Università degli Studi di Padova

Dipartimento di Scienze Biomediche

CORSO DI DOTTORATO DI RICERCA IN SCIENZE BIOMEDICHE

31° CICLO

MOLECULAR CHARACTERIZATION OF VHL FAMILY PROTEINS: WHEN SIMILARITIES MAKE THE DIFFERENCE

Coordinatore: Ch.mo Prof. Paolo Bernardi

Supervisore: Ch.mo Prof. Silvio C. E. Tosatto

Dottoranda: Raissa Bortolotto

Index

Figures Index

Supplementary Figures Index

Table Index

Supplementary Table Index

Abbreviations

2OG	2-oxoglutarate
3AT	3-Amino-1,2,4-Triazol
AD	activation domain
ADA2	adenosine deamidase 2
ADE2	phosphoribosylaminoimidazole carboxylase
ADGRG1	adhesion G protein-coupled receptor G1
ADH1	alcohol dehydrogenase ADH1 (promoter)
AKT	protein kinase B
AMP	ampicillin
APC	Adenomatus Polyposis Coli
ARM	armadillo domain
ARNT	aryl hydrocarbon receptor nuclear translocator
ATP	adenosine triphosphate
AUR1-C	inositol phosphoryl ceramide synthase
AURKA	Aurora kinase A
BGH	bovine growth hormone
bHLH	basic helix-loop-helix
BLAST	Basic Local Alignment Search Tool
Bluues	Electrostatic properties of proteins based on generalized Born radii
BRCA1	Breast Cancer Type 1 susceptibility protein
BSA	bovine serum albumin
C-	negative control

C+	positive control
Ca^{2+}	calcium
CBP	CREB-binding protein
ccRCC	clear cell renal cell carcinoma
CDI	cyclin-dependent kinase inhibitor (domain)
Cdk	Cyclin Dependent Kinase
CDKN1A	Cyclin Dependent Kinase Inhibitor 1A (or p21)
CDKN1B	Cyclin Dependent Kinase Inhibitor 1B (or p27)
CDKN1C	Cyclin Dependent Kinase Inhibitor 1C (or p57)
CDKN2A	cyclin-dependent kinase Inhibitor 2A
cfu	colony forming unit
CHAPS	3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate hydrate
CKI	Cyclin Dependent Kinase Inhibitor
CMV	cytomegalovirus
CNS	central nervous system
CO_2	Carbon dioxide
CODD	C-terminal oxygen-dependent degradation
Co-IP	Co-Immunoprecipitation
CREB	cAMP response element-binding protein
CT	carboxyl terminal region
C-TAD	C-terminal transactivation domain
Cul2	Cullin 2
DBD	DNA-binding domain
ddH_2O	Double-distilled water

DMEM	Dulbecco’s modified Eagle’s medium
DNA	Deoxyribonucleic acid
dsDNA	double-stranded DNA
DTT	dithiothreitol
E2F1	Transcription Factor 1
EGFR	Epidermal growth factor receptor
EloB	Transcription elongation factors B or Elongin B
EloC	Transcription elongation factors C or Elongin C
ELST	Endolymphatic sac tumor of the inner ear
EPO	Erythropoietin
ERK	Extracellular signal–regulated kinases
FBS	Fetal bovine serum
Fe	Iron
Fe^{2+}	Ferrous
FELLS	Fast Estimator of Latent Local Structure
FIH-1	Factor inhibiting HIF-1
FLAG-tag	Sequence motif DYKDDDDK
$G_{1\text{-}2}$ phase	Gap phase 1-2
Gal4	Transcription factor Gal4
GC-rich	Guanine-cytosine rich
GFP	Green fluorescence protein
Glu	Glutamic acid
GLUT1	Glucose transporter 1

GO	Gene Ontology
H1-4	α-helix 1-4
HA	Hemagglutinin
HB	Hemangioblastoma
HEK293T	Human embryonic kidney cells 293
HeLa	Henrietta Lacks (donor of the cell)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF-(1-3)α	Hypoxia-inducible factor (1-3) alpha
HIF-α	Hypoxia-inducible factor, α-subunit
HIF-β	Hypoxia-inducible factor, β-subunit
His	Histidine
HIS3	Imidazole glycerol-phosphate dehydratase 3
hnRNPA2B1	Heterogeneous nuclear ribonucleoprotein A2B1
HREs	Hypoxia-response elements
hsRPB7	Human RNA polymerase II seventh subunit
IDP	Intrinsically disordered proteins
IDR	Intrinsically disordered region
IGF2	Insulin-like growth factor 2
INK4	Cyclin-dependent kinase inhibitors
IP	Immunoprecipitate
IRAK2	Interleukin 1 receptor-associated kinase 2
KAN	Kanamycin
Ki67	Antigen KI-67

KID	Kinase inhibitory domain
LB	Luria-Bertani
LiAc	Acetate lithium
LM	Linear motif
LPP	Lipoma-preferred partner
Lys	Lysine
MAPK	MAP kinase
MCS	Multi-cloning site
MD	Molecular dynamic
MDM2	Mouse double minute 2 homolog
MEL1	Yeast α-galactosidase
MMP13	Matrix metalloprotease protein 13
M-phase	Mitosis phase
mRNA	Messenger RNA
NaCl	Sodium chloride
NeEMO	NEtwork Enthalpic MOdelling
NFkB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NLS	Nuclear localization signal
NODD	N-terminal oxygen-dependent degradation domain
NT	Amino-terminal region
N-TAD	N-terminal transactivation domain
O/N	Overnight

OD	Optical density
ODD	Oxygen-dependent degradation domain
p14ARF	Tumor suppressor protein 14 (or ARF)
p21-NT	p21 N-terminal region
p21-ΔN	p21 C-terminal region
p27-NT	p27 N-terminal region
p27-ΔN	p27 C-terminal region
TP53	p53 or tumor protein 53
p57-NT	p57 N-terminal region
p57-ΔN	p57 C-terminal region
PAPA	Region proline-alanine repeats region
PAS	Per-ARNT-Sim
PBS	Phosphate-buffered saline
PCC	Pheochromocytomas
PCNA	Proliferative cell nuclear antigen binding domain
PDB	Protein data bank
PDGF	Platelet-derived growth factor
PEG 4000	Polyethylene glycol 4000
PGL	Paraganglioma
PHD (1-3)	Prolyl-4-hydroxylase domain (1-3)
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
PIC	Protease inhibitors cocktail
PKM	Pyruvate kinase
PNS	Post-nuclear supernatant

Pol II	RNA polymerase II
PPI	Protein-protein interaction
PTEN	Phosphatase and TENsin homolog
PTM	Post-translational modification
QDO	Quadruple drop out
QMEAN	Qualitative Model Energy Analysis
Rb	Retinoblastoma
RBX1	RING-box protein1
RMSF	Root-mean-square fluctuation
RNA	Ribonucleic acid
ROS	Reactive oxygen species
Rpb1	RNA-Directed RNA Polymerase II Subunit
rpm	revolutions per minute
RT	Room temperature
SCF	Skp, Cullin, F-box containing complex
SD	Selective medium
SDS	Sodium dodecyl sulfate
SLC9A3	Solute Carrier Family 9 Member A3
SLIM	Short linear motif
SOCS	Box suppressor of cytokine signaling box
S-phase	Synthesis phase
STAT3	Signal transducer and activator of transcription 3
SV40	Simian virus 40

TAD	Trans-activating domain
TATA-box	Goldberg-Hogness box, thymine and adenine rich
TBS-T	Tris-buffered saline tween
TCA	Trichloroacetic acid
TGFB	Transforming growth factor beta
TGS	Tumor-suppressor gene
Trp	Tryptophan
TSS	Translation start site
Tyr	Tyrosine
Ub	Ubiquitin
UbH	Ubiquitin-homology
Ura	Uracil
UTR	Untranslated region
VCB	Complex (pVHL-elongin C-elongin B)-Cul-2 (Cullin-2)
VCB-CR	VHL-EloginC-EloginB complex and Cullin2-RBX1 complex
VEGF	Vascular Endothelial Growth Factor
VHL	von Hippel-Lindau
VHL172	pVHL172 or von Hippel-Lindau 172, isoform 172
VHL172ΔNT	pVHL172-ΔNT or von Hippel-Lindau protein, isoform 172 w/o N-terminal region
VHL19	pVHL19 or von Hippel-Lindau 19, shorter isoform 19
VHL30	pVHL or VHL von Hippel-Lindau 30, full-length isoform 30
VHLdb	von Hippel-Lindau database
VHL-L	VLP or von Hippel-Lindau like protein (paralog)
VHLα	von Hippel-Lindau α (isoform)

VHLβ	pVHLβ or von Hippel-Lindau protein β domain
w/o	without
Y2H	Yeast Two-Hybrid assay
Y2HGold	Gold Yeast two-hybrid strain
YDP	Yeast peptone dextrose
ZnF	Zinc finger
ΔNT	Amino-terminal region deletion

Summary

Mutations of the so-called *cancer-susceptibility genes* impair the biological function of key factors[1] as observed for the human von Hippel-Lindau (VHL) oncosuppressor gene. Individuals carrying loss-of-function mutations on the VHL gene develop a familial autosomal-dominant cancer predisposition referred as von Hippel-Lindau syndrome (VHL, OMIM 193300), characterized by benign or malign tumors and cysts diffused in patients organs[2]. The best-known pVHL30 function in healthy cells is its role in regulating the hypoxia-inducible transcription factor 1α (HIF-1α), which in turn mediates the adaptive response to hypoxia. This mechanism would sustain the cancer development whether altered. Although VHL syndrome is classified as an inherited disease, the genotype-phenotype correlation is not fully explained. Indeed, pVHL is classified as a multipurpose hub protein able to interact with more than 500 proteins[3] participating different cellular functions. The protein interacting network centered on pVHL has been deeply investigated over the years, in order to clarify the functional implications of these interactions. My PhD program was aimed to identify and characterize novel pVHL interactors through a combination of computational and experimental approaches. I identified a novel pVHL interaction with the cyclin-dependent inhibitor kinase 1 subfamily (CDKN1s). Their association resembles the pVHL/HIF-1α complex formation by involving both the pVHL30 β-domain and a CODD-like motif, which was found to localize within the linear CDKN1s N-terminal domain. The pVHL30 plasticity has been further investigated to include the pVHL172 isoform and VLP, the only known human pVHL paralog. In particular, interactions with HIF-1α, CDKN1s and MDM2, a novel pVHL interactor identified in a parallel work at my host laboratory, have been evaluated revealing isoform-specific bindings. Data generated during my PhD suggest a role for the pVHL disordered N-terminus in discriminating isoform-specific interactions. Confocal microscopy experiments aimed at characterizing the biochemical properties of pVHL-related proteins were then performed. Resulting data indicated pVHL30 and pVHL172 to assume a diffuse distribution in mammal cells, whereas both pVHL19 and VLP, which lack (or partially lack) of an extended N-terminus appear to localize in cytosolic punctate structures. The same behavior was also observed for the truncated mutant pVHL172-ΔNT, which indirectly supports a role for this accessory tail in regulating pVHL localization. Biochemical fractionation assays show that members of VHL-family presenting with the N-terminus are mostly recovered in the soluble fraction. Instead, N-terminus loss markedly increases the insoluble protein fraction. Collectively taken, these findings suggest the existence of distinct functional role for each member of the pVHL protein family, revealed in this work as isoforms-specific protein-protein association and intracellular localization. Intriguingly, the novel interactions and biochemical evidences were also supported by validation in

human cell. In light of these meaningful findings, I performed a tissue-specific genome-wide cDNA library screening to investigate the unknown VLP interactome. A total of 96 interacting fragments, each corresponding to a different protein, were collected and sequenced. A preliminary computational analysis revealed that VLP binding partners mostly localize in nucleus and cytosol, while participating three major functional pathways (i.e. regulation of extracellular adhesion, cytoskeleton remodeling, as well as cell survival and proliferation). In addition, a number of these new interactors are characterized by functional domains typically found in scaffold proteins, such as WD, LIM, ARM and zinc finger, thus suggesting a VLP role in participating multiple protein complexes. More in general, considering both the novelty of this tissue-specific library, ubiquitous pVHL-expression and their structural similarity, it can be speculated that at least a subset of these new VLP interactions can be also extended to pVHL, possibly representing an example of tissue-specific pVHL-partners. Their experimental validation, for example using VHL30 as *bait* in a similar environment, will be crucial to address this hypothesis. On the other hand, the VLP-specific proteome may derive by gene specialization evolved to account for tissue-specific requirements. If confirmed, these data may help to explain how functional impairment of the VHL-protein family promotes cancer insurgence.

Riassunto

Le mutazioni che colpiscono i cosiddetti "geni di suscettibilità al cancro" alterano la funzione biologica di alcune proteine chiave, come nel caso del gene oncosoppressore von Hippel-Lindau (*VHL*). Individui portatori di mutazioni *loss-of-function* sul gene *VHL* sviluppano una predisposizione famigliare autosomica dominante al cancro, chiamata sindrome di von Hippel-Lindau (VHL, OMIM 193300). La patologia è caratterizzata dalla diffusione di tumori o cisti, benigni o maligni, in differenti organi[2]. Nelle cellule sane, la funzione maggiormente nota di pVHL30 riguarda la regolazione del fattore di trascrizione HIF-1α, il quale a sua volta media e regola la risposta adattiva all'ipossia. Se alterato, questo meccanismo di regolazione può sostenere lo sviluppo tumorale. Nonostante la sindrome VHL sia classificata come famigliare predisposizione al cancro, la correlazione tra genotipo e fenotipo non risulta univoca nei pazienti. Ciò potrebbe trovare spiegazione nelle caratteristiche intrinseche della proteina in questione. pVHL30 viene infatti classificata come *hub* multifunzionale capace di interagire con più di 500 diversi interattori[3], partecipando così a differenti funzioni cellulari. La fitta rete di interazioni proteiche attorno a pVHL30 è stata ampiamente studiata negli anni, con lo scopo di chiarirne le implicazioni funzionali. Per questo, durante il mio percorso di dottorato, ho cercato di caratterizzare nuovi interattori di pVHL30 combinando approcci computazionali e sperimentali. Dagli esperimenti, è emersa una nuova associazione tra pVHL30 e la famiglia degli inibitori delle chinasi ciclina-dipendenti 1 (CDKN1), strutturalmente simile al legame tra pVHL30 ed HIF-1α. L'interazione infatti coinvolge sia il dominio β di pVHL30 sia la regione N-terminale delle proteine CDKN1, nella quale è contenuto il dominio CDI caratterizzato da un motivo *CODD-like*. Inoltre, la plasticità di pVHL30 è stata studiata sfruttando sia pVHL172, isoforma alternativa del gene *VHL*, sia la proteina paralogo VLP. In particolare, sono state valutate le interazione con HIF-1α, CDKN1 e MDM2, quest'ultima identificata come nuovo interattore di pVHL30 in un lavoro parallelo non ancora pubblicato. Le diverse interazioni hanno evidenziato una tendenza isoforma-specifica, suggerendo per la porzione N-terminale un ruolo di regione discriminante nella scelta degli interattori. Esperimenti di microscopia confocale in cellule di mammifero mostrano inoltre una distribuzione diffusa di entrambe le isoforme, pVHL30 e pVHL172. Diversamente l'isoforma pVHL19 e il paralogo VLP, che mancano interamente (o parzialmente) della regione N-terminale, sono localizzate nel citosol in strutture definite *punctate*. Lo stesso viene osservato con il mutante tronco dell'isoforma VHL172 (pVHL172-ΔNT), evidenziando indirettamente che la localizzazione delle strutture VHL è regolata dalla regione N-terminale. Mediante frazionamento biochimico, è stato osservato che le isoforme caratterizzate dalla coda N-terminale completa, sono maggiormente presenti della frazione cellulare solubile. Al contrario, l'assenza della regione N-terminale aumenta

marcatamente la presenza delle proteine nella frazione insolubile. Complessivamente, queste evidenze isoforma-specifiche suggeriscono un possibile ruolo funzionale distinto per ciascun membro della famiglia delle proteine VHL. In altre parole, queste strutture alternative di pVHL30 potrebbero essere frutto dell'evoluzione per sostenere funzioni tessuto-specifiche o competere per il legame con specifici *partners*. Entrambe le situazioni evidenziano ruoli biologici diversi al fine di regolare multiple o alternative vie di funzione nella cellula[4]. Alla luce di questi risultati significativi, è stato eseguito uno screening di libreria di cDNA tessuto-specifico in lievito per approfondire l'interattoma del paralogo VLP. Sono stati identificati e sequenziati 96 frammenti interagenti, ciascuno corrispondente ad una proteina diversa. Un'analisi computazionale preliminare ha rivelato che i *partners* di VLP sono principalmente localizzati nel nucleo e nel citosol, partecipando a tre principali funzioni cellulari (come regolazione dell'adesione extracellulare, rimodellamento del citoscheletro, nonché sopravvivenza e proliferazione cellulare). Inoltre, alcuni di questi interattori risultano caratterizzati da domini funzionali (come WD, LIM, ARM e *zinc finger*) comunemente presenti nelle proteine di *scaffold*, proponendo la possibile partecipazione di VLP in differenti complessi proteici. Più in generale, considerando la novità dei risultati ottenuti attraverso lo screening, la somiglianza strutturale con pVHL30 e l'espressione ubiquitaria di pVHL30, si può ipotizzare che queste nuove interazioni di VLP possano rappresentare un sottoinsieme di nuovi interattori di pVHL30. In altre parole, essi rappresentano un esempio di partner di pVHL30 tessuto-specifici. Sperimentalmente potrebbe essere possibile validare questa ipotesi eseguendo lo stesso screening utilizzando pVHL30 come esca (*bait*) del sistema doppio ibrido. D'altra parte, questo sottoinsieme di interattori specifici per VLP potrebbe essere frutto di un'evoluzione genica per rispondere alle esigenze specifiche del tessuto in questione. Se confermati, questi dati potrebbero aiutare a spiegare come l'alterazione funzionale della famiglia delle proteine VHL promuova l'insorgenza del cancro.

Introduction

Introduction

1.1 Cancer as a multifaced pathology

Cancer is a complex and multistep pathology, the second leading cause of death in the world[5]. Remarkable progress in cancer field highlighted most of the human cancers to share specific hallmarks[6]. These include sustained proliferative signaling, growth suppressors elusion, tissues invasion and metastasis cell death resistance, replicative immortality and induced angiogenesis through hypoxia[6]. More recently, this list has been updated with other biological capabilities, typically worsened in tumors. They consist in altered cellular metabolism, evasion of immune system, genomic instability and altered inflammation response (Figure 1)[7].

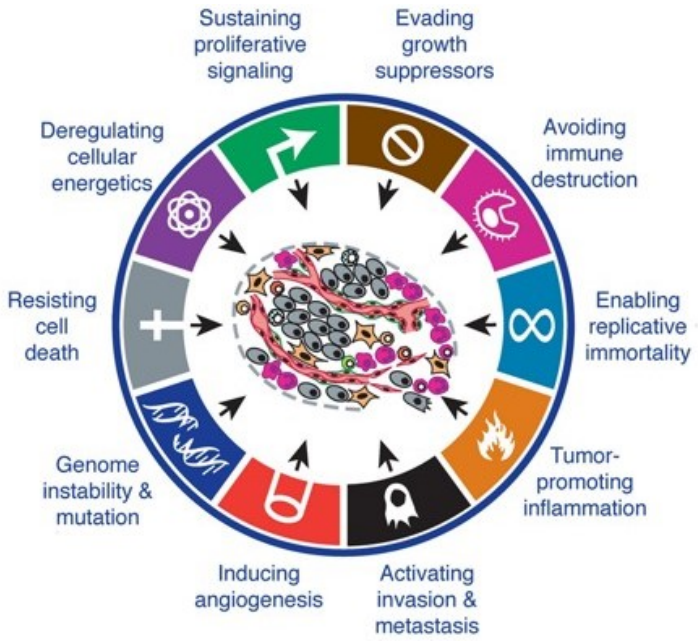


Figure 1. Cancer hallmarks. The original six cancer hallmarks[6] are updated including the capability to modify cellular metabolism, evade immunological destruction, genomic instability and inflammation responses. Modified from *Fouard, Y.A. & Aanei, C.*[1].

Although these non-genetic factors are recruited during tumors development, it is broadly accepted that genetic alterations are the fuse of this multi-step mechanism[8]. Cancer arises from mutations of multiple cancer-susceptibility genes, which are commonly grouped in three different classes: gatekeepers, caretakers[9] and landscapers genes[10]. Gatekeepers directly control pathways regulating cell growth and differentiation and comprise both oncogenes and tumor-suppressor genes (TGSs). Caretakers promote tumorigenesis indirectly leading to genetic instability with rapid accumulation of

changes in other genes, whereas landscapers generate alteration in stroma environment that contributes to neoplastic cell transformation[10]. Thus, cancer genetics has become intrinsic to cancer risk assessment, an essential component of the practice in preventive oncology. This feature increased the awareness of possible hereditary cancer syndrome spreading, often confused with cancer manifesting an apparent sporadic development. As an example of this concept, the hereditary retinoblastoma (Rb) is a paradigmatic syndrome for the study of inherited cancer susceptibility[11]. Nowadays, several genes are known to be associated with familiar cancer syndromes, such as BRCA1, PTEN, APC and p53[12]. Germline mutation in the *TP53* (p53) gene have been identified in 70% of families characterized by various form of cancers, later described as typical criteria of the hereditary Li-Fraumeni Syndrome[13]. Similarly, mutations of the von Hippel-Lindau protein (pVHL) is associated with a familiar increased cancer risk paired to alteration of hypoxia regulation and sustained angiogenesis.

1.2 The von Hippel-Lindau disease

The von Hippel–Lindau (VHL, OMIM 193300) disease is a rare, hereditary, autosomal dominant cancer disorder that affects approximately 1/36 000 people with an onset at a mean age of 26 years[14]. Historically, the German ophthalmologist Eugene von Hippel described some cases of familiar retinal angiomatosis in 1904[15]. Following these preliminary descriptions, the Swedish pathologist Arvid Lindau established the association between retinal angiomas and cerebellar hemangioblastoma (HB) in 1927[16]. The term von Hippel–Lindau disease was first used in 1936 and has been in common use since the 1970s. The identification of the human *VHL* gene, mapped in chromosome 3 in 1993, has been a remarkable progress leading to the understanding of the molecular basis of these lesions[17]. As an inherited multiple-neoplasia syndrome, VHL is characterized by highly vascular tumors and associated with an increased risk of development of both benign and malignant cysts spread in many organs (Figure 2). The main VHL manifestations include development of pancreatic cystadenoma[18] and endolymphatic sac tumor of the inner ear (ELSTs)[19], retinal or central nervous system (CNS) hemangioblastoma (HBs)[20], neuroendocrine tumor (pheochromocytomas, PCC and paraganglioma, PGL)[21], and renal cell carcinomas (RCCs) [22]. The latter, in particular, is considered the major cause of both morbidity and mortality for VHL patients due to its malignant features and diffuse radio-chemo-resistance[23]. Based on their histologic appearance, RCCs can be further subdivided into papillary renal carcinomas (15%), chromophobe tumors (5%), oncocytomas (5%) and clear-cell renal carcinomas (clear-cell RCC or ccRCC, 75%)[24] (Figure 2).

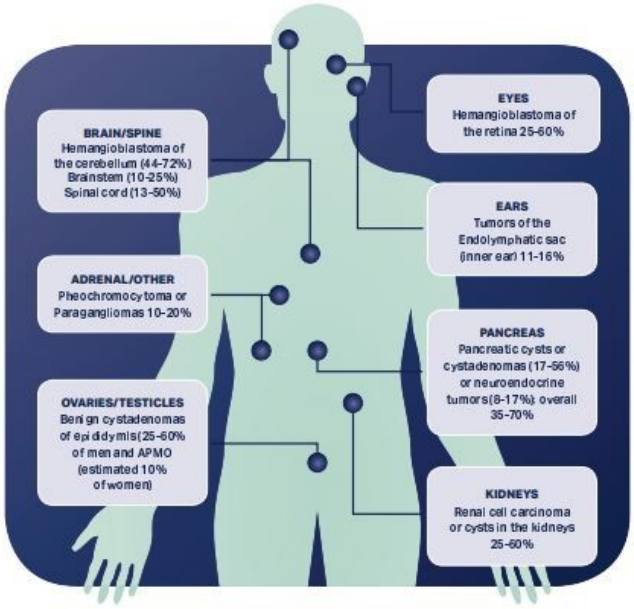


Figure 2. VHL occurrences in patients. VHL patients are characterized by multiple-neoplasia lesions spread into different organs (figure adapted from http://vhl-uk-ireland.org/). The highest percentage is related to retina and kidney tumors.

Clinical criteria for the diagnosis of VHL include presence: a) two or more HBs of the retina, spine, or brain; b) one CNS or retinal HB in addition to at least one of the typical visceral VHL manifestations including ccRCC or c) any one of the clinical manifestations along with a family history consistent with VHL[25]. Based on their propensity in developing HBs, PCC or RCC, VHL patients can be divided into several large subtypes accounting for different cancers risk. Type 1 VHL includes patients presenting hemangioblastoma and RCC, without pheochromocytomas[26]. Some data also suggested the existence of a Type 1B, in which patients develop CNS and/or retinal HBs, with low risk for both PCC and RCC[25]. Type 2 VHL is sub-divided as Type 2A, 2B or 2C. Respectively, Type 2A patients have a high risk of developing HBs and PCC in the absence of RCC, otherwise Type 2B has also RCC evidence. Finally, Type 2C category exclusively develops pheochromocytoma[27]. In addition, families harboring specific VHL mutations are characterized by autosomal-recessive polycythemia not necessary paired to an increased cancer risk. It has been proposed a Type 3 class grouping patients with this feature[28]. Although such phenotype classifications are promising, especially for medical patient management, it should be noted that a rigid subdivision in different VHL classes frequently fails in reliability due to VHL genotype-phenotype correlation, which is still evolving and far to be completely understood. As anticipated, initiation and progression of VHL-related tumors are connected to pVHL mutations or loss. Individuals with the hereditary form of these tumors inherit a single mutant *VHL* allele[29]. More than 1500 germline and somatic *VHL*

mutations are reported as linked to VHL disease[14]. These mutations are observed throughout the three exons of the VHL gene with 43.2% in exon 1, 17% in exon 2 and 39.8% in exon 3[14]. Excluding germline mutations linked to congenital polycythemias, the spectrum of mutation types is represented by 52% missense, 13% frameshift, 11% nonsense, 6% in-frame deletions/insertions, 11% large/complete deletions, and 7% splice mutations[30]. Truncating mutations or exon deletions usually display Type 1 disease features, whereas Type 2 patients are mostly linked to missense mutations[25]. Although *VHL* alteration is an almost frequent event in the pathogenesis of cancer, e.g. >70% of ccRCC, it is not sufficient to cause the disease onset. While a *VHL* allele is defective, a somatic inactivation has to disrupt the remaining wild-type copy. This may occur by a large deletion on the remaining allele or as a consequence of transcriptional silencing due to hypermethylation of both maternal and paternal alleles[31]. Thus, pVHL is proposed as a prototype of tumor suppressor gene consistent with the Knudson's two-hit model of tumorigenesis[32].

1.3 The VHL gene and protein structure

The human *VHL* gene is a 12-kb region located in the short arm of chromosome 3 (3p25-26.9)[33]. It consists of 3 exons encoding for a 213 amino-acid length protein namely VHL30 (or pVHL, full length), active both in nuclear and cytoplasmic compartments[34] and ubiquitously expressed in human tissues (Figure 3)[35]. *VHL* gene has three splice variants, 90 orthologues, and a paralog, which are associated with two different Ensemble families (i.e. von Hippel-Lindau disease tumor suppressor pVHL and interleukin 1 receptor-associated kinase 2 IRAK2). It is highly conserved among species, in particular in primates and rodents. However, homologs have been also identified in *C. elegans* and *Drosophila*[36]. The *VHL* promoter is GC-rich, lacking TATA or CCAAT motifs. Transcription initiates around a putative Sp1-binding site approximately 60 nucleotides upstream codon 1. Predicted binding sites for many transcription factors are present in the *VHL* promoter (NCBI Locus AF010238, HSU97187), although little is known about its transcriptional regulation[37]. The first exon spans codons 1–113 (nucleotides 1–340), while exon 2 codons 114–154 (nucleotides 341–463) and exon 3 codons 155–213 (nucleotides 464–642). Structurally, the 28-30 kDa protein product is formed by an acidic N-terminal domain (residues 1 to 53), a central β-domain core (residues 54 to 157) and a C-terminal α-domain (residues 158 to 213). In particular, the N-terminal tail is characterized by 8X5AA tandem repeats of G-[PAVG]-E-E-[DAYSLE] followed by a seven-stranded β sandwich (residues 63 to 154) and an α helix (H4; residues 193 to 204) that packs against one of the β sheets through hydrophobic interactions[38,39] (Figure 3). The α domain of VHL (residues 155 to 192) consists of three α helices (H1, H2, and H3), packed in a four-helix cluster (resembling a "folded leaf")[38]. The α and β domains are connected by two short polypeptide linkers (residues 154 to 156 and 189 to 194)[38,40].

Along the VHL amino-acid sequence three different interfaces are defined, i.e. a domain-domain (interface A), domain-peptide (interface B), and a lesser defined interface C[41]. These interfaces are binding sites for multiple VHL interactors. As reported in the VHLdb database[3], pVHL has indeed more than 500 different partners, related to numerous different cellular processes. However, many of these physical bindings are still not clear at the functional point of view.

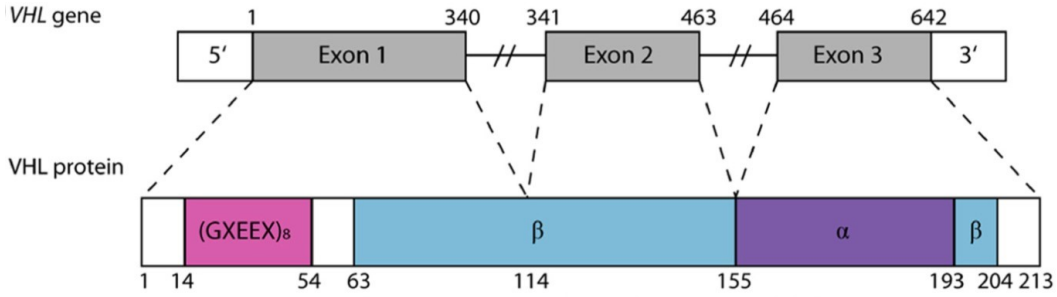


Figure 3. von Hippel-Lindau gene and protein structure. The VHL gene comprises three exons (gray). Nucleotide numbering is presented above gene structure. The VHL protein structure with (GXEEX)8 repeat motif (pink), α-domain (purple) and β-domain (blue). Codon numbering is indicated below[39].

1.3.1. VHL isoforms and paralog

The *VHL* gene encodes three different isoforms. The full-length pVHL30 includes exons 1, 2 and 3, (previously described). A shorter 160 amino-acid long isoform namely pVHL19[42] arises from an in-frame internal translation starting site (i.e. codon 54 within the *VHL* open reading frame). It completely lacks the N-terminal disordered region (1-53 amino acid), albeit preserving a biological activity[43]. Indeed, the reintroduction of either isoforms in *VHL-defective* ccRCC cells suppresses their ability to form tumors when xenografted in nude mice[42]. Both pVHL19 and pVHL30, bind elongin B, elongin C, and Cullin 2 in co-immunoprecipitation assays and reduce the hypoxia-inducible proteins levels such as vascular endothelial growth factor (VEGF), when reintroduced into renal carcinoma cells lacking a wild-type VHL allele[43]. They are ubiquitously expressed, although pVHL19 high expression levels are specifically observed in the urogenital system, brain, spinal cord, sensory ganglia, eyes and bronchial epithelium[44]. Some evidence, nevertheless, suggest the existence of asymmetric and differentiated functions of the two pVHL isoforms. As an example, pVHL19 is equally distributed in both the nucleus and cytoplasm when compared to the mainly cytoplasmic pVHL30 localization[43] as well as it does not associate cell membranes after biochemical fractioning.

This result may account for its failure to co-immunoprecipitate with fibronectin, a well-established pVHL30 interactor restricted to this sub-cellular compartment[45]. Moreover, a recent proteomic investigation linked specifically pVHL30 to the p53 pathway regulation through a complex formation with the p14ARF oncosuppressor[46]. This interaction is sustained by the N-terminal region of VHL full-length, which is absent in the pVHL19 isoform[4]. Over specific protein-protein interaction, pVHL19 isoform also shows an increased tendency to form homodimers further reinforcing the hypothesis of VHL isoform-specific binding specialization[4].

VHL172 isoform is produced by an alternative splicing event of exon 2 generating a 19 kDa-protein termed as VHL172 (or pVHL172)[47]. High expression level has been reported during human embryogenesis (i.e. 8-10 gestational weeks) in kidney, brain, spinal cord, eyes, testis and lung[48]. In some cases, VHL172 is reported to be the only *VHL* transcript detected in sporadic ccRCC samples or cell lines with the specific *VHL* point mutations in exon-intron boundaries or in the coding region[49]. This data suggests that VHL172 participates in tumor initiation or, at least, in progression events[49]. Recently, VHL172 isoform has been identified as a pro-tumorigenic protein, characterized by an antagonistic function whether compared to the pVHL30 ones[50]. Mice xenograft with pVHL172-expressing 786-O cells develop tumors with more extended sarcomatoid phenotype than tumours derived from parental 786-O cells. This tumor exacerbation effect has been proposed to be related to a hypoxia-independent pathway, considering that pVHL172 does not regulate hypoxia inducible transcriptional factor (HIF-2α), albeit retaining the ability to form an E3 ubiquitin ligase complex. Furthermore, pVHL172 upregulates the TGFB signaling and matrix metalloprotease protein 13 (MMP13), which are normally downregulated by pVHL30 reducing cancer aggressiveness. This role suggests that the expression balance of the two different isoforms may be critical in ccRCC initiation and progression[50].

Recently, a third translation variant of VHL30 has been identified and termed VHLα[51]. Its translation initiates from a second alternative start site, located upstream and in-frame with the VHL30 canonical ATG start codon. Thus, coding for a longer VHL isoform of about 35 kDa. VHLα ubiquitinates and degrades the hypoxia-inducible factor 1-alpha (HIF-1α) even though in a lesser manner compared to VHL30 and VHL19[52]. It is reported that VHLα preserves tumor suppressor activity interacting with heterogeneous nuclear ribonucleoprotein A2B1 (hnRNPA2B1), which consequently modulates pyruvate kinase (PKM) transcript splicing and reprograms cellular glucose metabolism (i.e. inhibition of Warburg effect)[51].

In 2004, the first human VHL paralog has been also identified and characterized. *VHL-like* gene (VHLL or VLP) is located on chromosome 1q21.2 between markers D1S2721 and D1S2635. The cDNA contains a predicted ORF of 676-bp encoding a 139-amino acid protein. Isolated by reverse transcription-PCR from human brain cerebellum, placenta and in several cancer cell lines, VLP has a 67.6% identity and 78% sequence similarity with the 1-157 VHL30 region. Despite of the extended truncation of the VHL-α domain required for enucleating the E3 ligase complex, VLP contains a highly conserved β domain and according to *in vitro* investigation it was proposed to maintain the association with HIF-1α in a hydroxylation-independent manner. The HIF-1α is consequently not degraded, rather it seems to be protected upon VLP expression. The accumulation of HIF-1α and overexpression of GLUT1, a glucose transporter whose expression is regulated by HIF-1α, are concrete evidence of this role. VLP acts as a dominant-negative VHL attenuating HIF-1α proteolysis in a dosage-dependent manner. In this way, the balance between VHL30 and VLP appears to be crucial in safeguarding the timely and appropriate level of transcriptional activity of HIF-1α[53].

1.3.2. VHL main role: HIF-1α and the oxygen-response regulation

The primary role of pVHL concerns the cellular response to oxygen levels and consequently it is related to renal cancer development, cardiovascular disease, ischemia, and chronic anemia[54]. Serving as the substrate recognition subunit of an E3 ubiquitin-protein ligase complex, in normoxic conditions VHL ubiquitinates and targets the α-subunit of hypoxia-inducible factors for a rapid oxygen-dependent proteolysis[55,56,57]. This mechanism puts into O_2-regulation play different proteins, such as HIFs family, PHDs and the VCB complex components (Figure 4).

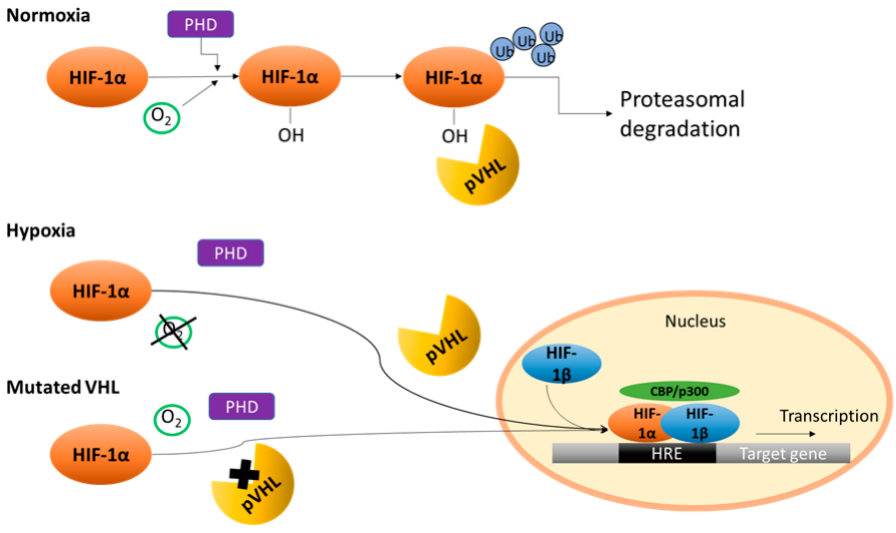


Figure 4. (Previous page) Regulation of HIF1α under normoxia and hypoxia conditions. Under normoxic conditions, the prolyl hydroxylase domain (PHD) uses molecular oxygen (O2) to hydroxylate HIF1α. Hydroxylated HIF1α is poly-ubiquitylated (Ub) for proteasomal degradation by binding the Von Hippel–Lindau protein (pVHL). Under hypoxic conditions, the activity of PHD is reduced or inactive. HIF1α is directly translocated to the nucleus where it binds HIF1β and CREB binding protein/p300 (CBP/p300) at the hypoxia response element (HRE) to act as transcription factor. Similarly, the entire mechanism results impaired when VHL gene is mutated, thus promoting an unregulated hypoxia response[57].

1.4 Hypoxia and HIFs family proteins

In the life of aerobic organisms, oxygen has a central role and its homeostasis is critical. Either insufficient or excess O_2 leads to increased levels of reactive oxygen species (ROS), which oxidize lipids, proteins and nucleic acids, yielding cell dysfunction and death. Thus, both the delivery and the consumption of O_2 are precisely regulated[58]. In addition, as the final acceptor of electrons in the mitochondrial respiratory chain, it generates adenosine triphosphate (ATP) maintaining cellular energy and viability. Under normoxia (or physoxia) a cell continuously preserves a high and constant ratio of cellular ATP/ADP to survive. A reduction of the normal oxygen supply will have consequences on the cell viability. This becomes relevant in particular during the non-physiological hypoxic status shared by a large number of pathological conditions, such as stroke, tissue ischemia, inflammation as well as during aggressive solid tumors growth[59]. It is generally accepted that the oxygen level in hypoxic tumor tissue is poorer than the oxygenation of the respective normal tissue and on average it is between 1%-2% O_2 and below. However, its specific value varies with the initial oxygenation of the tissue, size, stage and severity of the specific tumor. In healthy conditions, normoxia level ranges between approximately 9,5% O_2 in the renal cortex[60] to 4,6% O_2 in the brain[61]. As the most part of the biological mechanisms, hypoxia is not a binary stimulus. Rather, it's a complex mechanism with two different sides. During fetal development hypoxia represents a required positive stimulus[62], as the gradient in oxygen tension promotes a correct liver zonation[63]. The "Janus face" of hypoxia mechanism is also represented by tumor environment. Lower tissue O_2 concentration pushes anti-proliferative effects promoting differentiation, apoptosis and necrosis[64]. On the other hand, hypoxia also acts as an adaptive process leading some clones within the tumor to modify gene expression, acquire aggressiveness and promote malignacy[65]. Its complexity is related to the numerous intracellular signaling pathways acting to orchestrate this mechanism. Critical mediators of cellular adaptation to hypoxia are the hypoxia-inducible factors (HIFs), which are dimers consisting of two distinct subunits. The best studied are the HIF-α (α-subunit) and HIF-β (β subunit

also known as the aryl hydrocarbon receptor nuclear translocator, ARNT), respectively oxygen-sensitive and constitutively expressed. Both subunits belong to the bHLH-PAS protein family as their structures are related to two nuclear proteins found in *Drosophila* (Per and Sim, PAS) which present a basic-helix-loop-helix (bHLH) motif[66,67]. In general, the bHLH-PAS motifs are essential to allow heterodimerization between HIF-α and HIF-β subunits and for their binding to the hypoxia response element (HRE) on DNA sequence of target genes[68]. Only the HIF-α subunit contains the oxygen-dependent degradation domain (ODD domain or ODDD) (Figure 5)[69]. Under hypoxic conditions, HIF-α is binds HIF-β forming a dimer that specifically recognizes HRE and ultimately yielding active transcription of >100 target genes which mediate cell adaptation to low oxygen levels[70,71]. Three HIF-α paralogs are known in human: HIF-1α, located in chromosome 14 encoding three different isoforms (the longest one of 93 kDa); HIF-2α (also EPAS1 or HLF), located in chromosome 2 encoding a 97 kDa-protein and HIF-3α, located in chromosome 9 producing 7 different isoforms (the longest one of 73 kDa). HIF-1α and HIF-2α are the most widely studied. Unlike HIF-3α, HIF-2α shows relevant sequence similarity with HIF-1α (48% sequence identity); differences are observed at the level of the N-terminal transactivation domain (N-TAD). As a result, they are characterized by the different binding capabilities and target specificity[70], whereas the similar C-terminal transactivation domain (C-TAD) contributes to the transcription of their shared targets. These evidence could explain their common capability of heterodimerization with HIF-1β, but also their different tissue distributions and their incomplete overlapping roles[72]. HIF-1α is indeed ubiquitously expressed in the body, while HIF-2α expression is stricter to specific tissues (e.g. bone marrow macrophages and Kupffer cells)[73]. More importantly, HIF-1α and HIF-2α are reported to be mutually antagonistic in functions[74]. HIF-3α is the most different and subjected to extensive alternative splicing resulting in many isoform variants. Additionally, it has only one specific proline residue in the ODD domain and is lacking the C-TAD. Its transcription factor role has been doubted and in fact different HIF-3α variants have been shown to have inductive and suppressive effects on HIF targets[75]. In some case, it has been classified as a dominant-negative regulator over HIF-1/DNA binding ability[76]. Normally, the ODD domain contains two specific proline residues (Pro402 and 564 in human HIF-1α and Pro405 and Pro531 in human HIF-2α) that are hydroxylated under normoxic conditions by the HIF prolyl hydroxylase 1–3 (also known as EGLN2, 1, and 3 or HIF-P4H1-3)[77] and then rapidly targeted for degradation (Figure 5).

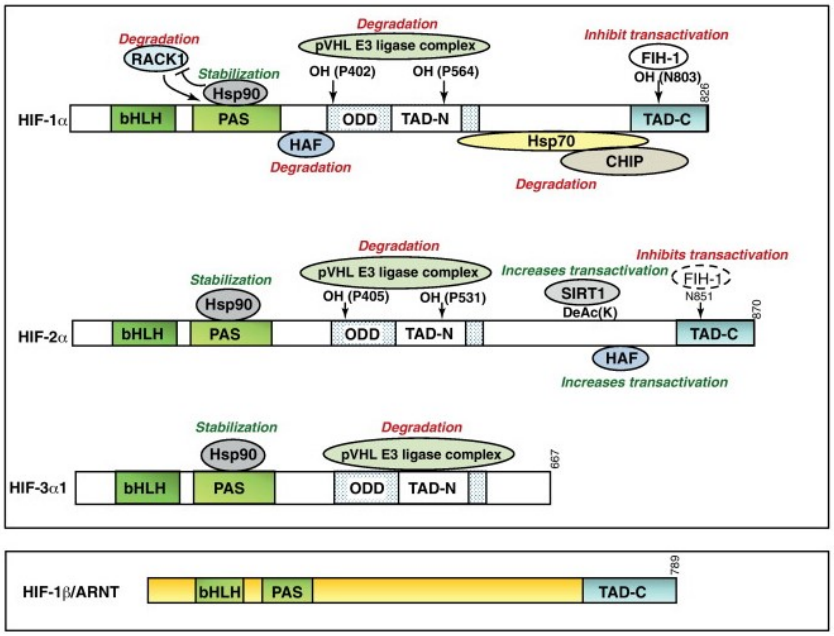


Figure 5. The structural domains of hypoxia-inducible factors. The basic helix–loop–helix (bHLH) and per-Arnt-SIM (PAS) domains are involved in DNA binding and heterodimerization. The oxygen-dependent degradation (ODD) domain is the oxygen-dependent hydroxylation, necessary for HIFs degradation. The N-terminal and C-terminal transactivation domains (TAD-N and TAD-C, respectively) are required for transcriptional activation. The von Hippel–Lindau protein (pVHL) E3 ligase complex regulates the oxygen-dependent degradation of all three major HIF-α subunits. Factor inhibiting HIF-1 (FIH-1) hydroxylates HIF-2α at a lower efficiency (broken oval) than HIF-1α (unbroken oval)[67].

Beside the HIF-regulation, other hypoxia-associated pathways include PI3K/AKT/mTOR[78], MAPK[79] also known as ERK pathways and NFkB[80]. Alone, these pathways are involved in cell proliferation, survival, apoptosis, metabolism, migration and inflammation.

1.4.1. Hydroxylation and the prolyl hydroxylases activity

Hydroxylation is a post-translational modification (PTM) catalyzed by hydroxylases, a family of ~70 enzymes that are strictly dependent on molecular oxygen, Fe(II), ascorbate, and the Kreb's cycle intermediate 2-oxoglutarate (2OG)[81]. These enzymes regulate fundamental cellular processes by catalyzing either the hydroxylation or demethylation (via hydroxylation) of DNA, RNA and proteins. As such, they have been implicated in various syndromes and diseases, but particularly in cancer. The potential importance of hydroxylation in cancer was first discussed in the 1990s following the discovery of asparaginyl- and aspartyl-hydroxylation in extracellular EGF repeats[82]. The prolyl hydroxylases (PHDs) are part of the family grouping the iron- and 2-oxoglutarate (2-OG)-dependent dioxygenases. In human, there are three PHD paralogs: PHD1, PHD2 and PHD3 (also known as

Elogin 1-3, EGLN 1-3). PHD1encodes a 44 kDa protein, which has also a shorter isoform lacking the first 33-amino acid. Whereas PHD2 is a 47 kDa protein with three different splicing variants and PHD3 a 28 kDa protein (UniProt data)[83]. The hydroxylation affinity of all the PHDs toward HIF-1α has been analyzed *in vitro* and reported as following PHD2 > PHD3 > PHD1. PHDs add a hydroxyl (OH−) group at the proline residue of HIFs but differ in their substrate specificity and distribution. The PHD2 hydroxylates more-efficiently HIF-1α, compared to HIF-2α and HIF-3α, whereas HIF-2α is more-efficiently hydroxylated by PHD1 and PHD3. PHD2 is specifically expressed in the adipose tissue, PHD1 in testis cells, whereas PHD3 is predominantly found in cardiac cells[84]. PHD2 remains localized to the cytoplasmic region of the cell, PHD3 is found in the nucleus, while PHD1 is present in cytoplasmic as well as in nuclear regions. In normoxic conditions, oxygen is abundant in cells enhancing the hydroxylation activity of PHDs[85], building up hydroxyl groups at the specific proline residues of the C-terminus of HIF-α. Chemically, hydroxylation involves the splitting of dioxygen into two oxygen atoms transferring one oxygen to a proline residue to form a hydroxyl group and another to the 2-OG to generate succinate and CO_2[86]. Fe^{2+} at the active site of the PHDs remains loosely bound by two histidine residues and one aspartic acid forming a 2-histidine-1-carboxylase coordination motif. Ascorbate helps to maintain iron (Fe) in the ferrous (Fe^{2+}) state fundamental to maintain and achieve the full activity of the PHDs[87]. Once HIF-α is hydroxylated at the proline residues by PHDs, it is further captured by pVHL and degraded through proteasomal pathway. In particular, HIF-1α is also hydroxylated by another cytoplasmic protein known as a factor inhibiting HIF-1α (FIH) at the aspartic residue of HIF-1α which further enhances degradation of HIF-1α and thereby reduces its transcriptional activation in tumor cells[88]. Indeed, FIH hydroxylates the asparagine residue within the C-TAD blocking HIF from binding to the p300-CBP coactivators and inhibits the transcriptional activation of HIF[89] (Figure 6).

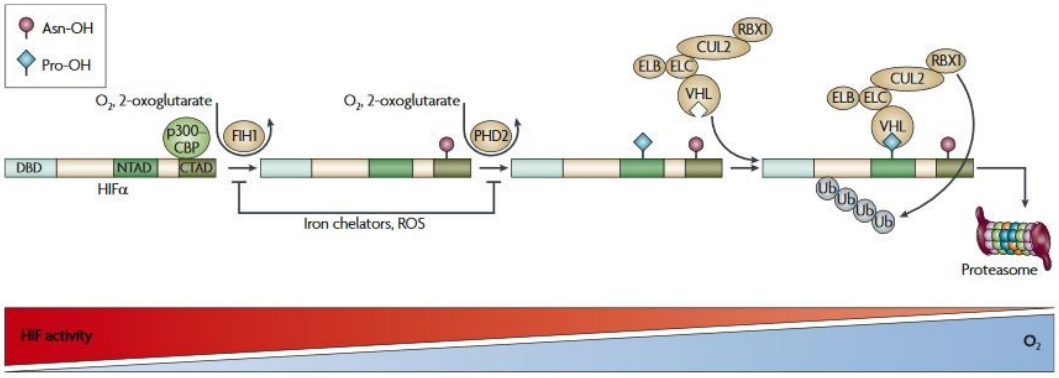


Figure 6. (Previous page) Control of hypoxia-inducible factor (HIF) by hydroxylation. When O_2 levels are low, HIFα is stable and the N-TAD and the C-TAD are active. At intermediate levels of hypoxia, the C-TAD is hydroxylated on a conserved asparagine (Asn) residue by factor inhibiting HIF (FIH1). This prevents recruitment of the co-activators p300 and CREB-binding protein (CBP). As O_2 is more plentiful HIFs are hydroxylated on two proline (Pro) residues near the N-TAD by prolyl hydroxylase 2 (PHD2), generating a binding site for the von Hippel–Lindau protein (VHL). Recruitment of the VHL ubiquitin (Ub) ligase complex containing Elongin C (ELC), Elongin B (ELB), cullin2 (CUL2) and RING-box protein 1 (RBX1) leads to the poly ubiquitination and proteasomal degradation of HIF. FIH1 and PHD2 require several cofactors, including 2-oxoglutarate and reduced iron, and are inhibited by reactive oxygen species (ROS)[59].

1.4.2. VHL protein associates EloC/EloB/Cul2 to regulate HIFs

Degradation of HIF-1α avoids its heterodimerization with HIF-1β inhibiting its nuclear translocation and target genes activation[90]. X-ray crystallographic studies of the pVHL–HIF-1α complex showed that pVHL presents an efficient binding pocket acting as a highly specialized cage for hydroxylated HIF-1α proline residues[40]. As member of the homonymous VHL-box, a specific E3 substrate receptors family, VHL cooperates with other proteins to perform its function. Similarly to the SOCS-box, the three helices structure is divided into a BC box and a cullin box[91]. Through the BC box, VHL forms a ternary complex with the transcription elongation factors C and B (also known as Elongin C, EloC and Elongin B, EloB) the so-called VCB complex[92,93]. Elongin C is a 112-amino-acid protein[94] which binds to the conserved F-box motif present in a large number of proteins participating in the regulation of multiple cellular processes including the cell cycle, gene transcription, and development. Elongin B is a 118-amino-acids protein composed of an 84-amino-acids amino-terminal ubiquitin-like (UbH) domain fused to a 34-amino-acids carboxy-terminal tail[95]. The Elongin BC complex was initially identified as a potent activator of RNA polymerase II (Pol II) elongation factor Elongin A, a transcription factor, stimulating RNA Pol II-dependent elongation rate by suppressing polymerase transient pausing along the DNA[96]. Through the Cullin Box, the VCB complex also associates with both cullin 2 (Cul2) and the RING-box protein (RBX1) forming the VCB–CR complex[97,98]. Cullin 2 serves as a complex's rigid scaffold contributing to catalysis by positioning the substrate against the ubiquitin-conjugating enzyme, with RBX1 mediating targets ubiquitination and their proteasomal degradation (Figure 7)[99].

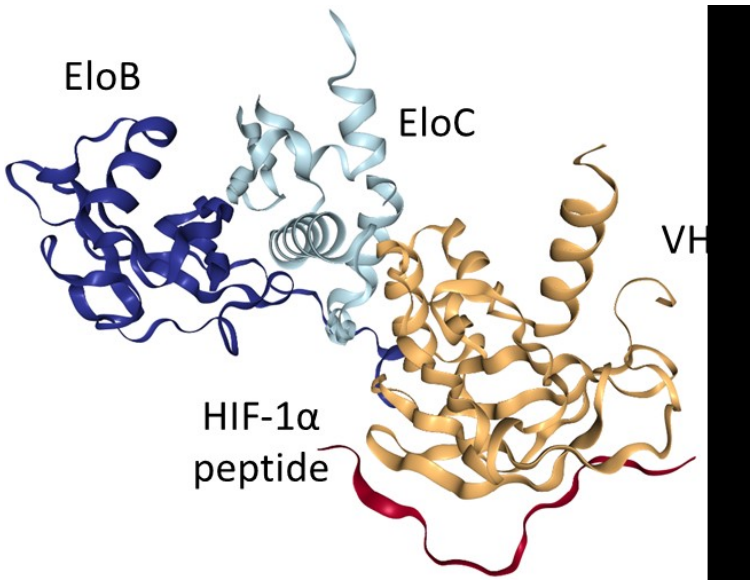


Figure 7. VCB crystal structure. HIF (represented as red colored peptide) ubiquitination and degradation is regulated by the VCB complex. VHL directly binds EloC through its helix domain, and in turn EloB is associated to EloC.

Besides the correct VCB-CR assembling, HIF-1α is poly-ubiquitinated and then degraded through the proteasomal pathway. Under hypoxic conditions (i.e. sub-optimal oxygen concentrations) VHL does not recognize HIF-1α leading to transcription of numerous downstream genes[100]. Similarly, several VHL mutations elude the ubiquitination of HIF-1α yielding VHL-derived diseases[101].

1.4.3. HIF-responsive gene products

The HIFs activation initiates a transcriptional program regulating hundreds of genes. A ChIP-seq analysis has identified more than 500 HIFs binding sites in the human genome[102]. HIFs target genes widely regulate cell homeostasis controlling cell survival, invasion, angiogenesis, metabolism, differentiation, senescence and epigenetics[103] (Figure 8). Thus, their pathways alteration hijack physiological cells to a tumour and malignant status[104]. For instance, the induction of robust angiogenesis in tumors appears to be due to overproduction of the vascular endothelial grow factor (VEGF), which is the most known HIF-responsive gene[105]. Linked to angiogenesis, there is also the platelet-derived growth factor (PDGF), similarly activating vessel sprouting[106]. Different growth factors and their receptors regulating cell cycle progression, i.e. transforming growth factor α (TGFα), epidermal growth factor receptor (EGFR), insulin-like growth factor 2 (IGF2), are also upregulated by HIFs. Further, HIF-1α upregulates cyclin D1 oncoprotein expression, which in turn inactivates pRB tumor suppressor via Cdk4 and Cdk6[107]. HIFs also regulate enzymes for glucose and fatty acid metabolism[108]. Especially, HIF-1α is thought to be a master mediator of metabolic switch from a highly efficient glucose oxidative phosphorylation to a less efficient glycolytic pathway aimed at

sustaining the energetic needs in the hypoxic environment. Enhancing the transactivation of glucose transporters (GLUT1) and glycolytic enzymes, glucose inside tumor cells is rapidly metabolized to pyruvate and then converted to lactic acid[109]. This in turn promotes an acidification of surrounding tumor environment paired with a sustained alkalinization of the tumor cell cytoplasm. Both these variations directly affect cell cycle, cell proliferation, differentiation and the cellular metabolism. Indeed, the intracellular alkaline pH stimulates proliferation, while the extracellular acidity activates degrading enzymes (i.e. metalloprotease) on the external matrix leading to tumor invasion and blocking the immune-mediated response. Finally, altered pH values also decrease the tumor access for certain chemotherapeutics[110] reducing the efficacy of tumor contrasting treatments. As a whole, when VHL regulation on HIFs is disrupted, the hypoxia modulator positively sustain tumor progression and spreading.

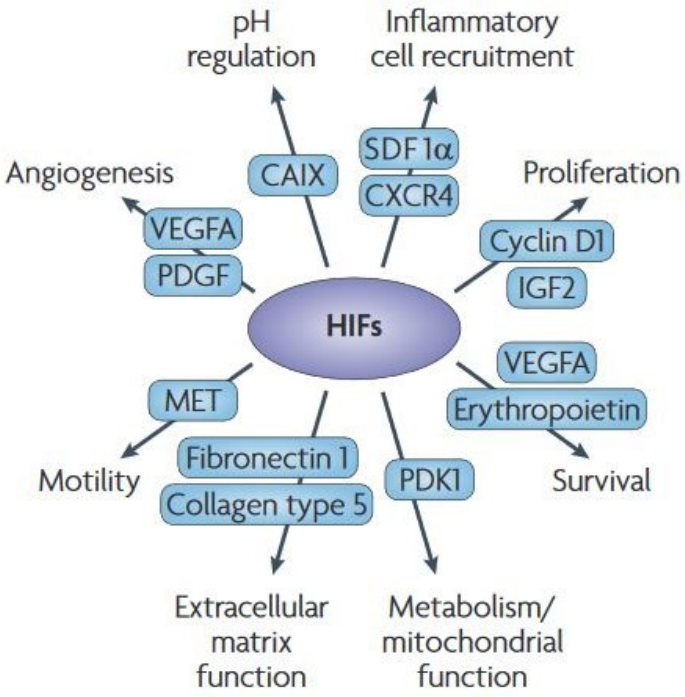


Figure 8. Genes activated by hypoxia-inducible factors (HIFs) involved in tumor progression. Genes encoding proteins involved in numerous aspects of tumor initiation, growth and metastasis are transcriptionally activated by either HIF-1α or HIF-2α[59].

1.4.4. HIF-independent VHL protein functions

Beside its role in regulating HIFs stability, pVHL has also multiple HIF-independent functions such as in regulating extracellular matrix assembly [111], receptors internalization[112], regulating RNA stability[113] and gene transcription[114], microtubule stability[115], phosphorylation enhancer and inhibitor[116,117], as well as modulating nerve growth factor signalling[34], cell senescence[118] and maintenance of the primary cilium[119,120]. Another important role concerns the apoptosis regulation. In particular, RCCs are known for their resistance to conventional cytotoxic chemotherapies, whose efficacy is tightly linked to p53-mediated apoptosis[121]. However, most RCCs present with neither p53 loss nor mutations, suggesting that alternative anti-apoptotic pathways are active[122]. Previous reports assess that both HIF-1α and pVHL affect p53 function. HIF-1α directly binds and modulates p53 activity[123], whereas pVHL-mediated regulation is HIF-independent and promotes MDM2 functional suppression yielding p53 stabilization and activation of its target genes[46]. Therefore, pVHL impairment could affect p53 regulation by both HIF-dependent and -independent effects. To complicate this scenario, pVHL deficient cells sharpen nuclear factor B (NF-B) activity mediating resistance to chemotherapy-induced apoptosis as well[124]. *In vivo* and *in vitro* loss of pVHL also induces cellular senescence, an irreversible growth arrest in response to DNA damage[125]. Contrarily to its well-known E3 ligase activity, many of these pVHL alternative functions are mediated by stabilization of its binding proteins. Thus, pVHL appears to be an adaptor protein that, depending on the interacting partners, can promote protein degradation or serve as a chaperon. In other words, being a small and metastable protein VHL is a typical example of a *hub* involved in a plethora of cellular pathways.

1.5 Hub proteins

Proteins associate other proteins to fulfill biological functions. They combine together permanently or transiently to form modules being part of a specific pathway[126]. Most proteins are believed to interact with only a few other proteins, while a limited number of proteins have multiple functional and biological partners. In cells, these highly connected proteins are considered as essential network "nodes" and their alteration will drastically change a biological pathway. As consequence, essential functions are disrupted and disease prevails[127]. Such proteins with a large number of interactors are classified as *hub* protein (e.g. p53 or VHL)[128]. Considering their important role in preserving cells homeostasis, *hub* proteins are extensively studied due to their deregulation frequently harmful. During the years, hubs in protein-protein interaction (PPI) have been differently classified into numerous categories. Han *et al*. first identified "date" and "party" hubs as proteins involved in a single interaction at a time (also, *transient hubs*) and proteins involved in multiple simultaneous

interactions (also, *obligate hubs*)[129], respectively. In following literature, hubs are divided into *stable*, *sociable* and *non-sociable hubs*. In particular, stable proteins are part of stable complexes, sociable proteins interact with three or more binding sites, and non-sociable proteins are bound to only one partner such as the transient one[130]. Beside the ability to interact with different proteins, *hubs* are also defined by peculiar structural features. One of the most significant factors is the structural flexibility, which is defined in these studies as the ability to fold into different structural conformations by presenting several loops or coils regions. In addition, the presence of intrinsically disordered regions supports protein flexibility feature[127]. Taking Calmodulin (CaM) as a prototype example of protein plasticity, the disordered region is represented by a flexible linker region connecting two structured domains. A 36 amino-acids length disordered region connects its two Ca^{2+}-binding domains allowing CaM to bind several unrelated targets[131]. In another case, the disordered region is itself the binding region such as the p53 N-terminal region, which is disordered and able to bind MDM2 ubiquitin ligase[132]. More recent literature suggests hubs to be also characterized by charged residues which are prone to engage electrostatic interactions as in the case of degradation tag protein Ubiquitin. It does not contain any disordered region, but its highly charged surface facilitates complex stabilization with other proteins[133]. Even if order domains often host the binding sites, the disordered region seems to be essential for hubs function[134].

1.6 Intrinsically disordered proteins (IDPs)

From structural point of view, proteins can be grossly divided in structured and either entirely or partially intrinsically disordered proteins (Figure 9). Structured proteins generally present with a single or multiple structured domain. According to a classical structure-function paradigm, protein function is critically dependent on the its well-defined ternary structures. This assumption is mostly derived by enzymes observations, which are characterized by high specificity for their proper cofactors and functions[135]. Nevertheless a large fraction of the proteome of any organism is enriched in non-structured proteins which are equally functional[136]. At physiologic conditions, such protein sequences are unable to fold into fixed globular structure but are dynamic and strictly disordered. These proteins rapidly fluctuate among several conformations and assemblies[137]. Some of them are demonstrated to be completely disordered, namely as intrinsically disordered proteins (IDPs). The majority of eukaryotic proteins are constituted of structured and intrinsically disordered regions (IDRs) and both are necessary for the protein role. The peculiarity is that the disordered features make a same protein to be adaptable to different cellular contexts[138]. Traditionally, IDRs were considered as passive segments in protein "linking" structured domains. However, it is now well established that IDRs are vigorously active part in the protein functionalities summarized as molecular recognition,

molecular assembly, protein modification and entropic chains[135]. According to these different activities, six different functional classes of disordered protein regions are obtained, such as entropic chains, display sites, chaperones, scaffolds, scavengers and effectors[135].

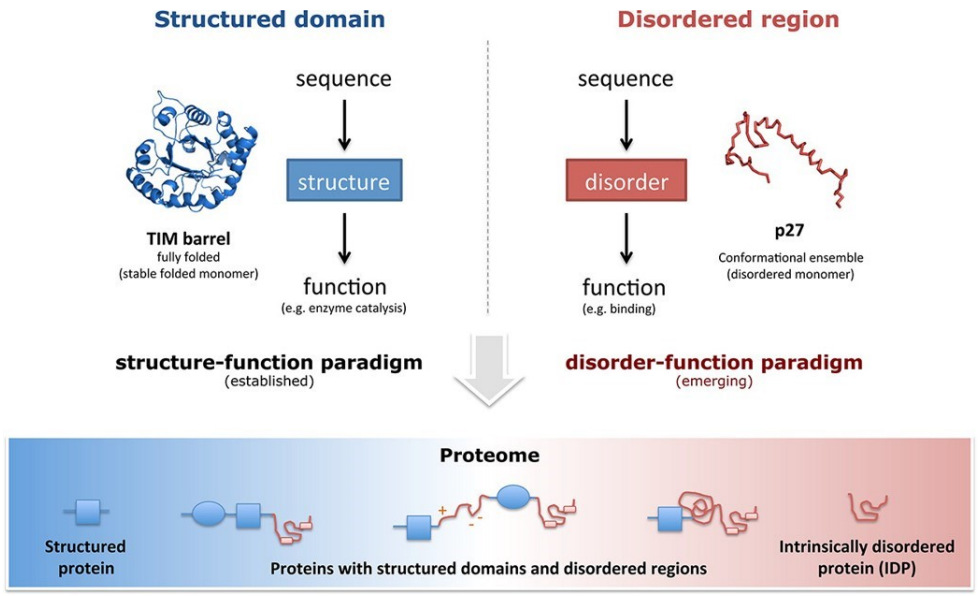


Figure 9. Structured domains and intrinsically disordered regions (IDRs). The scheme represents two fundamental classes of functional building blocks of proteins. Most proteins include both structured and disordered domains[135].

A common functional module within the IDRs is the concept of the linear motif, also known as short linear motif (SLiMs) or MiniMOtifs[139]. Annotated as usually 3-10 amino acids long, these sequences target proteins to a particular subcellular location, control the stability of a protein and promote recruitment of binding factors to facilitate complex formation. Of note, multiple enzymes recruit specific SLiMs thus modulating their chemical and functional state by post-translational modifications (PTMs). The short and flexible nature of many SLiMs allow them to sustain multiple interactions with different globular domains[135]. SLiMs are divided into two major families: SLiMs acting as modification sites and other acting as ligands, each one including sub-groups. The first family can be divided into three sub-groups. The post-translational processing events or proteolytic cleavage is the first one, in which the motif has a [ED]xxD[AGS] consensus sequence. Whereas the PTM moiety removal and addition is the second subgroups, composed by enzymes catalyzing PTMs on a precise sequence. This is a typical feature of the cyclin-dependent kinase recognition motif

[ST]Px[KR], which is modified to regulate cell cycle progression. Finally, the third sub-groups includes SLiMs acting as structural modification, generally driving the alteration of a peptide backbone[135]. SLiMs acting as ligands are also divided into three sub-classes: the complex promoting motifs, the docking motifs and the targeting motifs sub-class, which localize proteins towards subcellular organelles. Interestingly, docking motifs usually increase the protein binding specificity providing additional binding surfaces. Examples include the degron motifs[140] (KEN box, D box, F box), which act as recognition surfaces for ubiquitin ligases and often function to regulate protein degradation by the 26S proteasome. A clear example of this function is the hydroxy-degron recognized by pVHL[141].

1.7 Principles of the cell cycle progression

The cell cycle has a key role in regulating cells duplication and division. It is divided into synthesis phase (S-phase) and mitosis (M-phase), separated by two Gap phases (G_1-phase and G_2-phase). During G_1, a cell prepares itself for genome replication, assembling the various components forming the replication origins on chromatin. It is an extracellular environment-dependent progression, where adhesion to substratum and growth factors stimulate the checkpoint overcome leading to S-phase activation[142]. When growth-promoting signals are absent or, on the other hand, pro-differentiation signals are present, cells exit the cell cycle. This is defined as quiescence state or G_0 phase, a cellular reversible resting period finely regulated by Rb family proteins (i.e. Rb, Retinoblastoma; p107 and p130) phosphorylation and E2F (Elongation factor 2) binding[143]. Quiescent cells re-enter cell cycle in response to mitogenic inputs, leading to irreversible DNA replication in S-phase. At this level, a second Gap phase occurs. G_2 provides an appropriated DNA packaging, in order to assure a correct cells division in the M-phase[144]. These steps are timely controlled by a family of protein kinase, namely cyclin-dependent kinase (CDK) [145]. At the time of writing, CDK family accounts for >20 members [146], each characterized by a conserved catalytic core made up of an ATP-binding pocket, a PSTAIRE-like cyclin-binding domain and an activating T-loop motif. Collectively, these features are involved in associating CDK-cyclin via the PSTAIRE helix. Thus, the T-loop is displaced and the substrate-binding interface is exposed realigning critical active residues for the phospho-transfer reaction. Indeed, they structurally coordinate adenosine triphosphate (ATP) and transfer phosphate to a proper substrate. They are binary enzymes whose function requires their association to a specific family of allosteric activators[147]. Through the conserved cyclin box domain, each cyclin binds to specific CDKs during the different cell cycle phases. Although in most cases complexes rely on specific partnerships, some CDKs can form complexes with multiple cyclins. While the C-terminal is occupied by cyclin, the CDKs N-terminal regions are exposed to specific post-translation

modification on threonine and serine residues. In addition, CDKs are subjected to direct regulation by small inhibitory proteins defined as CDKs inhibitors (or CKIs). These regulators inactivate CDK-cyclin complexes by binding them[148]. There are two main families of CKIs, the Ink4 family (i.e. inhibitors of CDK4) which binds exclusively G_1 CDK4 and CDK6 and the CDKN1 family proteins, which regulates a broader range of CDK-cyclin complexes. The first family include four members, namely $p16^{Ink4a}$, $p15^{Ink4b}$, $p18^{Ink4c}$ and $p19^{Ink4d}$. The $p21^{Cip1}$, $p27^{Kip1}$, and $p57^{Kip2}$ form the second family.

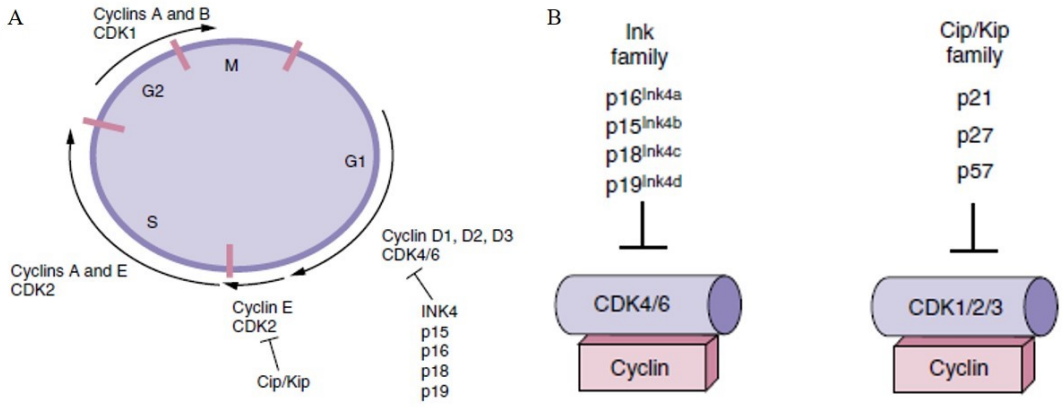


Figure 10. Regulation of the cell cycle. The cell cycle consists of four distinct phases: G1 phase, S phase (synthesis), G2 and M phase (mitosis or meiosis). The different steps are orchestrated by proper cyclin-CDK complexes (A). The cell cycle inhibition is controlled by two distinguished families of proteins, respectively named as Ink family and CDNK1 family (B).

1.7.1. General features of the CDKN1 family proteins

The CDKN1 family includes three different proteins (Figure 11): $p21^{Cip1/Waf1/Sdi1}$ (p21, encoded by *cdkn1a*)[149], $p27^{Kip1}$ (p27, *cdkn1b*)[150], and $p57^{Kip2}$ (p57, *cdkn1c*) [151]. Respectively, p21 gene is located in chromosome 6p21.2 and encodes a protein of 164 residues; p27 human gene is mapped to chromosome 12p13 and encodes a protein of 198 amino acids; and in chromosome 11p15.5 is present the p57 gene, a 316-protein length. They share a conserved N-terminal domain that mediates binding to cyclins and CDKs but diverge in the remainder of their sequences, suggesting that each of these proteins could have distinct functions and regulations. Within the amino-terminal region, they present several structured regions, such as the highly conserved kinase inhibitory domain (KID). In turn, the KID domain includes three sub-domains: the cyclin-binding subdomain (D1), the CDK-binding

subdomain (D2) and a linker subdomain that joins D1 and D2 (LH) [152]. Although they exhibit some similar features, the C-terminal regions differ in length and sequence features. The longest C-terminal tail belongs to p57 and contains two different functional domains: the QT box, significantly homologous with the C-terminal domain of p27 and the proliferative cell nuclear antigen binding domain (PCNA-binding domain) shared with the carboxy-terminal region of p21[153]. In addition, p57 has a central region rich in proline-alanine repeats (PAPA region), unique among the Cip/Kip and also absent in the murine isoform. These three proteins lack a fixed tertiary structure and are hence classified as intrinsically disordered proteins. Thus, they show a binding plasticity and a faster degradation rate[136].

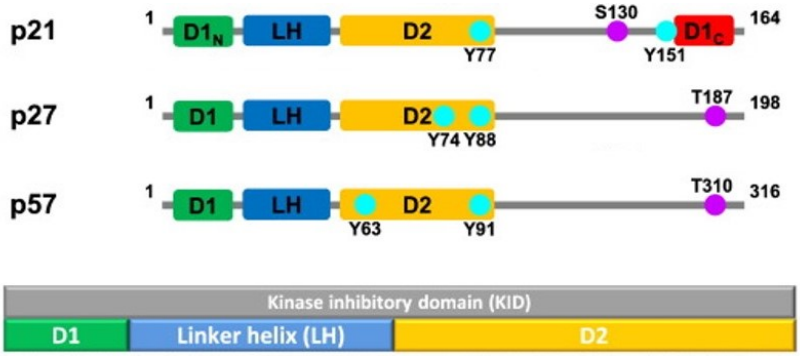


Figure 11. CDKN1 family proteins structures. The picture shows the different structural domains contained in p21, p27 and p57. In particular, in the amino-terminal region are included the cyclin-binding subdomain (D1), the CDK-binding subdomain (D2) and a linker subdomain that joins D1 and D2 (LH). Phosphorylation sites are highlighted with coloured dots.

Accordingly to their structural differences, CKIs display specific expression patterns, both *in vivo* and in *in vitro*. The p21 is mainly expressed in differentiated cells of adult tissues, while p27 and p57 are widely expressed in developing tissues. Although structurally closest members, p27 and p57 show mutually exclusive patterns. In particular, p27 expression prevails in thymus, spleen, retina, testis and ovary and persists at high levels during entire life. p57 instead prevails in intestine, cartilage, skeletal muscle, palate and pancreas and significantly decreases in human around the birth age to be finally restricted into specific organs. In adult, its expression is detectable in heart, brain, lung, kidney, pancreas, skeletal muscle, testis and placenta[153]. At cellular level, all are present into nucleus and eventually translocated into cytoplasm[154]. Their expression is also differently regulated. Especially, p21 is transcriptionally induced in response to DNA damage and p53 activation[155]. It is also controlled

through post-translational modifications in a p53-independent manner. Moreover, its turnover and sub-cellular localization is finely regulated by phosphorylation. For instance, phosphorylation on S130 leads to p21 poly-ubiquitination via SCF-Skp2 complex and subsequent degradation by the 26S proteasome. Whereas phosphorylation on T145, S146 or S153 is associated with cytoplasmic localization of p21. The p21 cytoplasmic localization, together with its expression deregulation, is common in human malignancies and correlates positively with aggressive tumors and poor prognosis[154]. Unlike, p27 mRNA levels are generally constant throughout the entire cell cycle. However, its expression is regulated by several different independent mechanisms. Among them, proteolytic degradation is mediated by CDK2-cyclin E via phosphorylation of Thr187 which facilitates recognition of p27 by the SCF-Skp2 complex[156]. Upon mitogenic stimuli during G_1 phase, p27 is also phosphorylated on Ser10 which induces its direct nuclear export. Alternatively, phosphorylation on Thr157, Tyr74 and Tyr88 respectively induces cytoplasmic localization and accumulation[157]. Its reduced expression or cytoplasmic localization is also associated with malignancy and worse outcomes. The p57 is the most complex member of the CDKN1 family. It has a highly cell-specific, temporal and spatial profile during fetal, postnatal, puberty and adult life. Moreover, it exposes to parental imprinting an epigenetic event regulated by the imprinted control region (ICR) acting in *cis*[158]. Similarly to the other CDKN1 members, its localization is regulated by phosphorylation. Functionally, p21 p27 and p57 inhibit the CDKs activity leading to cell cycle arrest, mainly at G_1 phase[153]. However, each members of this family is able to unroll proper cell cycle-dependent or independent functions. First, p21 inhibits CDK2–cyclin E complex[159]. Consequently, CDK2-dependent phosphorylation of RB and sequestration of E2F1 are inhibited, as cell progresses into and through S phase. Also, CDK2–cyclin A and CDK1–cyclin A complexes are disrupted by p21 altering the progression through S phase and into G_2, respectively. Additionally, p21 inhibits the kinase activity of CDK1–cyclin B1 inhibiting progression through G_2 and G_2/M. Through its carboxyl-terminal domain, p21 binds to and inhibits proliferative cell nuclear antigen (PCNA) modulating DNA synthesis and PCNA-dependent DNA repair[159] in a CDK2-independent manner. In addition to its roles in cell cycle regulation, p21 is involved in various other cellular processes including apoptosis, transcriptional factors activity (i.e. E2F1, STAT3 and MYC), cytoskeletal dynamics, stem cell commitment and differentiation[160]. This functional heterogeneity reflects to its conformational disordered structure able to bind multiple partners. p27 acts on CDK2-cyclin E complex during late G_1 and early S phase, avoiding phase transition. In addition, as transcriptional regulator, its reduction could deregulate many genes related to tumorigenesis. Since transcription is a nuclear affair, cytoplasmic p27 plays an oncogenic role activating cell motility, cell invasion and cytoskeleton modification[161]. For instance, in the cytosol, p27 can associate with RhoA, preventing

its activation and thereby modulating actin cytoskeleton dynamics and migration. Likewise, p57 best established function is to block cell cycle progression or promoting cell differentiation through the inhibition of G1/S CDKs. Like the other CKIs, p57 regulates cells by multiple CDK-inhibition independent mechanisms. For instance, p57 directly regulates differentiation interacting with some lineage-specific transcription factors, such as MyoD in skeletal muscle cells, Nurr1 in dopaminergic neurons, or NeuroD in neural stem cells. Being involved in the cell cycle regulation, the CDKN1 family deregulation is consequently related to human diseases. The primary consequence of an altered control includes failed or altered development and/or neoplastic/cancerous growth. In particular, the expression levels of the CDK-inhibitors are altered in a variety of cancers. For instance, p21 overexpression correlates positively with tumor grade, invasiveness and aggressiveness[162]. Similarly, the decrease of p27 seems to foster proliferation by deregulating its downstream pathways[163], while p57 appears increased in the initial stages of cancer and gradually decreased as malignancy progresses[153]. Thereby, they could be classified as tumors antagonists unrolling as oncosuppressor. Nevertheless, several evidences demonstrated that their subcellular localization cause an overturning of their activities. The CDKN1 proteins demonstrate a two-faced behavior in relation of their subcellular localization. For instance, cells with damaged or absent p53 may inhibit apoptosis, promote cell proliferation and facilitate migration, when p21 is moved and concentrated in the cytoplasm[164]. Low amount of nuclear p27 is frequently observed in human malignancies including carcinomas of the breast, brain, lung, and prostate associated with increased tumor aggressiveness[165]. Also, the cytoplasmic localization of p57 is related to an oncogenic-like behavior that has to be better clarified. The CDKN1 family proteins expression is considered as valuable prognostic marker for clinical outcome in human cancers[148]. The participation in cell cycle control is shared by CDKN1 proteins. They also function as modulators of apoptosis, cytoskeleton organization, DNA repair, cell migration and others. It was also observed that they also play a role as cell senescence regulators. This cellular phenomenon is an irreversible arrest of cell proliferation, which converts a mitotic cell into a post-mitotic one. Firstly described in 1961 by Hayflick and Moorhead, it was associated to a cellular “stopwatch” represented as a sort of cell aging mechanism[166]. Three decades later, the telomere shortening has been identified as the prime cause of the cell proliferation decline and senescence onset[167]. Considering the loss of proliferation capability, senescence resembles quiescence status. However, quiescence is reversible and necessary for cell maturation and differentiation, thus is the consequence of a proper developmental program. Unlike, the senescence response is triggered by stimuli, insults and stresses, such as oxidative damage, ionizing radiation, and oncogenes alteration[168]. As senescence is related to a stable growth arrest, proliferative markers (i.e. Ki67) are absent, cyclin-CDKs activity is strongly down-regulated and high level of CDKs-

inhibitors are revealed[169]. Thus, CDKN1 family members ablation overcome senescence activation supporting tumorigenesis[170].

1.7.2. p53 and MDM2 pathway

Besides the CDKN1s, the cell cycle is finely regulated by the p53-pathway[171]. Classified as the "guardian of the genome", p53 is a principal mediator of the growth arrest, senescence and apoptosis in response to a plethora of cellular damage, especially induced by damaged DNA[172]. In normal unstressed cells, p53 is characterized by low stability and short half-life. Its degradation is related to the human MDM2 protein, which harbors a self- and p53-specific E3 ubiquitin ligase activity[173]. The 491-amino acid length protein interacts with the N-terminal transactivation domain of p53 through its own N-terminal region. This interaction has been further elucidated by the crystal structure of the complex p53-MDM2[132]. Briefly, the amino terminal domain of MDM2 (25-109 residues) forms a deep hydrophobic pocket in which an amphipathic portion of p53 (18-26 residues) deeply inserts. As a result, the p53 transcription activity is blocked, the protein is poly-ubiquitinated and targeted for degradation. Accordingly, some p53 N-terminal mutants do not retain the interaction showing a degradation resistance[174]. Another way to stabilize and maintain p53 active is the direct regulation of MDM2. The cascade considers the Ras-dependent stimulation of p14ARF (or ARF), an alternative product of the INK4A tumor suppressor locus[175]. ARF binds the RING finger domain of MDM2, which is the site of its E3 ligase activity, sequestrating it into the nucleolus. Hence, p53 activates the transcription of its downstream genes, such as p21. This mechanism is crucial for cell cycle regulation and cancer suppression (Figure 12).

The p53-pathway is also reported to interconnect with the oxygen sensing regulation acting on both HIF-α and -β subunits. Under prolonged hypoxia, p53 accumulates in the cell downregulating HIF-1α transcriptional activity[176]. Whereas the p53-induced transcription of the miR-107 (microRNA-107) decreases the hypoxia signaling by suppressing expression of hypoxia inducible factor-1β (HIF-1β). As consequence, the neovascularization is reduced in human colon cancer cells[177]. Furthermore, HIF-1α is reported to directly bind MDM2 both *in vivo* and *in vitro*, suggesting that MDM2 may act as bridge in a ternary complex with p53 and HIF-1α. Given these data, HIF-1α could be degraded in a VHL-independent manner[178]. The direct interaction between p53 and pVHL resulting in both MDM2-induced suppression and nuclear p53 export[46] is also crucial. Like MDM2, pVHL may act as linker portion between MDM2 and p53 in a second possible biological ternary complex. Again, the dual role of VHL in inhibiting the ubiquitin-mediated degradation of either HIF-1α and p53 clearly sustain the functional interaction between cell cycle regulation and hypoxia.

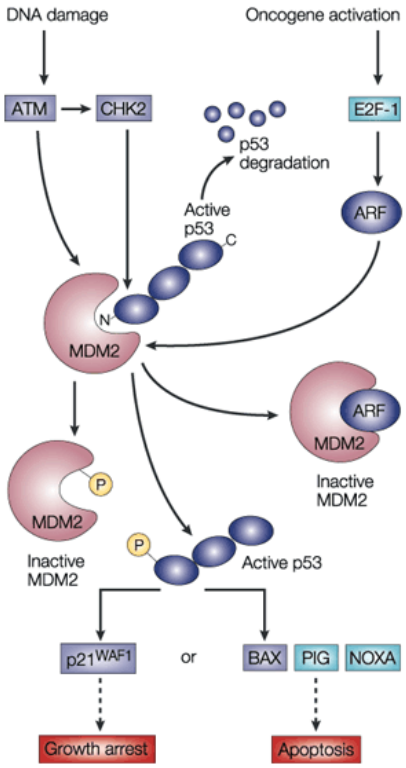


Figure 12. The p53 regulation. The transcription factor p53 is inactivated by MDM2, an ubiquitin ligase that targets p53 for degradation in the proteasome. In turn, MDM2 is regulated by p14ARF (ARF) and translocated into nucleolus.

Materials and Methods

Materials and Methods

2.1 Bacteria, yeast and mammalian cells

During experimental procedures, bacteria and different yeast strains were used respectively for cloning and *in vivo* analysis (Table 1).

Name	Organism	Genotype	Provider
TOP 10 *E. coli*	Bacteria	F mcrA Δ (mrr-hsdRMS-mcrBC) Φ80lac Δ ZM15 Δ lacX74 recA1araD139 Δ (araleu)7697 galU galK rpsL (StrR) end A1 nupG	Invitrogen
Y190	Yeast	MATa, gal4-542, gal80-538, his3, trp1-901, ade2-101, ura3-52, leu2-3, 112, URA3::GAL1-LacZ, Lys2::GAL1-HIS3cyhr	Euroscarf
Y187	Yeast	MATα, ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, met–, gal80Δ, URA3::GAL1UAS-GAL1TATA-lacZ	Clontech
Y2H Gold	Yeast	MATa, trp-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2::GAL2UAS-Gal1TATA-HIS3, GAL2UAS-Gal12TATA-Ade2, URA::MEL1UAS-Mel1TATA AUR1-c MEL1	Clontech

Table 1. Bacteria and yeast strains.

In addition, some experiments were also performed using mammalian cell lines (Table 2).

Cell line	Tissue	Morphology	Culture Properties	Features
HEK293T	Embryonic kidney	Epithelial	Adherent	Highly transfectable derivative of human embryonic kidney 293 cells (ATCC)
HeLa	Cervix	Epithelial	Adherent	Adenocarcinoma derived, suitable for transfection (ATCC)

Table 2. Cell lines features.

2.2 Media and solutions

Media components powders are correctly weight and resuspend in sterilized ddH_2O. The solution were then autoclaved at 121°C for 40 minutes to avoid any contamination.

2.2.1. Bacteria media

Bacteria cells were cultured in Luria-Bertani (LB) media. In particular, during the last step of transformation protocol (section 2.5) bacteria were plated on LB supplemented with Agar and

antibiotic. Whereas, liquid LB supplemented with antibiotic was used to inoculate bacteria before DNA extraction (section 2.6). The LB medium is composed as reported in the following Table 3:

Media	*Composition*
LB (Luria-Bertani)	1% Bacto Triptone (Difco) 0.5% Yeast Extract (Difco) 0.5% NaCl (Sigma) 2% Agar (Cat. 214030, Bacto™Agar, for solid media) * When necessary, LB is supplemented with Ampicillin (100 µg/mL, LBA) or Kanamycin (25 µg/mL, LBK)

Table 3. Bacteria media.

2.2.2. Yeast media

Yeast media were instead composed as reported below in Table 4. When necessary a liquid or solid media was employed during experiments supplemented with specific amino acids.

Media	*Composition*
YPD (rich medium)	1% Bacto Peptone (Difco) 1% Yeast Extract (Difco) 2% Dextrose (Sigma) 2% Agar(Cat. 214030, Bacto™Agar, for solid media) *When necessary, YPD is supplemented with adenine (Sigma, 80 mg/ml)
SD (selective medium)	0.17% Yeast Nitrogen Base (Difco) 2% Dextrose (Sigma) 0.5% Ammonium Sulfate (Sigma) 2,3 % Agar (Cat. 214030, Bacto™Agar, for solid media) * When necessary, SD is supplemented with Drop out (Cat. Y2001, Sigma), amino acids (Sigma, adenine 80 mg/ml; histidine 10 mg/ml; L-leucine 60 mg/ml, tryptophan 20 mg/ml and/or uracil 50 mg/ml), 3AT at different final concentration (30mM, 60mM or 90mM).

Table 4. Yeast media.

Whereas powder pouches were properly provided by Clontech to perform the library screening as reported in secion 3.3 (Results). The list of media are here presented in Table 5.

Yeast Media Pouches	*Clontech Cat. No.*	*Volume Media per pouches (mL)*
Rich media (for Routine Culturing of Untransformed Yeast)		
YPDA Broth	630306	0.5
YPDA with Agar	630307	0.5
Minimal Media Single Dropouts (SDO)		
SD-Trp Broth	630308	0.5
SD-Trp with Agar	630309	0.5
SD-Leu Broth	630310	0.5
SD-Leu with Agar	630311	0.5
Minimal Media Double Dropouts (DDO)		
SD-Leu/-Trp Broth	630316	0.5
SD-Leu/-Trp with Agar	630317	0.5
Minimal Media Quadruple Dropouts (QDO)		
SD-Ade/-His/-Leu/-Trp Broth	630322	0.5
SD-Ade/-His/-Leu/-Trp with Agar	630323	0.5

Table 5. Yeast library screening media. Yeast media set provided by Clontech for Matchmaker Gold Protocols. Clotech offers a "ready-mixed" foil pouches media. When necessary, media has been supplemented with Aureobasidin A (Cat. 630466, 500 µg/ml stock solution) and X-α-Gal (Cat. 630462, 20 mg/ml stock solution in dimethylformamide).

2.2.3. Mammalian cell media

Hela and HEK293T cells (ATCC) were grown at 37°C in a 5% CO_2 atmosphere in Dulbecco's modified Eagle's medium high glucose (DMEM; Euroclone), supplemented with 10% fetal bovine serum (FBS; Euroclone), L-glutamine 2mM (Euroclone), 100 U/ml penicillin and 100 mg/ml streptomycin (Euroclone). While Opti-MEM Reduced Serum Media (Gibco®) medium was used in transfection protocol (section 2.9).

2.3 Plasmid

2.3.1. Yeast vectors

- *pGBKT7 Gal4 DNA-BD, the bait vector*

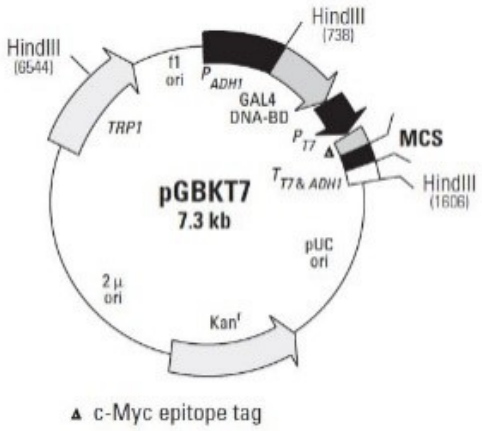


Figure 13. pGBKT7 map adapted from Clontech datasheet (PT3248-5)

pGBKT7 (Cat. 630489, Figure 13) is a yeast two-hybrid bait expression vector designed 7.3 kb length by Clontech's Matchmaker™ Systems. It expresses proteins fused to amino acids 1–147 of the GAL4 DNA-BD. In addition, the obtained bait proteins are also expressed as a fusion to a common c-Myc epitope tag. In yeast, fusion proteins are expressed at high levels from the constitutive ADH1 promoter (P_{ADH1}); transcription is terminated by the T7 and ADH1 transcription termination signals ($T_{T7\ \&\ ADH1}$). pGBKT7 also contains the T7 promoter, used in sequencing phase. The vector carries the Kan^r for selection in *E. coli* and the TRP1 nutritional marker for selection in yeast. pGBKT7 replicates autonomously in both *E. coli* and *S. cerevisiae* from the pUC and 2 µ ori, respectively[179].

- *pGADT7 Gal4 AD, the prey vector*

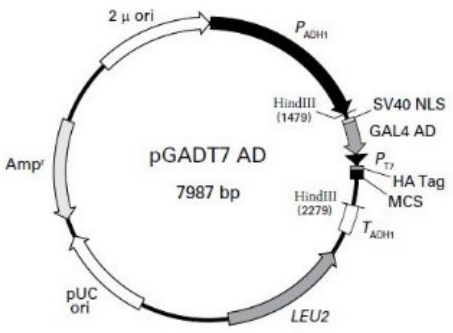


Figure 14. pGADT7 map adapted from Clontech datasheet (PT3249-5).

pGADT7 (Cat. 630442, Figure 14) is a yeast two-hybrid prey expression vector designed by Clontech's Matchmaker™ Systems. It expresses proteins fused to amino acids 768–881 of the GAL4 activation domain (AD). In addition, the obtained prey proteins are also expressed as a fusion to a common HA epitope tag, that allows the proteins to be easily detected with commercial HA-tag antibodies. In yeast, fusion proteins are expressed at high levels from the constitutive ADH1 promoter (PADH1); transcription is terminated by the T7 and ADH1 transcription termination signals (TT7 & ADH1). pGADT7 also contains the T7 promoter, exploited to sequencing phase. The vector carries the Ampr for selection in E. coli and the LEU2 nutritional marker for selection in yeast. pGADT7 replicates autonomously in both E. coli and S. cerevisiae from the pUC and 2 μ ori, respectively[179].

- *pGADT7-Rec, the library prey vector*

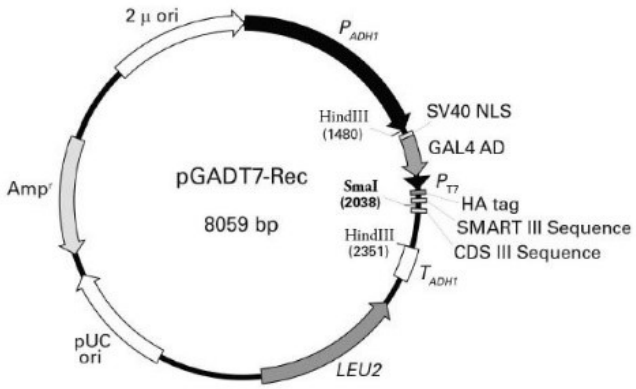


Figure 15. pGADT7-Rec map adapted from Clontech datasheet (102016).

Whether the *bait* is expressed by pGBKT7 Y2H vector, the *prey* plasmid is slightly different from pGADT7 Y2H vector. pGADT7-Rec (Cat. Sold as part of 630490, Figure 15) is engineered for the construction of GAL4 AD/cDNA libraries by homologous recombination in yeast in “Mate & Plate" Library System (Clontech). Similarly to pGADT7, it expresses proteins fused to amino acids 768–881 of the GAL4 activation domain (AD) and HA-tagged. In yeast, fusion proteins are expressed at high levels from the constitutive ADH1 promoter (P_{ADH1}); transcription is terminated by the T7 and ADH1 transcription termination signals ($T_{T7\ \&\ ADH1}$). pGADT7 also contains the T7 promoter, exploited to sequencing phase. The vector carries the Amp^r for selection in *E. coli* and the LEU2 nutritional marker for selection in yeast. pGADT7 replicates autonomously in both *E. coli* and *S. cerevisiae* from the pUC and 2 μ ori, respectively.

- *pBridge*^{TM}

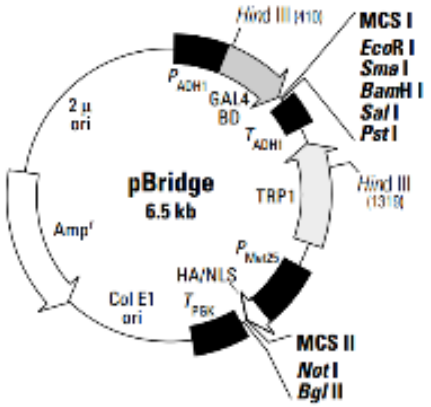


Figure 16. pBridgeTM map adapted from Clonotech datasheet (PT3212-5).

pBridge^{TM} (Cat. 630404, Figure 16) is a 6.5 kb vector, engineered for simultaneously expresses two proteins: a Gal4 DNA-binding domain fusion (DNA-BD; aa 1–147), and an additional protein. It thus allows to reconstitute a sort of three-hybrid systems when used in combination with an activation domain fusion vector for GAL4-based yeast two-hybrid systems[180]. Inserting the gene sequence of interest in the MCSI, the fusion protein is expressed in yeast host cells from the constitutive ADH1 promoter and transcription is terminated at the ADH1 transcription termination signal. The hybrid protein is targeted to the yeast nucleus by nuclear localization sequences (NLS). A second gene sequence can be then cloned into MCS II which is located downstream a second NLS. The resulting fusion protein is conditionally expressed from the Met25 promoter in response to methionine levels in the medium[181,182].

2.3.2. Mammalian cell vector: the pcDNA3.1(+)

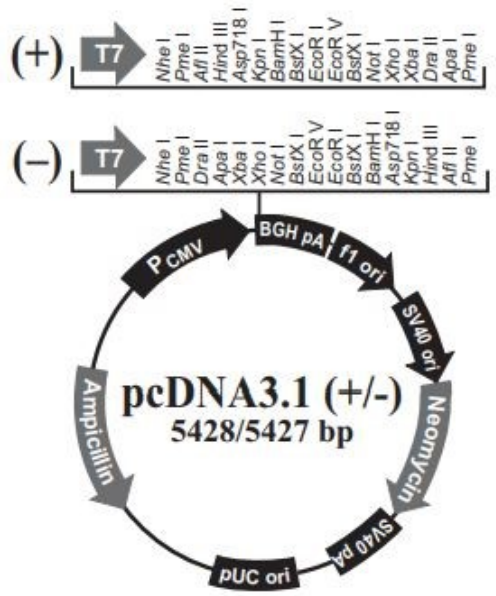


Figure 17. pcDNA3.1 family vectors. Both the forward-orientation pcDNA3.1(+) and the reverse-orientation pcDNA3.1(-) are represented in the image (ThemoFisher). In this work, the forward-orientation vector was used.

pcDNA™ 3.1(+) vector (ThemoFisher, V79020, Figure 17) is designed for high-level and constitutive protein expression in a variety of mammalian cell lines. It contains a Cytomegalovirus (CMV) promoter enhancer-promoter for high-level expression, a large forward-orientation (+) multiple cloning site and a Neomycin selectable marker. In addition, it has a SV40 origin for episomal replication and a Bovine Growth Hormone (BGH) polyadenylation signal and transcription termination sequence for enhanced mRNA stability. Ampicillin resistance gene and pUC origin are necessary for selection and maintenance in *E. coli*.

2.4 Cloning vectors

The described plasmids were then used to construct new vectors, employing the following *in vitro* approaches.

2.4.1. In-Fusion® protocol

In-Fusion® HD Cloning kit (Clontech; Cat. 011614) fuses DNA fragments (e.g., PCR-generated inserts and linearized vectors) efficiently by recognizing 15-bp overlaps at their ends. These 15-bp overlaps can be engineered by designing primers (Suppl. Table 1) for amplification of the desired

sequences. PCR reaction is performed using *CloneAmp HiFi* PCR Premix (Clontech, Cat. 63929). The general PCR parameters and reaction components used for amplification are reported in the following Table 6 and Table 7:

Components	***Volume***	***Final concentration***
Template	<100 ng	-
CloneAmp premix	12.5 µl	1X
Forward primer (10µ)	0.75 µl	0.2-0.3 µM
Reverse primer (10µ)	0.75 µl	0.2-0.3 µM
Nuclease free water	Up to 25 µl	-

Table 6. CloneAMp HiFi PCR components.

Temperature (°C)	***Step***	***Time***	***Number of cycles***
98	Denaturation	1 min	1
98	Denaturation	10 sec	30-35
55	Annealing	5 or 15 sec	
72	Extension	5 sec/kb	
72	Final extension	3 min	1
4	Cooling	forever	-

Table 7. CloneAmp HiFi PCR protocol.

Generally, PCR product contains ends homologous to a specific DNA regions surrounding the restriction enzyme site by which 3 µg of specific vector were linearized for 5h at 37°C and then purified by gel excision (section 2.7). After PCR purification (section 2.7), the In-fusion cloning was set up as follow (Table 8):

Components	***Volume***	***Time***
Purified PCR fragment	10-200 ng*	15 minutes at 50°C
Linearized vector	50-200 ng**	
5X In-Fusion HD Enzyme Premix	2 µl	
Nuclease free water	Up to 10 µl	

Table 8. In-Fusion reaction. In general, for optimal results under standard conditions use an insert to vector ratio of 2:1. *<0.5 kb: 10–50 ng, 0.5 to 10 kb: 50–100 ng, >10 kb: 50–200 ng; **<10 kb: 50–100 ng, >10 kb: 50–200 ng.

The reaction mixture was then transformed in bacteria, as previous reported in section 2.5. Recombinant plasmids were verified by Sanger sequencing, and subsequently the expression of the chimeric proteins was verified by immunoblot (section 2.13).

2.4.2. Ligation protocol

Alternatively, recombinant plasmids were obtained using the ligation method. The insert DNA (gene or fragment of interest) was cloned into a compatibly "sticky ends" digested vector backbone. Both insert and vector were obtained by digestion of 3 ug DNA at 37°C for 5h. The digested insert and the linearized vector were extracted from 1% agarose gel, purified using GenElute™ Gel Extraction Kit (Sigma, NA1010) and quantified using Nanodrop 2000 (Thermo Scientific). Ligation reaction was then catalyzes by the T4 DNA ligase (BioLabs, M0202). Reaction was set up as reported here (Table 9):

Components	***Volume***	***Time***
T4 DNA Ligase Buffer	2 µl	16 h at 16 °C
Vector DNA (4kb)	50 ng	
Insert DNA (1kb)	37.5 ng	
T4 DNA Ligase	1 µl	
Nuclease free water	Up to 20 µl	

Table 9. Example of T4 DNA Ligase (M0202) ligation protocol. Online NEBioCalculator calculate the correct ratio between vector and insert considering their sizes.

Using TOP10 *E. coli*, reaction mixture was transformed (section 2.5). Purified plasmid were then verified by Sanger sequencing.

2.4.3. Site directed mutagenesis

QuickChange II XL Site-directed mutagenesis Kit (Agilent Technologies) utilized a supercoiled double-stranded DNA (dsDNA) as template and two synthetic oligonucleotide primers, properly designed to contain a desired mutation. Thus, primers were each complementary to opposite strands of the vector leading to mutant amplification during a PCR reaction. The general PCR parameters used for amplification were reported in the following Table 10 and Table 11:

Components	*Volume*	*Final concentration*
Template	<10 ng	-
Reaction Buffer (10X)	5 µl	1X
Forward primer (10µM)	0.75 µl	0.2-0.3 µM
Reverse primer (10µM)	0.75 µl	0.2-0.3 µM
dNTP mix	1 µl	-
Quick Solution	3 µl	-
PfuUltra HF DNA polymerase	1 µl	2.5 U
Nuclease free water	Up to 50 µl	-

Table 10. QuickChange II XL Site-directed mutagenesis mix components.

Temperature (°C)	*Step*	*Time*	*Number of cycles*
95	Denaturation	2 min	1
95	Denaturation	20 sec	30
55	Annealing	30 sec	
65	Extension	30sec/kb	
65	Final extension	5 min	
4	Cooling	forever	1

Table 11. QuickChange II XL Site-directed mutagenesis PCR protocol.

Afterwards, the PCR product was treated with Dpn I digesting only the parental DNA template and selcting for mutation-containing synthesized DNA. Reaction was then transformed in bacteria (section 2.5). Mutated vectors were collected by DNA extraction (section 2.6) and verified by Sanger sequencing.

2.5 Bacteria transformation

One shot TOP 10 *E. coli* (Invitrogen, Table 1) were chemically competent cells (calcium chloride homemade protocol) that allow stable replication of high-copy number plasmids. 100 ng of a specific plasmid were added to 100 µl of competent bacteria and incubated on ice for 30 minutes[183]. Samples were exposed to heat shock at 42°C for 1 minute and again on ice for 2 minute. Afterwards bacteria were incubated at 37°C for 1 hour in 1 ml of LB medium. Finally cells were spread on LB-agar plate, supplemented with the proper antibiotic, i.e. ampicillin (AMP) 100 mg/ml or kanamycin (KAN) 50 mg/ml. The plates were incubated at 37°C O/N. Transformant colonies were used to perform plasmid purification, i.e. GenElute™ Plasmid Miniprep kit (Cat. PLN350, Sigma) or Pure Link™ HiPure plasmid Maxi prep kit (Cat. K210006, Invitrogen). The purified plasmid were quantified by

NanoDrop. Positive colonies were stocked in liquid LB supplemented with glycerol 20% (v/v) at -80° C.

2.6 Plasmid extraction

5 ml overnight recombinant *E. coli* culture was centrifuged and homogeneously resuspend with 200 µl of Resuspension Solution. 200 µl of Lysis Solution was added and samples were gently mixed by inversion (6-8 times) until the mixture becomes clear und viscous. Afterwards, cell debris were precipitated by adding 350 µl of neutralization solution gently mixing again by inversion. Meanwhile the floating particulates are centrifuged at ≥12,000g for 10 minutes, column were washed with 500 µl Column Preparation Solution. The cleared lysates were transferred into the column and centrifuged ≥12,000g for 1 minute. A wash step was performed adding 750 µl of Wash Solution and centrifuging ≥12,000g for 1 minute. The flow-through liquid was discard and two additional centrifuge were performed to remove ethanol excess. ddH_2O or Elution Solution was then used to eluted DNA by centrifuge at ≥12,000g for 10 minute. Protocol referred to GenElute™ Plasmid Miniprep Kit (Cat. PLN350, Sigma). When necessary, an up-scale plasmid extraction (i.e. maxi preparation) was performed referring to PureLink® HiPure Plasmid Purification Kit (Cat. K210006, Invitrogen).

2.7 DNA gel extraction and purification

DNA (i.e. linearized plasmid or PCR) is excised from a 1% agarose gel with a blade, minimizing the amount of agarose around the line. In a tared tube, the gel slice is weight. Three gel volumes of the Gel Solubilization Solution are then added to the gel slice and incubated at 55°C for 10 minutes, every 2-3 minutes the gel mixture is briefly vortexed. When adequately dissolved, add a volume of 100% isopropanol mixing until homogenous. To prepare the binding column, 500 µl of Column Preparation Solution are added centrifuging at ≥12,000g for 1 minute. Once the gel slice is completely dissolved, the mixture is loaded into the binding column and centrifuge at ≥12,000g for 2 minutes. A wash step is performed adding 700 µl of Wash Solution and centrifuging ≥12,000g for 1 minute. The flow-through liquid is discard and two additional centrifuge are performed to remove ethanol excess. For efficient recovery, Elution Solution is preheated at 65°C and then added to column membrane to eluted DNA by centrifuge at ≥12,000g for 10 minute. Protocol refers to GenElute™ Plasmid Miniprep Kit (Cat. NA1111, Sigma).

2.8 Yeast as a model

Yeast is a versatile and powerful tool for studying cell biology. Since 1988, *Botstein and others* speculate that yeast serves "*as a model for all eukaryotic biology derives from the facility with which*

the relation between gene structure and protein function can be established"[184]. As a model system, yeast has many advantages. It is cheap and simple to manipulate, characterized by a rapid growth compared to animal models and its genome has been completely sequenced[185]. Moreover, many signaling pathways are conserved between yeast and mammals. Thus, a wide array of molecular techniques are available to better understand diverse cellular processes, based on DNA microarrays, gene disruptions, protein localization, protein–protein interactions (PPI) and functional analysis by genetic interactions[186].

2.8.1. The yeast two-hybrid system (Y2H)

In this work, protein-protein interaction (PPI) are investigated through yeast two-hybrid assay by Matchmaker® Gold Two Hybrid System (Cat. 630489, ClonTech). As a genetic technique, the Y2H system allows detection of interacting proteins in living yeast cells[187]. The principle is based on the modular properties of eukaryotic transcription factor (TF), composed by two distinct domains: the DNA-binding domain (DBD) and the trans-activating domain (TAD). Although domains maintain their function independent of the presence of the other, the two-hybrid method exploits the fact that the DNA-binding domain is incapable of activating transcription unless physically, but not necessary covalently associated with an activating domain[188]. Considering these features, the TF distinct domains are fused to two different proteins or polypeptides to test whether they are able to interact. The most popular fusions use the N-terminal DBD and the C-terminal AD of the yeast TF Gal4[189]. It is part of a large fungal family of transcription factors, which positive regulates gene expression of galactose-induced genes[190]. The protein fused to the DBD is referred to as the 'bait', and the protein fused to the AD as the 'prey'. Upon interaction between the bait and the prey, the DBD and AD are brought in close proximity and the functional TF is reconstituted able to transcribe a reporter gene. Normally, this takes place in genetically modified yeast strains, in which the transcription of a reporter gene leads to a specific phenotype (i.e. growth on selective medium or colonies color change). The most popular reporter genes are HIS3 and LacZ, which respectively allow to select yeast on a medium lacking for histidine or to screen yeast in a colorimetric assay. Here, the HIS3 reporter gene is the one chosen to study novel interactions. HIS3 encodes a protein called imidazole glycerol-phosphate dehydratase, which catalyzes the sixth step in histidine biosynthesis. Whether bait and prey interact, HIS3 transcription allows the yeast cells to biosynthesize histidine and grow on selective medium (Table 4).

2.8.2. Yeast transformation

Exogenous DNA was introduced into intact yeast cells using one-step transformation method[191]. Briefly, an overnight (O/N) logarithmic phase culture was placed into 2 ml microtubes and centrifuged at 14,000 rpm for 15 minutes. Surnatant was discarded and 50 µg of boiled single-stranded DNA (DNA from Salmon Testes, Cat. D-1626, Sigma) were added to the pellet. Then, 1 µg DNA plasmid was added. Yeast were resuspend in 100 µl transformation solution (LiAc 2M pH 7.5, DTT 1M and PEG 4000 50% w/v). Lithium ions and heat shock promoted passage of DNA into the cell, while PEG promoted association of the transforming DNA with the surface of the cell[192]. Each tube was vortex and incubated at 45°C for 30 minutes. Hereafter samples were centrifuged, surnatant removed and pellet are resuspend in 200 µl NaCl 0,9% solution. Transformants were selected seeding yeast cells on selective media (section 2.2.2). Plates were incubated up-side-down at 30°C for 3-4 days. Generally, 3-5 positive colonies were spread patching out on selective media.

2.8.3. Yeast drop test

Three or more independent co-transformant colonies were tested using 10-fold serial dilutions on solid medium. In brief, interactions were detected spotting yeast cells on SD/-Leu/-Trp/-His medium supplemented with 3AT (3-Amino-1,2,4-Triazol) at different concentrations (i.e. 30 mM, 60 mM and 90 mM). Each plate contained a positive control strain, which co-expresses Gal4 AD-SV40 large T-antigen and Gal4 BD-murine p53 (fragment 72-390) and a negative control strain represented by empty vectors co-transformant yeast cells. In parallel, yeast cells were spotted on a permissive media (SD/-Leu/-Trp w/o 3AT) as healthy strain growth control. Plates were incubated up-side-down at 30°C and yeast growth was monitored for 8 days. Experiments were normally repeated three times, each one contains 3-5 independent colonies recovered from 3 independent transformations.

2.8.4. Yeast protein extraction

Protein extracts were obtained by TCA-based precipitation of yeast cells[193]. O/N yeast coltures were pelleted and were resuspended with 100 µl TCA 20% (trichloroacetic acid; Sigma). Cells were mechanically disrupted using glass beads (Cat. G8772, Sigma) and MagNA Lyser (Roche) for 30 sec at 6000 rmp. Then, surnatants were centrifuged at 14000rpm 4°C for 15 minutes after 500 µl TCA 5% addition. Obtained pellets were resuspended in 100 µl of sample buffer (0.1M Tris-HCl pH 8.8, 10% glycerol, 0.1M DTT, 2% SDS, 0,001% Blu Bromophenol) and incubated at 95°C for 5 minutes. Samples were quantified using Bradford assay[194] (Cat. B6916, Sigma).

2.8.5. Yeast library screening

The yeast library technology can be used to identify novel protein interactions, confirm putative interactions or define interacting domains. Similarly to Y2H assay (section 2.8.1), the principle is based on the *bait-prey* possible binding. Unlike, in a yeast two-hybrid screening a *library of prey* proteins is expressed to identify new binding partners of a chosen *bait* protein. The Matchmaker® Yeast Two-Hybrid (Clontech) offers a mate-and-plate library, constructed using homologous recombination-mediated cloning directly in the Y187 (MATa) yeast strain. The transformed cells are then mated to an opposite mating types Y2HGold strain (MATα) expressing a specific *bait* protein. Positive clone selection is highly stringent. Indeed, four different integrated reporter genes are under the control of three distinct Gal4-responsive promoters (i.e. G1, G2 and M1). These reporters are AUR1-C, HIS3, ADE2 and MEL1 (Figure 18).

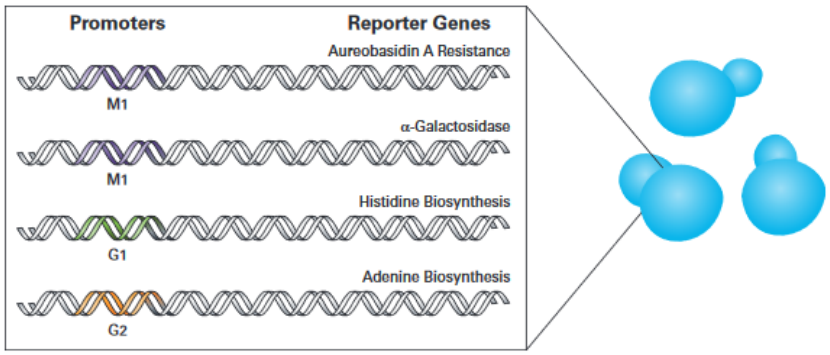


Figure 18. Interacting bait and prey fusion proteins drive the expression of four different reporters from three different GAL4-responsive promoters (M1, G1, and G2), stably integrated in Y2HGold genome. They confer growth selection in the presence of AbA and on histidine- and adenine-deficient media, while the α-galactosidase reporter produces blue colonies in the presence of X-alpha-Gal.

Briefly, AUR1-C is a drug reporter encoding the enzyme inositol phosphoryl ceramide synthase. In response to protein-protein interaction, it is expressed by Y2HGold strain leading to yeast survival. As already seen (section 2.8.1), HIS3 transcription activates histidine pathway leading to yeast survival on selective medium. Similarly, the ADE2 is the AIR-carboxylase enzyme[195] which catalyzes the sixth step of the purine (adenine) biosynthetic pathway that permit the yeast cells growth whether activate by proteins interaction. While MEL1 encodes a α-galactosidase (melibiase) secreted by yeast cells into the culture medium. In presence of the chromogenic substrate X-α-Gal, yeast colonies turn blue whether protein-protein interaction occurs. Only the simultaneous transcription of

these reporters allow the correct grow of diploid yeast cells containing a new protein-protein interaction (PPI).

Library protocol

Mating protocol was here divided in three phase: pre-mating, mating and post-mating phase.

I) The pre-mating phase

In the pre-mating phase, crucial steps were set up. Two-hybrid screening test was performed using a mate-and-plate human testis cDNA library (Cat. 630470, Clontech). Whereas the bait vector has been construct cloning the VLP encoding sequence in pGBKT7 using *in vitro* In-Fusion® method (section 2.4.1). The recombinant clone was correctly selected in bacteria and sequenced by Sanger method. After transformation (section 2.8.2) in Y2H Gold strain, the correct expression of VLP protein was determined by immunoblot (section 2.13). As a following step, it was imperative to test bait cells on library selective media to avoid auto-activation or reporter-independent growth. Thus, Y2H Gold expressing VLP was plated on SD/-Trp, SD/-Trp/X-α-Gal, SD/-Trp/X-α-Gal/AbA and DD/-Trp/X/A media. Thereafter a mating test has been performed to fine tune the protocol. A small amount of yeast library was scratched from the -80°C stored aliquot and patch out on SD/-Leu selective plates. Some cells were then picked up and inoculated in 50 ml SD-Leu liquid media, in parallel with 50 ml SD-Trp culture of *bait*-expressing cells and left grow at 30°C until 0.8 OD_{600} is reached. Both cultures were then centrifuge at 1000 g for 15 minutes and pellet were respectively resuspend in 1 ml SD-Leu and 5 ml SD-Trp to follow appropriately the mating protocol quantity. In a 2L flask, the two strains were co-culturing in 45 ml 2XYPDA liquid media at 30°C for at least 20h at slowly shaking (around 30-50 rpm). The overnight culture produced a set of diploid yeast clones, each co-expressing the *bait* protein with a different library *prey* protein. As suggested, after 20 hr a drop of co-culture was checked under a phase contrast microscope (40X) to observe or not the presence of zygotes. A zygote typically has a 3-lobed structure (Figure 19), which may resemble a cloverleaf or a “Mickey Mouse” face. If zygotes were not visible, it was necessary to continue the incubation. In this case, after 22h the mate culture has been collected. It has been centrifuge at 1000 g for 15 minutes and discard the supernatant. Different dilutions of co-culture were plated on SD/-Trp, SD/-Leu, SD/-Trp/-Leu and incubated at 30°C for 3-5 days.

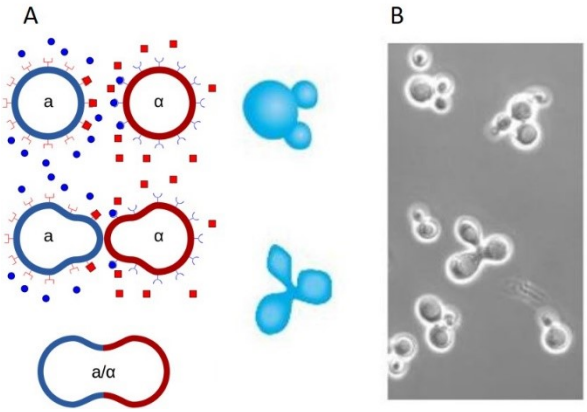


Figure 19. Yeast mating. Two haploid yeast of opposite mating (MATa and MATα) secrete pheromones, grow projections and mate (A). The obtained zygotes typically have a 3-lobed structure, similar to cloverleaf or "Mickey Mouse" face.

II) Mating phase

Using the set up protocol, 1 ml aliquot of cDNA human testis has been mated with an overnight culture of the *bait* strain. After 22h at 30°C, zygotes were present and the mated culture has been centrifuge at 1000 g for 15 minutes. The pelleted cells were resuspend in 10 ml of 0.5XYPDA liquid medium. An aliquot of the co-culture was used to spread 1/10, 1/100, 1/1000 and 1/10000 dilutions on SD/-Trp, SD/-Leu, SD/-Trp/-Leu. Incubated at 30°C for 3-5 days, plates were then used to calculate the number of the clones screened and the percentage of the mating efficiency. The remaining culture was plated on 150 mm SD/-Leu/-Trp/X/A (200 µl per plate) and incubated up-side-down at 30°C for 5 days monitoring colonies growth twice a day. Each colony has been marked by a progressive number.

III) Post-mating phase

All the blue positive colonies were patch out in parallel onto the higher stringency QDO/X/A and SD/-Leu/-Trp/Drop out agar plates. Whether the clones retains their blue color, they were inoculated in SD/-Leu/-trp/Drop out liquid media at 30°C O/N. From these cultures, plasmids were extracted using EUROGOLD Plasmid Miniprep Kit (Cat. EMR500200, Euroclone). Using GoTaq® Hot Start Polymerase (Table 12 and Table 13), a PCR was performed with 2-3 µl of extracted DNA, amplifying cDNA inserts with Gal4AD forward and reverse primers. PCR products were detected using 1.5 % agarose gel.

Components	*Volume*	*Final concentration*
Template	2-3 µl	-
Buffer	5 µl	1X
MgCl2 25mM	2.5 µl	1.0-4.0 mM
dNTPs mix 10mM	0.50 µl	5 mM
Forward primer (10µ)	0.75 µl	0.2-0.3 µM
Reverse primer (10µ)	0.75 µl	0.2-0.3 µM
Taq	0.25 µl	-
Nuclease free water	Up to 25 µl	-

Table 12. GoTaq® Hot Start Polymerase PCR mix components.

Temperature (°C)	*Step*	*Time*	*Number of cycles*
95	Denaturation	2 min	1
95	Denaturation	30 sec	30
55	Annealing	30 sec	
72	Extension	1min/kb	
72	Final extension	5 min	1
4	Cooling	forever	-

Table 13. GoTaq® Hot Start Polymerase protocol. Note that cDNA insert size are included in 0.5 kb to 4.0 kb size, thus normally 4-5 minutes are set for the extension step.

Hereafter, positive clones were amplified transforming 1-2 µl of extracted DNA in TOP10 *E. Coli* (section 2.5). Three colonies were stocked in LB supplemented with 20% glycerol at -80°C and two out of three were cut with EcoRI/BamHI restriction enzymes (NebCutter). Some peculiar clones showed more than one PCR product. In these cases, more bacteria colonies were screened out distinguish the two or more different population of plasmids. Finally, genuine positive clones were also confirmed using one-step transformation on Y190. Respectively, each selected clone was co-transformed with pGBKT7 empty vector (as auto-activation control) and with pGBKT7 VLP (as interaction control). For each transformation plate, a pool of colonies was patched out on SD/-Leu/-Trp/Drop out and SD/-Leu/-Trp/-His/30mM 3AT agar plates. To better distinguish the positive interactions, a spot test has been performed with all clones. Using T7 forward primer, samples were all sequenced by Sanger method.

2.9 Mammalian cell transfection

Twelve hours before transfection, mammalian cells were seeded onto 6- or 24-well plates and allowed to grow to 70%-80% of confluence. Using the Lipofectemine 2000 DNA transfection protocol (Invitrogen), cells were respectively transfected with 5 µg of total DNA to perform co-IP (section

2.10) or cellular fractionation (section 2.12), and 1 µg of total DNA for fluorescence microscopy experiments (section 2.11). After 5-6h post transfection, cell media was replaced with fresh media lacking for antibiotic mix. Plates were then incubated until samples collection.

2.10 Co-Immunoprecipitation

After 24h or 48h post-transfection onto 6-well plates at a density of $5{,}00x10^5$ cell/mm^2, samples were collected in 300 µl lysis buffer (20 mM HEPES-Na pH 7.4, 150 mM NaCl, 5 mM CHAPS)[196] supplemented with 1X protease inhibitors cocktail (PIC; Sigma). Lysates were centrifuged for 10 minutes at 600 rpm at 4°C and post-nuclear supernatant (PNS) are then collected. By immunoblot (section 2.13) the correct expression of tagged-transfected proteins was checked. To perform the co-immunoprecipitation, 5 µl of protein A magnetic beads (Pierce Thermoscientific) were washed once with ddH_2O, twice with TBS-T 0.05% and then pre-incubated with 2 µg of specific antibody (Table 14) 1h 500 rpm at RT. Washed three times with TBS-T 0.05%, beads were incubated with PNS for 4h at 500 rpm at 4°C. It could be useful to collect the flow-through surnatant. Beads were washed three times with lysis buffer and finally incubated for 5 minutes at 70°C in 30-35 µl of 1X NuPAGE LDS sample buffer (Invitrogen) supplemented with 0.1 M DTT. Using immunoblot method (section 2.13), both PNS and immunoprecipitate samples were detected.

2.11 Fluorescence microscopy in mammalian cells

After 24h post-transfection onto 13 mm glass coverslips at a density of $7{,}00x10^4$ cell/mm^{2*}, cells are fixed with 500 µl of 3.7% formaldehyde in phosphate-buffered saline (PBS; 140 mM NaCl, 2 mM KCl, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4, pH 7.4) for 20 minutes. Washed three times with PBS, coverslips are mounted using Mowiol 4–88 (Sigma) and observed at Leica TSC SP5 inverted confocal microscope using lasers wavelength of 488 nm and 100X. Images are acquired using the Leica AS software.

2.12 Cellular fractionation

After 24h post-transfection, cells were collected in 500 µl phosphate-buffered saline and centrifuged for 10 minutes at 1400 rpm 4°C. Pellet was resuspend in 200 µl of TEM buffer (20 mM NaCl, 20 mM Tris-HCl pH7.4, 5 mM $MgCl_2$, 0.1 mM EDTA) supplemented with 0.1% Triton X-100 and cells are mechanically lysate for 1 minute on ice using an electric dounce homogenizer. Samples were then incubated for 30 minutes on ice and centrifuge for 15 minutes at 16000 rfc at 4°C[197]. Surnatants were collected and pellet resuspend in O-solution (50mM Tris pH 6.8, SDS 2.3%, glycerol 2.3%). Respectively, they were considered the soluble and the insoluble part of cells lysis quantified by

Pierce™ BCA Protein assay kit (Cat. 23225, Thermo Scientific). Using immunoblot (section 2.13), 10-20 µg of each sample were detected.

2.13 Immunoblot

Proteins were loaded on a SDS-PAGE gel[198] in parallel with 5 µl of Prestained protein SHARPMASS™ VII (Protein Marker MW 6,5-270 kDa, Euroclone) as marker, after denaturating them for 5 minutes at 95°C. The electrophoretic course was led at 120 V in Trizma-Glycin-SDS Running Buffer 1X[199]. Hereafter samples were transferred in a Amersham™ Protan™ Premium 0.2 µm NC (Cat. 10600004, GE Healthcare) nitrocellulose membrane employing a constant voltage 350 mA for 90 minutes in Trizma-Glycin Transfert Buffer 1X. After transfer-phase, membrane were saturated using a milk-TBTS 5% solution for 1 hour at RT and then incubated in primary antibody diluted in BSA-TBTS 1% solution O/N at 4°C (Table 14). After three washing steps in TBST, membrane were incubated with secondary antibody diluted in BSA-TBTS 1% solution RT for 2 hour. The signal were revealed through Immobilon™ Western substrate (Cat. WI3KLS0500, Millipore) on UVITEC (Eppenford).

Antibody	***Dilution***	***Company***
Anti-HA clone HA-7	1:10000	Sigma (H9658)
Anti-Myc tag	1:4000	Cell Signaling (2276)
Anti-DDDDK tag	1:2000	Abcam (ab45766)
Anti-FLAG	1:4000	Sigma (F7425)
Anti-VHL (FL-181)	1:1000	Santa Cruz (sc-5575)
Anti-GFP (D5.1) XP®	1:1000	Cell Signaling (2956)
Anti-βActin	1:1000	Sigma (A5316)
Anti-Mouse IgG (whole molecule)-Peroxidase	1:20000	Sigma (A9044)
Anti-Rabbit IgG (whole molecule)-Peroxidase	1:10000	Sigma (A6154)

Table 14. Antibodies for immunoblot experiments.

2.14 Computational analysis

2.14.1. Interaction network and sequence feature analysis

VHL30 (P40337), p21 (P38936), p27 (P46527) and p57 (P49918) protein sequences were retrieved from UniProt[200] selecting the canonical sequence and visualized with Jalview[201]. Multiple sequence alignment was performed using T-Coffee[202] with default parameters. Disorder content was assessed with MobiDB[203,204] and DisProt[205], while functional domains were identified by Pfam[206] and

InterPro[207] scanning. A protein-protein interaction network centered around pVHL and CDKNs was built using Cytoscape[208] and importing interaction data from STRING[209], excluding text-mining and neighborhood interactions. The first shell of interactors was populated with VHL30, p21, p27 and p57 selecting the corresponding human proteins, and no more than 20 interactors were chosen for the second shell. The default interaction score and confidence parameter (0.400) were used.

2.14.2. Molecular dynamics simulations

The 1.8 Å crystal structure of VHL30 (PDB identifier: 1LM8)[210] and the 2.3 Å structure of p27 (PDB identifier: 1JSU)[211] were used as starting models. The VHL30/p27 complex were constructed by homology modeling using Modeller through superimposition to the VHL30/HIF-1α complex. All simulations were carried out with GROMACS[212] using the CHARMM27 force field and the TIP3p explicit solvent model. All simulation runs consisted of 100 conjugate gradient minimization steps, 100 ps in NVT conditions, and 50 ns of classic molecular dynamics simulation at 310 K and 1.01325 bar. Integration was based on the Verlet method[213] using a 2 fs time step. Trajectories were compared in terms of RMSD and root-mean-square fluctuation (RMSF). RING 2.0[214] was used to estimate variation in residue interaction network with strict distance thresholds and the "one interaction" option.

2.14.3. Selection and interpretation of mutations putatively affection VHL30/p27 interaction

Considering both sequence conservation with CODD motif and their specific association with cancer, the following cancer related p27 mutations, putatively affecting the VHL30/p27 interaction, were retrieved from COSMIC[215]: P35L, D37N, E40K and T42A. The first three localize in or are immediately close to conserved positions, while p.T42A has been chosen as it was found in patients with hematological malignancies[216]. Due to the novelty of the VHL30/p27 interaction and a lack of specific bibliographic data, both driver and passenger mutations affecting the CODD-like region were included in this study. Mutations were placed with Bluues[217] starting from the last MD simulation frame after removal of solvent and ion molecules. Bluues was also used to predict the electrostatic properties of wild-type and mutant p27. Stability variation was predicted for each mutant with NeEMO[218]. Mutations on the pVHL/p27 interaction were investigated predicting the electrostatic ΔGbind values with APBS[219], a method which evaluates the energy difference between the solvated unbound interactors and the solvated complex.

2.14.4. Homology modelling of pVHL172 and VLP

The crystal structure of pVHL bound to the CODD fragment of HIF-1α (PDB code: 1LM8) has been selected as a template to model both VHL172 and VLP structure. The human VHL172 and VLP sequences (accession code: and Q6RSH7) have been retrieved from Uniprot[200] and aligned with Jalview/T-Coffee using standard parameters[202]. Then models were built using SWISS-MODEL[220] and superimposition performed using Chimera[221]. Models quality was assessed with QMEAN[222] and TAP programs[223].

2.14.5. cDNA fragment analysis

The library screened cDNA fragments were analyzed by tBLASTx tool (https://blast.ncbi.nlm.nih.gov) using default parameters, and filtered for Homo sapiens-taxid 9606 against either the nucleotide collection (nr/nt) or Human RefSeq Gene dataset (RefSeq_Gen). Literature sub-localization data were then collected from Swiss-Prot (annotated group of interacting fragments) and enriched by manual mining of literature. A functional interaction network was derived from STRING[209] and expanded including *interactors-of-interactors*. In detail, query proteins were maintained in the first shell and >50 in the second shell imposing confidence of 0.40 and text-mining derived data included in the analysis. InterProScan[224] was used to detect the functional domains present in the interacting clones.

Results

Introductory Note

My PhD activity was first aimed at identifying and characterizing novel pVHL interactors pairing computational and experimental approaches. As first case study, I predicted, investigated and validated the association between pVHL and the cyclin-dependent kinase inhibitor 1 subfamily. Resulting data were presented in a scientific paper and described as following.

Results

3.1 Novel interactions of the von Hippel-Lindau (pVHL) tumor suppressor with the CDKN1 family of cell cycle inhibitors

Giovanni Minervini[1,*], Raffaele Lopreiato[1,*], Raissa Bortolotto[1], Antonella Falconieri[1], Geppo Sartori[1], Silvio C.E. Tosatto[1,2]

[1]Department of Biomedical Sciences, University of Padova, Viale G. Colombo 3, 35121, Padova, Italy; [2]CNR Institute of Neuroscience, Padova, Viale G. Colombo 3, 35121, Padova, Italy; [*]contributed equally

Acknowledgments

The authors are grateful to Fiorella Tonello for help with the co-IP protocols. This work was supported by Associazione Italiana per la Ricerca sul Cancro (AIRC) grants MFAG12740 and IG17753 to ST. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Author contributions

ST, GM, and RL conceived the experiments. GM, RB, and AF performed the experiments. GM, RL, and GS analyzed the data. GM, RL, and ST wrote the manuscript.

Abstract

Hypoxia is a common feature shared by the most active tumors, characterized by unregulated development and malignant progression[225]. As described in the introduction, the deregulation of this network is known to characterize the inherited von Hippel-Lindau syndrome[226]. Beyond neovascularization, hypoxia also plays a relevant role in regulating cell cycle progression and senescence[227], inducing growth arrest in cells exposed to prolonged oxygen deprivation. Hypoxia-dependent cellular senescence has a critical role in normal tumor suppressor response[227,228], modulating early malignant transformation[229] and drug resistance[228]. In particular, hypoxic induction of cell cycle arrest is linked to HIF-1α dependent transcription of the cyclin-dependent kinase inhibitors p21 (CDKN1A)[230] and p27 (CDKN1B)[231]. Together with p57 (CDKN1C), these form the kinase inhibitor family playing important roles in the negative regulation of the cell cycle[232]. Especially, p21 mediates G1 growth arrest[233] and its transcription is mainly regulated by the tumor suppressor axis p53-Mdm2 in response to DNA damage[234]. HIF-1α dependent transcription seems to regulate the activation of a specific genetic program to slow down the cell cycle in a p53-independent way leading to S phase cell cycle progression during hypoxia rather than apoptosis[231]. In parallel, the role of p53 in the regulation of HIF-1α is intriguing. Under prolonged hypoxia, p53 accumulates in the cell yielding repression of HIF-1α transcriptional activity in a pVHL-independent way[176,178]. Functional connections between hypoxia response and cell cycle regulation are also sustained by recent evidence linking pVHL and p14ARF[235,4], a modulator of Mdm2 function[236] arising from an alternative reading frame product of the CDKN2A locus encoding the p16ink4a1 tumor suppressor[237]. In the light of these considerations, I decided to investigate a possible novel and structural interaction between pVHL and the CDKN1 inhibitor family paring computational and experimental assays.

In silico investigation of the hypoxia response and apoptotic pathways suggests that pVHL and the CDKN1 protein family could be physically linked

An interaction network centered on the proteins pVHL, p53 and CDKN1s was generated with STRING[209] (Figure 20). Note that in literature, two almost functionally overlapping pVHL isoforms are reported[43–238] (pVHL30 and pVHL19, respectively). Here, both isoforms are collectively referred to as pVHL where not explicitly mentioned in the text. The resulting network is composed of 28 nodes connected by 155 edges, with an average node degree of 11.1 and a clustering coefficient of 0.671. The expected number of edges for a similarly populated network composed of random nodes is estimated to be 87. This finding suggests the proteins forming the network should be at least

partially biologically connected as a group. The network can be divided into six different clusters representing different biological processes (Figure 20).

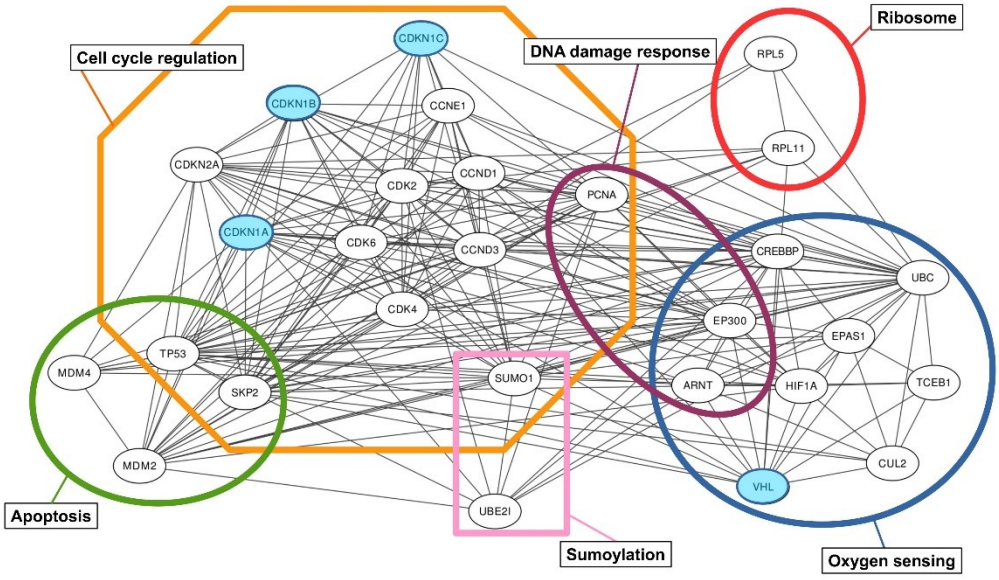


Figure 20. Prediction of pVHL/CDKN1 interactions. Protein-protein interaction network generated with STRING[209]. Functional connections between the pVHL and proteins involved in cell cycle regulation. Connections between nodes represent experimental evidence for interaction. Colored boxes group proteins participating in the same pathway or sharing similar function, e.g. ribosomal proteins. pVHL and CDKN1 are marked in blue.

Proteins directly involved in cell cycle regulation (e.g. the CDKN1 family) form the largest cluster, with 12 nodes. Three smaller clusters (4, 3 and 2 nodes) account for proteins involved in apoptosis, DNA damage response and sumoylation and share nodes with the largest cluster. The connection among these clusters is expected considering the functional role of the CDKN1 family in regulating cell cycle progression[232]. Less obvious is their connection with the second largest cluster (9 nodes), rich in proteins involved in oxygen sensing and DNA damage response. Functional enrichment in GO[239] terms shows "regulation of transcription from RNA polymerase II promoter in response to hypoxia" (GO:0061418) as the biological process best describing the entire network (count in gene set 10, false discovery rate 4.66 e^{-18}). Interactions between pVHL and RNA polymerase II subunits are already known[114,240]. In general, the functional connection with RNA polymerase II subunits seems to have multiple functional roles. pVHL is thought to modulate Rpb1 expression and is necessary for the oxidative-stress-induced interaction of Rpb1 with DNA[114], as well as to suppress hsRPB7-induced VEGF promoter transactivation[240]. Data together support a strong functional link

between oxygen sensing and cell cycle regulation. Further, pVHL was also proposed to mediate a HIF-1α independent senescence program[241], with p27 being upregulated in pVHL null cells. Since large interactome studies of binary protein-protein interactions reveal novel functional interactions among interactors-of-interactors[242], I decided to address the possible connection between pVHL and the CDKN1 family experimentally.

pVHL is able to interact with all CDKN1 proteins in Y2H and co-IP experiments

The protein-protein interaction (PPI) between pVHL and members of the CDKN1 family (p21, p27, and p57) has been investigated by the genetic two-hybrid system in yeast cells (Y2H). In particular, I used Y190 yeast strain to co-transform bait- and prey-expressing vectors (section 2.8). In detail, the coding sequences of the CDKN1 proteins were fused to Gal4 DNA-Binding Domain (BD, bait), while the pVHL protein was inserted in-frame to Gal4 Activation Domain (AD, prey). Upon checking that the fusion proteins were both properly co-expressed in yeast cells (Figure 21), I then selected three or more independent colonies to assay the interaction using the drop test (section 2.8.3).

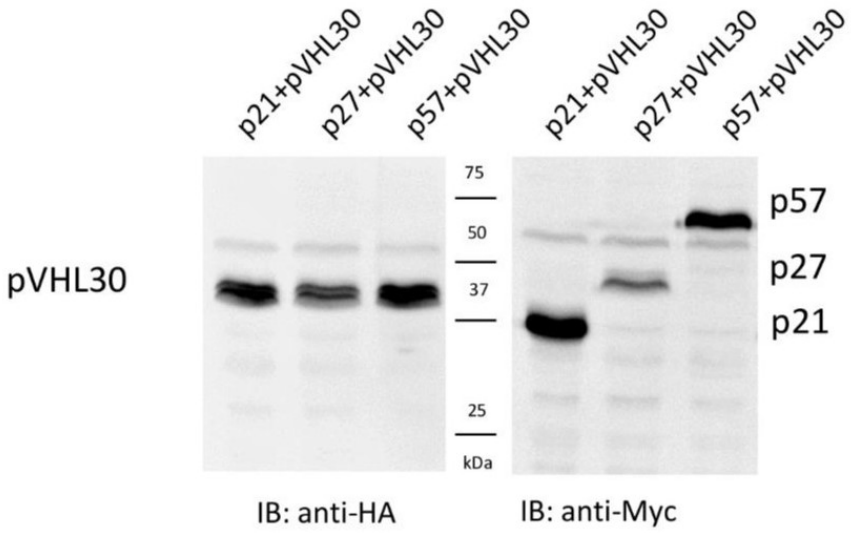


Figure 21. Western blot analysis of total protein lysates. Yeast cells co-expressing pVHL30 (HA-tagged), and each CDKN1 (Myc-tagged), as Gal4-fusion proteins are shown. Membranes were immunoblotted with either anti-HA (left), and anti-Myc (right) antibodies, revealing the presence of fusion proteins at the expected molecular weight, as indicated. Lane 1: p21-pVHL30; lane 2: p27-pVHL30; lane 3: p57-pVHL30.

As shown in Figure 22, clones were in parallel spotted both on permissive and selective media. In particular, interactions have been determined as the ability to support the growth of yeast cells in absence of histidine (supplemented with the 3AT inhibitor) activating the Gal4-mediated transcription

of a *HIS3* reporter gene (Figure 22). In all experiments, positive (C+) and negative (C-) internal controls were spotted (Figure 22). Provided by the manufacturer, the positive strain simultaneously expressed Gal4 DBD-murine p53 and Gal4 AD-SV40 large T-antigen representing a strong level of binding. While the negative control carried the Y2H empty vectors resulting as a reference of binding absence. Yeast cells expressing either p21, p27, or p57 alone in a selective medium almost failed to grow, as well as with pVHL30 only (Suppl. Figure 1 and Suppl. Figure 2). These were reported as protein-related controls necessary to avoid interaction-independent activation of the HIS3 reporter gene. Whereas in presence of pVHL30, cell growth was markedly improved indicating that the proteins were able to associate *in vivo*. Interestingly, these data were obtained in absence of prolyl hydroxylase 3 (PHD3), necessary for the recognition of HIF-1α by VCB complex. Lacking for many eukaryotic proteins, the used of the yeast model suggests a PHD3-independent interaction. Co-immunoprecipitation (Co-IP) experiments in human cells were performed to confirm the interactions in a more physiological context. Thus, a series of plasmids able to over-express either HA-tagged pVHL30 protein or the Flag-tagged CDKN1 proteins in mammalian cells were constructed (section 2.3.2 and 2.10). These recombinant plasmids were used to transiently transfect HEK293T mammalian cells and perform Co-IP assays from total cell lysates using anti-VHL antibody.[243,47] As a result (Figure 22, bottom part), all three CDKN1 proteins were able to maintain the interaction with pVHL30, as demonstrated by their presence in the immunoprecipitate (IP) revealed with the anti-Flag antibody (*bottom panels*). Taken together, experiments show that these proteins can form at least binary complexes in human, and notably kidney cells. Furthermore, the qualitative comparison among the three cyclin-dependent kinase inhibitors showed a similar interaction trend to pVHL suggesting a shared putative interaction site.

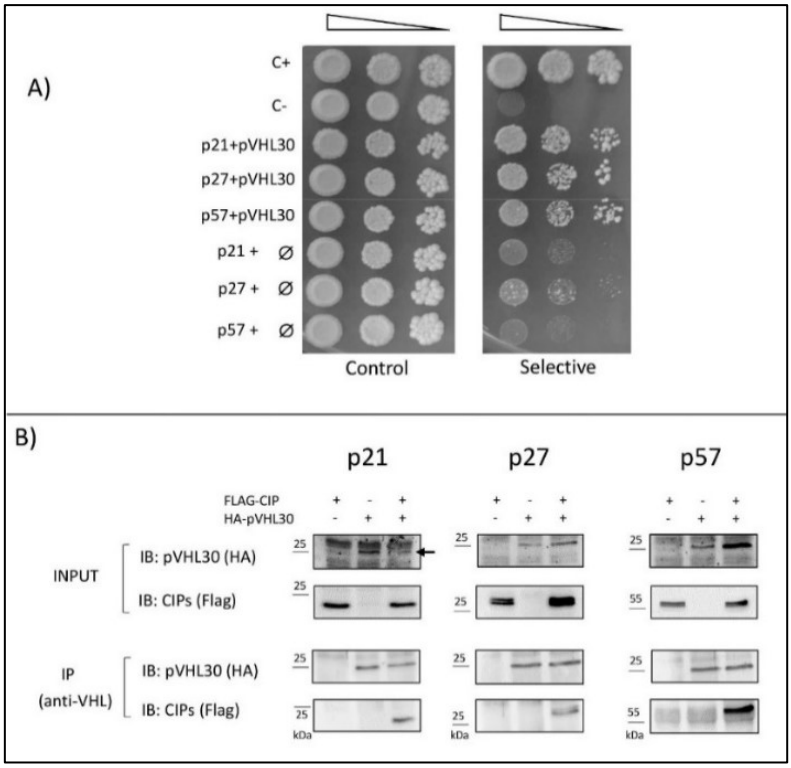


Figure 22. Validation of pVHL/CDKN1 interactions. (A) Serial dilutions of yeast cells were spotted on both permissive (left) and selective (right) media, and incubated for 8 days at 30°C. C+ and C- are the positive and negative controls, while ∅ indicates an empty vector (i.e. auto activation control). The image is representative of four independent experiments, each with 2-3 different clones analyzed. (B) Human HEK293T cells were transiently transfected with plasmids overexpressing Flag-tagged CDKN1 proteins and/or HA-tagged pVHL30 protein, as indicated on the top row. Recombinant proteins have been revealed in total cell lysates (Input) by immunoblotting with either anti-HA, or anti-Flag antibodies. Upon pVHL immunoprecipitation with a specific antibody, presence of the CDKN1 proteins in the immunoprecipitates (IP) was finally verified using the anti-Flag antibody (bottom panel). In the p21 panel (Input), an arrow indicates the band corresponding to pVHL30, which is partially confused by surrounding unspecific signals.

A conserved motif in the CDKN1 domain resembles HIF-1α, and MD simulations suggest the pVHL β-domain to drive the interaction

A closer investigation of multiple amino acidic sequence alignment of the CDKN1 members partially sustained the hypothesis of a common interaction site. Indeed, they share a conserved 48-residue domain located at the N-terminus identified in Pfam as the CDI protein family (pfam02234). The alignment showed relevant sequence similarity (64%) with the CODD[210] motif of HIF-1α, a value high enough to assume conserved features among the sequences. In particular, the CDI seems to be chemically compatible with amino acids 556-574 of HIF-1α (Figure 23), a functional region responsible for the interaction with pVHL[210]. In particular, in the blue square the highlighted region (Figure 23B) shows similar residues among the aligned sequence. Residue P (proline) is highly

conserved, whereas the other residues are similarly hydrophobic (i.e. leucine and methionine), non-polar (i.e. alanine and leucine) or charged.

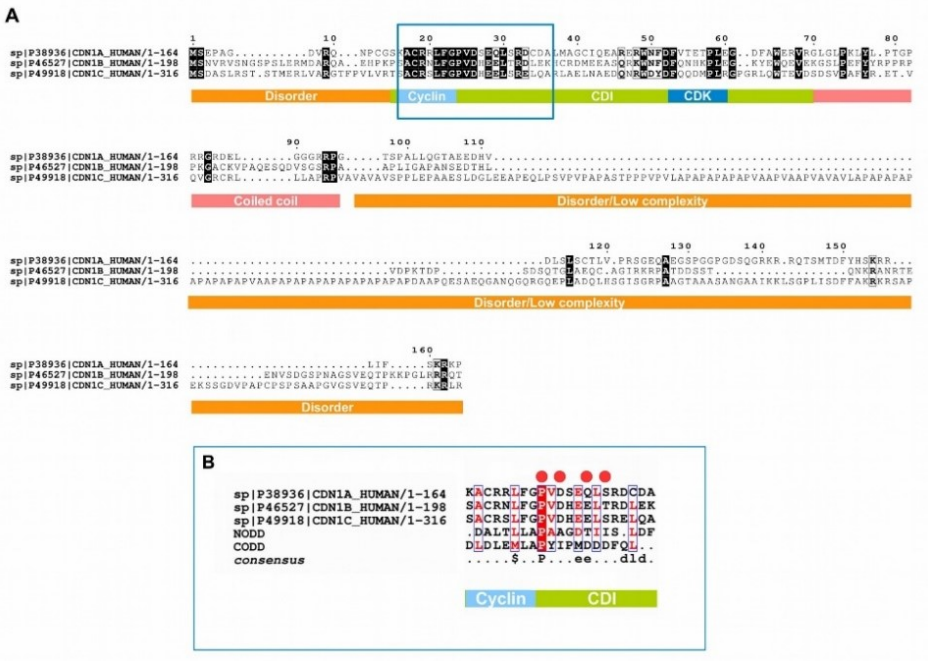


Figure 23. Overview of the CDKN1 sequence features. (A) Multiple sequence alignment of the CDKN1 protein family. The functional motif organization is presented as a colored bar below. (B) Close-up of sequence conservation between the CDI domain and NODD/CODD motifs of HIF-1α. The CODD-like region localizes to the CDI domain (green bar), partially overlapping the cyclin A recognition element (light blue). Red dots represent cancer-related mutations in the CODD-like region of p27. The consensus sequence is shown below, with $ used for hydrophobic residues and e for charged residues.

Data in the literature[244,245,205], as well as secondary structure prediction of CDI, performed with FELLS[246] (Figure 24) suggested this short segment to adopt random coil structure when not in complex to cyclin A. Combining these findings, CDI may sustain the interaction with pVHL in a HIF1α like fashion.

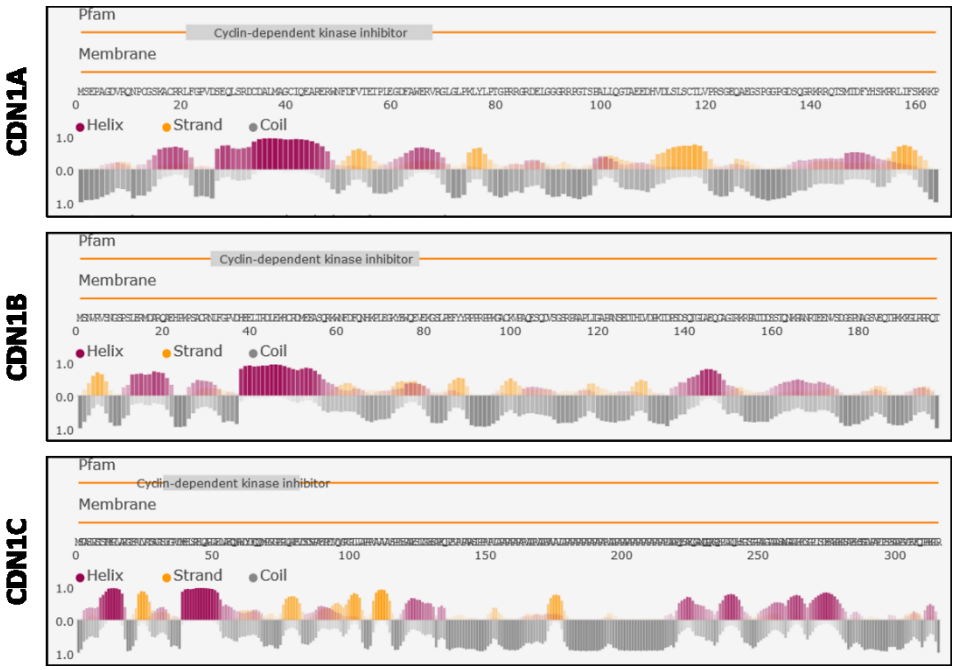
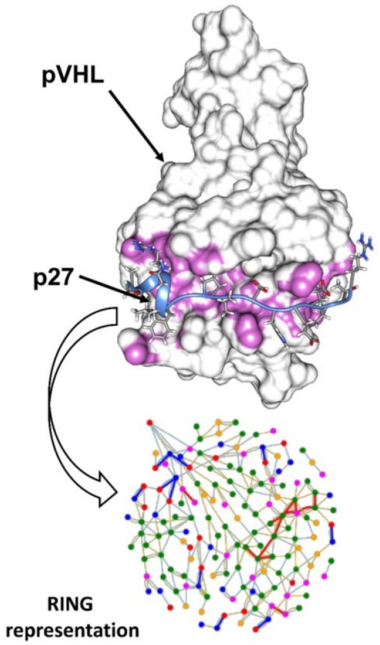


Figure 24. FELLS analysis. CDN1A_human, CDN1B_human and CDN1C_human sequences were analyzed using FELLS tool[246].In particular, the conserved 48-residue domain located at the N-terminus were characterized by random coil structure propensity when not in complex to cyclin A.

Furthermore, p27 crystal structure (PDB identifier: 1JSU) has been inspected. Residues 25-37 partially overlaps the putative CODD motif forming a short coil binding cyclin A. This finding reinforced the idea that this segment may adopt an extended conformation able to interact with pVHL under physiological conditions. Fascinatingly, four cancer-related p27 mutations were found to affect the CODD-like region[215]. Whereas multiple experimental pieces of evidence describe the pVHL/CODD interaction[210,40], no structural data is available for the interactions investigated here. The corresponding p27 N-terminal region (p27-NT or p27-CODD-like, residues 27-51) was modeled, using the HIF-1α CODD crystal structure (PDB code: 1LM8, chain H) as a template. The obtained model was used to perform molecular dynamics (MD) simulations to investigate whether such a motif is able to sustain the interaction. Our simulations showed p27-CODD-like bound in a stable way to pVHL after 50 ns at 310K. The p27-NT backbone was predicted to assume a β-sheet-like conformation binding the fourth pVHL β-strand through two hydrogen bonds with pVHL residues I109 and H110. The interaction was further sustained by a salt bridge between p27-E40 and pVHL-R107 as well as an additional van der-Waals interaction between the pVHL-H115 backbone and p27-G34. Additional investigation with RING 2.0[214] showed that p27-E40 is able to form interactions with other residues on the same interface of the pVHL β-domain (Figure 25). The MD simulations also predicted the interaction to be stable in the absence of Elongin B and Elongin C, suggesting a

possible proteasome-independent function. Although MD simulations are not necessarily representative of physiological conditions, the *in silico* data collectively suggest binding between pVHL and the CODD-like domain of the CDKN1 proteins.



Residue p27	Type of interaction	Residue pVHL
H:29 CYS	VDW:MC_SC	V:91 PHE
H:30 ARG	VDW:SC_SC	V:90 ASN
H:30 ARG	HBOND:SC_SC	V:92 ASP
H:33 PHE	VDW:SC_SC	V:69 ARG
H:33 PHE	VDW:SC_SC	V:91 PHE
H:33 PHE	PIPISTACK:SC_SC	V:112 TYR
H:33 PHE	VDW:MC_SC	V:112 TYR
H:34 GLY	VDW:MC_SC	V:115 HIS
H:35 PRO	VDW:SC_SC	V:88 TRP
H:35 PRO	VDW:SC_SC	V:98 TYR
H:35 PRO	VDW:SC_SC	V:115 HIS
H:35 PRO	VDW:SC_SC	V:117 TRP
H:36 VAL	HBOND:MC_MC	V:110 HIS
H:36 VAL	VDW:SC_SC	V:110 HIS
H:37 ASP	VDW:SC_SC	V:109 ILE
H:38 HIS	HBOND:MC_SC	V:107 ARG
H:38 HIS	HBOND:MC_MC	V:108 ARG
H:38 HIS	VDW:SC_SC	V:108 ARG
H:38 HIS	PIPISTACK:SC_SC	V:110 HIS
H:38 HIS	VDW:SC_SC	V:110 HIS
H:39 ASP	IONIC:SC_SC	V:108 ARG
H:39 ASP	HBOND:SC_SC	V:108 ARG
H:39 ASP	VDW:SC_SC	V:108 ARG
H:40 GLU	HBOND:SC_SC	V:107 ARG
H:40 GLU	VDW:SC_SC	V:107 ARG
H:40 GLU	HBOND:MC_SC	V:108 ARG
H:41 LEU	VDW:SC_SC	V:75 ILE
H:41 LEU	VDW:SC_SC	V:77 CYS
H:41 LEU	HBOND:MC_MC	V:106 GLY
H:41 LEU	VDW:SC_MC	V:106 GLY
H:41 LEU	VDW:SC_SC	V:108 ARG
H:41 LEU	VDW:SC_SC	V:148 PHE
H:42 THR	VDW:SC_MC	V:106 GLY

Figure 25. Binding model of pVHL/p27 after 50 ns of MD simulations. Residues of pVHL forming the interaction surface are shown in purple and a cartoon representation is used for p27-CODD like region (residues 21-42). Below the interaction network representing residues of interacting each other, while interaction types are described on the right side.

In vivo dissection of the pVHL-CDKN1 interacting regions

The regions involved in binding were mapped using the Y2H system to verify the computational results. The pVHL30 protein was first dissected in three parts, N-terminal disordered tail (residues 1-53), β-domain (54-157), and α-domain (158-213). Mutant plasmids expressing the different pVHL fragments were generated and yeast cells transformed to determine binding with each CDKN1 protein. Our results show the pVHL N-terminus not interacting, as yeast cells expressing it together with any CDKN1 protein are unable to grow in selective medium (Suppl. Figure 3). The pVHL β-domain is able to bind any member of the CDKN1 family, as indicated by yeast cell growth (Suppl. Figure 3). The results also confirmed that the pVHL19 isoform, lacking the N-terminus is able to interact with all CDKN1 proteins like the pVHL30 protein (Suppl. Figure 3). The data also suggests

that the pVHL α-domain, while not strictly required, may be important for proper CDKN1 binding, possibly by stabilizing the pVHL structure (Suppl. Figure 3). On the other side, the CDKN1 region involved in pVHL binding was mapped based on the conservation shared by the three CDKN1 sequences previously observed. Thus, Y2H plasmids expressing either the three similar CDKN1 N-terminal tails containing the CDI domain (p27-NT residues: 1-60; p21-NT: 1-49; p57-NT: 1-61) or the corresponding leftover C-terminal moiety (p27-ΔN residues: 61-198; p21-ΔN: 50-164; p57-ΔN: 62-316) were also generated. As shown in Figure 26, loss of the p27 N-terminus completely disrupts its ability to associate with pVHL30, as yeast cells expressing the C-terminus of p27 were all unable to grow in selective medium. Similar data have been also obtained for both p21 and p57 (Suppl. Figure 4), strongly supporting the notion that regions responsible for CDKN1 binding to pVHL30 are included in CDKN1 CDI domains. Thus, I directly assayed the interaction between pVHL30 and the N-terminus of each CDKN1 protein. As shown in Figure 26 (Suppl. Figure 4), the N-terminus of p27 was able to sustain the growth of pVHL30-expressing yeast cells even more efficiently than full-length p27 protein. Similar results have been also obtained for the N-terminus of p57 (Suppl. Figure 4). Although at sequence level all CDKN1 N-terminus appeared highly similar, data on the p21-NT fragment cannot be considered due to an unspecific activation of the reporter gene (i.e. auto-activation, Suppl. Figure 4). Together results suggest that removal of the CDKN1 C-terminus may have positive effects on pVHL binding, as judged by increased yeast cell growth in selective medium expressing the N-terminus.

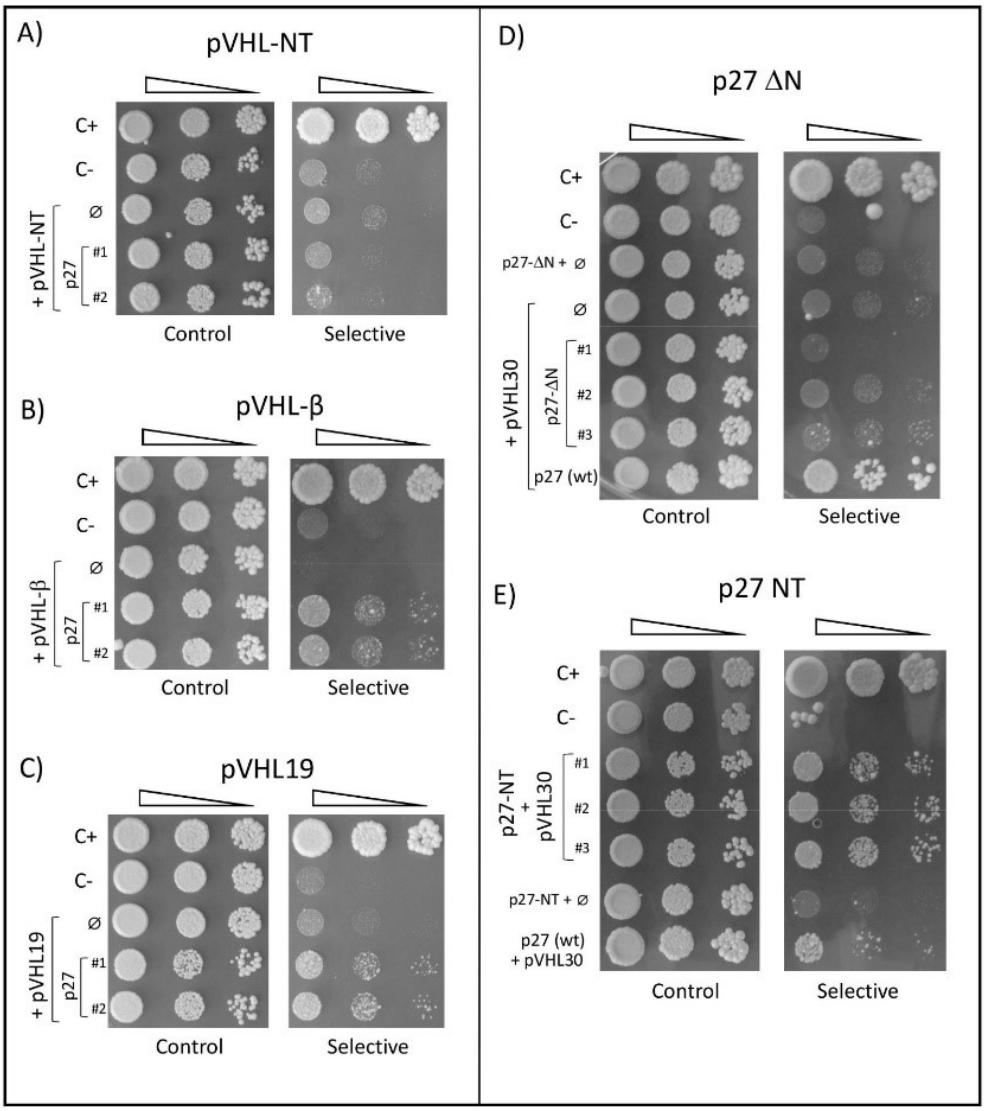


Figure 26. Y2H dissection map of pVHL binding to p27. Yeast two hybrid (Y2H) assays are shown of pVHL binding to p27, including fragments of either protein. The assayed interaction is shown on the left, with Ø used for an empty vector (i.e. negative control). C+ and C- are entirely positive and negative controls. The image is representative of four independent experiments, each with 2-3 different clones analyzed. Plates were incubated up-side-down for 8 days at 30°C (A) Y2H assay of pVHL-NT (residues 1-53) with p27 shows no growth on selective medium. (B) Y2H assay of pVHL-β (residues 54-157) with p27 shows growth on selective medium. (C) Y2H assay of pVHL19 (residues 54-213) with p27 is indicative of their binding. (D) Removal of the first 60 residues of p27 (p27-ΔN) abolishes interaction with pVHL30. In this case, plates were incubated at 30°C for longer time to confirm the absence of yeast growth. (E) pVHL30 were tested for binding by Y2H assay with the p27 N-terminus (NT, residues: 1-60). On selective medium, p27-NT yeast cells display increased growth rate with respect to the full-length protein (WT), possibly reflecting a negative impact of the p27 C-terminus on pVHL binding. These images represent only 4 days of incubation at 30°C.

Pathological p27 mutations may influence pVHL binding

As already highlighted, four cancer-related p27 mutations were found to affect the CODD-like region shared by all CDKN1 proteins. Consequently, I further inserted several N-terminal p27 missense mutations by site-directed mutagenesis reaction on pGBKT7 p27 full-length. The substitutions P35L, D37N, E40K and T42A (Figure 23) have been selected considering the conservation of the CODD motif, their position on the interaction surface as well as their pathological relevance in cancer[215], i.e. hematological malignancies compatible with the deregulation of the pVHL/HIF-1α axis. Notably, none of these had been characterized with respect to the normal function of p27 as a CDK inhibitor. Before the *in vivo* analysis, *in silico* replacements have been performed. Many times a single amino-acid change has been sufficient to alter protein structure, binding region, and function[241]. For the evaluation of stability changes on mutants the NEtwork Enthalpic MOdelling (NeEMO)[218] tool has been used. Both D37N and P35L were predicted to stabilize (or have a negligible impact) on the pVHL/p27 binding moiety (Table 15), while the contrary is predicted for E40K and partially for T42A.

Variant	Δ_{bind}G kJ/mol	Δ_{bind}G kcal/mol	NeEMO kJ/mol	Bluues kJ/mol
Wild-Type	-90.7	-21.52	-	-
p.P35L	-117.35	-28.04	-0.49	-0.89
p.D37N	-94.4	-22.56	-0.20	-8.83
p.E40K	170.98	40.86	0.48	183.33
p.T42A	-86.4	-20.65	0.59	3.58

Table 15. In silico prediction of variations between wild-type (wt) and mutant pVHL/p27. Free binding energy (ΔG_{bind}, electrostatic component) is calculated with APBS[219], while stability prediction and electrostatic solvation free energy of mutations affecting p27 (CDKN1b) were calculated with NeEMO[218] and Bluues[217].

MD simulations provided similar predictions. Notably, trajectory inspection suggests P35L to bind slightly better as it is facing a pVHL hydrophobic pocket. Conversely, E40K is predicted to negatively perturb the interaction, which could generally reflect the impact of charge inversions at the pVHL/p27 binding interface. To better address this behavior, the electrostatic properties of the putative pVHL/p27 complex wild-type (wt) in respect to the complexes formed with p27 mutant-forms has

been characterized using Bluues[217] (Figure 27). As expected, almost no variation in the solvation energy is predicted for P35L (-0.89 kJ/mol), with a modest variation predicted for D37N (-8.83 kJ/mol). Being exposed to the solvent and mutated into asparagine, the p27 D37 residue was predicted to little stabilize the complex. Conversely, E40K was predicted to markedly destabilize the interaction. Previously, the formation of a salt bridge between p27 E40 and pVHL R107 has been observed during MD simulation. A change in lysine clearly abolishes this interaction, suggesting a possible repulsive effect on the interaction surface. The predicted difference in electrostatic solvation energy of 183.33 kJ/mol for this mutant has been considered high enough to destabilize the complex. A slightly positive variation (3.58 kJ/mol) is predicted instead for T42A. Partially located in the hydrophobic pocket at the pVHL/p27 interface, T42 should have a modest impact on complex formation. The Poisson-Boltzmann Analytical-Method (APBS)[219] was used to quantify the binding free energy electrostatic component, required for pVHL/p27 complex formation (Table 15). The binding free energy predicted for the wild-type complex is $\Delta_{bind}G$ -22.07 kcal/mol, a favorable value for complex formation. Similar values are predicted for mutants D37N and T42A ($\Delta_{bind}G$ -22.56 and -20.65 kcal/mol, respectively) and $\Delta_{bind}G$ -28.05 kcal/mol is predicted for P35L. Instead, a positive value is predicted for E40K ($\Delta_{bind}G$ 40.86 kcal/mol). Collectively, these findings suggest that E40K is able to disturb the pVHL/p27 interaction, while a modest effect is predicted for the other mutants.

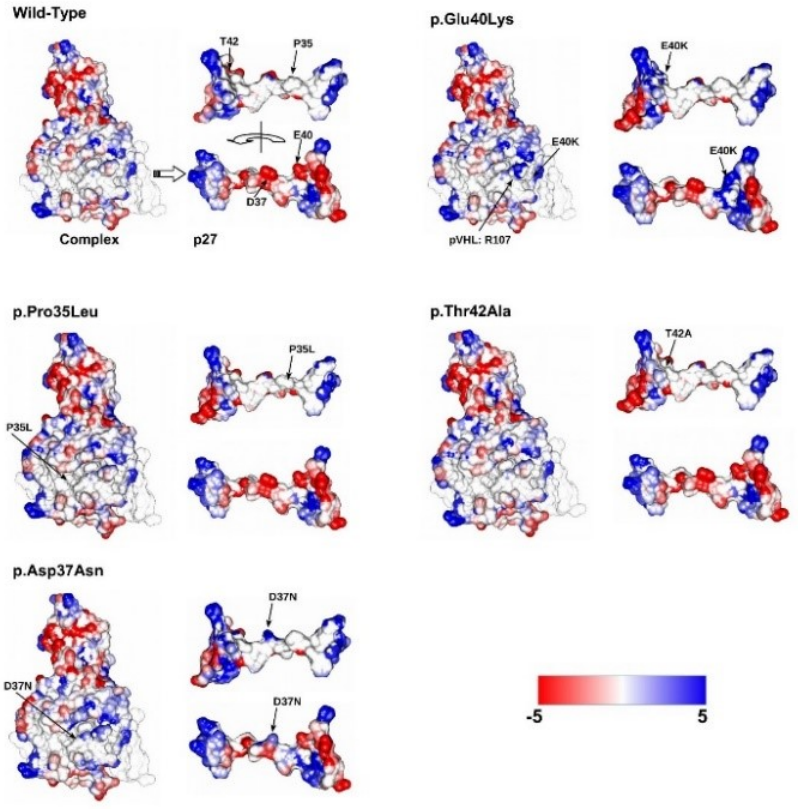


Figure 27. (Previous page) In silico prediction of mutations putatively affecting pVHL/p27 binding. Solvent accessible surface representation of a predicted pVHL/p27-NT complex colored by electrostatic potential. Blue represents positively charged areas, red negative. For each mutation, the full complex is presented with p27 alone next to it. Both the front (accessible surface) and rear (interaction interface) views are shown. The electrostatic potential was generated with Bluues[217].

The predictions have been then experimentally tested by Y2H assay (Figure 28). Both p27 P35L and D37N mutants retained the ability to associate pVHL30 almost completely, as indicated by similar yeast cell growth in selective medium. Although its possible modest predicted impact on the complex, the Y2H data further showed that T42A does not perturb pVHL-binding. Indeed, its introduction did not affect the growth rate of mutant yeast cells. Differently, E40K sensitively reduced p27 binding to pVHL30, as yeast cells expressing the p27 almost failed to grow in selective medium. Collectively, these results on p27 mutants further support that p27/pVHL binding may be described by the *in silico* model.

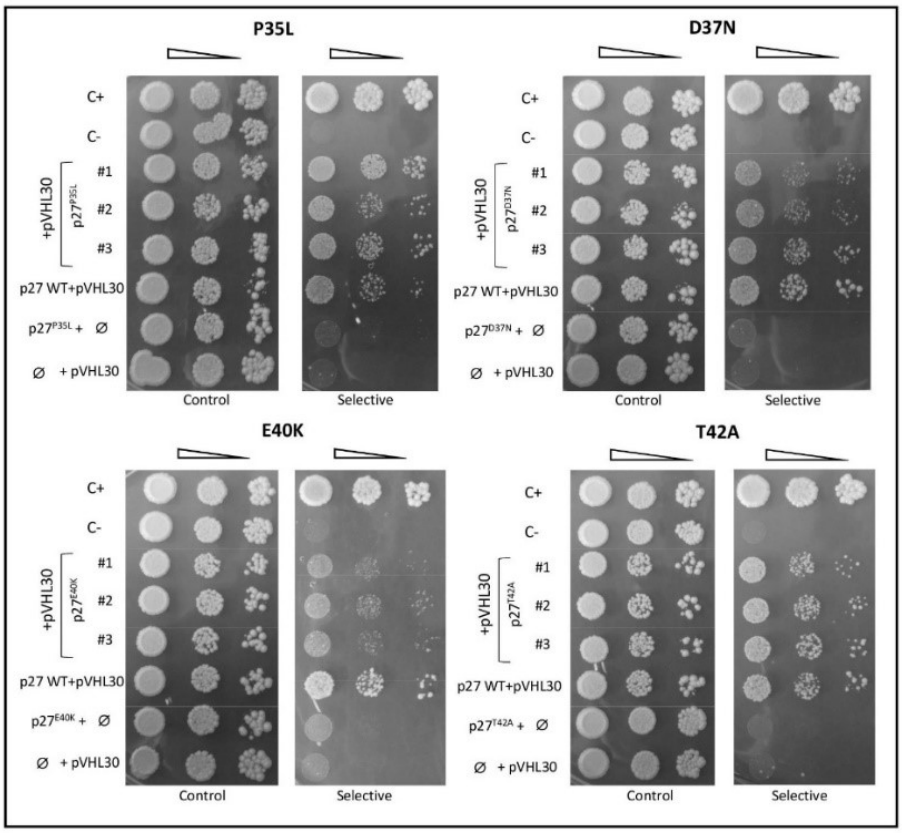


Figure 28. (Previous page) Effects p27 mutagenesis on pVHL binding. Y2H analysis of the p27 mutations P35L, D37N, E40K and T42A. Yeast cells co-expressing pVHL30 together with the indicated p27 mutants were assayed in Y2H. Three independent mutant clones are shown. Yeast cells have been grown in selective medium (right) supplemented with 60 mM 3AT, in order to increase the stringency of the binding assay. In all experiments, C+ and C- are positive and negative controls of the assay, respectively. The images are representative of three independent experiments, where 3-5 clones were analyzed.

Discussion

Supported by bioinformatic analysis, the starting point for this work lies in the observation that the Pfam CDI domain (pfam02234), which identifies the cyclin-dependent kinase inhibitor (CDKN1) protein family as sharing common elements with the CODD motif of HIF-1α, is involved in its binding with the pVHL protein. I personally provided evidence that all CDKN1 proteins (p21, p27 and p57) are able to associate with pVHL *in vivo*, as first indicated by Y2H assays and confirmed in mammalian cells by Co-IP. In this way, data pointed out a novel connection between the regulation of cell proliferation and transduction of multiple signals related to the HIF-1α and p53-dependent pathways. The functional association with pVHL may have important effects on the CDKN1 proteins, possibly influencing their regulatory functions. Recent findings show regulation of p27, in particular, to be linked with other components of the oxygen sensing pathway, such as PHD3.[247] This is thought to drive cell cycle entry at the G1/S transition by decreasing p27 stability. In this scenario, interaction with the von Hippel-Lindau protein could serve as a further modulator of p27 half-life. Together data highlight the common features between the interactions of the pVHL-β domain with both the CDKN1-CDI and HIF-1α-CODD regions, as supported by dissection of their binding elements. The pVHL/p27 interaction has been modeled *in silico* and investigated with respect to p27 pathological mutations found in COSMIC[215] and associated with cancer development. Collectively, *in silico* calculations and *in vivo* validations support that the association is maintained under hypoxic conditions, i.e. proline hydroxylation is not necessary, and the interaction is potentially able to exert its functional roles within the cells. Cell cycle progression, where the CDKN1 proteins act as natural inhibitors may be functionally linked *via* pVHL to the HIF-1α and HIF-2α dependent pathways, i.e. hypoxia/angiogenesis response, which is particularly relevant in cancer development. It is well known that under prolonged hypoxia, HIF-2α plays a key role in promoting genomic integrity and cell cycle regulation by stimulating c-Myc-mediated activation of cyclin D2 and the E2F1 transcription factor with concomitant repression of p21 and p27[248]. Importantly, since the same pVHL interaction surface may be involved, alternative binding to either HIF-1/2α, or each CDKN1 protein, might be mutually exclusive, and possibly competitive. In this biological context, particularly relevant is the observation that the pVHL/CDKN1s interaction is possible also in the absence of a hydroxylated proline. The pVHL interface B is known to bind different proteins[114,240,117] in a CODD-like fashion, however, all these interactions require the previous hydroxylation of the binding partner. In other words, it can be assumed that these proteins collectively compete with HIF-1α for the same binding interface at physiological oxygen concentrations. In both mild and prolonged hypoxia, PHD-dependent proline hydroxylation is inhibited[55,100], reducing competition for interface B. Very recently, a novel

hydroxylation-independent interaction between pVHL and Aurora kinase A (AURKA)[249], a serine/threonine kinases essential for cell proliferation was reported. In this context, interaction with the CDKN1 protein family may have evolved to allow transmission of a generic hypoxic signal to other signaling pathways, re-using the same pVHL adaptor protein. Intriguingly, the same pVHL region[41,3] mediates its association with over 40 proteins beyond HIF-1/2α[3], implying multiple factors in dynamic competition to alternatively associate unbound pVHL in the cell. These associations likely have to be also regulated at the post-translational level and may be related to different signals/pathways, as observed for HIF-1α hydroxylation by the proline hydroxylase (PHD) enzymes. Since yeast cells are devoid of PHD activity, our data strongly suggest that hydroxylation should not be involved in regulating CDKN1 binding, as further supported by the p27 P35L mutant. Experimental evidence suggested a negative impact of the CDKN1 C-terminus on association with pVHL, pointing to regulatory effects on the binding upon introduction of specific post-translational modifications, known to occur within the C-terminal region of the CDKN1 proteins[250,251,154,157]. Although preliminary correlations between hypoxia response and the CDKN1 proteins have been reported [252,253,247], the consequences of their association with pVHL are far from understood. As a member of the VCB complex, pVHL could also regulate CDKN1 degradation, similarly to HIF-1α, although published data on p27 turnover indicate the involvement of the Skp2 system[254]. Interestingly, a direct interaction between Skp2 and pVHL, mediated by its β-region, has been reported to stimulate Skp2 proteasomal degradation, independently from the pVHL-VCB E3-ligase activity[255]. Increased p27 levels have been observed in several pVHL-mutant cell lines[256], further pointing to their functional association. Future work is needed to address if (and how) binding with pVHL/VCB could impact both the stability and function of the CDKN1 proteins in cancer and health. The relative affinity of pVHL binding (i.e. K_d values) should be determined. Moreover, mammalian cells should be used to investigate the competitive nature of the binding, to determine which are the functional implications of the protein-protein interactions characterized here. Site-specific mutations results, combining both computational and experimental approaches, contribute to shed light on the effects of pathogenic variants towards the association of p27 with pVHL, which could help to clarify the relationship between clinical phenotype and functional defects caused by CDKN1 mutations.

3.2 Investigation of the human von Hippel-Lindau paralog VLP and isoforms.

Structural comparison between the VHL isoforms and VLP protein by in silico investigation.

Upon its transcription, primary VHL mRNA molecule can be translated encoding either full-length isoform[257] (i.e. pVHL30, 213 residues) and/or a shorter isoform (160 residues) referred as pVHL19[43] which is produced by an alternative TSS (translation start site) located at methionine-54. At the level of exon 2, moreover, the same mRNA may be alternatively spliced producing a third pVHL172 isoform[47] (172 residues). Collectively taken, these isoforms differ from each other for the presence/absence of specific regions, which are known to engage in multiple protein-protein interactions. In addition to this complex scenario, a VHL paralog gene (*VHLL*) located in the human chromosome 1 (1q21.2) encodes for VHL-Like Protein (VLP), 139 residues protein sharing large sequence identity (68%) with pVHL[53]. Both pVHL172 and VLP are poorly described by literature and their biological role is far to be clear. In light of these peculiarities, I decided to investigate these proteins *in silico* and validate resulting prediction by *in vivo* experiments. The *VHL-related* proteins (Uniprot[200] identifiers: P40337-1 pVHL30; P40337-3 pVHL19; P40337-2 pVHL172) were compared by multiple sequence alignment. All isoforms share two highly conserved regions spanning pVHL30 residues 54-113 and 155-213, predicted to form β-sheet and α-helix, respectively. The pVHL19 isoform, as already known, completely lacks of the N-terminal region[258] (residues 1-54), which is maintained in pVHL172 (Figure 29A). However, this isoform entirely lacks residues 114-154[49]. Inspection of pVHL30 3D structure shows these residues to form part of the β-domain. Severe structural rearrangements are expected for pVHL172. Similarly, I analysed the VLP sequence (Uniprot identifier: Q6RSH7). The alignment shows VLP to be characterized by a shorter N-terminal tail, which albeit being acidic, presents less repeated units than pVHL30 N-terminus. Indeed, only three xEEx repetitions are present in the sequence, contrary to pVHL30 accounting for eight repeated GxEEx motif[257]. Furthermore, VLP sequence lacks the C-terminal tail (residues 158-213), otherwise conserved in all three pVHL isoforms.

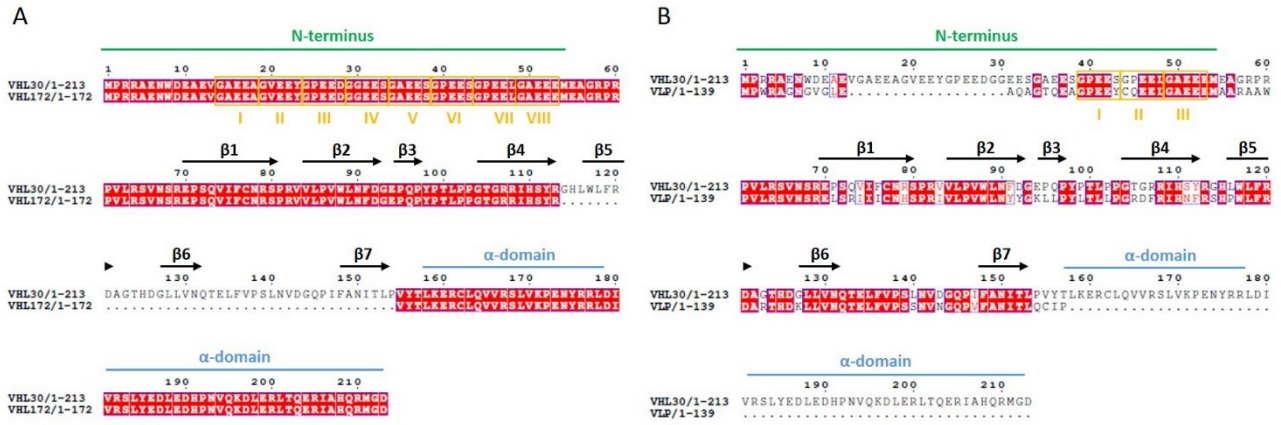


Figure 29. Multiple-sequences alignment of VHL isoforms and VLP. (A) pVHL30 and pVHL172 alignment and (B) pVHL30 and VLP alignment. In both the N-terminus (green), GxEEx repetitions (yellow boxes) and secondary structures are highlighted (α-helix in light blue and β-sheet in black).

Considering the overall sequence similarity among VLP and pVHL isoforms, I predicted their putative 3D-structures by homology modeling. The crystal structure of pVHL30 bound to HIF-1α together with ElonginB and ElonginC (VCB complex)[40] was used as template (PDBID: 1LM8[259]). For pVHL172, an overall calculated QMEAN score was 0.16, while TAP estimated a 0.7830 confidence value. The pVHL172 3D model preserves the pVHL α-helices organization of the C-terminal domain (Figure 30), which is known to participate the VCB[40]. Although regions spanning residues 60-113 and 155-209 are conserved in the pVHL172 primary sequence, the resulting β-region appears largely compromised with the β1- β4- β5- and β6- strands structurally disrupted and markedly assuming coil organization. The region corresponding to β2- and β3- strands is partially preserved and predicted to form a short β-hairpin (Figure 30). Similarly, VLP 3D shows an almost conserved β-domain folding organization, while totally missing the α-domain. In particular, the highest conservation (100% identity) is found for the β2- and β6- strands and a short loop, which in pVHL structure correspond to a disordered region. The VLP model quality was also assessed with QMEAN (score 0.65) and TAP 0.8008 confidence value[223] indicating an overall acceptable model quality.

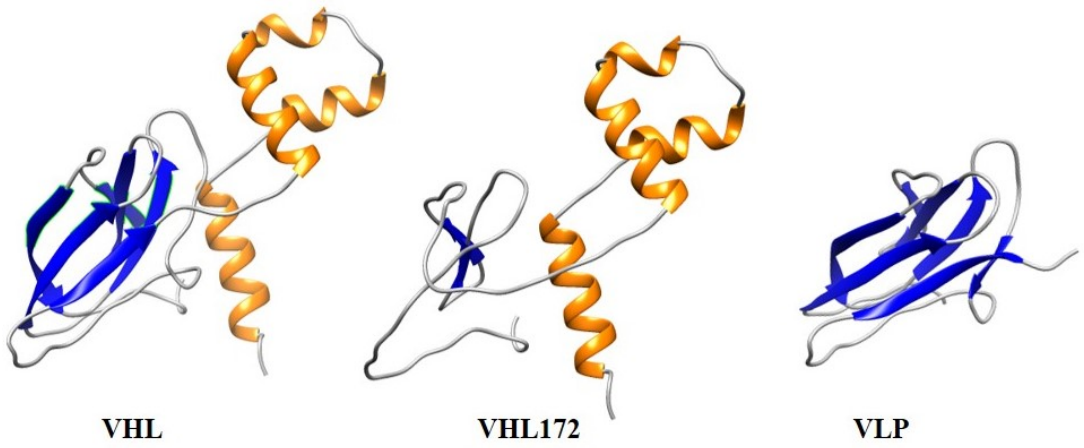


Figure 30. VHL172 and VLP model. By superimposition on VHL protein crystal structure, the 3D structure of pVHL172 and VLP has been predicted. Respectively, coil regions are colored in dark gray, α-helices are highlighted in orange and β-sheets in blue. Models were obtained using SWISS-MODEL and Chimera[221].

Although sharing common structural elements, data support that each pVHL isoform (i.e. 30, 19, 172) as well as VLP are however endowed with specific structural features, possibly reflecting their different functional involvements. Therefore, I further addressed this biological question investigating their ability to interact with other proteins and their behavior in human cell models.

PPI (protein-protein interaction) analysis for the pVHL isoforms and the VLP paralog.

Given its ability to interact with >500 different interactors[260], full-length pVHL30 can be considered as a functional hub. Among them, the best studied is its interaction with the transcription factor HIF-1α, modulating cellular response to hypoxia. Acting as an SRP (substrate recognition particle) of the VCB multiprotein complex (together with Elongin B, Elongin C, Cullin-2 and RBX1), pVHL30 targets HIF-1α to the ubiquitin-mediated proteasome degradation pathway[261]. The binding involves the oxygen-dependent degradation domain of HIF-1α (ODD, residues 556-574), which assumes a linear conformation while fitting the cleft on the pVHL30 β-region by multiple interactions (i.e. hydrophobic, van der Waals, ionic)[210]. At physiologic conditions, this association requires the hydroxylation of a conserved proline residue in the ODD domain mediated by the activity of specific oxygen-dependent prolyl-hydroxylase enzymes (PHD1-3). As *in silico* structural data predict the VLP β-domain to be globally maintained while altered in pVHL172, I firstly assayed their ability to interact with HIF-1α. For this purpose a modified version of the Y2H has been used, as already

described by *Bex et al. 2007*[262]. In detail, a single Y2H vector (i.e., pBridge) allowing the simultaneous expression in yeast cells of the *bait* protein (pVHL172 or VLP fused to Gal4-DBD) with full-length human prolyl-hydroxylases 3 (PHD3)[262] was tested. The plasmid was then co-transformed in Y190 yeast cells with the *prey* expressing vector (i.e., pGADT7 HIF-1α). At least three independent clones were selected and assayed, where yeast cells growth on selective medium is indicative of binding between bait and prey proteins. As expected, yeast cells expressing HIF-1α and pVHL30 β-domain markedly grew on selective medium (Figure 31 and Suppl. Figure 5), whereas yeast expressing either VLP or pVHL172 were not able to survive, indicating no binding with HIF-1α for both proteins, also in the presence of PHD3 enzyme.

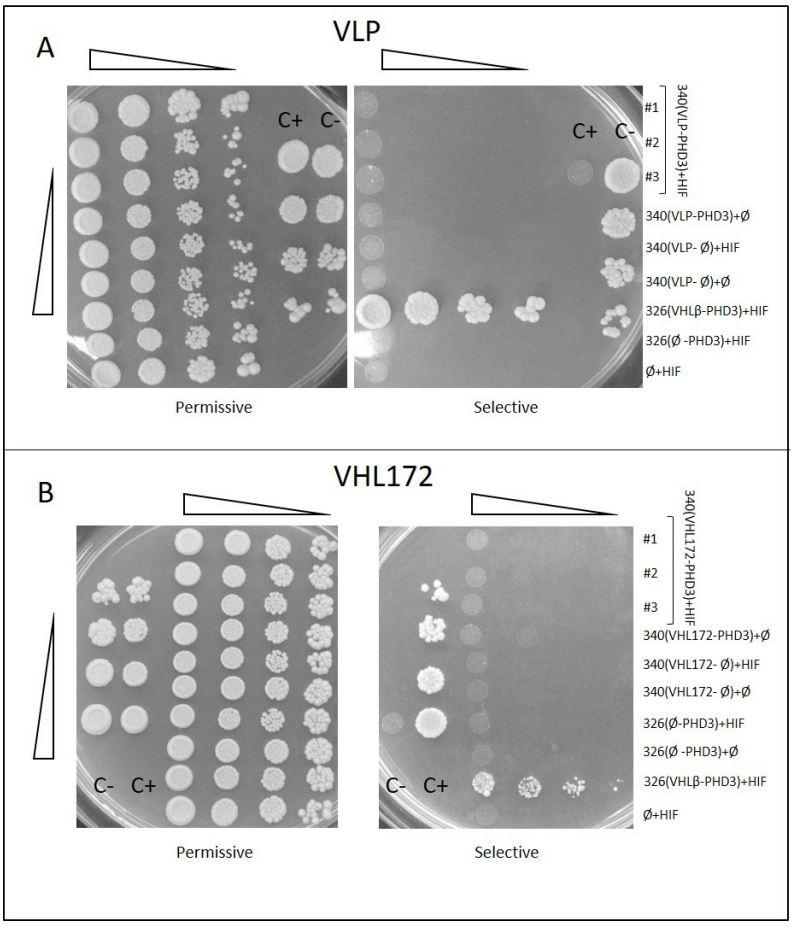


Figure 31. HIF-1α does not retain the interaction with VLP or VHL172, even in PHD3 presence. (Top) In the Y2H reconstituted system, yeast cells co-expressing VLP, PHD3 and HIF-1α proteins show no growth on selective medium after 8 days at 30°C. Whereas the simultaneous expression of VHLβ, PHD3 and HIF-1α proteins markedly allow yeast growth in absence of histidine. (Bottom) Similarly, the reconstituted Y2H expressing VHL172 is lethal for the yeast on selective media. In both panels, C+ and C-are positive and negative controls of the assay. In addition, different auto-activation control exclude independent activation of the report gene HIS3. Images are representative of at least three independent experiments, where three different clones have been analyzed (n≥9). Others details on *Suppl. Figure 5*.

Absence of interaction between HIF-1α and pVHL172 is interpreted as due to marked alterations of the β-domain structure, as also highlighted by *in silico* analysis. Accordingly, it has been reported pVHL172 does not regulate HIF-1α stability *in vivo*, although being able to participate VCB complex[50]. Unexpectedly, VLP is also unable to interact with HIF-1α, despite its largely conserved β-domain. However, a closer analysis of the primary sequence highlighted some crucial differences within β domain residues. In particular, the serine 111 (S111) of VHL30, a polar residue known to engage direct interaction with HIF-1α, is turned into asparagine (N) in VLP (Figure 29B). Asparagine is characterized by a longer and bulkier lateral chain suggesting that the substitution may induce local structural reorganization. More importantly, it has been already reported that several pVHL30 mutations affecting S111 abolish HIF-1α binding[56]. Together with Trp88, Tyr98, His115 and Trp117, S111 is highly conserved in human and necessary to coordinate interaction with hydroxyl-proline 564 within the HIF-1α CODD motif[40]. Among other possibilities, this result could be due to a weak binding, undetectable by Y2H system used here. It can be also proposed that VLP/HIF-1α interaction may require additional post-translational modifications (PTMs) beyond proline hydroxylation, which were not available in Y2H system. In the literature[53] is reported that a weak association may occur in human cells between VLP and HIF-1α, independently from PHD activity, which could be consistent with the undetectability of the binding by our Y2H assay.

pVHL30 shares with both VLP and pVHL172 common mechanisms to bind the CIPs proteins.

Over HIF-1α/pVHL30, the β-domain also mediates the binding between pVHL30 and all members of the CDKN1 protein subfamily (i.e. p21, p27, p57)[263]. As reported in chapter 1, the presence of a CODD-like linear motif at the N-terminus of the CDKN1 leads the binding with pVHL30 in a proline hydroxylation-independent fashion[263]. In light of sequence and structural similarities, I wondered whether this association is also mediated by pVHL172 and VLP. As shown in Figure 32 (p21 (A), p27 (B), p57 (C), Suppl. Figure 6 and Suppl. Figure 8) yeast cells co-expressing any CDKN1 protein together with either VLP and pVHL172 protein are viable and actively growing on selective medium, with respect to negative control cells, actually demonstrating their ability to interact.

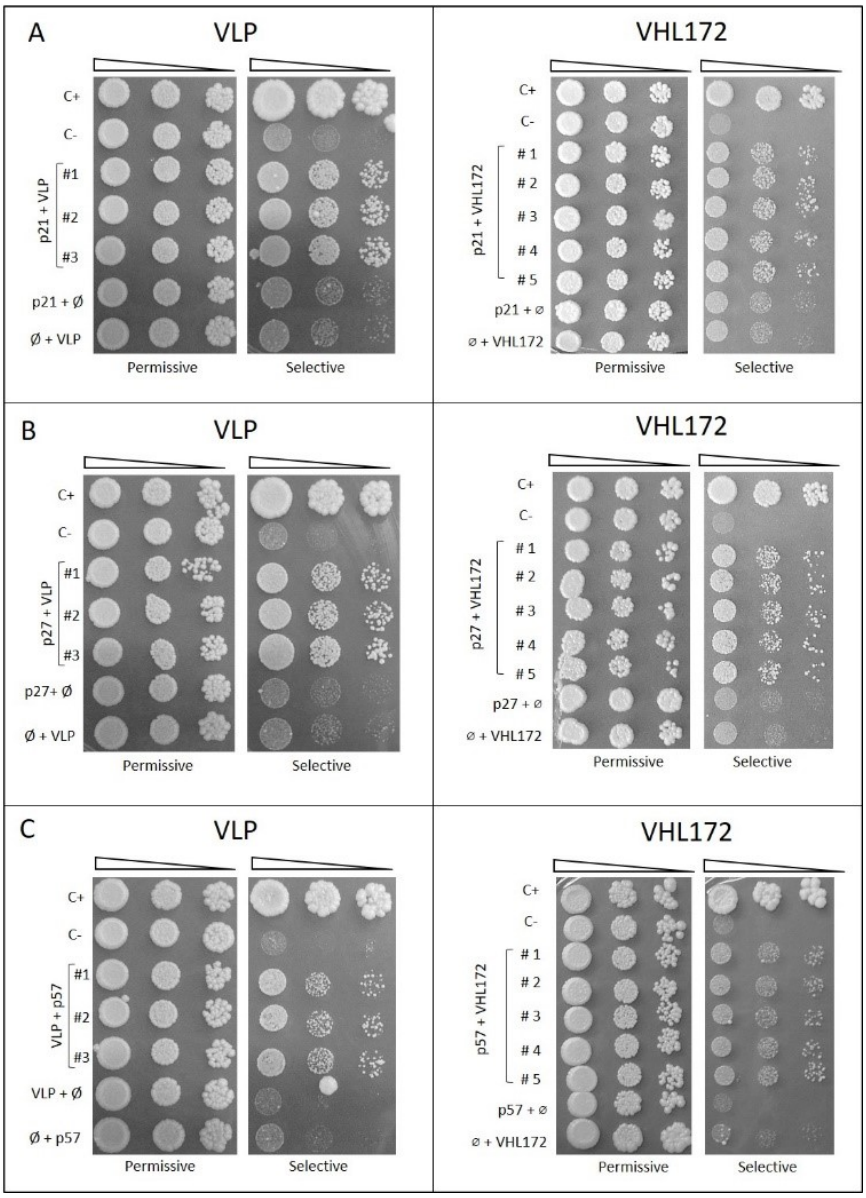


Figure 32. The CDKN1 members interact with both VLP and VHL172. Yeast two hybrid (Y2H) assay of VLP (left panel) and VHL172 (right panel) binding to the CDKN1 members (A: p21; B: p27 and C: p57). 10-fold serial dilutions of yeast cells co-expressing the indicated Gal4BD/AD-fusion proteins are spotted on both permissive (+HIS, left) and selective (-HIS, +3AT, right) media, and growth is constantly monitored for 8 days at 30°C. In all plates, C+ and C- are positive and negative internal controls of the Y2H system, indicative of strong, or no binding, respectively. Yeast cells co-transformed with the empty vector (Ø) are controls of the auto-activation by fusion proteins alone. Images are representative of at least three independent experiments, where 3-5 different clones have been analyzed (n≥15). See also in *Suppl. Figure 6 and Suppl. Figure 8* the protein expression.

However, data also indicate that despite predicted structural alterations, the pVHL172 β-domain can yet organize to bind the CDKN1 proteins when expressed in yeast cells. This suggest that the protein may assume a 3D organization which is not easily predicted by homology modelling. In order to confirm these interesting interactions in a more physiological context, I performed a series of co-immunoprecipitation experiments in human cells. As *proof-of-principle*, I overexpressed in HEK293T cells epitope-tagged proteins (i.e. VLP or pVHL172 together with the CDKN1C protein, i.e. p57), and total cell lysates were subjected to co-IP assay. As shown in Figure 33, p57 is able to associate both VLP (8A) and pVHL172 (8B), as revealed by their presence in the immunoprecipitates, demonstrating that also in human cell the pVHL172 β-domain can properly bind CDKN1 proteins. These data suggest that the regions shared by pVHL172 and VLP may form a CDKN1 binding site. Further analysis are however required to confirm this evidence.

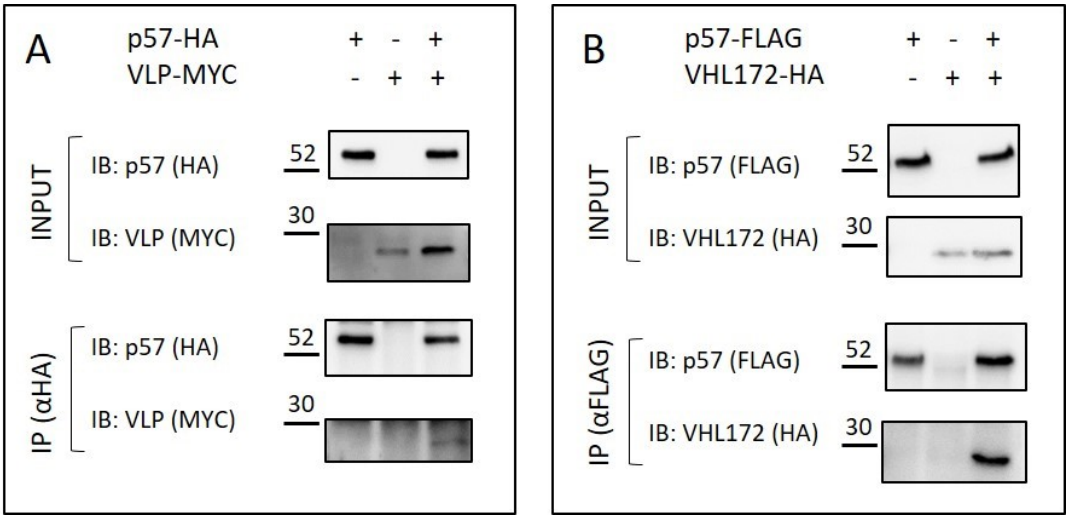


Figure 33. Validation of p57 and VHL172 interaction in mammalian cells. The co-immunoprecipitation assay is performed using HEK293T cells lysates. By cells transfection, Flag-tagged p57 and Myc-tagged VLP (A) and Flag-tagged p57 and HA-tagged VHL172 (B) are transiently overexpressed. Recombinant proteins expression is firstly detected by immoblotting with either anti-HA or anti-Flag or anti-Myc antibodies (Input, upper panel). Afterwards p57 is immunoprecipitated respectively with anti-HA antibody or anti-Flag antibody and VLP or VHL172 are detected in the immunoprecipitates (IP, bottom panel).

Then, I wondered whether the binding requires the CKI domain, localizing at the CDKN1s N-terminus, similarly to what observed for their association with pVHL30[263]. Excluding p21, which cannot be analyzed due to aspecific activation of the reporter gene, I tested binding between the N-

terminal tails of both p27 (1–60 residues) and p57 (1–61 residues) with VLP and pVHL172. Yeast cells expressing either VLP or pVHL172 together with any N-terminal region of CDKN1 proteins maintain survival and growth (Figure 34 and Suppl. Figure 7-10), therefore demonstrating that the interaction is mediated by the same motif proposed in [263].

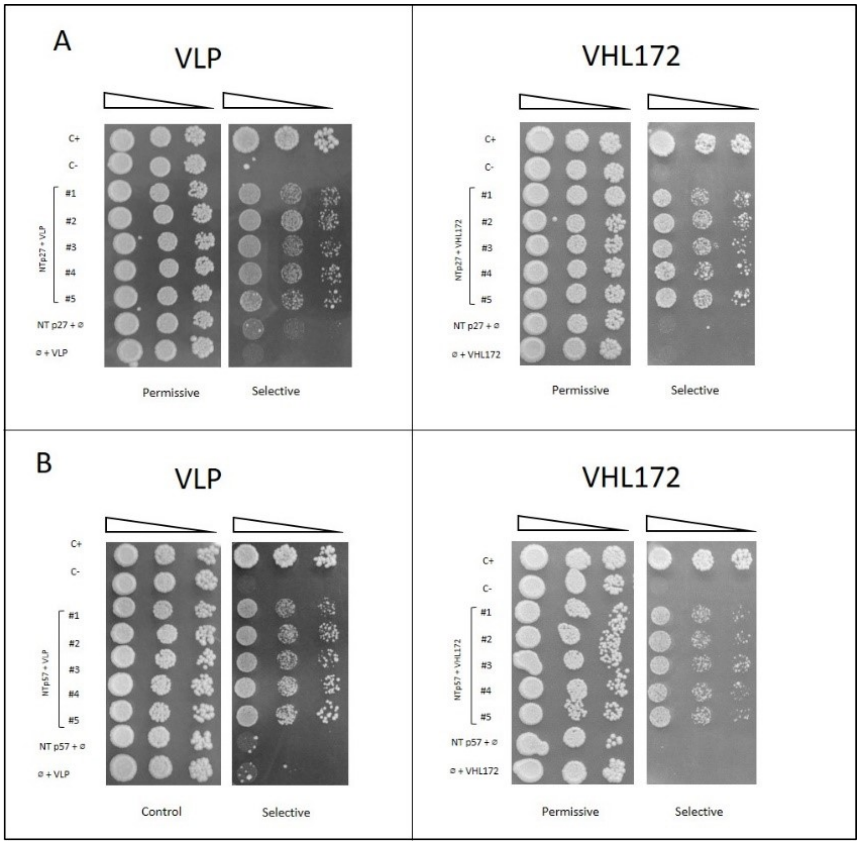


Figure 34. N-terminal tail of p27 and p57 is necessary to bind VLP and VHL172. Y190 strain is co-transformed with the N-terminus region of p27 or p57 proteins (A: NTp27, residues 1-60 and B: NTp57, residues 1-61) and VLP or VHL172. Standard drop test is performed on permissive (+HIS) and on selective (-HIS, +3AT) media. While co-transformant with empty vector (⌀) are evidence of auto-activation absence, C+ and C- are respectively positive and negative controls. Images are representative of at least three independent experiments, where 3-5 different clones have been analyzed (n≥15).

Accordingly to pVHL30, removal of the N-terminus disrupts the association (Figure 35 and Suppl. Figure 7-10), as yeast cells do not survive on selective medium when co-expressing VLP or pVHL172 together with any CDKN1 C-terminal regions, i.e. p27-CT (residues 61–198), and p57-CT (residues 62–316). Due to autoactivation activity, yeast co-expressing VLP or pVHL172 together with p21-CT are not considered (Suppl. Figure 11).

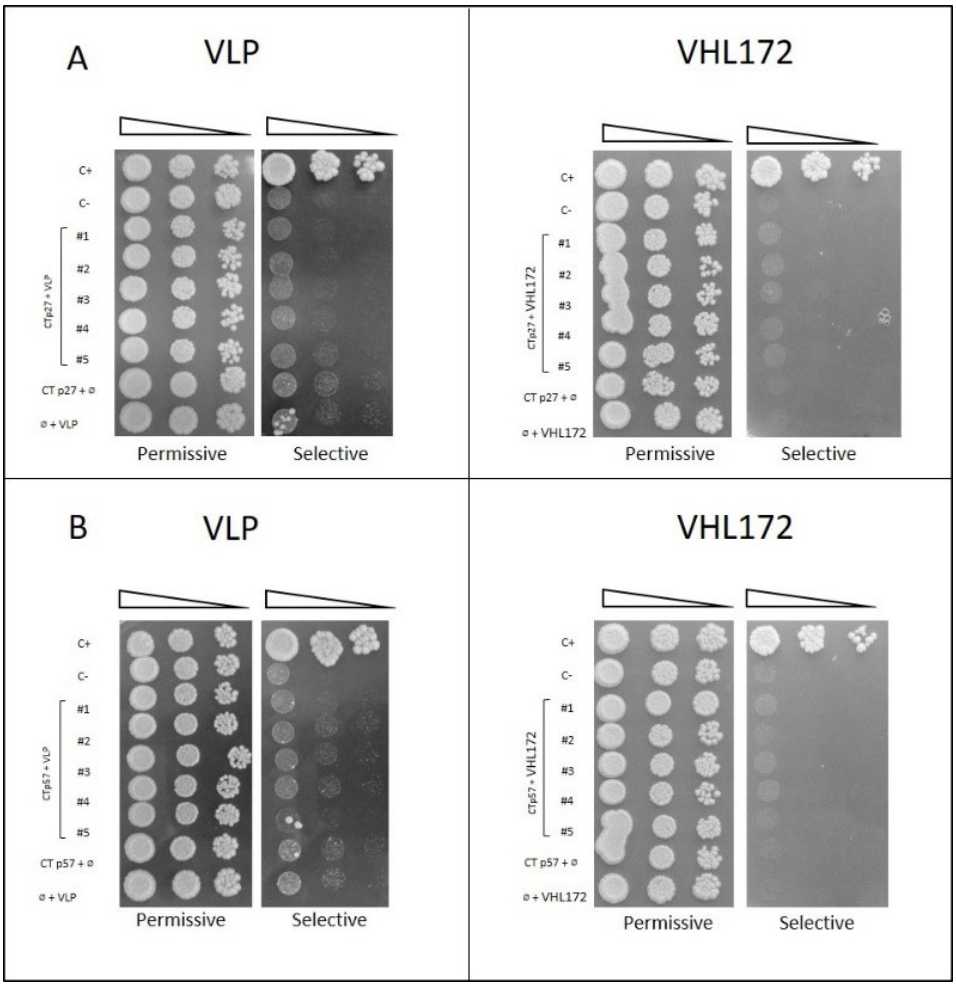


Figure 35. C-terminus region of CDKN1 members does not retain the interaction with VLP. Using Y2H assay, the absence of interaction between the different C-terminus regions of CDKN1 proteins (A: CTp27, residues 61-198 and B: CTp57, residues 62-316) and VLP or VHL172 are detected. Each plate has a positive (C+) and a negative (C-) control, enriched by auto-activation controls (yeast cells co-transformed with empty vector - ⌀). The spot test experiment is performed both on a control plate (+HIS) and on a selective one (-HIS, +3AT). At 30°C, the assay lasts 8 days and it is repeated at least in three independent experiments (clones n≥15).

Finally, the p27 pathological mutation Glu40Lys located within the N-terminus CDKN1 [216] was tested as it was already demonstrated to inhibit p27 association with pVHL30[263]. As shown in Figure 36 (Suppl. Figure 7 and 9), the p27-E40K mutation also negatively impacts p27 interaction with both VLP and pVHL172, as indicated by the reduced growth rate of the yeast cells on selective medium respect to the wild-type p27 expressing cells.

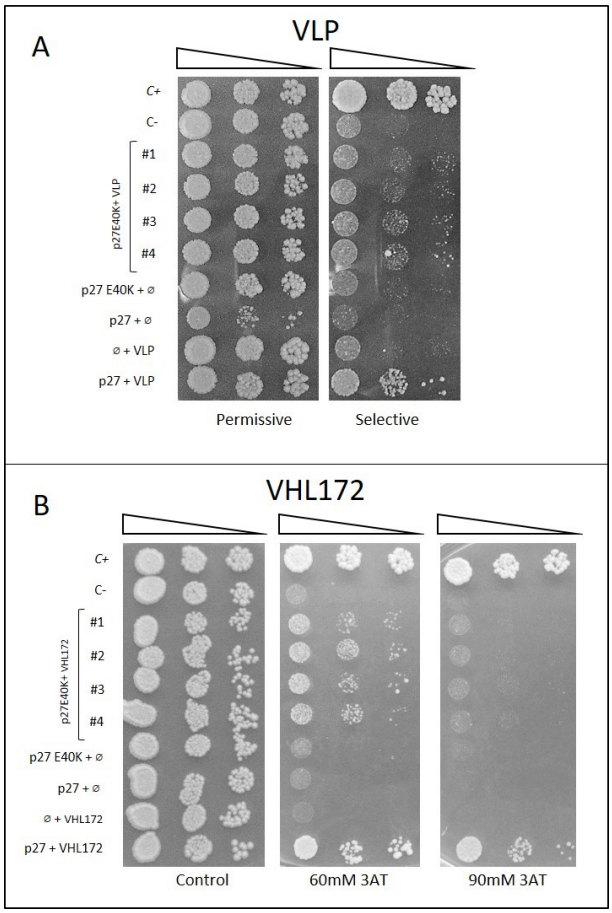


Figure 36. The pathological mutation on p27 alters the interaction with VLP. The interaction between mutated p27 (residue E40>K) with VLP (A) and VHL172 (B) are analyzed using Y2H technique. Several yeast cells dilutions are spotted on permissive media (+HIS, left) and on selective media (-HIS, +3AT at different concentration when necessary, right). The experiment contains several internal controls necessary for a qualitative growth examination, i.e. positive (C+) and negative (C-) controls, auto activation (yeast co-transformed with empty vector, Ø) and p27 wild-type interaction are present. Images are representative of at least three independent experiments, where four different clones have been analyzed (n≥12).

Collectively, experimental data support that the conserved CDKNs motif also sustain their interaction with both VLP and pVHL172, similarly to the binding mechanism observed for pVHL30[263].

MDM2 interacts to pVHL172 and VLP

Recently the association of pVHL30 with the Mouse double minute 2 homolog (MDM2) oncosuppressor has been revealed in our laboratory (*unpublished data*). Unlikely to CDKN1/pVHL interaction, this association also involves the N-terminal disordered tail of pVHL30, indicating an

isoform-specific binding. Notably, experimental evidence supported that the binding occurs by multiple interactions, involving MDM2 residues located onto 329-452 region. In particular, a predicted linear motif within the MDM2 interval 351-452 is recognized by the pVHL β-region, while a MDM2 positively charged stretch immediately upstream (residues 329-350) the first interacting motif, is proposed to form multiple salt-bridges with the negative charged N-tail of pVHL30 (*unpublished data*). In light of their conserved acidic tails, I thus checked whether both pVHL172 and VLP maintain the ability to interact with MDM2 by the Y2H assay. First, I tested their binding with the full-length MDM2. As shown in Figure 37, co-expressing MDM2 with either pVHL172 or VLP allows yeast cells to survive and grow on selective medium, with respect to negative control cells, mirroring the behavior registered for pVHL30. Although the Y2H spot test cannot be considered quantitatively, growth rate of yeast cells on selective medium may be indicative of the binding affinity. As judged by the experimental data, yeast cells co-expressing MDM2 and pVHL172 were more viable than those co-expressing VLP, suggesting a difference in binding affinity. Further, these findings are consistent with a specific role of the pVHL disordered N-terminus in associating MDM2, as interaction, i.e. yeast grow rate, is maintained in pVHL172 and reduced in the VLP.

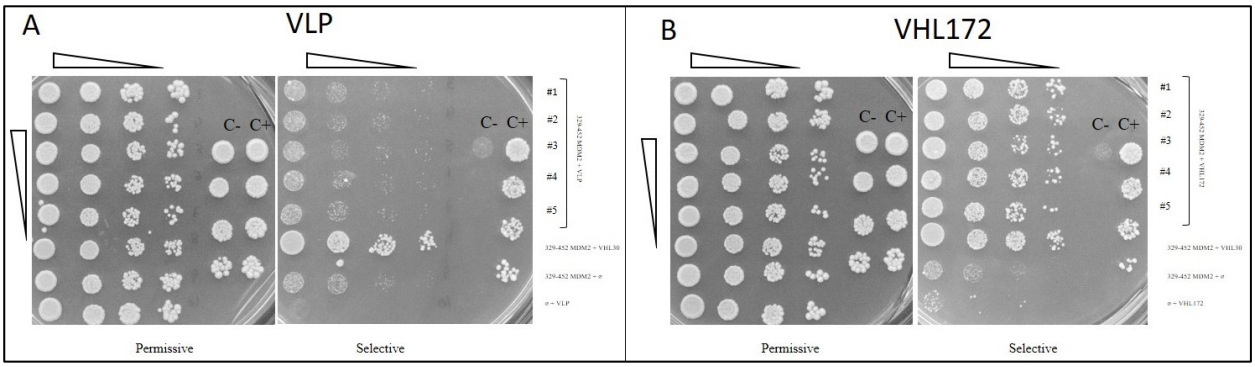


Figure 37. The MDM2 dissection unveiled the preference for VHL172 and the presence of N-terminal region. Yeast cells co-expressing VLP (A) or VHL172 (B) together with 329-452 MDM2 portion are assayed by Y2H system. Serial dilution of yeast cells is seeded on permissive and on selective media. C+ and C- (disposed on the right of each plates) are positive and negative assay controls present in all experiments. Each panel contains auto activation controls represented by yeast strain co-transformed with the empty vector (Ø). Images are representative of at least three independent experiments (n clone≥15).

Collectively taken, these experimental evidences point to a role for the pVHL N-terminal tail in forming a further pVHL binding surface putatively driving isoform-specific protein-protein interactions.

The N-terminus has a role for pVHL protein stability in human cells.

In order to correlate the different structure of the VHL proteins with their behavior within mammal cells, I investigated their intracellular distributions. I inserted each coding sequence in fusion with the Green Fluorescent Protein (GFP), originating a series of plasmids able to transiently express the fusion proteins in human cells. Then, HeLa cells were transfected and GFP-tagged proteins visualized by confocal microscopy after 24 hours post-transfection. Representative images presented as Figure 38, show that cellular localization is somewhat different, whether compared to GFP alone (i.e. control experiment). In particular, pVHL30 and pVHL172 share a similar diffused distribution, subtly enriched in the nuclear compartment, whereas VLP appears to localize into punctate structures widely distributed within the cytosol. Interestingly, the pVHL19 isoform also shows a punctate localization, strongly suggesting a role for the intrinsically disordered N-terminal tail in driving pVHL localization when present. Indeed, this region is minimal in VLP, absent in pVHL19, while it is fully maintained in pVHL172 which share a comparable profile with pVHL30. Accordingly, removing the pVHL172 N-terminus (pVHL172-ΔNT) modifies its fluorescence pattern to a punctate distribution, similar to what observed for both VLP and pVHL19 proteins. Notably, I observed similar fluorescence profiles for all VHL proteins also when the same experiments are performed in HEK293T cells confirming that these behaviors are also maintained with different cell-type and genetic background.

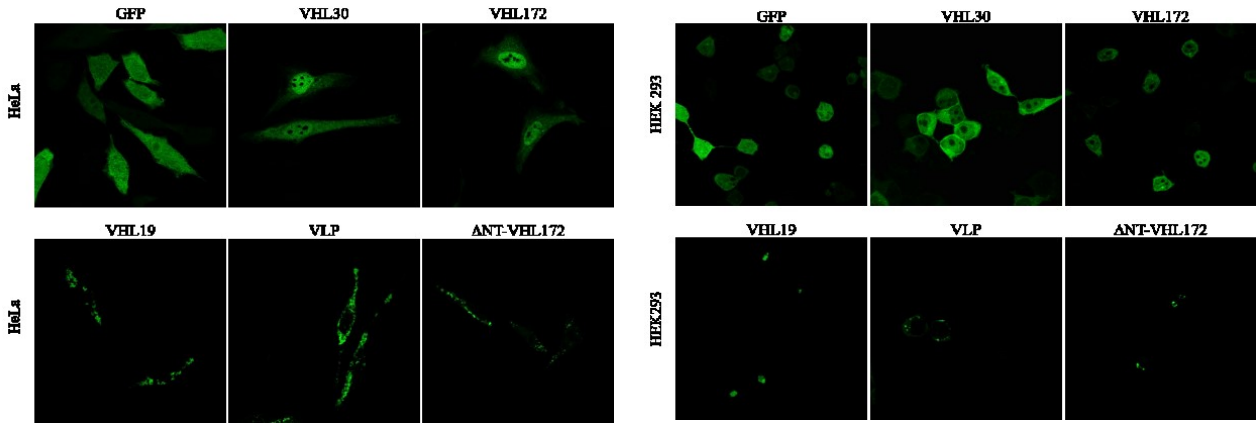


Figure 38. (Previous page) The N-terminal portion may regulate the VHL-structures localization with punctate distribution. Representative image of GFP-VHL-proteins localization in Hela (left panel) and HeK293 (right panel) mammalian cells. Transiently transfected, cells were then fixed and observed using Leica SP5 confocal microscopy. A diffused distribution was associated to VHL-proteins N-terminal region-equipped, similarly to GFP control. Whereas, the partial or complete absence of the amino-terminal portion resulted in punctate aggregation into cytosol.

Such punctate distribution cannot be easily associated to known-established cellular compartments (e.g. ER, Golgi or mitochondria), I checked whether it could be related to formation of protein aggregates, possibly modulated by the pVHL N-terminus. I therefore performed biochemical experiments on total protein lysates from HeLa cells overexpressing each GFP-tagged protein. By cellular fractionation[197], I separated insoluble/aggregated from soluble proteins and subjected each sample to SDS-PAGE and Western blotting. Finally, immunoblots have been revealed by mean of a GFP antibody. Resulting data (Figure 39) show that VHL proteins endowed of extended acidic N-terminal tail, i.e. pVHL30 and pVHL172, are mainly distributed in soluble fractions similarly to GFP alone. Differently, the N-terminus absence clearly unbalances the amount of insoluble protein, as indicated by higher bands intensity detected by western blots assaying pVHL19, VLP, and pVHL172-ΔNT.

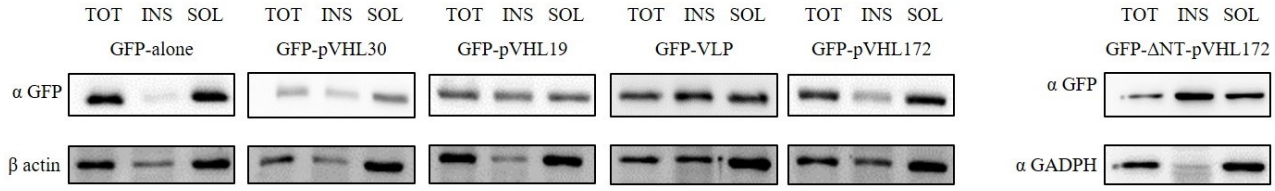


Figure 39. The N-terminal region may unbalance the insoluble/soluble amount of proteins.

These findings, although obtained in non-physiological conditions (i.e. transient overexpression), collectively suggest for the first time a role for the N-terminal region in regulating pVHL stability in cellular environment, which differentiates pVHL30/172 biological properties from pVHL19 and VLP, either lacking or partially missing the disordered N-terminus. Further exciting future experiments are required to elucidate the biological meaning and their impact on cell functions.

Discussion

In human cells, the *VHL* gene is expressed into different protein isoforms (pVHL30, pVHL19 and pVHL172), by either alternative translation and/or splicing mechanisms[50]. Additionally, the first paralog of VHL protein (VLP) has been recently identified in human[53]. Among these proteins, only full-length pVHL30 and the shorter pVHL19 isoform have been so far extensively studied, while little is known about the structural and functional properties of both pVHL172 and VLP. In this work, using the full length pVHL30 as pVHL prototype I comparatively investigated some of their biological features and highlighted isoform-specific interactions. Deciphering the isoform-specificity of the pVHL-related proteins could indeed be helpful to clarify their biological competence within the cells, in health and disease. Full-length pVHL30 (residues 213) can be structurally subdivided into three main regions: a N-terminal disordered tail (residues 1-53), a β-domain (residues 54-157), and an α-domain (residues 158-213). Primary sequence inspection by multiple sequence alignment shows the pVHL172 isoform to preserve both N- and C-terminus of pVHL30, whereas a 40 residues large deletion affects the central β-region (residues 114-154). Compared to pVHL30, the VLP paralog is endowed with a shorter acidic N-terminus and completely missing the alpha domain, whereas its β-domain displays high levels of sequence conservation. Investigation *in silico* predicts human VLP to assume a fold resembling pVHL30 β-domain, further pointing out that their main difference resides in the complete absence of the α-domain, which is required to bind specific partners (e.g. VCB members). Conversely, this latter domain is completely preserved in pVHL172, which is however characterized by an altered β-domain paired to an extended intrinsically disordered long N-terminus. These characteristics collectively taken suggest that the different members forming the pVHL protein family differ each other by the absence of a single specific binding interface. In other words, it can be proposed that the three pVHL isoforms may be evolved to specifically account different exclusive functions, e.g. regulating specific subset of pVHL30 interactors. Consistently, experimental evidence supports that in human cells pVHL172 retains the ability to form the VCB complex[50], while the presence of an altered HIF-1α binding site suggests that the regulation of its stability may be exerted by competing pVHL30/19 to form VCB complex. Collectively taken, my investigations also pointed out structural and functional similarity. In this context, it has been investigated by the Y2H system the ability of both pVHL172 and VLP to associate with specific partners (i.e. CDKN1 members, HIF-1α, MDM2). Notably, these interactions are mainly mediated by multiple contacts between the pVHL30 β-domain and at least a linear motif originated on target protein. However, isoform-specialization in associating specific binding partners can be distinguished. The binding of pVHL30 with HIF-1α is well-known to require the hydroxylation of specific proline within N- and C-ODD degradation motifs. This post translational modification is not a requirement for associating the

CDKN1 protein family. Further, the pVHL30 N-terminus is found to have a relevant role in binding MDM2 suggesting this interaction to be exclusive of pVHL isoforms endowed of a disordered N-terminal tail. Interestingly, despite the pVHL172 β-domain is predicted to be largely altered, experimental data seems to indicate that this region can properly associate specific partners in multiple cellular environments, as also confirmed by co-immunoprecipitation assays in HEK293T and HeLa cells. Reconstituted Y2H experiments indicated that pVHL17 does not bind HIF-1α pointing to isoform-specializations distinct from full-length pVHL30. Accordingly, it has been reported that the ability of pVHL172 to form the VCB complex appears unrelated to HIF-2α stability regulation[50]. This independent observation is also in agreement with my data which suggest both pVHL172 and VLP to not be not able to associate HIF-1α. In that regard, VLP data are particularly puzzling considering the structural conservation of its β-domain. A very weak interaction with HIF-1α cannot be, however, completely excluded. Indeed an *in vitro* association between VLP and HIF-1α has been reported, even if noticeably weaker than pVHL30[53]. My experimental results in this sense, may be related to either a limited Y2H sensitivity or other PTM-mediated regulations which are absent in yeast cells. Further, Y2H experiments confirmed that both pVHL172 and VLP associate full-length MDM2, similar to pVHL30 (*unpublished data*). Notably, data also demonstrated the relevance of the VHL N-terminus, already related to isoform-specific properties/functions[264]. Together these experimental evidence strongly supports that the structural differences among the VHL-related proteins (i.e. pVHL30, 19, 172, and VLP) may be linked to peculiar biological properties, as suggested by their differential ability to interact with specific VLP partners. Possible involvement in different roles is further suggested by the intracellular localization analysis. When transiently overexpressed in human cells, the VHL-structures are indeed characterized by different distributions. Compared to GFP-alone, pVHL30 and pVHL172 appear globally diffused in mammalian cells, whereas VLP is restricted to punctate elements in the cytosol (similarly to pVHL19). Data suggest a role for the N-terminal sequence in regulating their intracellular distribution, which is further supported by the results from partially ablated pVHL172 mutant.

My findings also validated that the intracellular localization correlates with the distribution of the VHL-related proteins into either soluble or insoluble/aggregated. Collectively taken, data suggest that the N-terminal tail regulates pVHL solubility, which would become prone to aggregation when the N-terminus is lost (pVHL19), or reduced (VLP). Interestingly, VLP expression seems to be enriched in hypoxia-sensitive tissue, i.e. cerebellum, placenta and testis. High pVHL172expression levels have been observed during human embryogenesis in kidney, brain, spinal cord, eyes, testis, lung, and, in some cases, it is reported to be the only VHL transcript detected in sporadic ccRCC samples[47].

These observations are suggestive of two concepts, i.e. redundancy-for-robustness and redundancy-for-differentiate. The biological robustness is a property of complex organisms that allow a system to maintain its functions despite external and internal perturbations[265]. Being a hub protein, pVHL is classified as a conserved core protein in the robustness view, modulated by proteins acting as subordinated signals (modulators)[265]. In other word, pVHL well fits with the so-called bow-tie system, in which a large and highly connected core cluster is interfaced with less connected clusters[266]. The core of the structure is generally considered the Achilles' heel of the network, as if sufficiently perturbed or hijacked for alternative uses, it has a harmful effect on cellular homeostasis[267]. To guarantee system stability, opposing to external aggression such as mutations and metabolic imbalance, the central core elements frequently evolve in specialized isoforms with partially overlapping functions. The p53-like protein family is an example of this system[268]. The p53, p63 and p73 genes are a transcription factors family acting as regulator of the cell cycle progression. These are partially redundant even if each members owns its unique functions. A similar hypothesis might be traced for pVHL and its family members. When pVHL role is attacked, they (pVHL isoforms and VLP) co-operate to avoid failure; unlike, they maintain their own pathways. In this context, as an example, competitive binding to specific cellular partners could play a relevant role in determining alternative cell fates. Indeed, the different VHL-related proteins seem feature by structural alterations, in order to carry out different interactions and unrolling different cellular functions. For instance, the functional impact of the alternative isoforms has been negatively correlated in cancer-contest[269]. Recently, subset of transcript isoforms has been proposed as alternative-splicing drivers (AS drivers) characterized by independent-oncogenic activity. These driver-like properties may explain the reported pVHL172 expression in sporadic ccRCC[47]. Cell-specific isoforms switch may play a crucial role in driving neoplastic development and cancer phenotype-differentiation, especially in the case of VHL familial syndrome, where a same mutation does not even correlated to the same pathological-features in patients.

3.3 High-throughput cDNA Y2H library screening

As previously described, in 2004 the first human VHL paralog (VHL-like or VLP) has been identified[53]. Despite this, the paralog is still poorly studied leaving many questions open. Compared to pVHL30, I have already predicted structural similarities and validated its ability to maintain the interaction with the CDKN1 family and MDM2. However, it is also evident that the missing of a pVHL α-domain counterpart in VLP is a relevant element of differentiation. A high throughput proteome screening aimed at identifying new VLP interactors was then performed by Gal4-based yeast two-hybrid system library. In parallel, this screening was also designed to address whether VLP and pVHL30 share relevant interactors, indicative of possible overlapping pathways and functions. In particular, taken into the account VLP expression profile, I tested the Mate & Plate™ Human testis cDNA library (Clontech). In detail, the Y187 (MATa) *prey library* has been mated to the opposite mating types Y2HGold strain (MATα) expressing VLP as specific *bait* protein. The *prey library* was provided by Clontech, while the *bait* has been constructed (a detailed description is reported in section 2.8.5 materials and methods). Thus, VLP sequence has been cloned in-frame with the Gal4-DBD in yeast two-hybrid pGBKT7 vector using the In-Fusion® procedure. Verified by Sanger sequencing, the obtained clone has been transformed into Y2HGold yeast strain and the recombinant protein expression consequently determined by immunoblot assay. As reported in Suppl. Figure 12, the chimera protein (Gal4-DBD-VLP) was correctly detected at 32 kDa. Notably, bait expression did not alter either the yeast colony growth-rate or their dimension. These observations suggest that the bait is not toxic for yeast cells. It was also crucial to define that in absence of the prey protein the bait does not autonomously activate the reporter genes. Thus, pGBKT7 VLP alone has been transformed in Y2HGold strain and cells were plated on selective agar media (SD/-Trp, SD/-Trp/X and SD/-Trp/X/A). As expected, in presence of AbA drug no colony is visible on agar plate, while white colonies are present in SD/-Trp and SD/-Trp/X (data not shown). Before the mating experiment, a small-scale mating has been also performed. The overnight culture has been monitored for about 40 hours to estimate the mating duration. After 16 hours of incubation and then every 4 hours, a drop deriving from the co-culture has been inspected by a phase contrast microscope (40X). According to the manufacturer protocol, 3-lobed structure (also known as “Mickey Mouse” face-shape) may be visible, whether the mating appropriately occurs. This peculiar yeast shape represents two haploid parental cells and the budding diploid cell. After 20 hours, the 3-lobed-shape yeast cells have been detected. Of note, the presence of these structures is slightly controversial. Indeed when visible, the mating is clearly going on however the operator does not know in which mating-phase they are. In

other words, it is not possible to know whether the mating event is just started or it is almost finished. To address this problem, I decided to recover a co-culture aliquot after 24 hours, measured it by spectrometer and then spread a dilution on different media, i.e. YPGA, SD/-Trp/-URA, SD/-Leu/-Ura and SD/-Trp/-Leu/-Ura. According to spectrometer measurements, the incubated co-culture had reached OD_{600} ~13,60 containing about 40,98x10^4 cell/µl (OD_{600} ~ 3x10^7 cell/ml). Figure 40 shows that the mating culture has been growing healthy on the rich media (YPGA). On specific selective media, I observed the single-culture growth-rate, apparently diverging between the two strains. In particular, on SD/-Trp/-Ura a high number of bait-expressing colonies were grown, whereas library-expressing colonies were many fold less populated on SD/-Leu/-Ura. This could be reasonable considering that only a small amount of yeast library aliquot has been employed for this trail. Although this, few diploid colonies were present on SD/-Trp/-Leu/-Ura, confirming that mating correctly happens. A same procedure has been also repeated after 36-38 hours. The co-culture apparently stopped growing, with fewer diploids detectable on SD/-Trp/-Leu/-Ura suggesting that the incubation time should be shorter (plates not shown). Evaluating the performed control experiments and the user protocol suggestions, I chose to incubate the co-culture for 22-24 hours.

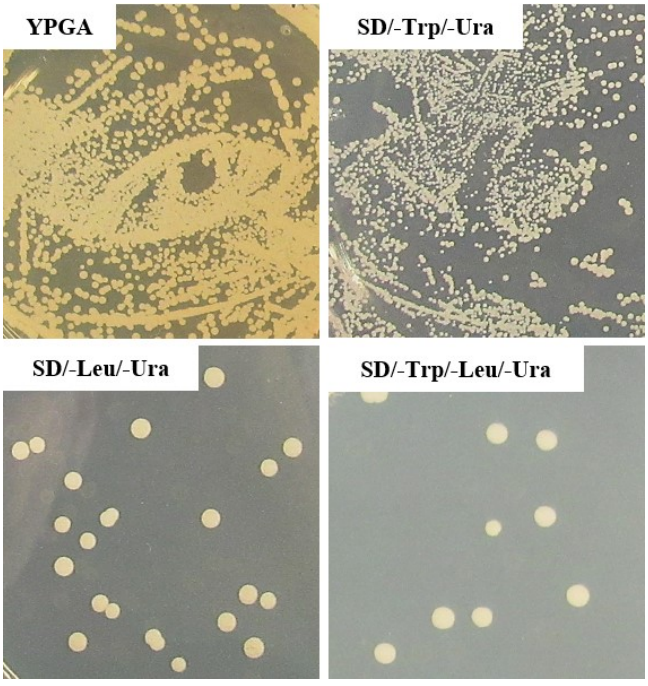


Figure 40. Small-scale mating controls. A mated culture aliquot has been spread on permissive and selective agar plates. Here is reported the 1:1000 dilution plate.

Afterward the strongly recommended control experiments, the large-scale mating has been performed as reported in Materials and Methods (section 2.8.5). The used library aliquot has been properly titer spreading 100 µl of 1/10, 1/100, 1/1000 and 1/10000 dilutions on SD/-Leu agar plates (Figure 41). Considering the 1/10000 dilutions, ~$8x10^4$ cells/µl were counted consistent with protocol expectation (i.e. > $2x10^7$ cells/ml).

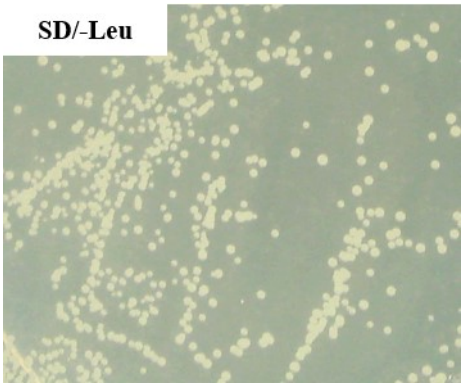


Figure 41. Titration of the library aliquot. Different dilution were spread on SD/-Leu medium in order to titre the library aliquot. Here is reported the 1:10000 dilution, used to count cells.

In addition, it was essential to calculate the number of screened clones in order to determine the mating efficiency. Thus, an aliquot of the 22 hours-mated culture has been also spread on SD/-Trp, SD/-Leu, SD/-Trp/-Leu and colonies were manually counted (Figure 42).

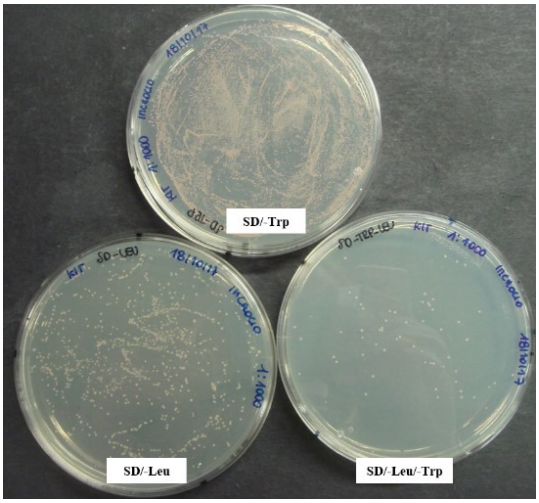


Figure 42. Diploids count. After 22 hours of mating, a co-culture aliquot was spread on selective agar plates in order to calculate the number of screened clones. Here is reported the 1:1000 dilution plate.

Data were collected in the following Table 16 :

Media	*N. clones*	*Viability*	*Resusp. Vol.*	*Plating Vol.*	*Dilution*	*Screened clones*
SD/-Leu	134	Prey Library	11.4 ml	100 µl	1:10000	$1{,}53x10^8$
SD/-Trp	2400	Bait	11.4 ml	100 µl	1:10000	$2{,}74x10^9$
SD/-Leu/-Trp	62	Diploids	11.4 ml	100 µl	1:1000	$7{,}0x10^6$

Table 16. Cell viability measurements. Bait, prey library and diploids viabilities are respectively calculated multiplying the number colonies on the plates with the resuspension volume before plating the mated culture, with the plating volume and the considered dilution.

Using the viability measures was possible to calculate the mating efficiency (percentage of diploids):

$$\frac{No.\,of\ cfu/ml\ of\ diploids}{No.\,of\ cfu/ml\ of\ limiting\ partner}\ x\ 100 = \%\ diploids$$

The strain with the lowest viability was the "limiting partner", in this case the prey library. Thus, VLP screening achieved 4,57% mating efficiency accordingly with the estimated range of 2-5%. In the meanwhile, the mated plates were monitored for 5 days (twice a day). As reported in the Table 17**Table 1**, we collected 134 blue positive colonies overall. According to the recorded data (Figure 43), the colonies growth "peak" occurred in the first two days decreasing progressively.

Day	N. Clone	N. Clone/day	Total
2	1-32	57	57
	33-57		
3	58-84	48	105
	85-105		
4	106-117	22	127
	118-127		
5	128-130	7	**134**
	131-134		

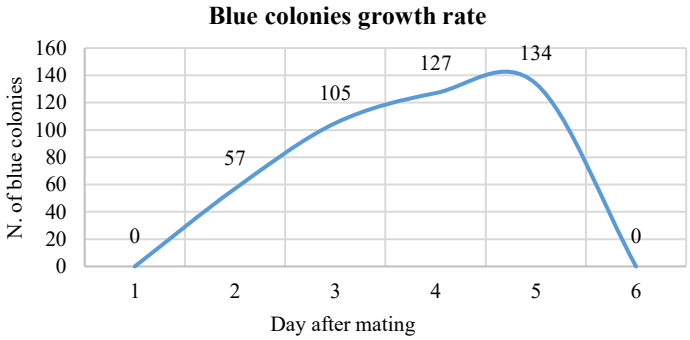


Figure 43. (Previous page, on the right) Positive blue colonies growth rate. In the graph 'Number of clones' is plotted to 'day after mating' showing an increasing number of positives, collapsed at the 6th day.

Table 17. (Previous page, on the left) Clones monitoring. Blue colonies selection was collected for 5 days, twice a day. Here, the number of clones per day and the final total number are reported.

Screened clones were then validated on more stringent agar plates (QDO). As a result (Figure 44 and Suppl. Figure 13), 104 out of 134 clones retained the ability to grow and turn blue on selective media. The 30 false positive clones were phenotypically distinguished into different groups: i) able to retain the blue color, but unable to grow further; ii) growing but not showing blue color (i.e. one of the four reporter genes was not activated); iii) unable to grow on selective media albeit retaining blue color.

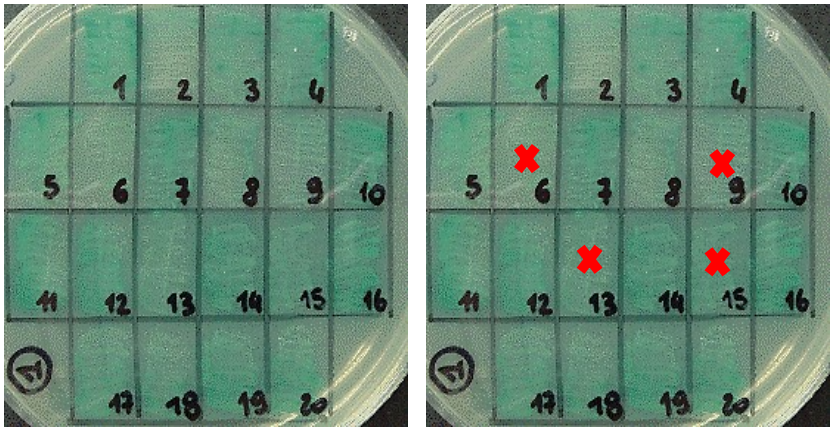


Figure 44. Colony patches. The blue colonies were patched out on higher stringency agar plates (QDO/X/A) in order to avoid false positive interactions. Here, an exemplificative picture (other clones are in *Suppl. Figure 13*) representing clone 1 to 20 patches (left) is reported. Red crosses indicate four false positive clones that were not further analysed (right).

At this level, it was necessary to find out the DNA sequences of the confirmed positives. Thus, yeast DNA (i.e. genome, bait and prey plasmids) was extracted from the screened clones and analyze by PCR reaction. The PCR products were obtained using Gal4 AD forward and reverse (Suppl. Table 1) primers able to pair the vector sequences upstream and downstream MCS of pGADT7-Rec. Results were detected running PCR products on 0.8% agarose gels. Figure 45 shows the first group of positive clones cDNA inserts (see also Suppl. Figure 14), which were included in 0.5 kb to 4.0 kb range as expected by the quality control of the original cDNA library. Notably, some clones were characterized

by double inserts presence (i.e. clone 17 and clone 18), suggesting that a single colony encloses two different yeast populations (namely "*mixed samples*"). As consequence, it was necessary to distinguish the two different VLP interactors expressed by these samples and whether both were able or not to sustain the interaction. Following the manufacturer instructions, this PCR step might be useful to eliminate duplicates clones containing the same AD/library plasmid. Although this suggestion, it seems to be not possible to suppose *a priori* that samples with equal-length inserts are really duplicated. For this reason, I carried on the further analysis with the all set of positives.

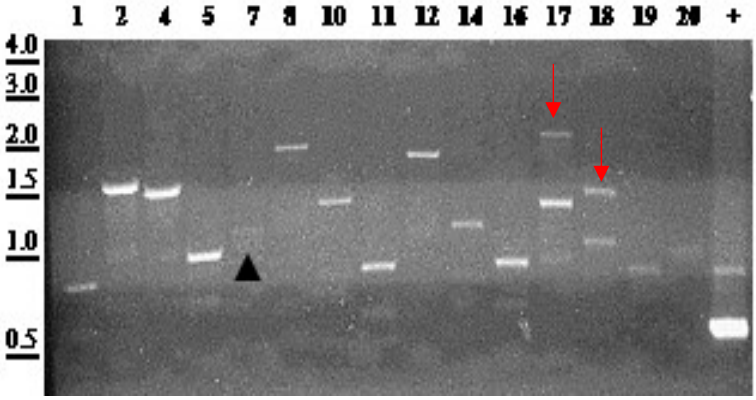


Figure 45. AD/plasmid insert amplifications. The first 15 PCR products were differently separated on 0.8% agarose gel. (Top) Each line has a number, which corresponds to a selected blue clone, whereas + is the positive control (i.e. a pGADT7 sample). The negative control was loaded in a second part of the gel (see also *Suppl. Figure 14*). The black triangle indicates insert amplification of clone 7, which seems to be less visible than the others do. Red arrows point clones characterized by double inserts, i.e. 17 and 18.

Rescued yeast DNA was transformed in *E. coli* in the following step in order to amplify and discern the AD/library plasmids from yeast genome and *bait* plasmid insert. A further step of plasmid purification was performed and different pGADT7-Rec collected. For each sample, I recovered three independent clones from LBA agar plates and two out of three were analyzed by restriction enzyme reaction (i.e. using EcoRI and BamHI enzymes). As expected, the two independent clones show the same restriction pattern on 0.8% agarose gel running (Figure 46 and Suppl. Figure 14). Differently, five independent clones were handled in the case of *mixed samples* in order to discriminate between the two different pGADT7-Rec populations. For instance, Figure 47 shows that clone 17 and clone 18 were *mixed samples* presenting two AD/plasmid with different restriction pattern.

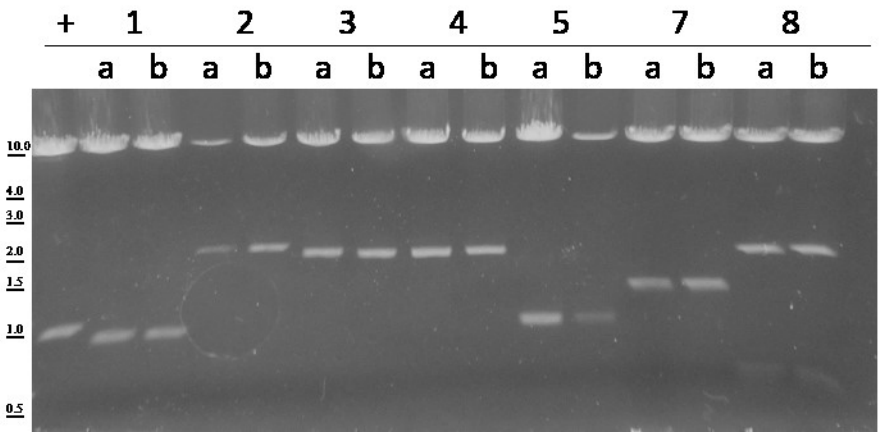


Figure 46. Clones restriction analysis. For each positive, a couple of clones (i.e. a and b) were analyzed by restriction reaction using EcoRI and BamHI enzymes at 37°C. Each line has a number, which properly corresponds to a selected clone, whereas positive control is indicated with + (see also *Suppl. Figure 14*).

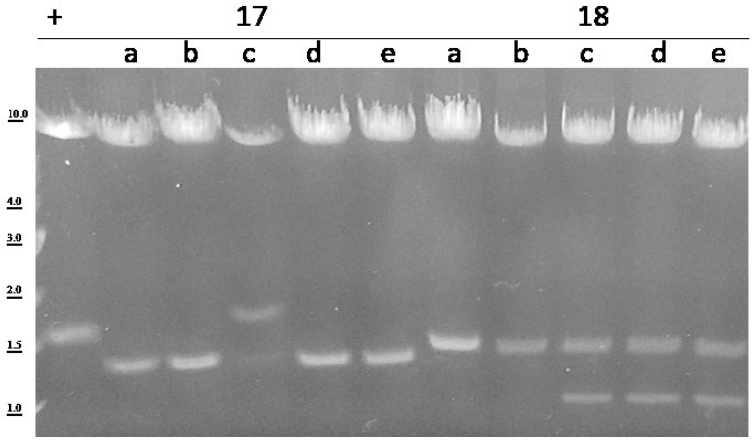


Figure 47. Mixed clones restriction pattern. Exemplificative picture of two mixed clones, i.e. 17 and 18. Five independent clones (i.e. a, b, c, d and e) were analysis by EcoRI and BamHI restriction enzymes, revealing in both cases the presence of two distinct AD/plasmid populations.

At this point, interactions have to be further verified. Although following the library user manual the incidence of false positive was reduced to a minimum compared to another system, it is known that with every two-hybrid screening there is a chance of detecting false positives. For this purpose, it was necessary to validate the novel putative bindings by new yeast transformations. Changing genetic background, yeast transformations were carried out in Y190 strain, largely utilized in the previous results section chapters (section 3.1 and 3.2) to perform a canonical Y2H assay. Therefore, each

collected plasmid was co-transformed with either pGBKT7-VLP or pGBKT7-empty (Ø) vector (i.e. auto-activation control) and spread on SD/-Trp/-Leu agar plates. Then, from each transformation, three independent colonies were patched out on permissive (i.e SD/-Trp/-Leu) and selective media (i.e. SD/-Trp/-Leu/-His supplemented with 30mM 3AT) as shown in Figure 48 (and Suppl. Figure 16). Where interactions were genuine, co-transformed yeast growth was strongly sustained on selective media suggesting the correctly HIS3 reporter gene activation. As shown in Figure 48, the three independent clones of 1.1 and 3.1 samples (i.e. a, b, and c) markedly bind to VLP on selective media (*right*), similarly to the positive control (C+). On the contrary, 2.1a, 2.1b and 2.1c clones were characterized by a weaker interaction. Indeed, compared to their auto activation (in particular the auto-activation clone a) and VLP auto activation (clone b), it was difficult to certainly assess whether a real interaction occurred.

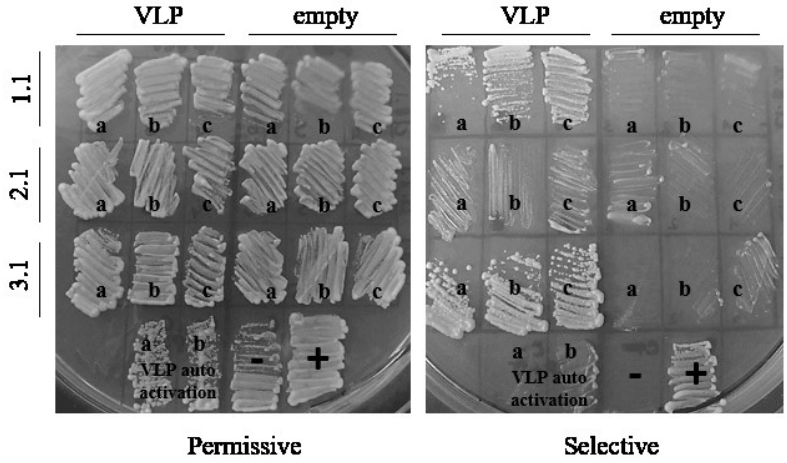


Figure 48. Yeast patches. On permissive and selective media, three independent clones were tested. The numbers indicate the analyzed clones (i.e. 1.1, 2.1, 3.1) and the letters (i.e. a, b and c) point out the three clones picked up from the same co-transformation plates containing either pGBKT7 VLP or pGBKT7 empty (Ø) vectors. Standard Y2H positive and negative controls are + and –, whereas VLP auto activation are yeast cells co-transformed with pGKBT7 VLP and pGADT7 empty vectors. In *Suppl. Figure 16* are reported all the other screened clones.

Similar to sample 2.1, some other outcomes were controversial due to the high background of the negative control or high level of auto-activation. Thus, the drop test has been employed on permissive and selective media (i.e. supplemented with 30mM 3AT and 60mM 3AT) to fix debated results. Serial dilutions of yeast cells pool were compared with positive, negative and auto-activation controls as in a canonical drop test. Figure 49 confirmed patches results, highlighting that the clone 2.1 interacts with VLP in a weaker-manner. As a result, 96 out of 104 novel VLP interactors were verified, whereas

eight samples (i.e. clones 18.1, 18.2, 45, 70, 76.2, 76.5, 88.3 and 108.2) were classified as false positives due to growth absence on selective medium.

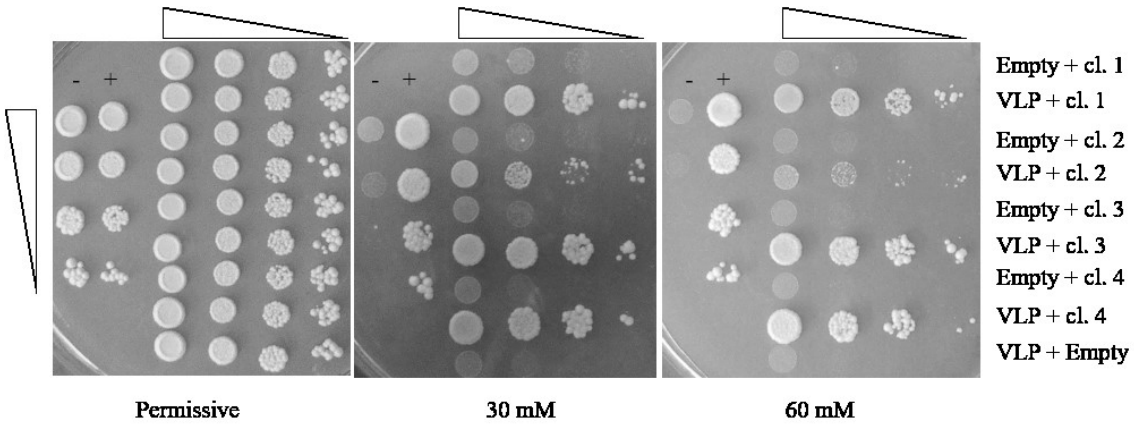


Figure 49. Drop test assay on library clones. Yeats cells serial dilutions were spotted on permissive and selective media (i.e. supplemented with 30mM 3AT and 60mM 3AT). Positive (+) and negative (-) controls properly indicate yeast cells growth and lethality. Similarly, auto-activation controls (i.e. empty vector co-transformed with VLP or a library clone) point out that the detected interactions are related to simultaneous bait and prey proteins expression. On 60mM 3AT, yeast cells co-transformed with VLP (as bait) and a library clone (i.e. 1, 2, 3 and 4) are able to survive in absence of histidine due to their interaction. In *Suppl. Figure 18* are collected all the clones screened.

Once these interactions were also validated in another yeast strain, *prey* inserts were sequenced by the Sanger method using T7 sequencing forward primer (GATC service). Sequencing outcomes were examined by the translated BLAST tool tblastx (NCBI, https://blast.ncbi.nlm.nih.gov) as reported in section # (Material and Methods). Most of the sequences have been detected only once, suggesting that in testis VLP associate different protein partners. Resulting sequences were organized in two sub-datasets, annotated- and unannotated-dataset, containing respectively 50 and 33 clones (

Suppl. Table **3**). The former groups proteins for which manually curated data are available in public database, whereas the latter includes unnoted human proteins. The remaining 13 clones were considered false positive as sequence analysis showed either undeterminable correspondence with human proteins (e.g. very short sequence) or vast sequencing errors. I believe these samples to derive from aberrant sequences generated during the library preparation. Indeed, as reported in the manufacturer protocol, this problem could not be reduced due to yeast naturally tolerate translational frameshifts to survive. Few protein responsible for the interaction correspond to large ORF in the

wrong reading frame, while other sequences (e.g. clones 44, 46, 74, 92, 98 or 103) seemed to correspond to a known gene region, but not included in the coding-sequence (CDS). Thus, they completely or partially included 5'- or 3'-UTR regions. In other cases, obtained sequences were too short (such as clones 8 or 53.2) or corresponded to pseudogenes (as clone 107). Re-clone the insert in-frame in a new vector could be reasonable to verify whether the four different reports are still transcriptionally activated in presence of the *bait*-protein. For a few of them, the sequenced fragment resulted too short or characterized by 5'- or 3'-untranslated region (UTR) to be correctly identified in a homology-based screening. Especially, the sequencing reactions of 2 out of 16 clones (i.e. 25 and 118.4) failed twice, even if sequencing primers were changed (i.e. Gal4-AD forward and reverse). As mentioned by the user manual, it could be that these positive library clones were transcribed in the reverse orientation from a cryptic promoter within the ADH1 terminator on the *bait* plasmid[270]. Collectively taken, these samples were defined as non-classified clones. The annotated-dataset samples have been further characterized by computational approaches. First an analysis of their sub-cellular localization (Figure 50) revealed nucleus and cytosol to be the preferred localizations of VLP binding partners, with 18 and 12 proteins each respectively. Groups accounting for proteins localizing in either membrane or mitochondrion (ca. 6 proteins) resulted less populated. A limited number of interactors were associated to lysosome, endosome, Golgi as well as endoplasmic reticulum (ca. 2 clones for each organelle). Interestingly, two entries encoded secreted proteins, i.e. adenosine deamidase 2 (ADA2) and adhesion G protein-coupled receptor G1 (ADGRG1 or GPR56). ADA2 regulates adenosine levels controlling a variety of cell signaling. Poorly studied, it behaves in a growth factor-manner acting on inflammation and cancer[271]. GPR56 belongs to the highly conserved GPCR family, implicating both in cell-adhesion and G protein-coupled signaling. In melanoma progression, it regulates angiogenesis through VEGF inhibition[272]. The GPR56 interaction may reinforce the *dominant-negative* role already connected to VLP[53], blocking angiogenesis along the HIF-1α downstream-pathway. Future investigations may be required to shed light on this possible cancer-related mechanism.

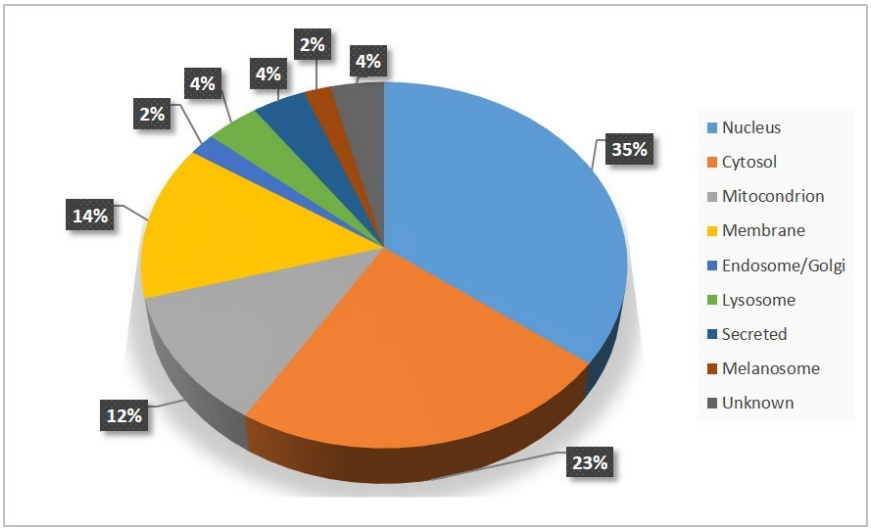


Figure 50. Sub-cellular localization of VLP interactors.

A preliminary protein interaction network built on annotated-dataset of interactors showed no trivial connections among them (Suppl. Figure 19). This data reinforces the idea of VLP acting as a *hub* protein. The interaction network was hence expanded taking into the account VLP *interactors-of-interactors* (Figure 51). The functional network obtained shows 92 node and 238 edges, with an average node degree of 4.96. Cluster analysis of the resulting network shows VLP partners to be associated to multiple relevant cellular pathways. Three major and populated clusters indicate VLP to be connected with pathways regulating extracellular adhesion (9 proteins), cytoskeleton remodeling (7 proteins), as well as to interact with a number of transcription factor mediating cell survival and proliferation (15 proteins). In particular, the microtubule and external matrix homeostasis are shared functions with the full-length VHL[273]. In light of VLP structural peculiarity, these findings suggest that this alternative pVHL functions could be sustained by the pVHL β-domain. Thus, it could be interesting to test pathogenic mutations affecting this domain in future investigations. Other less populated functional clusters were related to RNA regulation, sumoylation, cell motility, sugar metabolism and flagellar remodeling and spermatogenesis. Interestingly, no hypoxia-related pathways interactors were found. Rather than being unexpected, this finding can be due to the specific oxygen requirements of testis. On the other hand, this difference may indicate a binding preference clearly distinguishing VLP from pVHL30-partners.

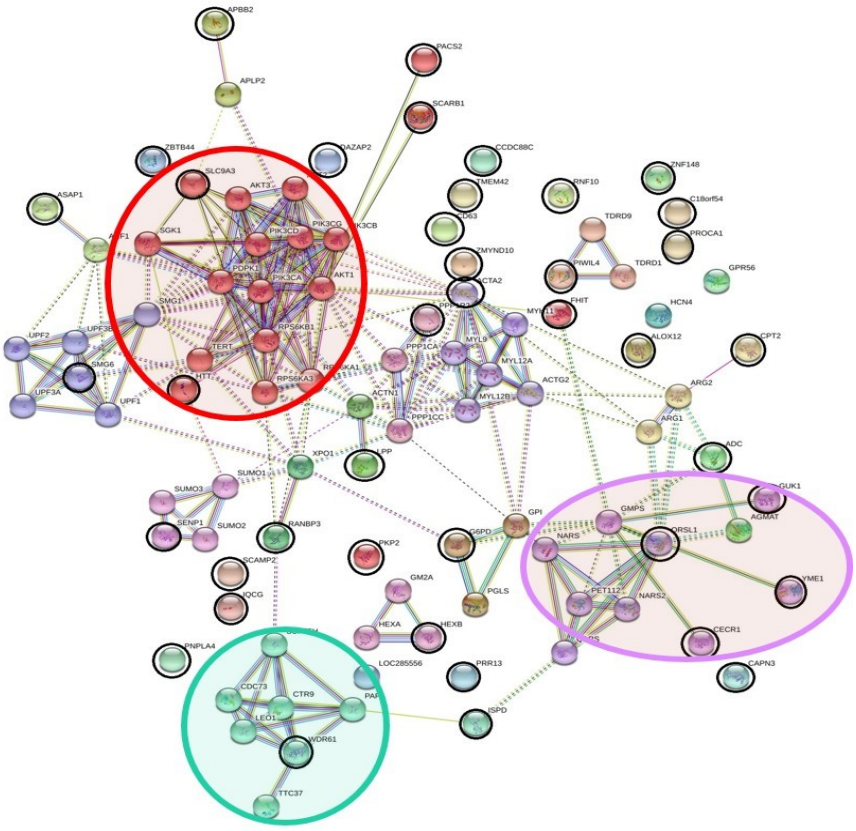


Figure 51. STRING functional network. Three populated clusters are highlighted and they represent pathways regulating extracellular adhesion (9 proteins, pink circle), cytoskeleton remodelling (7 proteins, light blue circle), as well as to interact with a number of transcription factor mediating cell survival and proliferation (15 proteins, red circle). A black circle marked the VLP annotated-dataset interactors.

The interacting fragments were further examined with InterProScan tool. The software package allows sequences to be scanned against InterPro[207] database, which predicts the presence of functional domains classifying protein sequences into families. Similar to pVHL30, VLP recognizes functional domains, such as armadillo repeated domain (ARM), LIM, WD and zinc finger (ZnF). Interestingly, these domains are all classified as scaffold structure able to interact and/or bind multiple proteins tethering them into complexes regulating crucial cell signaling pathways. Respectively, the ARM domain is characterized by 40-repeated residues, typically in tandemly repeated copies[274]. An armadillo unit is composed by a pair of alpha helices forming a hairpin structure, where many repeated units form an alpha solenoid structure[274]. For instance, the ARM domain is included both in α-importin[275] and β-catenin[276] proteins. Differently, LIM is a protein structural module composed of two contiguous zinc finger domains separated by a two-amino acid residue hydrophobic linker. LIM-domain containing proteins are a large family emerging as key molecules in a wide variety of human cancer[277]. WD domain instead folds into a β-propeller structure, providing a platform for the

interaction and assembly of several proteins[278], while ZnF contain multiple finger-like protrusions that make tandem contacts with their target molecule displaying an extraordinary binding plasticity to bind RNA, lipids, proteins, and protein post-translational modifications (PTMs)[279]. Notably, VLP shows also a putative preference for linear binding motifs localizing in the N- and/or C-terminal protein regions. Considering the structural similarity with pVHL30, I further compared the identified VLP interactors to VHLdb, a database collecting VHL interactors and mutations[3]. I found that the lipoma-preferred partner (LPP)[280], a LIM-domain containing protein implicated in cell motility and cancer, is a VLP/pVHL30 shared interactor. In particular, LPP is proposed to play a structural role at sites of cell adhesion in maintaining cell shape and motility[280]. Considering both the novelty of this tissue-specific library and the ubiquitous pVHL30-expression, it can be speculated that most of the interactions emerged in this screening can be also extended to pVHL30, thus representing a first example of tissue-specific pVHL-partners. Their experimental validation, for example using VHL30 as *bait* in a similar screening, will be crucial to address this hypothesis. On the other hand, the VLP-dataset may represent a gene specialization evolved to account for tissue-specific requirements. As arguments supporting this alternative hypothesis, I found VLP to interact with Solute Carrier Family 9 Member A3 (SLC9A3). To the best of our knowledge, no evidence of SLC9A3 interaction with pVHL30 were found in the literature. Whereas SLC2A1 and SLC3A2, two other members belonging from the same protein family, were experimentally validated pVHL30 interactors, also reported in VHLdb. These three proteins are respectively associated to Na/H^+, glucose and light chain amino acid transport and their differential interactions may respond to tissue-specific specializations. In other words, it can be proposed that a conserved mechanism, i.e. the general association of pVHL30 β-domain with solute carrier transporters, is tissue specific and embracing multiple pathways. If confirmed, these evidences can help to explain how functional impairment of the VHL-protein family promotes cancer insurgence.

Discussion and Conclusions

Recently we described a novel association between the pVHL tumor suppressor and CDKN1 family members[263]. This interaction is sustained by the pVHL β domain, which recognizes a binding motif conserved among CDKN1s. We proposed that this interaction links, somewhat more widely, the cell cycle progression to the hypoxia adaptive response[263]. Importantly, this binding points out an alternative pVHL function related to the regulation of the cell cycle[46]. The pVHL is a small proteins presenting with a limited number of binding surfaces[41]. Due to this physical constrain, its interaction with different partners should be forcedly regulated in both time and space, as well as taking under consideration the physiological or pathological cell conditions. In this complex scenario, pVHL seems to play an essential role in the cellular decision-making process. In healthy conditions, pVHL is known to bound HIF-1α regulating its stability and transcriptional activity[210]. The novel interaction presented in chapter 1 of this manuscript, however, opens a new possible scenario. The CDKN1s associate pVHL in a hydroxylation-independent fashion, as yeast cells used to identify and dissect their interactions are devoid of specific metazoan enzymes[263]. It could be thus assumed that in both mild and prolonged hypoxia pVHL, being not strictly bound to HIF-1α, could be prone to sustain alternative interactions possibly relevant for other cellular pathways. In this way, the pVHL "multipurpose" interactions are in turn indirectly handled by oxygen level in cells. These alternative decisional states could play a crucial role in tumor onset. Malignant solid tumor mass is characterized by an O_2 concentration gradient and a combination of different tumor-derived cell subtypes that respond differently to an ever-changing cellular microenvironment. Indeed, the tumor mass generally presents with a central region formed by cordons of necrotic cells experiencing severe hypoxia, while tumor cells maintaining expanding abilities and sustained motility are prominent gradually moving to the proximity of capillaries and vascularized areas[281]. As a *protein hub,* the numerous fluctuating pVHL interactions could thus drive cells forward several fates, for example gradually shaping the tumor progression. Furthermore, the presence of pVHL isoforms and paralog further complicate this possible scenario[53,47,51]. These proteins are clearly characterized by both structural similarities and differences whether compared to pVHL30. As reported in chapter two, these features are associated to isoform-specific interactions, assayed by Y2H screening system, as well as to their peculiar intracellular distribution. Borrowing a paradigm from the system biology, pVHL could be proposed to play the double role of family-master regulator and conserved *hub* element driving crucial cell functions. Indeed, upon interacting with different partners (or when mutated in pathological conditions), its activity could be orchestrated by the expression of its alternative isoforms, which in turn may participate or compete pVHL30 for specific functions. The majority of human genes are

alternatively spliced[282] and many isoforms are also associated to cancer progression, either supporting the tumor therapeutic resistance or playing as biomarkers in particular types of cancers[283]. In this sense, the p53 tumor-suppressor, the so-called "guardian of the genome", is characterized by 12 isoforms differently truncated in N- or C-terminus and is probably the best-known example[284]. It has been observed that abnormal and altered expression of a particular p53 isoform drives insurgence of a specific cancer by promoting functional inhibition of the main full-length p53[285]. A similar condition could be proposed for the pVHL family members, however, further investigations are necessary to validate this hypothesis. Collectively taken, the results presented in this work, highlighting multiple pVHL isoform-specific specializations, may represent a step forward in this direction. Extensive future studies will be necessary to shed light on the multifunctional pVHL and its dense network of interactions.

References

1. Fouad, Y. A. & Aanei, C. Revisiting the hallmarks of cancer. *American Journal of Cancer Research* **7,** 1016–1036 (2017).
2. Chittiboina, P. & Lonser, R. R. Von Hippel–Lindau disease. in *Handbook of Clinical Neurology* **132,** 139–156 (2015).
3. Tabaro, F. *et al.* VHLdb: A database of von Hippel-Lindau protein interactors and mutations. *Sci. Rep.* **6,** (2016).
4. Minervini, G. *et al.* Isoform-specific interactions of the von Hippel-Lindau tumor suppressor protein. *Nat. Publ. Gr.* 1–9 doi:10.1038/srep12605
5. Global Burden of Disease. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* **390,** 1151–1210 (2017).
6. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100,** 57–70 (2000).
7. Hanahan, D. *et al.* The Hallmarks of Cancer. *Cell* **100,** 57–70 (2014).
8. Michor, F., Iwasa, Y. & Nowak, M. A. Dynamics of cancer progression. *Nature Reviews Cancer* (2004). doi:10.1038/nrc1295
9. Kinzler, K. W. & Vogelstein, B. Gatekeepers and caretakers. *Nature* (1997). doi:10.1038/386761a0
10. Bissel, M. J. & Radisky, D. Putting tumors in context. *Nat. Rev. Cancer* (2001). doi:10.1038/35094059.PUTTING
11. Friend, S. H. *et al.* A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* (1986). doi:10.1038/323643a0
12. Garber, J. E. & Offit, K. Hereditary cancer predisposition syndromes. *J. Clin. Oncol.* (2005). doi:10.1200/JCO.2005.10.042
13. Varley, J. M. Germline TP53 mutations and Li-Fraumeni syndrome. *Human Mutation* (2003). doi:10.1002/humu.10185
14. Nordstrom-O'Brien, M. *et al.* Genetic analysis of von Hippel-Lindau disease. *Hum. Mutat.* **31,** n/a-n/a (2010).
15. Hippel, E. Über eine sehr seltene Erkrankung der Netzhaut. *Albr. von Græfe's Arch. für*

Ophthalmol. **59,** 83–106 (1904).

16. LINDAU, A. ZUR FRAGE DER ANGIOMATOSIS RETINae UND IHRER HIRNKOMPLIKATIONEN. *Acta Ophthalmol.* **4,** 193–226 (2009).

17. Latif, F. *et al.* Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science* (1993). doi:10.1002/anie.201102909

18. Van Asselt, S. J. *et al.* Pancreatic cyst development: Insights from von Hippel-Lindau disease. *Cilia* **2,** (2013).

19. Choo, D. *et al.* Endolymphatic sac tumors in von Hippel-Lindau disease. *J. Neurosurg.* **100,** 480–7 (2004).

20. Abd Hamid, D., Abdullah, J., Ariff, A., Muhamad, M. & Madhavan, M. Cerebellar hemangioblastoma in a patient with von hippel-lindau disease : a case report. *Malays. J. Med. Sci.* **7,** 43–8 (2000).

21. Rebollar GonzÃ¡lez, S. *et al.* Pheochromocytoma associated with von Hippel-Lindau disease. *Rev. del Lab. Clin.* **9,** (2016).

22. Clark, P. E. The role of VHL in clear-cell renal cell carcinoma and its relation to targeted therapy. *Kidney International* **76,** 939–945 (2009).

23. Gallou, C. *et al.* Mutations of the VHL gene in sporadic renal cell carcinoma: Definition of a risk factor for VHL patients to develop an RCC. *Hum. Mutat.* **13,** 464–475 (1999).

24. Lopez-Beltran, A. *et al.* 2009 update on the classification of renal epithelial tumors in adults. *International Journal of Urology* (2009). doi:10.1111/j.1442-2042.2009.02302.x

25. Nielsen, S. M. *et al.* Von Hippel-Lindau disease: Genetics and role of genetic counseling in a multiple neoplasia syndrome. *J. Clin. Oncol.* (2016). doi:10.1200/JCO.2015.65.6140

26. Chen, F. *et al.* Germline mutations in the von Hippel-Lindau disease tumor suppressor gene: correlations with phenotype. *Hum Mutat* **5,** 66–75 (1995).

27. Maher, E. R. & Kaelin Jr., W. G. von Hippel-Lindau disease. *Med.* **76,** 381–391 (1997).

28. Ang, S. O. *et al.* Disruption of oxygen homeostasis underlies congenital Chuvash polycythemia. *Nat. Genet.* **32,** 614–621 (2002).

29. Prowse, A. H. *et al.* Somatic inactivation of the VHL gene in Von Hippel-Lindau disease tumors. *Am. J. Hum. Genet.* (1997).

30. Nordstrom-O'Brien, M. *et al.* Genetic analysis of von Hippel-Lindau disease. *Human Mutation*

(2010). doi:10.1002/humu.21219

31. Beroukhim, R. *et al.* Patterns of gene expression and copy-number alterations in von-Hippel Lindau disease-associated and sporadic clear cell carcinoma of the kidney. *Cancer Res.* (2009). doi:10.1158/0008-5472.CAN-09-0146

32. Knudson, A. G. Mutation and Cancer: Statistical Study of Retinoblastoma. *Proc. Natl. Acad. Sci.* **68,** 820–823 (1971).

33. Richards, F. M. *et al.* Detailed genetic mapping of the von Hippel-Lindau disease tumour suppressor gene. *J Med Genet* **30,** 104–107 (1993).

34. Lee, S. *et al.* Neuronal apoptosis linked to EglN3 prolyl hydroxylase and familial pheochromocytoma genes: Developmental culling and cancer. *Cancer Cell* **8,** 155–167 (2005).

35. Okuda, H. *et al.* The von Hippel-Lindau tumor suppressor protein mediates ubiquitination of activated atypical protein kinase C. *J. Biol. Chem.* **276,** 43611–43617 (2001).

36. Woodward, E. R. *et al.* Comparative sequence analysis of the VHL tumor suppressor gene. *Genomics* **65,** 253–65 (2000).

37. Kuzmin, I. *et al.* Identification of the promoter of the human von Hippel-Lindau disease tumor suppressor gene. *Oncogene* **10,** 2185–2194 (1995).

38. Stebbins, C. E., Jr, W. G. K. & Pavletich, N. P. Structure of the Complex : Implications for VHL Tumor Suppressor Function. (2000).

39. Dandanell, M., Friis-Hansen, L., Sunde, L., Nielsen, F. C. & Hansen, T. V. O. Identification of 3 novel VHL germ-line mutations in Danish VHL patients. *BMC Med. Genet.* (2012). doi:10.1186/1471-2350-13-54

40. Min, J.-H. *et al.* Structure of an HIF-1alpha -pVHL Complex: Hydroxyproline Recognition in Signaling. *Science (80-.).* **296,** 1886–1889 (2002).

41. Leonardi, E., Murgia, A. & Tosatto, S. C. E. Adding structural information to the von Hippel-Lindau (VHL) tumor suppressor interaction network. *FEBS Lett.* **583,** 3704–3710 (2009).

42. Schoenfeld, A., Davidowitz, E. J. & Burk, R. D. A second major native von Hippel-Lindau gene product, initiated from an internal translation start site, functions as a tumor suppressor. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 8817–8822 (1998).

43. Iliopoulos, O., Ohh, M. & Kaelin, W. G. pVHL19 is a biologically active product of the von Hippel-Lindau gene arising from internal translation initiation. *Proc. Natl. Acad. Sci. U. S. A.*

95, 11661–6 (1998).

44. Richards, F. M., Schofield, P. N., Fleming, S. & Maher, E. R. Expression of the von Hippel-Lindau disease tumour suppressor gene during human embryogenesis. *Hum Mol Genet* **5,** 639–644 (1996).

45. Giancotti, F. G. & Ruoslahti, E. Elevated levels of the alpha 5 beta 1 fibronectin receptor suppress the transformed phenotype of Chinese hamster ovary cells. *Cell* **60,** 849–859 (1990).

46. Roe, J.-S. *et al.* p53 Stabilization and Transactivation by a von Hippel-Lindau Protein. *Mol. Cell* **22,** 395–405 (2006).

47. Chesnel, F. *et al.* The von Hippel–Lindau tumour suppressor gene: uncovering the expression of the pVHL172 isoform. *Br. J. Cancer* **113,** 336–344 (2015).

48. Taylor, C., Craven, R. A., Harnden, P., Selby, P. J. & Banks, R. E. Determination of the consequences of VHL mutations on VHL transcripts in renal cell carcinoma. *Int. J. Oncol.* **41,** 1229–1240 (2012).

49. Martella, M. *et al.* Molecular analysis of two uncharacterized sequence variants of the VHL gene. *J. Hum. Genet.* **51,** 964–8 (2006).

50. Hascoet, P. *et al.* The pVHL 172 isoform is not a tumor suppressor and up-regulates a subset of pro-tumorigenic genes including TGFB1 and MMP13. (2017).

51. Liu, Y. *et al.* A novel VHLα isoform inhibits Warburg effect via modulation of PKM splicing. *Tumor Biol.* **37,** 13649–13657 (2016).

52. Kamura, T. *et al.* Activation of HIF1alpha ubiquitination by a reconstituted von Hippel-Lindau (VHL) tumor suppressor complex. *Proc. Natl. Acad. Sci.* **97,** 10430–10435 (2000).

53. Qi, H. *et al.* Molecular Cloning and Characterization of the von Hippel-Lindau-Like Protein. **2,** 43–52 (2004).

54. Kaelin, W. G. The von Hippel-Lindau tumour suppressor protein: O2 sensing and cancer. *Nature Reviews Cancer* **8,** 865–873 (2008).

55. Haase, V. H. The VHL tumor suppressor: master regulator of HIF. *Curr. Pharm. Des.* **15,** 3895–3903 (2009).

56. Ohh, M. *et al.* Ubiquitination of hypoxia-inducible factor requires direct binding to the beta-domain of the von Hippel-Lindau protein. *Nat. Cell Biol.* **2,** 423–427 (2000).

57. Schanza, L. M. *et al.* MicroRNAs associated with Von Hippel-Lindau pathway in renal cell

carcinoma: A comprehensive review. *International Journal of Molecular Sciences* (2017). doi:10.3390/ijms18112495

58. Prabhakar, N. R. & Semenza, G. L. Oxygen Sensing and Homeostasis. *Physiology (Bethesda).* **30,** 340–8 (2015).

59. Bertout, J. A., Patel, S. A. & Simon, M. C. The impact of O2 availability on human cancer. *Nat. Rev. Cancer* (2008). doi:10.1038/nrc2540

60. Müller, M. *et al.* Renocortical tissue oxygen pressure measurements in patients undergoing living donor kidney transplantation. *Anesth. Analg.* **87,** 474–6 (1998).

61. Hoffman, W. E., Charbel, F. T. & Edelman, G. Brain tissue oxygen, carbon dioxide, and pH in neurosurgical patients at risk for ischemia. *Anesth. Analg.* (1996). doi:10.1097/00000539-199603000-00027

62. Semenza, G. L. Hypoxia, clonal selection, and the role of HIF-1 in tumor progression. *Critical Reviews in Biochemistry and Molecular Biology* (2000). doi:10.1080/10409230091169186

63. Jungermann, K. & Kietzmann, T. Oxygen: Modulator of metabolic zonation and disease of the liver. *Hepatology* (2000). doi:10.1002/hep.510310201

64. Vaupel, P. & Mayer, A. Hypoxia in cancer: Significance and impact on clinical outcome. *Cancer and Metastasis Reviews* (2007). doi:10.1007/s10555-007-9055-1

65. Axelson, H., Fredlund, E., Ovenberger, M., Landberg, G. & Påhlman, S. Hypoxia-induced dedifferentiation of tumor cells - A mechanism behind heterogeneity and aggressiveness of solid tumors. *Seminars in Cell and Developmental Biology* (2005). doi:10.1016/j.semcdb.2005.03.007

66. Jiang, B. H., Rue, E., Wang, G. L., Roe, R. & Semenza, G. L. Dimerization, DNA binding, and transactivation properties of hypoxia-inducible factor 1. *J. Biol. Chem.* **271,** 17771–8 (1996).

67. Koh, M. Y. & Powis, G. Passing the baton: The HIF switch. *Trends in Biochemical Sciences* (2012). doi:10.1016/j.tibs.2012.06.004

68. Masoud, G. N. & Li, W. HIF-1α pathway: Role, regulation and intervention for cancer therapy. *Acta Pharmaceutica Sinica B* **5,** 378–389 (2015).

69. Bruick, R. K. & McKnight, S. L. A conserved family of prolyl-4-hydroxylases that modify HIF. *Science (80-.).* **294,** 1337–1340 (2001).

70. Hu, C.-J., Sataur, A., Wang, L., Chen, H. & Simon, M. C. The N-terminal transactivation domain confers target gene specificity of hypoxia-inducible factors HIF-1alpha and HIF-2alpha. *Mol. Biol. Cell* **18,** 4528–42 (2007).

71. Greer, S. N., Metcalf, J. L., Wang, Y. & Ohh, M. The updated biology of hypoxia-inducible factor. *EMBO Journal* **31,** 2448–2460 (2012).

72. Wang, V., Davis, D. A., Haque, M., Huang, L. E. & Yarchoan, R. Differential gene up-regulation by hypoxia-inducible factor-1alpha and hypoxia-inducible factor-2alpha in HEK293T cells. *Cancer Res.* **65,** 3299–3306 (2005).

73. Talks, K. L. *et al.* The expression and distribution of the hypoxia-inducible factors HIF-1α and HIF-2α in normal human tissues, cancers, and tumor-sssociated macrophages. *Am. J. Pathol.* **157,** 411–21 (2000).

74. Yuan, G. *et al.* Mutual antagonism between hypoxia-inducible factors 1 and 2 regulates oxygen sensing and cardio-respiratory homeostasis. *Proc. Natl. Acad. Sci.* **110,** E1788–E1796 (2013).

75. Heikkila, M., Pasanen, A., Kivirikko, K. I. & Myllyharju, J. Roles of the human hypoxia-inducible factor (HIF)-3a variants in the hypoxia response. *Cell. Mol. Life Sci.* **68,** 3885–3901 (2011).

76. Makino, Y. *et al.* Inhibitory PAS domain protein is a negative regulator of hypoxia-inducible gene expression. *Nature* **414,** 550–554 (2001).

77. Yu, F., White, S. B., Zhao, Q. & Lee, F. S. HIF-1 binding to VHL is regulated by stimulus-sensitive proline hydroxylation. *Proc. Natl. Acad. Sci.* **98,** 9630–9635 (2001).

78. Courtnay, R. *et al.* Cancer metabolism and the Warburg effect: the role of HIF-1 and PI3K. *Molecular biology reports* (2015). doi:10.1007/s11033-015-3858-x

79. Sanchez, A. *et al.* p38 MAPK: A mediator of hypoxia-induced cerebrovascular inflammation. *J. Alzheimer's Dis.* (2012). doi:10.3233/JAD-2012-120829

80. Koong, a C., Chen, E. Y. & Giaccia, a J. Hypoxia causes the activation of nuclear factor kappa B through the phosphorylation of I kappa B alpha on tyrosine residues. *Cancer Res.* (1994).

81. Loenarz, C. & Schofield, C. J. Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends in Biochemical Sciences* **36,** 7–18 (2011).

82. Stenflo, J. *et al.* Hydroxylation of aspartic acid in domains homologous to the epidermal growth factor precursor is catalyzed by a 2-oxoglutarate-dependent dioxygenase. *Proc. Natl. Acad. Sci. U. S. A.* **86,** 444–7 (1989).

83. UniProt Consortium, T. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* (2018). doi:10.1093/nar/gky092

84. Wong, B. W., Kuchnio, A., Bruning, U. & Carmeliet, P. Emerging novel functions of the oxygen-sensing prolyl hydroxylase domain enzymes. *Trends in Biochemical Sciences* **38,** 3–11 (2013).

85. Nguyen, L. K. *et al.* A dynamic model of the hypoxia-inducible factor 1α (HIF-1α) network. *J. Cell Sci.* **126,** 1454–63 (2013).

86. Wielockx, B. & Meneses, A. PHD2: from hypoxia regulation to disease progression. *Hypoxia* 53 (2016). doi:10.2147/HP.S53576

87. Nandal, A. *et al.* Activation of the HIF prolyl hydroxylase by the iron chaperones PCBP1 and PCBP2. *Cell Metab.* **14,** 647–657 (2011).

88. Tarhonskaya, H. *et al.* Kinetic investigations of the role of factor inhibiting hypoxia-inducible factor (FIH) as an oxygen sensor. *J. Biol. Chem.* **290,** 19726–19742 (2015).

89. Lyssiotis, C. A. & Kimmelman, A. C. Metabolic Interactions in the Tumor Microenvironment. *Trends in Cell Biology* **27,** 863–875 (2017).

90. Chilov, D. *et al.* Induction and nuclear translocation of hypoxia-inducible factor-1 (HIF-1): heterodimerization with ARNT is not necessary for nuclear accumulation of HIF-1alpha. *J Cell Sci* **112 (Pt 8,** 1203–1212 (1999).

91. Lonser, R. R. *et al.* Von Hippel-Lindau disease. in *Lancet* **361,** 2059–2067 (2003).

92. Duan, D. R. *et al.* Inhibition of transcription elongation by the VHL tumor suppressor protein. *Science (80-.).* **269,** 1402–1406 (1995).

93. Kibel, A., Iliopoulos, O., DeCaprio, J. A. & Kaelin Jr., W. G. Binding of the von Hippel-Lindau tumor suppressor protein to Elongin B and C. *Science (80-.).* **269,** 1444–1446 (1995).

94. Garrett, K. P. *et al.* Molecular cloning of an essential subunit of RNA polymerase II elongation factor SIII. *Proc Natl Acad Sci U S A* **91,** 5237–41. (1994).

95. Garrett, K. P. *et al.* Positive regulation of general transcription factor SIII by a tailed ubiquitin homolog. *Proc. Natl. Acad. Sci.* **92,** 7172–7176 (1995).

96. Aso, T., Lane, W. S., Conaway, J. W. & Conaway, R. C. Elongin (SIII): A multisubunit regulator of elongation by RNA polymerase II. *Science (80-.).* **269,** 1439–1443 (1995).

97. Lonergan, K. M. *et al.* Regulation of hypoxia-inducible mRNAs by the von Hippel-Lindau tumor suppressor protein requires binding to complexes containing elongins B/C and Cul2. *Mol. Cell. Biol.* **18,** 732–741 (1998).

98. Kamura, T. *et al.* Rbx1, a component of the VHL tumor suppressor complex and SCF ubiquitin ligase. *Science (80-.).* **284,** 657–661 (1999).

99. Scott, D. C. *et al.* Two Distinct Types of E3 Ligases Work in Unison to Regulate Substrate Ubiquitylation. *Cell* **166,** 1198–1214.e24 (2016).

100. Ivan, M. *et al.* HIFalpha Targeted for VHL-Mediated Destruction by Proline Hydroxylation: Implications for O2 Sensing. *Science (80-.).* **292,** 464–468 (2001).

101. Kaelin, W. G. Molecular basis of the VHL hereditary cancer syndrome. *Nature Reviews Cancer* **2,** 673–682 (2002).

102. Schödel, J., Mole, D. R. & Ratcliffe, P. J. Pan-genomic binding of hypoxia-inducible transcription factors. *Biological Chemistry* (2013). doi:10.1515/hsz-2012-0351

103. Shen, C. & Kaelin, W. G. The VHL/HIF axis in clear cell renal carcinoma. *Seminars in Cancer Biology* **23,** 18–25 (2013).

104. Semenza, G. L. Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics. *Oncogene* (2010). doi:10.1038/onc.2009.441

105. Gnarra, J. R. *et al.* Post-transcriptional regulation of vascular endothelial growth factor mRNA by the product of the VHL tumor suppressor gene. *Proc. Natl. Acad. Sci. U. S. A.* (1996). doi:10.1073/pnas.93.20.10589

106. Benjamin, L. E., Hemo, I. & Keshet, E. A plasticity window for blood vessel remodelling is defined by pericyte coverage of the preformed endothelial network and is regulated by PDGF-B and VEGF. *Development* (1998).

107. Zatyka, M. *et al.* Identification of cyclin D1 and other novel targets for the von Hippel-Lindau tumor suppressor gene by expression array analysis and investigation of cyclin D1 genotype as a modifier in von Hippel-Lindau disease. *Cancer Res.* (2002).

108. Quaegebeur, A. *et al.* Deletion or inhibition of the oxygen sensor PHD1 protects against ischemic stroke via reprogramming of neuronal metabolism. *Cell Metab.* (2016). doi:10.1016/j.cmet.2015.12.007

109. Solaini, G., Baracca, A., Lenaz, G. & Sgarbi, G. Hypoxia and mitochondrial oxidative metabolism. *Biochim. Biophys. Acta - Bioenerg.* (2010). doi:10.1016/j.bbabio.2010.02.011

110. Xu, J. *et al.* Na+/H+exchanger 1, Na+/Ca2+exchanger 1 and calmodulin complex regulates interleukin 6-mediated cellular behavior of human hepatocellular carcinoma. *Carcinogenesis* (2015). doi:10.1093/carcin/bgw004

111. Stickle, N. H. *et al.* pVHL modification by NEDD8 is required for fibronectin matrix assembly and suppression of tumor development. *Mol Cell Biol* (2004). doi:10.1128/MCB.24.8.3251

112. Champion, K. J., Guinea, M., Dammai, V. & Hsu, T. Endothelial function of von Hippel-Lindau tumor suppressor gene: Control of fibroblast growth factor receptor signaling. *Cancer Res.* (2008). doi:10.1158/0008-5472.CAN-07-6003

113. Datta, K. *et al.* Role of elongin-binding domain of von hippel lindau gene product on HuR-mediated VPF/VEGF mRNA stability in renal cell carcinoma. *Oncogene* **24,** 7850–7858 (2005).

114. Mikhaylova, O. *et al.* The von Hippel-Lindau tumor suppressor protein and Egl-9-Type proline hydroxylases regulate the large subunit of RNA polymerase II in response to oxidative stress. *Mol. Cell. Biol.* **28,** 2701–17 (2008).

115. Wang, Y. *et al.* Regulation of endocytosis via the oxygen-sensing pathway. *Nat. Med.* (2009). doi:10.1038/nm.1922

116. Yang, H. *et al.* pVHL Acts as an Adaptor to Promote the Inhibitory Phosphorylation of the NF-κB Agonist Card9 by CK2. *Mol. Cell* **28,** 15–27 (2007).

117. Guo, J. *et al.* pVHL suppresses kinase activity of Akt in a proline-hydroxylation-dependent manner. *Science (80-.).* **353,** 929–932 (2016).

118. Young, A. P. *et al.* VHL loss actuates a HIF-independent senescence programme mediated by Rb and p400. *Nat. Cell Biol.* (2008). doi:10.1038/ncb1699

119. Lutz, M. S. & Burk, R. D. Primary cilium formation requires von Hippel-Lindau gene function in renal-derived cells. *Cancer Res.* (2006). doi:10.1158/0008-5472.CAN-06-0501

120. Frew, I. J. & Krek, W. PVHL: A multipurpose adaptor protein. *Science Signaling* **1,** (2008).

121. Johnstone, R. W. *et al.* Apoptosis: a link between cancer genetics and chemotherapy. *Cell* (2002). doi:10.1016/S0092-8674(02)00625-6

122. Tomasino, R. M. *et al.* p53 expression in human renal cell carcinoma: an immunohistochemical

study and a literature outline of the cytogenetic characterization. *Pathologica* (1994).

123. Sendoel, A., Kohler, I., Fellmann, C., Lowe, S. W. & Hengartner, M. O. HIF-1 antagonizes p53-mediated apoptosis through a secreted neuronal tyrosinase. *Nature* (2010). doi:10.1038/nature09141

124. An, J. & Rettig, M. B. Mechanism of von Hippel-Lindau protein-mediated suppression of nuclear factor kappa B activity. *Mol. Cell. Biol.* (2005). doi:10.1128/MCB.25.17.7546-7556.2005

125. Kim, W. Y. & Sharpless, N. E. VHL Inactivation: A New Road to Senescence. *Cancer Cell* (2008). doi:10.1016/j.ccr.2008.03.012

126. Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. From molecular to modular cell biology. *Nature* **402,** C47–C52 (1999).

127. Patil, A., Kinoshita, K. & Nakamura, H. Hub promiscuity in protein-protein interaction networks. *Int. J. Mol. Sci.* **11,** 1930–43 (2010).

128. Jeong, H., Mason, S. P., Barabási, A.-L. & Oltvai, Z. N. Lethality and centrality in protein networks. *Nature* **411,** 41–42 (2001).

129. Han, J. D. J. *et al.* Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430,** 88–93 (2004).

130. Higurashi, M., Ishida, T. & Kinoshita, K. Identification of transient hub proteins and the possible structural basis for their multiple interactions. *Protein Sci.* **17,** 72–8 (2008).

131. Wilson, M. A. & Brunger, A. T. The 1.0 Å crystal structure of Ca2+-bound calmodulin: An analysis of disorder and implications for functionally relevant plasticity. *J. Mol. Biol.* **301,** 1237–1256 (2000).

132. Kussie, P. H. *et al.* Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science (80-.).* **274,** 948–953 (1996).

133. Patil, A. & Nakamura, H. Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Lett.* **580,** 2041–2045 (2006).

134. Patil, A., Kinoshita, K. & Nakamura, H. Domain distribution and intrinsic disorder in hubs in the human protein-protein interaction network. *Protein Sci.* **19,** 1461–1468 (2010).

135. van der Lee, R. *et al.* Classification of Intrinsically Disordered Regions and Proteins. *Chem.*

Rev. **114,** 6589–6631 (2014).

136. Wright, P. E. & Dyson, H. J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *Journal of Molecular Biology* (1999). doi:10.1006/jmbi.1999.3110

137. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology* **6,** 197–208 (2005).

138. Eisenberg, D., Marcotte, E. M., Xenarios, I. & Yeates, T. O. Protein function in the post-genomic era. *Nature* **405,** 823–6 (2000).

139. Davey, N. E., Haslam, N. J., Shields, D. C. & Edwards, R. J. SLiMSearch 2.0: Biological context for short linear motifs in proteins. *Nucleic Acids Res.* **39,** (2011).

140. Mészáros, B., Kumar, M., Gibson, T. J., Uyar, B. & Dosztányi, Z. Degrons in cancer. *Science Signaling* **10,** (2017).

141. Fuxreiter, M., Simon, I., Friedrich, P. & Tompa, P. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J. Mol. Biol.* **338,** 1015–1026 (2004).

142. Lim, S. & Kaldis, P. Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development* **140,** 3079–93 (2013).

143. Weinberg, R. A. The retinoblastoma protein and cell cycle control. *Cell* (1995). doi:10.1016/0092-8674(95)90385-2

144. Kastan, M. B. & Bartek, J. Cell-cycle checkpoints and cancer. *Nature* (2004). doi:10.1038/nature03097

145. Control, C. C. & Death, C. Cell Cycle Control, Cell Death and Cancer. *Cell Cycle* 127–136 (2007).

146. Malumbres, M. & Barbacid, M. Cell cycle, CDKs and cancer: A changing paradigm. *Nature Reviews Cancer* (2009). doi:10.1038/nrc2602

147. Malumbres, M. Cyclin-dependent kinases. *Genome Biol.* (2014). doi:10.1186/gb4184

148. Besson, A., Dowdy, S. F. & Roberts, J. M. CDK Inhibitors: Cell Cycle Regulators and Beyond. *Developmental Cell* (2008). doi:10.1016/j.devcel.2008.01.013

149. Gu, Y., Turck, C. W. & Morgan, D. O. Inhibition of CDK2 activity in vivo by an associated 20K regulatory subunit. *Nature* (1993). doi:10.1038/366707a0

150. Toyoshima, H. & Hunter, T. p27, a novel inhibitor of G1 cyclin-Cdk protein kinase activity, is

related to p21. *Cell* (1994). doi:10.1016/0092-8674(94)90573-8

151. Matsuoka, S. *et al.* Imprinting of the gene encoding a human cyclin-dependent kinase inhibitor, p57KIP2, on chromosome 11p15. *Proc. Natl. Acad. Sci.* (1996). doi:10.1073/pnas.93.7.3026

152. Iconaru, L. I. *et al.* Discovery of Small Molecules that Inhibit the Disordered Protein, p27Kip1. *Sci. Rep.* (2015). doi:10.1038/srep15686

153. Pateras, I. S., Apostolopoulou, K., Niforou, K., Kotsinas, A. & Gorgoulis, V. G. p57KIP2: "Kip"ing the cell under control. *Mol. Cancer Res.* **7,** 1902–19 (2009).

154. Child, E. S. & Mann, D. J. The intricacies of p21 phosphorylation: Protein/protein interactions, subcellular localization and stability. *Cell Cycle* **5,** 1313–1319 (2006).

155. Mercer, W. E. *et al.* WAF1/CIP1 Is Induced in p53-mediated G1 Arrest and Apoptosis. *Cancer Res.* (1994). doi:10.1158/0008-5472.can-09-4024

156. Abukhdeir, A. M. & Park, B. H. p21 and p27. *Expert Rev. Mol. Med.* (2008). doi:10.1017/S1462399408000744

157. Chu, I. *et al.* p27 Phosphorylation by Src Regulates Inhibition of Cyclin E-Cdk2. *Cell* **128,** 281–294 (2007).

158. Ben-Porath, I. & Cedar, H. Imprinting: Focusing on the center. *Current Opinion in Genetics and Development* **10,** 550–554 (2000).

159. Luo, Y., Hurwitz, J. & Massagué, J. Cell-cycle inhibition by independent CDK and PCNA binding domains in p21cip1. *Nature* (1995). doi:10.1038/375159a0

160. Huang, Y., Yoon, M. K., Otieno, S., Lelli, M. & Kriwacki, R. W. The activity and stability of the intrinsically disordered Cip/Kip protein family are regulated by non-receptor tyrosine kinases. *J. Mol. Biol.* (2015). doi:10.1016/j.jmb.2014.11.011

161. Hnit, S. S. T. *et al.* $P27^{Kip1}$ signaling: Transcriptional and post-translational regulation. *Int. J. Biochem. Cell Biol.* (2015). doi:10.1016/j.biocel.2015.08.005

162. Georgakilas, A. G., Martin, O. A. & Bonner, W. M. p21: A Two-Faced Genome Guardian. *Trends in Molecular Medicine* **23,** 310–319 (2017).

163. Bachs, O. *et al.* Role of $p27^{Kip1}$as a transcriptional regulator. *Oncotarget* **9,** (2018).

164. Abbas, T. & Dutta, A. P21 in cancer: Intricate networks and multiple activities. *Nature Reviews Cancer* **9,** 400–414 (2009).

165. Besson, A. *et al.* A pathway in quiescent cells that controls p27Kip1 stability, subcellular

localization, and tumor suppression. *Genes Dev.* **20,** 47–64 (2006).

166. Hayflick, L. & Moorhead, P. S. The serial cultivation of human diploid cell strains. *Exp. Cell Res.* **25,** 585–621 (1961).

167. McEachern, M. J., Krauskopf, A. & Blackburn, E. H. Telomeres and Their Control. *Annu. Rev. Genet.* **34,** 331–358 (2000).

168. Bertschmann, J., Thalappilly, S. & Riabowol, K. The ING1a model of rapid cell senescence. *Mechanisms of Ageing and Development* (2018). doi:10.1016/j.mad.2018.06.004

169. Narita, M. *et al.* Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* **113,** 703–716 (2003).

170. Quereda, V., Porlan, E., Canãmero, M., Dubus, P. & Malumbres, M. An essential role for Ink4 and Cip/Kip cell-cycle inhibitors in preventing replicative stress. *Cell Death Differ.* **23,** 430–441 (2016).

171. Bretones, G., Delgado, M. D. & León, J. Myc and cell cycle control. *Biochim. Biophys. Acta* (2015). doi:10.1016/j.bbagrm.2014.03.013

172. Ananiev, J., Tchernev, G., Patterson, J. W., Gulubova, M. & Ganchev, G. p53 - 'The Guardian of Genome'. *Acta Medica Bulgarica* (2011). doi:10.1038/358015a0

173. Moll, U. M. & Petrenko, O. The MDM2-p53 interaction. *Mol. Cancer Res.* **1,** 1001–1008 (2003).

174. Haupt, Y., Maya, R., Kazaz, A. & Oren, M. Mdm2 promotes the rapid degradation of p53. *Nature* **387,** 296–299 (1997).

175. Gallagher, S. J., Kefford, R. F. & Rizos, H. The ARF tumour suppressor. *International Journal of Biochemistry and Cell Biology* **38,** 1637–1641 (2006).

176. Ravi, R. *et al.* Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1α. *Genes Dev.* **14,** 34–44 (2000).

177. Yamakuchi, M. *et al.* P53-induced microRNA-107 inhibits HIF-1 and tumor angiogenesis. *Proc. Natl. Acad. Sci.* (2010). doi:10.1073/pnas.0911082107

178. Chen, D. Direct Interactions between HIF-1alpha and Mdm2 Modulate p53 Function. *Journal of Biological Chemistry* **278,** 13595–13598 (2003).

179. Louvet, O., Doignon, F. & Crouzet, M. Stable DNA-binding yeast vector allowing high-bait expression for use in the two-hybrid system. *Biotechniques* **23,** 816–820 (1997).

180. Brachmann, R. K. & Boeke, J. D. Tag games in yeast: The two-hybrid system and beyond. *Current Opinion in Biotechnology* **8,** 561–568 (1997).

181. Tirode, F. *et al.* A conditionally expressed third partner stabilizes or prevents the formation of a transcriptional activator in a three-hybrid system. *J. Biol. Chem.* **272,** 22995–9 (1997).

182. Osborne, M. A., Dalton, S. & Kochan, J. P. The Yeast Tribrid System—Genetic Detection of trans-phosphorylated ITAM-SH2-Interactions. *Nat. Biotechnol.* **13,** 1474–1478 (1995).

183. Hanahan, D., Jessee, J. & Bloom, F. R. Plasmid transformation of Escherichia coli and other bacteria. *Methods Enzymol.* **204,** 63–113 (1991).

184. Botstein, D. & Fink, G. R. Yeast: An experimental organism for modern biology. *Science (80-.).* (1988). doi:10.1126/science.3287619

185. Botstein, D., Chervitz, S. A. & Cherry, J. M. Yeast as a model organism. *Science* (1997). doi:10.1126/science.277.5330.1259

186. Oliver, S. G. From DNA sequence to biological function. *Nature* (1996). doi:10.1038/379597a0

187. Brückner, A., Polge, C., Lentze, N., Auerbach, D. & Schlattner, U. Yeast Two-Hybrid , a Powerful Tool for Systems Biology. 2763–2788 (2009). doi:10.3390/ijms10062763

188. Van Criekinge, W. & Beyaert, R. Yeast Two-Hybrid: State of the Art. *Biol. Proced. Online* **2,** 1–38 (1999).

189. Marmorstein, R., Carey, M., Ptashne, M. & Harrison, S. C. DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* **356,** 408–414 (1992).

190. Elliott, D. a & Brand, A. H. The GAL4 system : a versatile system for the expression of genes. *Methods Mol. Biol.* (2008). doi:10.1007/978-1-59745-583-1_5

191. Chen, D. C., Yang, B. C. & Kuo, T. T. One-step transformation of yeast in stationary phase. *Curr. Genet.* **21,** 83–4 (1992).

192. Chen, P. *et al.* Visualized investigation of yeast transformation induced with Li+ and polyethylene glycol. *Talanta* **77,** 262–268 (2008).

193. Wright, A. P. H., Bruns, M. & Hartley, B. S. Extraction and rapid inactivation of proteins fromSaccharomyces cerevisiae by trichloroacetic acid precipitation. *Yeast* **5,** 51–53 (1989).

194. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248–54 (1976).

195. Giaever, G. *et al.* Functional profiling of the Saccharomyces cerevisiae genome. *Nature* (2002). doi:10.1038/nature00935

196. Hakhverdyan, Z. *et al.* Rapid, optimized interactomic screening. *Nat. Methods* (2015). doi:10.1038/nmeth.3395

197. Nivon, M. *et al.* NF-κB regulates protein quality control after heat stress through modulation of the BAG3-HspB8 complex. *J. Cell Sci.* (2012). doi:10.1242/jcs.091041

198. LAEMMLI, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227,** 680–685 (1970).

199. J. Sambrook, E.F. Fritsch, T. M. *Molecular cloning : a laboratory manual / J. Sambrook, E.F. Fritsch, T. Maniatis. - Version details - Trove*. (Cold Spring Harbor Laboratory Press, 1989).

200. UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res.* **43,** D204–D212 (2015).

201. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25,** 1189–1191 (2009).

202. Notredame, C., Higgins, D. G. & Heringa, J. T-coffee: a novel method for fast and accurate multiple sequence alignment 1 1Edited by J. Thornton. *J. Mol. Biol.* **302,** 205–217 (2000).

203. Potenza, E., Domenico, T. Di, Walsh, I. & Tosatto, S. C. E. MobiDB 2.0: an improved database of intrinsically disordered and mobile proteins. *Nucleic Acids Res.* **43,** D315–D320 (2015).

204. Piovesan, D. *et al.* MobiDB 3.0: more annotations for intrinsic disorder, conformational diversity and interactions in proteins. *Nucleic Acids Res.* **46,** D471–D476 (2018).

205. Piovesan, D. *et al.* DisProt 7.0: a major update of the database of disordered proteins. *Nucleic Acids Res.* **45,** D219–D227 (2017).

206. Punta, M. *et al.* The Pfam protein families database. *Nucleic Acids Res.* **40,** D290–D301 (2012).

207. Finn, R. D. *et al.* InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res.* **45,** D190–D199 (2017).

208. Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **13,** 2498–2504 (2003).

209. Franceschini, A. *et al.* STRING v9.1: protein-protein interaction networks, with increased

coverage and integration. *Nucleic Acids Res.* **41,** D808–D815 (2012).

210. Hon, W.-C. *et al.* Structural basis for the recognition of hydroxyproline in HIF-1α by pVHL. *Nature* **417,** 975–978 (2002).

211. Russo, A. A., Jeffrey, P. D., Patten, A. K., Massagué, J. & Pavletich, N. P. Crystal structure of the p27Kip1 cyclin-dependent-kinase inibitor bound to the cyclin A–Cdk2 complex. *Nature* **382,** 325–331 (1996).

212. Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **4,** 435–447 (2008).

213. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).

214. Piovesan, D., Minervini, G. & Tosatto, S. C. E. The RING 2.0 web server for high quality residue interaction networks. *Nucleic Acids Res.* **44,** W367–W374 (2016).

215. Forbes, S. A. *et al.* COSMIC: Exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **43,** D805–D811 (2015).

216. Papaemmanuil, E. *et al.* Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122,** 3616–3627 (2013).

217. Walsh, I. *et al.* Bluues server: electrostatic properties of wild-type and mutated protein structures. *Bioinformatics* **28,** 2189–2190 (2012).

218. Giollo, M., Martin, A. J., Walsh, I., Ferrari, C. & Tosatto, S. C. NeEMO: a method using residue interaction networks to improve prediction of protein stability upon mutation. *BMC Genomics* **15,** S7 (2014).

219. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci.* **98,** 10037–10041 (2001).

220. Biasini, M. *et al.* SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42,** (2014).

221. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).

222. Benkert, P., Tosatto, S. C. E. & Schomburg, D. QMEAN: A comprehensive scoring function

for model quality assessment. *Proteins* **71,** 261–277 (2008).

223. Tosatto, S. C. E. & Battistutta, R. TAP score: Torsion angle propensity normalization applied to local protein structure evaluation. *BMC Bioinformatics* **8,** (2007).

224. Quevillon, E. *et al.* InterProScan: Protein domains identifier. *Nucleic Acids Res.* (2005). doi:10.1093/nar/gki442

225. Wilson, W. R. & Hay, M. P. Targeting hypoxia in cancer therapy. *Nature Reviews Cancer* **11,** 393–410 (2011).

226. Iliopoulos, O., Kibel, A., Gray, S. & Kaelin, W. G. Tumour suppression by the human von hippel-lindau gene product. *Nat. Med.* **1,** 822–826 (1995).

227. Welford, S. M. & Giaccia, A. J. Hypoxia and Senescence: The Impact of Oxygenation on Tumor Suppression. *Mol. Cancer Res.* **9,** 538–544 (2011).

228. Sullivan, R., Pare, G. C., Frederiksen, L. J., Semenza, G. L. & Graham, C. H. Hypoxia-induced resistance to anticancer drugs is associated with decreased senescence and requires hypoxia-inducible factor-1 activity. *Mol. Cancer Ther.* **7,** 1961–1973 (2008).

229. Ohtani, N., Mann, D. J. & Hara, E. Cellular senescence: Its role in tumor suppression and aging. *Cancer Sci.* **100,** 792–797 (2009).

230. Koshiji, M. *et al.* HIF-1alpha induces cell cycle arrest by functionally counteracting Myc. *EMBO J.* **23,** 1949–56 (2004).

231. Goda, N. *et al.* Hypoxia-inducible factor 1alpha is essential for cell cycle arrest during hypoxia. *Mol. Cell. Biol.* **23,** 359–69 (2003).

232. Nakayama, K. I. & Nakayama, K. Cip/Kip cyclin-dependent kinase inhibitors: Brakes of the cell cycle engine during development. *BioEssays* **20,** 1020–1029 (1998).

233. Brugarolas, J. *et al.* Radiation-induced cell cycle arrest compromised by p21 deficiency. *Nature* **377,** 552–557 (1995).

234. Macleod, K. F. *et al.* p53-Dependent and independent expression of p21 during cell growth, differentiation, and DNA damage. *Genes Dev.* **9,** 935–944 (1995).

235. Lai, Y., Song, M., Hakala, K., Weintraub, S. T. & Shiio, Y. Proteomic dissection of the von hippel-Lindau (VHL) interactome. *J. Proteome Res.* **10,** 5175–5182 (2011).

236. Lomazzi, M., Moroni, M. C., Jensen, M. R., Frittoli, E. & Helin, K. Suppression of the p53- or pRB-mediated G1 checkpoint is required for E2F-induced S-phase entry. *Nat. Genet.* **31,**

190–194 (2002).

237. Duro, D., Bernard, O., Della Valle, V., Berger, R. & Larsen, C. J. A new type of p16INK4/MTS1 gene transcript expressed in B-cell malignancies. *Oncogene* **11,** 21–29 (1995).

238. Ratcliffe, P. J. New insights into an enigmatic tumour suppressor. *Nat. Cell Biol.* **5,** 7–8 (2003).

239. Consortium, T. G. O. Gene Ontology Annotations and Resources. *Nucleic Acids Res.* **41,** D530–D535 (2013).

240. Na, X. *et al.* Identification of the RNA polymerase II subunit hsRPB7 as a novel target of the von Hippel-Lindau protein. *EMBO J.* **22,** 4249–4259 (2003).

241. Young, A. P. *et al.* VHL loss actuates a HIF-independent senescence programme mediated by Rb and p400. *Nat. Cell Biol.* **10,** 361–369 (2008).

242. Tenga, M. J. & Lazar, I. M. Proteomic study reveals a functional network of cancer markers in the G1-Stage of the breast cancer cell cycle. *BMC Cancer* **14,** (2014).

243. Liu, W., Xin, H., Eckert, D. T., Brown, J. A. & Gnarra, J. R. Hypoxia and cell cycle regulation of the von Hippel-Lindau tumor suppressor. *Oncogene* **30,** 21–31 (2011).

244. Galea, C. A., Wang, Y., Sivakolundu, S. G. & Kriwacki, R. W. Regulation of Cell Division by Intrinsically Unstructured Proteins: Intrinsic Flexibility, Modularity, and Signaling Conduits †. *Biochemistry* **47,** 7598–7609 (2008).

245. Yoon, M.-K., Mitrea, D. M., Ou, L. & Kriwacki, R. W. Cell cycle regulation by the intrinsically disordered proteins p21 and p27. *Biochem. Soc. Trans.* **40,** 981–988 (2012).

246. Piovesan, D., Walsh, I., Minervini, G. & Tosatto, S. C. E. FELLS: Fast estimator of latent local structure. *Bioinformatics* **33,** 1889–1891 (2017).

247. Högel, H., Miikkulainen, P., Bino, L. & Jaakkola, P. M. Hypoxia inducible prolyl hydroxylase PHD3 maintains carcinoma cell growth by decreasing the stability of p27. *Mol. Cancer* **187,** 1–14 (2015).

248. Gordan, J. D., Bertout, J. A., Hu, C. J., Diehl, J. A. & Simon, M. C. HIF-2α Promotes Hypoxic Cell Proliferation by Enhancing c-Myc Transcriptional Activity. *Cancer Cell* **11,** 335–347 (2007).

249. Hasanov, E. *et al.* Ubiquitination and regulation of AURKA identifies a hypoxia-independent E3 ligase activity of VHL. *Oncogene* **36,** 3450–3463 (2017).

250. Rodier, G. *et al.* p27 cytoplasmic localization is regulated by phosphorylation on Ser10 and is not a prerequisite for its proteolysis. *EMBO J.* **20,** 6672–6682 (2001).

251. Dash, B. C. & El-Deiry, W. S. Phosphorylation of p21 in G2/M Promotes Cyclin B-Cdc2 Kinase Activity. *Mol. Cell. Biol.* **25,** 3364–3387 (2005).

252. Atkins, D. J. *et al.* Concomitant deregulation of HIF1α and cell cycle proteins in VHL-mutated renal cell carcinomas. *Virchows Arch.* **447,** 634–642 (2005).

253. Mack, F. A., Patel, J. H., Biju, M. P., Haase, V. H. & Simon, M. C. Decreased growth of Vhl-/- fibrosarcomas is associated with elevated levels of cyclin kinase inhibitors p21 and p27. *Mol. Cell. Biol.* **25,** 4565–78 (2005).

254. Starostina, N. G. & Kipreos, E. T. Multiple degradation pathways regulate versatile CIP/KIP CDK inhibitors. *Trends in Cell Biology* **22,** 33–41 (2012).

255. Roe, J. S., Kim, H. R., Hwang, I. Y., Cho, E. J. & Youn, H. D. Von Hippel-Lindau protein promotes Skp2 destabilization on DNA damage. *Oncogene* **30,** 3127–3138 (2011).

256. Sgambato, A. *et al.* Deregulated expression of p27(Kip1) in human breast cancers. *Clin. Cancer Res.* **3,** 1879–1887 (1997).

257. Gossage, L., Eisen, T. & Maher, E. R. VHL, the story of a tumour suppressor gene. *Nat. Rev. Cancer* **15,** 55–64 (2015).

258. Blankenship, C., Naglich, J. G., Whaley, J. M., Seizinger, B. & Kley, N. Alternate choice of initiation codon produces a biologically active product of the von Hippel Lindau gene with tumor suppressor activity. *Oncogene* **18,** 1529–1535 (1999).

259. Rose, P. W. *et al.* The RCSB protein data bank: Integrative view of protein, gene and 3D structural information. *Nucleic Acids Res.* **45,** (2017).

260. Zhang, Q., Yang, H., Wei, W. & Wei, J.-J. The Roles of VHL-Dependent Ubiquitination in Signaling and Cancer. (2018). doi:10.3389/fonc.2012.00035

261. Maxwell, P. H. *et al.* The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* **399,** 271–275 (1999).

262. Bex, C., Knauth, K., Dambacher, S. & Buchberger, A. A yeast two-hybrid system reconstituting substrate recognition of the von Hippel-Lindau tumor suppressor protein. **35,** (2007).

263. Minervini, G., Lopreiato, R., Bortolotto, R. & Falconieri, A. Novel interactions of the von

Hippel-Lindau (pVHL) tumor suppressor with the CDKN1 family of cell cycle inhibitors. *Nat. Publ. Gr.* 1–12 (2017). doi:10.1038/srep46562

264. Lai, Y., Song, M., Hakala, K., Weintraub, S. T. & Shiio, Y. Proteomic Dissection of the von Hippel–Lindau (VHL) Interactome. *J. Proteome Res.* **10,** 5175–5182 (2011).

265. Kitano, H. Biological robustness. *Nature Reviews Genetics* (2004). doi:10.1038/nrg1471

266. Ma, H. W. & Zeng, A. P. The connectivity structure, giant strong component and centrality of metabolic networks. *Bioinformatics* (2003). doi:10.1093/bioinformatics/btg177

267. Csete, M. & Doyle, J. Bow ties, metabolism and disease. *Trends in Biotechnology* (2004). doi:10.1016/j.tibtech.2004.07.007

268. Murray-Zmijewski, F., Lane, D. P. & Bourdon, J. C. p53/p63/p73 isoforms: An orchestra of isoforms to harmonise cell differentiation and response to stress. *Cell Death and Differentiation* (2006). doi:10.1038/sj.cdd.4401914

269. Climente-González, H., Porta-Pardo, E., Godzik, A. & Eyras, E. The Functional Impact of Alternative Splicing in Cancer. *Cell Rep.* (2017). doi:10.1016/j.celrep.2017.08.012

270. Chien, C. T., Bartel, P. L., Sternglanz, R. & Fields, S. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci.* (1991). doi:10.1073/pnas.88.21.9578

271. Zavialov, A. V *et al.* Human adenosine deaminase 2 induces differentiation of monocytes into macrophages and stimulates proliferation of T helper cells and macrophages. *J. Leukoc. Biol.* **88,** 279–290 (2010).

272. Yang, L. *et al.* GPR56 regulates VEGF production and angiogenesis during melanoma progression. *Cancer Res.* **71,** 5558–5568 (2011).

273. Robinson, C. M. & Ohh, M. The multifaceted von Hippel-Lindau tumour suppressor protein. *FEBS Lett.* **588,** 2704–2711 (2014).

274. Peifer, M., Berg, S. & Reynolds, A. B. A repeating amino acid motif shared by proteins with diverse cellular roles. *Cell* (1994). doi:10.1016/0092-8674(94)90353-0

275. Herold, a, Truant, R., Wiegand, H. & Cullen, B. R. Determination of the functional domain organization of the importin alpha nuclear import factor. *J. Cell Biol.* (1998). doi:10.1083/jcb.143.2.309

276. Huber, a H., Nelson, W. J. & Weis, W. I. Three-dimensional structure of the armadillo repeat

region of beta-catenin. *Cell* (1997). doi:10.1016/S0092-8674(00)80352-9

277. Matthews, J. M., Lester, K., Joseph, S. & Curtis, D. J. LIM-domain-only proteins in cancer. *Nature Reviews Cancer* (2013). doi:10.1038/nrc3418

278. Jain, B. P. & Pandey, S. WD40 Repeat Proteins: Signalling Scaffold with Diverse Functions. *Protein J.* 1–16 (2018). doi:10.1007/s10930-018-9785-7

279. Vilas, C. K., Emery, L. E., Denchi, E. L. & Miller, K. M. Caught with One's Zinc Fingers in the Genome Integrity Cookie Jar. *Trends in Genetics* (2018). doi:10.1016/j.tig.2017.12.011

280. Kuriyama, S. *et al.* LPP inhibits collective cell migration during lung cancer dissemination. *Oncogene* (2016). doi:10.1038/onc.2015.155

281. Brahimi-Horn, M. C., Chiche, J. & Pouysségur, J. Hypoxia and cancer. *Journal of Molecular Medicine* (2007). doi:10.1007/s00109-007-0281-3

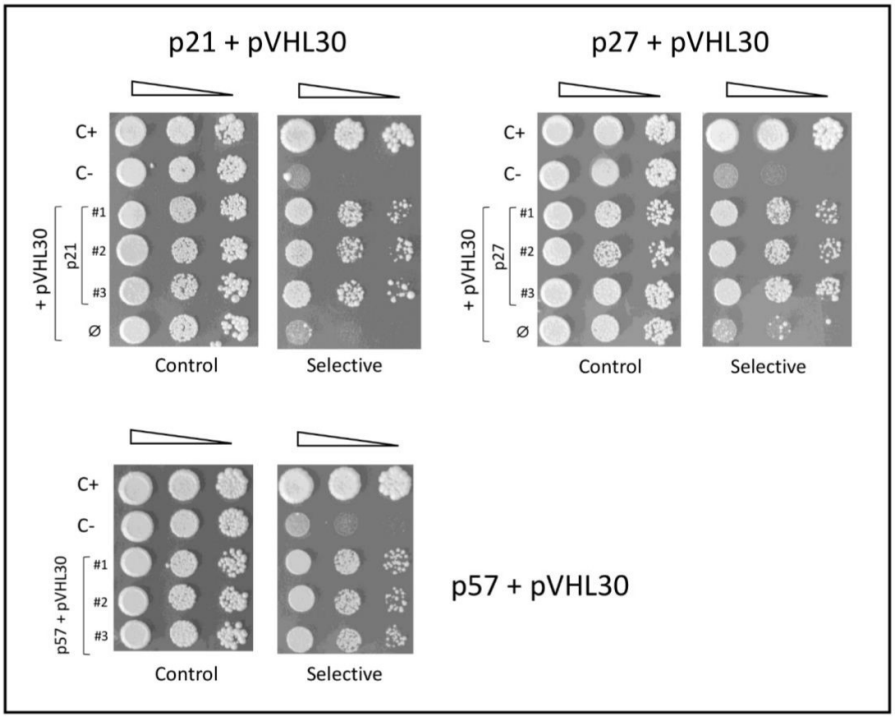
282. Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* (2008). doi:10.1038/ng.259

283. Oltean, S. & Bates, D. O. Hallmarks of alternative splicing in cancer. *Oncogene* (2014). doi:10.1038/onc.2013.533

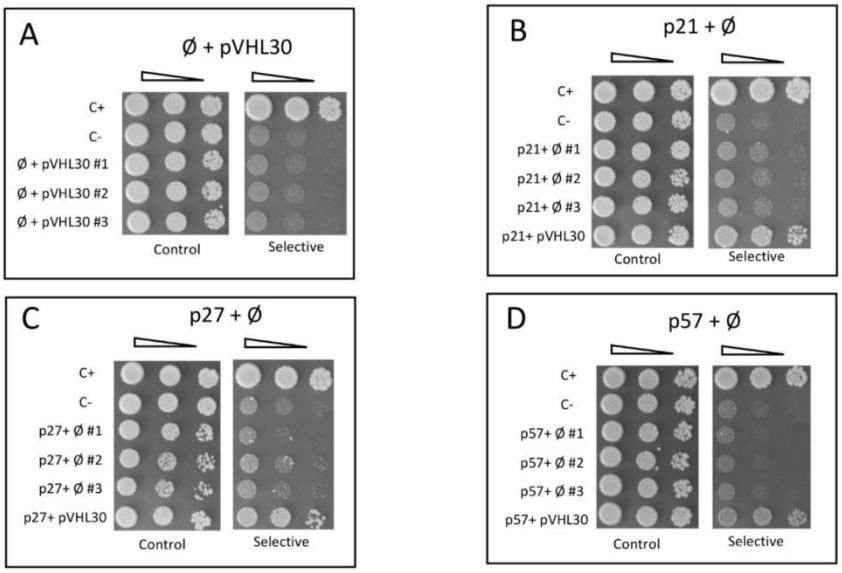
284. Vieler, M. & Sanyal, S. p53 Isoforms and Their Implications in Cancer. *Cancers (Basel).* **10,** 288 (2018).

285. Hafsi, H., Santos-Silva, D., Courtois-Cox, S. & Hainaut, P. Effects of Δ40p53, an isoform of p53 lacking the N-terminus, on transactivation capacity of the tumor suppressor protein p53. *BMC Cancer* (2013). doi:10.1186/1471-2407-13-134

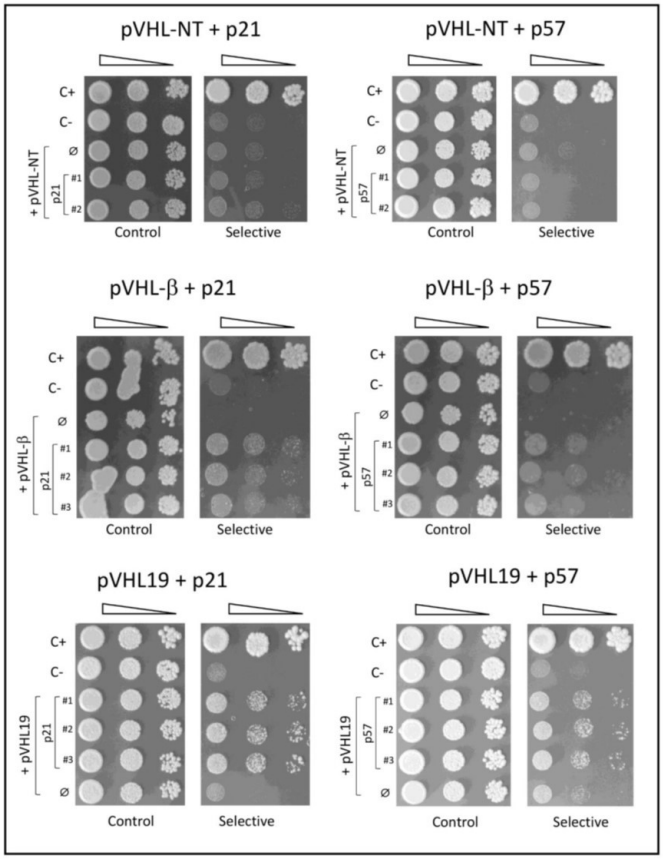
Supplementary figures



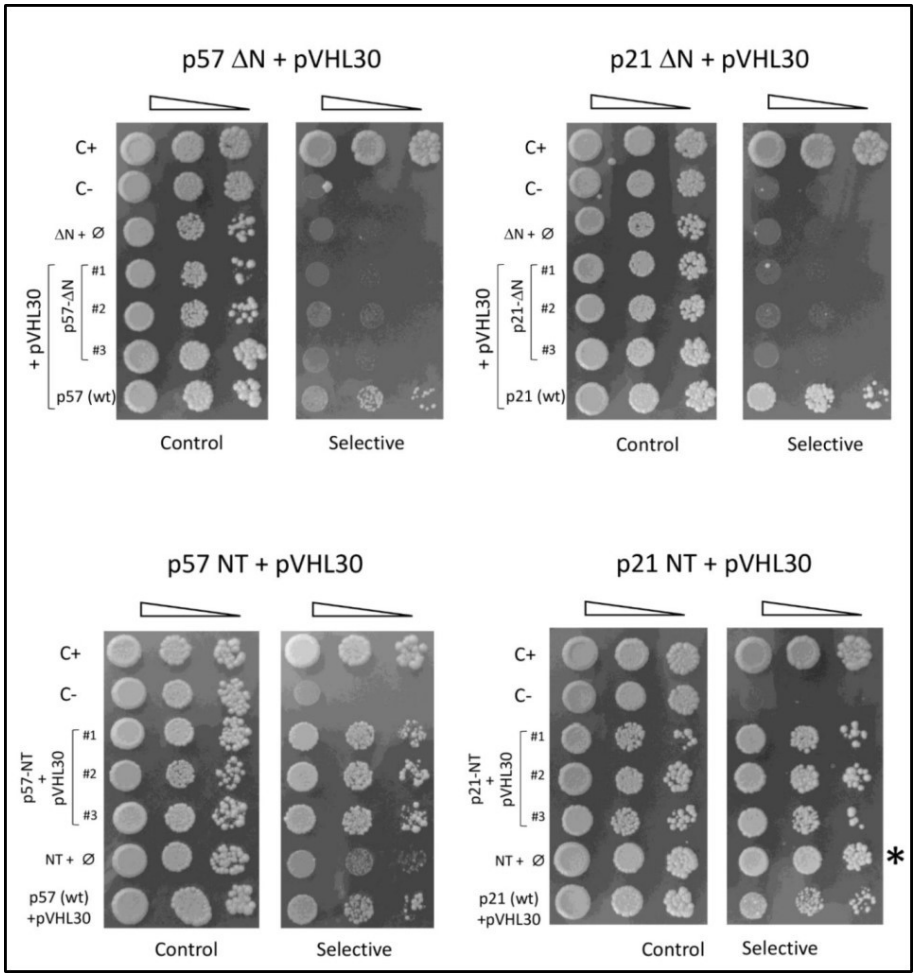
Suppl. Figure 1. pVHL is able to bind the CDKN1 proteins in Y2H assay. The images are representative of 3 independent experiments, where 3 different clones co-expressing either p21, p27, or p57 with pVHL30 have been analysed by standard drop test. Cell growth on selective medium indicates positive interaction. Auto-activation by a Gal4BD-pVHL30 fusion protein has been excluded by the lethality of yeast cells co-transformed with the Gal4AD empty vector (Ø). C+ and C- are positive and negative internal controls of the assay.



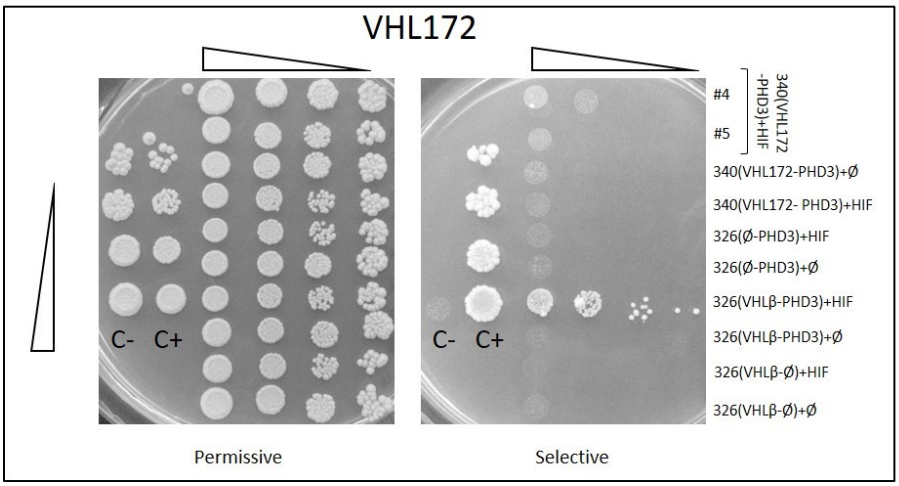
Suppl. Figure 2. (Previous page) Auto-activation by the pVHL30 and CDKN1 proteins in Y2H assay. The pVHL30 (A) and the CDKN1 (B: p21; C: p27; D: p57) proteins alone do not activate the reporter genes in a yeast two hybrid (Y2H) assay. Serial dilutions of yeast cells expressing either pVHL30 or CDKN1 (i.e. co-transformed with the empty vector, Ø) were spotted on both permissive (left) and selective (right) media, and incubated for several days at 30°C. Absence of cell growth on selective medium indicates lack of HIS3 reporter gene activation. C+ and C-are positive and negative controls of the assay.



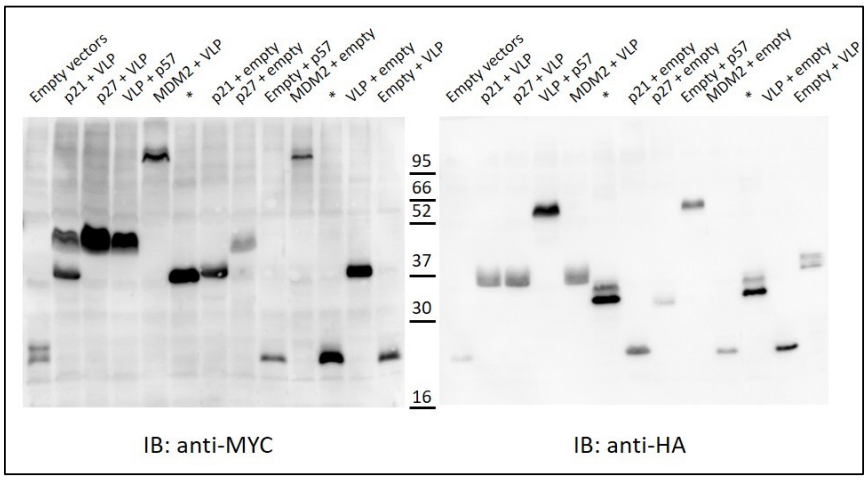
Suppl. Figure 3. pVHL dissection to map binding with p21 and p57 proteins. The pVHL N-terminus has no role in binding p21 and p57 (top). Yeast cells co-expressing pVHL-NT (residues 1-53) with either p21 and p57 (2 clones shown), or Gal4AD alone as control (empty vector, Ø), were tested for their binding in a Y2H assay. No cell growth on selective medium can be observed, indicating absence of interaction. (In the middle) The pVHLβ region maintains binding with both p21 and p57. Yeast cells expressing pVHLβ (54-157 residues) together with either p21 and p57 are able to grow on selective medium. Whereas cells expressing pVHLβ alone (empty vector, Ø), cannot survive. Three independent clones for each CDKN1 are shown. (Bottom) The pVHL19 isoform interacts with p21 and p57. Yeast cells co-expressing pVHL19 (54-213 residues) with either p21 and p57 (3 independent clones each CDKN1) were spotted on permissive and selective medium, where yeast growth is indicative of their binding. Absence of growth for pVHL19 alone (empty vector, Ø) is also shown, excluding aspecific reporter activation.



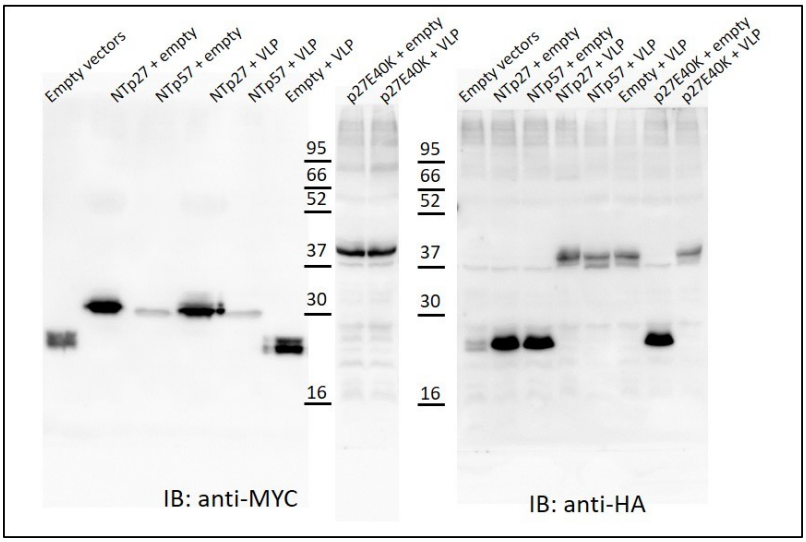
Suppl. Figure 4. p57 and p21 dissection to map binding with pVHL. (Top) The N-terminal tail of the p57 and p21 proteins is essential for binding. Removal of the p57 and p21 N-terminus (ΔN, 1-60 and 1-49 residues, respectively) abolishes interaction with pVHL30. No growth in selective medium was observed for 3 clones of yeast cells co-expressing pVHL30 and ΔN mutant proteins, unlike the full-length sequences (WT). Plates were incubated at 30°C for more time (8 days) to confirm the absence of yeast growth. (Bottom) The N-terminal region the p57 protein is able to bind pVHL. Yeast cells expressing pVHL30 together with the p57 N-terminus (NT, 1-60 residues) were tested for binding by Y2H assay. On selective medium, yeast cells expressing p57-NT grow better than cells carrying the full-length protein (WT). Plates were incubated at 30 °C for 5 days. Data on p21-NT (1-49 residues) cannot be considered, due to its ability to activate the HIS3 reporter gene in presence of the empty vector (Ø), as indicated by the asterisk (*). In all panels, C+ and C-are positive and negative controls of the assay. The images are representative of at least 3 independent experiments, where 2-4 different clones were analyzed.



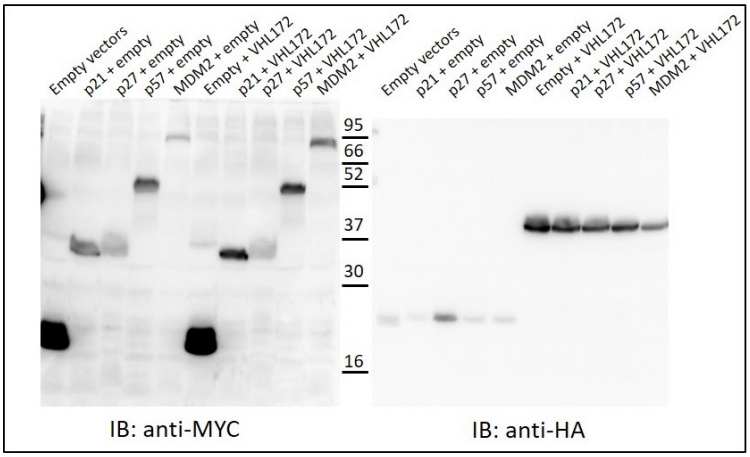
Suppl. Figure 5. HIF-1α does not interact to VHL172. A second independent experiment has been performed, where others auto-activation have been excluded by the lethality of the yeast cells co-transformed. Two independent clones sustain the absence of interaction between VHL172 and HIF-1α compared to yeast cells expressing VHLβ and HIF-1α. Again, C+ and C-are respectively positive and negative controls of the assay. Images are representative of at least three independent experiments, where 2-3 different clones have been analyzed (n≥9).



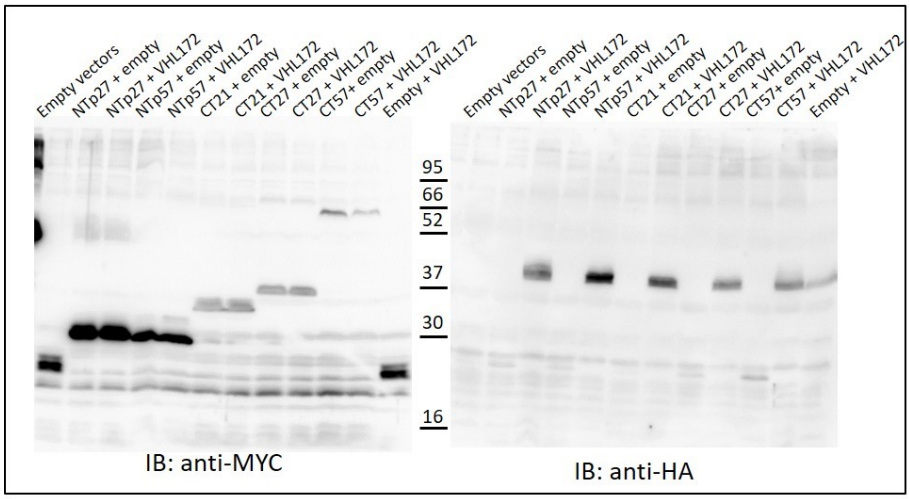
Suppl. Figure 6. Western blot analysis of TCA protein extracts. Yeast cells co-expressing VLP (HA-tagged), CDKN1 (Myc-tagged) and MDM2 (Myc-tagged), as Gal4-fusion proteins are here shown. Membranes are incubated with both anti-Myc (left), and anti-HA (right) antibodies, revealing the presence of fusion proteins at the expected molecular weight, as indicated. Lane 1: Ø+Ø; lane 2: p21+VLP; lane 3: p27+VLP, lane 4: VLP+p57, lane 5: MDM2+VLP, lane 7: p21+Ø, lane 8: p27+Ø, lane 9: Ø+p57, lane 10: MDM2+Ø, lane 12: VLP+Ø and lane 13: Ø+VLP (lane 6 and 11 contain proteins not discuss in this section).



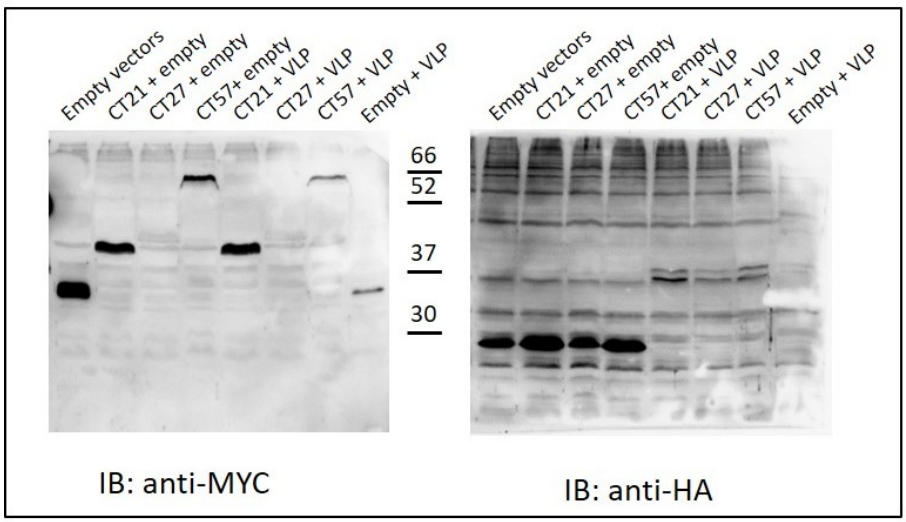
Suppl. Figure 7. Total yeast protein lysates detection by western blot. Total protein extracts of yeast cells co-expressing VLP (HA-tagged) and N-terminus region of p27 and p57 (Myc-tagged) and mutated p27 (p27 E40>K, Myc-tagged). Proteins are respectively fused to Gal4-AD and Gal4DBD regions. Membranes are incubated with either anti-Myc (left), and anti-HA (right) antibodies, revealing the presence of fusion proteins. As indicated by Sharpmass VII marker (Euroclone), proteins have expected molecular weight. Lane 1: Ø+Ø; lane 2: NTp27+Ø; lane 3: NTp57+Ø, lane 4: NTp27+VLP, lane 5: NTp57+VLP, lane 7: Ø+VLP, lane 8: p27E40K+Ø, lane 9: p27E40K+VLP.



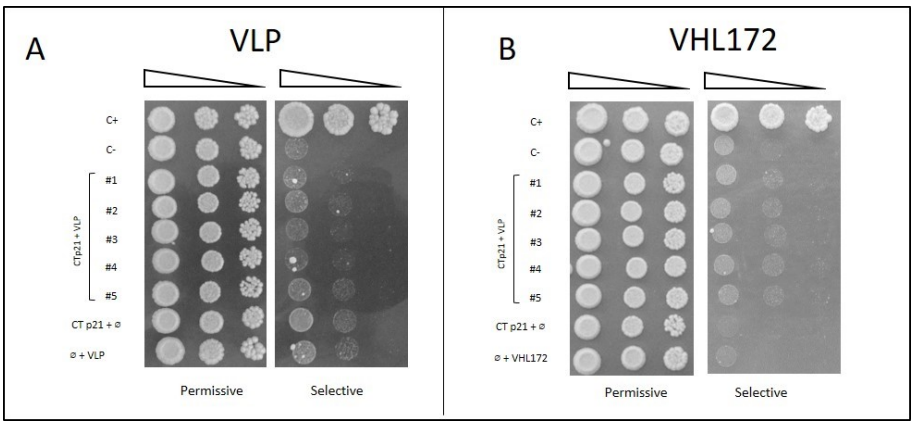
Suppl. Figure 8. Yeast proteins analysis by immunoblot. The correct co-expression of VLP (HA-tagged), CDKN1 (Myc-tagged) or MDM2 (Myc-tagged), as Gal4-fusion proteins, is here shown. The presence of Gal4-fused proteins is revealed incubating respectively with either anti-Myc (left), and anti-HA (right) antibodies. Lane 1: Ø+Ø; lane 2: p21+Ø; lane 3: p27+Ø, lane 4: p57+Ø, lane 5: MDM2+Ø, lane 6: Ø+VHL172, lane 7: p21+VHL172, lane 8: p27+VHL172, lane 9: p57+VHL172, lane 10: MDM2+VHL172. Referring to Sharpamass VII (Euroclone) marker, proteins are at the expected molecular weight.



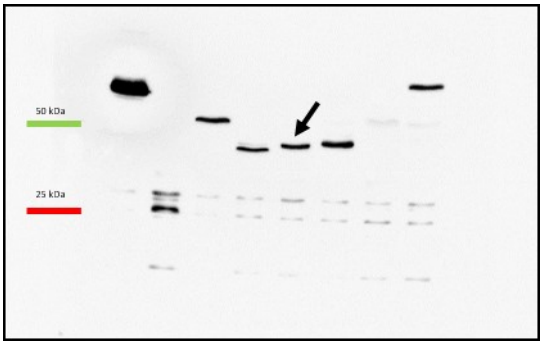
Suppl. Figure 9. Recombinant proteins detection by immunoblot. Total protein extracts of yeast cells co-expressing VHL172 (HA-tagged) and N-terminus/C-terminus region of p27 and p57 (Myc-tagged) and mutated p27 (p27 E40>K, Myc-tagged). Proteins are respectively fused to Gal4-AD and Gal4DBD regions. Membranes are incubated with either anti-Myc (left), and anti-HA (right) antibodies, revealing the presence of fusion proteins at the expected molecular weight. Lane 1: Ø+Ø; lane 2: NTp27+Ø; lane 3: NTp27+VHL172, lane 4: NTp57+Ø, lane 5: NTp57+VHL172 lane 6: CTp21+Ø, lane 7: CTp21+VHL172, lane 8: CTp27+Ø, lane 9: CTp27+VHL172, lane 10: CTp57+Ø, lane 11: CTp57+VHL172, lane 12: Ø+VHL172.



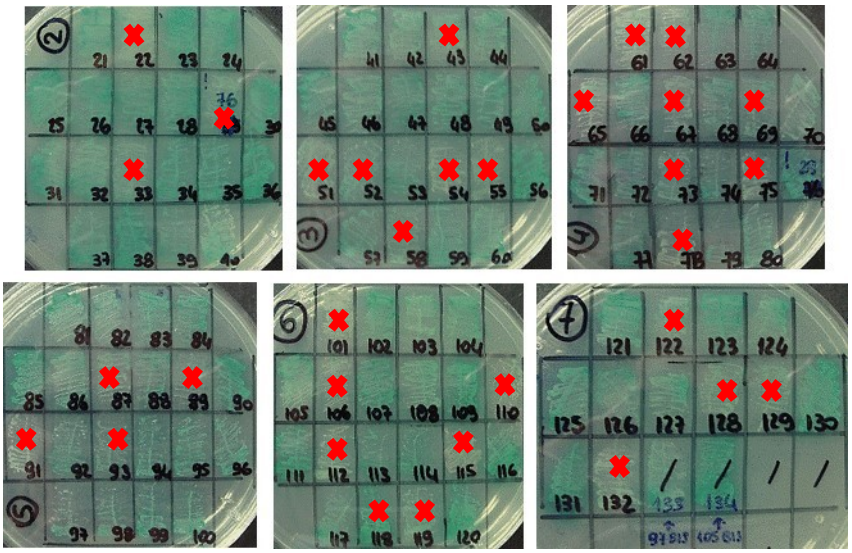
Suppl. Figure 10. Yeast protein lysates detection by immunoblot. Total protein extracts of yeast cells co-expressing VLP (HA-tagged) and C-terminus region of p27 and p57 (Myc-tagged). Proteins are respectively fused to Gal4-AD and Gal4DBD regions. Membranes are incubated with either anti-Myc (left), and anti-HA (right) antibodies, revealing the presence of fusion proteins at the expected molecular weight. Lane 1: Ø+Ø; lane 2: CTp21+Ø; lane 3: CTp27+Ø, lane 4: CTp57+ Ø, lane 5: CTp21+VLP, lane 7: CTp27+VLP, lane 8: CTp57+VLP, lane 9: Ø+VLP.



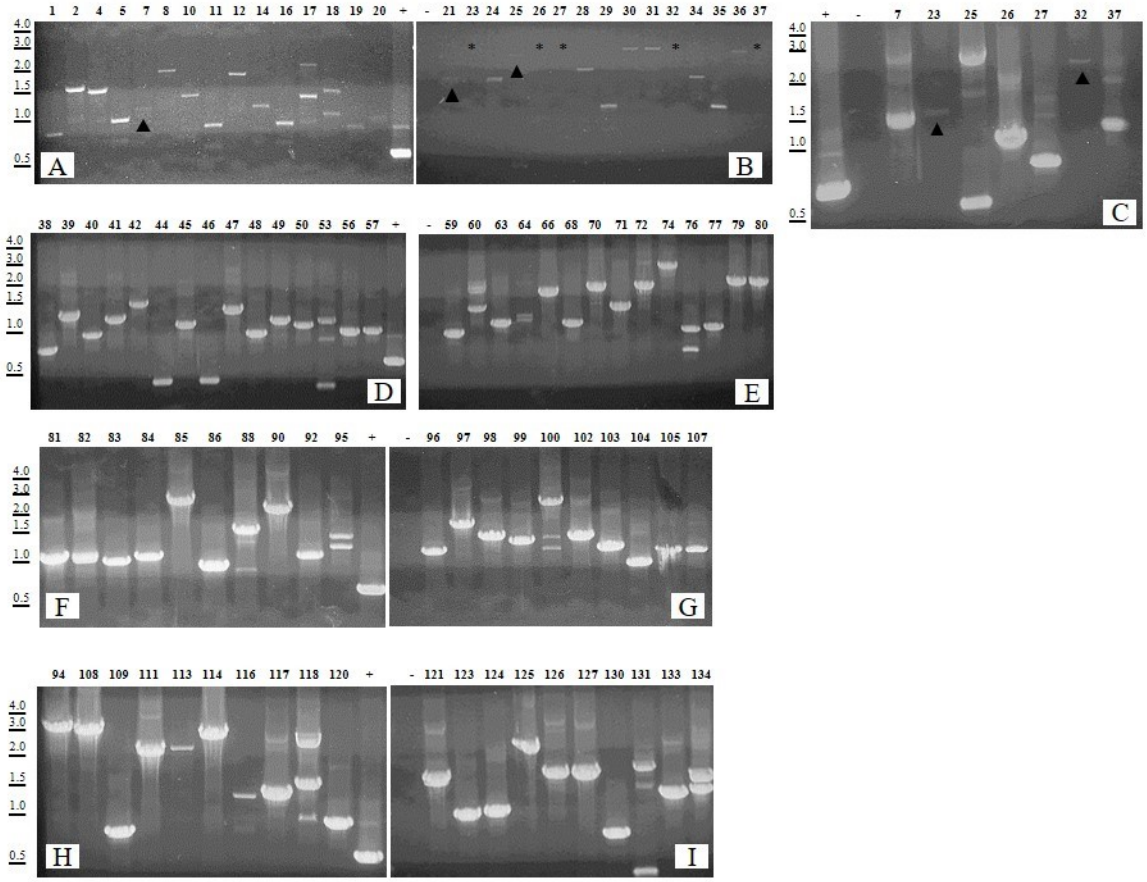
Suppl. Figure 11. p21 C-terminus region doesn't bind to VLP and VHL172. Yeast cells co-expressing VLP (A) or VHL172 (B) together with C-terminus region p21 (CT p21, residues 50-164) are assayed by Y2H system. Serial dilution of yeast cells are seeded on permissive (+HIS, left) and on selective media (-HIS, + 30mM 3AT, right). C+ and C- are positive and negative assay controls present in all experiments. Each panel contains auto activation controls represented by yeast strain co-transformed with the empty vector (Ø). Images are representative of at least three independent experiments (n clone≥15).



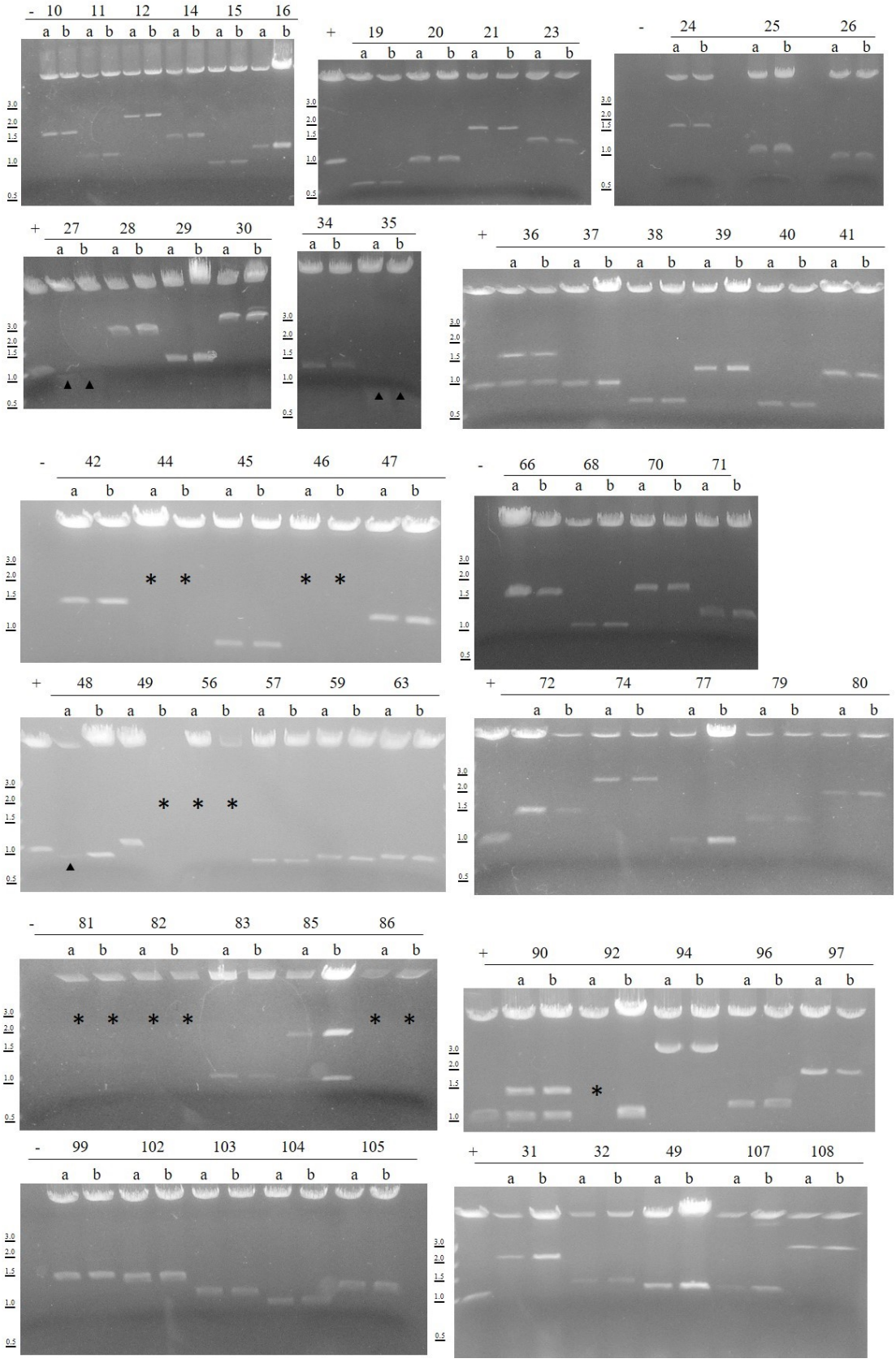
Suppl. Figure 12. The chimera bait-protein expression. The 32 kDa Gal4-DBD-VLP is pointed by the black arrow. (Whereas in the other lines are detected proteins not discuss in this section).

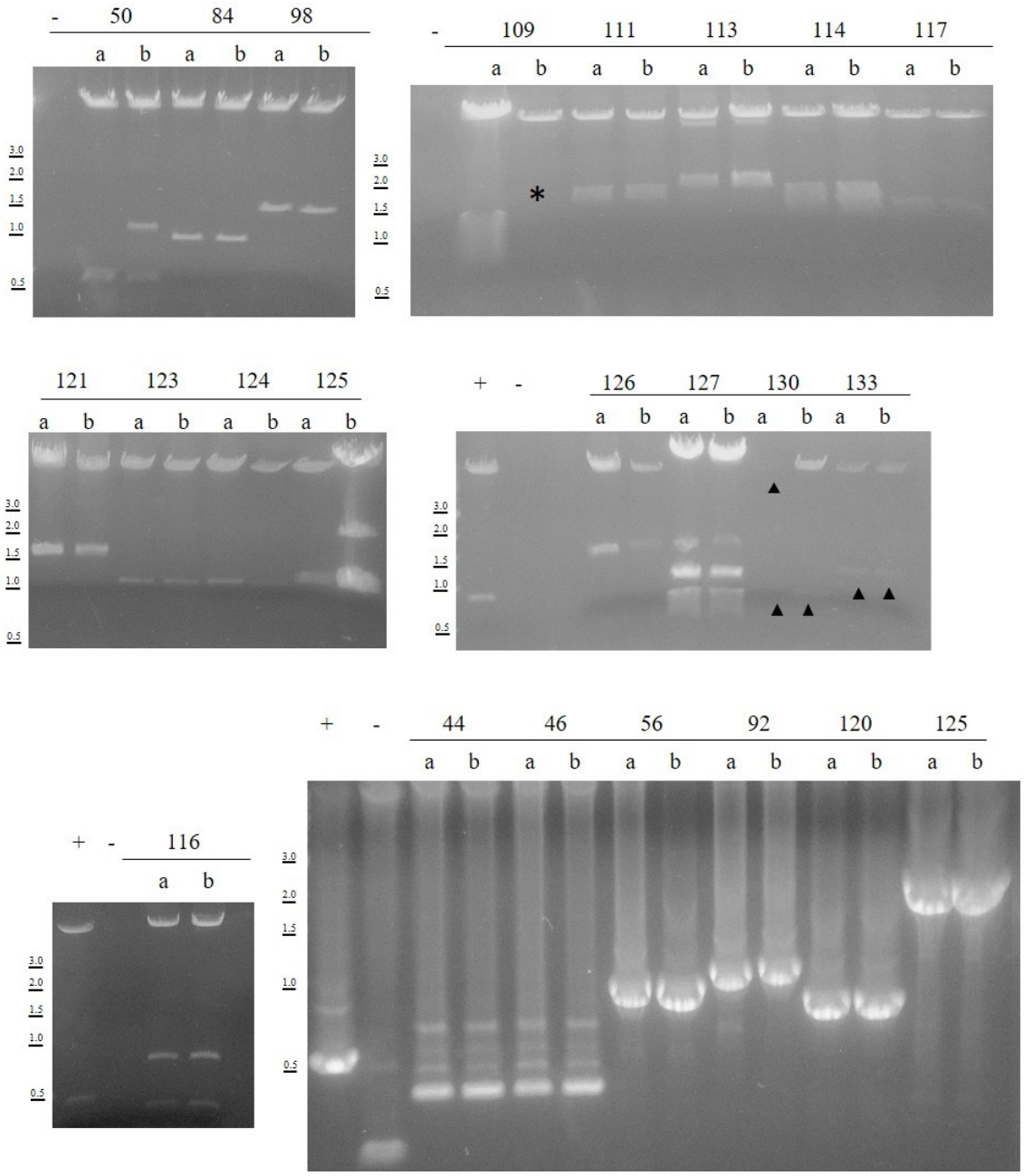


Suppl. Figure 13. Colony patches. Six days post-mating, all the blue colonies were patched out on higher stringency agar plates (QDO/X/A) in order to avoid false positive interactions. Red crosses indicate false positive clones that were not further analyzed.

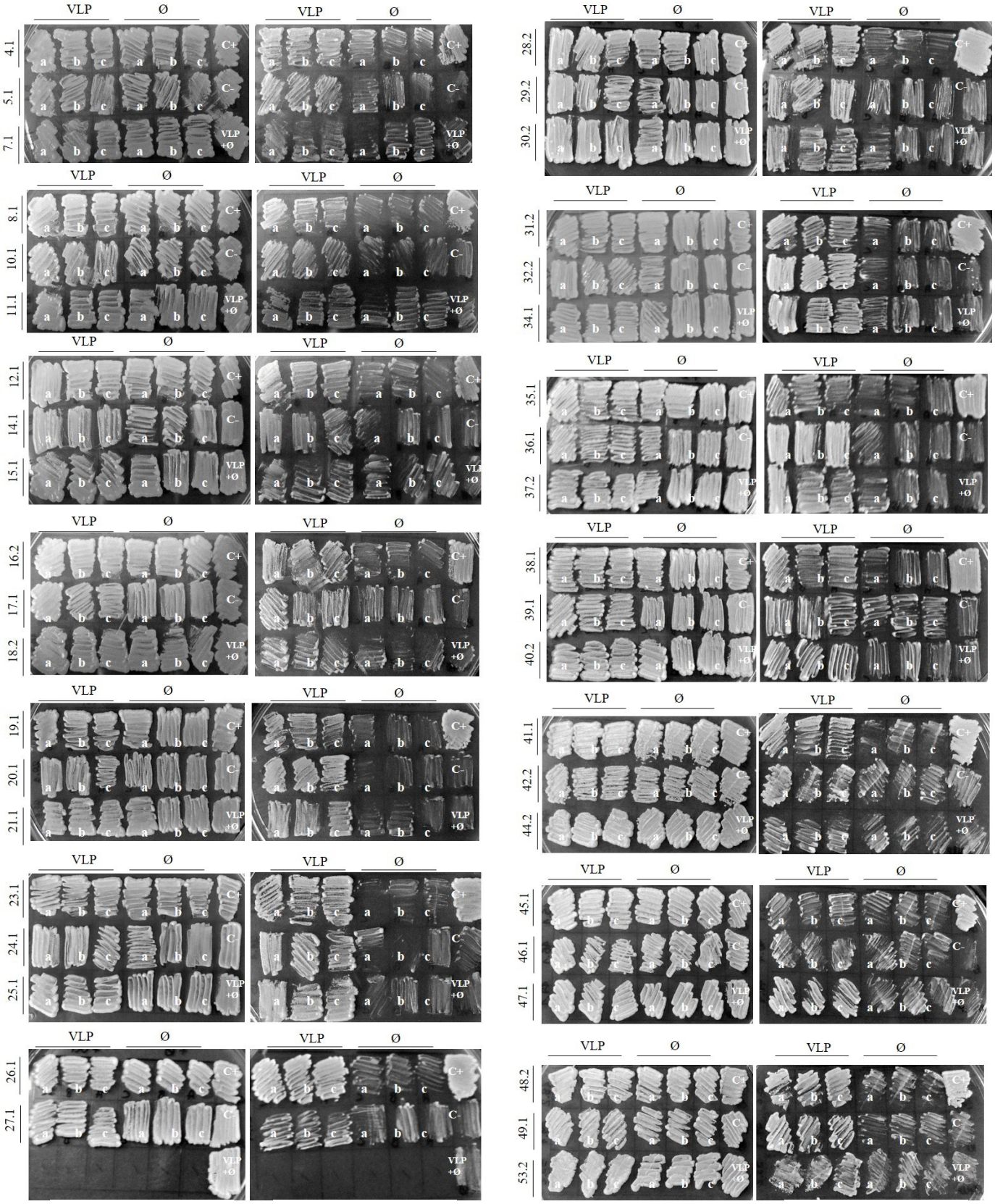


Suppl. Figure 14. (Previous page) AD/plasmid insert amplifications. PCR products were differently separated on 0.8% agarose gel. Each line has a number, which corresponds to a selected blue clone, whereas + is the positive control (i.e. a pGADT7 sample) and – the negative one (i.e. reaction mix without any plasmid). The black triangle are used to pint out insert bands less visible. While * indicates failed amplification, which are then repeated in other independent experiment.

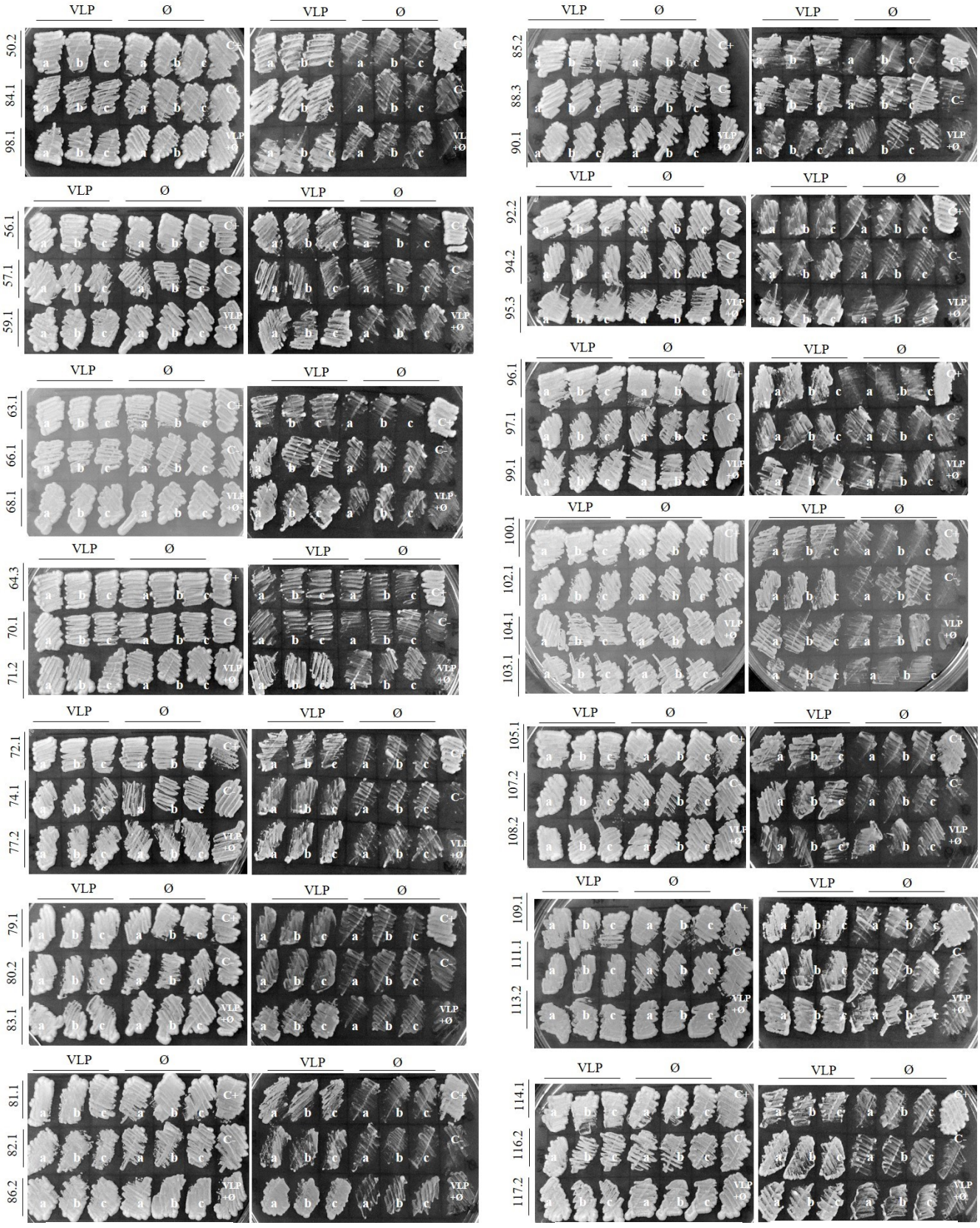


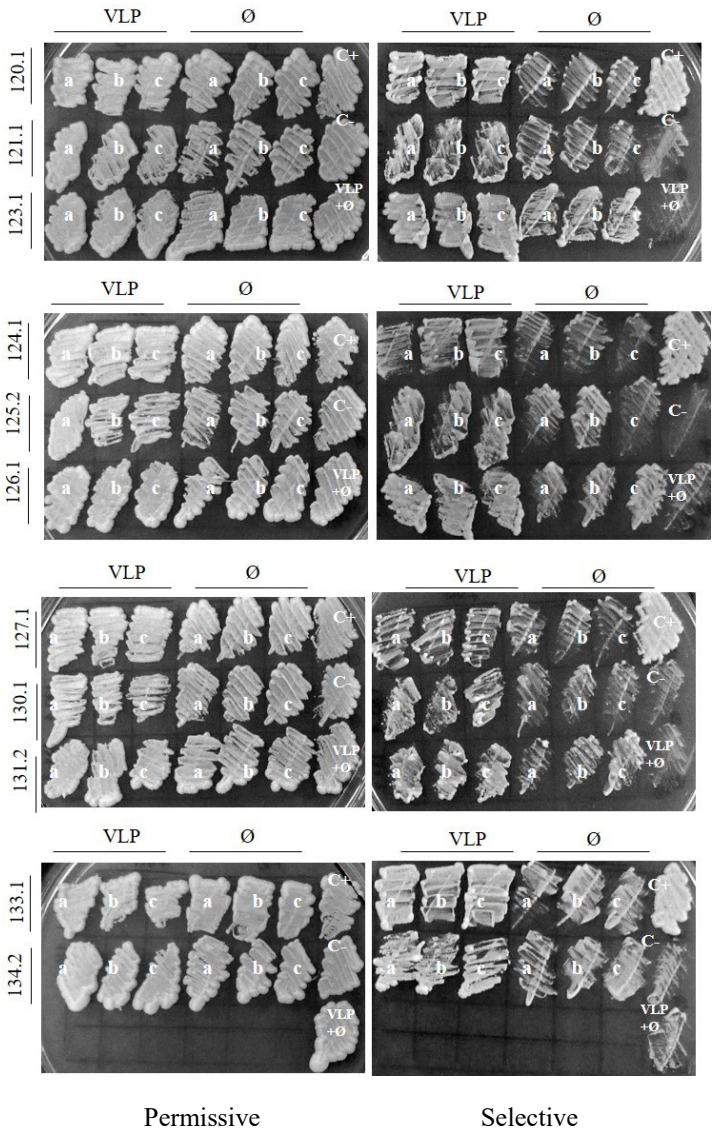


Suppl. Figure 15. Restriction analysis of the selected AD/plasmids. For each positive clones, a couple of clones (i.e. a and b) were analyzed by restriction reaction using EcoRI and BamHI enzymes at 37°C. Each line has a number, which properly corresponds to a selected clone, whereas positive control is indicated with + and negative with -. The black triangle is used to point out bands less visible. While * indicates failed restriction reaction, which is then analyzed by PCR.

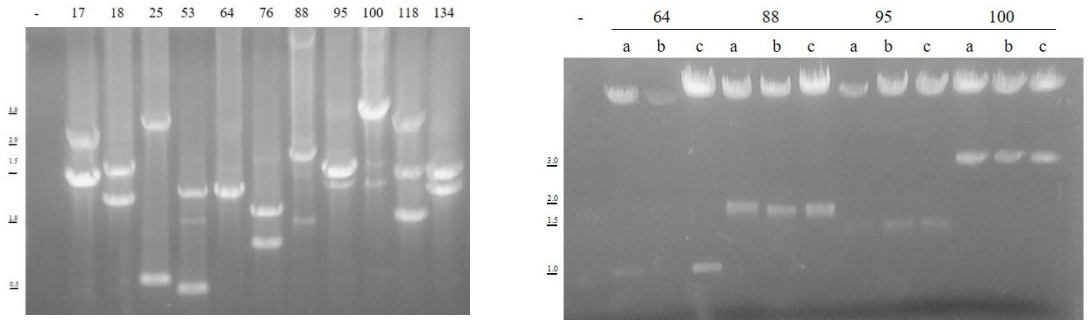


Permissive Selective Permissive Selective

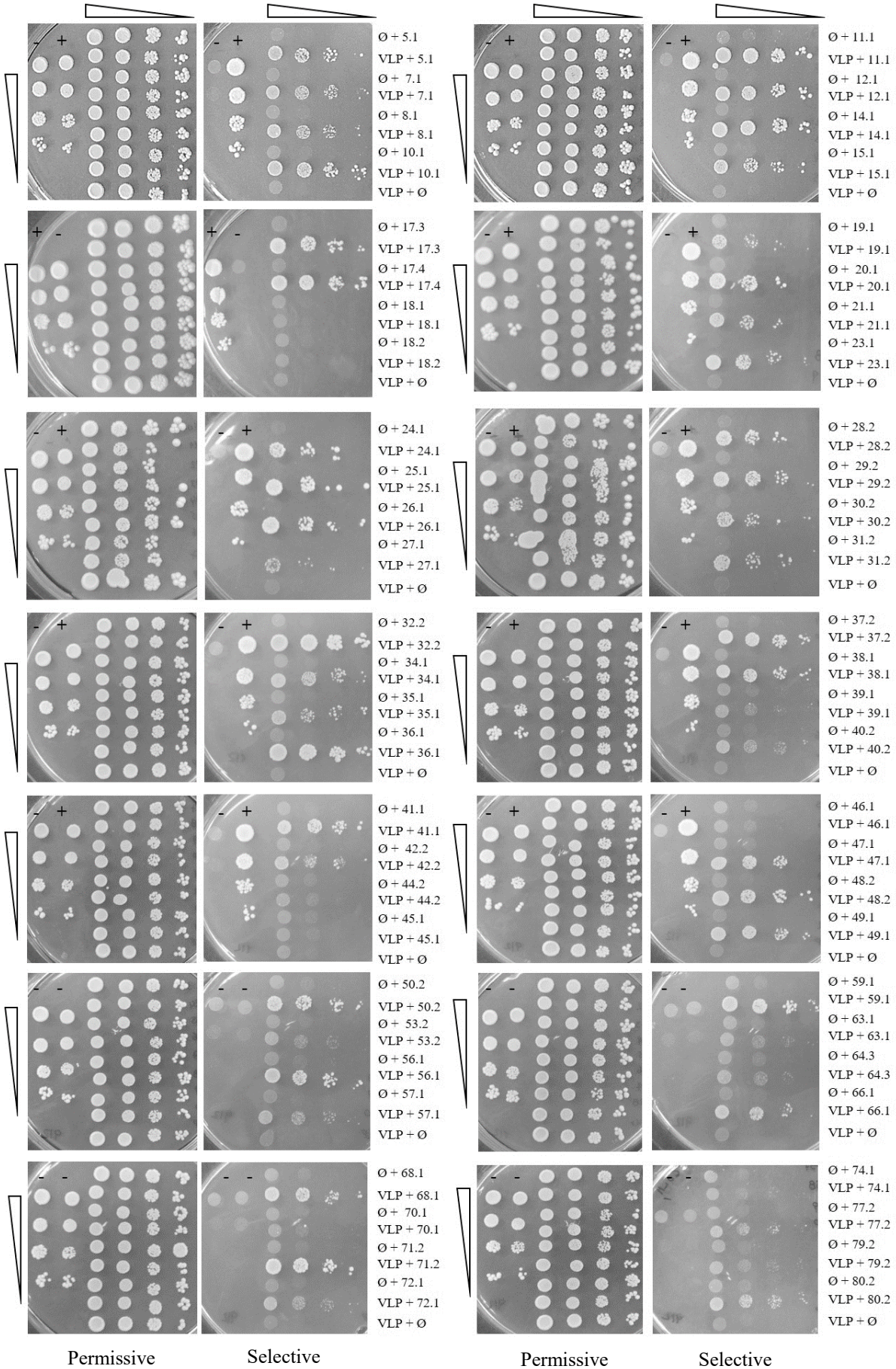

VLP
Ø
50.2
84.1
98.1
56.1
57.1
59.1
63.1
66.1
68.1
64.3
70.1
71.2
72.1
74.1
77.2
79.1
80.2
83.1
81.1
82.1
86.2
85.2
88.3
90.1
92.2
94.2
95.3
96.1
97.1
99.1
100.1
102.1
104.1
103.1
105.1
107.2
108.2
109.1
111.1
113.2
114.1
116.2
117.2
C+
C-
VLP +Ø
Permissive
Selective
Permissive
Selective

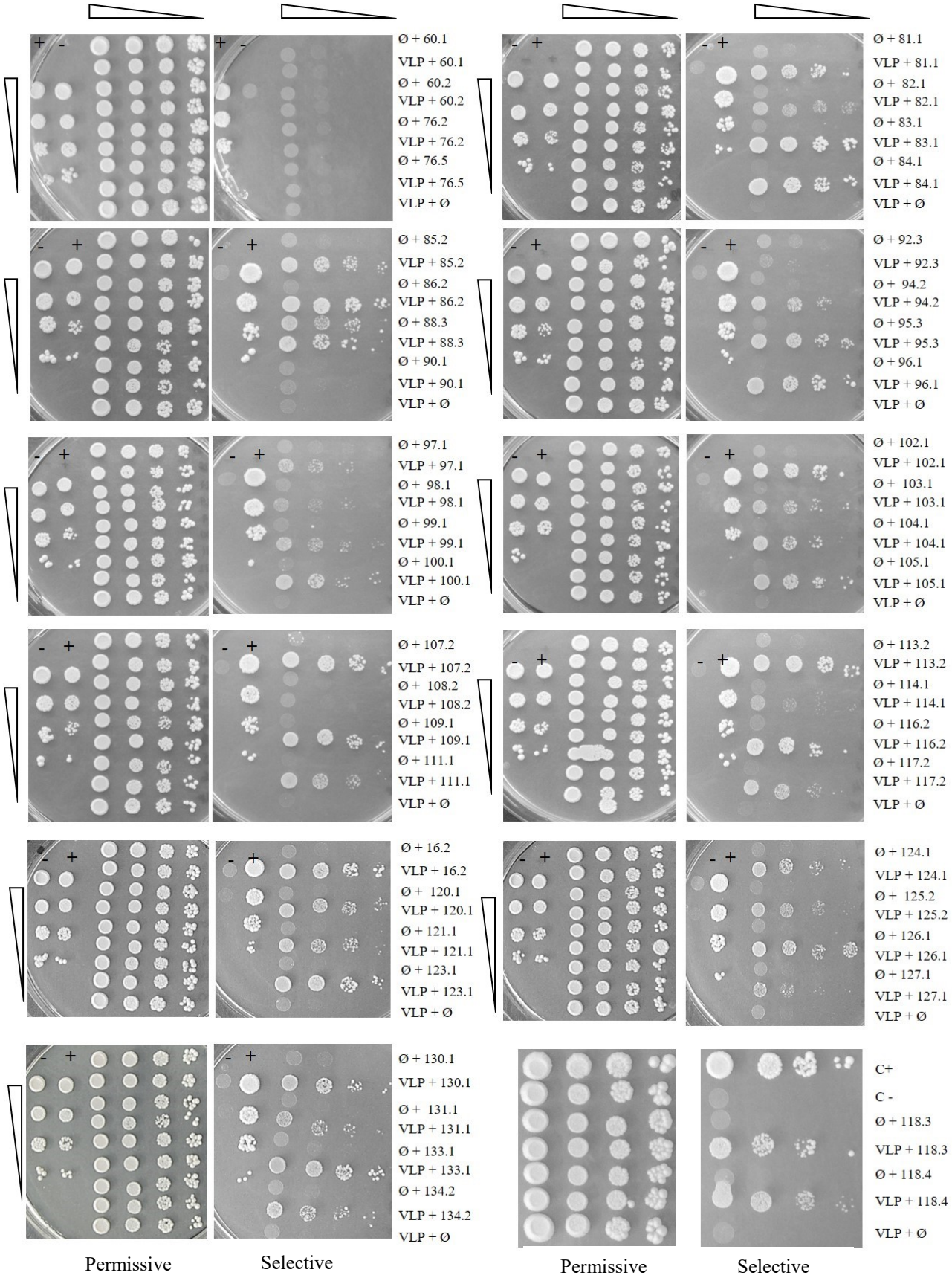


Suppl. Figure 16. Yeast patches. On permissive (right) and selective media (left), three independent clones were tested. The numbers indicate the analyzed clones and the letters (i.e. a, b and c) point out the three clones picked up from the same co-transformation plates containing either pGBKT7 VLP or pGBKT7 empty (Ø) vectors. Standard Y2H positive and negative controls are + and –, whereas VLP auto activation are yeast cells co-transformed with pGKBT7 VLP and pGADT7 empty vectors.

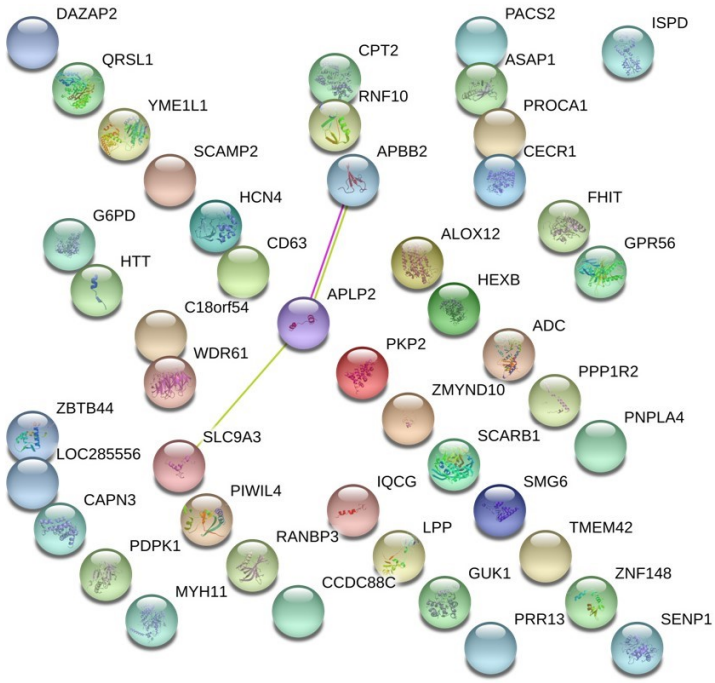


Suppl. Figure 17. PCR on mixed AD/plasmid population. (Right) Bacteria transformation were repeated with clones previously classified as 'mixed'. For each transformation, a pool of clones was collected to again purified plasmids, then analyzed by PCR. As previously, the majority of clones confirmed to contain a mix of plasmids. Whereas an insert seem to prevail on the other on lines 64, 88, 95 and 100. (Left) These were then analyzed by EcoRI/BamHI restriction enzyme reaction, observing that three independent clone had the same pattern. In both picture, - is a negative control of the reactions.


Ø + 5.1
VLP + 5.1
Ø + 7.1
VLP + 7.1
Ø + 8.1
VLP + 8.1
Ø + 10.1
VLP + 10.1
VLP + Ø
Ø + 11.1
VLP + 11.1
Ø + 12.1
VLP + 12.1
Ø + 14.1
VLP + 14.1
Ø + 15.1
VLP + 15.1
VLP + Ø
Ø + 17.3
VLP + 17.3
Ø + 17.4
VLP + 17.4
Ø + 18.1
VLP + 18.1
Ø + 18.2
VLP + 18.2
VLP + Ø
Ø + 19.1
VLP + 19.1
Ø + 20.1
VLP + 20.1
Ø + 21.1
VLP + 21.1
Ø + 23.1
VLP + 23.1
VLP + Ø
Ø + 24.1
VLP + 24.1
Ø + 25.1
VLP + 25.1
Ø + 26.1
VLP + 26.1
Ø + 27.1
VLP + 27.1
VLP + Ø
Ø + 28.2
VLP + 28.2
Ø + 29.2
VLP + 29.2
Ø + 30.2
VLP + 30.2
Ø + 31.2
VLP + 31.2
VLP + Ø
Ø + 32.2
VLP + 32.2
Ø + 34.1
VLP + 34.1
Ø + 35.1
VLP + 35.1
Ø + 36.1
VLP + 36.1
VLP + Ø
Ø + 37.2
VLP + 37.2
Ø + 38.1
VLP + 38.1
Ø + 39.1
VLP + 39.1
Ø + 40.2
VLP + 40.2
VLP + Ø
Ø + 41.1
VLP + 41.1
Ø + 42.2
VLP + 42.2
Ø + 44.2
VLP + 44.2
Ø + 45.1
VLP + 45.1
VLP + Ø
Ø + 46.1
VLP + 46.1
Ø + 47.1
VLP + 47.1
Ø + 48.2
VLP + 48.2
Ø + 49.1
VLP + 49.1
VLP + Ø
Ø + 50.2
VLP + 50.2
Ø + 53.2
VLP + 53.2
Ø + 56.1
VLP + 56.1
Ø + 57.1
VLP + 57.1
VLP + Ø
Ø + 59.1
VLP + 59.1
Ø + 63.1
VLP + 63.1
Ø + 64.3
VLP + 64.3
Ø + 66.1
VLP + 66.1
VLP + Ø
Ø + 68.1
VLP + 68.1
Ø + 70.1
VLP + 70.1
Ø + 71.2
VLP + 71.2
Ø + 72.1
VLP + 72.1
VLP + Ø
Ø + 74.1
VLP + 74.1
Ø + 77.2
VLP + 77.2
Ø + 79.2
VLP + 79.2
Ø + 80.2
VLP + 80.2
VLP + Ø
Permissive
Selective
Permissive
Selective


Ø + 60.1
VLP + 60.1
Ø + 60.2
VLP + 60.2
Ø + 76.2
VLP + 76.2
Ø + 76.5
VLP + 76.5
VLP + Ø
Ø + 81.1
VLP + 81.1
Ø + 82.1
VLP + 82.1
Ø + 83.1
VLP + 83.1
Ø + 84.1
VLP + 84.1
VLP + Ø
Ø + 85.2
VLP + 85.2
Ø + 86.2
VLP + 86.2
Ø + 88.3
VLP + 88.3
Ø + 90.1
VLP + 90.1
VLP + Ø
Ø + 92.3
VLP + 92.3
Ø + 94.2
VLP + 94.2
Ø + 95.3
VLP + 95.3
Ø + 96.1
VLP + 96.1
VLP + Ø
Ø + 97.1
VLP + 97.1
Ø + 98.1
VLP + 98.1
Ø + 99.1
VLP + 99.1
Ø + 100.1
VLP + 100.1
VLP + Ø
Ø + 102.1
VLP + 102.1
Ø + 103.1
VLP + 103.1
Ø + 104.1
VLP + 104.1
Ø + 105.1
VLP + 105.1
VLP + Ø
Ø + 107.2
VLP + 107.2
Ø + 108.2
VLP + 108.2
Ø + 109.1
VLP + 109.1
Ø + 111.1
VLP + 111.1
VLP + Ø
Ø + 113.2
VLP + 113.2
Ø + 114.1
VLP + 114.1
Ø + 116.2
VLP + 116.2
Ø + 117.2
VLP + 117.2
VLP + Ø
Ø + 16.2
VLP + 16.2
Ø + 120.1
VLP + 120.1
Ø + 121.1
VLP + 121.1
Ø + 123.1
VLP + 123.1
VLP + Ø
Ø + 124.1
VLP + 124.1
Ø + 125.2
VLP + 125.2
Ø + 126.1
VLP + 126.1
Ø + 127.1
VLP + 127.1
VLP + Ø
Ø + 130.1
VLP + 130.1
Ø + 131.1
VLP + 131.1
Ø + 133.1
VLP + 133.1
Ø + 134.2
VLP + 134.2
VLP + Ø
C+
C -
Ø + 118.3
VLP + 118.3
Ø + 118.4
VLP + 118.4
VLP + Ø
Permissive
Selective
Permissive
Selective

Suppl. Figure 18. (Previous page) Spot test on screened clones. Yeats cells serial dilutions were spotted on permissive and selective media (i.e. supplemented with 60mM 3AT). Positive (+) and negative (-) controls properly indicate yeast cells growth and lethality. Similarly, auto-activation controls (i.e. empty vector co-transformed with VLP or a library clone) point out that the detected interactions are related to simultaneous bait and prey proteins expression. On 60mM 3AT, yeast cells co-transformed with VLP (as bait) and library clone are able to survive in absence of histidine due to their interaction. Some clones do not survive suggesting that these are false positives.



Suppl. Figure 19. Functional network. A preliminary protein interaction network built on the annotated-dataset of clones using STRING[209]. It contains only two edges with an average node degree of 0.0087 and an avg. local clustering coefficient of 0.0435. The PPI enrichment p-value is 0.825.

Supplementary Tables

Plasmid	*Primer*	*Primer sequence (5'-3')*	*PCR Template*	*Starting vector*	*Restriction enzyme*	*Method*	*Protein*	*Technique*
pVHL30-pGADT7	VHL30-AD For	ggaggccagtgaattcATGCCCCGGAGGGCGGAGAA	VHL30-pcDNA3.1+	Empty pGADT7	EcoRI	In-fusion	Gal4AD-pVHL30 (1-213)	Y2H
	VHL30-AD Rev	cacccgggtggaattgTCAATCTCCCATCCGTTGAT						
pVHL-NT-pGADT7	VHL30-AD For	ggaggccagtgaattcATGCCCCGGAGGGCGGAGAA	VHL30-pcDNA3.1+	Empty pGADT7	EcoRI	In-fusion	Gal4AD-pVHL-NT (1-53)	Y2H
	VHLNT-AD Rev	cacccgggtggaattgTCACTCCTCCTCGGCGCCCA						
pVHL-β-pGADT7	VHL19-AD For	ggaggccagtgaattcATGGAGGCCGGGCGGCCGCG	pVHL19-pGADT7	Empty pGADT7	EcoRI/BamHI	In-fusion	Gal4AD-pVHL-β (54-157)	Y2H
	VHLβ-AD Rev	cacccgggtggaattgCTAAGTATACACTGGCAGTGTGATATTGGC						
pVHL19-pGADT7	VHL19-AD For	ggaggccagtgaattcATGGAGGCCGGGCGGCCGCG	VHL30-pcDNA3.1+	Empty pGADT7	EcoRI	In-fusion	Gal4AD-pVHL19 (54- 213)	Y2H
	VHL30-AD Rev	cacccgggtggaattgTCAATCTCCCATCCGTTGAT						
p21-pGBKT7	p21-BD For	catggaggccgaattcATGTCAGAACCGGCTGGGGATGTCC	p21-pcDNA3.1+	Empty pGBKT7	EcoRI	In-fusion	Gal4BD-p21 (1- 164)	Y2H
	p21-BD Rev	ggatccccgggaattgCTAGGGCTTCCTCTTGGAGAAGATCAGC						
p21-NT-pGBKT7	p21-BD For	catggaggccgaattcATGTCAGAACCGGCTGGGGATGTCC	p21-pGBKT7	Empty pGBKT7	EcoRI	In-fusion	Gal4BD-p21-NT (1-49)	Y2H
	p21NT-BD Rev	ggatccccgggaattgCTACCATCGCTCACGGGCCTCCTGGATGC						
p21-CT-pGBKT7	p21CT-BD For	catggaggccgaattcAACTTCGACTTTGTCACCGAGACACC	p21-pGBKT7	Empty pGBKT7	EcoRI	In-fusion	Gal4BD-p21- CT (50-164)	Y2H
	p21-BD Rev	ggatccccgggaattgCTACCATCGCTCACGGGCCTCCTGGATGC						
p27-pGBKT7	p27-BD For	catggaggccgaattcATGTCAAACGTGCGAGTGTC	p27-pcDNA3.1+	Empty pGBKT7	EcoRI	In-fusion	Gal4BD-p27 (1-198)	Y2H

	p27-BD Rev	ggatccccgggaattgTCACGTTTGACGTCTTC TGAG						
p27-NT-pGBKT7	p27-BD For	catggaggccgaattcATGTCAAACGTGCGAGT GTC	p27-pGBKT7	Empty pGBKT7	EcoRI	In-fusion	Gal4BD-p27-NT (1-60)	Y2H
	p27NT-BD Rev	ggatccccgggaattgCTACCACTTGCGCTGGC TCGCCTCTTCC						
p27-CT-pGBKT7	p27CT-BD For	catggaggccgaattcAATCACAAACCCCTAGA GGGCAAGTACG	p27-pGBKT7	Empty pGBKT7	EcoRI	In-fusion	Gal4BD-p27-CT (61-198)	Y2H
	p27CT-BD Rev	ggatccccgggaattgCATGGTGGCGGATCCGA GCTCGGTACC						
p57-pGBKT7	p57-BD For	catggaggccgaattcATGTCCGACGCGTCCCT CCG	p57-pcDNA3.1+	Empty pGBKT7	EcoRI	In-fusion	Gal4BD-p57 (1-316)	Y2H
	p57-BD Rev	ggatccccgggaattgTCACCGCAGCCTCTTGC GCG						
p57-NT-pGBKT7	p57NT-BD For	catggaggccgaattcATGTCCGACGCGTCCCT CCGCAGC	p57-pGBKT7	Empty pGBKT7	EcoRI	In-fusion	Gal4BD-p57-NT (1-61)	Y2H
	p57NT-BD Rev	ggatccccgggaattgCTACCAGCGGTTCTGGT CCTCGGCGTT C						
p57-CT-pGBKT7	p57CT-BD For	catggaggccgaattcGATTACGACTTCCAGCA GGACATGC	p57-pGBKT7	Empty pGBKT7	EcoRI	In-fusion	Gal4BD-p57-CT (62-316)	Y2H
	p57CT-BD Rev	ggatccccgggaattgCTACCGCAGCCTCTTGC GCGGGGTCTGC						
P35L-p27-pGBKT7	P35L-p27 For	CTTCGTGGTCCACCAGGCCGAAGAGGT TC	p27-pGBKT7	-	-	Mutagenesis	Gal4BD-p27-P35L	Y2H
	P35L-p27 Rev	GAACCTCTTCGGCCTGGTGGACCACGA AG						
D37N-p27-pGBKT7	D37N-p27 For	TTAACTCTTCGTGGTTCACCGGGCCGA AGAG	p27-pGBKT7	-	-	Mutagenesis	Gal4BD-p27-D37N	Y2H
	D37N-p27 Rev	CTCTTCGGCCCGGTGAACCACGAAGAG TTAA						
E40K-p27-pGBKT7	E40K-p27 For	CCGGGTTAACTTTTCGTGGTCCACCGG G	p27-pGBKT7	-	-	Mutagenesis	Gal4BD-p27-E40K	Y2H
	E40K-p27 Rev	CCCGGTGGACCACGAAAAGTTAACCCG G						
T42A-p27-pGBKT7	T42A-p27 For	AAGTCCCGGGCTAACTCTTCGTGGTCC ACC	p27-pGBKT7	-	-	Mutagenesis	Gal4BD-p27-T42A	Y2H

	T42A-p27 Rev	GGTGGACCACGAAGAGTTAGCCCGGGACTT						
HA-VHL30-pcDNA3.1^{+}	HA-F For	taccgagctcggatcATGGAGTACCCATACGACGTACCAGATTACG	VHL30-pGADT7	Empty pcDNA3.1+	EcoRI/BamHI	In-fusion	HA-pVHL30 (1-213)	Cells
	HA-R Rev	gatatctgcagaattTCAATCTCCCATCCGTTGATGTGCAATGCGC						
HA-VHL19-pcDNA3.1^{+}	HA-F For	taccgagctcggatcATGGAGTACCCATACGACGTACCAGATTACG	VHL19-pGADT7	Empty pcDNA3.1+	EcoRI/BamHI	In-fusion	HA-pVHL19 (54-213)	Cells
	HA-R Rev	gatatctgcagaattTCAATCTCCCATCCGTTGATGTGCAATGCGC						
HA-VLP-pcDNA3.1	-	-	-	Empty pcDNA3.1+	BamHI/XhoI	In-fusion	HA-VLP (1-139)	Cells
	-	-		HA-VLP-pGADT7 (insert)	BglII/XhoI			
VHL172-pGADT7	Del114-154 For	cgcatccacagctaccgagtgtatactctgaaagag	pVHL30-pGADT7	-	-	Mutagenesis	HA-VHL172 (1-172)	Y2H
	Del114-154 For	ctctttcagagtatacactcggtagctgtggatgcg						
HA-VHL172 pcDNA3.1+	HA-F For	taccgagctcggatcATGGAGTACCCATACGACGTACCAGATTACG	VHL172-pGADT7	Empty pcDNA3.1+	EcoRI/BamHI	In-fusion	HA-VHL172 (1-172)	Cells
	HA-R Rev	gatatctgcagaattTCAATCTCCCATCCGTTGATGTGCAATGCGC						
GFP-HA-VHL30-pcDNA3.1^{+}	GFP-VHL For	TTAAACTTAAGCTTGGTACCATGGTGAGCAAGGGCGAGGAGCTGTTCACC	GFP-pcDNA3.1+	HA-VHL30-pcDNA3.1^{+}	KpnI	In-fusion	GFP-HA-VHL30	Cells
	GFP-VHL Rev	catGATCCGAGCTCGGTACCaCTTGTACAGCTCGTCCATGCCGAGAGTGATCCC						
GFP-HA-VHL19-pcDNA3.1^{+}	GFP-VHL For	TTAAACTTAAGCTTGGTACCATGGTGAGCAAGGGCGAGGAGCTGTTCACC	GFP-pcDNA3.1+	HA-VHL19-pcDNA3.1^{+}	KpnI	In-fusion	GFP-HA-VHL19	Cells
	GFP-VHL Rev	catGATCCGAGCTCGGTACCaCTTGTACAGCTCGTCCATGCCGAGAGTGATCCC						
GFP-HA-VLP-pcDNA3.1^{+}	GFP-VHL For	TTAAACTTAAGCTTGGTACCATGGTGAGCAAGGGCGAGGAGCTGTTCACC	GFP-pcDNA3.1+	HA-VLP-pcDNA3.1	KpnI	In-fusion	GFP-HA-VLP	Cells
	GFP-VHL Rev	catGATCCGAGCTCGGTACCaCTTGTACAGCTCGTCCATGCCGAGAGTGATCCC						

GFP-HA-VHL172-pcDNA3.1^{+}	GFP-VHL For	TTAAACTTAAGCTTGGTACCATGGTGAGCAAGGGCGAGGAGCTGTTCACC	GFP-pcDNA3.1+	HA-VHL172 pcDNA3.1+	KpnI	In-fusion	GFP-HA-VHL172	Cells
	GFP-VHL Rev	catGATCCGAGCTCGGTACCaCTTGTACAGCTCGTCCATGCCGAGAGTGATCCC						
HA-p57-pcDNA3.1^{+}	HA-F For	taccgagctcggatcATGGAGTACCCATACGACGTACCAGATTACG	p57-pGADT7	Empty pcDNA3.1+	EcoRI/BamHI	In-fusion	HA-p57	Cells
	HA-p57 Rev	gatatctgcagaattTCACCGCAGCCTCTTGCGCGGGGTCTGC						
Myc-VLP-pcDNA3.1^{+}	Myc-VLP For	tactgagtcggatccATGGAGGAGCAGAAGCTG	VLP-pGBKT7	Empty pcDNA3.1+	EcoRI/BamHI	In-fusion	Myc-VLP	Cells
	Myc-VLP Rev	GATATCTGCAGAATTCTTAGGGTATACACTGCAGTGTGATG						
HIF-1α pGADT7	HIF-For	ggaggccagtgaattcATGGAGGGCGCCGGCGGCGC	HIF-1α - pcDNA3.1+	Empty pGADT7	EcoRI	In-fusion	HA- HIF-1α	Y2H
	HIF-Rev	cacccgggtggaattgTCAGTTAACTTGATCCAAAG						
GFP-HA-ΔNT-VHL172-pcDNA3.1^{+}	Del114-154 For	cgcatccacagctaccgagtgtatactctgaaagag	GFP-HA-VHL19-pcDNA3.1^{+}	-	-	Mutagenesis	GFP-HA-ΔNT-VHL172	Cells
	Del114-154 For	ctctttcagagtatacactcggtagctgtggatgcg						
MCSI:βVHL, MCSII:PHD3 pAB326	FLAG-VHLβ-Bridge-For	tgtatcgccggaattcGATTACAAGGATGACGACGATAAGgccCTGCGCTCGGTGAACTCGCGCGAGC	MCSI:Ø, MCSII:PHD3 pAB326	VHL30 pGADT7	EcoRI on MCSI	In-fusion	βVHL and PHD3 simultaneously	Reconstituted Y2H
	VHLβ-Bridge-Rev	ggatccccgggaattAAGTATACACTGGCAGTGTGATATTGGC						
MCSI:VLP, MCSII:PHD3 pAB326	VHLL-Bridge-For	tgtatcgccggaattcGATTACAAGGATGACGACGATAAGATGCCCTGGAGAGCGGGGAACGGG	MCSI:Ø, MCSII:PHD3 pAB326	VLP pGADT7	EcoRI on MCSI	In-fusion	VLP and PHD3 simultaneously	Reconstituted Y2H
	VHLL-Bridge-Rev	ggatccccgggaattAGGGTATACACTGCAGTGTGATGTTGGC						
ODD pGADT7	Oligomer paired	ggaggccagtgaattcGACACAGATTTAGACTTGGAGATGTTAGCTCCCTATATCCCAATGGATGATGACTTCCAGTTACGTTaattccaccgggtg	Empty pGADT7	-	EcoRI	In-fusion	HA-ODD	Reconstituted Y2H

		cacccgggtggaattAACGTAACTGGAAGTCATCATCCATTGGGATATAGGGAGCTAACATCTCCAAGTCTAAATCTGTGTCgaattcactggcctcc						
MCSI:VLP, MCSII:Ø pAB340	-	-	MCSI:VHL19, MCSII:Ø pAB340	VLP pGADT7	EcoRI/BamHI on MCSI	Ligation	HA-VLP	Reconstituted Y2H
MCSI:VLP, MCSII:PHD3 pAB340	-	.	MCSI:VLP, MCSII:Ø pAB340	MCSI:Ø, MCSII:PHD3 pAB326	NotI/BglII on MCSI for pAB340 and MCSII for pAB326	Ligation	VLP and PHD3 simultaneously	Reconstituted Y2H
MCSI:βVHL, MCSII:Ø pAB340	-	-	MCSI:VHL19, MCSII:Ø pAB340	βVHL pGADT7	EcoRI/BamHI on MCSI	Ligation	βVHL and PHD3	Reconstituted Y2H
MCSI:VHL172, MCSII:Ø pAB340	-	-	MCSI:VLP, MCSII:Ø pAB340	VHL172 pGADT7	EcoRI/BamHI on MCSI	Ligation	HA-VHL172	Reconstituted Y2H
MCSI:VHL172, MCSII:PHD3 pAB340	pBridge-VHL172-For	tgtatcgccggaattcatgccccggagggcggagaactgggacgaggcc	MCSI:Ø, MCSII:PHD3 pAB326	VHL172 pGADT7	EcoRI on MCSI	In-fusion	VHL172 and PHD3 simultaneously	Reconstituted Y2H
	pBridge-VHL172-Rev	ggatccccgggaattgtcaatctcccatccgttgatgtgcaatgcgctc						
MDM2 329-452 pGADT7	MDM2 (329-433) For	catggaggccgaattcTGGGCCCTTCGTGAGAATTGGCTTCC	MDM2 pcDNA3.1+	Empty pGADT7	EcoRI	In-fusion	HA-MDM2 (329-452)	Y2H
	MDM2 (351-452) Rev	ggatccccgggaattgTTAATGGACAATGCAACCATTTTTAGGTCG						
MDM2 pGADT7	MDM2 For	catggaggccgaattcATGGTGAGGAGCAGGCAAATGTGCAATACC	MDM2 pcDNA3.1+	Empty pGADT7	EcoRI on MCSI	In-fusion	HA-MDM2 (1-491)	Y2H
	MDM2 Rev	ggatccccgggaattgttaGGGGAAATAAGTTAGCACAATCATTTGAATTGG						
GAL4-AD	GAL4-AD For	TACCACTACAATGGATG	-	-	-	Sequencing	-	-
GAL4-AD	GAL4-AD Rev	GGTTTTTCAGTATCTACG	-	-	-	Sequencing	-	-
T7 promoter	T7	taatacgactcactataggg				Sequencing		

Suppl. Table 1. Recombinant plasmids.

Plasmid	*Cat. n.*
p21-pcDNA3.1^{+}	GeneScript: OHu26670
p27-pcDNA3.1^{+}	GeneScript: OHu27895
p57-pcDNA3.1^{+}	GeneScript: OHu27234
pVHL30-pcDNA3.1+	GeneScript: OHu23297
MDM2-pcDNA3.1^{+}	GeneScript: OHu28568
HIF1α-pcDNA3.1^{+}	GeneScript: OHu27176
VHLL-pcDNA3.1^{+}	GeneScript: OHu21555
MCSI:Ø, MCSII:PHD3 pAB326	Received from Bex's Lab
MCSI:VHL19, MCSII:Ø pAB340	Received from Bex's Lab

Suppl. Table 2. Read-to-use plasmids.

Clone	PCR (ca. bp)	PCR Primers	Sequencing Primer	Dataset	tblastx	Uniprot	Localization
1.1	900	Gal4AD F/R	T7 For	Annotated	AZIN 2	Q96A70	Nucleus
2.1	1500	Gal4AD F/R	T7 For	Annotated	FHIT	P49789	Nucleus/Mitocondrion
10.1	1500	Gal4AD F/R	T7 For	Annotated	RANBP3	Q9H6Z4	Nucleus
14.1	1300	Gal4AD F/R	T7 For	Annotated	RNF10	Q8N5U6	Nucleus
19.1	1000	Gal4AD F/R	T7 For	Annotated	PIWIL4	Q7Z3Z4	Nucleus
23.1	1500	Gal4AD F/R	T7 For	Annotated	PRR13	Q9NZ81	Nucleus
26.1	1200	Gal4AD F/R	T7 For	Annotated	ZNF148	Q9UQR1	Nucleus
27.1	1000	Gal4AD F/R	T7 For	Annotated	RNF10	Q8N5U6	Nucleus
28.2	2100	Gal4AD F/R	T7 For	Annotated	PKP2	Q99959	Nucleus
32.2	2800	Gal4AD F/R	T7 For	Annotated	DAZAP2	Q15038	Nucleus
84.1	1200	Gal4AD F/R	T7 For	Annotated	ZBTB44	Q8NCP5	Nucleus
90.1	2200	Gal4AD F/R	T7 For	Annotated	SENP1	Q9P0U3	Nucleus
96.1	1200	Gal4AD F/R	T7 For	Annotated	G6PD	P11413	Nucleus/cytosol
100.1	3000/1600/1400	Gal4AD F/R	T7 For	Annotated	APLP2	Q06481	Nucleus/membrane
104.1	1100	Gal4AD F/R	T7 For	Annotated	LPP	Q93052	Nucleus
111.1	2100	Gal4AD F/R	T7 For	Annotated	APBB2	Q92870	Nucleus
114.1	2800	Gal4AD F/R	T7 For	Annotated	HTT	P42858	Nucleus
123.1	1000	Gal4AD F/R	T7 For	Annotated	WDR61	Q9GZS3	Cytosol
4.1	1500	Gal4AD F/R	T7 For	Annotated	PPP1R2	P41236	Cytosol
11.1	1000	Gal4AD F/R	T7 For	Annotated	GUK1	Q16774	Cytosol
34.1	1700	Gal4AD F/R	T7 For	Annotated	ISPD	A4D126	Cytosol
41.1	1000	Gal4AD F/R	T7 For	Annotated	SMG6	Q86US8	Cytosol
47.1	1400	Gal4AD F/R	T7 For	Annotated	PROCA1	Q8NCQ7	Cytosol

49.1	1200	Gal4AD F/R	T7 For	Annotated	ALOX12	P18054	Cytosol/membrane
53.2	1200/900/400	Gal4AD F/R	T7 For	Annotated	ASAP1	Q9ULH1	Cytosol
59.1	1000	Gal4AD F/R	T7 For	Annotated	CCDC88C	Q9P219	Cytosol
66.1	1800	Gal4AD F/R	T7 For	Annotated	IQCG	Q9H095	Cytosol
77.2	1100	Gal4AD F/R	T7 For	Annotated	CAPN3	P20807	Cytosol
86.2	1000	Gal4AD F/R	T7 For	Annotated	CCDC88C	Q9P219	Cytosol
127.1	1700	Gal4AD F/R	T7 For	Annotated	PDPK2P	Q6A1A2	Mitocondrion
5.1	1100	Gal4AD F/R	T7 For	Annotated	QRSL1	Q9H0R6	Mitocondrion
35.1	1200	Gal4AD F/R	T7 For	Annotated	PACS2	Q86VP3	Mitocondrion
37.2	1300	Gal4AD F/R	T7 For	Annotated	YME1-like 1	Q96TA2	Mitocondrion
95.3	1500/1300	Gal4AD F/R	T7 For	Annotated	CPT2	P23786	Mitocondrion
118.3	2800/1600/1000	Gal4AD F/R	T7 For	Annotated	PNPLA4	P41247	Mitocondrion
120.1	900	Gal4AD F/R	T7 For	Annotated	PNPLA4	P41247	Membrane
16.2	1100	Gal4AD F/R	T7 For	Annotated	TMEM42	Q69YG0	Membrane
24.1	1700	Gal4AD F/R	T7 For	Annotated	SCARB1	Q8WTV0	Membrane
36.1	2900	Gal4AD F/R	T7 For	Annotated	HCN4	Q9Y3Q4	Membrane
79.2	2200	Gal4AD F/R	T7 For	Annotated	ZMYND10	O75800	Membrane
116.2	1700	Gal4AD F/R	T7 For	Annotated	SLC9A3	P48764	Membrane
40.1	900	Gal4AD F/R	T7 For	Annotated	HEXB	P07686	Lysosome
124.1	1000	Gal4AD F/R	T7 For	Annotated	CD63	P08962	Lysosome
71.2	1500	Gal4AD F/R	T7 For	Annotated	SCAMP2	O15127	Endosome + Golgi
102.1	1600	Gal4AD F/R	T7 For	Annotated	SCAMP2	O15127	Endosome + Golgi
7.1	1200	Gal4AD F/R	T7 For	Annotated	ADGRG1	Q9Y653	Secreted
15.1	1500	Gal4AD F/R	T7 For	Annotated	ADA2	Q9NZK5	Secreted
20.1	1200	Gal4AD F/R	T7 For	Annotated	MYH11	P35749	Melanosome
64.3	1300/1200	Gal4AD F/R	T7 For	Annotated	C4orf54	D6RIA3	Unknown
74.1	3000	Gal4AD F/R	T7 For	Annotated	C18orf54	J3KS56	Unknown
3.1	2000	Gal4AD F/R	T7 For	Unannonated	Homo sapiens protein phosphatase 1, regulatory (inhibitor) subunit 2, mRNA		
12.1	2000	Gal4AD F/R	T7 For	Unannonated	AMZ2P1		
17.3	1500/2500	Gal4AD F/R	T7 For	Unannonated	BAC clone RP11-784K9		
21.1	1700	Gal4AD F/R	T7 For	Unannonated	BAC clone RP11-175A7		
29.2	1200	Gal4AD F/R	T7 For	Unannonated	Homo sapiens MORC2 antisense RNA 1		
31.2	3000	Gal4AD F/R	T7 For	Unannonated	POM121L4P		
38.1	800	Gal4AD F/R	T7 For	Unannonated	BAC clone RP11-155B1		
39.1	1100	Gal4AD F/R	T7 For	Unannonated	Non-coding protein (IPW)		

42.2	1400	Gal4AD F/R	T7 For	Unannonated	Clone RP11-1049H7		
48.2	1000	Gal4AD F/R	T7 For	Unannonated	Homo sapiens mRNA; cDNA DKFZp781F21103		
50.2	1100	Gal4AD F/R	T7 For	Unannonated	Non-coding protein (IPW)		
56.1	1000	Gal4AD F/R	T7 For	Unannonated	Clone RP1-9E21 on chromosome 1q24-25		
57.1	1000	Gal4AD F/R	T7 For	Unannonated	Hexosaminidase B		
63.1	1900/1400	Gal4AD F/R	T7 For	Unannonated	Non-coding protein (IPW)		
68.1	1100	Gal4AD F/R	T7 For	Unannonated	Non-coding protein (IPW)		
72.1	1900	Gal4AD F/R	T7 For	Unannonated	Cosmid F9933, complete sequence		
80.2	2200	Gal4AD F/R	T7 For	Unannonated	PREDICTED: Homo sapiens uncharacterized LOC105376156 (LOC105376156), transcript variant X2, ncRNA		
81.1	1200	Gal4AD F/R	T7 For	Unannonated	P1 clone 702A10 (LBNL H56)		
82.1	1200	Gal4AD F/R	T7 For	Unannonated	PREDICTED: Homo sapiens chromosome 2 open reading frame 40 (C2orf40), transcript variant X3, mRNA		
83.1	1100	Gal4AD F/R	T7 For	Unannonated	Clone CTD-2563K22		
85.2	2900	Gal4AD F/R	T7 For	Unannonated	PREDICTED: Homo sapiens uncharacterized LOC112268371 (LOC112268371), ncRNA		
94.2	3100	Gal4AD F/R	T7 For	Unannonated	Homo sapiens clone 3684 breakpoint junction genomic sequence		
97.1	2000	Gal4AD F/R	T7 For	Unannonated	RUN and FYVE domain containing 2		
99.1	1500	Gal4AD F/R	T7 For	Unannonated	Clone CH17-7N6, complete sequence		
105.1	1300	Gal4AD F/R	T7 For	Unannonated	H.sapiens mRNA for cysteine-rich secretory protein-2/type I		
109.1	900	Gal4AD F/R	T7 For	Unannonated	BAC clone RP11-155B1		
113.2	2100	Gal4AD F/R	T7 For	Unannonated	Homo sapiens genomic DNA, chromosome 22q11.2, Cat Eye Syndrome region, clone:c91G6		
117.2	1800	Gal4AD F/R	T7 For	Unannonated	Homo sapiens chromosome 19 clone RP11-157B13, complete sequence		
121.1	1500	Gal4AD F/R	T7 For	Unannonated	CTCL tumor antigen HD-CL-04 mRNA, complete cds		
126.1	1700	Gal4AD F/R	T7 For	Unannonated	FOSMID clone ABC8-41107200L10		
130.1	800	Gal4AD F/R	T7 For	Unannonated	MAF1 homolog (S. cerevisiae)		
133.1	1300	Gal4AD F/R	T7 For	Unannonated	H.sapiens mRNA for cysteine-rich secretory protein-2/type I		
134.2	1500/1300	Gal4AD F/R	T7 For	Unannonated	Homo sapiens mRNA for cysteine and glycine-rich protein 1 variant protein		
8.1	2000	Gal4AD F/R	T7 For	Not-classified			

17.4	1500/2500	Gal4AD F/R	T7 For	Not-classified			
25.1	3000/500	Gal4AD F/R	T7 For	Not-classified			
30.2	3000	Gal4AD F/R	T7 For	Not-classified			
44.2	400	Gal4AD F/R	T7 For	Not-classified			
46.1	400	Gal4AD F/R	T7 For	Not-classified			
92.2	1200	Gal4AD F/R	T7 For	Not-classified			
98.1	1700	Gal4AD F/R	T7 For	Not-classified			
103.1	1400	Gal4AD F/R	T7 For	Not-classified			
107.2	1300	Gal4AD F/R	T7 For	Not-classified			
118.4	2800/1600/1000	Gal4AD F/R	T7 For	Not-classified			
125.2	2300	Gal4AD F/R	T7 For	Not-classified			
131.1	1600/1400/400	Gal4AD F/R	T7 For	Not-classified			

Suppl. Table 3. List of cDNA fragments interacting with VLP protein.