	11342	O43567	381		1	E3	E3-RING	RING-type zinc finger
RNF130	55819	Q2HIY3	276			E3	E3-RING	RING-type zinc finger
RNF133	168433	Q8WVZ7	376			E3	E3-RING	RING-type zinc finger
RNF135	84282	Q8IUD6	432		1	E3	E3-RING	RING-type zinc finger
RNF138	51444	Q8WVD3	245		1	E3	E3-RING	RING-type zinc finger
RNF139	11236	Q8WU17	664			E3	E3-RING	RING-type zinc finger
RNF14	9604	Q9UBS8	474		1	E3	E3-RING	RING-type zinc finger
RNF141	50862	Q8WVD5	230			E3	E3-RING	RING-type zinc finger
RNF144A	9781	P50876	292			E3	E3-RING	RING-type zinc finger
RNF144B	255488	P50876	303			E3	E3-RING	RING-type zinc finger
RNF145	153830	Q96MT1	663		1	E3	E3-RING	RING-type zinc finger
RNF146	81847	Q9NTX7	358		2	E3	E3-RING	RING-type zinc finger
RNF148	378925	Q8N7C7	305			E3	E3-RING	RING-type zinc finger
RNF150	57484	Q9ULK6	405	Deletion 1–92	3	E3	E3-RING	RING-type zinc finger
RNF151	146310	Q2KHN1	245			E3	E3-RING	RING-type zinc finger
RNF152	220441	Q8N8N0	203			E3	E3-RING	RING-type zinc finger
RNF166	115992	Q96A37	237		1	E3	E3-RING	RING-type zinc finger
RNF167	26001	Q9H6Y7	350			E3	E3-RING	RING-type zinc finger
RNF168	165918	Q8IYW5	571			E3	E3-RING	RING-type zinc finger
RNF170	81790	Q96K19	116		5	E3	E3-RING	RING-type zinc finger
RNF175	285533	Q8N4F7	328		1	E3	E3-RING	RING-type zinc finger
RNF180	285671	Q86T96	416		2	E3	E3-RING	RING-type zinc finger
RNF181	51255	Q9P0P0	153			E3	E3-RING	RING-type zinc finger
RNF182	221687	Q8N6D2	247			E3	E3-RING	RING-type zinc finger

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
RNF185	91445	Q96GF1	192		1	E3	E3-RING	RING-type zinc finger
RNF186	54546	Q9NXI6	227			E3	E3-RING	RING-type zinc finger
RNF19A	25897	Q9NV58	838		1	E3	E3-RING	RING- and IBR-type zinc finger
RNF19B	127544	Q6ZMZ0	732	Deletion 1–185	1	E3	E3-RING	RING- and IBR-type zinc finger
RNF2	6045	Q99496	336		1	E3	E3-RING	RING-type zinc finger
RNF204	148066	Q8WWF5	429			E3	E3-RING	RING-type zinc finger
RNF206	10293	Q9BWF2	469			E3	E3-RING	RING-type zinc finger
RNF208	727800	Q9H0X6	261	Deletion 1–81		E3	E3-RING	RING-type zinc finger
RNF209	339976	Q8N9V2	468	Deletion 1–276		E3	E3-RING	B-box domain & RING-type zinc
RNF212	285498	Q495C1	280		4	E3	E3-RING	RING-type zinc finger
RNF213	57674	Q63HN8	1063		3	E3	E3-RING	RING-type zinc finger
RNF214	257160	Q8ND24	703			E3	E3-RING	RING-type zinc finger
RNF24	11237	Q9Y225	148		1	E3	E3-RING	RING-type zinc finger
RNF25	64320	Q96BH1	459			E3	E3-RING	RING-type zinc finger
RNF26	79102	Q9BY78	433			E3	E3-RING	RING-type zinc finger
RNF32	140545	Q9H0A6	235		3	E3	E3-RING	RING-type zinc finger
RNF34	80196	Q969K3	372		1	E3	E3-RING	RING-type zinc finger
RNF38	152006	Q9H0F5	432		3	E3	E3-RING	RING-type zinc finger
RNF40	9810	O75150	1001		1,4	E3	E3-RING	RING-type zinc finger
RNF41	10193	Q9H4P4	317		1	E3	E3-RING	RING-type zinc finger
RNF43	54894	Q68DV7	783		1	E3	E3-RING	RING-type zinc finger
RNF44	22838	Q7L0R7	432		1	E3	E3-RING	RING-type zinc finger
RNF5	6048	Q99942	180			E3	E3-RING	RING-type zinc finger
RNF6	6049	Q9Y252	685		1	E3	E3-RING	RING-type zinc finger
RNF7	9616	Q9UBF6	113		1	E3	E3-RING	RING-type zinc finger
RNF8	9025	O76064	485		1	E3	E3-RING	RING-type zinc finger
RNFT1	51136	Q5MZT0	435	Deletion 1–37	3	E3	E3-RING	RING-type zinc finger
RNFT2	84900	F8W6L4	321			E3	E3-RING	RING-type zinc finger
RSPRY1	89970	Q96DX4	576		1	E3	E3-RING	RING-type zinc finger
RWDD3	25950	Q9Y3V2	185		3	Other	Sumoylation enhancer	
RWDD3	25950	Q9Y3V2	195		2	Other	Sumoylation enhancer	
SAE1	10055	Q9UBE0	346		1	E1	E1(SUMO)	
SH3RF1	57630	Q7Z6J0	888		1	E3	E3-RING	RING-type zinc finger
SH3RF2	153769	Q8TEC5	729		1	E3	E3-RING	RING-type zinc finger
SH3RF2	153769	Q8TEC5	220		2	E3	E3-RING	RING-type zinc finger
SHKBP1	92799	Q8TBC3	707		1	E3	E3 CRL adaptor	BTB (POZ) domain
SIAH1	6477	Q8IUQ4	282			E3	E3-RING	RING-type zinc finger
SIAH2	6478	O43255	324			E3	E3-RING	RING-type zinc finger
SKP1	6500	P63208	163		1	E3	E3 CRL component	
SKP2	6502	Q13309	424		2	E3	E3 CRL adaptor	F-box domain
SLX4	84464	Q8IY92	1834	Deletion 1–683		E3	E3 CRL adaptor	BTB (POZ) domain
SMURF2	64750	Q9HAU4	748			E3	E3-HECT	HECT
SOCS2	8835	O14508	198			E3	E3 CRL adaptor	SOCS box domain
SOCS3	9021	O14543	225			E3	E3 CRL adaptor	SOCS box domain
SOCS4	122809	Q8WXH5	440			E3	E3 CRL adaptor	SOCS box domain
SOCS5	9655	O75159	536			E3	E3 CRL adaptor	SOCS box domain
SOCS6	9306	O14544	535			E3	E3 CRL adaptor	SOCS box domain
SPOP	8405	O43791	374			E3	E3 CRL adaptor	BTB (POZ) domain.
SPOPL	339745	Q6IQ16	392			E3	E3 CRL adaptor	BTB (POZ) domain

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
SPSB1	80176	Q96BD6	273			E3	E3 CRL adaptor	SOCS box domain
SPSB2	84727	Q99619	263			E3	E3 CRL adaptor	SOCS box domain
SPSB3	90864	Q6PJ21	355			E3	E3 CRL adaptor	SOCS box domain
STAM2	10254	O75886	525			Other		
STAMBP	10617	O95630	424			DUB	DUB	JAMM/MPN
STAMBPL1	57559	Q96FJ0	436	Deletion 1–15		DUB	DUB	JAMM/MPN
SYVN1	84447	Q86TM6	617			E3	E3-RING	RING-type zinc finger
TAX1BP1	8887	Q86VP1	789			Other		
TBL1X	6907	O60907	577			E3	E3 CRL adaptor	F-box like domain
TCEB2	6923	Q15370	118			E3	E3 CRL adaptor	SOCS box domain
TLE3	7090	Q04726	772			E3	E3 CRL adaptor	WD repeat
TNFAIP1	7126	Q13829	316			E3	E3 CRL adaptor	BTB (POZ) domain
TNFAIP3	7128	P21580	790			E3/DUB	DUB/E3 RING	OTU, A20 type zinc finger
TRAF1	7185	Q13077	416			E3	E3-RING	
TRAF2	7186	Q12933	501			E3	E3-RING	RING-type zinc finger
TRAF3	7187	Q13114	568			E3	E3-RING	RING-type zinc finger
TRAF3IP2	10758	O43734	565			Other		
TRAF4	9618	Q9BUZ4	470			E3	E3-RING	RING-type zinc finger
TRAF5	7188	O00463	557			E3	E3-RING	RING-type zinc finger
TRAF6	7189	Q9Y4K3	522			E3	E3-RING	RING-type zinc finger
TRIM10	10107	Q9UDY6	395		Beta	E3	E3-RING	B-box and RING-type zinc fingers
TRIM11	81559	Q96F44	468	Deletion 1–266		E3	E3-RING	B-box and RING-type zinc fingers
TRIM13	10206	O60858	407			E3	E3-RING	B-box and RING-type zinc fingers
TRIM15	89870	Q9C019	465			E3	E3-RING	B-box and RING-type zinc fingers
TRIM17	51127	Q9Y577	477			E3	E3-RING	B-box and RING-type zinc fingers
TRIM2	23321	Q9C040	744			E3	E3-RING	B-box and RING-type zinc fingers
TRIM21	6737	P19474	475			E3	E3-RING	B-box and RING-type zinc fingers
TRIM22	10346	Q8IYM9	498			E3	E3-RING	B-box and RING-type zinc fingers
TRIM23	373	P36406	574		Alpha	E3	E3-RING	B-box and RING-type zinc fingers
TRIM24	8805	O15164	1050			E3	E3-RING	B-box and RING-type zinc fingers
TRIM25	7706	Q14258	630			E3	E3-RING	B-box and RING-type zinc fingers
TRIM26	7726	Q12899	539			E3	E3-RING	B-box and RING-type zinc fingers
TRIM27	5987	P14373	513		Alpha	E3	E3-RING	B-box and RING-type zinc fingers
TRIM28	10155	Q13263	835		1	E3	E3-RING (SUMO/UB)	B-box and RING-type zinc fingers
TRIM3	10612	O75382	744			E3	E3-RING	B-box and RING-type zinc fingers
TRIM32	22954	Q13049	653			E3	E3-RING	B-box and RING-type zinc fingers
TRIM35	23087	Q9UPQ4	493	Deletion 207–493	2	E3	E3-RING	B-box and RING-type zinc fingers

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
TRIM36	55521	Q9NQ86	728		1	E3	E3-RING	B-box and RING-type zinc fingers
TRIM37	4591	O94972	964		1	E3	E3-RING	B-box and RING-type zinc fingers
TRIM38	10475	O00635	465			E3	E3-RING	B-box and RING-type zinc fingers
TRIM39	56658	Q9HCM9	488		2	E3	E3-RING	B-box and RING-type zinc fingers
TRIM4	89122	Q9C037	294		Gamma	E3	E3-RING	B-box and RING-type zinc fingers
TRIM41	90933	Q8WV44	210		3	E3	E3-RING	B-box and RING-type zinc fingers
TRIM41	90933	Q8WV44	630	Deletion 1–85	1	E3	E3-RING	B-box and RING-type zinc fingers
TRIM42	287015	Q8IWZ5	723		1	E3	E3-RING	B-box and RING-type zinc fingers
TRIM43	129868	Q96BQ3	446			E3	E3-RING	B-box and RING-type zinc fingers
TRIM47	91107	Q96LD4	400		2	E3	E3-RING	B-box and RING-type zinc fingers
TRIM48	79097	Q8IWZ4	208			E3	E3-RING	B-box and RING-type zinc fingers
TRIM49	57093	P0CI25	452			E3	E3-RING	B-box and RING-type zinc fingers
TRIM5	85363	Q9C035	347		Gamma	E3	E3-RING	B-box and RING-type zinc fingers
TRIM50	135892	Q86XT4	487		Alpha	E3	E3-RING	B-box and RING-type zinc fingers
TRIM51	84767	Q9BSJ1	309	Deletion 1–16	2	E3	E3-RING	B-box and RING-type zinc fingers
TRIM52	84851	Q96A61	297			E3	E3-RING	B-box and RING-type zinc fingers
TRIM54	57159	Q9BYV2	358		1	E3	E3-RING	B-box and RING-type zinc fingers
TRIM55	84675	Q9BYV6	452		2	E3	E3-RING	B-box and RING-type zinc fingers
TRIM56	81844	Q9BRZ2	755		1	E3	E3-RING	B-box and RING-type zinc fingers
TRIM57	286827	Q8IWR1	403			E3	E3-RING	B-box and RING-type zinc fingers
TRIM60	166655	Q495X7	471			E3	E3-RING	B-box and RING-type zinc fingers
TRIM61	391712	Q5EBN2	209			E3	E3-RING	B-box and RING-type zinc fingers
TRIM62	55223	Q9BVG3	475		1	E3	E3-RING	B-box and RING-type zinc fingers
TRIM69	140691	Q86WT6	341		2	E3	E3-RING	B-box and RING-type zinc fingers
TRIM72	493829	Q6ZMU5	269		2	E3	E3-RING	B-box and RING-type zinc fingers
TRIM74	378108	Q86UV6	249		2	E3	E3-RING	B-box and RING-type zinc fingers
TRIM9	114088	Q9C026	550		5	E3	E3-RING	B-box and RING-type zinc fingers

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
UBA3	9039	Q8TBC4	463		1	E1	E1 (for NEDD8)	
UBA5	79876	Q9GZZ9	404		1	E1	E1-like	
UBAC1	10422	Q9BSL1	405			E3	E3-RING	UB-like, UBA domains
UBB	7314	P0CG47	229	Stop at aa 78		Other	Ub	Ub-like domain
UBC9	7329	P63279	158			E2	E2 (SUMO)	UBC domain
UBD	10537	O15205	165			Other	Ub	Ub-like domain
UBE2A	7319	P49459	152		1	E2	E2	UBC domain
UBE2B	7320	P63146	152			E2	E2	UBC domain
UBE2D1	7321	P51668	147			E2	E2	UBC domain
UBE2D3	7323	P61077	146		1	E2	E2	UBC domain
UBE2D4	51619	Q9Y2X8	147			E2	E2	UBC domain
UBE2E	7325	Q96LR5	201			E2	E2	UBC domain
UBE2E3	10447	Q969T4	207			E2	E2	UBC domain
UBE2F	140739	Q969M7	185		1	E2	E2	UBC domain
UBE2G1		P62253	170			E2	E2	UBC domain
UBE2G2	7327	P60604	165		1	E2	E2	UBC domain
UBE2H	7328	P62256	183		1	E2	E2	UBC domain
UBE2J1	51465	Q9Y385	318			E2	E2	UBC domain
UBE2K	3093	P61086	200		1	E2	E2	UBC domain
UBE2L6	9246	O14933	153		1	E2	E2	UBC domain
UBE2M	9040	P61081	183			E2	E2 (nedd8)	UBC domain
UBE2O	63893	Q9C0C9	1292	Deletion 1–551		E2	E2	UBC domain
UBE2Q2	92912	Q8WVN8	375		1	E2	E2	UBC domain
UBE2R2	54926	Q712K3	238			E2	E2	UBC domain
UBE2S	27338	Q16763	222			E2	E2	UBC domain
UBE2T	29089	Q9NPD8	197			E2	E2	UBC domain
UBE2U	148581	Q5VVX9	226		2	E2	E2	UBC domain
UBE2V1		Q13404	147		3	E2	E2	UBC domain
UBE2V2	7336	Q13404	145			E2	E2	UBC domain
UBE2W		Q96FI0	151			E2	E2	UBC domain
UBE2Z	65264	Q9H832	246		2	E2	E2	UBC domain
UBE3B	8991 0	Q7Z3V4_3	244		3	E3	E3-HECT	HECT
UBE3C	9690	Q15386	1083		3	E3	E3-HECT	HECT
UBL4A	8266	P11441	157			Other		Ub-like domain
UBL7	84993	Q96S82	380			Other		Ub-like, UBA domains
UBOX5	22888	O94941	541		1	E3	E3-U box	U box domain
UBR2	23304	Q8IWV8	439		2	E3	E3-RING	RING-type zinc finger
UBTD2	92181	B2R886	234			Other	UB-like	Ub domain
UCHL1	7345	P09936	223			DUB	DUB	UCH
UCHL3	7347	P15374	230			DUB	DUB	UCH
UCHL5	51377	Q9Y5K5	328		3	DUB	DUB	UCH
UCHL5	51377	Q9Y5K5	326		4	DUB	DUB	UCH
UEVLD	55293	Q8IX04	215	Deletion 1–38	5	E2	E2	UBC domain
UEVLD	55293	Q8IX04	357		3	E2	E2	UBC domain
UFC1	51506	Q9Y3C8	167			E2	E2 like	UBC domain
UFD1L	7353	Q92890	307		Short	E3	E3-RING	RING-type zinc finger
UFM1	51569	P61960	85		1	Other	UB-like	Ub fold
UNKL	64718	Q9H9P5	277		5	E3	E3-RING	RING-type zinc finger
USP1	7398	O94782	785			DUB	DUB	USP

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
USP10	9100	Q14694	798		1	DUB	DUB	USP
USP11	8237	P51784	963			DUB	DUB	USP
USP12	219333	O75317	370			DUB	DUB	USP
USP14	9097	P54578	494		1	DUB	DUB	USP
USP15	9958	Q9Y4E8	981			DUB	DUB	USP
USP16	10600	Q9Y5T5	823		2	DUB	DUB	USP
USP18	11274	Q9UMW8	372		1	DUB	DUB	USP
USP19	10869	O94966	1318		1	DUB	DUB	USP
USP20	10868	Q9Y2K6	914	Deletion 310–910		DUB	DUB	USP
USP21	27005	Q9UK80	565	Deletion 1–184	1	DUB	DUB	USP
USP22	23326	Q9UPT9	513		2	DUB	DUB	USP
USP25	29761	Q96B65	450			DUB	DUB	USP
USP25	29761	Q9UHP3	1055		A	DUB	DUB	USP
USP26	83844	Q9BXU7	913			DUB	DUB	USP
USP28	57646	Q96RU2	583		3	DUB	DUB	USP
USP29	57663	Q9HBJ7	922	Deletion 913–922		DUB	DUB	USP
USP3	9960	Q9Y6I4	520		1	DUB	DUB	USP
USP30	84749	Q70CQ3	517	Deletion 1–9		DUB	DUB	USP
USP32	84669	Q8NFA0	390		2	DUB	DUB	USP
USP33	23032	Q8TEY7	828		3	DUB	DUB	USP
USP35	57558	A2RRA6	604			DUB	DUB	USP
USP36	57602	Q9P275	1121			DUB	DUB	USP
USP37	57695	Q86T82	979	Deletion 288–979	1	DUB	DUB	USP
USP38	84640	Q8NB14	1042		1	DUB	DUB	USP
USP39	10713	Q53GS9	565		1	DUB	DUB	USP
USP4	7375	Q13107	963		1	DUB	DUB	USP
USP42	84132	Q9H9J4	1316		2	DUB	DUB	USP
USP44	84101	Q9H0E7	712			DUB	DUB	USP
USP46	64854	P62068	366		1	DUB	DUB	USP
USP47	55031	Q96K76	157		3	DUB	DUB	USP
USP48	84196	Q86UV5	1035	Deletion 435–1035		DUB	DUB	USP
USP49	25862	Q70CQ1	640		2	DUB	DUB	USP
USP5	8078	P45974	835		Short	DUB	DUB	USP
USP53	54532	Q70EK8	1073	Deletion 1–885		DUB	DUB	USP
USP54	159195	Q70EL1	725		6	DUB	DUB	USP
USP7	7874	Q93009	1102		1	DUB	DUB	USP
USP8	9101	P40818	1118		1	DUB	DUB	USP
USPL1	10208	Q5W0Q7	1092		1	DUB	DUB	USP
VHL	7428	P40337	172		2	E3	E3-CRL component	VHL domain
WDR82	80335	Q6UXN9	313			E3	E3 CRL adaptor	WD repeats
WDTC1	23038	Q8N5D0	667		1,4	Other		
WWP1	11059	Q9H0M0	922		1	E3	E3-HECT	HECT
WWP2	11060	O00308	870			E3	E3-HECT	HECT
XIAP	331	P98170	497			E3	E3-RING	RING-type zinc finger
ZBTB1	22890	Q9Y2K1	713		1	E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB10	65986	Q96DT7	698		3	E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB12	221527	Q9Y330	459			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB14	7541	O43829	449			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB16	7704	Q05516	673			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB20	26137	Q9HC78	668		2	E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB21	49854	Q9ULJ3	1066			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB22	9278	O15209	634	Deletion 374–634		E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB24	9841	O43167	333		2	E3	E3 CRL adaptor	BTB (POZ) domain

Table 1. (Continued).

SYMB	GeneID	Uniprot	Protein length	UPS clone	Isoform	Protein class	Family	InterPro (domain/motif)
ZBTB25	7597	P24278	435			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB26	57684	Q9HCK0	441			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB32	27033	Q9Y2Y4	487	Deletion 294–487		E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB37	84614	Q5TC79	361		2	E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB38	253461	Q9H6F0	326			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB39	9880	O15060	712			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB42	100128927	B2RXF5	422			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB43	23099	O43298	467			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB45	84878	Q96K62	511			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB48	3104	P10074	688			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB49	166793	Q6ZSB9	765		1	E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB49	166793	Q32MK9	248		2	E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB5	9925	O15062	677			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB6	10773	Q15916	424			E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB8A	653121	Q96BR9	441		1	E3	E3 CRL adaptor	BTB (POZ) domain
ZBTB9	221504	Q96C00	473			E3	E3 CRL adaptor	BTB (POZ) domain
ZFAND4	93550	Q86XD8	727	Deletion 1–75		Other		Ub-like domain
ZFAND5	7763	O76080	213			Other	Proteasome associated protein	A20-type zinc finger
ZIBRA (RNF31)	55072	Q96EP0	1072	Deletion 1–484		E3	E3-RING (LUBAC complex)	RING- and IBR-type zinc fingers
ZNF131	7690	P52739	589	Deletion 1–80	2	E3	E3 CRL adaptor	BTB (POZ) domain
ZNF219	51222	Q9P2Y4	722			E3	E3 CRL adaptor	BTB (POZ) domain
ZNF238	10472	Q99592	522		1	E3	E3 CRL adaptor	BTB (POZ) domain
ZNF598	90850	Q86UK7	904			E3	E3-RING	RING-type zinc finger
ZNF645	158506	Q8N7E2	425			E3	E3-RING	RING-type zinc finger
ZNRF1	84937	Q8ND25	227		1	E3	E3-RING	RING-type zinc finger
ZSWIM2	151112	Q8NEG5	633			E3	E3-RING	RING-type zinc finger

of the confidence interval calculated from the RRS (see Materials and methods and Table S4). By using these criteria, 14 binary interactions of the 81 pairs (15%) and 33 of 65 pairs (53%) were validated for E6 and E7, respectively (Table 2). TRAF3IP2 and UBB were selected with both viral proteins, which is most likely due to their high background signals as described above. Importantly, most of the UPS factors identified are specific to either E6 or E7, with 12 UPS factors targeted specifically by E6 and 31 UPS factors targeted by E7. Among these UPS factors, we recovered the already described UPS partners of E6, namely E6AP/UBE3A [9], USP15 [17], and LNX3/PDZRN3 [31].

The analysis of the domain composition of the UPS targets highlighted distinct patterns of UPS interaction. HPV16 E7 showed a strong preference to some BTB domain-carrying proteins, which are considered as potential SRFs of Cullin 3-based CRL, whereas HPV16 E6 binds to only one HECT Ub ligase (E6AP) and some RING-type Ub ligases, but no SRF of CRL. Interestingly, each viral protein binds specific DUBs, including USP15 for E6 and MPND, USP26, USP29, and USP33 for E7.

Validation of protein pairs by co-immunoprecipitation

We tested all the potential interacting protein pairs detected by the GPCA approach by using an orthogonal co-immunoprecipitation (IP) experiment. In this case, the Gaussia-tagged UPS factors were co-expressed with FLAG-tagged viral bait (E6, E7) in HEK293T cells. The protein complexes were purified using ANTI-FLAG beads and analyzed by gel electrophoresis and western blot using anti-Gaussia luciferase antibody. This low-throughput interactomic approach allowed us to consolidate some of the

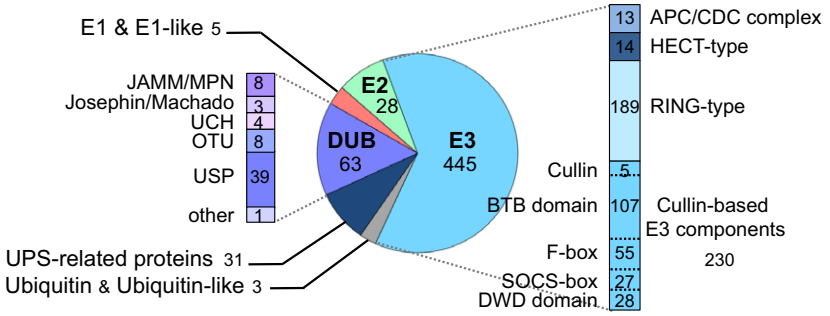


Fig. 1. Composition of the UPS library. The UPS library contains 575 unique entries of cDNAs encoding E1 or E1-like enzymes, E2-conjugating enzymes, E3-ligating enzymes, Ub and Ub-like proteins, the subunits of APC/C complex, and some UPS-related proteins. The E3 Ub ligases are divided into three families: HECT-type Ub ligases, RING-type Ub ligases, and Cullin-based RING Ub ligases (CRL). The DUBs contain DUBs with JAMM/MPN domain, MJD, UCH domain, OTU domain, USP domain.

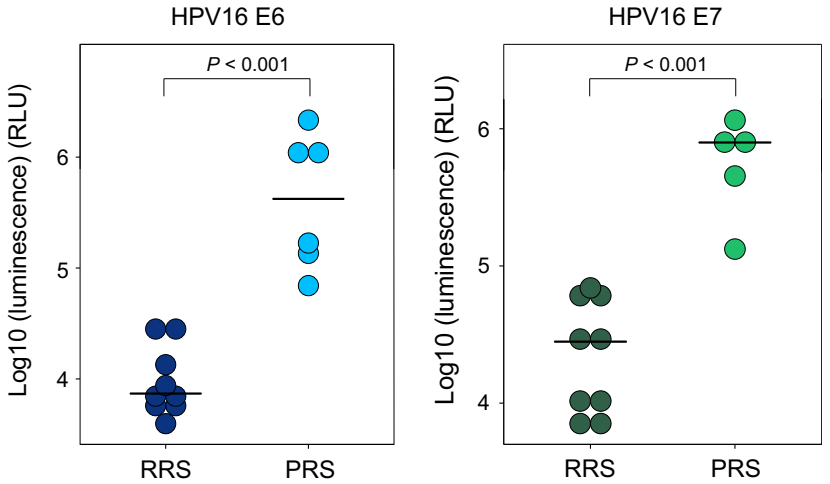


Fig. 2. GPCA signal of HPV16 E6 and E7 proteins tested against the proteins of the PRS and the RRS. HEK293T cells were cotransfected with plasmids encoding Gluc2-16E6 or Gluc2-16E7 and the Gluc1-RRS proteins or Gluc1-PRS. The PRS was composed of UBE3A/E6AP, hADA3, MAGI-1, SCRIB, DLG3, IRF3 for E6 and EEF1A1, PKM2, PTPN14, RB1, RBL1 for E7. Twenty-four hours after transfection, the cells were lysed and the luciferase activity was measured. The log10 of the luminescence value for each protein pair is represented as a dot, and the median as a black line. Statistical analysis was performed with two-tailed *t*-test with unequal population variances ($P < 0.001$).

protein–protein interactions (PPI) identified by the GPCA. We selected UPS factors with an IP ratio > 3 (i.e., Flag-E6/Ctrl and Flag-E7/Ctrl ratio in Fig. 4, Table S9). We observed a strong association of E6 with the Ub ligases E6AP, LNX3, and LNX4 (IP ratio > 5) and a more modest one with MGRN1 and USP15 (Fig. 4A). For E7, we observed strong interactions with the RING Ub ligases NEURL1 and RNF135, the BTB domain protein TNFAIP1, the WD40-containing protein DCAF15, and the DUBs USP26, USP29, USP33 (IP ratio > 5). Weaker interactions were instead observed between E7 and the RING Ub ligases (TRAF2, TRAF5, TRIM32, TRIM72), the BTB domain proteins (BTBD15, KCTD13, ZBTB9, ZBTB32, SHKBP1, ZBTB48), and the DUB MPND (Fig. 4B).

Identification of the domain/motif interactions implicated in UPS targeting by E6 and E7

The HPV16 E6 contains two zinc-binding domains and a PDZ-binding motif (PBM) at its COOH terminus. The structure of HPV16 E6 in complex with a peptide from UBE3A/E6AP revealed the presence of a charged hydrophobic pocket in E6 hosting the LxxLL motif [10]. We verified the expression of E6 mutants (E6 L50E containing a mutation within the LxxLL-binding pocket and E6ΔPBM containing a deletion of

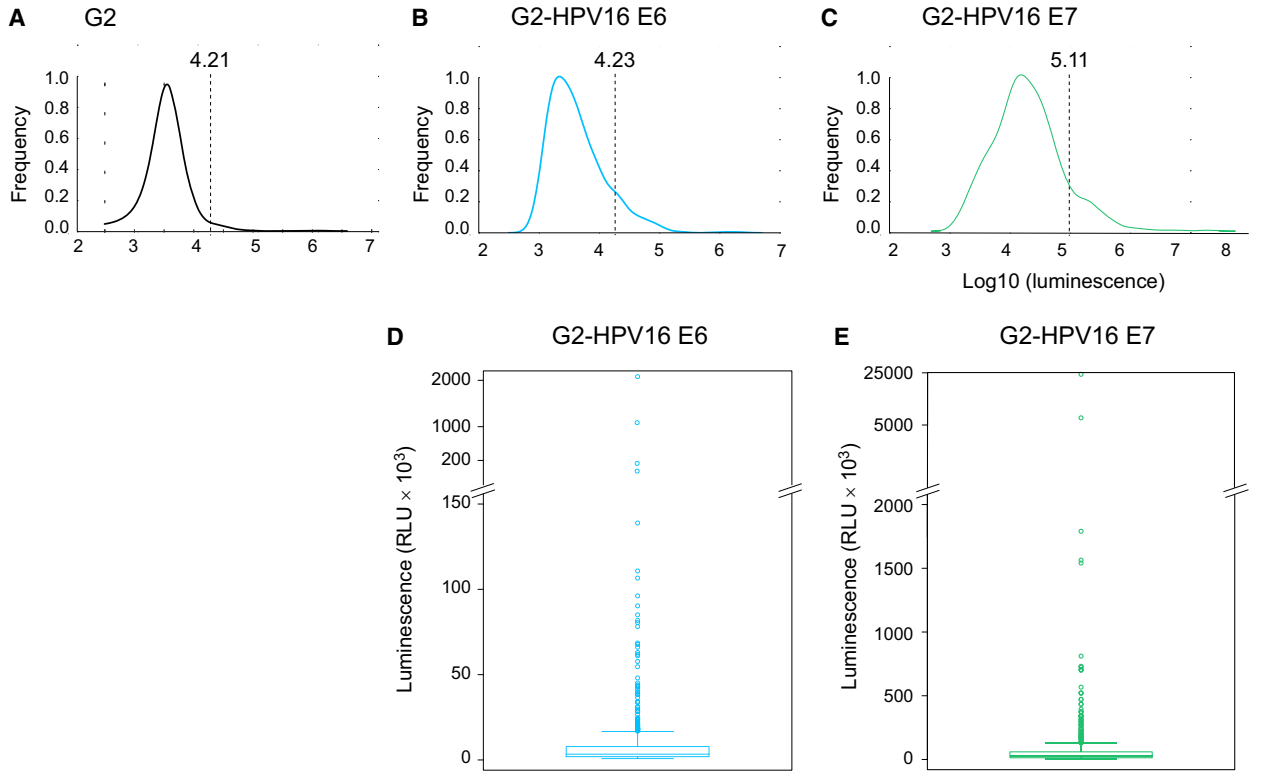


Fig. 3. Screening of the UPS library with Gluc2, Gluc2-16E6, and Gluc2-16E7 by using GPCA. HEK293T cells were cotransfected with plasmids pSPICA-N1-UPS encoding Gluc1-UPS and the pSPICA-N2, pSPICA-N2-16E6, pSPICA-N2-16E7 encoding Gluc2, Gluc2-16E6, Gluc2-16E7. Twenty-four hours after transfection, the cells were lysed and the luciferase activity was measured. The log10 of the luminescence values of Gluc1-UPS factors tested against Gluc2 (A), Gluc2-16E6 (B), or Gluc2-16E7 (C) are represented according to the frequency in the KDE plot representation. On the curve, the dashed line indicates the position of the 'shoulder' (cutoff value). Whisker plot representation of the luminescence values of Gluc1-UPS factors tested against Gluc2-16E6 (D) and Gluc2-16E7 (E). The outlier luminescence values (circles) fall above the threshold $T = Q3 + 1.5xIQR$.

the last four amino acids inhibiting the binding to PDZ domain-containing proteins) by western blot (Fig. 5A). We tested E6 L50E and E6ΔPBM mutants against all UPS factors validated by co-IP using the GPCA (Fig. 5B, Table S5). The E6 L50E was unable to interact with E6AP as expected, but it was still able to interact with LNX3, LNX4, and MGRN1. For USP15, the GPCA signals obtained with wild-type or mutant E6 proteins were too weak to allow us to conclude on the impact of the mutation. E6ΔPBM was unable to bind MAGI-1 but it could still interact with LNX4, MGRN1, and E6AP. The deletion of the PBM also inhibited the binding of E6 to LNX3, which is consistent with previous reports that LNX3/E6 interaction requires the PBM of E6 [31]. Surprisingly, the E6ΔPBM was still able to interact with our construct of LNX4 that corresponds to isoform 2.

The E7 protein contains three regions: the unfolded CR1 and CR2 regions (the latter one containing the LxCxE motif required for RB1 interaction [32]) and the C-terminal region containing a zinc-binding domain [33]. We analyzed by western blot the expression of the three constructs of HPV16 E7: the CR1 + CR2 region (1–36 amino acids), the C-terminal region (37–98 amino acids), and the HPV16 E7 C24G/E26G construct mutated within the LxCxE motif (Fig. 6A). We noticed that at the protein level the Gluc2-CR1 + CR2 construct was lower than the Gluc2-C-terminal construct of E7. As the CR1 and CR2 regions are intrinsically disordered [34], the construct Gluc2-CR1 + CR2 is maybe less expressed or more prone to aggregation/degradation.

The E7 constructs were tested against all selected UPS factors by GPCA and co-IP (Fig. 6B; Table S6). As expected, we confirmed that the interaction between HPV16 E7 and RB1 or RBL1 requires the integrity of the LxCxE motif within the CR2 region [29]. Indeed, we observed that the C24G/E26G mutations inhibited completely the binding to RB1 and, to some extent, the binding to RBL1. This observation

Table 2. UPS factors interacting with HPV16 E6 and E7 proteins in GPCA or co-IP experiments. UPS factors that give outliers luminescence values when tested against the Gluc2 fragment alone are indicated. UPS factors selected by GPCA (squares) or co-IP (asterisks) for binding to E6 or E7 proteins are indicated.

	HPV16 E6		HPV16 E7		G2			
SYMB	GPCA	Co-IP	GPCA	Co-IP	GPCA	UPS class	Family	Domain/motif
UBB	■		■		■	UB		Ub
LNX3	■	✻				E3	RING	RING-type zinc finger PDZ domain
LNX4	■	✻				E3	RING	RING-type zinc finger PDZ domain
MGRN1	■	✻				E3	RING	RING-type zinc finger
RNF25	■					E3	RING	RING-type zinc finger
RNF40	■					E3	RING	RING-type zinc finger
TRAF5	■		■	✻		E3	RING	RING-type zinc finger
TRAF6	■					E3	RING	RING-type zinc finger
E6AP	■	✻				E3	HECT	HECT
ITCH	■					E3	HECT	HECT
AIMP2	■					Other		
TAX1BP1	■					Other		
TRAF3IP2	■		■		■	Other		TRAF-binding domain
USP15	■	✻				DUB		USP
UBE2A			■		■	E2		UBC
NEURL1			■	✻		E3	RING	RING-type zinc finger
RNF135			■	✻		E3	RING	RING-type zinc finger
SH3RF1			■			E3	RING	RING-type zinc finger
TRAF2			■	✻		E3	RING	RING-type zinc finger
TRAF3			■		■	E3	RING	RING-type zinc finger
TRAF4			■		■	E3	RING	RING-type zinc finger
TRIM9			■			E3	RING	RING-type zinc finger
TRIM22			■		■	E3	RING	RING-type zinc finger
TRIM32			■	✻		E3	RING	RING-type zinc finger
TRIM54			■			E3	RING	RING-type zinc finger
TRIM72			■	✻		E3	RING	RING-type zinc finger
ZNF219			■			E3	RING	RING-type zinc finger
ZNF598			■			E3	RING	RING-type zinc finger
BTBD15			■	✻		E3	Cul3-CRL	BTB
KCTD13			■	✻		E3	Cul3-CRL	BTB
NACC1			■			E3	Cul3-CRL	BTB
SHKBP1			■	✻		E3	Cul3-CRL	BTB
TNFAIP1			■	✻		E3	Cul3-CRL	BTB
ZBTB9			■	✻		E3	Cul3-CRL	BTB
ZBTB20			■			E3	Cul3-CRL	BTB
ZBTB32			■	✻		E3	Cul3-CRL	BTB
ZBTB42			■			E3	Cul3-CRL	BTB
ZBTB43			■			E3	Cul3-CRL	BTB
ZBTB48			■	✻		E3	Cul3-CRL	BTB
DCAF15			■	✻		E3	Cu 4-CRL	WD repeats
MPND			■	✻	■	DUB		JAMM
USP26			■	✻		DUB		USP
USP29			■	✻		DUB		USP
USP33			■	✻		DUB		USP

suggests that there are potential additional binding sites for RBL1 in HPV16 E7. In contrast, the mutation of the LxCxE motif did not affect the binding to the UPS factors. The GPCA signals obtained with RB1 or RBL1 and with the CR1 + CR2 construct were not as high as the ones obtained with the full-length E7 protein, possibly due to the lower expression and/or stability of the CR1 + CR2 construct. The association with PTPN14 was mediated by the C-terminal region of E7 as previously reported [30]. All UPS factors seem to interact with the C-terminal region of E7, albeit at different extents. For some UPS factors (such

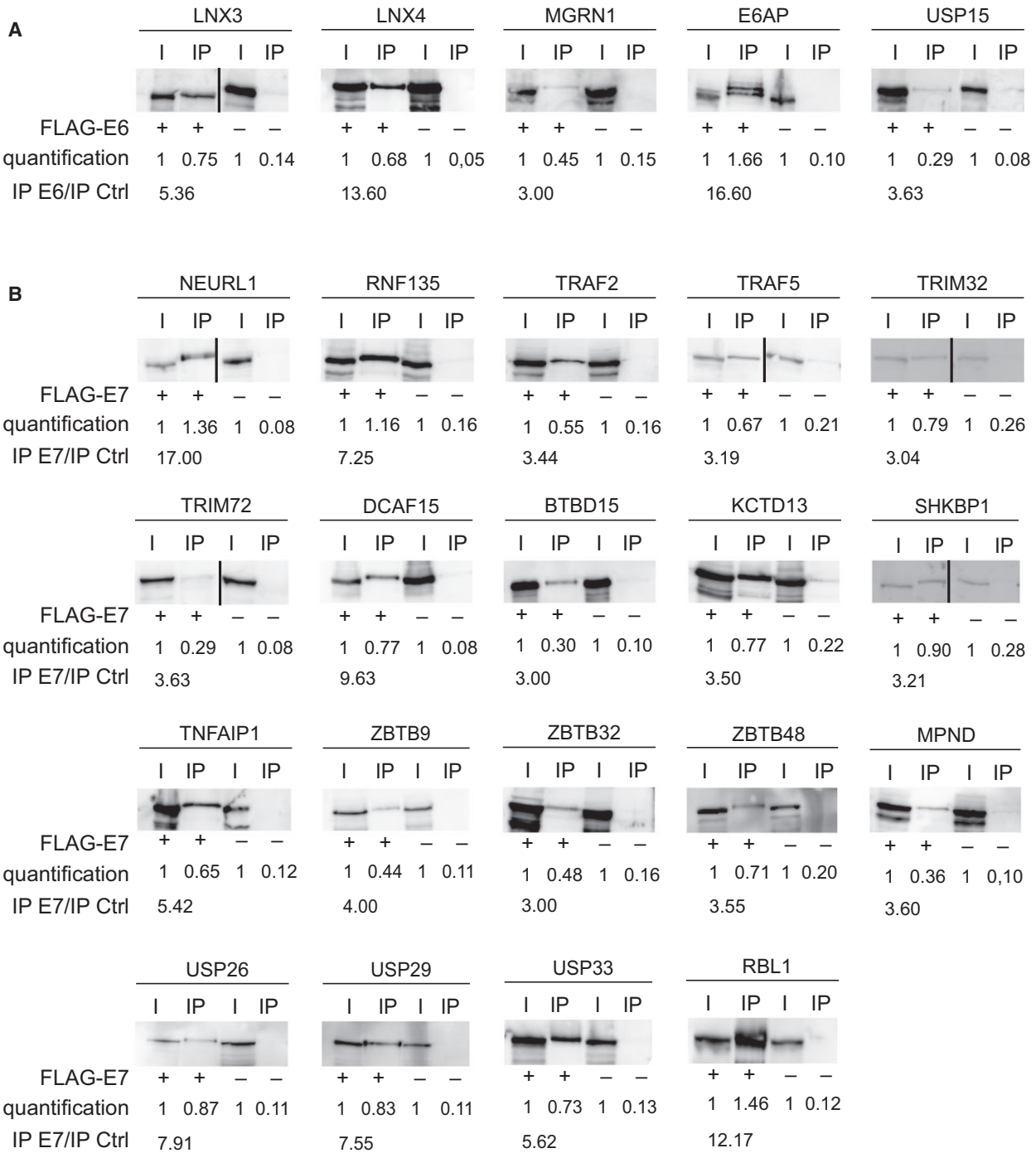


Fig. 4. Co-IP of UPS factors with HPV16 E6 and E7 proteins. HEK293T cells were cotransfected with pCineo-3xFLAG-16E6 (A), pCineo-3xFLAG-16E7 (B), or empty vector (–) and pSPICA-Gaussia-UPS encoding UPS proteins fused to Gaussia protein. RBL1 was also used as a positive control. Lysates were incubated with anti-FLAG magnetic beads. Total proteins (20% input, IN), and IP proteins were analyzed by western blot by using anti-Gluc antibodies. Protein band intensities were quantified using IMAGE J software (open source, Wayne Rasband, NIH) (Table S9). Band intensities in the IP lanes were normalized according to intensities in the input lanes. For each Gluc-UPS protein, we calculated the IP (E6 or E7)/IP (Ctrl) ratio, where IP (E6 or E7) corresponds to the normalized intensity in the E6 or E7 IP lane and IP (Ctrl) corresponds to the normalized intensity in the control IP lane. UPS factors with an IP ratio < 3 are not represented (data not shown). Note: For LNX3, NEURL1, TRAF5, TRIM32, TRIM72, SHKBP1 blots, samples were run on the same gel but intervening lanes were removed as indicated in the figure.

as TNFAIP1 and USP33), we observed a weak binding for the isolated C-terminal region, suggesting that these interactions required the entire folded E7 protein.

Binding of UPS factors to E6 or E7 from different HPV types

We analyzed the binding profiles of validated UPS factors with E6 and E7 proteins from high-risk mucosal HPV (HPV16, HPV18, HPV33), low-risk mucosal HPV (HPV6), and cutaneous HPV (HPV8, HPV38) types by GPCA. The expression of E6 proteins (Fig. 7A) and E7 proteins (Fig. 8A) from these different HPV types was verified by western blot. We observed some slower-migrating bands for E6 proteins from HPV16, HPV18, and HPV33 types, which likely correspond to E6* produced from the splicing of the intronic sequence within the E6 gene. To estimate the background GPCA signal produced by the different E6 and E7 proteins, the RRS was also assessed against each protein (Figs 7B and 8B; Tables S7 and S8). We observed that the E7 proteins from HPV6, HPV8, and HPV38 gave higher NLR values with the RRS than the E7 proteins from HPV16, HPV18, and HPV33 types. We thus considered an interaction to be positive if the NLR of the protein pair was above the upper limit of the confidence interval calculated from the RRS for each E6 or E7 protein of the different HPV types.

We confirmed that E6AP is targeted by E6 from high-risk and low-risk mucosal HPV, whereas MAGI-1 and SMAD3 are specific targets of E6 from high-risk mucosal HPV and cutaneous HPV, respectively, as previously published [29] (Fig. 7C, Table S7). Considering the UPS factors, USP15 and MGRN1 interacted

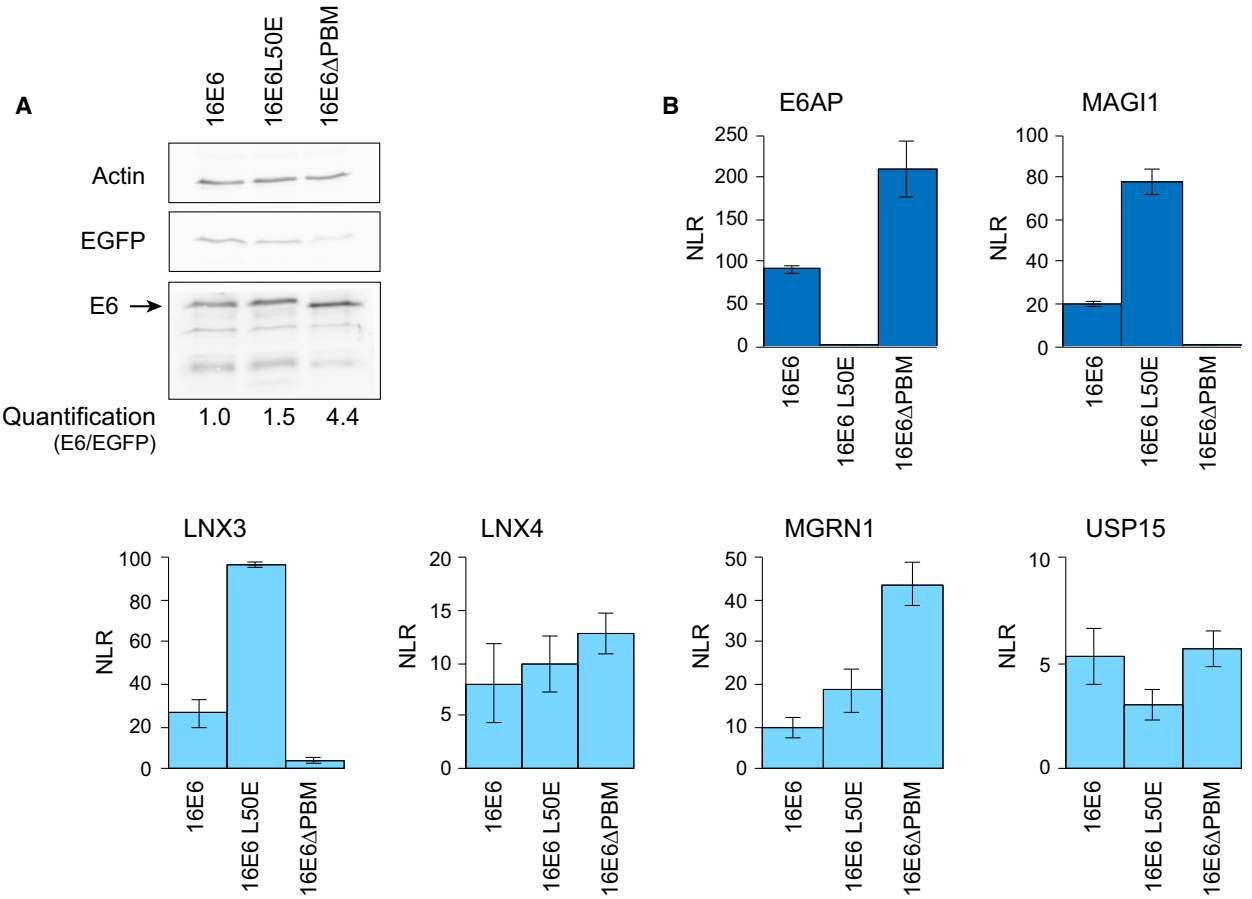


Fig. 5. Binding of HPV16 E6 mutants to UPS factors in GPCA. (A) HEK 293T cells (4×10^5) were transfected with pSPICA-N2-16E6, pSPICA-N2-E6 L50E, and pSPICA-N2-E6ΔPDZ plasmids and pEGFP-C3 plasmid as internal control. The cleared cellular lysates were analyzed by western blot using anti-Gluc, anti-actin, and anti-GFP antibodies. (B) HEK 293T (28×10^3) cells were transfected with pSPICA-N2-E6 and pSPICA-N1-MAGI-1, pSPICA-N1-SMAD3, and pSPICA-N1-UPS. After transfection, the Gaussia luciferase activity was measured. Representative NLR values for each protein pair are shown.

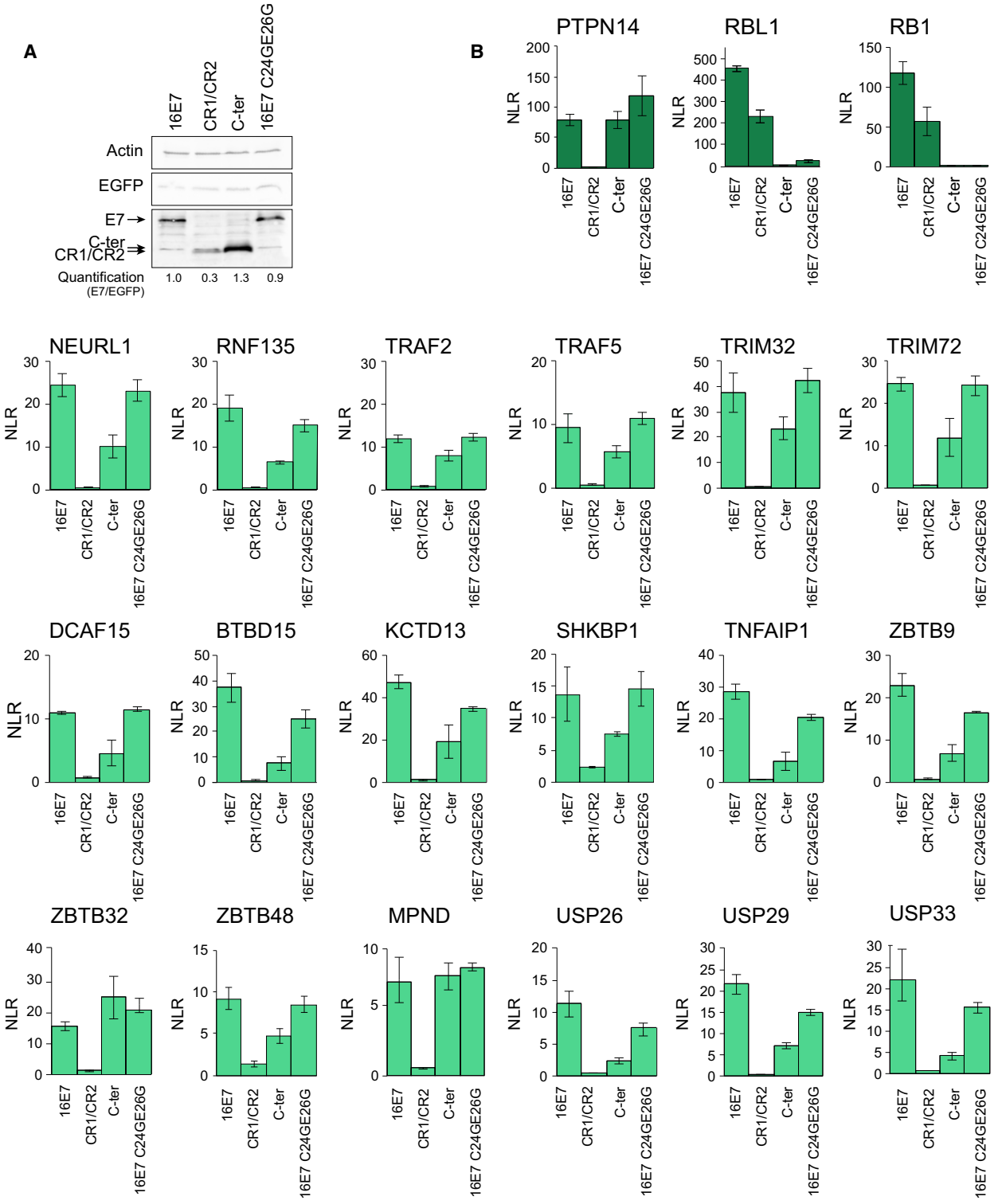


Fig. 6. Binding of HPV16 E7 domains and HPV16 E7 mutant to UPS factors in GPCA. (A) HEK 293T cells (4×10^5) were transfected with pSPICA-N2-16E7, pSPICA-N2-16E7 (C24G/E26G) and pSPICA-N2-E7(CR1 + CR2), pSPICA-N2-E7-C-terminal plasmids, and pEGFP-C3 plasmid as internal control. The cleared cellular lysates were analyzed by western blot using anti-Gluc, anti-actin, and anti-GFP antibodies. (B) HEK293T cells (28×10^3) were cotransfected with pSPICA-N1-PTPN14, pSPICA-N1-RB1, pSPICA-N1-RBL1, pSPICA-N1-UPS plasmids, and pSPICA-N2-E7. After transfection, the Gaussia luciferase activity was measured. Representative NLR values for each protein pair are shown.

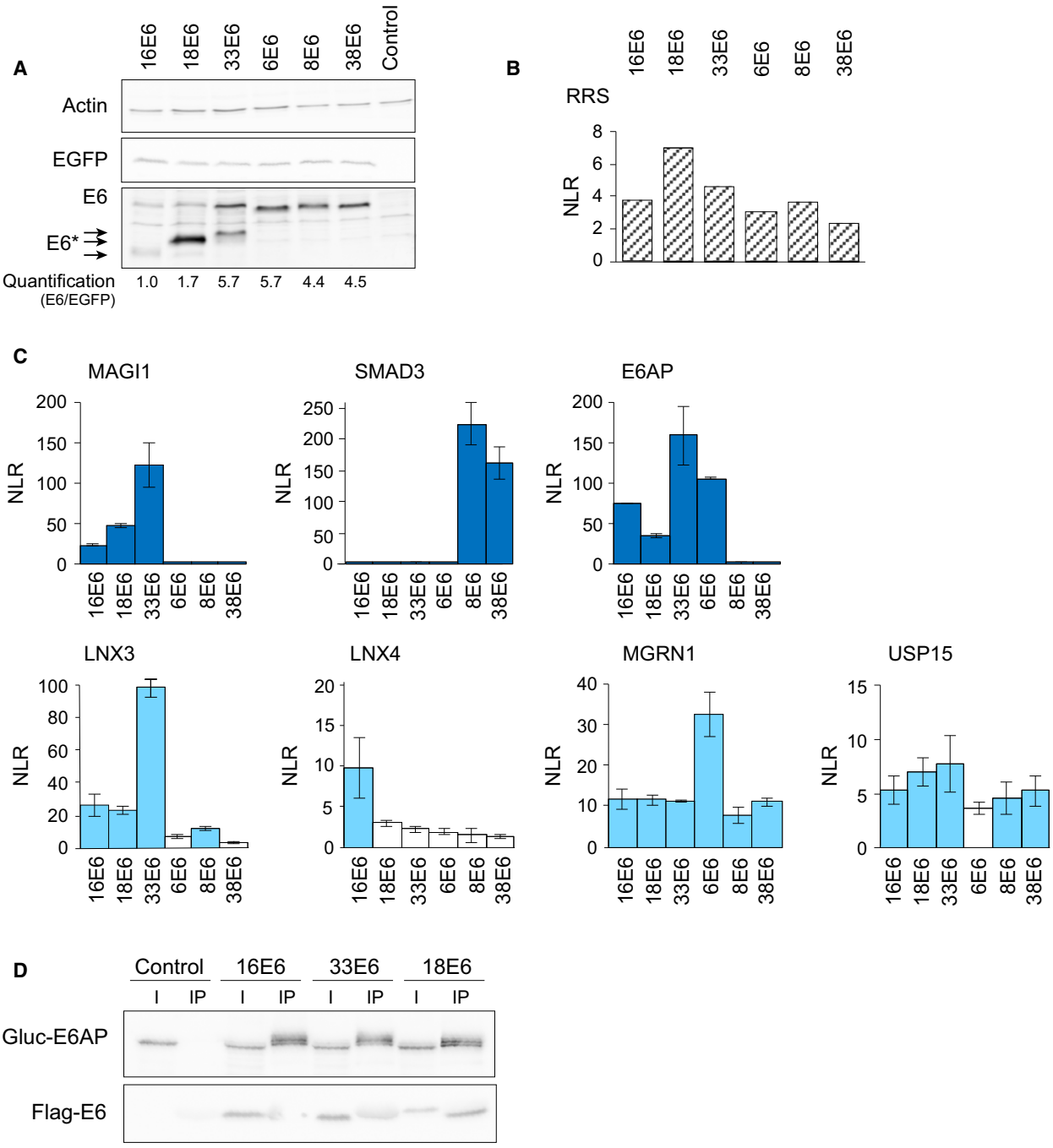
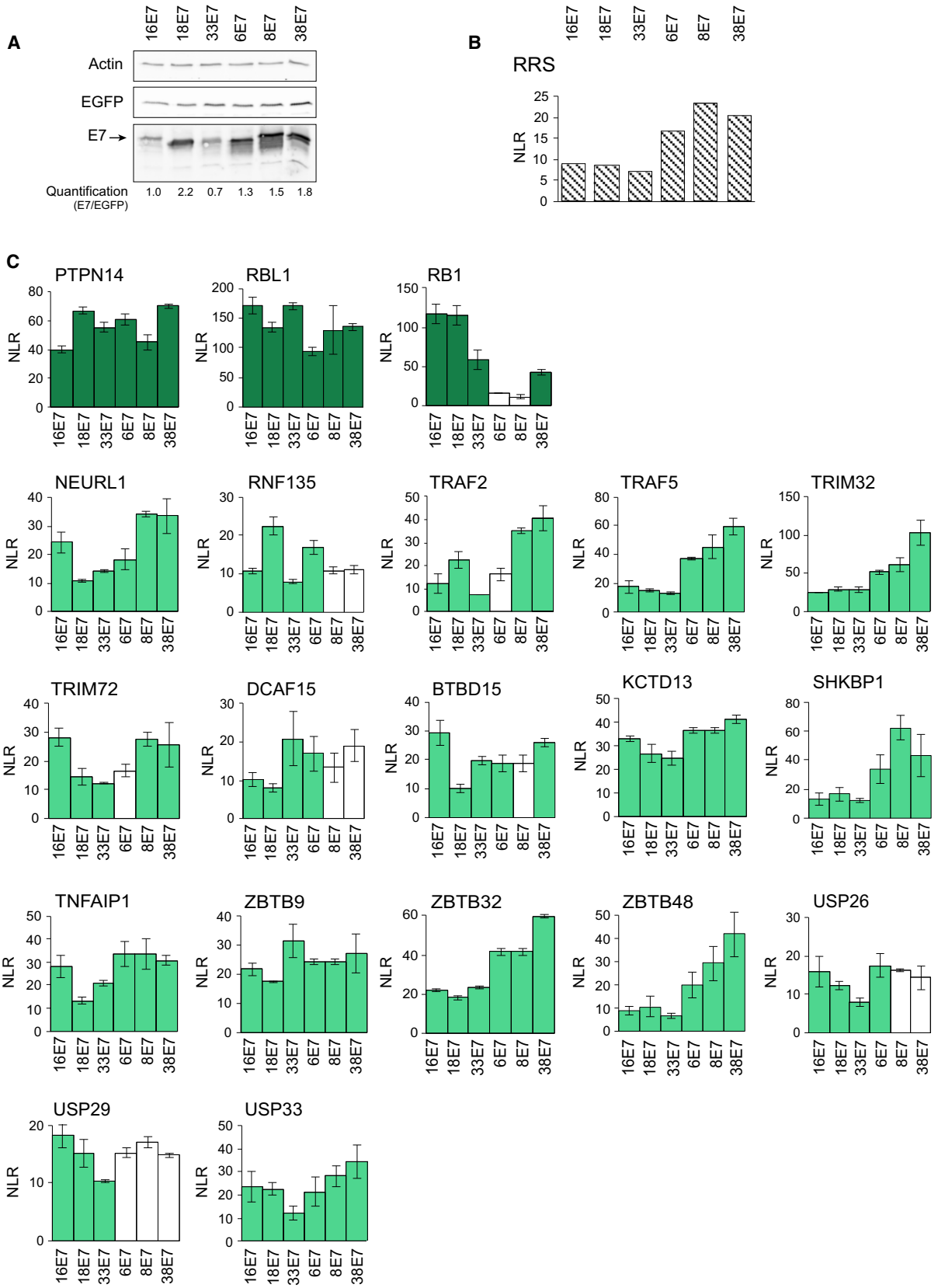


Fig. 7. Binding of UPS factors with E6 proteins from six different HPV types in GPCA. (A) HEK 293T cells (4×10^5) were transfected with pSPICA-N2-16E6, pSPICA-N2-18E6 pSPICA-N2-33E6, pSPICA-N2-6E6, pSPICA-N2-8E6 or pSPICA-N2-38E6 plasmids, and pEGFP-C3 plasmid as internal control. The cleared cellular lysates were analyzed by western blot using anti-Gluc, anti-actin, and anti-GFP antibodies. HEK293T cells (28×10^3) were cotransfected with plasmids pSPICA-N1-MAGI-1, pSPICA-N1-SMAD3, pSPICA-N1-RRS, pSPICA-N1-UPS and pSPICA-N2-16E6, pSPICA-N2-18E6, pSPICA-N2-33E6, pSPICA-N2-6E6, pSPICA-N2-8E6, or pSPICA-N2-38E6. (B) NLR values representing the upper limit of the confidence interval calculated for the different E6/RRS pairs. (C) Representative NLR values for interaction between E6 proteins and MAGI-1, SMAD3, and UPS factors. The boxes are colored in blue (positive interaction) when the NLR was above the upper limit of the confidence interval defined with RRS. (D) HEK293T cells were cotransfected with pCineo3xFlag-16E6, pCineo-3xFlag-18E6 or pCineo-3xFlag-33E6, and pSPICA-N1-E6AP. The lysates were incubated with anti-FLAG magnetic beads. Total proteins (input) and IP proteins were analyzed by western blot by using anti-Gluc antibody and anti-Flag antibody. Note: The E6 proteins migrate at the same size than the light chain of immunoglobulins in the IP samples.


A
16E7
18E7
33E7
6E7
8E7
38E7
Actin
EGFP
E7
Quantification (E7/EGFP)
1.0
2.2
0.7
1.3
1.5
1.8
B
RRS
NLR
C
PTPN14
RBL1
RB1
NEURL1
RNF135
TRAF2
TRAF5
TRIM32
TRIM72
DCAF15
BTBD15
KCTD13
SHKBP1
TNFAIP1
ZBTB9
ZBTB32
ZBTB48
USP26
USP29
USP33

with E6 proteins from all the tested HPV types, whereas interactions with LNX4 were restricted to HPV16. In contrast, the interaction with LNX3 was strong with E6 proteins from high-risk mucosal HPV types. The high GPCA signal obtained with HPV33 E6 and E6AP is likely to be a consequence of its higher expression by comparison to HPV16 and HPV18 E6 proteins. Indeed, E6AP was equally IP by HPV16, HPV18, or HPV33 E6 proteins (Fig. 7D).

For E7 proteins, we found a difference in the strength of binding to RB1 and RBL1 (Fig. 8C, Table S8). The binding of E7 proteins from HPV6 and HPV38 with RB1 was much weaker than for E7 proteins from high-risk mucosal HPV types. In our assay, the HPV8 E7 did not seem to interact with RB1. In contrast, we observed an interaction between RBL1 and all E7 proteins. Nevertheless, the association appeared stronger with high-risk mucosal HPV E7 proteins. The phosphatase PTPN14 was able to interact with all E7 proteins as previously published [30]. We observed a positive interaction for all E7 proteins with the RING-type Ub ligases (NEURL1, TRAF5, TRIM32), the BTB-containing proteins (TNFAIP1, KCTD13, ZBTB9, ZBTB32, ZBTB48, SHKBP1), and the DUB USP33.

Discussion

We used the split *G. princeps* luciferase protein complementation assay (GPCA) to identify interactions between HPV E6 and E7 proteins and effectors of the UPS. Here, GPCA is used for the first time as a PPI screening approach and not just as a validation assay [29,35]. GPCA is a sensitive unbiased approach (in which each potential interacting protein is equally overexpressed in the mammalian cell), complementary to the most popular interactomic methods, namely yeast two-hybrid and AP/MS. GPCA mostly detects direct binary interactions. Indeed, the protein pair HPV16 E6 and P53, which interacts via E6AP, gave no GPCA signal (data not shown). This is probably due to the fact that cellular E6AP protein level is not sufficiently high to bridge the overexpressed Gluc2-E6 and Gluc1-P53.

We assembled and screened a UPS library consisting in 590 cDNAs encoding human proteins and covering about 50% of the human ubiquitination system. To minimize the selection of nonspecific interactions, we adopted a stepwise GPCA-based screening strategy. First, we measured luminescence values for all Gluc1-UPS constructs in the presence of the Gluc2 fragment, in order to determine the 'background' GPCA signal. We found that only few Gluc1-fused proteins (mainly TRAF3IP2 and UBB) produced a high level of background GPCA signal. Interestingly, these proteins did not co-IP with either E6 or E7, arguing that they might be some false positives of the GPCA. Of note, this determination of the 'background' signal of each UPS factor will serve for any further screening of the UPS library. Second, the potential interacting pairs issued from the first global screen were retested in a second round of GPCA experiment. Third, we tested the HPV E6 and E7 proteins against supposedly non-interacting proteins from a RRS set, which enabled the calculation of a confidence interval. Hence, a protein pair was considered as interacting if the NLR was above the upper limit of the confidence interval calculated for the RRS set. Fourth, we used co-IP as an orthogonal approach to confirm some PPIs detected by the GPCA approach. About 25% of the targets identified in the second round of GPCA experiments were validated as strong binding partners by co-IP. This overlap score for PPI methods is in the same range of what has been reported in a previous study, in which only 25% of well-documented positive interactions have been detected by a single assay and not by the four other interactomic approaches [36].

In our assay, the G2 tag was fused to the N terminus of E6 and E7 proteins. Indeed, the position of the tag might interfere with protein interaction. It is known that tagging E6 at its C terminus inhibits its interaction with PDZ domain proteins, whereas tagging E7 at its N terminus suppresses binding to UBR4 [37]. Thus, ideally, screening should be performed twice with both N- and C-terminal tagged bait protein.

To be able to place our findings in the biological context, we searched in the Genotype Tissue

Fig. 8. Binding of UPS factors with E7 proteins from six different HPV types in GPCA. (A) HEK 293T cells (4×10^5) were transfected with pSPICA-N2-16E7, pSPICA-N2-18E7, pSPICA-N2-33E7, pSPICA-N2-6E7, pSPICA-N2-8E7, or pSPICA-N2-38E7 plasmids and pEGFP-C3 plasmid as internal control. The cleared cellular lysates were analyzed by western blot using anti-Gluc, anti-actin, and anti-GFP antibodies. (B) HEK293T cells (28×10^3) were cotransfected with plasmids pSPICA-N1-PTPN14, pSPICA-N1-RB1, pSPICA-N1-RBL1, pSPICA-N1-RRS or pSPICA-N1-UPS and pSPICA-N2, pSPICA-N2-16E7, pSPICA-N2-18E7, pSPICA-N2-33E7, pSPICA-N2-6E7, pSPICA-N2-8E7, or pSPICA-N2-38E7 and assessed for GPCA. (B) NLR values representing the upper limit of the confidence interval calculated for the different E7/RRS pairs. (C) Representative NLR values for interaction between E7 proteins and PTPN14, RBL1, RB1, and UPS factors. The boxes are colored in green (positive interaction) when the NLR was above the upper limit of the confidence interval defined with RRS.

Expression resource (Gtex Project, www.gtexportal.org) for the expression of all the selected UPS factors and we found that these factors are expressed to some extent in most tissues, in particular tissues that are relevant to HPV biology (cervix, esophagus mucosa, vagina, skin). Only TRIM72, USP26, and USP29 seem to be highly expressed in skeletal muscle and testis, respectively.

In our screen, we recovered known partners of E6 such as UBE3A/E6AP, USP15, and LNX3, arguing in favor of the validity of our strategy. We identified two novel RING Ub ligases (MGRN1 and LNX4). Mainly, two modes of interaction have been described for HPV16 E6: capture of cellular LxxLL motifs, via an hydrophobic pocket and binding to PDZ domains, via its C-terminal PDZ-binding motif (PBM). We found that LNX3, LNX4, USP15, and MGRN1 proteins bind equally well to the wild-type and L50E E6 proteins, showing that these interactions are not mediated via LxxLL recognition. Consistently, we could not detect any LxxLL motif in predicted exposed regions of these proteins (data not shown). Interestingly, binding to USP15 and MGRN1 was common to E6 from all the six HPV types tested. The binding to LNX3 was instead dependent on PDZ/PBM interaction. As expected, LNX3 interacted preferentially with high-risk mucosal HPV E6 proteins (HPV16, HPV18, HPV33). Concerning the E6/LNX4 interaction, we previously suggested that it proceeds via the PDZ1 domain of LNX4, which recognizes the PBM of E6 [24]. However, here, we showed that also isoform 2 of LNX4, which lacks the PDZ1 domain, could interact with E6 via a mechanism that is independent of the PBM motif and specific to HPV16 E6.

We found that the HPV16 E7 protein binds to seven BTB domain-containing proteins (BTBD15, KCTD13, TNFAIP1, SHKBP1, ZBTB9, ZBTB32, and ZBTB48) among the 107 present in the UPS library. The functions of these BTB proteins are largely unknown but some of them are identified substrate adaptors of some Cullin 3-RING Ub ligases such as TNFAIP1 and its homolog KCTD13 [38]. This is in agreement with a previous mass spectrometry-based proteomic analysis showing that E7 proteins were associated with Cullin 3–protein complexes [13]. In our screen, the GPCA signal for E7 and Cullin 3 (that is not NEDDylated in our assay) was very low, suggesting that the two proteins did not interact directly but probably via some BTB proteins. Some of BTB proteins contain also a C2H2-type zinc finger (ZBTB9, ZBTB32, BTBD15, ZBTB48) frequently associated with transcriptional activation. This is consistent with the role of E7 as a main regulator of transcription [39]. The E7 interactions with BTB domains should be further characterized by biochemical and structural approaches.

The HPV16 E7 protein was also shown to bind to (a) RING finger Ub ligases implicated in innate immunity, such as RNF135 [40], TRAF2 [41], TRAF5 [42], TRIM32 [43], (b) the RING finger Ub ligase NEURL1 implicated in Notch signaling [44], (c) the substrate adaptor DCAF15 of Cullin 4-DDB1-RING Ub ligase (CRL4) [45], (d) the DUBs USP26, USP29, and USP33. In contrast to the pocket protein family (RB1 and RBL1), the UPS factors interacting with E7 did not target the LxCxE motif. The C-terminal region of HPV16 E7 protein appeared to interact with all UPS factors albeit at different extents.

We found that most of the identified UPS factors bind to E7 proteins across the different HPV types, in contrast to E6 proteins, which display a more specific binding profile according to HPV types.

Human papillomavirus E6 and E7 proteins can hijack some host Ub ligases and DUBs by redirecting them to novel substrates. The best example is the rewiring of E6AP by E6 to novel substrates such as P53 [11] and Scribble [23]. On the other hand, by interacting with some Ub ligases or DUBs, E6 and E7 proteins might interfere with the ubiquitination of their cellular substrates. Up to now, only few substrates of Ub ligases and DUBs have been clearly identified. The DUB proteins are expected to have a much broader substrate specificity as they are eightfold less abundant. USP15 appears to have many different substrates, among them HPV16 E6 itself [17], the Ub ligases TRIM25 [46], and MDM2 [47]. USP29 has been involved in the stabilization of the p53 protein [48] and the checkpoint activator claspin [49]. This latter substrate appears consistent with previous observation showing that HPV16 E7 induces premature degradation of claspin, attenuating the DNA damage checkpoint [50]. USP33 regulates centrosome biogenesis by stabilizing the centriolar protein CP110 [51] and controls the switch between autophagy and innate immune responses by stabilizing RAL-B protein [52].

Analysis of the interactomes of E6 and E7 with UPS revealed that two major cellular pathways are targeted: the p53 and the innate immune pathways. Indeed, p53 is polyubiquitinated by E6AP/E6 [9], TRIM32 [53] and stabilized by USP29 [48]. Ubiquitination and especially nondegradative types of poly-Ub chains (such as K63-linked polyubiquitin chains) are known to regulate the innate immune-sensing pathways [54]. Viral RNA species are detected in the cytosol of infected cells by RIG-1 which is activated by K63-linked polyubiquitination mediated by RNF135 [40] and TRIM25 [55], which is a substrate of USP15

[46]. Sensing of intracellular viral DNAs is mediated by the cGAS/STING pathway. The signaling activity of STING is controlled by K63-linked ubiquitination mediated by TRIM32 [43]. Interestingly, HPV18 E7 has recently been found to interact with the STING protein [56]. The interplay between E7/STING and TRIM32 will require further investigation.

In conclusion, we assembled a library of ORFs encoding UPS factors and described a robust screening strategy to identify direct binary interactions with the UPS. We provide a first characterization of the mechanisms of E6–E7/UPS interactions and of their conservation across the different families. We believe that our systematic interactomic approach will facilitate the deciphering of the Ub system landscape for any further pathogen or host factors.

Materials and methods

DNA constructs

GPCA vectors pSPICA-N1 and pSPICA-N2 (both derived from the pCiNeo mammalian expression vector), respectively, express the Gluc1 and Gluc2 complementary fragments of the *G. princeps* luciferase linked to the N-terminal ends of tested proteins by a flexible hinge polypeptide of 20 amino acid residues.

The ORFs encoding for the E6 and E7 proteins from HPV16 were amplified by PCR and cloned into vector pDONR207 by recombination cloning (Gateway system; Invitrogen, Carlsbad, CA, USA). The resulting entry clones were then transferred into Gateway-compatible GPCA destination vector pSPICA-N2.

The ORFs encoding UPS factors were obtained from the Human ORFeome collections v3.1, 5.1, 7.1, and 8.1 (CCSB, Dana-Farber Institute, Boston, MA, USA) as entry clones into the pDONR223 or pDONR207 vector. They were transferred into the destination vector pSPICA-N1 by Gateway cloning. The resulting pSPICA-N1-UPS plasmids were sequenced in both directions by using forward primer (5′-CAGCTCTTAAGGCTAGAGTAC-3′) and reverse primer (5′-CACTGCATTCTAGTTGTGGTTTGTCC-3′). The sequences were compared to the human genome by performing a blast search to identify the cDNA and the isoform when it is known (Table 1). The presence of truncation or mutation was annotated and the in-frame fusion with the Gluc1 fragment of Gaussia luciferase was verified.

The RRS contains human ORFs encoding proteins randomly picked from the human ORFeome and *a priori* not interacting with the viral proteins [57]: LRCC28, NXP2, NFE2L1, GSTT1, GYPA, UGT3A1, DBH, PLEKHA9, NXPH1. The PRS contains human ORFs encoding a fragment of E6AP (291–875 amino acids) [9], hADA3 [21], MAGI-1 [22], SCRIB [23], DLG3 [24], an homolog of DLG1 [25], IRF3 [26] for E6 and EEF1A1 [29], PKM2 [29], PTPN14 [30], RB1 [27], RBL1 [28] for E7. The PRS and RRS ORFs were cloned into the pSPICA-N1 vector.

Cell line

HEK-293T cells were grown and maintained in Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% FBS at 37 °C with 5% CO_2.

GPCA

HEK-293T cells were seeded in white 96-well plates at a concentration of $2{,}5x10^4$ cells per well. After 24 h, cells were transfected using JetPEI® (Polyplus transfection, Illkirch, France) with 100 ng of pSPICA-N2-E6 or E7 and 100 ng of pSPICA-N1-UPS factors. At 24 h post-transfection, cells were washed with 50 μL of PBS and lysed with 40 μL of Renilla Lysis Buffer (Promega, Madison, WI, USA, E2820) for 30 min. *Gaussia princeps* luciferase enzymatic activity was measured using a Berthold Centro LB960 luminometer by injecting 50 μL of luciferase substrate reagent (Promega, E2820) per well and counting luminescence for 10 s.

Results were expressed as relative luciferase activity (RLU) or as a fold change normalized over the sum of controls, specified herein as NLR. For a given protein pair A/B, NLR = (Gluc1-A + Gluc2-B)/[(Gluc1-A + Gluc2) + (Gluc1 + Gluc2-B)]. Cassonnet *et al.* used an interaction matrix consisting of human proteins of the NFκB and interferon signaling pathways and determined that a cutoff of 3.5 for NLR discriminates interacting pairs from noninteracting pairs with a false-positive background below 2.5% [19]. For the NLR validation experiment, each protein pair was assessed three times or more. The NLR of protein pair was considered as 'validated' if above a threshold value of 3.5 and above the upper limit of the confidence interval defined for the RRS/bait pairs.

UPS library screening using GPCA

For Gluc2, Gluc2-16E6, and Gluc2-16E7, the entire UPS library was screened in one experiment. The distribution of raw luminescence values was represented by KDE plot and Whisker plot. In the latter representation, the median and the interquartile range (IQR) were calculated allowing the selection of outliers.

Expression of HPV E6 and E7 proteins and co-immunoprecipitation experiment

HEK293T cells (in six-well plate) were cotransfected with 1.8 μg of pSPICA-N2-E6 or pSPICA-N2-E7 and 0.2 μg of pEGFP-C3 as an internal transfection efficiency control. After transfection, cells were harvested and lysed in 100 μL of lysis buffer (50 mM Tris/HCl pH 7.4, 150 mM NaCl,

1 mM EDTA, 1% TRITON X-100, and cocktail proteases inhibitors). After centrifugation at 16 000 ***g*** for 20 min at 4 °C, the cellular lysates were analyzed by gel electrophoresis and western blot using anti-Gaussia antibody (NEB, Ipswich, MA, USA E8023S), anti-actin antibody (Sigma-Aldrich, Saint-Louis, MO, USA A2066), and anti-GFP antibody (Roche, Basel, Switzerland).

The E6- and E7-coding sequences were cloned into the pCi-neo-3xFLAG vector allowing expression of the respective proteins with a Nter-3xFlag tag. HEK-293T cells (in six-well plates) were transfected with 1 μg of plasmid Flag-E6, Flag-E7, or empty vector and 1 μg of plasmid expressing UPS factors fused to the Gaussia protein. Cells were harvested after 24 h and lysed by adding 100 μL of lysis buffer (50 mM Tris/HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% TRITON X-100, and cocktail proteases inhibitors). After centrifugation at 16 000 ***g*** for 20 min at 4 °C, 80 μL of the cleared lysates was incubated overnight with 10 μL ANTI-FLAG M2® magnetic beads (Sigma) at 4 °C. Beads were washed five times with lysis buffer and were resuspended with SDS sample buffer. Total proteins (input corresponding to 20%) and IP proteins on ANTI-FLAG magnetic beads were analyzed by gel electrophoresis and western blotting. The poly(vinylidene difluoride) membrane was incubated with rabbit anti-Gaussia luciferase (NEB, E8023S) as primary antibody and with goat HRP-conjugated anti-rabbit IgG as secondary antibody (Advansta, Menlo Park, CA, USA). The Gluc-tagged UPS proteins were revealed by chemiluminescence using WesternBright™ Sirius™ (Advansta), visualized by a LAS 4000 camera. Band intensities in the IP lanes were normalized according to intensities in the input lanes. For each Gluc-UPS protein, we calculated the IP (E6 or E7)/IP (Ctrl) ratio, where IP (E6 or E7) corresponds to the normalized intensity in the E6 or E7 IP lane and IP (Ctrl) corresponds to the normalized intensity in the control IP lane.

Statistical analysis of the data

Whisker plot representation

Whisker plot representation provides statistical distribution of luminescence values. A box encompassing the lower quartile (Q1), the upper quartile (Q3) and whiskers were shown. Whisker length corresponds to 1.5 times the IQR that is equal to the difference between the upper and lower quartiles (IQR = Q3–Q1). The outliers were defined as values falling above a standard threshold (T) defined by the third quartile (Q3) plus 1.5-fold the IQR, $T = Q3 + 1.5(IQR)$.

KDE plot representation

Kernel density plot representation consists of placing a Gaussian function at each luminescence intensity value (represented on the x-axis) and to sum the contributions of all Gaussian functions to obtain a smooth curve. The curve is then plotted using a log scale for intensity values as described previously (Charpentier and Flachaire. SSNR, 2014, https://doi.org/10.2139/ssrn.2514882). The bandwidth of Gaussian function is considered to be equal to the experimental uncertainty, which was estimated at 25% of the intensity based on data reproducibility. This approach performs well enough as judged by the absence of undersmoothed or oversmoothed aspect in the profiles.

Determination of the confidence interval

To estimate the significance of a NLR value for a given protein–protein pair by comparison to the RRS sampling signal, we calculated a confidence interval for the RRS dataset considering the estimated standard error SE and a confidence level of 99.73% (i.e., a risk $\alpha = 0.27\%$) by using the following expression: $\mu - t.SE \pm \mu + t.SE$, where t is the critical value for α two-sided Student's test and for $(n-1)$ degrees of freedom. In a first approximation, we considered the NLR value of a protein pair as statistically significantly different from the RRS if its value is above the upper limit of the confidence interval determined for the RRS dataset.

Acknowledgements

This work was supported by the CNRS, the University of Strasbourg, the Ligue Nationale Contre le Cancer (CCIR-GE), ANR (INFECT-ERA, HPV-motiva project), Alsace Contre le Cancer. JP was a recipient of a PhD fellowship of the Ligue Nationale Contre le Cancer and Alsace Contre le Cancer. EB was a recipient of a MENRT PhD fellowship. We thank the Center for Cancer Systems Biology (CCSB) at the Dana-Farber Cancer Institute (Boston, MA, USA) for providing us with ORFs. We thank Katja Luck for helpful discussion and Noëlla Mukobo for technical assistance.

Author contributions

JP, EB, MLS, and PC performed the experimental work; YN and LJ performed biocomputing and statistical analysis; SW, KZ, and GT performed a critical reading of the manuscript; YJ, CD, and MM performed the planning and the analysis of the experiments; MM wrote the manuscript.

Conflict of interest

The authors have no conflict of interest to declare.

References

1 Hershko A & Ciechanover A (1998) The ubiquitin system. *Annu Rev Biochem* **67**, 425–479.

2 Sowa ME, Bennett EJ, Gygi SP & Harper JW (2009) Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389–403.
3 Li W, Bengtson MH, Ulbrich A, Matsuda A, Reddy VA, Orth A, Chanda SK, Batalov S & Joazeiro CAP (2008) Genome-wide and functional annotation of human E3 ubiquitin ligases identifies MULAN, a mitochondrial E3 that regulates the organelle's dynamics and signaling. *PLoS ONE* **3**, e1487.
4 Gustin JK, Moses AV, Früh K & Douglas JL (2011) Viral takeover of the host ubiquitin system. *Front Microbiol* **2**, 161.
5 zur Hausen H (2002) Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer* **2**, 342–350.
6 Michaud DS, Langevin SM, Eliot M, Nelson HH, Pawlita M, McClean MD & Kelsey KT (2014) High-risk HPV types and head and neck cancer. *Int J Cancer* **135**, 1653–1661.
7 Daling JR, Madeleine MM, Johnson LG, Schwartz SM, Shera KA, Wurscher MA, Carter JJ, Porter PL, Galloway DA & McDougall JK (2004) Human papillomavirus, smoking, and sexual practices in the etiology of anal cancer. *Cancer* **101**, 270–280.
8 Tommasino M (2017) The biology of beta human papillomaviruses. *Virus Res* **231**, 128–138.
9 Scheffner M, Huibregtse JM, Vierstra RD & Howley PM (1993) The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **75**, 495–505.
10 Zanier K, Charbonnier S, Sidi AOMO, McEwen AG, Ferrario MG, Poussin-Courmontagne P, Cura V, Brimer N, Babah KO, Ansari T *et al.* (2013) Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science* **339**, 694–698.
11 Martinez-Zapien D, Ruiz FX, Poirson J, Mitschler A, Ramirez J, Forster A, Cousido-Siah A, Masson M, Vande Pol S, Podjarny A *et al.* (2016) Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature* **529**, 541–545.
12 Huh K, Zhou X, Hayakawa H, Cho J-Y, Libermann TA, Jin J, Harper JW & Munger K (2007) Human papillomavirus type 16 E7 oncoprotein associates with the cullin 2 ubiquitin ligase complex, which contributes to degradation of the retinoblastoma tumor suppressor. *J Virol* **81**, 9737–9747.
13 White EA, Sowa ME, Tan MJA, Jeudy S, Hayes SD, Santha S, Munger K, Harper JW & Howley PM (2012) Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses. *Proc Natl Acad Sci USA* **109**, E260–E267.
14 Reinstein E, Scheffner M, Oren M, Ciechanover A & Schwartz A (2000) Degradation of the E7 human papillomavirus oncoprotein by the ubiquitin-proteasome system: targeting via ubiquitination of the N-terminal residue. *Oncogene* **19**, 5944–5950.
15 Wang J, Sampath A, Raychaudhuri P & Bagchi S (2001) Both Rb and E7 are regulated by the ubiquitin proteasome pathway in HPV-containing cervical tumor cells. *Oncogene* **20**, 4740–4749.
16 Stewart D, Kazemi S, Li S, Massimi P, Banks L, Koromilas AE & Matlashewski G (2004) Ubiquitination and proteasome degradation of the E6 proteins of human papillomavirus types 11 and 18. *J Gen Virol* **85**, 1419–1426.
17 Vos RM, Altreuter J, White EA & Howley PM (2009) The ubiquitin-specific peptidase USP15 regulates human papillomavirus type 16 E6 protein stability. *J Virol* **83**, 8885–8892.
18 Lin C-H, Chang H-S & Yu WCY (2008) USP11 stabilizes HPV-16E7 and further modulates the E7 biological activity. *J Biol Chem* **283**, 15681–15688.
19 Cassonnet P, Rolloy C, Neveu G, Vidalain P-O, Chantier T, Pellet J, Jones L, Muller M, Demeret C, Gaud G *et al.* (2011) Benchmarking a luciferase complementation assay for detecting protein complexes. *Nat Methods* **8**, 990–992.
20 Gao T, Liu Z, Wang Y, Cheng H, Yang Q, Guo A, Ren J & Xue Y (2012) UUCD: a family-based database of ubiquitin and ubiquitin-like conjugation. *Nucleic Acids Res* **41**, D445–D451.
21 Kumar A, Zhao Y, Meng G, Zeng M, Srinivasan S, Delmolino LM, Gao Q, Dimri G, Weber GF, Wazer DE *et al.* (2002) Human papillomavirus oncoprotein E6 inactivates the transcriptional coactivator human ADA3. *Mol Cell Biol* **22**, 5801–5812.
22 Glaunsinger BA, Lee SS, Thomas M, Banks L & Javier R (2000) Interactions of the PDZ-protein MAGI-1 with adenovirus E4-ORF1 and high-risk papillomavirus E6 oncoproteins. *Oncogene* **19**, 5270–5280.
23 Nakagawa S & Huibregtse JM (2000) Human scribble (Vartul) is targeted for ubiquitin-mediated degradation by the high-risk papillomavirus E6 proteins and the E6AP ubiquitin-protein ligase. *Mol Cell Biol* **20**, 8244–8253.
24 Vincentelli R, Luck K, Poirson J, Polanowska J, Abdat J, Blémont M, Turchetto J, Iv F, Ricquier K, Straub M-L *et al.* (2015) Quantifying domain-ligand affinities and specificities by high-throughput holdup assay. *Nat Methods* **12**, 787–793.
25 Gardiol D, Kühne C, Glaunsinger B, Lee SS, Javier R & Banks L (1999) Oncogenic human papillomavirus E6 proteins target the discs large tumour suppressor for proteasome-mediated degradation. *Oncogene* **18**, 5487–5496.
26 Ronco LV, Karpova AY, Vidal M & Howley PM (1998) Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity. *Genes Dev* **12**, 2061–2072.

27 Dyson N, Howley PM, Munger K & Harlow E (1989) The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* **243**, 934–937.
28 McIntyre MC, Ruesch MN & Laimins LA (1996) Human papillomavirus E7 oncoproteins bind a single form of cyclin E in a complex with cdk2 and p107. *Virology* **215**, 73–82.
29 Neveu G, Cassonnet P, Vidalain P-O, Rolloy C, Mendoza J, Jones L, Tangy F, Muller M, Demeret C, Tafforeau L *et al.* (2012) Comparative analysis of virus-host interactomes with a mammalian high-throughput protein complementation assay based on *Gaussia princeps* luciferase. *Methods* **58**, 349–359.
30 White EA, Munger K & Howley PM (2016) High-risk human papillomavirus E7 proteins target PTPN14 for degradation. *MBio* **7**, e01530–16.
31 Thomas M & Banks L (2015) PDZRN3/LNX3 is a novel target of human papillomavirus type 16 (HPV-16) and HPV-18 E6. *J Virol* **89**, 1439–1444.
32 Lee JO, Russo AA & Pavletich NP (1998) Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7. *Nature* **391**, 859–865.
33 Ohlenschläger O, Seiboth T, Zengerling H, Briese L, Marchanka A, Ramachandran R, Baum M, Korbas M, Meyer-Klaucke W, Dürst M *et al.* (2006) Solution structure of the partially folded high-risk human papilloma virus 45 oncoprotein E7. *Oncogene* **25**, 5953–5959.
34 Lee C, Kim D-H, Lee S-H, Su J & Han K-H (2016) Structural investigation on the intrinsically disordered N-terminal region of HPV16 E7 protein. *BMB Rep* **49**, 431–436.
35 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**, e1002761.
36 Braun P, Tasan M, Dreze M, Barrios-Rodiles M, Lemmens I, Yu H, Sahalie JM, Murray RR, Roncari L, de Smet A-S *et al.* (2009) An experimentally derived confidence score for binary protein-protein interactions. *Nat Methods* **6**, 91–97.
37 Huh KW, DeMasi J, Ogawa H, Nakatani Y, Howley PM & Munger K (2005) Association of the human papillomavirus type 16 E7 oncoprotein with the 600-kDa retinoblastoma protein-associated factor, p600. *Proc Natl Acad Sci* **102**, 11492–11497.
38 Chen Y, Yang Z, Meng M, Zhao Y, Dong N, Yan H, Liu L, Ding M, Peng HB & Shao F (2009) Cullin mediates degradation of RhoA through evolutionarily conserved BTB adaptors to control actin cytoskeleton structure and cell movement. *Mol Cell* **35**, 841–855.
39 Songock WK, Kim S-M & Bodily JM (2017) The human papillomavirus E7 oncoprotein as a regulator of transcription. *Virus Res* **231**, 56–75.
40 Oshiumi H, Matsumoto M, Hatakeyama S & Seya T (2009) Riplet/RNF135, a RING finger protein, ubiquitinates RIG-I to promote interferon-beta induction during the early phase of viral infection. *J Biol Chem* **284**, 807–817.
41 Sasai M, Tatematsu M, Oshiumi H, Funami K, Matsumoto M, Hatakeyama S & Seya T (2010) Direct binding of TRAF2 and TRAF6 to TICAM-1/TRIF adaptor participates in activation of the Toll-like receptor 3/4 pathway. *Mol Immunol* **47**, 1283–1291.
42 Tang ED & Wang C-Y (2010) TRAF5 is a downstream target of MAVS in antiviral innate immune signaling. *PLoS ONE* **5**, e9172.
43 Zhang J, Hu M-M, Wang Y-Y & Shu H-B (2012) TRIM32 protein modulates type I interferon induction and cellular antiviral response by targeting MITA/STING protein for K63-linked ubiquitination. *J Biol Chem* **287**, 28646–28655.
44 Koutelou E, Sato S, Tomomori-Sato C, Florens L, Swanson SK, Washburn MP, Kokkinaki M, Conaway RC, Conaway JW & Moschonas NK (2008) Neuralized-like 1 (Neurl1) targeted to the plasma membrane by N-myristoylation regulates the Notch ligand Jagged1. *J Biol Chem* **283**, 3846–3853.
45 Lee J & Zhou P (2007) DCAFs, the missing link of the CUL4-DDB1 ubiquitin ligase. *Mol Cell* **26**, 775–780.
46 Pauli E-K, Chan YK, Davis ME, Gableske S, Wang MK, Feister KF & Gack MU (2014) The ubiquitin-specific protease USP15 promotes RIG-I-mediated antiviral signaling by deubiquitylating TRIM25. *Sci Signal* **7**, ra3, 1–11.
47 Zou Q, Jin J, Hu H, Li HS, Romano S, Xiao Y, Nakaya M, Zhou X, Cheng X, Yang P *et al.* (2014) USP15 stabilizes MDM2 to mediate cancer-cell survival and inhibit antitumor T cell responses. *Nat Immunol* **15**, 562–570.
48 Liu J, Chung H-J, Vogt M, Jin Y, Malide D, He L, Dundr M & Levens D (2011) JTV1 co-activates FBP to induce USP29 transcription and stabilize p53 in response to oxidative stress. *EMBO J* **30**, 846–858.
49 Martín Y, Cabrera E, Amoedo H, Hernández-Pérez S, Domínguez-Kelly R & Freire R (2015) USP29 controls the stability of checkpoint adaptor Claspin by deubiquitination. *Oncogene* **34**, 1058–1063.
50 Spardy N, Covella K, Cha E, Hoskins EE, Wells SI, Duensing A & Duensing S (2009) Human papillomavirus 16 E7 oncoprotein attenuates DNA damage checkpoint control by increasing the proteolytic turnover of claspin. *Cancer Res* **69**, 7022–7029.
51 Li J, D'Angiolella V, Seeley ES, Kim S, Kobayashi T, Fu W, Campos EI, Pagano M & Dynlacht BD (2013) USP33 regulates centrosome biogenesis via

deubiquitination of the centriolar protein CP110. *Nature* **495**, 255–259.
52 Simicek M, Lievens S, Laga M, Guzenko D, Aushev VN, Kalev P, Baietti MF, Strelkov SV, Gevaert K, Tavernier J *et al.* (2013) The deubiquitylase USP33 discriminates between RALB functions in autophagy and innate immune response. *Nat Cell Biol* **15**, 1220–1230.
53 Liu J, Zhang C, Wang XL, Ly P, Belyi V, Xu-Monette ZY, Young KH, Hu W & Feng Z (2014) E3 ubiquitin ligase TRIM32 negatively regulates tumor suppressor p53 to promote tumorigenesis. *Cell Death Differ* **21**, 1792–1804.
54 Davis ME & Gack MU (2015) Ubiquitination in the antiviral immune response. *Virology* **479–480**, 52–65.
55 Gack MU, Shin YC, Joo C-H, Urano T, Liang C, Sun L, Takeuchi O, Akira S, Chen Z, Inoue S *et al.* (2007) TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature* **446**, 916–920.
56 Lau L, Gray EE, Brunette RL & Stetson DB (2015) DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway. *Science* **350**, 568–571.
57 Munier S, Rolland T, Diot C, Jacob Y & Naffakh N (2013) Exploration of binary virus-host interactions using an infectious protein complementation assay. *Mol Cell Proteomics* **12**, 2845–2855.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. GPCA raw luminescence values of HPV16 E6 and E7 proteins tested against proteins of the RRS and the PRS.

Table S2. GPCA raw luminescence values of UPS factors against Gluc2 fragment.

Table S3. GPCA raw luminescence values of UPS factors tested against HPV16 E6 or E7 proteins.

Table S4. NLR values of UPS factors tested against HPV16 E6 (A) or E7 (B) proteins.

Table S5. NLR values of wild-type/mutant HPV16 E6 tested against MAGI-1 and UPS factors.

Table S6. NLR values of wild-type/mutant HPV16 E7 and its regions (CR1 + 2, C terminus) tested against PTPN14, RB1, RBL1 and UPS factors.

Table S7. NLR values of E6 proteins (HPV 16, 18, 33, 6, 8, 38) tested against RRS proteins, MAGI-1, SMAD3 and UPS factors.

Table S8. NLR values of E7 proteins (HPV 16, 18, 33, 6, 8, 38) tested against RRS proteins, PTPN14, RB1, RBL1 and UPS factors.

Table S9. Quantification of the band intensities on the western blots and calculation of the IP ratio.

Review

Virologie

Virologie 2016, 20 (6) : E32-E48

Structure resolution of the trimeric RNA-dependent RNA polymerase of influenza viruses: impact on our understanding of polymerase interactions with host and viral factors

Elise Biquand[1,2]
Caroline Demeret[1,2]

[1] Unité de génétique moléculaire des virus à ARN, CNRS, UMR 3569, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris cedex 15, France
[2] Université Paris Diderot, Sorbonne Paris Cité, Institut Pasteur, 28, rue du Dr. Roux, 75724 Paris cedex 15, France

Abstract. Influenza viruses are segmented negative-sense RNA viruses whose RNA dependant RNA polymerase (RdRp) multiple activities are central for the viral life cycle. The RdRp is composed of three subunits, PB1, PB2 and PA. It binds to the extremities of each vRNA segments encapsidated with multiple copies of the Nucleoprotein (NP), altogether constituting the viral ribonucleoproteins (vRNPs). The RdRp performs both vRNA transcription and replication in the context of vRNP in the nuclei of infected cells. The temporal regulation of RdRp-associated activities is essential for the successful completion of the virus life cycle, but its understanding has been limited by the lack of structural information about the polymerase complex. The atomic-resolution of polymerase complexes from influenza virus type A, type B and type C came out in the past two years. We compile here the data provided by the near-concomitant resolution of several influenza polymerase crystal structures. We will highlight how structural information can contribute to our understanding of the interactions between the RdRp and viral or host factors.

Key words: inlfuenza virus, RNA-dependent RNA polymerase, structure, transcription, replication

Influenza viruses are segmented negative-stranded RNA viruses of long-term human health concerns, which still remain largely unresolved despite the therapeutic advances provided by vaccines and anti-viral therapeutic drugs. A specific difficulty to durably combat influenza infection resides in the high variability of influenza viruses, which leads to the continuous apparition of variants that resist antiviral drugs or escape vaccine-induced immunity. The low fidelity of the viral RNA-dependent RNA polymerase (RdRp) accounts for such variability. The RdRp is a heterotrimer composed of the polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2) and polymerase acidic protein A (PA) proteins. The complex is bound to conserved sequences of the Non-Coding Regions (NCR) at the extremities of each viral RNA segment. The vRNA wraps around oligomers of the NP protein, thereby forming encapsidated structures designed viral ribonucleoproteins (vRNPs). The RdRp catalyzes both the transcription and the replication of viral RNA segments in the nucleus of infected cells, and is thus able to perform distinct activities on the vRNA templates. The RdRp operates only in the context of vRNP, which are the functional units for RdRp transcription and replication activities (for reviews see [1-3]).

Upon vRNPs entry into the nucleus a first round of transcription occurs, where the polymerase transcribes vRNA into viral messenger RNAs (mRNA) [1]. The vRNA 5'NCR and 3' NCR constitute the promoter for transcription of each viral segment, which rely on the primed synthesis of the mRNA using host capped oligonucleotides. To initiate transcription, the viral polymerase first captures selected capped non-coding or pre-messenger host RNA associated with the host transcribing RNA-polymerase II, *via* the Cap-binding domain of PB2. Then PA cleaves it 8-14 nucleotides from the cap *via* its endonuclease activity. The resulting short-capped oligonucleotides serve as primers for the synthesis of viral mRNAs, which are last poly-adenylated by the stuttering of the polymerase on a poly-U sequence located near the 5'end of the RNA template.

Tirés à part : C. Demeret

doi:10.1684/vir.2016.0671

Pour citer cet article : Biquand E, Demeret C. Structure resolution of the trimeric RNA-dependent RNA polymerase of influenza viruses: impact on our understanding of polymerase interactions with host and viral factors. *Virologie* 2016; 20(6) : 32-48 doi:10.1684/vir.2016.0671

The viral replication is taking place in two steps. The RNA segment of each vRNP is first copied in complementary cRNA strand of positive polarity, which forms cRNP after the loading of NP oligomers and the association of a polymerase complex. Each cRNP serves as template for the synthesis of vRNA segments, giving rise to progeny vRNPs (reviewed in [4, 5]). Newly synthesized vRNPs are either transcribed to provide high levels of viral protein expression (secondary transcription), or are exported from the nucleus to be incorporated into novel virions. For vRNA replication, the RdRp operates an unprimed RNA polymerisation, and this process is dependent upon newly synthetized PB1, PB2, PA and NP proteins. The PB1 subunit performs the RNA polymerization, and harbors functional motifs common to all RNA-dependent RNA polymerases [6]. The influenza polymerase is devoid of proof reading activity (or 3'-5' exonuclease activity), accounting for the high mutation rate in the replicated viral segments. Other RNA replication independent activities of the polymerase complex are likely to be involved in the viral cycle, such as a role in the splicing of NS1 viral mRNA [7], or in the trafficking of vRNP through the cytoplasm (reviewed in [3]).

Influenza virus RdRp therefore appears as a multifaceted protein complex, performing different activities directed towards the vRNAs, whose temporal coordination ensures the proper execution of the viral cycle. The processes driving these activities and their regulations are poorly understood at the molecular level so far. Intrinsic characteristics of the polymerase complex, as well as interaction with different sets of viral or host factors, are thought to determine the capacity of influenza virus polymerase keep to proceed to replication or transcription. This notion is nevertheless supported by few experimentally derived pieces of evidence, hampered in particular by the lack of structural knowledge regarding the polymerase complex. Until recently, information was available for the structure of the PA endonuclease N-terminal domain (PA-ENDO), and for PA C-terminal domain in complex with the sixteen N-terminal residues of PB1 [8, 9]. For PB2, the structure of the Cap binding domain (PB2-CAP, [10], and of two C-terminal domains, termed the 627K and NLS domains [11, 12] had been resolved. In addition, the extreme N-terminus of PB2 had been co-crystalized in complex with the C-terminus of PB1 [13]. Nothing was known about the structure of the PB1 subunit except its extreme N and C termini associated with the PA and PB2 subunit respectively. The atomic resolution of influenza viruses type A, type B and type C RdRp crystal structures came out in a remarkably short period of time, concluding long-lasting intensive attempts that confronted challenges of producing sufficient quantities of purified polymerase subunits to be crystalized [14-17].

Two papers published in the *Nature* issue of the 18th of December, 2014, described the structure of the influenza virus polymerases co-crystalized with a synthetic 5'+3' vRNA, mimicking the viral promoter. One was on the RdRp from a bat influenza A virus [16] the other on a human influenza B virus RdRp [15]. Later on came out the structure of the RdRp from the influenza virus C polymerase without any bound viral RNA (apo-FluC) [14]. Then a study published in the early 2016 still incremented the picture, in particular by providing the structure of RdRp from influenza virus B bound to a 5'cRNA [17]. These studies first provided some insights into the mechanistic aspects of the cap-snatching, the primed and unprimed RNA polymerization, which are reviewed in the broader context of negative-strand viral RNA-dependent RNA polymerases in [18]. They also enabled the immediate appreciation that influenza virus polymerase complex can adopt different conformations (*figure 1A*). We will describe here the main characteristics of influenza polymerase structures that have been brought out, and focus on polymerase interaction with viral or host factors.

The RdRp Core: a central invariant fold

The RdRp organizes around an invariant central polymerase body made of the PB1 protein, the N-terminal part of PB2 (PB2-N) and the C-terminal part of PA (PA-C) (*figure 1B*). The PB1 subunit adopts in its central region a canonical RNA polymerase fold, conserved among RNA-dependent RNA polymerases [19]. It consists in a palm sub-domain, exhibiting the functional motifs for RNA polymerization, as well as fingers and thumbs set on both sides of the palm. A large internal cavity provides the catalytic chamber for RNA binding and polymerization. A priming loop emerges within the central cavity from the thumb of PB1 (*figure 2A*) and is involved in the terminal initiation of unprimed RNA synthesis [20]. This fold is flanked on both sides by PB2-N and PA-C (*figure 2B*), involving an intricate intertwining of the polymerase subunits. The buried surface of the central RdRp core is far more extended that envisioned from former crystallographic studies with subunit domains [9, 13, 21]. This results in a somewhat compact structure, from which a long β ribbon of PB1 is extruding, that contains the bipartite NLS of the influenza A and B PB1 proteins (*figure 2B*).

Binding of the vRNA promoter to the core polymerase

The bat influenza virus A (FluA) and the human influenza virus B (FluB) polymerases were both crystallized with

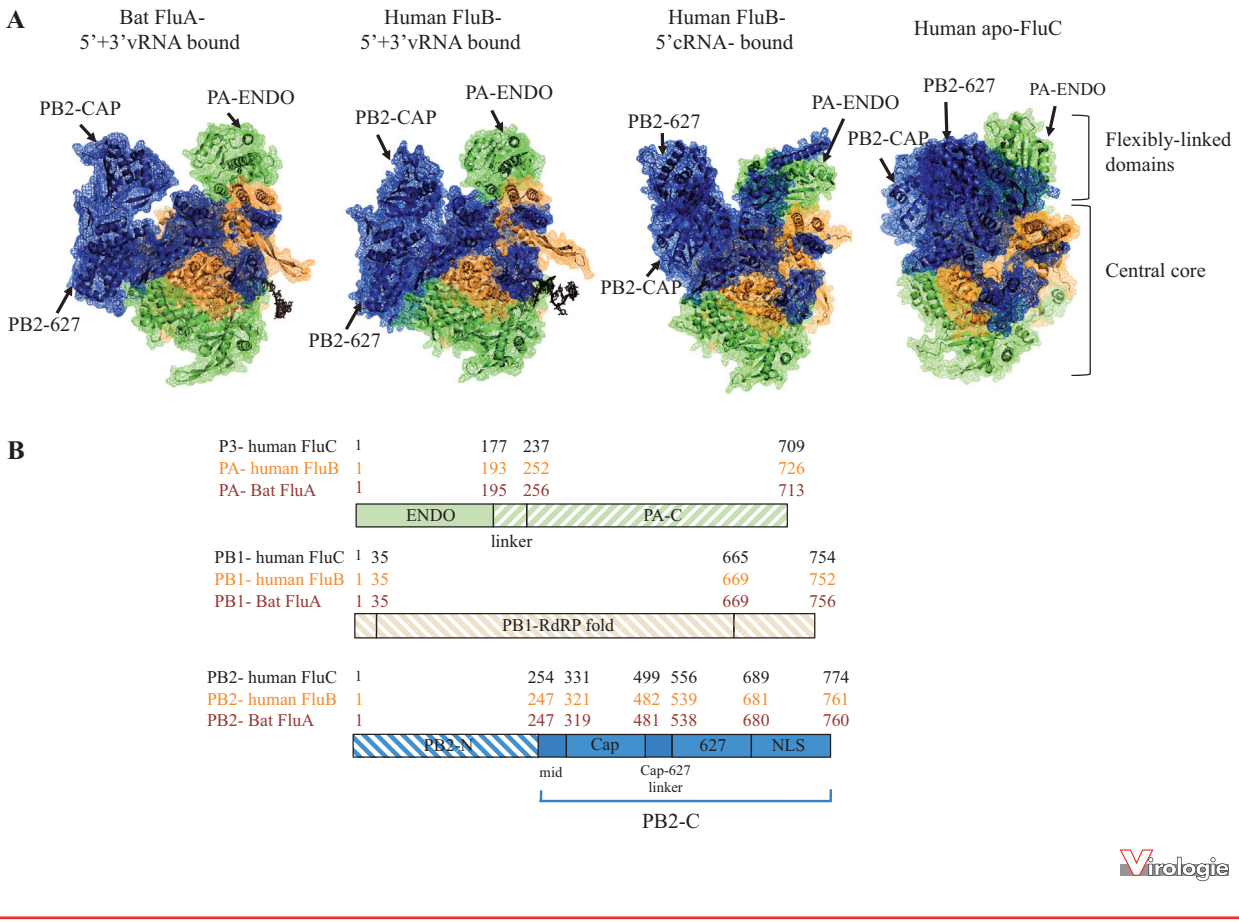


Figure 1. A. Overview of the different polymerase structures. The overall shape of the polymerase complexes in four resolved structures are shown with a fixed position of the central core. The PB1 subunit is colored beige, PB2 blue, PA green. The pdb are as follows: bat FluA bound to 5'+3'vRNA pdb 4WSB [16]; human FluB bound to 5' + 3'vRNA pdb4WSA [15]; human FluB bound to 5'cRNA pdb 5EPI [17], human apo-FluC pdb5d9 [14]. **B.** Shematics of the PA /P3, PB1 and PB2 domain organization. Numbering is given for human bat FluA, human FluB and human FluC as indicated. Hatched domains are those constituting the central core polymerase.

bound vRNA 3' and 5' extremities mimicking the viral promoter, thereby providing a picture of viral promoter binding to the influenza virus polymerase. The distal region of the vRNA promoter is base-paired and not associated with the polymerase, while 3' and 5' extremities are separated (*figure 3A*), and bind to different interfaces of the polymerase core.

The binding of 3' vRNA involves all three subunits and numerous base-specific RNA-protein interactions (*figure 3B*). The bound 3' vRNA adopts an arc shape, its 3' extremity lies close to the distal double-stranded portion of the promoter (*figure 3C*). It is predictable that the 3'end of the RNA template changes orientation to enter into the catalytic site, which probably involves modifications of polymerase core conformation at the vicinity of RNA binding interfaces. The 5' promoter end forms a stem loop structure through internal base-pairing (*figure 3A*) essentially as predicted [22]. It binds to a pocket formed by the PB1 and PA subunits through multiple amino acids-base interactions (*figure 3D*). The binding characteristics of 5' RNA were found similar when only a 5'cRNA or 5'vRNA is bound to the polymerase [17]. In all, The binding of viral RNA to the core turns out to be full part of the polymerase structures.

Distinctively arranged domains of RdRp

The PB2 two-third C-terminal (PB2-C) and PA N-terminal endonuclease domain (PA-ENDO) turn out to be variously

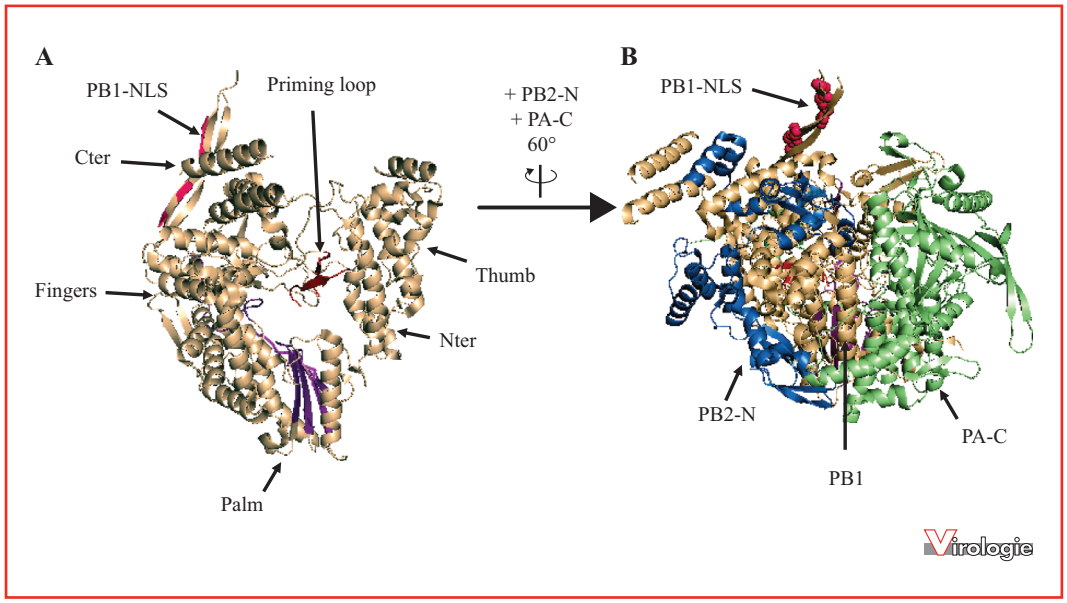


Figure 2. The catalytic core of influenza polymerase. A polymerase fold (pdb 4WRT) displaying the unique resolved structure of the priming loop. Example of the polymerase fold of the flu structure is given (pbd 4WRT) which is the only containing a resolved structure of the priming loop. **A.** Ribbon diagram of the PB1 subunit, with the modules typical of an RNA dependent RNA polymerase fold indicated. The conserved functional motifs of RNA polymerase lying in the palm domain are colored in purple, the priming loop in red and the PB1 bipartite NLS in rust. **B.** Diagram representing the polymerase fold of fluA PB1 flanked by the PA-C on one side and by the PB2-N on the other side of the thumb. PB1, PB2-N and PA-C are colored beige, blue and green respectively the NLS of PB1 is represented with rust spheres.

arranged according to the polymerase structures, highlighting a noteworthy plasticity relative to the central core, and resulting in notably differing overall shapes of the polymerase complexes (*figure 1A*). These domains are thus likely to be key players of the versatility of the polymerase. The individual structural modules that can be identified within PB2-C, consisting in mid, cap-binding, cap-627 linker, 627 and NLS domains [17], remain individually unchanged, but differ in their respective position. Indeed, the cap-binding and 627 domains are differently rotated around a rigid element made of the mid and cap-627 linker (*figure 4*). The PB2 subunit therefore adopts distinct conformations according to the polymerase configurations. The same conformation as the one found in the c5'RNA bound/Apo polymerase conformation was also recovered upon the expression of isolated PB2-C (pdb 5FMM and pdb 5FML) [17]. However the resolution of the NLS domain was not resolved in this latter structure.

Transcription-competent polymerase conformation

The polymerases bound to the 5' + 3' vRNA mimicking viral promoter (bat FluA pdb 4WSB and human FluB pdb 4WSA) [15, 16], or to v5'only-bound polymerase (pdb 5FMZ, [17]), adopt a similar U-shape fold (*figure 1*).

In this configuration, both PB2-CAP and PA-ENDO form protruding arms facing each other across a solvent-exposed channel and the channel entry joins the catalytic center of the core polymerase (*figure 5A*). Two alternative positions of the PB2-CAP can be distinguished. In the bat fluA structure, the cap-binding site is pointing to the endonuclease active site across the separating channel (*figure 5A*). In such orientation, upon binding of the host pre mRNA cap, the 10th to 15th nt would face the endonuclease catalytic site of PA, and cleavage could occur at these positions. This is in agreement with the observed lengths of capped oligonucleotides released by PA-mediated cleavage.

The cap-binding domain of PB2 in the promoter-bound fluB polymerase structure is rotated by 70 Å compared to its position in the fluA polymerase, so that the cap-binding site of PB2 is directed towards the polymerase RNA catalytic cavity, and away from the endonuclease domain of PA (*figure 5A* and *B*). In this configuration, the 3' end of the capped oligonucleotide would be channeled into the polymerase catalytic center, providing polymerase configuration assumed to be competent for an initiation of primed viral mRNA synthesis [15]. The differences in both promoter-bound structures suggest that a conformational switch must occur for viral mRNA synthesis during transcription [15].

Based on modeling with the template-product elongation complex from the poliovirus polymerase [23], the template

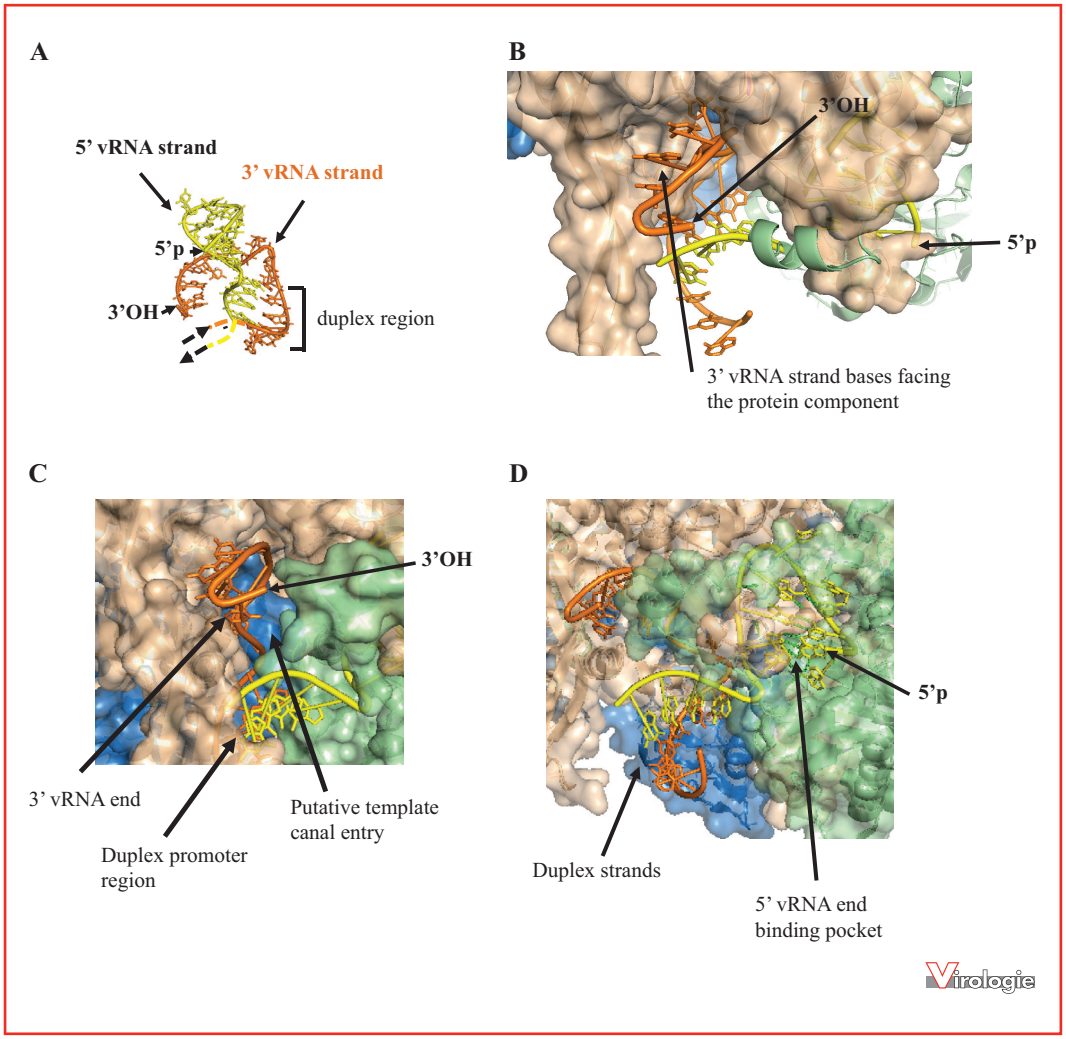


Figure 3. RNA promoter binding to the polymerase. All images are generated from the FluB structure pdb 4WRT [15] which is the only crystal containing full-length 5' and 3' vRNA strands used as surrogate of viral promoter. **A.** Representation of the 5' and 3' vRNA promoter, highlighting the base pairing of the distal region and configuration of single-strand extremities. **B.** Surface representation of the polymerase promoter-binding interface, showing the RNA-protein binding interface with the 3' vRNA extremity. The 5' and 3' ends of vRNA strands are indicated. **C.** Surface representation of the vRNA 3'end at the entry of the putative template canal modeled from the superimposition with the primer-template structure of the Norwalk virus polymerase [15]. The 3' end of the 3' vRNA strand is indicated. **D.** Surface representation of the polymerase promoter-binding interface highlighting the binding 5'vRNA hook in a pocket formed by PB1 and PA. The 5' end of the 5' vRNA strand is indicated.

exit would be blocked by an α helical bundle of PB2-N named helical lid [15]. This lid has been suggested to participate to product-template strands separation, which in turn should exit the core polymerase through distinct tunnels [23]. The transcript strand has been proposed to leave the polymerase core between the PB2-CAP, and 627-domains of PB2 [15] (*figure 5C*). When elongation proceeds close to the 5'end of RNA template, template strand progression would be opposed by the tight binding of the 5' extremity to the core, thereby provoking the stuttering of the polymerase on the oligo-U sequence 17-22 nucleotides upstream the 5' extremity, which has been shown to generate addition of poly-A to the viral mRNA [24].

5'cRNA and apo-FluC conformations

In the polymerase bound to a 5'cRNA end [17], there is an entirely different arrangement of the PB2 subunit. In particular the PB2-627 and PB2-NLS domains are

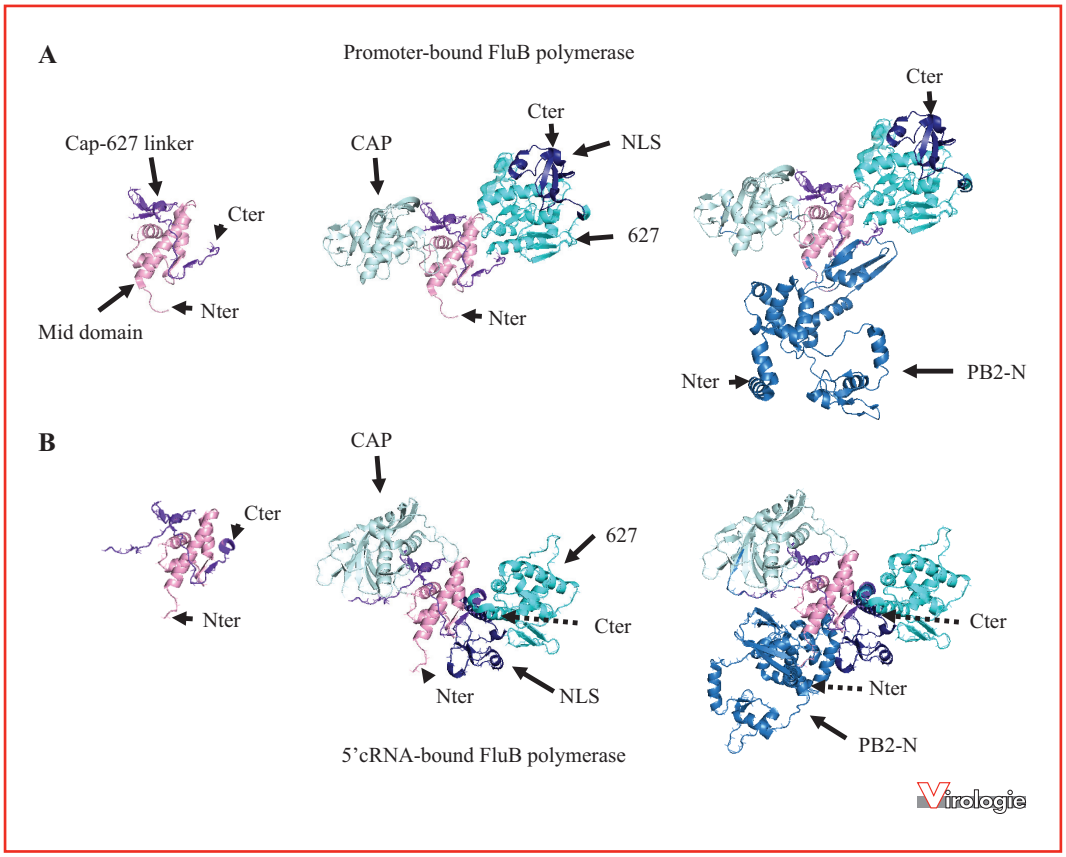


Figure 4. Cartoon representation of PB2 folding in promoter-bound (shown for FluB pdb 4WSA) and in the 5'cRNA bound (FluB pdb 5EPI) polymerase configurations. The step-wise representation of the PB2 structure starts with the rigid axis constituted by the cap-627 linker (purple) and mid domain (pink), then with the cap (light blue), 627 (cyan) and NLS (violet) subdomains, last in full PB2 protein (PB2-N in sky blue). N- and C-terminus extremities of the corresponding polypeptides are indicated.

displaced relative to the transcription-competent polymerase (*figure 6A*).

The PB2-CAP contacts both the PB1 subunit and PB2 cap-627 linker. The cap-binding site is packed against the PB2 cap-627 linker and consequently is not accessible (*figure 6B*). A similar organization is observed in the apo-FluC polymerase structure (*figure 6A*) [14]. The obstruction of the cap-binding site of PB2 suggests a debilitated cap-snatching activity, which has been confirmed experimentally [14, 17] and argues that the primed synthesis of viral mRNA is not possible in such configuration. In contrast, the 5'cRNA polymerase configuration might be operational for the cRNA to vRNA step of replication [17]. The apo-Flu conformation has been proposed to be a closed conformation [14]. It might nevertheless perform replication upon stabilization of the priming loop, which is disordered in the resolved structure [14].

In the 5' cRNA-bound fluB polymerase, the NLS domain of PB2 is tightly apposed to PA-ENDO through a large interaction interface, and this domain packing also involves a rotation of PA-ENDO relative to the promoter bound structure. In this configuration, the PB2 NLS domain lies apart from PB2 627 domain. In the apo-fluC, similar interaction interfaces exist between PB2-NLS and PA-ENDO (P3 in fluC), in an orientation where PB2-NLS and PB2-627 lies together (*figure 6A*).

The remarkable flexibility of influenza virus polymerase

Divergent structural organizations between the promoter-bound and the 5'cRNA-bound or apo-polymerases are emerging from the various cristal structures resolved so far. Analysis of the behavior in solution of different polymerase complexes highlights various degrees of compaction [17]. Indeed, the apopolymerase demonstrates an extended conformation, which was also detected for a polymerase complex bound to only a 3' vRNA end, and seems to be flexible owing to the conformational heterogeneity

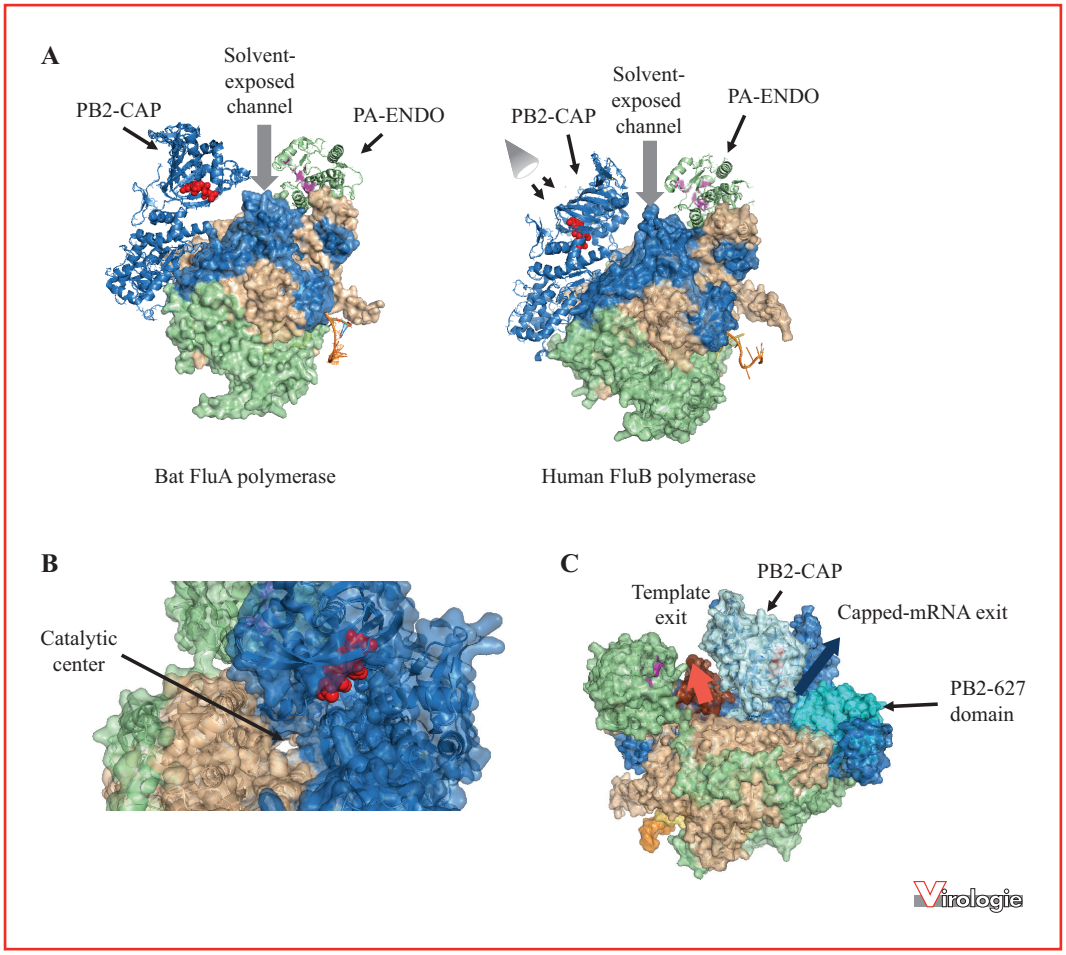


Figure 5. The transcription-competent polymerases. **A.** Structures of 5' + 3' vRNA bound polymerases bat FluA (pdb 4WSB) and human FluB (pdb 4WSA), shown with the PB2-C and PA-ENDO domains in ribbon diagram and the rest of the polymerase in a surface view. A cap bound to the PB2 cap-binding domain is shown in red, and the catalytic site of PA endonuclease domain is colored purple. Color code for PB1, PB2 and PA subunits are as in *figure 1*. **B.** Surface view of the human-fluB polymerase conformation (pdb 4WSA) highlighting the orientation of a cap bound to the PB2 cap binding site facing the catalytic center. The view position is across the PB2 cap binding domain as indicated in the panel a. **C.** Surface representation of the transcription-competent polymerase (from pdb 4WSA) highlighting the position of template exit indicated by the location of the obstructing PB2 helical lid (colored brown), and of the putative mRNA exit tunnel.

observed in solution. Other polymerase forms, consistently including a bound 5'RNA, are more compact in solution. The hypothesis is that the high-affinity binding of a 5'vRNA (in the range of 2 nM [25]) or 5'cRNA extremity to the core rigidifies the polymerase in a compacted conformation and somehow shapes different polymerase conformations according to the type of 5' RNA end (5'vRNA or 5'cRNA) binding [17]. A wider range of alternative polymerase conformation has been detected in cross-linking experiments than in the crystal structures. The structural modules of PB2-C (*i.e.* mid, cap binding, cap-627-linker, NLS) seem to adopt a large variety of relative dispositions, both at the level of intra-protein subdomains packing and relative to other subunits PB1 and PA [17].

Potential higher ordered polymerase forms

An additional level of complexity arises from the possible involvement of higher oligomerzation states of the polymerase trimeric complex. Indeed, cryo-electron microscopy studies with the polymerase complex of an influenza A H5N1 virus devoid of the PB2 C-terminal domain (PB2 N 1-130) showed that this incomplete polymerase, termed subcomplex I, assembled into dimers in solution, and further formed tetramers upon binding of 5'v or cRNA strands [26]. These data led to the proposal that binding of the vRNA promoter may regulate higher degree of polymerase oligomerization. The level of resolution reached in these

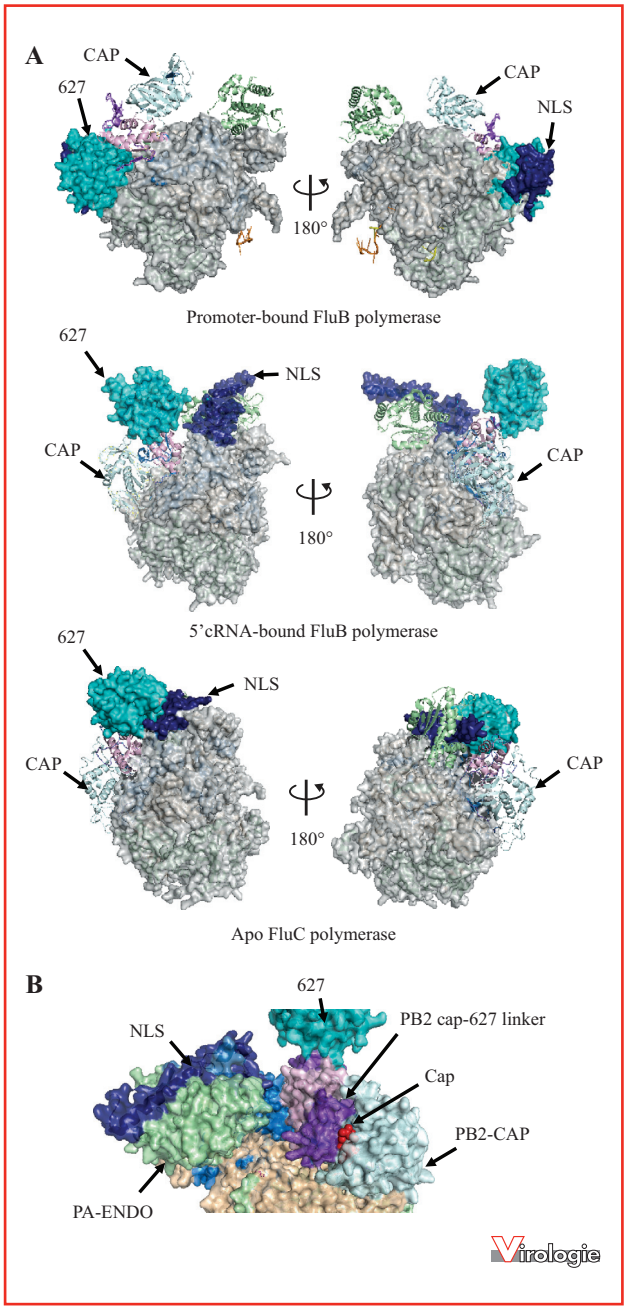


Figure 6. The different arrangements of the PB2 subunit in polymerase configurations. **A.** Representation of the structures of the FluB polymerase when bound to the 5'+3' vRNA viral promoter (pdb 4WSA) or to the 5'cRNA only (pdb 5EPI), and of the apo-FluC polymerase (pdb 5d9). The polymerase core is represented in grey surface, the PB2-C is colored in a ribbon representation for cap domain (light blue), cap-627 linker (purple) and mid domain (pink). PB2-627 (cyan) and NLS (dark blue) domains are in surface representation. The PA-ENDO is shown in ribbon and colored green. Polymerase structures are represented with similar orientation of the polymerase core to highlight the rearrangements of the PB2-C domains in the polymerase complex. **B.** Surface representation of PB2-CAP bound to a cap represented in red spheres in the 5'cRNA polymerase, highlighting the obstruction of the cap-binding site through apposition to the PB2 cap-627 linker.

studies allowed to specify the oligomerization interfaces, but did not provide all the precise locations of the involved residues. It nevertheless indicated that the dimer interface involves α helices from the PA-C (aa 292/291 to 355/352, coordinates for the H5N1 human FluA and human FluB respectively) and PB2-N domains (attributed to aa 86/88-130/132 for H5N1 FluA and human FluB respectively [26]), which are lying adjacent in the cryo-EM derived structure (*figure 7A*) [26]. A weakest interaction mainly involving PB1-N α helical fingers would mediate tetramer formation, which however could not be precisely characterized except for the participation of PB1 aa 180-205 and 355-375.

Oligomerization of polymerase is compatible with the transreplication model, which proposes the recruitment of a soluble polymerase to the cRNA through interactions with the resident RNP-bound polymerase [27]. In contrast, the synthesis of viral mRNA would be carried out by the RNP-bound polymerase (*cis* conformation) [4, 27]. Polymerase oligomerization state transitions are proposed to take part to the regulation between the transcription and replication activities and the polymerase conformation compatible with dimer and tetramer formation to correspond to the replication active state [26].

However, in the c5'RNA-bound/apo-polymerases, the participating PB2 and PA α helices are not contiguous and the PB2 α helices are partly embedded, thus not fully accessible to the polymerase surface (*figure 7B*). Overall, the crystal structures obtained so far, including the promoter-bound polymerase conformation, are not compatible with the cryo-EM model for dimer formation. Nevertheless, it still can be proposed that polymerase conformations that are competent for dimer (and possibly tetramer) formation have not yet been captured by crystallographic studies, owing to the polymerase flexibility.

Polymerase in the context of the RNP

The functional transcription/replication unit of influenza viruses is the RNP. Assembly of viral RNA segments into RNPs is necessary for the viral replication process, which requires the binding of NP to nascent vRNA or cRNA products. In contrast, the viral mRNAs are not encapsidated by NP upon synthesis during the transcription process. For both replication and transcription, it is thought that the template RNA strand locally disassemble from NP to enter

Figure 6. (*Continued*)

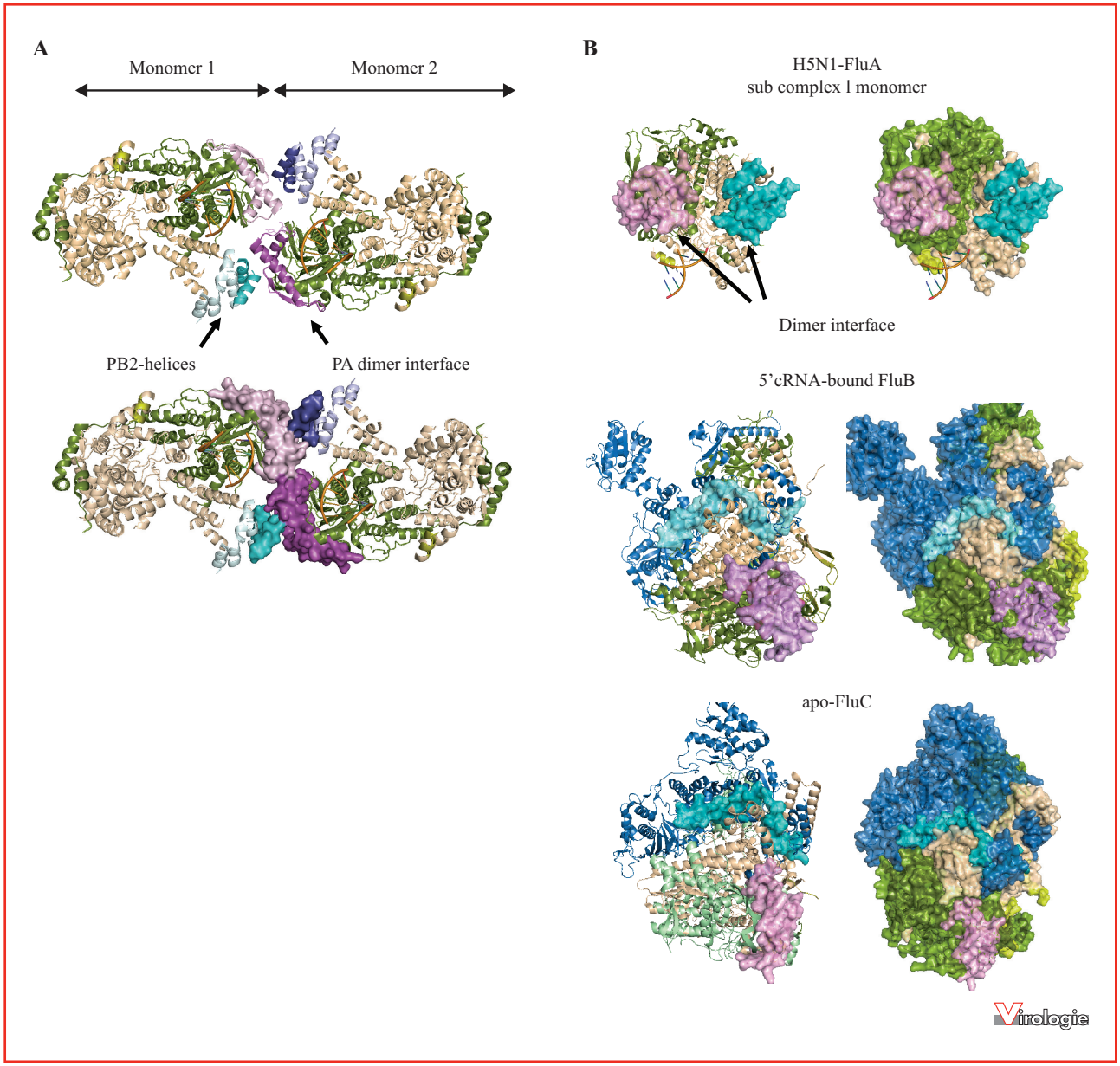


Figure 7. Polymerase dimer formation **A.** Dimer observed for FluA H5N1 subcomplex 1 (pdb3j9). Top: Ribbon representation of the dimer, PB1 is colored beige, PA is colored green, PB2-αhelices (attributed to aa 86 to 130) are shown in cyan (monomer 1) or blue (monomer 2). The PA α-helices involved in the dimer interface are shown in light pink (monomer 1) and dark pink (monomer2), PB2 helices involved in dimer interface are shown in cyan (monomer 1) and dark blue (monomer 2). Amino acids of PB1 proposed to be involved in tetramer formation are in colored yellow. Bottom: the surface of the dimer interface highlights the contacts between each monomer. **B.** Structural elements proposed for dimer interface are shown in surface representation colored as in one monomer of fluA H5N1 subcomplex 1, in the 5'cRNA bound form of fluB polymerase (pdb 5 EPI) and in the apo-fluC polymerase (pdb5d9). They are shown with the rest of the polymerase in ribbon (left) and in surface (right) representation to highlight both the contiguity and the accessibility of the intervening α-helices in the dimer formation.

the catalytic core of the polymerase, and binds back to NP upon exit. The NP protein is also proposed to be involved in the regulation of transcription-replication switch of the polymerase activities [28].

The cryo-EM based three dimensional reconstruction of a mini-RNP, made with shorten vRNA, indicated protein contacts between the promoter-proximal NP monomers and the polymerase complex, without precise identification of

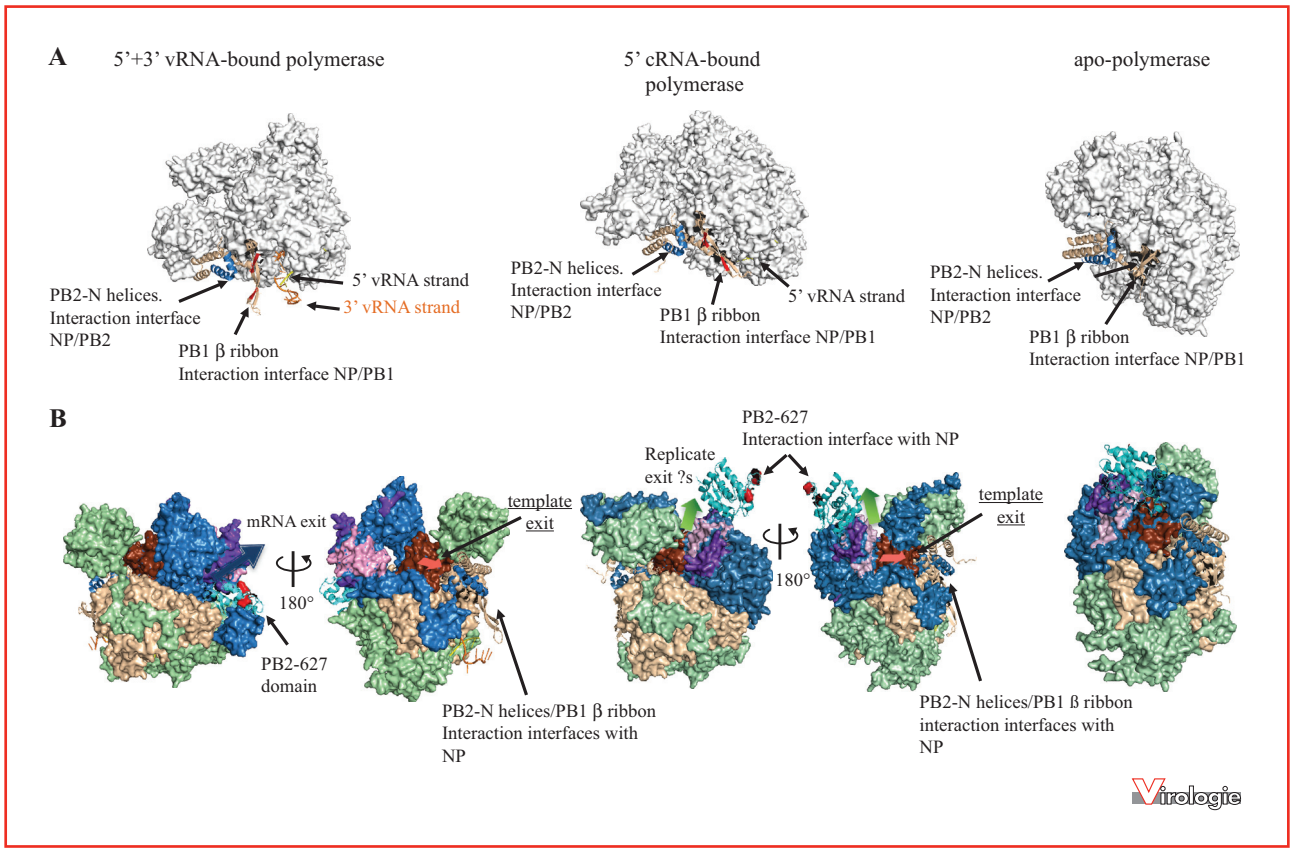


Figure 8. Interaction with NP. **A.** Interaction interface with NP. The PB1 β ribbon interacting with NP is colored beige, 5' v or cRNA in yellow, 3'vRNA in orange, PB2 N-terminal α helices involved in NP binding colored blue, the PB1-C-terminal α helices involved in the PB2-N/PB1-C interaction interface observed in isolated domains colored beige. The position of the PB1 NLS in the PB1β ribbon is labeled red for the FluB 5' + 3'vRNA-bound and 5'c-RNA bound polymerases, but not in the FluC apo polymerase structure where it is not conserved. Both PB1 β ribbon and RNA strands are shown in ribbon relative to the rest of the polymerase in white surface. Images are built from pdb 4WSA for the FluB promoter-bound polymerase, pdb 5EPI for 5'cRNA FluB bound polymerase, pdb 5j9 for apo-FluC polymerase. **B.** Location of the interaction interfaces with NP in the promoter-bound FluB (left, pdb 4WSA), the 5'cRNA-bound fluB (right, pdb 5EPI) and the apo-FluC (pdb5d9) polymerases relative to the putative exit tunnels for products (dark blue arrow for mRNA product, red arrow for replicate product) and to the putative position of template exit, indicated by the location of the obstructing PB2 helical lid (colored brown). Note that the putative template exit is indicated on the face of the polymerase where the template is expected to emerge. The PB2-627K domain colored is cyan, residues involved in interaction with NP are in red. The PB2 627K, PB1C/PB2N helices bundle and PB1 β ribbon are in ribbon representation.

the molecular interfaces [29, 30]. The structure nevertheless suggested specific interactions between the PB1 and PB2 polymerase subunits and the two molecules of NP adjacent to the polymerase, in line with previous studies showing interaction of NP with PB1 and PB2, but not with PA [31, 32]. By fitting the polymerase structure into the mini-RNP pseudo-atomic model, the interaction interface between PB1 and NP has been located to the flexible β ribbon of extruding from the core polymerase, which contains the PB1 NLS at least for influenza A and B and also takes part to the binding interface of the polymerase with the 3' vRNA end [15]. This PB1 β ribbon projects away from the polymerase core in the promoter-bound polymerases (*figure 8A*) where it is predicted to lie in close proximity to the promoter-proximal NP protein. In the absence of 3'vRNA, this β ribbon is packed on the polymerase core [15, 17].

The PB2 subunit lacked electron density in the pseudo atomic structure of mini-RNP, so that only the first α helices (aa 1-24) of PB2 could be fitted in the mini RNP structure, lying close to NP [15]. These extreme N-terminal PB2 α helices are consistently at the surface of the polymerase core, close to the PB1 β-ribbon binding the other NP monomer (*figure 8A*). They also correspond to the PB2-N/PB1-C interaction interface resolved from isolated domain [13], which may explain why a strong overlap had been detected between the interaction domains with PB1 and with NP in the PB2 protein [33]. The PB2 N-terminal α

helices are lying close to the exit of the template strand (*figure 8B*), suggesting that they constitute NP binding interfaces relevant for the loading of outgoing RNA template with NP (*figure 8B*). Additional PB2/NP binding interfaces have been observed in the mini-RNP structure but could not be mapped [29, 30].

On the opposite side of the polymerase, the PB2 627K domain has been shown to interact with NP [32, 33]. The host adaptive 627 residue (generally a lysine in human polymerases and a glutamic acid in avian polymerases) together with amino acid 630 have been shown to modulate PB2 binding with NP and to affect RNP activities [32], although it has been challenged [34]. In the c5'RNA-bound configuration, the 627K domain is well accessible, protruding from the rest of the polymerase, with both 627 and 630 residues exposed. Considering the possibility that this configuration can perform the c to v RNA replication, the binding of NP to PB2-627K domain would be involved in the loading of NP on nascent RNA replicates, which would imply that the exit region of replicated RNA strand lies in the vicinity of PB2 627K domain. In such context, it can be postulated that the exit region of the outgoing replicated strand follows the tunnel delimited on one side by the packed PB2-NLS/PA-ENDO, and on the other side by the mid/cap-627K linker subdomains of PB2-C (*figure 8B*). The 627K domain is more closely packed against the polymerase core in other polymerase conformations, suggesting a reduced flexibility and a limited surface availability for interaction with NP (*figure 8B*).

Amino acids involved in host adaptation

Mammalian adaptation of influenza A viruses is a field of extended research, and several residues have been found implicated in the host adaptation process. Compiling residues identified in a series of studies as being under adaptation pressure pointed to 24 residues in PB2, 6 on PB1, PB2 and 3 on PA proteins [35, 36]. These residues, designed thereafter as adaptative residues, are thought to be involved in optimized interactions with factors of the novel host. The problematic of host specificity is more relevant for the influenza A viruses, but the position of the adaptative mutations will be shown in FluB polymerase structures, because they are the only ones enabling a comparison between transcription-competent and 5'cRNA-bound polymerases. Given the high similarity between the structures of FluA and FluB polymerases, it can be assumed that the positions of adaptive residues are transferable to the FluA polymerase.

Two residues of PA playing a role in the host adaptation are located on the surface of the core polymerase domain (M352....) and the endonuclease domain (FluB E98, FluA T97) (*figure 9*). For PB1, six residues have been described, two of them are exposed at the surface of the polymerase (FluB 52K, 583K; FluA 52K, 584K) whereas others are embedded into polymerase (105A and 516 V FluB, 105N and 517 V in fluA). Of note, one residue, 13P (fluB and Flu A) is hidden in the transcription configuration but well exposed in the 5'cRNA-bound configuration, while the reverse applies to D677 (fluB)/S678 (fluA) (*figure 9A*). The adaptative residues are though to be involved in species specific interactions required for optimized replication of influenzaviruses in humans.

Most of the adaptative mutations concern the PB2 subunit. In the N-terminal part of PB2 within the polymerase core, several adaptative mutations are lining the interaction interface with PB1 (*figure 9B*). This suggests that these residues could be involved in interactions of the isolated PB2 subunit with host factors. They may alternatively be necessary for an efficient polymerase complex formation through optimized interaction with a mammalian PB1 subunit. However, equivalent mutations lining inter subunit interfaces are not found in the PB1 and PA subunits. The PB2 adaptative mutations located in the C-terminal 627K and NLS domains *i.e.* E570/T569, M589/T588, Q592/R591, K627/K627, Q664/M661, R686/G682, D705/D701, P706/K702 (fluB/fluA coordinates) are mostly lying at the surface. Their exposure may nevertheless vary according to the polymerase configuration. In the transcription competent form (*figure 9C*, left), only E570 is readily accessible since turned toward the exterior of the polymerase structure, while M589/Q592/K627 form a cluster somewhat embedded in exit mRNA tunnel, and R686 lies close to the PA subunit (*figure 9C*). The D705 (fluB) D701 (fluA) residue is well exposed on the PB2 NLS domain only in the transcription form of the polymerase, where it forms an exposed cluster with the P706 (fluB) in the vicinity of the PA subunit (*figure 9C*). One residue under selection pressure, Leu 476, lies within the cap binding of domain of PB2, suggesting that some host-specific characteristics may be involved in the recognition of capped mRNA.

The overall accessibility of the adaptative residues is more obvious in the replication competent form of the polymerase, owing to the outward position of PB2-627K. Indeed, the M589/Q592/K627 cluster is facing the exterior, as Q664, E570, and 718E (*figure 9C*). Two adaptive residues in the extreme C-terminal NLS containing peptide, E718 and V762, are lying on the exposed face of the NLS packed against the PA endonuclease domain in the replication competent polymerase form. The corresponding peptide is not resolved in the transcription competent polymerase, which probably reflects an extended conformation (*figure 9C*). The

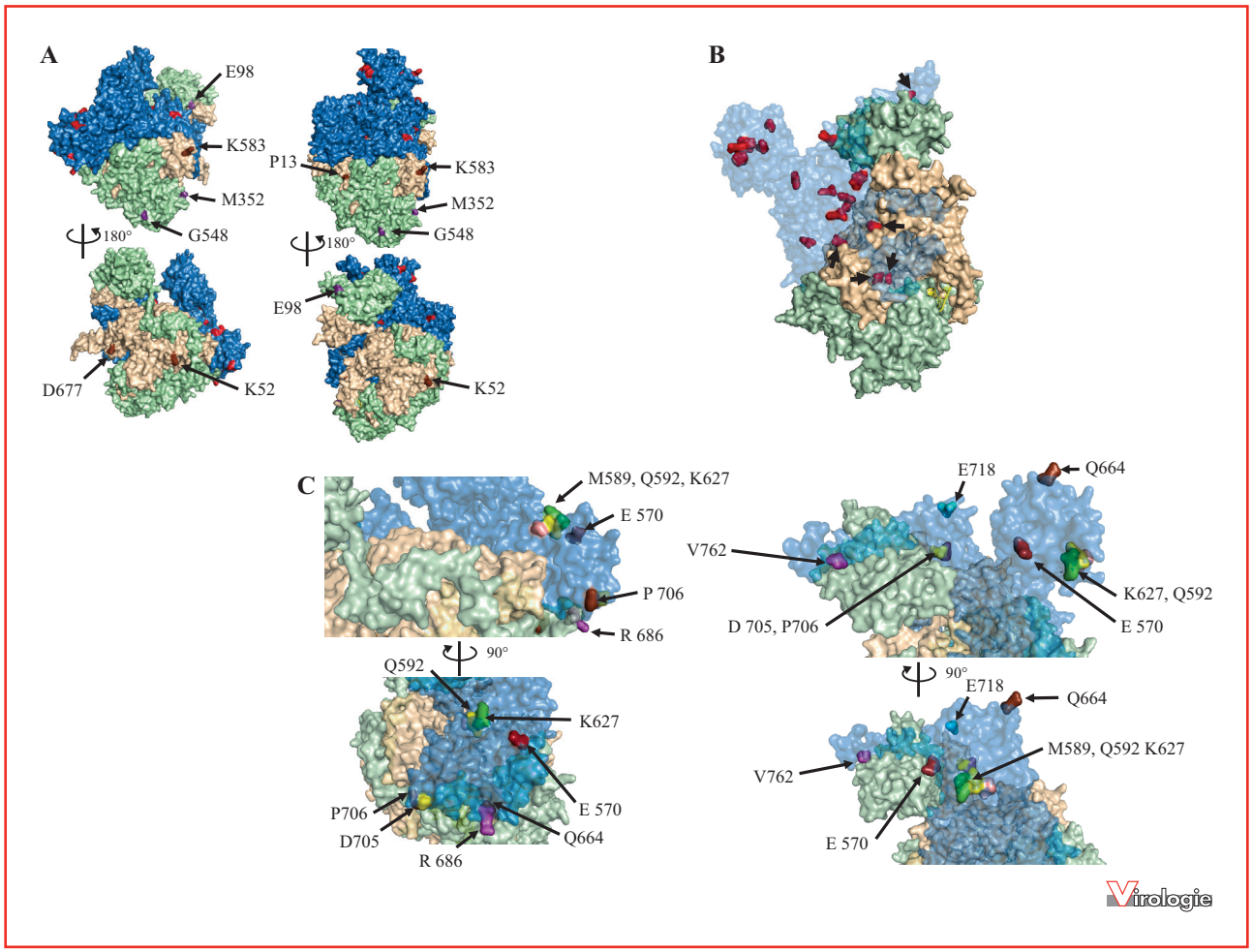


Figure 9. Adaptative Residues in the polymerase **A.** Adaptative residues of influenza polymerase. Residues are shown at the surface of the transcription-competent (pdb 5WSA, left) and 5'cRNA-bound (pdb 5EPI, right) forms of the FluB polymerase. Adaptative residues are colored red in PB2, purple in PA and brown in PB1, positions are given for the PA and PB1 subunits. **B.** Adaptative residues of the PB2 subunit colored in red are shown in a surface representation of the 5'cRNA-bound FluB polymerase, with transparency to highlight residues lying along the interaction interface between PB2 and PB1 or PA (indicated by arrows). **C.** Adaptative resides in the 627/NLS domains of PB2 are colored distinctly, their accessibility is shown in the transcription competent (left), and 5'-cRNA bound (right) forms of the fluB polymerase.

particularly well accessible positions of PB2-627K adaptive residues in the replication form favor the hypothesis that they engage interactions mostly involved in the replication process.

Amino acids important for host factor binding

It is hypothesized that the influenza virus polymerase exploits host cell factors to perform the transcription and replication of vRNA segments in a regulated manner. A series of host factors have been identified as being involved in viral life cycle, through high throughput studies of targeted depletion strategies using si-RNA or sh-RNA [37-40]. On another side, interaction mapping led to the identification of potential host factors interacting with the polymerase [41-44]. However, the overlaps between functional and interaction mapping are limited, and only few of the polymerase/host factors interactions have been mapped precisely. In most cases, it is even not known whether interaction involves the PB1/PA/PB2 trimer or individual subunits. The available structures now provide some insight on the interplay between polymerase and host factors.

Interaction with importins

Upon synthesis in the cytoplasm, the PB2 subunit from one side, and the PB1-PA dimer on another side are directed to the nucleus [11, 45]. Nuclear import of the PB1-PA dimer relies on its binding of the PB1 subunit to RanBP5, involving a protein interface which contain the residues of its bipartite NLS located in the flexible β ribbon extruding from the polymerase complex [46]. In the vRNPs, the same β-ribbon seems to be engaged in an interaction with one promoter-proximal NP monomer, it is therefore likely involved in RanBP5-mediated nuclear import only in the context of the PB1-PA dimer.

Five amino acids have been described as being involved in the interaction of the PB2 C-terminal peptide with α-importins [47]: 701D 737R, 738K, 752K and 755R (FluA) or 705D, 740R, 741K, 756K and 759R (FluB). It has been suggested that the binding of PB2 to importin α is involved in the efficiency of the polymerase activity of 627K-human adapted polymerase, independently of its role in nuclear import of PB2 [48]. In the replication-competent form, the PB2 NLS-containing peptide folds as a long α helix packing on the endonuclease domain of PA, and the residues interacting with importin α are all accessible except 705D (*figure 10A*). However, it has been shown that unfolding of the C-terminal PB2 NLS peptide is required for its efficient binding to importin α [11], and consequently would not be possible in the 5'-cRNA bound polymerase. By contrast, the binding to importin α could probably occur in the transcription-competent polymerase configuration, where the extreme C-terminal NLS peptide of PB2 could be in a flexible extended state as deduced from its lack of electron density, suggesting that importins would impact on the transcription process mediated by the human adapted 627K polymerase rather than on the replication process. The PB2/importin interaction in the promoter-bound polymerase configuration could also take part for the nuclear import of the incoming vRNP, which however has been shown to be mainly mediated by that the NP protein [3]. Interaction of PB2 with importin has been suggested to be part of the host adaptation mechanism, as importin isoforms from human or avian hosts are differentially bound to PB2 from avian or human viruses [49]. This most probably impacts on the efficiency of nuclear translocation of the isolated PB2.

Interaction with the cellular RNA polymerase II and transcription modulator hCLE

The viral influenza polymerase complex has been shown to interact with the C-terminal domain of transcriptionally-active RNA polymerase II (polII) [50]. In the context of vRNP, this is thought to enable the binding to host capped-mRNA required for cap-snatching. A direct interaction with RNA polII has been recently detected, while its binding interface on the influenza polymerase complex is not elucidated [50]. The association of polymerase complex with RNA polII had been proposed earlier to be mediated by the human transcription modulator hCLE [51]. In the influenza virus polymerase complex, two regions of PA have been identified as interaction domains with hCLE [52] and are lying in a pocket close to the vRNA binding site, and are therefore not accessible (*figure 10B*). This does not support the proposed role in recruiting RNA polymerase II to facilitate cap-snatching. The location of PA interface with hCLE suggests an involvement in processes involved in both the replication and the transcription.

The degradation of RNA polymerase II during influenza viral infection was identified as a determinant of pathogenicity, with residues 504 in PB2 and 550 in PA (FluA coordinates) found to be involved in the ability of the virus to induce such degradation [53]. The PA 550 is near the published hCLE interaction interface, in a well-exposed position (*figure 10B*), suggesting a functional link between hCLE binding and induction of RNA pol II degradation. The PB2 504 residue is positioned on the opposite side of the polymerase, in the cap-627 linker proximal to the cap-binding site (*figure 10B*). Given the distance between these two residues, an additional factor could be recruited by the viral polymerase to induce the degradation of RNA polymerase II. Such degradation is associated with an increased ubiquitination of RNA polII, suggesting that a factor of the host Ubiquitin-Proteasome System could be involved [54]. The PB2 504 seems more accessible in the transcription competent polymerase form of FluB polymerase than in 5'c-RNA bound form, where it is lying in a pocket of PB2, whereas it is buried in the apo-conformation.

Host factors interaction from interactomics studies

A number of other host factors have been identified as acting on the viral life cycle *via* an interaction with the polymerase [3]. More recently, high-throughput interaction mappings led to the identification of numerous host factors interacting either with the RNP in an infectious context [41, 43], or with isolated polymerase subunits [38, 42, 44]. Despite limited overlap in the polymerase host partners identified, several cell functions emerged as being targeted by the viral polymerase, among which protein chaperoning, RNA metabolism, and nucleo-cytoplasmic transport. The functional links between these interactions and the activities of the polymerase or of its isolated subunits are however almost not deciphered.

From the currently known structures, one can deduce which binding interfaces are differently exposed according to the polymerase configuration. In the 5' + 3' vRNA-bound polymerase, exposed surfaces are provided by one face of the

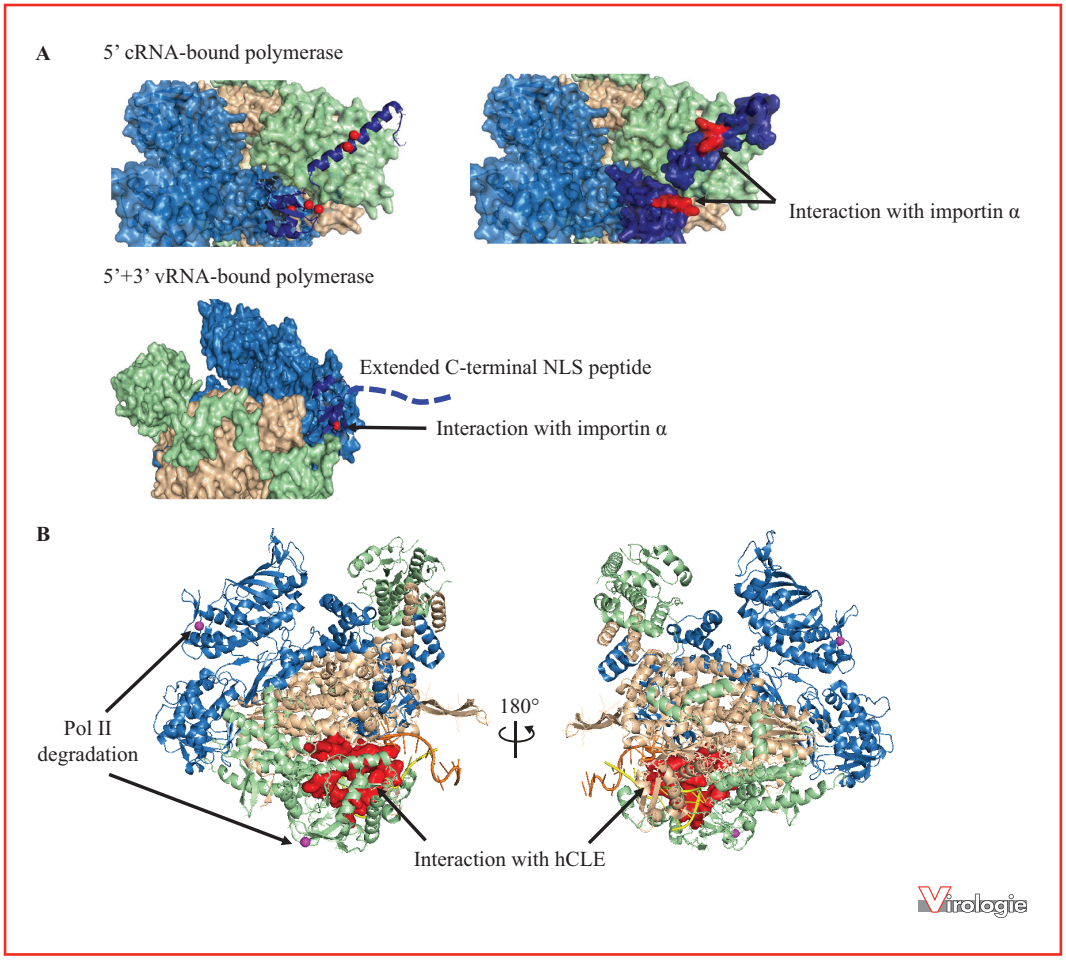


Figure 10. Polymerase interactions with host factors. **A.** Interaction with importin α. Top (Left) Ribbon representation of the NLS domain of 5'cRNA-bound FluB polymerase (pdb 5EPI) colored dark blue, with the rest of the polymerase shown in a surface representation. The residues involved in interaction with importin α are colored red. (right) surface representation of the PB2 NLS in the same polymerase configuration. Bottom-the NLS domain in the promoter-bound FluB conformation (pdb 4WSA), only partly resolved, is shown in ribbon representation, colored dark blue. The extreme C-terminus peptide containing the NLS lacks electron density reflecting, a flexible non-folded state and is represented with a dashed line. **B.** Interaction with hCLE. The interaction interface of PA with the human factor hCLE is shown in surface representation and colored in red in the FluB polymerase associated to 5'+3' vRNA promoter (pdb 4WSA). Residues PB2 504 and PA 550 involved in RNA polII degradation are represented in purple spheres.

Cap binding domain and amino acids patches of PB2-mid (*figure 11*). These surfaces are not accessible because they are packed within the polymerase complex in the 5'cRNA-bound or apo-Flu polymerase configurations (*figure 11*). They consequently could mediate interactions engaged by the polymerase bound to 5' and 3' vRNA ends. One such interaction could be with the rab11 protein, which has been shown to bind the vRNP through PB2 [55]. This interaction is involved in the cytoplasmic transport of the produced vRNP unit, and thus should engage with newly synthetized polymerases bound to 5'+3' vRNA. Similarly, the α helix 6 of PA endonuclease domain provides a well-exposed surface in the polymerase bound to 5'+3' vRNA ends, while being buried in the 5'cRNA-bound configuration owing to its packing against the NLS domain of PB2 (*figure 11*).

The PB2 627 domain is well separated and highly accessible from the rest of the polymerase only in the 5'cRNA-bound polymerase, while in the transcription-competent forms it is packed by contacts with the PB1 and PA subunit in a somewhat buried position along the mRNA exit pathway (*figure 11*). The PB2-627 domain could be an important driver of interactions with host factors instrumental for the replication activity of the polymerase. For example, the ANP32 protein has recently been shown to be involved

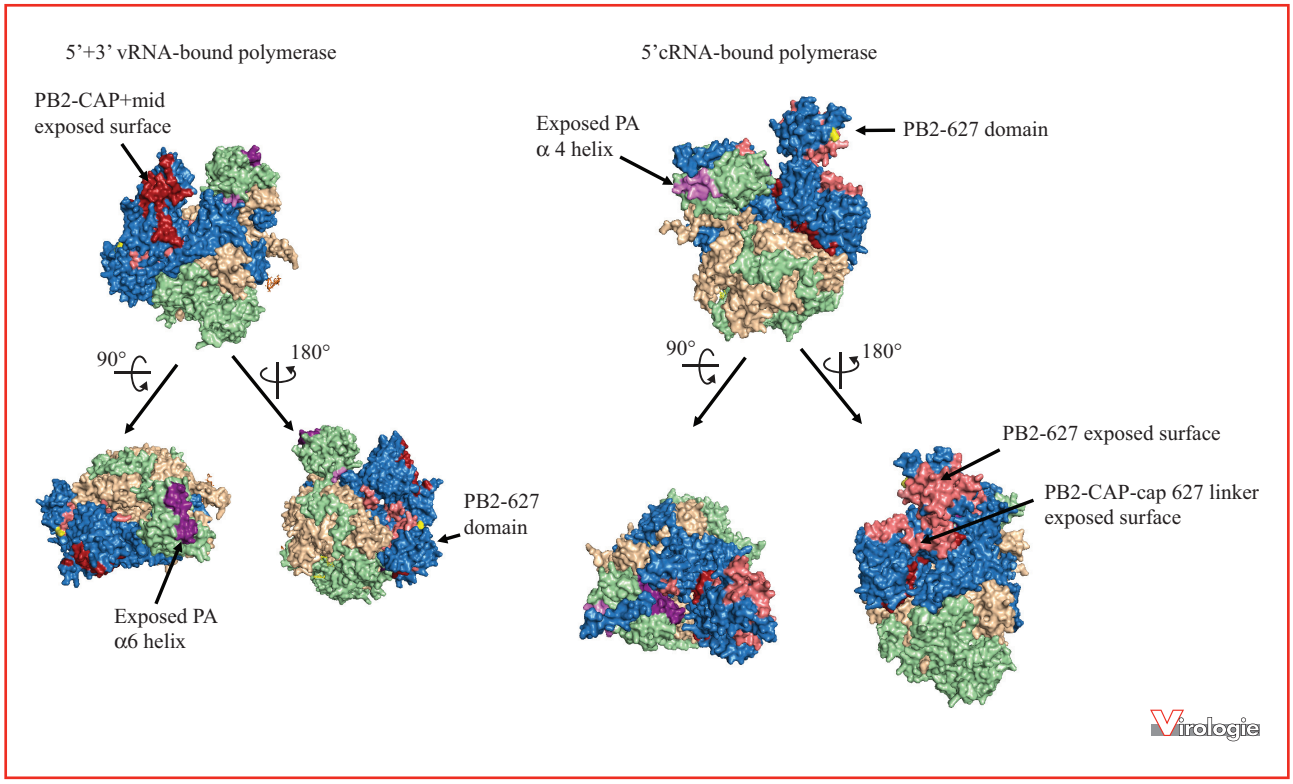


Figure 11. Potential distinctive binding interfaces. Surface representation of the FluB polymerase bound 5' + 3' vRNA (pdb4WSA, right) or to 5'cRNA bound (pdb 5EPI right). Surfaces specifically exposed in the 5' + 3'vRNA polymerase are colored brown (PB2) and purple (PA), surfaces more exposed in the 5'cRNA-bound form are colored salmon (PB2) and pink (PA). The PB2 residue 627K is colored yellow.

in the replication competence of influenza polymerase in human cells, depending upon the 627 residue of PB2 [56]. The role of ANP32 could thus be mediated by an interaction with the PB2 627K domain in the replication-competent polymerase. The fourth α helix on the PA subunit is exposed only in the 5'cRNA-bound polymerase, lying under the NLS C-terminal peptide of PB2, and may represent an interaction interface for factors involved in the replication activity (*figure 11*). One such interaction has been described involving PA and several subunits of the human MCM (minichromosome maintenance element) [57].

In addition, a number of residues are exposed in all forms of polymerase resolved so far, but harbor different locations as a result of the distinctive PB2-C and PA-ENDO arrangements. One can therefore propose that a number of polymerase activity-specific interactions could be mediated through specific patterns of exposed residues patches. In contrast, the exposed surfaces of the polymerase core essentially remain unchanged, and possibly mediate interactions with host factors required for all polymerase activities.

Concluding remarks

The recently characterized influenza virus polymerase structures highlight a number of biological key points. From a general point of view, the influenza polymerase encompasses a central catalytic fold, shared by other RNA-dependent RNA polymerases, which contributes to an emerging picture of the catalytic fold for negative strand RdRps [18].

One remarkable feature of the influenzavirus RdRP is that the flexibly linked domains (PB2-C and PA-ENDO) can adopt stable and distinctive packing relative to a core polymerase, depending upon viral RNA ends binding.

The stable and distinctive packing of that the flexibly linked domains (PB2-C and PA-ENDO) can adopt relative to the constant polymerase core, depending upon viral RNA binding, constitute a remarkable feature of the influenzavirus RdRP. The rules governing such switches still remain to be deciphered. The binding of RNA segments extremities to the polymerase core seems to take an active part in determining polymerase conformations, possibly through

structural changes which somehow transmit to the distal PB2-C and PA-ENDO flexibly linked domains.

One may speculate the association of host factors would assist polymerase conformation remodeling and contribute to stabilize alternative polymerase conformations. This is in line with the different patches of interaction motifs that are exposed according to the type of RNA bound to the polymerase. A simultaneous contribution of viral RNA binding and host-protein interactions is in fact likely to underlie the multiple conformations, hence to the multiple activities, of the polymerase complex. Studies of the polymerase-host interplays revealed a diversified panel of cellular factors associating with influenza virus polymerase, but only few host factors have been implicated in a specific activity of the polymerase. Further studies will be required to identify sets of interacting host factors impacting on specific polymerase activities. The advances in the structure knowledge of influenza virus polymerase represent a powerful framework for such studies.

Acknowledgments. We are grateful to Nadia Naffakh for her helpful comments and suggestions.

Conflicts of interest : none.

References

1. Fodor E. The RNA polymerase of influenza a virus: mechanisms of viral transcription and replication. *Acta Virol* 2013 ; 57 : 113-22.

2. Martin-Benito J, Ortin J. Influenza virus transcription and replication. *Adv Virus Res* 2013 ; 87 : 113-37.

3. Eisfeld AJ, Neumann G, Kawaoka Y. At the centre: influenza A virus ribonucleoproteins. *Nat Rev Microbiol* 2015 ; 13 : 28-41.

4. York A, Fodor E. Biogenesis, assembly, and export of viral messenger ribonucleoproteins in the influenza A virus infected cell. *RNA Biol* 2013 ; 10 : 1274-82.

5. Resa-Infante P, Jorba N, Coloma R, Ortin J. The influenza virus RNA synthesis machine: advances in its structure and function. *RNA Biol* 2011 ; 8 : 207-15.

6. Biswas SK, Nayak DP. Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1. *J Virol* 1994 ; 68 : 1819-26.

7. Fournier G, Chiang C, Munier S, *et al*. Recruitment of RED-SMU1 complex by influenza A virus RNA polymerase to control viral mRNA splicing. *PLoS Pathog* 2014 ; 10 : e1004164.

8. Dias A, Bouvier D, Crepin T, *et al*. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* 2009 ; 458 : 914-8.

9. He X, Zhou J, Bartlam M, *et al*. Crystal structure of the polymerase PA(C)-PB1(N) complex from an avian influenza H5N1 virus. *Nature* 2008 ; 454 : 1123-6.

10. Guilligay D, Tarendeau F, Resa-Infante P, *et al*. The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat Struct Mol Biol* 2008 ; 15 : 500-6.

11. Tarendeau F, Boudet J, Guilligay D, *et al*. Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nat Struct Mol Biol* 2007 ; 14 : 229-33.

12. Tarendeau F, Crepin T, Guilligay D, Ruigrok RW, Cusack S, Hart DJ. Host determinant residue lysine 627 lies on the surface of a discrete, folded domain of influenza virus polymerase PB2 subunit. *PLoS Pathog* 2008 ; 4 : e1000136.

13. Sugiyama K, Obayashi E, Kawaguchi A, *et al*. Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. *EMBO J* 2009 ; 28 : 1803-11.

14. Hengrung N, El Omari K, Serna Martin I, *et al*. Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. *Nature* 2015 ; 527 : 114-7.

15. Reich S, Guilligay D, Pflug A, *et al*. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* 2014 ; 516 : 361-6.

16. Pflug A, Guilligay D, Reich S, Cusack S. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* 2014 ; 516 : 355-60.

17. Thierry E, Guilligay D, Kosinski J, *et al*. Influenza polymerase can adopt an alternative configuration involving a radical repacking of PB2 domains. *Mol Cell* 2016 ; 61 : 125-37.

18. Reguera J, Gerlach P, Cusack S. Towards a structural understanding of RNA synthesis by negative strand RNA viral polymerases. *Curr Opin Struct Biol* 2016 ; 36 : 75-84.

19. O'Reilly EK, Kao CC. Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. *Virology* 1998 ; 252 : 287-303.

20. Te Velthuis AJ, Robb NC, Kapanidis AN, Fodor E. The role of the priming loop in Influenza A virus RNA synthesis. *Nat Microbiol* 2016 ; 1 : 1-7.

21. Obayashi E, Yoshida H, Kawai F, *et al*. The structural basis for an essential subunit interaction in influenza virus RNA polymerase. *Nature* 2008 ; 454 : 1127-31.

22. Pritlove DC, Poon LL, Devenish LJ, Leahy MB, Brownlee GG. A hairpin loop at the 5' end of influenza A virus virion RNA is required for synthesis of poly(A)+ mRNA *in vitro*. *J Virol* 1999 ; 73 : 2109-14.

23. Gong P, Peersen OB. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc Natl Acad Sci U S A* 2010 ; 107 : 22505-10.

24. Poon LL, Pritlove DC, Fodor E, Brownlee GG. Direct evidence that the poly(A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template. *J Virol* 1999 ; 73 : 3473-6.

25. Tomescu AI, Robb NC, Hengrung N, Fodor E, Kapanidis AN. Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. *Proc Natl Acad Sci U S A* 2014 ; 111 : E3335-42.

26. Chang S, Sun D, Liang H, *et al*. Cryo-EM structure of influenza virus RNA polymerase complex at 4.3 A resolution. *Mol Cell* 2015 ; 57 : 925-35.

27. Jorba N, Coloma R, Ortin J. Genetic trans-complementation establishes a new model for influenza virus RNA transcription and replication. *PLoS Pathog* 2009 ; 5 : e1000462.

28. Mena I, Jambrina E, Albo C, *et al*. Mutational analysis of influenza A virus nucleoprotein: identification of mutations that affect RNA replication. *J Virol* 1999 ; 73 : 1186-94.

29. Coloma R, Valpuesta JM, Arranz R, Carrascosa JL, Ortin J, Martin-Benito J. The structure of a biologically active influenza virus ribonucleoprotein complex. *PLoS Pathog* 2009 ; 5 : e1000491.

30. Martin-Benito J, Area E, Ortega J, *et al*. Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO Rep* 2001 ; 2 : 313-7.

31. Biswas SK, Boutz PL, Nayak DP. Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. *J Virol* 1998 ; 72 : 5493-501.

32. Ng AK, Chan WH, Choi ST, *et al*. Influenza polymerase activity correlates with the strength of interaction between nucleoprotein and PB2 through the host-specific residue K/E627. *PLoS One* 2012 ; 7 : e36415.

33. Poole E, Elton D, Medcalf L, Digard P. Functional domains of the influenza A virus PB2 protein: identification of NP- and PB1-binding sites. *Virology* 2004 ; 321 : 120-33.

34. Cauldwell AV, Moncorge O, Barclay WS. Unstable polymerase-nucleoprotein interaction is not responsible for avian influenza virus polymerase restriction in human cells. *J Virol* 2013 ; 87 : 1278-84.

35. Tamuri AU, Dos Reis M, Hay AJ, Goldstein RA. Identifying changes in selective constraints: host shifts in influenza. *PLoS Comput Biol* 2009 ; 5 : e1000564.

36. Taft AS, Ozawa M, Fitch A, *et al*. Identification of mammalian-adapting mutations in the polymerase complex of an avian H5N1 influenza virus. *Nat Commun* 2015 ; 6 : 7491.

37. Karlas A, Machuy N, Shin Y, *et al*. Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* 2010 ; 463 : 818-22.

38. Shapira SD, Gat-Viks I, Shum BO, *et al*. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* 2009 ; 139 : 1255-67.

39. Konig R, Stertz S, Zhou Y, *et al*. Human host factors required for influenza virus replication. *Nature* 2010 ; 463 : 813-7.

40. Watanabe T, Watanabe S, Kawaoka Y. Cellular networks involved in the influenza virus life cycle. *Cell Host Microbe* 2010 ; 7 : 427-39.

41. York A, Hutchinson EC, Fodor E. Interactome analysis of the influenza A virus transcription/replication machinery identifies protein phosphatase 6 as a cellular factor required for efficient virus replication. *J Virol* 2014 ; 88 : 13284-99.

42. Watanabe T, Kawakami E, Shoemaker JE, *et al*. Influenza virus-host interactome screen as a platform for antiviral drug development. *Cell Host Microbe* 2014 ; 16 : 795-805.

43. Munier S, Rolland T, Diot C, Jacob Y, Naffakh N. Exploration of binary virus-host interactions using an infectious protein complementation assay. *Mol Cell Proteomics* 2013 ; 12 : 2845-55.

44. Bradel-Tretheway BG, Mattiacio JL, Krasnoselsky A, *et al*. Comprehensive proteomic analysis of influenza virus polymerase complex reveals a novel association with mitochondrial proteins and RNA polymerase accessory factors. *J Virol* 2011 ; 85 : 8569-81.

45. Fodor E, Smith M. The PA subunit is required for efficient nuclear accumulation of the PB1 subunit of the influenza A virus RNA polymerase complex. *J Virol* 2004 ; 78 : 9144-53.

46. Swale C, Monod A, Tengo L, *et al*. Structural characterization of recombinant IAV polymerase reveals a stable complex between viral PA-PB1 heterodimer and host RanBP5. *Sci Rep* 2016 ; 6 : 24727.

47. Resa-Infante P, Jorba N, Zamarreno N, Fernandez Y, Juarez S, Ortin J. The host-dependent interaction of alpha-importins with influenza PB2 polymerase subunit is required for virus RNA replication. *PLoS One* 2008 ; 3 : e3904.

48. Hudjetz B, Gabriel G. Human-like PB2 627K influenza virus polymerase activity is regulated by importin-alpha1 and -alpha7. *PLoS Pathog* 2012 ; 8 : e1002488.

49. Gabriel G, Klingel K, Otte A, *et al*. Differential use of importin-alpha isoforms governs cell tropism and host adaptation of influenza virus. *Nat Commun* 2011 ; 2 : 156.

50. Martinez-Alonso M, Hengrung N, Fodor E. RNA-free and ribonucleoprotein-associated influenza virus polymerases directly bind the serine-5-phosphorylated Carboxyl-terminal domain of host RNA polymerase II. *J Virol* 2016 ; 90 : 6014-21.

51. Rodriguez A, Perez-Gonzalez A, Nieto A. Cellular human CLE/C14orf166 protein interacts with influenza virus polymerase and is required for viral replication. *J Virol* 2011 ; 85 : 12062-6.

52. Rodriguez-Frandsen A, de Lucas S, Perez-Gonzalez A, *et al*. hCLE/C14orf166, a cellular protein required for viral replication, is incorporated into influenza virus particles. *Sci Rep* 2016 ; 6 : 20744.

53. Llompart CM, Nieto A, Rodriguez-Frandsen A. Specific residues of PB2 and PA influenza virus polymerase subunits confer the ability for RNA polymerase II degradation and virus pathogenicity in mice. *J Virol* 2014 ; 88 : 3455-63.

54. Vreede FT, Chan AY, Sharps J, Fodor E. Mechanisms and functional implications of the degradation of host RNA polymerase II in influenza virus infected cells. *Virology* 2010 ; 396 : 125-34.

55. Amorim MJ, Bruce EA, Read EK, *et al*. A Rab11- and microtubule-dependent mechanism for cytoplasmic transport of influenza A virus viral RNA. *J Virol* 2011 ; 85 : 4143-56.

56. Long JS, Giotis ES, Moncorge O, *et al*. Species difference in ANP32A underlies influenza A virus polymerase host restriction. *Nature* 2016 ; 529 : 101-4.

57. Kawaguchi A, Nagata K. *De novo* replication of the influenza virus RNA genome is regulated by DNA replicative helicase, MCM. *EMBO J* 2007 ; 26 : 4566-75.

Abstract

An estimated 10%-20% of the world's population is affected each year by seasonal epidemic influenza, causing about 250,000 to 500,000 fatal cases. The pandemic risk reinforces the trait of influenza A virus (IAV) infection as a public health issue. The virus life cycle critically relies on its ability to manipulate the host proteome. Besides, the ubiquitin-proteasome system (UPS) is involved in many regulatory processes in mammalian cells by inducing protein degradation, mediating protein activation or shaping their sub-cellular localisation. Therefore, UPS is a prime target hijacked by viruses. Recent evidence indicates that an intricate regulatory network involving viral proteins and the cellular UPS is likely to contribute to viral replication and immune evasion of influenza A viruses. However, usurpation of the host UPS by IAV is far from being comprehensively deciphered.

To gain better understanding, we assessed the interplay between the human UPS and the PB2 subunit of the influenza A virus polymerase through a global proteomic profiling approach. For that purpose, an UPS-dedicated library of 590 human cDNAs, comprising 63% of the whole human UPS, was constituted and characterised. In an initial screening, UPS factors were challenged using a high-throughput split luciferase assay for interaction with the PB2 protein from 5 influenza A strains of different pathogenicity in human. A total of 80 UPS factors emerged as potential PB2 partners, of which 42 were validated as high-confidence PB2 partners for at least one of the strains. Further comparison of interaction profiles of the 5 PB2 with the UPS by hierarchical clustering revealed an interaction dendrogram fitting with the circulation time in the human population.

Functional importance of interactors was tested by siRNA-mediated knock down experiments using luciferase tagged recombinant IAV viruses. Depletion of 36 out of the 42 tested UPS factors showed an effect on the infection with all or a subset of IAV strains, underlying the strong functional output of the developed methodology. Among these factors three deubiquitinases (DUBs) were further studied to decipher their involvement in IAV viral cycle. We have shown that they are involved in early and late stages of the infection and began to draw their function in viral cycle. We demonstrated with our colleagues in Hong-Kong that OTUB1 is involved in the host cytokine response and most probably in virus assembly. OTUD6A was also shown to be implicated in late stages of the infection but we still don't know its exact role. Contrariwise, the inactive DUB PAN2, which is part of poly-adenylation complexes, is implicated in the early phase of IAV infection, but surprisingly apparently not through viral mRNA regulation. More work is on going to specify the mechanisms of these DUBs implication in IAV infection.

Keywords: Influenza A viruses, polymerase, PB2 protein, Ubiquitin Proteasome System, deubiquitinases, comparative interactomics, protein complementation assay, host-pathogen interactions.