Università degli Studi di Padova
Dipartimento di Scienze Biomediche

CORSO DI DOTTORATO DI RICERCA IN SCIENZE BIOMEDICHE
32° CICLO

INSIGHT INTO NON-CANONICAL pVHL FUNCTIONS

Coordinatore: Ch.mo Prof. Paolo Bernardi

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

Dottoranda: Antonella Falconieri

Index

Abbreviations

2-OG	2-oxoglutarate
3-AT	3-amino 1,2,4-triazol
AD	activation domain
Ade	adenine
ADRB2	beta 2 adrenergic receptor
Akt	protein kinase B (or PKB)
AMP	adenosine monophosphate
Amp	Ampicillin
APC	Adenomatous Polyposis Coli
AR	androgen receptor
ARE	androgen responsive element
ARNT	aryl hydrocarbon receptor nuclear transport
ATM	serine/threonine kinase
ATP	adenosine triphosphate
AUR1-C	Aureobasidin A (Aba) resistance gene
BCA	protein quantification based on bicinchoninic acid
bHLH	basic helix-loop-helix
BLAST	basic local alignment search tool
BRCA1	breast cancer type 1 susceptibility protein
BSA	bovine serum albumin
C-	negative control (or Ø)
C+	positive control
CAIX	carbonic anhydrase 9
CBP	CREB binding protein
ccRCC	clear cell renal cell carcinoma (or RCC)
CDK	cyclin-dependent kinase

cDNA	complementary DNA
CFU	colony forming unit
CHAPS	3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate hydrate
CHX	cycloheximide
CKI	CDKs inhibitors
CMV	cytomegalovirus
c-Myc	Myc proto-oncogene protein
CNS	Central nervous system
CO_2	carbonic anhydride
CODD	oxygen-dependent domain C-terminal
Co-ip	co-immunoprecipitation
COSMIC	Catalogue of Somatic Mutations in Cancer
CT	C-terminal
C-TAD	transactivation domain C-terminal
Cul2	cullin 2
DAPI	nuclear marker 4,6-diamidino-2-phenylindole
DBD	DNA binding domain
ddH_2O	bi-distilled water
DDO	double dropout
DHT	dihydrotestosterone
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
DTT	1,4 dithiothreitol
ECM	extracellular matrix
ELM	eukaryotic linear motifs predictor
Elo B/C	elongin B/C (or EGLN)
ELSTs	endolymphatic sac tumors
EPOR	erythropoietin receptor

ERK	extracellular signal-regulated kinase
EtOH	ethanol
FBS	fetal bovine serum
Fe^{2+}	iron
FELLS	Fast estimator of latent local structure
FGF	fibroblast growth factor
G	gap phase
GAL4	transcription factor Gal4
Gal4-AD	activation domain of Gal4
Gal4-BD	binding domain of Gal4
GC-rich	guanine-cytosine rich
GFP	green fluorescent protein
GLUT (1-3)	glucose transporter (1-3)
HB	hemangioblastoma
HEK293T	human embryonic kidney cells 293
HeLa	Henrietta Lacks (donor of cell)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF (α/β)	Hypoxia-inducible factor, α or β subunit
His	histidine
Hn	helix number
HRE	HIF responsive element
HSPs	heat shock proteins
IDRs	intrinsically disordered regions
IGF2	insulin-like growth factor 2
INK4	inhibitors of CDK4
JAK	janus kinase
Kan	kanamycin
LB	Luria-Bertani medium

LBD	ligand-binding domain
LDS	NuPAGE LSD Sample Buffer Invitrogen
Leu	leucine
LiAc	litium acetate
LOF	loss of function
M	mitosis
MAFFT	multiple alignment using fast Fourier transform
MAPK	mitogen-activated protein kinase
MCS	multiple cloning site
MDM2	mouse double minute 2 homolog
MdmX	double minute 4 protein (or MDM4)
MEL1	alpha-galactosidase gene 1
MMP (1-13)	matrix metalloprotease (1-13)
MobiDB	database of protein disorder and mobility annotations
MYBBP1A	Myb-binding protein 1A
NaCl	sodium chloride
NES	nuclear export signal
NLS	nuclear localization signal
NODD	oxygen-dependent domain N-terminal
NT	N-terminal
o/n	over night
O_2	oxygen
OD	optical density
ODD	oxygen-dependent domain
OMA	Orthologous Matrix
OMIM	Online Mendelian Inheritance in Man
OOF	out-of-frame
p14ARF	tumor suppressor protein 14

p300	histone acetyltransferase p300
p53	tumor suppressor p53
PAS	Per-Arnt-SIM
PBS	phosphate buffer saline
PCC	pheochromocytoma
PCDC5	programmed cell death 5 protein
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
PEI	polyethylenimine
PGL	paraganglioma
PHDs	prolyl hydroxylase domain enzymes
PNS	post-nuclear surnatant
PolII	polymerase 2
PRMT3	protein arginine methyltransferase 3
Pro	proline
PROTAC	proteolysis targeting chimera
PTEN	phosphatase and TENsin homolog
PTM	post translational modification
QDO	quadruple dropout
RBX1	Ring-box protein 1
RIPA	radio-immunoprecipitation assay buffer
RNA	ribonucleic acid
ROS	reactive oxygen species
RT	room temperature
RT-PCR	real-time PCR
S	synthesis phase
SBMA	spinal bulbar muscular atrophy

SCF	complex (Skp, Cullin, F-box containing complex)
SD	selective medium
SDS	sodium dodecyl sulfate
Ser	serine
SLIMs	short linear motifs
Spry2	protein sprout homolog 2
SRP	substrate recognition particle
STAT	signal transducers and activators of transcription
STRING	search tool for the retrieval of interacting genes/proteins
SV40	simian virus 40
TAD	transactivation domain
TATA	Goldberg-Hogness box, thymine and adenine rich
TBST	Tris-buffered saline tween
TCA	trichloroacetic acid
TD-NEM	transcriptional-dependent nuclear export motif
TF	transcription factor
TGFB1	transforming growth factor beta 1
TGFα	transforming growth factor alpha
TGS	gatekeepers
Thr	threonine
Tip60	histone acetyl transferase
Trp	tryptophan
UPS	ubiquitin-proteasome system
USP7	ubiquitin-specific processing protease 7
UTR	untranslated region
UV	ultraviolet
VCB	complex (pVHL-elonginC-elonginB)
VCB-Cul2	complex (pVHL-elonginC-elonginB) + Cullin2

VEGF	vascular endothelial growth factor
VEGFR	vascular endothelial growth factor receptor
VHL	von Hippel-Lindau
VHL172	pVHL172 or von Hippel-Lindau 172, isoform 172
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
VHLα	von Hippel-Lindau α (isoform)
VLP	von Hippel-Lindau like protein (paralog)
w/o	without
WB	Western blot
X-α-Gal	X-α-galactose
Y2H	yeast two-hybrid
Y2HGold	Gold yeast two-hybrid strain
Y3H	yeast three-hybrid
YPD	rich medium
ZnF	Zinc finger
α	alpha
β	beta

Summary

One of the most frequent key factors driving tumorigenesis is the functional impairment of tumor suppressor proteins. Mutations of the von Hippel-Lindau tumor suppressor protein (pVHL) are causative of the von-Hippel Lindau syndrome, a hereditary predisposition to develop cancers. The protein exists in two main isoforms: the full-length pVHL30 and a shorter pVHL19 lacking the N-terminal portion. Both these proteins exert their function as oncosuppressor by regulating the hypoxia-inducible factor 1-alpha (HIF-1α) stability. Besides its well-characterized role in HIF-1α regulation, pVHL is proposed as a multifunctional adaptor protein involved in multiple different cellular processes. These additional functions are collectively referred as non-canonical pVHL functions. Aim of this thesis is the further characterization of these HIF-independent functions. A multidisciplinary strategy combining experimental and bioinformatics approaches was used to address three different biological questions: the relationship between pVHL and cell-cycle regulation, the role of pVHL30 in testis-tissue and the existence of a functional link between pVHL30 and the androgen receptor (AR). My work identified novel protein-protein interactions that further link pVHL to cell-cycle control. In particular, I identified a new direct interaction between pVHL and the Cip/Kip protein family (i.e. p21,p27,p57), mediated by the pVHL30 β-domain and a COOD-like motif located into the linear CDKN1s N-terminal domain. In parallel, my efforts in dissecting pVHL also demonstrated pVHL30 to bind the human Mouse double minute 2 homolog (MDM2), an E3-ubiquitin ligase involved in the degradation of the Tumor protein p53 (p53). This interaction was found to be pVHL30-specific and mediated by its N-terminal tail, with the shorter pVHL19 lacking this binding property. The molecular dissection distinguished the pVHL30 β-domain and the N-terminal tail as the regions interacting with a long disordered portion located at the MDM2 C-terminus. On note, the association was found to be mediated by two specific intrinsically disordered portions confirming the pVHL ability to recognize and bind linear motifs. Then using a Y2H genome-wide screening approach I investigated the unknown pVHL30 interactome in testis-tissue. This analysis identified 260 positive clones further classified into 2 main groups addressing 61 human proteins and 118 non coding peptides, respectively. Among the first group I found 6 already known pVHL30 interactors (e.g. Elongin C) that confirmed the approach reliability, whereas the remaining 55 novel interactors were never associated to pVHL30 so far. These testis-specific interactors include proteins involved in relevant cellular pathways such as cell-cycle regulation, DNA

damage repair, apoptosis and cytoskeleton regulation. On the other hand, a preliminary protein-protein interacting network generated using these data showed limited internal connections at the pathway level, suggesting a broader involvement of pVHL30 in other pathways not yet identified. Finally considering the relevance of AR in testis-tissue, I investigated the functional relationship between pVHL30 and this protein. Multiple cellular biology techniques were employed to identify a physical interaction between pVHL30 and AR wild-type (AR24Q) or poly-Q expanded (AR65Q). The expansion of the AR poly-Q stretch is linked to protein dysfunction that leads to pathological conditions. My findings demonstrated pVHL30 to inhibit the AR transcriptional activity as well as influencing AR protein stability by accelerating its turnover. In summary, findings reported in this thesis identify 4 novel general interactors that point to pVHL30 involvement in cell-cycle regulation, while 55 new interactors were found to be testis-specific. Moreover, important pieces of evidence shedding light on the functional relationship between pVHL and AR are also discussed. Collectively taken, all of these evidence suggest pVHL30 to have isoform-specific functions as well as putative testis-specific and sex-linked functions.

Riassunto

Uno degli eventi alla base del processo di tumorigenesi è l'alterazione funzionale delle proteine che funzionano da oncosoppressori. Mutazioni nell'oncosoppressore von-Hippel Lindau (pVHL) causano la sindrome di Von-Hippel Lindau, una predisposizione ereditaria allo sviluppo di cancro in differenti tessuti. Esistono due isoforme principali della proteina: un'isoforma completa chiamata pVHL30 e un'isoforma più corta mancante della regione N-terminale nota come pVHL19. Entrambe le isoforme svolgono la loro azione come oncosoppressori principalmente regolando la stabilità del fattore regolato da ipossia HIF-1α. Oltre a regolare HIF-1α, pVHL agisce anche da adattatore multifunzionale coinvolto in molteplici processi cellulari diversi. Queste funzioni aggiuntive sono definite come funzioni di pVHL non-canoniche. Lo scopo di questa tesi è la caratterizzazione ulteriore di queste funzioni indipendenti da HIF. Una strategia multidisciplinare che combina approcci sperimentali e bioinformatici è stata utilizzata per analizzare tre differenti aspetti: la relazione tra pVHL e la regolazione del ciclo cellulare, il ruolo di pVHL30 nel tessuto di testicolo e il link funzionale esistente tra pVHL30 e il recettore degli androgeni (AR). Il mio lavoro ha identificato nuove interazioni proteina-proteina che collegano ulteriormente pVHL al controllo del ciclo cellulare. Nel dettaglio, è stata identificata una nuova interazione tra pVHL30 e la famiglia delle proteine Cip/Kip (ovvero p21,p27 e p57) mediata dal dominio β di VHL30 e il motivo *COOD-like* situato nel dominio lineare N-terminale degli inibitori delle chinasi ciclina-dipendenti 1 (CDKN1s). Parallelamente, ulteriori studi per la caratterizzazione di pVHL hanno mostrato che pVHL30 è in grado di legare MDM2, una E3 ubiquitina ligasi coinvolta nella degradazione del soppressore tumorale p53. I dati hanno evidenziato che questa interazione è mediata dalla regione N-terminale di pVHL che è presente in pVHL30 ma non in pVHL19 che infatti non lega MDM2. La dissezione molecolare ha identificato la coda N-terminale e il dominio β come le regioni interagenti con una lunga porzione disordinata presente nella regione C-terminale di MDM2. Questa interazione coinvolge due regioni intrinsecamente disordinate, a conferma dell'abilità di pVHL di riconoscere e legare motivi lineari tipici di queste regioni. Durante il secondo anno, uno screening ad ampio raggio basato sul doppio ibrido è stato utilizzato per caratterizzare l'interattoma di pVHL30 nel tessuto di testicolo. Questa analisi ha identificato 260 cloni positivi poi ulteriormente classificati in due gruppi corrispondenti, rispettivamente, a 61 proteine umane e 118 peptidi non codificanti. Tra le proteine, sono stati trovati 6 interattori di pVHL30 già noti (ad esempio l'elonghina C) a

conferma della validità dell'approccio usato, e altri 55 nuovi interattori mai associati prima d'ora a pVHL30. Questi interattori tessuto-specifici comprendono proteine coinvolte in pathways cellulari rilevanti come la regolazione del ciclo cellulare, la riparazione del danno al DNA, l'apoptosi e la regolazione del citoscheletro. Dall'altro laro, però, la rete preliminare di interattori generata usando questi dati ha mostrato un numero limitato di connessioni a livello di pathway suggerendo il coinvolgimento di pVHL30 anche in altri pathways non ancora caratterizzati. Infine, considerando l'importanza del recettore degli androgeni nel tessuto di testicolo, ho studiato la relazione funzionale tra pVHL30 e questa proteina. Diverse tecniche di biologia cellulare sono state utilizzate per identificare l'interazione fisica tra pVHL30 e la forma wild-type del recettore (AR24Q) o la forma mutata (AR65Q). La forma mutata è caratterizzata da un'espansione della regione poli-Q. Questa alterazione genera una disfunzione della proteina che porta a condizioni patologiche. I miei risultati hanno dimostrato che pVHL30 inibisce l'attività trascrizionale di AR e influenza la stabilità della proteina accelerandone il turnover. Per riassumere, i risultati descritti in questa tesi identificano 4 nuovi interattori che collegano pVHL30 alla regolazione del ciclo cellulare e 55 nuovi interattori specifici per il tessuto di testicolo. Inoltre, vengono descritte importanti evidenze che fanno luce sulla relazione funzionale tra pVHL e AR. Considerati nell'insieme i dati suggeriscono che pVHL30 è caratterizzato da funzioni isoforma specifiche così come da putative funzioni specifiche per il testicolo e funzioni legate al sesso.

Introduction

1 – Introduction

1.1 The hallmarks of cancer

Cancer is the second cause of death in the world[1]. It can be considered as a complex disease arising from mutational events in cancer-susceptibility genes[2] causing unregulated cells division and consequent cells spreading into the body. Cancer development arises to genetic alterations on genes classified into three different classes: gatekeepers[3], caretakers[3] and landscapers[4]. Gatekeepers include oncogenes and tumor suppressors (TGSs) directly involved in cell growth control. Alterations in these genes leads to uncontrolled cell proliferation. On the contrary caretakers are indirectly involved in proliferation control throughout maintenance of genome integrity. In this case alterations promote genetic instability. Finally, landscapers generate alterations in stroma environment that contributes to neoplastic cell transformation. Thus, cancer is essentially a genetic disease, the pathogenesis of which is influenced by both hereditary and environmental factors[5]. In 2000s, in order to rationalize the complexities of neoplastic disease, Hanhan and Weinberg proposed six essential alterations in cell physiology shared by cancer cells[6]. They correspond to self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis[6]. Each physiologic change is acquired by normal cells during the multistep process that leads them to become tumorigenic. This list of hallmarks, however did not consider the contribution of "tumor microenvironment" to tumorigenesis. For this reason, in 2011 it was further extended by adding two new hallmarks, i.e. reprogramming of energy metabolism and evading immune destruction, and two enabling traits, i.e. genome instability and mutation, and tumor-promoting inflammation[7](figure 1).

Cancer can be distinguished in sporadic and hereditary. The 80% of all cancers are thought to be sporadic in which mutation of tumor-associated genes occur by chance. The exact causes for sporadic cancers are largely unknown, whereas they may include environmental and lifestyle risk factors (e.g. smoking, UV exposure) as well as being the result of natural process of aging. On the contrary, in about 5-15% of cancers, the underlying cause is linked to an inherited gene mutation that increases the risk to develop certain types of cancer during the lifetime. Although the itself mutation is not cancerous *per se*, it increases the chance for other random mutations to occur and accumulate in the cells, leading to the development of cancer. Studies in cancer genetics identified several genes associated with

hereditary cancer syndromes like BRCA1, PTEN, APC and p53[8]. Germline mutations in p53 are paired with various type of cancers associated to the Li-Fraumeni Syndrome[9]. Similarly, mutations of the von Hippel-Lindau tumor suppressor protein (pVHL) cause the von Hippel-Lindau syndrome, a hereditary predisposition to develop cancer due to alteration of oxygen sensing mechanism and sustained angiogenesis.

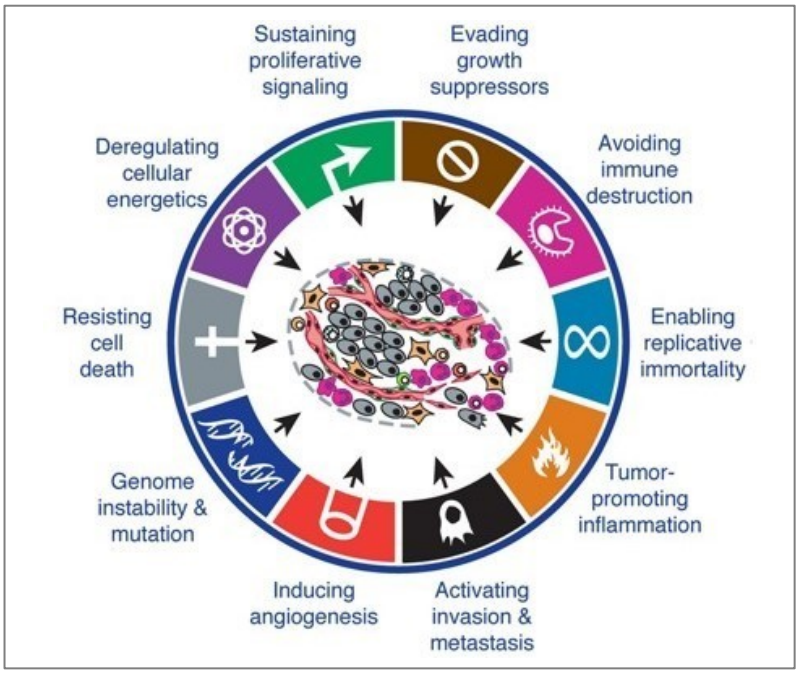


Figure 1 Cancer hallmarks. The original six cancer hallmarks are updated including the capability to modify cellular metabolism, evade immunological destruction, genomic instability and inflammation responses. Modified from Hanahan and Weinberg, Cell 144, March 4, 2011.

1.2 The von Hippel-Lindau syndrome

The von Hippel-Lindau (VHL) disease (OMIM number 193300) is a rare autosomal dominantly inherited genetic disorder[10]. The disease was firstly described separately by von Hippel in 1911[11] and by Lindau in 1926[12]. The first, Eugen von Hippel, a German ophthalmologist, described same cases of retinal hemangioblastomas. Instead, 15 years later, Arvid Lindau, a Swedish pathologist recognized the association between retinal lesions and cerebellar hemangioblastomas in addition to finding in other organs, as part of a familiar syndrome. The term von Hippel–Lindau disease was first used in 1936 and has been in common use since the 1970s. This cancer disorder have an incidence of 1:36000 live births and it has a penetrance of > 90% by age 65 years with a mean age at tumor diagnosis of 26 years[13]. As summarized in picture 2, it is a multiple-neoplasia syndrome

associated with the development of multiple vascular tumors affecting different tissues like eyes, brain, kidneys, adrenal glands and paraganglia, pancreas, endolymphatic sac, epididymis and broad ligaments. Therefore, pVHL-associated tumors are several, such as retinal and CNS heamangioblastoma[14] (HBs), clear cell renal cell carcinoma (RCC)[15], pheochromocytoma[16] (PCC) and paraganglioma[17] (PGL), pancreatic cystoadenomas[18], endolympathic sac tumors (ELSTs)[19], epididymal cysts[20] and broad ligament cystoadenoma[21]. In particular, RCC occurs in about 70% of VHL affected individuals.

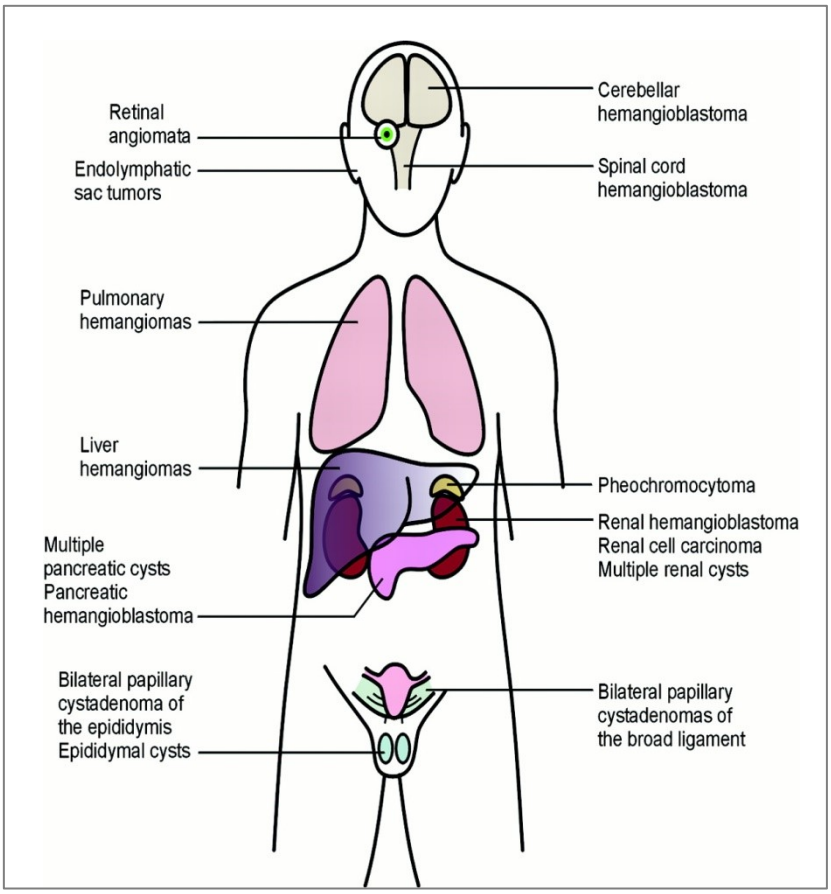


Figure 2 VHL occurrences in patients. VHL patients are characterized by multiple-neoplasia lesions spread into different organs throughout the body. From Friedrich Human Molecular Genetics Vol. 10, No. 7763–767 (2011).

Despite the disease heterogeneity, taken into account the propensity to develop PCC the syndrome can be classified into VHL type 1 or VHL type 2[22]. Type 1 VHL includes patients affected by HBs and RCC, without pheochromocytoma. Upon pheocromocytoma insurgence the classification become type 2 subdivided in 2A, 2B and 2C according to the risk of RCC development. Type 2A is associated with HBs and PCC whereas RCC is extremely rare. On the contrary type 2B is characterized by the occurrence of RCC paired

with HBs and/or PCC. Finally, type 2C is connected to the occurrence of PCC only[22]. Although routinely used for the initial assessment of VHL patients, this rigid classification can generate ambiguous assignments due to the absence of an univocal genotype-phenotype correlation. Indeed, clinical VHL manifestations are frequently variable, with different phenotypes associated to a same mutation in different families or even in the same family[23]. VHL syndrome arise from pathogenic inactivation of homonymous VHL gene, mapped to 3p25-26 in 1993[24]. Individuals with the hereditary form inherit a single mutant VHL allele, and tumor development occurs when the second wild-type copy is spontaneously lost or inactivated. This second hit occurs through point mutation, deletion or promoter hypermethylation[25]. Data collected from literature linked VHL disease to more than 1500 germline and somatic mutations[13]. These mutations spanning throughout the VHL gene with 43,2% in exon 1, 17% in exon 2 and 39,8% in exon 3[13]. In particular, the spectrum of mutation types is formed by 52% missense, 13% frameshift, 11% nonsense, 6% in-frame deletions/insertions, 11% large/complete deletions, and 7% splice mutations[13]. Genotype-phenotype correlation studies have associated VHL Type 1 mainly with protein truncating mutations and deletions, whereas VHL Type 2 with missense variant[26].

1.3 The VHL gene and protein structure

The VHL tumor suppressor gene is a 10 kb region located in the short arm of chromosome 3 (3p25-26)[24]. As reported in picture 3, the gene consists of 3 exons: the first 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). The VHL gene is highly conserved throughout the evolution among human, primate and rodent. A pVHL homologue was also found in *Caenorhabditis elegans* and *Drosophila*[27]. Evolutionary studies shown a good conservation also for pVHL promoter among primates. The VHL promoter is GC-rich, lacking either TATA or CCAAT motifs. Transcription initiates around a putative Sp1-binding site approximately 60 nucleotides upstream codon 1[28]. VHL encodes two protein products: a 30 kDa full-length protein named pVHL30 and a shorter isoform named pVHL19 which is generated by an alternative translation initiation site at methionine 54[29]. Structurally, the full-length protein can be subdivided in three main regions: the N-terminal portion (aa 1-53), the central β-domain (aa 54 to 157) and a C- terminal α-domain (aa 158 to 213). In particular, the N-terminal tail, present only in pVHL30, is disordered and characterized by

acidic repeats G-[PAVG]-E-E-[DAYSLE]. The number of repeats is 8 in human and higher primates and progressively decrease to one in rodent[27]. The β-domain is composed by seven-stranded β sandwich (aa 63 -154) and an α helix (H4; aa 193-200) that packs against one of the β sheets through hydrophobic interactions[30].

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")[30]. The α and β domains are connected by two short polypeptide linkers (residues 154 to 156 and 189 to 194)[31]. Data collected in VHL database[32], indicate pVHL to interact with more than 500 different partners involved in multiple cellular processes. Considering the protein-protein interaction properties the VHL surface can be subdivided in three specific areas defined interface A, B and C[33]. Interface A contains the ElonginC (EloC) and Culin2 (Cul2) binding sites while interface B includes HIF-1α binding site. Interestingly, interface C contains the nuclear export signal (NES) and appears to influence pVHL localization into nucleus and cytoplasm[33].

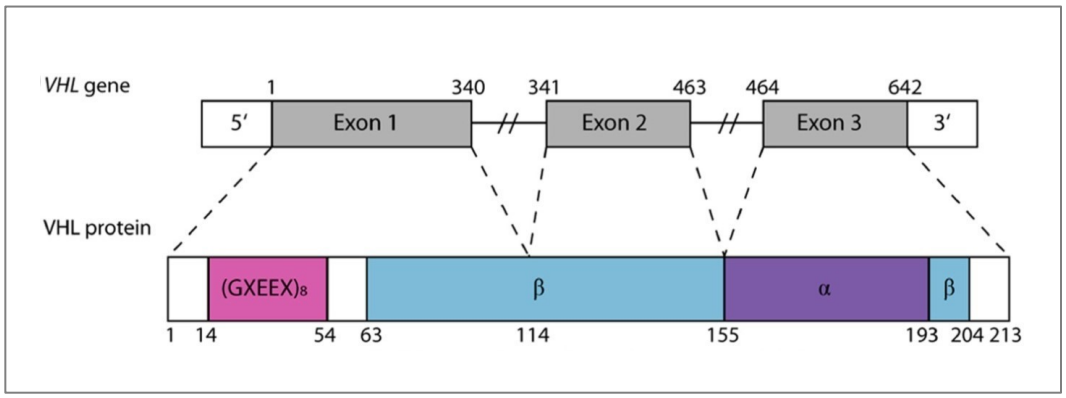


Figure 3 von Hippel-Lindau gene and protein structure. The VHL gene composed by 3 exons colored in grey. Numbers correspond to nucleotides positions. Bottom, a schematization of the protein structure characterized by N-terminus (pink), β-domain (light blue) and α-domain (purple). Numbering correspond to amino acids positions.

1.3.1 VHL protein isoforms

As already described the VHL gene transcription generates two major alternatively-spliced isoforms namely pVHL30 and pVHL19. The pVHL19 is shorter and lacks the first 53 N-terminal residues[29]. The expression of the two isoforms is ubiquitous with high levels in the urogenital system, brain, spinal cord, sensory ganglia, eyes and bronchial epithelium[34]. Both isoforms overlap in displaying tumor suppressor abilities[35], inhibiting cancer development when a wild-type copy is reintroduced in ccRCC cells[36]. They are biologically active and form an ubiquitin E3 ligase complex known as VCB, binding

elongins B and C and cullin 2[37,38]. Despite the conserved function, pVHL30 and pVHL19 differ in their subcellular localization. The full-length isoform is mainly cytosolic whereas pVHL19 is equally distributed among cytosol and nucleus. Furthermore, pVHL19 does not associate with cell membranes as confirmed by its failure in fibronectin coimmunoprecipitation[29]. Recently, data suggested a possible functional isoform specialization underling an important role of the pVHL N-terminal tail as further interaction interface[39]. Indeed pVHL30, unlike pVHL19, was shown to associate with p14ARF, a tumor suppressor linked to p53 pathway regulation[39]. These data suggest a connection between pVHL30 and the tumor suppressor p53. Conversely, pVHL19 isoform shows, if compared to pVHL30, an increased tendency to form homodimers[39] further reinforcing the hypothesis of VHL isoform-specific binding specialization.

In this thesis the attention is focused on these two canonical isoforms. However, it's important to mention that two other pVHL isoforms are also described. The first, named pVHL172 is produced by alternative splicing of exon 2[40]. It is expressed at high levels during human embryogenesis (i.e. 8-10 gestational weeks) in kidney, brain, spinal cord, eyes, testis and lung[41]. The absence of part of the β domain (aa 114-154) modifies the number of beta sheets in its structure with structural and functional consequences[40]. In particular, pVHL172 does not act as tumor suppressor. Instead, it plays a role in renal carcinoma in up-regulating a subset of pro-tumorigenic genes i.e. TGFB1, MMP1 and MMP13[42]. The last pVHL isoform, named pVHLα is the result of an alternative translation site located upstream and in-frame with the VHL30 canonical ATG start codon[43]. This isoform maintains the tumor suppressor activity inhibiting the Warburg effect[43].

1.4 VHL main function: HIF-1α and the hypoxia-response system

pVHL is the substrate recognition particle (SRP)[38] of the E3-ubiquitin ligase complex (VCB) involved in the degradation of hypoxia-inducible factor subunit 1 alpha (HIF-1α), a process finely regulated by oxygen availability[44]. In normoxic conditions, HIF-1α is hydroxylated, recognized by VHL and poly-ubiquitinated for proteosomal degradation. Instead, in hypoxic conditions, HIF-1α is stabilized and translocated to the nucleus where dimerize with HIF-1β and activates the transcription of genes involved in hypoxic response. pVHL loss or mutations disrupting this oxygen-sensing regulation system are driver of VHL syndrome insurgence[45]. This oxygen-regulated mechanism is highly

integrated and implicates different proteins, such as HIFs family, PHDs and the VCB complex components (Figure 4).

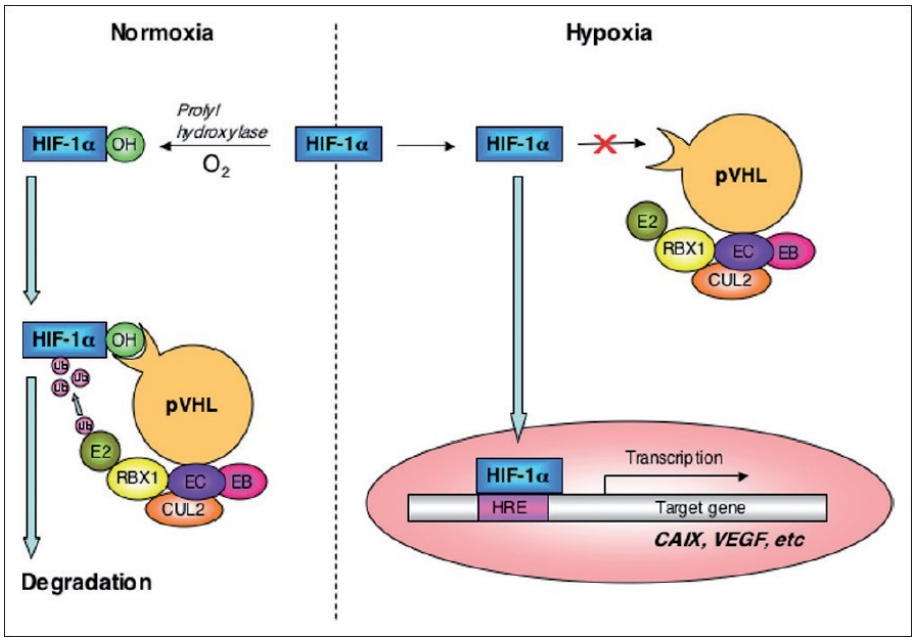


Figure 4 Oxygen-dependent regulation of HIF-1α pathway. Under normoxia, HIF-1α is hydroxylated by a prolyl hydroxylase and bound to pVHL, which forms a complex with elongin B (EB), elongin C (EC), cullin2 (CUL2), RBX1 and a ubiquitin-conjugating enzyme (E2). Then HIF-1α is ubiquitinated and degraded. In contrast, under hypoxia, non-hydroxylated HIF-1α enters to the nucleus (dimerize with HIF-1β, not shown) and activates target genes, like CAIX, VEGF, etc. From Indian Journal of Urology.

1.4.1 HIFs family protein and hypoxia

The maintenance of oxygen levels in aerobic organisms is important and tightly regulated. Indeed, an increase (hyperoxia) or decrease (hypoxia) in cellular oxygen levels leads to the activation of cellular signaling pathways that allow cells and organisms to maintain oxygen homeostasis[46,47]. Oxygen (O_2) is required for aerobic metabolism to maintain intracellular bioenergetics and serves as an electron acceptor in many organic and inorganic reactions. Both delivery and consumption of O_2 are precisely regulated. Dysregulation of these systems is a common feature of several human pathological conditions[47]. Data indicated that oxygen level in hypoxic tumor is lower when compared to healthy tissues, reaching 1-2% on average. However, its specific value varies with the initial oxygenation of the tissue, size, stage and severity of the specific tumor[48]. In healthy conditions, normal oxygen levels ranges between 9,5% to 4,6%. Hypoxia can be considered a mechanism with two different "sides". During fetal development it represents a positive stimulus, as the gradient in oxygen tension promotes a correct liver zonation[49]. On the other hand, hypoxia increases

HIF transcriptional activity, which in same case appears to play a critical role in the development of invasive and metastatic properties associated to cancer lethality [50]. During period of reduced oxygen supply, the most profound changes in genes expression are mediated by transcription factors known as hypoxia-inducible factors (HIFs)[51]. In human, there are three different transcriptional complexes: HIF-1, HIF-2 and HIF-3 composed by the constitutive HIF-1β (also known as ARNT – aryl hydrocarbon receptor nuclear transport) associated respectively with the oxygen-dependent subunits HIF-1α, HIF-2α and HIF-3α[52]. Of the three α-subunits, HIF-1α and HIF-2α are the best studied. All the three HIF subunits share high similarity in the basic helix-loop-helix (bHLH) and Per-Arnt-SIM (PAS) domains mediating heterodimer formation and DNA binding[53]. In addition, HIFα subunits contain oxygen-dependent (ODD) domains and transactivation domains (TAD) to promote the expression of target genes[54] (figure 5). Despite their structural similarities, including the DNA recognition motif, HIF-1 and HIF-2 are activated with different kinetics and bind to different cell-specific sites across the genome[55]. Conversely, HIF-3α cannot induce the expression of hypoxia-inducible target genes to the same extent as HIF-1α and HIF-2α as it lacks the C-terminal TAD[56] (figure 5). It acts as suppressor of HIF-dependent genes expression by competing with HIF-1α or HIF-2α to bind HIF-1β and the promotors of target genes[46]. The hypoxic response is primarily mediated by HIF-1 and HIF-2 which have both overlapping and unique target genes. HIF-1α is ubiquitously expressed in the body, while HIF-2α expression is restricted to specific tissues[57]. It appears that in some cell lines, HIFs have specific temporal and functional roles. HIF-1 drives the initial response to hypoxia (< 24 hours) while HIF-2 drives the chronic response (> 24 hours)[58]. Thus, during the hypoxic response, either in physiologic or pathological conditions, HIF-1 and HIF-2 play divergent while complementary roles.

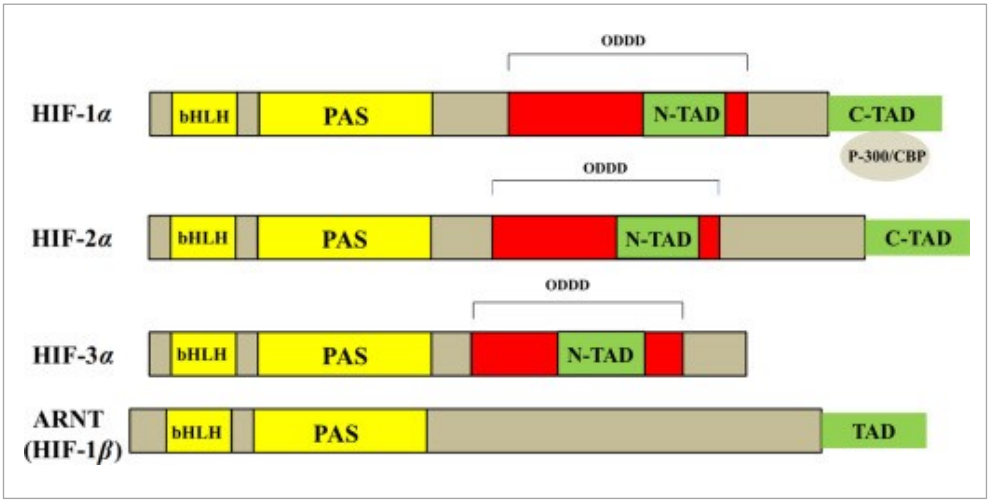


Figure 5 Structural domains of HIFs. All HIFs present both bHLH and PAS functional domains. HIF-1α and HIF-2α share high degree of amino acid sequence similarities and both of them have two distinct TADs (C-TAD and N-TAD). In contrast, HIF-3α has only the N-TAD. All HIFα subunits contain the oxygen-dependent degradation domain (ODD). From Masoud and Li, Acta Pharmaceutica Sinica B, 5: 378-389 (2015)

1.4.2 Prolyl hydroxylases and HIFα hydroxylation

Protein hydroxylation is a post-translational modification catalyzed by 2-oxoglutarate (2-OG)-dependent dioxygenases. These iron-containing enzymes pair substrate oxidation to the conversion of 2-OG[59]. In details, hydroxylation consists in the splitting of dioxygen into two oxygen atoms. One oxygen is added to an amino acidic residue (i.e. proline) to form the hydroxyl group and the other used to convert 2-OG into succinate and carbon dioxide[60]. $Fe2^{+}$ at the active site remains loosely bound by two histidine residues and one aspartic acid forming a 2-histidine-1-carboxylase coordination motif[60]. Thus, molecular oxygen, 2-OG and iron(II) are required for the catalytic activity. Hydroxylation can take place on various amino acids including proline, lysine, asparagine, aspartate and histidine. A typical example of this modification is HIF-1α hydroxylation. Under normoxia, HIF-1α is hydroxylated at either an asparagine residue by the factor inhibiting HIF (FIH)[61] or at specific proline residues by prolyl-4-hydroxylase domains (PHDs)[62]. FIH hydroxylates the asparagine residue within the C-TAD blocking HIF from binding to the p300-CBP coactivators and inhibiting in this way its transcriptional activation[63]. Otherwise, PHDs hydroxylate either the proline residue on the oxygen-dependent degradation domain at the N- or C-termini (NODD and CODD, respectively) of HIF1α and HIF2α. Hydroxylation occurs at Pro402-Pro564 and Pro405-Pro531 of HIF1α and HIF2α respectively [64,65]. This modification allows their recognition by pVHL, ubiquitination and proteosomal degradation. As HIF hydroxylases have a strict requirement for molecular oxygen, this

process is suppressed in hypoxia allowing the HIF-alpha subunits to escape destruction, dimerize with HIF-beta and activate genes target transcription. Furthermore, PHDs turnover is highly sensitive to O_2 concentration, making these protein well-suited oxygen sensors[66]. In human, PHDs is a family of three protein named PHD1, PHD2 and PHD3 (also known as EGLN 1-3)[62]. All isoforms possess a high sequence identity at the C-terminal domain, while their expression pattern is specific. Indeed, PHD2 is highly expressed in the adipose tissue, PHD1 in testis cells, whereas PHD3 is predominantly found in cardiac cells[67]. Also the subcellular localization is distinct. Although PHD1 is reported to be exclusively nuclear, PHD2 localizes in the cytoplasm, while PHD3 is equally distributed throughout both cell compartments[68]. It's also known that each isoform has a different affinity for the inhibition of HIFs subunits. Under normal conditions, PHD1 shows a preference for HIF-2α whereas PHD2 has a stronger substrate preference for HIF-1α. PHD3, in turn, preferentially degrades HIF2α under hypoxia[69]. This isoform is itself a target of HIF-1α.

1.4.3 The VCB-Cul2 complex

In normoxic conditions pVHL associates with Elongin B, Elongin C, Cullin2 and Rbx1 to form the VCB-Cul2 complex involved in HIF-1α degradation. HIF-1α recognition by pVHL is specifically driven by HIF proline hydroxylation. The crystallographic structure of HIF-1α peptide – VCB complex shows that HIF-1α binds to pVHL in an extended β strand–like conformation[31] (figure 6). The hydroxyproline inserts into a gap in the pVHL hydrophobic core, with its 4-hydroxyl group recognized by serine and histidine residues. Although the β sheet–like interactions contribute to the stability of the complex, the hydroxyproline contacts are central in the VCB complex function[31]. Indeed, the hydroxyproline site contact in pVHL is an hotspot for several tumor-associated mutations[70]. In the VCB complex, Elongin C binds Elongin B and pVHL across two distinct interfaces, whereas pVHL and Elongin B do not interact[30]. In particular, Elongin C interacts with pVHL α domain and a small part of the β domain (figure 6). The VCB complex, in turn, associates with both Cullin 2 and the RING-H2 finger protein Rbx1[71,72]. Cullin 2 acts as scaffold contributing to catalysis by positioning the substrate against the ubiquitin-conjugating enzyme, with Rbx1 mediating targets ubiquitination and their proteosomal degradation.

Very interesting, the pVHL binding partners share homology to component of the SCF multiprotein complex that targets many cell cycle regulatory proteins for ubiquitin-mediated proteolysis[73]. Elongin C has sequence similarity to the SCF Skp1 protein[74], Elongin B to ubiquitin[75] and Cul2 to the SCF Cul1 protein[76]. Similarly to the VCB-Cul2 complex, abnormal regulation of SCF contributes to uncontrolled proliferation, genomic instability, and cancer[77].

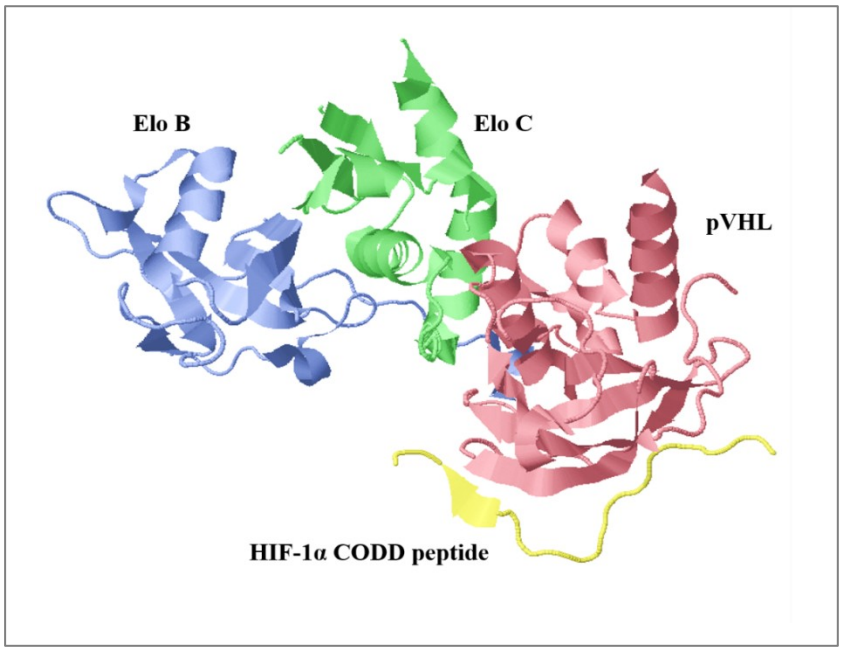


Figure 6 VCB cristal structure. PDB 6gfx visualized with Jalview. COOD of HIF-1α (yellow) containing an hydroxyproline is associated to VCB complex. In blue Elongin B, in green Elongin C and in red pVHL.

1.4.4 HIF responsive elements (HREs) and genes target

The VCB-Cul2 complex, described above, is formed and acts only in normoxic conditions. Instead, during hypoxia HIF-1α is not hydroxylated, translocates to the nucleus where it dimerizes with HIF-1β. As transcriptional factor, the heterodimeric HIF-1 recognizes and binds the consensus sequence 5'-(A/G)CGTG-3' named hypoxia-responsive elements (HREs) to activate genes target transcription[78]. To date, data published identified more than 100 direct HIF-1α genes target encoding proteins involved in cell homeostasis, angiogenesis, cell survival, glucose metabolism and invasion (figure 7). One of the major HIF-1α target genes is the vascular endothelial grow factor (VEGF) which is the most potent endothelial-specific mitogen to directly participate in angiogenesis[79]. In hypoxic conditions, this growth factor interacts with its receptor VEGFR stimulating endothelial

cells proliferation and migration, required for generation of new blood vessels and oxygenation[80]. Several other growth factors, most notably insulin-like growth factor-2 (IGF2)[81] and transforming growth factor (TGFα)[82], are also HIF-1α target genes. Binding of IGF2 and TGFα factors to their receptors activate signal transduction pathways that lead to cell survival and proliferation[83]. Under hypoxic conditions, energetic cells requirements change with cells switching to oxygen-independent pathways, i.e. glycolysis, to produce ATP[84]. HIF-1 regulates the expression of all the enzymes involved in the glycolytic pathway as well as the expression of glucose transporters GLUT1 and GLUT3 that mediate cellular glucose uptake[85]. As result, the increase in glycolytic metabolism enhances lactate production with consequent pH decreasing in cell microenvironment. This phenomenon occurs in several tumors yielding acidification of tumor environment and alkalization of tumor cell cytoplasm. Furthermore, HIF-1 regulates the expression of genes encoding cathepsin D, matrix metalloproteinase 2, fibronectin 1 and keratins; all proteins with an established role in the tumor invasion[51]. Altogether, HIF-1 regulates target genes that allow cells adaptation to hypoxic environment. The sustained activation of these pathways is associated with several tumours suggesting HIF to play a critical role in various stages of carcinogenesis from angiogenesis necessary for tumour survival to invasion at metastatic state.

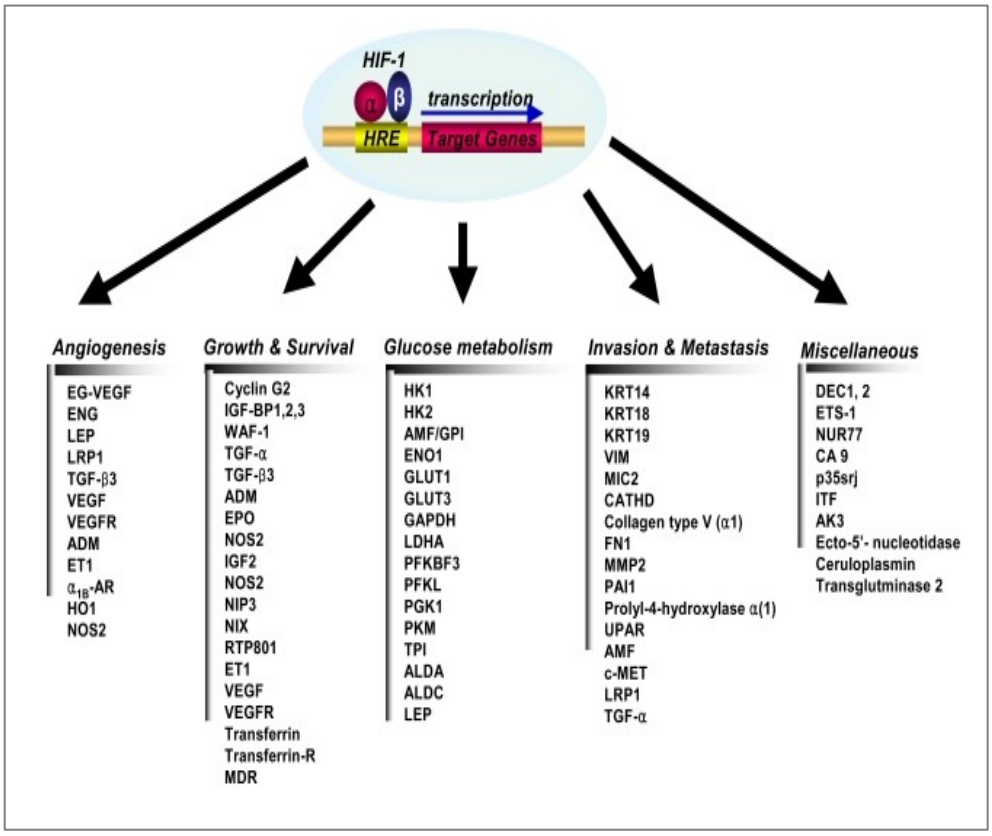


Figure 7 HIF-1 target genes. HIF regulates the transcription of several genes involved in different cellular processes. Its target genes encode for proteins involved in several aspect of tumorigenic process, such as initiation, tumor growth, metabolism, metastasis and invasion. From Hong et al., Cancer Research and Treatment 36(6): 343-353.

1.5 HIF-independent pVHL functions

Besides its well-characterized role in HIF-1α regulation, pVHL has also many HIF-independent activities such as apoptosis regulation[86], cell senescence control[87], microtubule stabilization[88], maintenance of the primary cilium[89] and regulation of extracellular matrix formation[90] and cell-cell adhesion[91]. Experimental evidences show that some of these functions are mediated through VHL targets stabilization, in contrast with its known E3 ligase activity[92]. Taken into account these two pVHL "faces", this protein can be considered as a multifunctional adaptor[93]. Depending on the interacting partner pVHL can promote protein degradation or act as chaperon. Furthermore, it is known that pVHL interacts with more than 500 different partners[32] suggesting its function as *hub* in a complex network of cellular pathways. Some alternative pVHL functions are described in details below.

1.5.1 pVHL associates with intrinsically disordered regions

In according to protein hubs classification published by Han et al. in 2004[94], pVHL can be considered a dynamic hub (also known as date hub) which binds its interaction partners at different time and/or localization, consistently with their large number. This ability is strictly associated with structural flexibility, necessary to mediate different kind of protein-protein interactions. The structural flexibility is defined as the ability to change folding into different structural conformations by presenting several loops or coils regions[95]. It is positively influenced by the presence into the protein of regions that lack a clear structure, generally referred as intrinsically disordered regions (IDRs). These unstructured regions are proposed to have a double function: they are important for flexible and rapidly reversible interactors binding and also serve as linkers between structured domains[96,97]. Also the solvent-exposed charges seem to help hubs in sustaining multiple interactions[98]. In the case of pVHL, this protein contains a long intrinsically disordered and negatively charged N-terminal tail.

It is well established that IDRs actively participate in different functions mediated by proteins. IDRs are frequently regulated by post-translational modifications (PTM)[99,100] and expose short linear motifs (SLIMs)[101] of about 3-10 amino acids that are able to sustain multiple interactions with structured domains of other proteins. SLIMs can target proteins to a particular subcellular localization, control the stability of a protein and promote recruitment of binding factors thus facilitating multiprotein-complex formation[102]. SLIMs can be grouped in two large major families, with those acting as modification sites opposed to SLIMs acting as ligands[103]. Among this second group, can be listed different functional behavior such as protein-complex promoting motifs, docking motifs and targeting motifs. In particular, docking motifs increase the specificity and efficiency of modification events (i.e. addition or removal of PTM) generating additional binding surfaces. Example are the degron motifs[104] (i.e. KEN box and D box) which act as recognition surfaces for ubiquitin ligases regulating protein degradation by the 26S proteasome. To this group belong the hydroxy-degrons recognized and bound by pVHL.

As component of the E3-ubiquitin ligase complex, pVHL seems specifically to bind short intrinsically disordered regions containing a hydroxylated proline residue. The original linear motif involved in pVHL – HIF-1α recognition is the LxxLAP motif within the so-called NODD and CODD fragments of HIF-1α[105]. The same motif is also highly specific for PHD-mediated hydroxylation[106]. New evidences have extended these initial findings

suggesting that different pVHL interactions are mediated upon proline hydroxylation. Collectively, the linear motifs involved in pVHL binding can be referred as hydroxy-degrons. As mentioned before, this particular motifs require activation by PTM, in this case promoted by 4-prolyl-hydroxylase enzymes. The activity of these enzymes requires molecular oxygen, thus differences in oxygen concentration in different tissues may differently impair hydroxy-degrons function and drive cancer progression. Examples of hydroxy-degrons recognized by pVHL are found in Spry2, EPOR, ADRB2 and MYBBP1A proteins.

Spry2 (protein sprout homolog 2) is a modulator of MAPK/ERK pathway[107] acting as growth factors antagonist. It mediates the quiescence and the barrier integrity in endothelial cells[108] preventing cancer progression. Indeed, Spry2 expression is repressed in cancer of the breast, liver, lung, prostate and in lymphoma[109]. PHD hydroxylate Spr2 on three proline residues (i.e. 18,144 and 160) to form three putative hydroxy-degrons recognized by pVHL[110]. Very interestingly, Spry protein family is reported to inhibit activation of ERK in response to fibroblast growth factor (FGF)[111] and vascular endothelial growth factor (VEGF)[112]. This evidence suggests pVHL to perform an additional regulation of VEGF in HIF-1α-independent manner.

EPOR Another example of hydroxy-degron was found in erythropoietin receptor (EPOR) in 2016[113]. EPOR is hydroxylated on Pro419 and Pro426 by PHD3, the resulting degrons have no sequence similarity to LxxLAP motif, while localize within a proline rich region. Authors found pVHL mutants that maintain the ability to bind and regulate HIF-1α while show severe defect in binding EPOR. In particular, pVHL mutations associated to VHL sub-type 2C show defect of the protein to bind hydroxylated EPOR[113]. EPOR gene transcription is regulated by HIF-2α[114]. This receptor mediates erythropoietin-induced erythroblast proliferation and differentiation. In particular, the Epo receptor activates the JAK-STAT pathway: upon ligand binding, the receptor-associated tyrosine kinase JAK2 is phosphorylated and in turn phosphorylates the transcription factor STAT5[115,116]. In this system, pVHL is involved in JAK2 ubiquitination and degradation. When mutated, pVHL fails in JAK2 degradation leading to prolonged JAK-STAT activation and insurgence of polycythemia, an overexpansion of erythrocytes[117]. This pathological conditions is also linked to EPOR mutations[118] and hypoxia sensing impairment[119].

ADRB2 In this case the hydroxy-degron is associated to hypoxia-regulated beta2-adrenergic receptor (ADRB2) degradation[120]. ADBR2 is known to modulate the

intracellular oxygen homeostasis modulating AMP/ATP ratio, ROS production and PHD activity[121]. PHD3 hydroxylates ADR2 in position Pro382 and Pro395 generating two pVHL binding sites. These hydroxy-degron present low similarity with the canonical LxxLAP motif of HIF-1α confirming that pVHL β domain is able to recognize different hydroxy motifs. ADRB2 is demonstrated to have a critical role in modulating the hypoxia response suggesting the PHD3-ADRB2-pVHL axis involvement in cancer progression.

MYBBP1A (Myb-binding protein 1A) may activate or repress transcription via interactions with sequence specific DNA-binding proteins. In particular, it is known to regulate AhR-dependent gene expression[122], suppress mitochondrial respiration[123] and together with CRY1 repress the transcription of the core circadian clock component PER2 in mammals[124]. Hydroxylation of Pro695 of MYBBP1A activates a degron signal which promotes its pVHL-mediated degradation[125]. MYBBP1A is linked to the regulation of anoikis[126], a programmed cell death that occurs when a specific cell detach from the extracellular matrix (ECM). pVHL is known to regulate ECM deposition [127]. Defects in the hypoxia-depend regulation of MYBBP1A may have a role in cancer progression.

Taken together, these examples prove that the pVHL ability to recognize and bind hydroxy-degrons links the protein to different biological functions.

1.5.2 pVHL and cell-cycle regulation

The cell-cycle is the series of events that take place in a cell leading to DNA replication, division and daughter cells production. It consists mainly in DNA synthesis (S phase) and mitosis (M phase) separated by two "gap" phases named G_1 and G_2. In the first gap the cell is preparing for DNA synthesis, whereas in the second cell prepares for mitosis. When cell is not actively in the cycle, this step is referred as G_0 phase[128]. The transition from one cell phase to another is finely regulated by different proteins. Key regulatory proteins are the cyclin-dependent kinases (CDK), a family of serine/threonine kinases that are activated at specific check-points. Until now, five CDKs are known to be active during the cell cycle; in particular during G_1 (CDK4, CDK6 and CDK2), S (CDK2), G_2 and M (CDK1)[129]. CDK protein levels are stable during the cell-cycle, whereas those of the cyclins (CDK activating proteins) periodically changes[130]. Different cyclins act at different cell-cycle phases: D-type cyclins are essential for entry in G_1, cyclin E regulates progression from G_1 into S phase, cyclin A is required during S phase and promote entry into M phase regulated also by cyclin B[131]. In addition, CDK activity is also regulated by small inhibitory proteins defined as CDKs inhibitors (CKI) which bind to CDK or CDK-cyclin complex[132]. There

are two distinct families: the INK4 family and Cip/Kip family. The first family includes p15 (INK4b), p16 (INK4a), p18 (INK4c) and p19 (INK4d) which specifically inactivate G_1 CDK (CDK4 and CDK6). Instead, the second one include p21 (Cip1), p27 (Cip2) and p57 (kip2) that inhibit the G_1 CDK-cyclin complexes (figure 8).

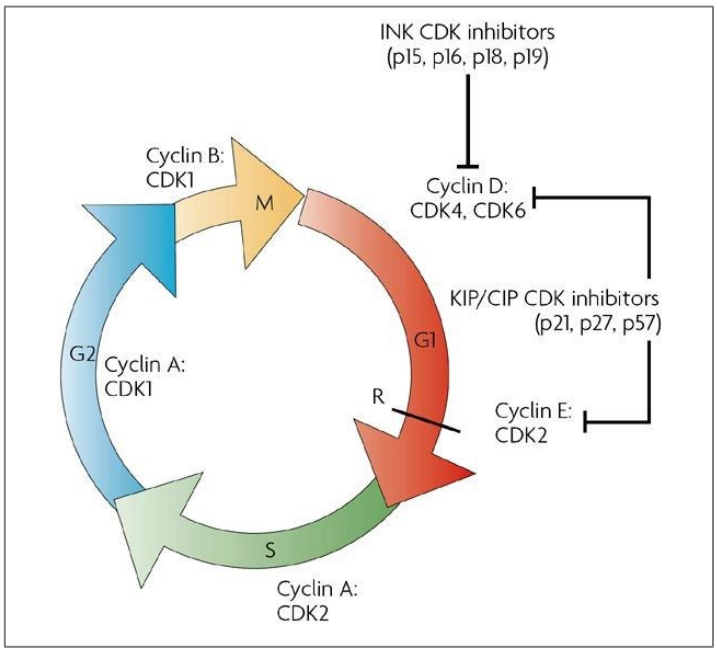


Figure 8 Cell-cycle regulation. The cell cycle consists of four distinct phases: G1 phase, S phase (synthesis), G2 and M phase (mitosis). Each step is regulated by specific cyclin-CDK complexes inhibited by CDK inhibitors (INK and Cip/Kip families). From Colette Dehay and Henry Kennedy, Nature vol.8 438-450

The right sequence of events in the cell-cycle is also regulated by additional controls and checkpoints like DNA damage checkpoints that arrest the progression in order to provide time for DNA repair. These kind of checkpoints are positioned before the cell enters in S phase (G_1-S checkpoint) or after DNA replication (G_2-M checkpoint) inducing arrest in G_1 and G_2 phases. The G_1 arrest is p53-dependent whereas the p53 role in G_2 is controversial[133]. At physiological conditions p53 level is low but DNA damage leads to induction of p53 activity[134]. In particular, p53 blocks the cell-cycle activating p21 that inhibits G_1 CDKs. p53 regulates also itself through a negative feedback loop in which stimulates the transcription of MDM2 (Mouse double minute 2 homolog), the E3-ubiquitin ligase mainly involved in its ubiquitination and degradation[135]. Furthermore, p53 levels are also regulated by ARF protein, encoded by the ARF-INK4 locus, which binds MDM2 preventing MDM2-mediated p53 degradation[136]. When these regulation mechanisms fail, p53 induces cell death activating genes involved in apoptotic signaling.

The complex system of cell-cycle control is also indirectly regulated by multiple kinases. Among them, Akt is known to modulate the function of numerous substrates related to cell cycle progression at the G_1/S and G_2/M transitions, either by direct phosphorylation of the target proteins or indirectly by regulating protein expression levels (figure 9). In more details, growth-factors stimulation triggers Akt activation that leads to increase in c-Myc transcription causing cells exit from G_0 by inducing D-type cyclins and suppressing the expression of p21 and p27[137]. Akt phosphorylates p21 at Thr145, within its nuclear localization signal (NLS) inducing its translocation to the cytoplasm. In this way Akt suppress p21 inhibitory effect on cell cycle progression[138]. Similarly Akt downregulates p27 transcription by phosphorylation at Ser10, Thr187 and Thr198 and consequent translocation to the cytoplasm[139]. Also the p27 protein stability is controlled by Akt-dependent phosphorylation of Skp2, a component of the E3-ubiquitin complex SCF involved in p27 degradation[140]. Furthermore, in response to growth factor Akt binds and phosphorylates MDM2 at Ser166 and Ser186 to enhance protein stability and facilitate p53 degradation[141]. The MDM2 stabilization is also due to MdmX binding, another Akt target protein. All these regulatory mechanisms occur at the G_1/S phase. Similarly, Akt also regulates G_2/M phase through phosphorylation of both Cdk1 activators and inhibitors.

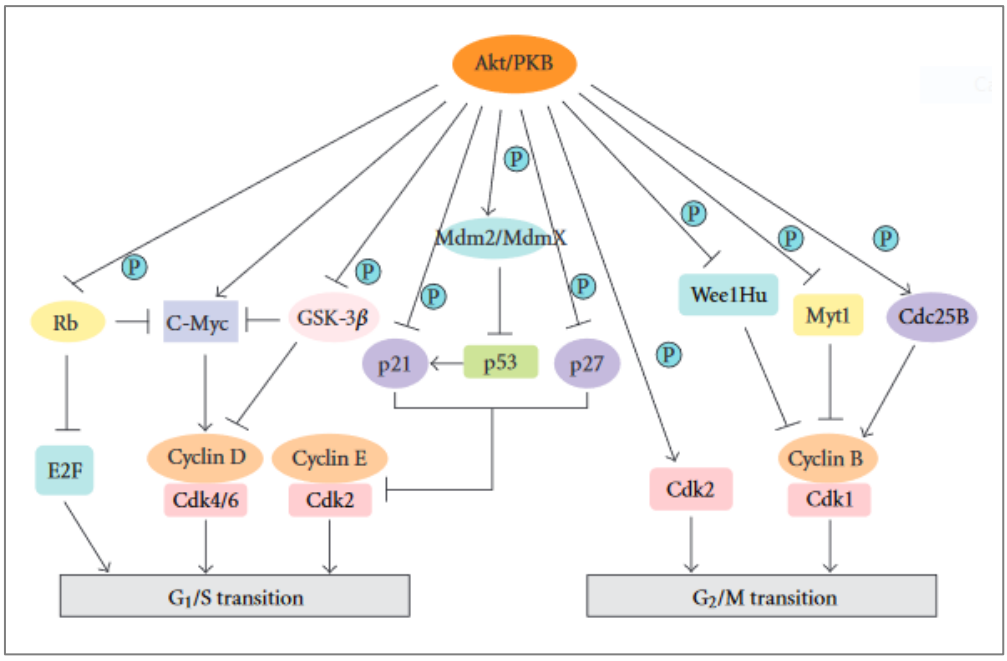


Figure 9 Role of Akt in cell-cycle progression. Activated Akt kinase regulates a number of proteins involved in cell cycle progression at the G1/S and G2/M transitions, either by direct phosphorylation of the target proteins themselves or indirectly regulating protein expression levels. From Xu et al., Journal of Oncology Volume 2012, Article ID 951724,15 pages

Very interesting, several evidences link pVHL to cell-cycle regulation. In 2006, Roe at al., demonstrated pVHL to directly bind and stabilize p53 by suppressing its MDM2-mediated ubiquitination and degradation in a HIF-1α independent manner[142]. In particular, the pVHL α domain was found to associate with the DNA binding domain of p53 located in the central part of the protein. The pVHL-p53 interaction enhances p53 transactivation through its acetylation. On the contrary, in 2015, Essers et al., showed that MDM2 degradation is regulated upon pVHL interaction with PDCD5 (programmed cell death 5 protein), a negative MDM2 regulator[143]. In this case, in contrast with the pervious findings, the p53 stabilization is indirect and occurs in absence of pVHL. As mentioned before p53 levels are also regulated by ARF. This protein was demonstrated to associate specifically with pVHL30 and not with the shorter isoform pVHL19[144]. p14ARF was found to release pVHL30 from the E3 ligase complex, promoting the binding of pVHL30 to a protein arginine methyltransferase, PRMT3. This finding suggests an important role of the pVHL N-terminal tail in pVHL30-ARF association. Further studies link pVHL to Akt. In particular, Akt1 was found to associate with pVHL upon Akt hydroxylation by PHD2, which in turn is the main regulator of HIF-1α hydroxylation[145]. The interaction promotes Akt1 functional inhibition without effect on protein degradation. Authors showed also that pVHL residues driving the binding to Akt1 and HIF-1α partially overlap but are not identical suggesting that the pVHL β-domain may evolved to exert other functions over the HIFs α subunits recognition. The pVHL itself is also substrate of different kinases[146,147]. Youn et al., reported that the phosphorylation of pVHL Ser111 by Chk2 (checkpoint kinase 2) enhances pVHL-mediated transactivation of p53 on DNA damage[147] by recruiting p300 and Tip60 to the chromatin of p53 target genes.

All these examples demonstrate that pVHL exerts its function as tumor suppressor also regulating cell-cycle in HIF-1α independent manner. Considering that hypoxia-induced cell cycle arrest is also due to HIF-1α dependent transcription of Cip/Kip CDK inhibitors, it could be interesting to evaluate whether pVHL regulates these proteins also directly.

1.5.3 pVHL associates with the Androgen receptor

Sex is a key factor affecting the etiology, pathogenesis and prognosis of specific types of cancer. The 2-to-4-fold higher risk of males to develop cancer compared to females is independent of socio-economic conditions and geographical origins suggesting that this sex discrepancy is due to genetic biological factors, still unknown[148,149]. The renal cell carcinoma (RCC) is an example of disease characterized by sex discrepancy. This evidence

indicates that sex hormones and their receptors could have a role in the insurgence and progress of the cancer. The androgen receptor (AR) and its natural ligands, testosterone and dihydrotestosterone, are very important in the development of primary and secondary sexual characteristics during development. Mutations in AR correlated with a spectrum of androgen-related diseases. Loss of function mutations leads to androgen insensitivity syndrome[150], whereas gain of function mutations cause prostate cancer[151]. Structurally, AR is composed of three domains, the amino-terminal domain, the DNA-binding domain (DBD) and the hinge region, and the ligand-binding domain (LBD)[152]. In the N-terminus portion, it is characterized by a polyglutamine (poly-Q) tract whose length affects AR function leading to androgen-dependent disorders. Epidemiologic studies demonstrated that subjects with a poly-Q shorter than 20 residues present an higher risk to develop prostate cancer compared with individuals with poly-Q longer than 26 aa[153,154]. Furthermore, AR poly-Q expansions over 28 residues are associated with male infertility[155] whereas repeats longer than 38 aa leads to the spinal bulbar muscular atrophy (SBMA)[156], a neuromuscular disease. In both cases alterations in AR poly-Q length induce AR dysfunction. AR is a transcription factor regulated by androgens binding. In its inactive state, AR localizes into the cytoplasm associated with heat shock proteins (HSPs). Instead, upon hormones binding, AR dissociates from HSPs, translocates to the nucleus and binds the androgen-responsive elements (ARE) to regulate the expression of its target genes. Androgen binding also induces post-translational modifications, such as phosphorylation, which regulate different AR aspects like degradation, subcellular localization and ligand binding affinity[157].

Very interesting, experimental evidences showed AR to associate pVHL suggesting a functional relationship between a tumor promoting gene (AR) and a tumor suppressor (pVHL). pVHL is able to form a complex with AR in cultured cells[158,159]. In particular, in 2013 was demonstrated pVHL β domain to associate with the AR DBD and hinge regions inducing AR transactivation. According to the authors, the pVHL–AR association stimulates AR transcriptional activity upon dihydrotestosterone (DHT) treatment [159]. On the contrary, one year later Wang et al., identified the AR LBD as the portion involved in pVHL β domain binding. In this case pVHL was shown to inhibit AR transcriptional activity without altering the AR turnover. The authors provided evidence to show that pVHL enhances AR de-ubiquitination, suggesting a non-canonical function of pVHL in the regulation of AR[158]. Moreover, additional post-translational modifications, such as phosphorylation and prolyl hydroxylation, may contribute to regulate the interaction

between AR and pVHL. This association offers therapeutic opportunity based on a promising strategy named PROTAC[160], at the moment in phase I clinical trial. This technology is based on the creation of bivalent molecules that bring together an E3-ubiquitin ligase and a target protein to degrade (figure 10). It was already employed using MDM2 to degrade AR[161]. Thus, pVHL could be used as alternative E3-ubiquitin ligase to the same purpose. Taken into account the possible therapeutic relevance and the lack into mechanistic insight of pVHL-AR interaction could be interesting further investigate the functional role of this association.

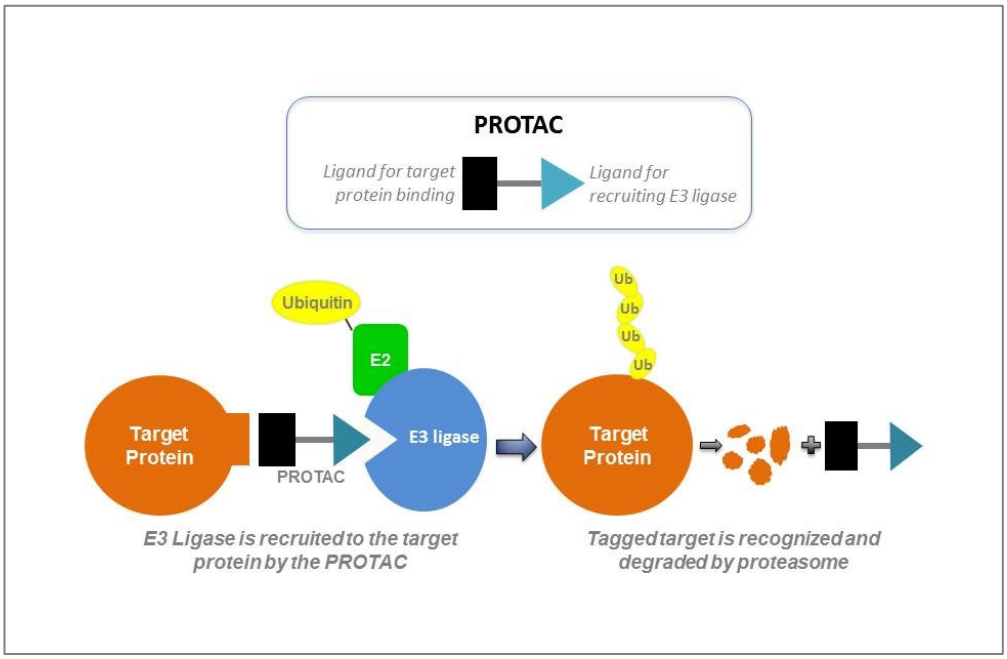


Figure 10 The PROTAC mechanism of action. Synthetic molecules are used to bring together an E3-ubiquitin ligase and a specific protein target. From the official blog of American Association for Cancer Research

Taken into account all the information reported in this introduction, it is possible conclude that pVHL exerts its function as tumor suppressor in two ways. It is the main regulator of the hypoxia response through HIF-1α but it is also involved in several HIF-1α independent activities. This second pVHL "face" is less understand despite in the last years high-throughput approaches identified a huge amount of data about pVHL interactors. There are, however, several aspects that are not well-described such as the functional role of these associations or the tissue specificity of pVHL interactions. Other interesting point to deeper investigate is the recently proposed idea of pVHL isoform-specialization that could reveal functional differences among pVHL30 and pVHL19. Considering these aspects my

PhD project focuses on the investigation of pVHL non-canonical functions considering both the pVHL19 and pVHL30 isoforms and it tissue specificity. To this purpose a combination of multiple different approaches including bioinformatics, yeast biology and mammalian cells biology were used to identify and dissect new interactions, reveal pVHL30 interaction partners in testis tissue and try to shed light on the functional connection between pVHL30 and the human androgen receptor.

Materials and Methods

2 - Materials and Methods

2.1 Bacteria, yeast and mammalian cells

Plasmid construction and Yeast two-hybrid experiments were performed using bacteria and different yeast strains; all details are reported in the table below. (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 Characteristics of bacteria and yeast strains

In addition, mammalian cell lines were used during experimental procedures. Details in 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 Characteristics of mammalian cell lines

2.2 Culture media

According to medium composition, powders were mixed and resuspended in deionized water (ddH2O). Solutions were then sterilized at 121°C for 40 minutes.

2.2.1 Bacteria media

Bacteria were cultured in Luria-Bertani (LB) medium. Composition is reported in Table 3.

Medium	*Composition*
LB (Luria-Bertani)	1% Bacto Triptone (Difco) 0.5% Yeast Extract (Difco) 0.5% NaCl (Sigma)

Table 3 Bacteria medium composition

When required, LB was complemented with Agar 2% (Cat. 214030, Bacto™Agar) and/or antibiotics such as Ampicillin (100 μg/mL) or Kanamycin (50 μg/mL).

2.2.2 Yeast media

Yeasts were cultured mainly in rich or selective medium. Compositions are listed in Table 4.

Medium	Composition
YPD (rich medium)	1% Bacto Peptone (Difco) 1% Yeast Extract (Difco) 2% Dextrose (Sigma)
SD (selective medium)	0.17% Yeast Nitrogen Base (Difco) 2% Dextrose (Sigma) 0.5% Ammonium Sulfate (Sigma)

Table 4 Yeast media composition

Solid media were obtained adding Agar (2% in YPD and 2,3% in SD) at the solutions. When necessary, YPD was complemented with Adenine (Sigma, 80 mg/ml). Instead, SD was supplemented with Drop-out (Sigma Y2001), amino acids (Sigma, Adenine 80 mg/ml; Histidine 10 mg/ml; L-leucine 60 mg/ml; Tryptophan 20 mg/ml) or 3-amino-1,2,4 triazole (3-AT) at different final concentrations (i.e. 30 mM, 60 mM, 90 mM).

Media used to perform the library screening described in the Results section were purchased from Clontech. Details are listed in Table 5 (following page).

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	SD-Leu Broth 630309	0.5
SD-Leu with Agar	630310	0.5
	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 were provided by Clontech as a "ready-mixed" foil pouches, so media were resuspended in deionized water and then sterilized before the use. When required media were supplemented with Aureobasidin A (Cat. 630466, 200ng/ml) and X-α-Gal (Cat. 630462 40μg/ml).

2.2.3 Cell media

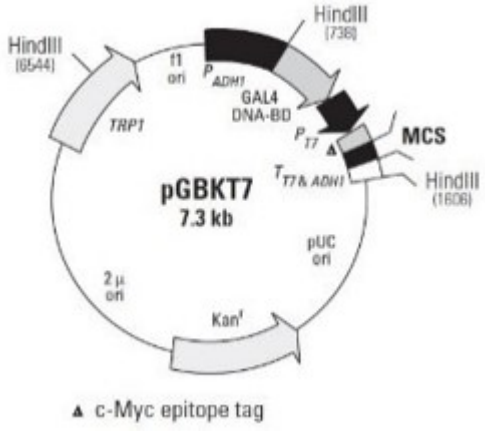
HEK293T (ATCC), and HeLa (ATCC) cells were cultured in Dulbecco's modified Eagles's medium high glucose (DMEM; Euroclone) with 10% of heat inactivated fetal bovine serum (FBS,Euroclone), penicillin/streptomycin (100 mg/ml,Euroclone) and L-glutamine (2 mM,Euroclone) at 37°C in a humidified atmosphere containing 5% CO_2.

2.3 Plasmids

2.3.1 Yeast vectors

All the experiments performed in yeast were based on yeast two-hybrid assay (Y2H). For this reason, all genes were cloned using two main plasmids: the bait vector pGBKT and the prey vector pGADT7. Empty vectors were provided by Clontech as components of Clontech's Matchmaker™ Systems.

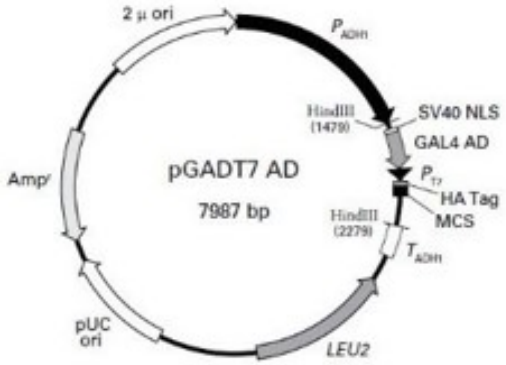
The bait vector: pGBKT7



pGBKT7 map adapted from Clontech datasheet (PT3248-5)

pGBKT7 (Cat. 630489) is a bait vector of 7,3 kb length that allows the expression of a protein fused to amino acids 1–147 of the GAL4 DNA-BD. The protein is expressed in frame with c-Myc epitope tag useful to detection in Western blot (WB). The gene transcription is regulated by a constitutive ADH1 promoter (PADH1) and by the T7 and ADH1 termination signals (TT7 & ADH1). It's also present a T7 promoter, used to do sequencing analysis. The vector carries the Kanr 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.

The prey vector: pGADT7

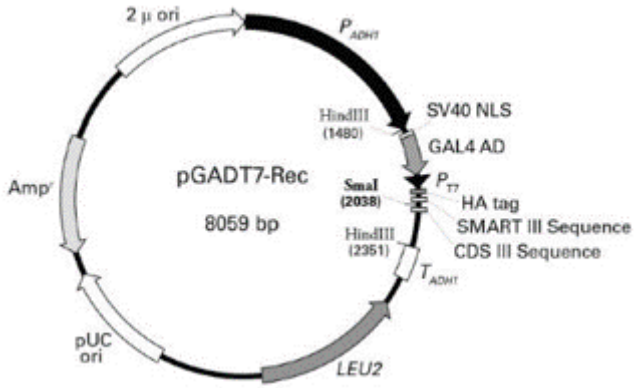


pGADT7 map adapted from Clontech datasheet (PT3248-5)

pGADT7 (Cat. 630442) is a prey vector of 7,9 kb length that allow the expression of a protein fused to amino acids 768–881 of the GAL4 activation domain (AD). In addition, the prey protein is fused to a common HA epitope tag, that allows the protein detection in WB. Transcription promoters and terminators are the same of pGBKT7. 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.

The library screening is also based on Y2H assay. In this case the bait was expressed used pGBKT7 described above, instead the prey proteins were expressed in pGADT7-Rec vector that is slightly different from pGADT7.

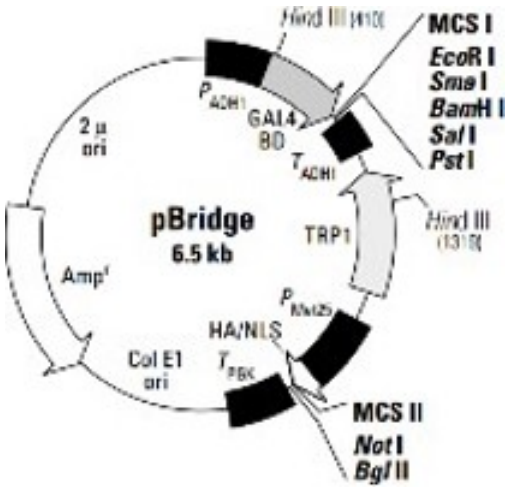
The library prey vector: pGADT7-Rec



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

pGADT7-Rec (Cat. Sold as part of 630490) is a vector of 8 kb length, it 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. All the others characteristics (promoters, terminators, markers and replication origins) are the same of pGADT7 vector.

The bait vector in Y3H system: pBridge

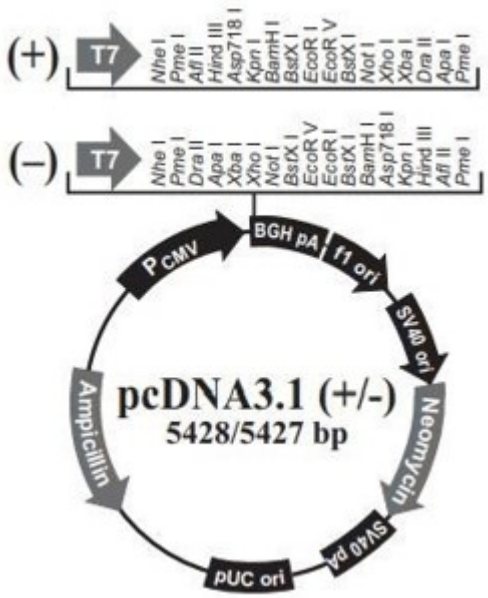


pBridgeTM map adapted from Clonotech datasheet (PT3212-5).

pBridgeTM (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[162]. 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[163].

2.3.2 Mammalian cell vectors

pCDNA3.1+



pcDNA3.1 vectors map from TermoFisher. Both the forward-orientation pcDNA3.1(+) and the reverse-orientation pcDNA3.1(-) are represented in the image

pcDNA™ 3.1(+) vector (ThemoFisher, V79020) is a vector of 5,5 kb length, it 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*.

All plasmids expressing Androgen receptor protein were previously generated in Prof. Pennuto laboratory and gently provided for this work. In particular AR24Q, AR65Q, AR24Q mutants were cloned into pCMV5 vector; instead AR flag constructs were cloned into pCMV6 vector.

2.4 Cloning vectors

The described plasmids were manipulated to create new vectors, employing the following *in vitro* approaches.

2.4.1 In-Fusion® protocol

In-Fusion® HD Cloning kit (Clontech; Cat. 011614) is a commercial kit used to create recombinant plasmids. This system is based on in vitro recombination due to recognition of 15 bp present at the end of both vector and insert. Generally, 15 bp already present in the vector (protruding ends after digestion with restriction enzyme), are added into the primers for insert amplification. 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 Tables (Tables 6 and 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.

PCR product was purified as described later (section 2.7). To perform the In-fusion reaction it was also necessary digest 3 µg of vector for 5h at 37 °C according to restriction enzyme protocol and then purify it by agarose gel excision (section 2.7).

At this point In-fusion reaction was assembled as reported in Table 8 (following page).

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 into bacteria, as reported in section 3.5. Recombinant plasmid obtained was finally controlled by Sanger sequencing.

2.4.2 Ligation protocol

Alternatively, recombinant plasmids were obtained using the traditional ligation method. In this case, the insert and the vector (digested or amplified in order to have compatible sticky ends) were fused using the T4 DNA ligase (BioLabs, M0202). The ratio plasmid: insert used was from 1:3 up to 1:10 according to their length in base pairs.

Reaction was assembled as reported in 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 DNA Ligase (M0202) ligation protocol.

The reaction was then used to transform TOP10 *E. Coli* (as described in section 2.5). Recombinant plasmids was finally controlled by Sanger sequencing.

2.5 Bacteria transformation

Bacteria cells (Calcium chloride competent cells homemade or commercial competent cells) were transformed using 100 ng of DNA. To start DNA was added to bacteria and the mixture was incubated for 30 minutes on ice[164]. The samples were exposed to heat shock at 42 °C for 1 minute and again on ice for 2 minutes. At this point, 1 ml of LB medium was added to each sample and bacteria were incubated for 1 h at 37 °C. Finally cells were spread on LB-agar plate supplemented with the proper antibiotic. The plates were incubated at 37 °C o/n.

2.6 Plasmids extraction

2.6.1 Bacteria plasmids extraction

Each plasmid was obtained from 5 ml of o/n *E.Coli* culture using Plasmid Miniprep Kit (Cat. PLN350, Sigma). After centrifugation, bacteria pellet was resuspended in 200 µl of resuspension buffer. 200 µl of Lysis buffer was added and the sample was gently mixed by inversion (6-8 times) until the mixture becomes clear and viscous. At this point, alkaline lysis was stopped by the addition of 350 µl of neutralization buffer that allows the formation of cell debris. To separate cell debris, each sample was centrifugated at ≥12000 g for 1 minute. The cleared lysate was transferred into the column (previously washed with 500 µl of Column preparation solution) and centrifugated at ≥ 12000 g for 1 minute. A wash step was performed adding 750 µl of wash buffer and centrifuging ≥ 12000 g for 1 minute. The flow-through was eliminated and the column was centrifugated again to remove ethanol excess. Elution buffer or ddH_2O (50 µl) was used to elute the DNA by centrifugation at ≥ 12000 g for 1 minute. When necessary, an up-scale plasmid extraction (i.e. maxi preparation) was performed to PureLink® HiPure Plasmid Purification Kit (Cat. K210006, Invitrogen).

In the case of library screening, considering the high number of clones analyzed, bacteria plasmids extraction was performed using the Zyppy TM -96 Plasmid Miniprep kit (Cat. D4041) provided by Zymo Research. It is the faster and simplest high-throughput method available for efficient isolation of plasmid DNA from *E. Coli*. The kit is based on a modified Pellet-Free alkaline system that bypass centrifugation, pelleting and re-suspension common to conventional procedures. In this case, bacteria were inoculated

directly in a 96-well block (750 µl /well) and the lysis buffer was added directly to the 96-well block without previous centrifugation. All the steps are similar to common procedures. The use of this kit allowed to obtain 96 different plasmids in the same extraction.

2.6.2 Yeast plasmids extraction

To start the analysis of positive clones obtained from Y2H library screening it was necessary extract DNA from yeast *S. cerevisiae* cells. Considering the high number of positive clones, also in this case DNA extraction was performed using a commercial kit (Zymoprep™-96Yeast Plasmid Miniprep- Cat D2005) provided by Zymo Research.

This kit is designed for the simple, rapid and high-throughput (96-well) isolation DNA from tough-to-lyse fungi including *S. cerevisiae.* It is based on the classic *E.coli* alkaline method (described above) with addition of Zymolyase in the first buffer; so elimination of yeast cell wall is based on digestion. DNA plasmids isolation can be performed starting from colonies, patches or liquid cultures. In our case, DNA plasmids were obtained from liquid cultures.

2.7 DNA gel extraction and purification

Linearized vectors or PCR products were excised from 1% agarose gel with a blade, minimizing the amount of agarose around the bound. The purification was performed according to the protocol of GenElute ^{TM} Gel Extraction Kit Sigma (Cat. NA1111). To start the gel piece was weight. Three gel volumes of the Gel Solubilization solution were added to the gel and the mixture was incubated and mixed at 55 °C for 10 minutes. When the gel was dissolved, one volume of isopropanol was added and mixed to the solution. To prepare the column, 500 µl of Column preparation Solution were added to the column and centrifugated at ≥12000 g for 1 minute. At this point the mixture was loaded into the column and centrifugated at ≥12000 g for 2 minutes. A wash step was performed adding 700 µl of Wash Solution, followed by centrifugation at 12000 g for 1 minute. The flow-through was eliminated and the column was centrifugated again to remove ethanol excess. Elution solution or ddH_2O (50 µl) was used to elute the DNA by centrifugation at ≥ 12000 g for 1 minute. To increase DNA plasmid recovery Elution Solution wasp re-heated at 65 °C.

2.8 Yeast as a model

Yeast is an ideal experimental organism, a powerful tool for studying cell biology. In 1988 Botstein and Fink affirmed that "the reason that yeast could serve as a model for all eukaryotic biology derives from the facility with which the relation between gene structure and protein function can be established". Yeast is a useful model for many reasons: it is cheap and simple to manipulate, it can exist both in haploid or diploid form, it grows in aerobic and anaerobic conditions. Furthermore, in 1996, yeast genome has been completely sequenced and further studies demonstrated that the 46% of yeast genes have an orthologue in human; in particular 290 yeast genes are orthologues of human diseases genes. This evidence underlines that many signalling pathways are conserved between yeast and mammals. For this reason , yeast is used to elucidate several cellular processes using molecular techniques based on DNA microarrays, gene disruption, protein localization, protein-protein interactions (PPi) and functional analysis by genetic interactions. In this thesis, yeast was used as model organism to study protein-protein interactions.

2.8.1 The Yeast two-hybrid system (Y2H)

This system allows to identify binary protein-protein interactions in living yeast cells. It is based on the modularity of the eukaryotic transcription factors like GAL4, composed by the Gal4 DNA-binding domain (DBD) and the Gal4 DNA activation domain (AD). This two domains can exist separately but the Gal4 transcription factor is active only when the two subunits are linked together. To test whether two proteins are able to interact, they are expressed one fused to the Gal4-DBD and the other fused to the Gal4-AD. The protein fused to the Gal4-DBD is referred to as the "bait", while the other fused to the Gal4-AD is referred as the "prey". Upon interaction between bait and prey proteins, the Gal4 subunits are close, the TF Gal4 is reconstituted and activates the transcription of its target genes. Typically, this technique is applied in genetically modified yeast strains, in which some genes (i.e. HIS3,LEU2,TRP1) that encode key enzymes for amino acids biosynthesis, are mutated in order to create nutritional auxotrophic markers. As result, these yeast strains are unable to synthetize some amino acids and to grow on medium without these nutrients (selective medium). In this contest, the gene target transcription by the Gal4 TF leads to auxotrophies compensation and the yeast strain grows on selective medium. In this thesis, HIS3 was the most employed reported gene choose to study protein-protein interaction.

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 without this amino acid.

2.8.2 Yeast transformation

Yeast transformation was performed according to one-step protocol [165]. An overnight yeast culture was divided into 2 ml microtubes and centrifugated at 14000 rpm for 10 minutes. Surnatant was discarded and 5 µl of DNA carrier from Salmon Testes (previously heated at 95 °C for 5 minutes) and 1 µg of plasmidic DNA were added to each sample. At this point, the pellet was resuspended with 100 µl of Transformation Solution (LiAc 2M pH 7.5, DTT 1M and PEG 4000 50% w/v). During all these passages the sample was kept on ice. Lithium ions and heat shock promote passage of DNA into the cell, while PEG promote association of the transforming DNA with the surface of the cell[166]. Sample was incubated at 45°C for 30 minutes. Then, it was centrifugated, surnatant was eliminated and the pellet was resuspended in 150 µl of NaCl 0,9% solution. Transformants were selected seeding yeast cells on selective medium. Plates were incubated at 28 °C for 3 days and generally for each transformation 5 colonies were patched out and analysed.

2.8.3 Yeast spot assay

This assay can be used to compare the cell growth rate of yeast under different growth conditions. It involves the serial dilution and spotting of yeast single colonies on plates. For each transformation three or more independent colonies were tested. Commonly, yeast cells were spotted on permissive and selective medium. In the first condition (SD – Leu – Trp) all yeast strains grow. Conversely, the selective medium was employed to detect interactions. It consists in (SD – Leu – Trp – His) with the addition of 3AT (3-Amino-1,2,4-Triazol) at different concentration (i.e. 30 mM, 60 mM, 90 mM). Plates were incubated at 28/30 °C and yeast growth was monitored for 8 days. Each experiment was repeated three times using yeast colonies from three independent transformations.

2.8.4 Yeast protein extraction

Proteins were extracted using a protocol TCA-based[167]. Yeast cultures were centrifugated and surnatants were eliminated. Pellets were resuspended with 100 µl of TCA 20% (Trichloroacetic acid; Sigma) and were kept on ice. Yeast cells wall was mechanically disrupted using glass beads (Cat. G8772, Sigma) and MagNA Lyser (Roche) for 30 seconds at 6000 rpm. Then 500 µl of TCA 5% were added and samples were centrifugated at 14000 rpm for 15 minutes. Surnatants were discarded and each pellet was resuspended using 100 µl of Sample Buffer (0.1M Tris-HCl pH 8.8, 10% glycerol, 0.1M DTT, 2% SDS, 0,001% Blu Bromophenol). Samples were incubated at 95°C for 5 minutes. Quantification was performed using Bradford assay[168] (Cat B6916,Sigma).

2.8.5 Yeast library screening

The yeast library screening is a method used to identify a large number of new binding partners of a protein; it is based on Y2H system described above. In this case, there is the bait protein fused to the Gal4-DBD and a library of prey proteins expressed fused to Gal4-AD. The Matchmaker® Gold Yeast Two-Hybrid library (Clontech) used, is a commercial library generated to perform the screening using yeast mating. According to this, library is expressed in Y187 yeast strain (MATa), conversely the bait protein must be expressed in a MATα yeast strain like Y2HGold. When bait and library (prey) fusion proteins interact, the DNA-BD and AD are brought into proximity to activate transcription of four independent reporter genes (AUR1-C, HIS3, ADE2 and MEL1). As result, positive clones selection is highly stringent.

Briefly, AUR1-C is a drug reporter encoding the enzyme inositol phosphoryl ceramide synthase. In response to protein-protein interaction, it is expressed and it confers resistance to Aureobasidin A (AbA). HIS3 transcription permits the cell to biosynthesize histidine and grow on – His selective medium. Similarly, the expression of ADE2 allows the yeast growth on -Ade selective medium. Instead, MEL1 encodes the α-galactosidase, an enzyme secreted by yeast cells into the culture medium. In presence of the chromogenic substrate X-α-Gal, yeast colonies turn blue whether there is protein-protein interaction.

When protein-protein interaction occurs, all the four reporter genes are activated and diploid cells generate blu colonies on selective medium without histidine and adenine, supplemented with Aureobasidin.

Library protocol

Library-screening protocol was divided in three phases: pre-mating, mating and post-mating phase. The pre-mating phase included set up experiments. Two-hybrid screening test was performed using a mate-and-plate human testis cDNA library (Cat. 630470, Clontech). Instead, the bait vector was obtained cloning the VHL30 cds sequence in pGBKT7 using in vitro In-Fusion® method. The recombinant clone was selected in bacteria and sequenced by Sanger method. After transformation in Y2H Gold strain, the correct expression of VHL30 protein was verified by Western blot. As following step, the bait was tested in order to exclude its autoactivation and its toxicity on selective media. Thus, Y2HGold expressing VHL30 was plated on SD - Trp, SD - Trp + X-α-Gal, SD - Trp + X-α-Gal + AbA. At this point a pilot experiment was performed in order to define the best mating conditions.

A small amount of testis library stored at - 80 °C was patched on SD – Leu plate. Some cells were inoculated in 50 ml of SD – Leu liquid medium; whereas yeasts expressing VHL30 were inoculated in 50 ml of SD – Trp liquid medium. When 0,8 OD600 was reached, both the cultures were centrifugated and pellets were resuspended in 1 ml of SD – Leu and 4 ml of SD – Trp, respectively. In a 3L flasks, the two strains were co-culturing in 45 ml 2X YPDA liquid medium at 30°C for at least 20h at slowly shaking (around 30-50 rpm). As suggested on the protocol, after 24 hr a drop of co-culture was checked under a phase contrast microscope (40X) to observe or not the presence of zygotes. The presence of zygote indicates that mating occurred. If zygotes were not visible, it was necessary to continue the incubation. In this case, after 24h and 30h the mated culture was collected. Pellets were resuspended and plated on SD -Trp, SD -Leu, SD -Trp -Leu and plates were incubated at 30°C for 3-5 days. The number of zygotes (diploids) was higher after 30 hours of mating; thus this incubation time was chosen to perform the real experiment.

During the mating-phase an overnight culture of pVHL30 expressing strain was crossed with 1 ml aliquot of cDNA human testis library. After 30 h at 30°C the mated culture was centrifugated and the pellet was resuspended in 10 ml of 0,5X YPDA liquid medium.

At this point, serial dilutions 1:10, 1:100, 1:1000 and 1:10000 of a small aliquot were plated on SD- Trp, SD – Leu, SD -Trp -Leu and incubated at 30°C for 3-5 days. The number of colonies were counted and used to calculate the number of the clones screened and the percentage of the mating efficiency. The remaining mated culture was plated on 150 mm SD - Leu -Trp + X-α-Gal + AbA (200 µl per plate) and plates were incubated at

30°C for 5 days monitoring colonies growth twice a day. Progressive numbers were used to mark blue colonies.

The post-mating phase corresponded to the identification of positive clones. To start, all positive clones (933) were patched out on more stringency medium QDO + X-α-Gal + AbA in order to avoid false positive. After this step the number of colonies was reduced to 607. Considering the high number of positive clones, they were classified in 3 groups according to dimensions and blue intensity. To facilitate the manipulation colonies were subdivided in 96 multi-well. DNA was extracted from yeast cells using the Zymoprep™-96Yeast Plasmid Miniprep (Zymo Research Cat D2005) and 2 µl of each DNA were used to transform TOP10 *E.Coli*. At this point 5 colonies mix of each transformation was analysed. In this case, plasmidic DNAs were extracted from bacteria cells using Zyppy ™ - 96 Plasmid Miniprep kit (Cat. D4041) and used as templates to amplify cDNA inserts with Gal4AD for and Gal4AD Rev primers. PCRs were performed according to Wondertaq protocol (Tables 10 and 11) and visualized on 0,8% agarose gel.

Components	**Volume**	**Final concentration**
Template	1 µl	-
Premix 5X	5 µl	1X
Forward primer (10µM)	1 µl	0.4 µM
Reverse primer (10µM)	1 µl	0.4 µM
Taq	0.25 µl	-
Nuclease free water	Up to 25 µl	-

Table 10 Wonder Taq Polymerase PCR mix components.

Temperature (°C)	**Step**	**Time**	**Number of cycles**
95	Denaturation	5 min	1
95	Denaturation	15 sec	35
55	Annealing	15 sec	
72	Extension	20 sec/kb	
72	Final extension	2 min	1
4	Cooling	forever	-

Table 11 Wonder Taq Polymerase protocol. Note that cDNA insert size are included in 0.5 kb to 4.0 kb, thus normally 1'20'' are set for the extension step.

Some clones showed more than one PCR product and were excluded from the identification process. On the other hand, positive clones with only one PCR product were transformed in Y190 yeast strain. Each clone was co-transformed with pGBKT7 empty vector (as auto-activation control) and with pGBKT7 VHL30 (as interaction control). For each transformation plate, a pool of colonies was patched out on permissive medium and on selective ones (SD- Leu -Trp -His + 30mM or 60 mM 3AT). Finally, using the Gal4AD For primer, all positive clones were sequenced by Sanger method.

2.9 Mammalian cell protocols

2.9.1 Cell transfection

For Co-ip experiments reported in the first chapter of results section, HEK293T were transfected using 4 µg of total DNA and Lipofectamine 2000 (Invitrogen) according to manufacturer's protocol. After 5 hours from transfection, cells medium was replaced with fresh DMEM complete. Conversely, in the others experiments HEK293T were transfected with polyethylenimine (PEI) linear MW 25 KDa (Sigma-Aldrich) according to well dimension. DNA:PEI ratio was 1:1 and the cells medium was replaced after 3 hours from transfection.

HeLa cells, used for immunofluorescence experiments, were transfected with Calcium phosphate (Sigma-Aldrich) using 3 µgr of total DNA/well. In this case cells medium was changed after 16 hours from transfection.

2.9.2 Co-immunoprecipitation

After 48 hours from transfection, samples were collected in 300 µl lysis buffer/well (20 mM HEPES-Na pH 7.4, 150 mM NaCl, 5 mM CHAPS)[169] supplemented with 1X protease inhibitors cocktail (PIC; Sigma). Lysates were centrifugated at 600 rpm for 15 minutes and the post-nuclear surnantants (PNS) quantified using Bradford or BCA assays. At this point, 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 for 1 h at RT. After 3 washes with TBS-T 0,05%, beads were incubated with 0,3 mg of cell lysate for 4 h at 4°C. Then, immune complexes were washed 3 times with 300 µl of lysis buffer and

eluted by incubating the beads 5 minutes at 70°C in 30 µl of 1X NuPAGE LDS buffer (Invitrogen) supplemented with 0.1 M DTT. Both input (1/20 of cell lysates immunoprecipitated) and Co-ip samples were visualized by Western Blot. For the Androgen Receptor (AR) experiments, samples were exposed to dihydrotestosterone (DHT) 10 nM or vehicle for 24h after transfection and before lysis. Co-immunoprecipitations were performed using anti-FLAG (Sigma, F7425) or anti-AR 441 (Santa Cruz, sc7305).

2.9.3 Protein extraction

After wash in ice-cold PBS 1X, cells were scraped in 100 µl of RIPA lysis buffer/well (25 mM Tris pH 7.5, 150 mM NaCl, 0.1% SDS, 0,5% sodium deoxycholate, 1% NP-40), supplemented with protease inhibitor cocktail (Roche). Each lysate was sonicated for 10 seconds at intensity 2, centrifugated at 14000 rpm for 15 minutes at 4°C. Surnants were collected and quantified using BCA assay. Finally, 20 µg of protein extract for each sample were visualized by Western blot.

Samples expressing the AR were treated with DHT 0.5 nM or vehicle for 24 hours before the protein extraction.

2.9.4 RNA extraction and cDNA synthesis

After wash in ice-cold PBS 1X, cells exposed to vehicle or DHT 10 nM for 24 hours, were scraped with 500 µl/well of Trizol Reagent (Invitrogen Cat. n 15596018), incubated on ice 2 minutes and 15 minutes at RT. 100 µl of Chloroform was added to each sample. Samples were mixed by inversion and centrifugated at 15000 rpm for 20 minutes at 4°C. The centrifugation allowed the phase separation; the aqueous one containing the RNA was transferred in a new tube. At this point an equal amount of isopropanol was added, samples were inverted and incubated on ice for 30 minutes. Finally, after centrifugation at 4 °C for 20 minutes, pellet washing with EtOH 70%, the RNA was resuspended in 30 µl of ddH_20 RNAse free/sample. Total RNA was quantified using Nanodrop 2000 (Thermo Scientific). RNAs were extracted from samples exposed to DHT 10 nM or vehicle for 24 hours before the extraction. At this point 2 µg of total RNA was retro-transcribed in cDNA using the Superscript ^{TM} III Reverse Transcriptase (Invitrogen Cat n 18080) and oligo $(dT)_{15}$. The reaction was assembled according to the manufacturer's protocol.

2.9.5 Real-time PCR

cDNA samples were diluted at 12,5 ng/µl and 1 µl was used as template in real-time PCR. This PCR was performed using Sybr Green (Sso Advanced™ Universal SYBR® Green Supermix,Biorad 172-5274) and reaction components for each well were assembled as indicated in the Table 12.

Components	Volume	Final concentration
Template	1 µl	-
Sybr green 2X	5 µl	1X
Primers Mix (50µM)	0,1 µl	0.5 µM
Nuclease free water	Up to 10 µl	-

Table 12 real-time PCR mix components.

The parameters for the amplification are reported in Table 13.

Temperature (°C)	Step	Time	Number of cycles
95	Denaturation	5 min	1
95	Denaturation	15 sec	35
55	Annealing	15 sec	
72	Extension	20 sec/kb	
72	Final extension	2 min	1
4	Cooling	forever	-

Table 13 Sybr Green protocol

Data were exported using Bio-Rad CFX Manager program and used to calculate a comparative ΔΔCT.

2.9.6 Transcriptional assay

HEK293T cells were transfected (PEI) with AR vectors, VHL isoforms vectors together with pARE-Luciferase and pRL-TK (TK promoter-Renilla luciferase construct as control). Then, cells were treated with DHT 0.5 nM or vehicle and 48h after transfection used to perform the assay. Cells were washed in PBS 1X at RT and lysed in passive lysis buffer (100 µl/well) coming from dual-Luciferase assay system (Promega). Cell lysates were used

to measure the luciferase activity with a luminometer (Tecan instrument). To start, 25 µl/well of luciferase reagent were added for measuring the firefly luciferase activity. Then, 25 µl/well of Stop-Glo reagent were added to measure the Renilla luciferase activity. Data were normalized to Renilla luciferase activity.

2.9.7 Protein turnover assay

HEK293T cells were transfected with 1 µgr/well of exogenous DNAs. After 24h from transfection, *de novo* protein synthesis was blocked with cycloheximide (CHX) 50 µM for different time points (i.e. 2h,4h,8h,12h). This treatment allowed to analyse the proteins degradation during the time. Thus, samples were collected at each time-point and then visualized by Western blot.

2.9.8 Immunofluorescence

HeLa cells were seeded on 13 mm glass coverslips and transfected with 1 µg/ well of DNA and Calcium phosphate. After 16 h, medium was changed. Then, cells were fixed with 4% paraformaldehyde (PFA) + 20% sucrose in PBS for 10 minutes after treatment with DHT 0.5 nM or vehicle for 16 hours. Washed three times with PBS, cells were permeabilized with 0,5% Triton X-100 in PBS for 10 minutes. At this point coverslips were blocked with blocking solution (3% BSA +0,1% Triton X-100 in PBS) for 1 h and then incubated with primary antibody o/n at 4 °C. Washed 3 times with PBS, coverslips were incubated with secondary antibody for 1 h and with DAPI for 15 minutes. Finally, after extensive washing, coverslip were mounted with commercial mounting solution and observed at fluorescence microscopy using 40X objective. Images were acquired using the Leica AS software. In these experiments the antibodies used were: antiAR 441 (SantaCruz- sc7305) diluted 1:500 and anti-mouse Alexa Fluor® 546 (Invitrogen - A11030) diluted 1:300. Antibodies were diluted in blocking solution.

2.10 Western blot

Protein samples were extracted as described before, boiled at 95 °C or 70°C for 5 minutes and then separated by SDS-polyacrylamide gel electrophoresis. Samples were loaded together a 5 µl of SHARPMASS ™ VII (Protein Marker MW 6,5-270 KDa, Euroclone) as marker for bounds size. The electrophoretic course was led at 120 V in Trizma-Glycin-

SDS Running Buffer 1X[170]. Then samples were transferred to a nitrocellulose membrane (Amersham™ Protan™ premium 0,2 µm – GE Healthcare Cat n 10600004) employing a costant voltage 350 mA for 90 minutes in Trizma-Glycin Transfer Buffer 1X. At this point, membranes were saturated using BSA-TBST 1% or milk-TBST 5% for 1h at RT and incubated o/n at 4°C with primary antibodies. The day after, they were washed three times with PBS and incubated with secondary antibodies 1h at RT. Two kind of secondary antibodies were used: antibodies conjugated with fluorescence molecules or antibodies conjugated with the peroxidase. For this reason, signals were revealed using Odissey Imaging System (LI-COR Bioscences) or using Immobilon™ Western substrate (Millipore, Cat n WI3KLS0500) and UVITEC machine as detector (Eppendorf). Antibodies used are listed in the Table 14.

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-Tubulin	1:5000	Abcam (ab21057)
Anti-Androgen receptor	1:5000	Abcam (ab9474)
Anti-Mouse Alexa Fluor® 546	1:10000	Invitrogen (A11030)
Anti-Rabbit Alexa Fluor® 488	1:10000	Invitrogen (A32790)
Anti-Mouse IgG (whole molecule)-Peroxidase	1:10000	Sigma (A9044)
Anti-Rabbit IgG (whole molecule)-Peroxidase	1:10000	Sigma (A6154)

Table 14 Antibodies for immunoblot experiments.

2.11 Computational analysis

2.11.1 VHL30 and MDM2 interaction

A network of interacting proteins was built from STRING[171] and visualized with Cytoscape[172] with medium confidence (0.400), no more than 5 and 10 interactors for the first and second shells were selected, while only text mining-, experiments- and databases-derived data were considered to reduce number of false positive interaction. The final number of nodes was 20, connected by 52 edges, with an average node degree of 5.2,

average local clustering coefficient of 0.721 and expected number of edges of 25. Human MDM2 (Accession number: Q00987) and pVHL (Accession number P40337) sequences were retrieved from Uniprot[173]and visualized with Jalview[174]. Orthologous sequences were retrieved from OMA[175] browser and aligned with a T-Coffee[176] (default parameters). Sequence features and secondary structure content were investigated with FELLS[177], functional domains mapped on the sequence using Interpro[178] and Pfam[179]. Linear motifs were predicted with ELM[180] and intrinsic disorder with MobiDB 3.0[181]. Ab initio protein structures prediction was performed with Rosetta 3.8[182] using specific protocol to model intrinsically disordered regions[183]. For each fragment 5,000 decoys were generated and clustered using Rosetta clustering module. DSSP[184] was used to calculate secondary structure content for each 3D structure model, while networks of interacting residues were predicted with RING 2.0[185].

2.11.2 cDNA fragments analysis

The library screened cDNA fragments were translated using ExPASy[186] translate (https://web.expasy.org/translate/) to identify the correct proteins open reading frame. For each clone, vector sequence upstream the cDNA fragment was chosen as reference to select the correct reading frame. Resulting amino acidic sequences were used to perform a search against UniprotKB/ Swiss-prot database using BLASTP with default parameters and filtered for *Homo sapiens* (TaxID 9606). Whether identification process identified no significant results, cDNA fragments were analyzed using tBLASTx (https://blast.ncbi.nlm.nih.gov) against either the Human RefSeq Gene or nucleotide collection (nr/nt) as database, whereas filtering for *Homo sapiens* was maintained in all searches. Identified proteins were used to generate a functional protein-protein interaction network using Cytoscape[172] and retrieving interaction data from STRING[171]. Due to low intra-network connectivity, network was expanded including interactors-of-interactors. In detail, query proteins were maintained in the first shell and >50 in the second shell imposing confidence of 0.35. Text-mining derived data was manually filtered and excluded from the analysis.

2.11.3 AR proline hydroxylation prediction

The sequence of human AR was retrieved from the UniProt databank (accession ID P10275). Only the canonical sequence was included in this analysis. Orthologous sequences of AR were retrieved from the OMA[175] browser database and used to perform a multiple sequences alignment with T-Coffee[176]. Conserved proline residues were predicted for hydroxylation using a consensus strategy based on multiple predictors, such as iHyd-PseCp[187], OH-PRED[188] and RF-Hydroxysite[189]. Cancer-related mutations affecting putative hydroxylation sites were retrieved by COSMIC[190] database. Sequence feature and disorder content were predicted with FELLS[177].

Results

3 – Results

My PhD activity aimed at characterizing the von-Hippel Lindau tumor suppressor. In particular, a combination of experimental and bioinformatics approaches was used to shed light on pVHL30 non-canonical functions. Results are organized in four different sections as following. In the first two sections (3.1 and 3.2) I present data about pVHL and proteins involved in cell cycle regulation.

3.1 The E3 ubiquitin-protein ligase MDM2 is a novel interactor of the von Hippel-Lindau tumor suppressor

At physiological conditions, pVHL is translated in two main isoforms: the full-length pVHL30 and a shorter pVHL19 lacking 54 residues forming the N-terminal tail[24,29]. Both isoforms act as tumor suppressor and overlap in modulating the adaptive response to hypoxia by regulating the hypoxia-inducible factor 1-alpha (HIF-1α) stability[191]. For this reason, since their identification, the two pVHL isoforms were considered as overlapping in functions. As of 2014 however, it was first proposed pVHL30 to form an isoform-specific interaction with p14ARF[39]. This evidence suggested a functional asymmetry between pVHL products and a relevant role for the pVHL30 N-terminal tail in sustaining further protein-protein associations.

The p14ARF tumor suppressor exerts its function mediating regulation of p53. It sequesters MDM2 (mouse double minute 2 homolog) into the nucleoli thus preventing p53 degradation[192]. Similarly, pVHL is also thought to induce p53 stabilization. In zebrafish model it was demonstrated that MDM2 degradation is regulated upon pVHL interaction with PDCD5 (programmed cell death 5 protein), a negative MDM2 regulator[143]. All of these evidences suggest a functional connection between pVHL and the p53-MDM2 at the pathway level. I wondered whether pVHL may sustain a novel direct association with MDM2. To tackle this question, a combined computational and experimental approach was designed.

Protein-protein interacting network identification *in silico*

An interaction network centered on pVHL, HIF-1α, p14ARF, p53 and MDM2 was generated using STRING[171]. The resulting network is composed by 20 nodes connected by 52 edges, each of them representing experimentally validated interactions (Figure 11). The network is significantly more connected than expected by chance, suggesting these

proteins to be functionally interconnected. The network can be divided in three different clusters linked to different biological processes (Figure 11).

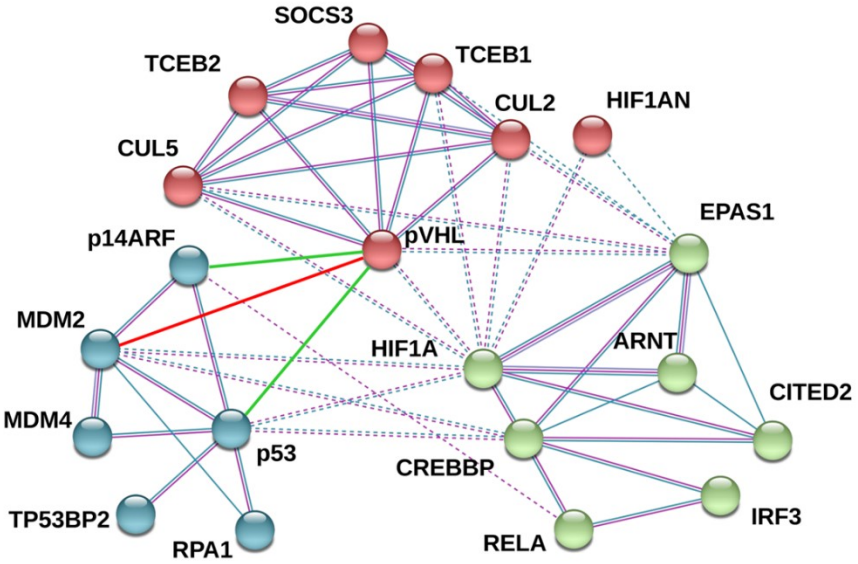


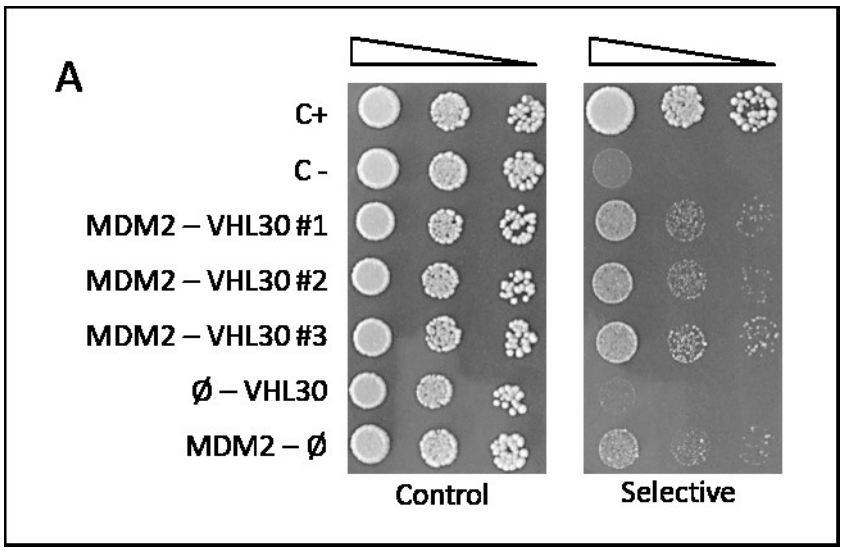
Figure 11 Network of interacting proteins. Connections between nodes represent experimental evidence for interaction, while colored nodes are for proteins forming the network. Green edge represents the interaction between pVHL30 and p14ARF, currently not yet reported in STRING, while the red line is for the predicted association of pVHL30 with MDM2.

The first cluster (blue nodes) composed by six proteins linked to apoptosis, includes: p53, MDM2 and p14ARF. Gene ontology terms predicts this cluster to be mainly localized in the nucleus and participating in the regulation of transcription in response to hypoxia and cellular stresses. The second cluster (red nodes) includes seven proteins, among them pVHL, linked to protein ubiquitination. The last cluster (green nodes) is the most heterogeneous including proteins associated to both hypoxia- and stress-response, e.g. nutrient deprivation and viral aggression. Of note, several experimentally validated interactions among all these three clusters exist, suggesting the functional interconnection between the hypoxia response and the cell cycle regulation. Manually curated enrichment of protein-protein interactions across network nodes further increases the connection between clusters I and II. Since interactome studies of binary protein-protein interaction often reveal also interaction between "interactors of interactors"[193] and the already validated physical interactions between pVHL, p53 and p14ARF, I decided to validate the possible physical connection between pVHL and MDM2 experimentally.

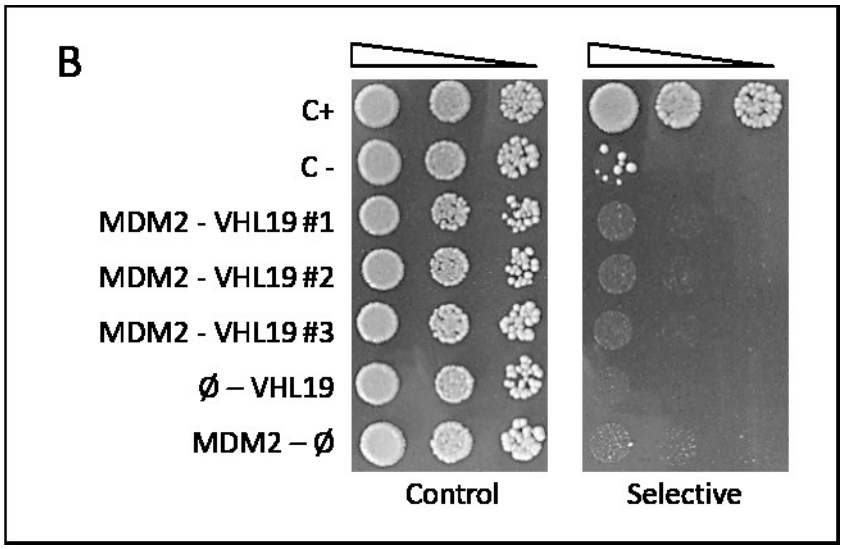
MDM2 establishes an isoform-specific interaction with pVHL30

The predicted association between pVHL and MDM2 was first validated using yeast two-hybrid assays (Y2H). In particular, I used Y190 yeast strain to express pVHL30 or pVHL19 as bait fused to the Gal4 DNA-binding domain (DBD) while MDM2 as prey fused to the Gal4 activation domain (AD). Upon protein expression check (Figure S1), I used at least three independent colonies of each transformation to evaluate the interaction by spot assay. In all experiments yeast cells were spotted on permissive and selective medium. As mentioned in material and methods section, interaction between bait and prey proteins allows the activation of *HIS3* gene reporter leading to the yeast growth on selective medium without histidine and supplemented with 3-AT. To evaluate yeast growth, in each experiment one positive (C+) and one negative (C-) control were spotted as reference. The positive yeast strain expressed Gal4 DBD-murine p53 and Gal4 AD-SV40 large T antigen, two proteins strongly interacting. As marker of binding absence, negative control carrying the Y2H empty vectors was used.

Yeast expressing pVHL30 or MDM2 alone correctly failed growing on selective conditions, excluding their ability to independently activate the gene reporter transcription (Figure 12A). Contrarily, when co-expressed both pVHL30 and MDM2, yeast was able to grow on selective medium thus demonstrating the *in vivo* association of the two proteins in the yeast model. Although Y2H data cannot be considered strictly quantitative, these findings suggest a weak interaction as yeast growth rate was less sustained when compared with positive control.



The Y2H assay was also performed expressing the shorter isoform pVHL19. In this case, any interaction was detected (Figure 12B). This result suggests an important role of the pVHL30 N-terminal tail in sustaining this association.



.

Figure 12 Identification of pVHL30-MDM2 interaction by Y2H. Serial dilution of yeast cells were spotted on both permissive (left) and selective (right) medium. C+ and C- corresponds to positive and negative controls. Ø indicates an empty vector (i.e. expressing only the Gal4-AD or the Gal4-DBD) used to control the absence of auto-activation of the paired fusion proteins. The image is representative of three independent experiments, each with 3 different clones analyzed. (A) Y2H assay of pVHL30 binding to MDM2 – (B) Y2H assay of pVHL19 binding to MDM2.

To validate the pVHL30-MDM2 interaction in a more physiological context, Co-immunoprecipitation (co-IP) experiments were performed in human cells transiently over-expressing both proteins of interest. HEK293T cells were transfected with recombinant plasmids expressing MDM2 Flag-tagged and VHL30 (or VHL19) HA-tagged. Total cell lysates were immunoprecipitated using anti-FLAG antibody against Flag-MDM2 protein. As reported in the Figure 13, co-IP experiments agree with Y2H-derived results. The ability of pVHL30 to associate MDM2 was confirmed (Fig.13 panel A), whereas pVHL19 failed to interact with MDM2 (Fig. 13 panel B). In particular, pVHL19 co-IP was not specifically driven by MDM2, as the anti-flag antibody also immunoprecipitated pVHL19 expressed alone and there is not an enrichment of the immunoprecipitated pVHL19-HA in the presence of MDM2-Flag protein. Collectively, these findings indicate that pVHL30 and MDM2 can form at least binary complexes in human, notably in kidney cells. Interestingly, only pVHL30, the full-length isoform of the protein, associates with MDM2 suggesting an isoform-specificity of this interaction and a key role of the pVHL N-terminal portion in MDM2 binding.

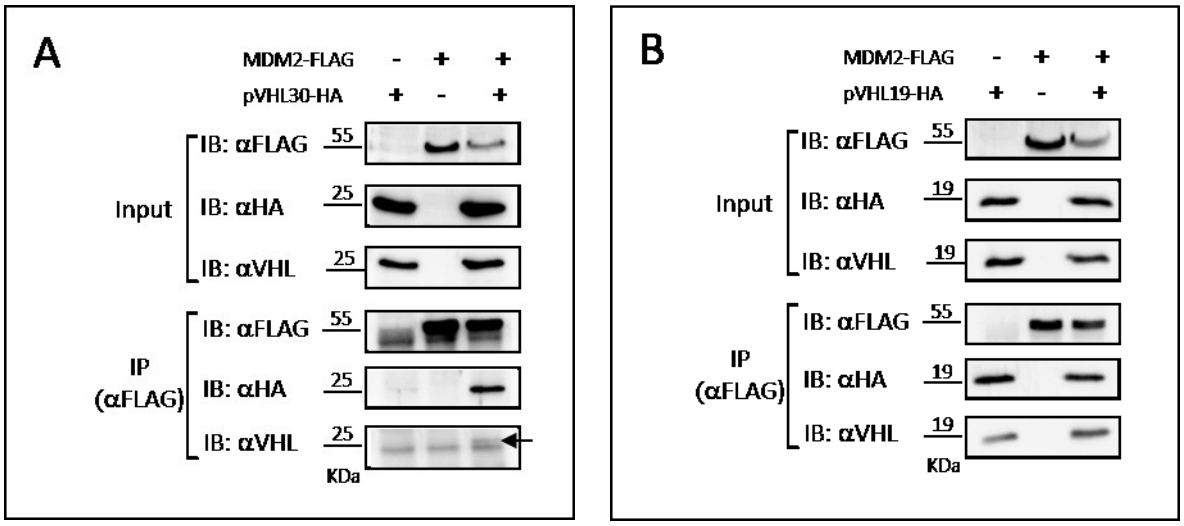


Figure 13 Validation of pVHL30-MDM2 interaction in cells. Human HEK293T cells were transiently transfected with plasmids overexpressing Flag-tagged MDM2 proteins and/or HA-tagged pVHL proteins, as indicated on the top row. Recombinant proteins have been revealed in total cell lysates (Input) by immunoblotting with either anti-HA, anti-VHL, or anti-Flag antibodies. Upon MDM2 immunoprecipitation with anti-FLAG antibody, presence of the VHL proteins in the immunoprecipitates (IP) were finally verified using the anti-HA and anti-VHL antibodies. (A) Co-ip assay of MDM2 and pVHL30 – (B) Co-ip assay of MDM2 and pVHL19

Molecular dissection of pVHL30 binding region

As reported in the last revision of the VHL database, the VHL protein is able to interact with more than 500 partners[32] suggesting the coexistence of multiple binding surface[33], each responsible for specific protein-protein interactions. To identify the pVHL30 region involved in MDM2 binding, I generated multiple plasmids expressing different truncated forms of pVHL30 protein and each fragment was tested for its ability to interact with MDM2 in Y2H assay (see scheme in Fig. 20B). Firstly, three fragments were generated in according to pVHL30 domain organization, consisting of a pVHL N-terminal portion (aa 1-53), the so-called beta-region (aa 54-157) and C-terminal alpha domain (aa 158-213). As can be seen in Figure 14, none of these isolated domains is able to interact with MDM2 in Y2H assay. These results suggested that the pVHL30 N-terminal tail is essential but not sufficient for MDM2 binding. The N-terminal tail of pVHL is classified as an intrinsically disordered region, so I performed an analysis with FELLS[177], looking at the disorder content of the protein. As reported in Supplementary materials (Fig S2) the N-terminal disordered portion can be extended to the first 70 amino acid residues of pVHL. Thus, I generated a plasmid expressing this enlarged N-terminus (aa 1-70) and tested its ability to

bind MDM2 in Y2H assay. As showed in figure S2, yeast co-expressing pVHL N-terminus (aa 1-70) and MDM2 failed to grow on selective medium, indicating no interaction.

All these evidences indicate that pVHL30 – MDM2 association is cooperatively mediated by the pVHL disordered N-terminal portion in association with other pVHL domains.

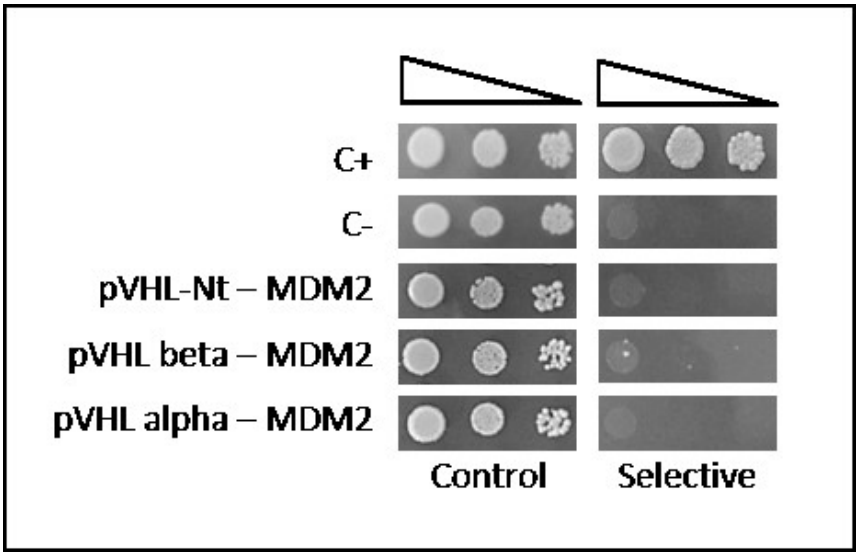


Figure 14 Molecular dissection of pVHL30 binding region. Yeast cells co-expressing MDM2 together with truncated forms of pVHL were serially diluted and spotted on permissive (left) and selective (right) medium and incubated for 6 days at 30°C. C+ and C- are positive and negative controls. The image is representative of three independent experiments, each with 3 different clones analyzed. The entire subset of data generated for one experiment is reported as Supplementary materials (fig. S3). N=3 pVHL N-terminus (aa 1-53) – pVHL beta portion (aa 54-157) – pVHL alpha domain (aa 158-213)

Molecular dissection of MDM2 binding region

MDM2 was reported to associate with a number of different functional partners[194]. In order to identify the MDM2 region sustaining pVHL30 binding, I firstly generated Y2H recombinant plasmids expressing three different MDM2 fragments (see scheme Fig. 20A) corresponding to MDM2 N-terminal portion (aa 1-150), the central part (aa 151-350) and the MDM2 C-terminus (aa 351-491). Each plasmid was co-transformed with pVHL30 expressing plasmid in yeast and assayed in Y2H. Spot assays showed the inability of both the MDM2 N-terminal and central region to interact with pVHL30, whereas the MDM2 C-terminal portion maintains its ability to bind pVHL30. Indeed, as reported in figure 15, only the expression of MDM2 C-terminal portion allows yeast viability on selective medium demonstrating that an interaction occurs.

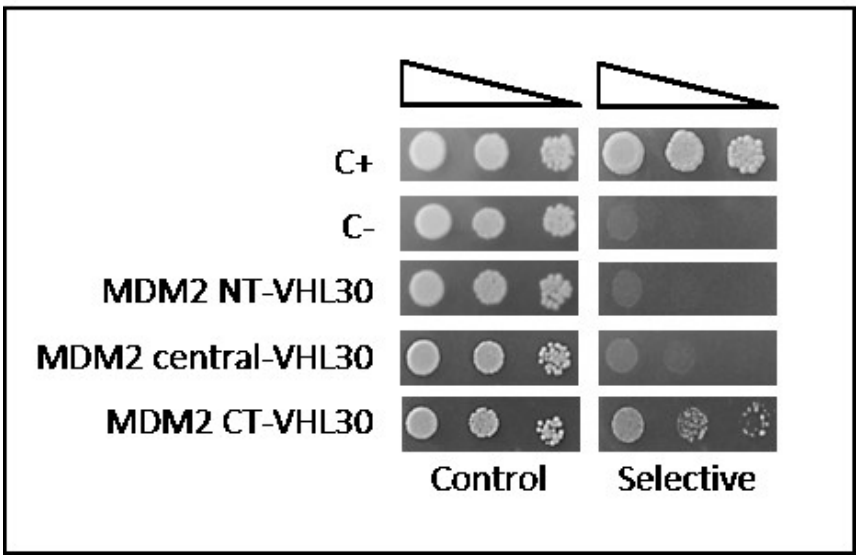


Figure 15 Molecular dissection of MDM2 binding region. Yeast cells co-expressing VHL30 together with mutants of MDM2 were serially diluted and spotted on permissive (left) and selective (right) medium. C+ and C- are positive and negative controls. The image is representative of three independent experiments, each with 3 different clones analyzed. The entire subset of data generated for one experiment is reported as Supplementary materials (fig. S4). N=3 MDM2 N-terminus (aa 1-150) – MDM2 central part (aa 151-350) – MDM2 C-terminal (aa 351-491)

Identification of the minimal protein region supporting MDM2 – pVHL30 binding

At this point, the molecular dissection established the pVHL30 N-terminal tail as necessary but not sufficient for MDM2 binding, as well as the MDM2 C-terminal (aa 351-491) portion relevant for this interaction. Several Y2H assays were additionally performed to identify a possible MDM2 specific motif involved in pVHL binding. A more accurate analysis of MDM2 C-terminal sequence showed hallmark of an intrinsically disordered region spanning residues 329 to 433. The disordered region terminates before the RING domain (aa 434-491), a well-known C3HC4-type Zinc finger, known to mediate protein-protein and protein-RNA interactions of MDM2[195].

To understand whether the MDM2 disordered part or the structured ones have a key role in pVHL binding, I expressed all the different sub-fragments of the MDM2 C-terminal region described in Fig. 20A and tested their ability to bind pVHL30 in Y2H assay. As reported in figure 16, yeast co-expressing the MDM2 RING domain (aa 434-491) and pVHL30 were unable to grow on selective medium. This observation excluded the involvement of the RING domain in pVHL30 binding. Similarly, also the MDM2 disordered portion (aa 351-433) alone failed to interact with pVHL30 as confirmed by absence of yeast growth on selective conditions (Fig 16). However, the MDM2 disordered portion restored its ability

to associate with pVHL30 when the C-terminal end of this fragment (aa 351-433) was enlarged toward the first part of the RING domain (aa 351-451). Indeed, yeast co-expressing pVHL30 and MDM2 fragment (aa 351-451) survived as shown in figure 16. Taken together, data suggested that at least a portion of MDM2 RING domain can participate and/or modulate pVHL30-MDM2 binding.

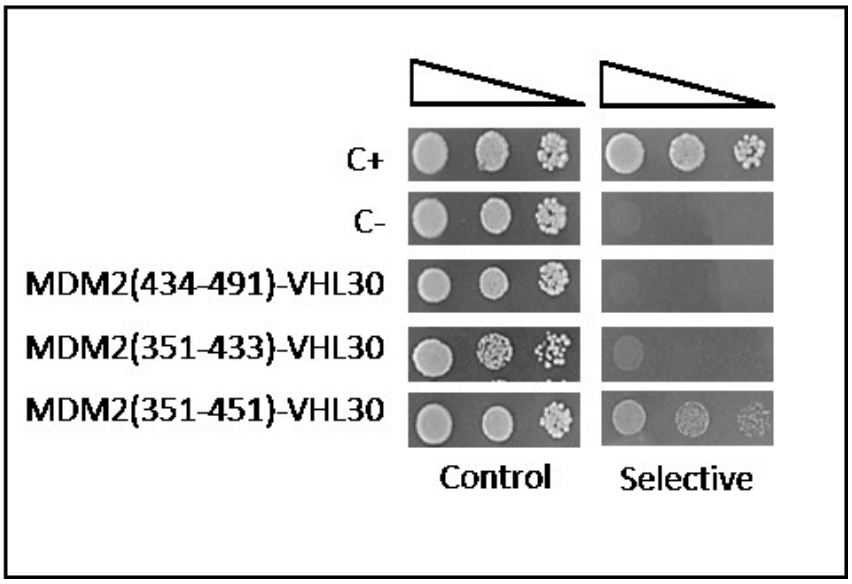


Figure 16 Dissection of MDM2 C-terminal region. Yeast cells co-expressing VHL30 together with mutants of MDM2 were serially diluted and spotted on permissive (left) and selective (right) medium. C + and C- correspond to positive and negative controls. This picture is representative of nine independent experiments, three for each MDM2 fragment. The entire subset of data generated for one experiment is reported in Fig S5. N=3

The MDM2 disordered region is poorly characterized, with no functional data available in the literature. Thus, an *in silico* characterization with FELLS[177] was performed. The pVHL30 N-terminal tail, which is involved in the interaction, is mostly acidic and characterized by the presence of eight GxEEx repetitions[39]. Considering this aspect, I wondered whether clusters of positively charged residues within the MDM2 disordered region can mediate pVHL30 binding.

As reported in figure 17, characterization *in silico* identified a region spanning residues 329-350 as a good candidate for pVHL30 binding. This region is predicted as fully disordered, with a typical alternation of charged residues. Of note, the position of five lysine residues in the sequence suggested these to form a positively charged cluster, which could form electrostatic interactions with the acidic pVHL30 N-terminal tail.

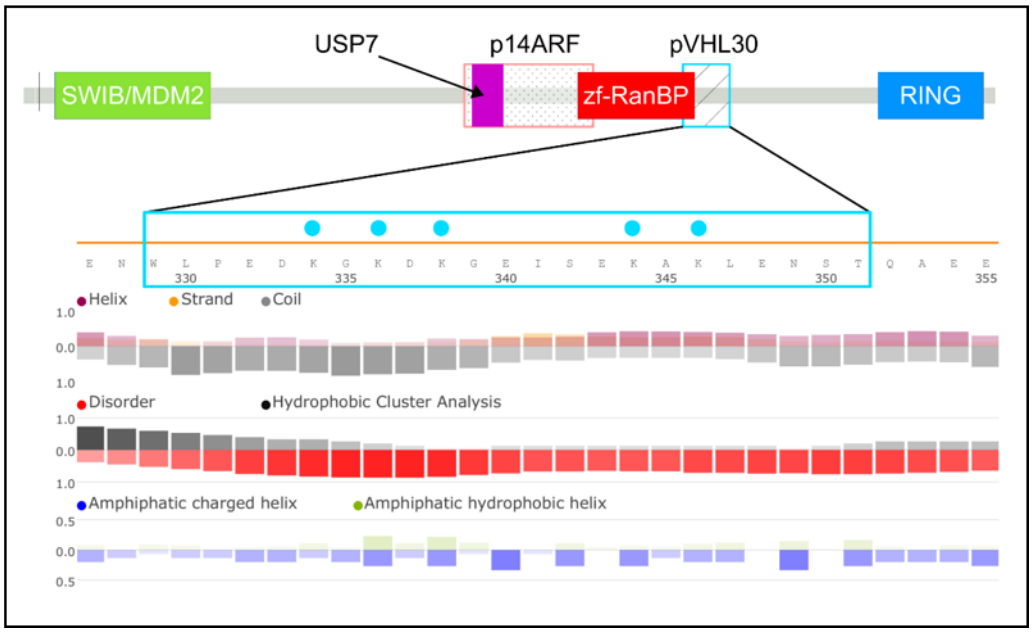


Figure 17 Graphical overview of MDM2 sequence features. The different binding domains are presented as colored rectangles. Zoomed view and FELLS representation of the binding motif driving pVHL30-MDM2 association. Blue dots represent positively charged residues.

To validate these predictions, new recombinant plasmids were prepared, checked for proteins expression (fig S1) and tested for their ability to bind pVHL30.

Y2H assays showed that MDM2 fragment 329-433 interacts with pVHL30, indeed as reported in figure 18, yeast co-expressing the two proteins survive on selective medium. This observation confirmed the relevance of MDM2 residues 329-350 for binding, as also suggested by the loss of pVHL30 binding upon its removal in the fragment 351-433 (see Fig. 16). On the other hand, the stretch aa 329-350 *per se* was not able to sustain the binding, as confirmed by the inability of the MDM2 central portion to associate pVHL30 (fig. 15).

At this point, data indicated that different elements in MDM2 disordered region can cooperatively contribute to pVHL30 binding. I expressed a longer MDM2 fragment (aa 329-451) including the portion of the RING domain previously shown implicated in the interaction, and tested its ability to bind pVHL30. As reported in figure 18, the co-expression of this fragment was found to improve the growth rate of yeast on selective conditions, supporting the idea that residues 329-350 and residues 434-451 are elements both important and necessary to improve the binding stability. To confirm this hypothesis, I further compared the growth rate in the same plate of yeast co-expressing pVHL30 and three different MDM2 fragments, i.e. residues 329-433 or 329-451 or 351-451. As showed in figure 18, yeast cells expressing MDM2 fragment (aa 329-451) grow better compared to

the other ones, confirming the role of the flanking sequences to the binding stability or regulation.

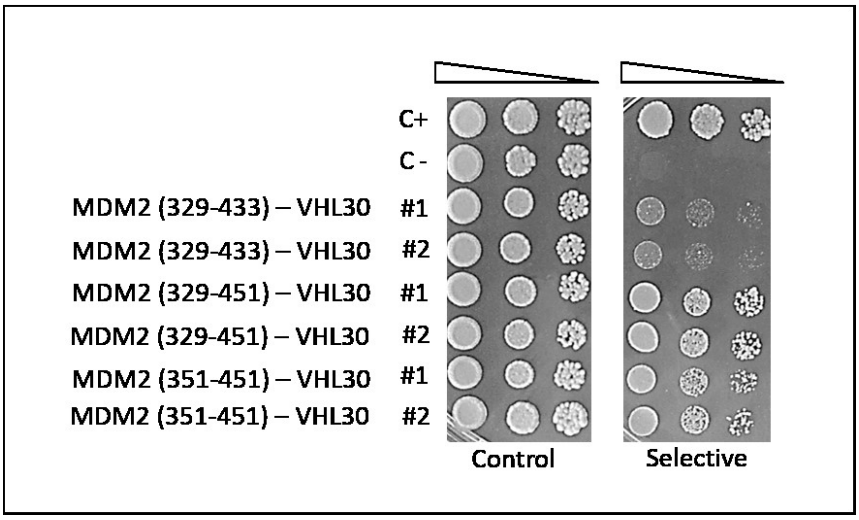


Figure 18 MDM2 fragments binding properties. The growth of two independent clones for each MDM2 interacting fragment is showed on the same plate to better appreciate semi-quantitative differences in the pVHL30 binding. Yeast cells co-expressing VHL30 together with mutants of MDM2 were serially diluted and spotted on permissive (left) and selective (right) medium. C + and C- correspond to positive and negative controls.

As described above, the pVHL30 dissection suggested the involvement of both N-terminus and other pVHL domain in MDM2 binding. To verify it, I generated a pVHL mutant corresponding to N-terminus and beta domain (aa 1-157), checked the protein expression (fig S1) and tested its interacting ability. Interestingly, Y2H showed (fig 19) that yeast co-expressing this pVHL mutant and the MDM2 fragment 329-433 were able to survive on selective medium. This evidence confirmed that both the N-terminal tail and beta domain of pVHL30 are strictly required for MDM2 binding.

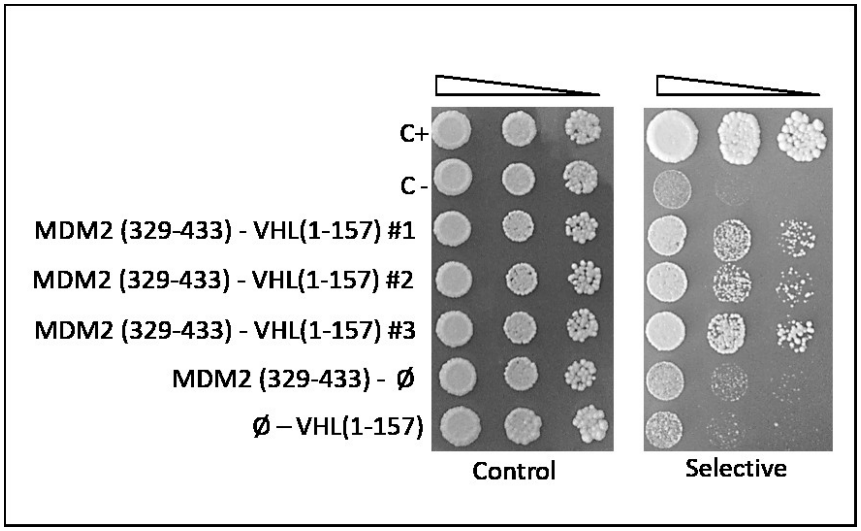


Figura 19 (Previous page) Minimal interacting fragments. Yeast co-expressing MDM2 and VHL fragments were spotted on permissive (left) and selective (right) medium. C+ and C- correspond to positive and negative controls. This picture is representative of three independent experiments.

To summarize, experimental data indicate that the interaction occur between the MDM2 C-terminal region and the pVHL30 N-terminus and beta domain. In particular, the binding is cooperatively mediated by multiple interaction motifs and by an electrostatic component. The interaction motifs are located within the MDM2 disordered portion (aa 350-433) and the pVHL beta portion. Instead, the electrostatic force occurs between the MDM2 329-350 residues positively charged and the pVHL N-terminus negatively charged region. The pVHL30-MDM2 association is also positively regulated by the first part of MDM2 RING domain. A schematic summary of all interactions tested is shown in figures 20A and 20B.

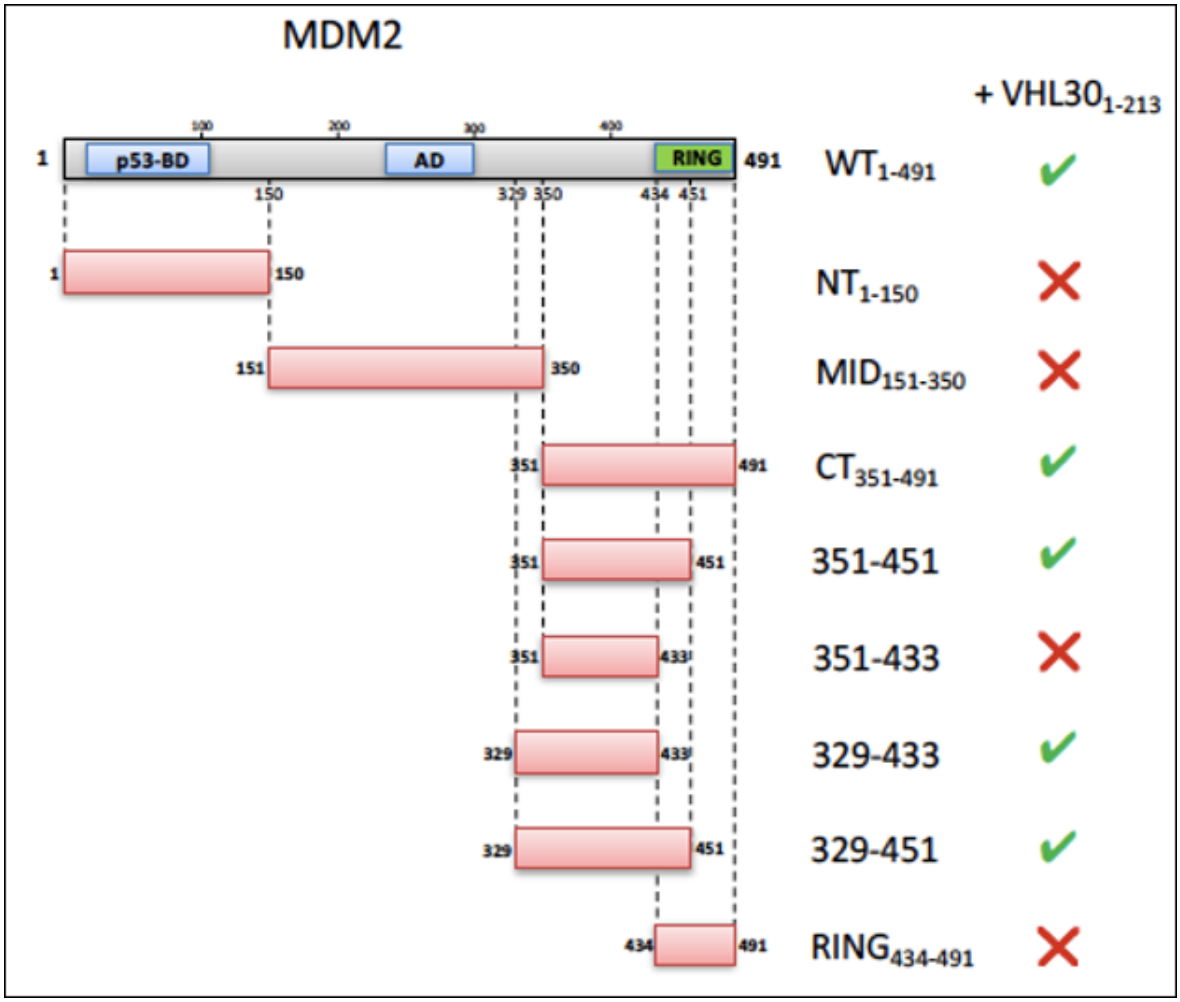


Figure 20A Schematic view of the MDM2 fragments tested for interaction by Y2H.

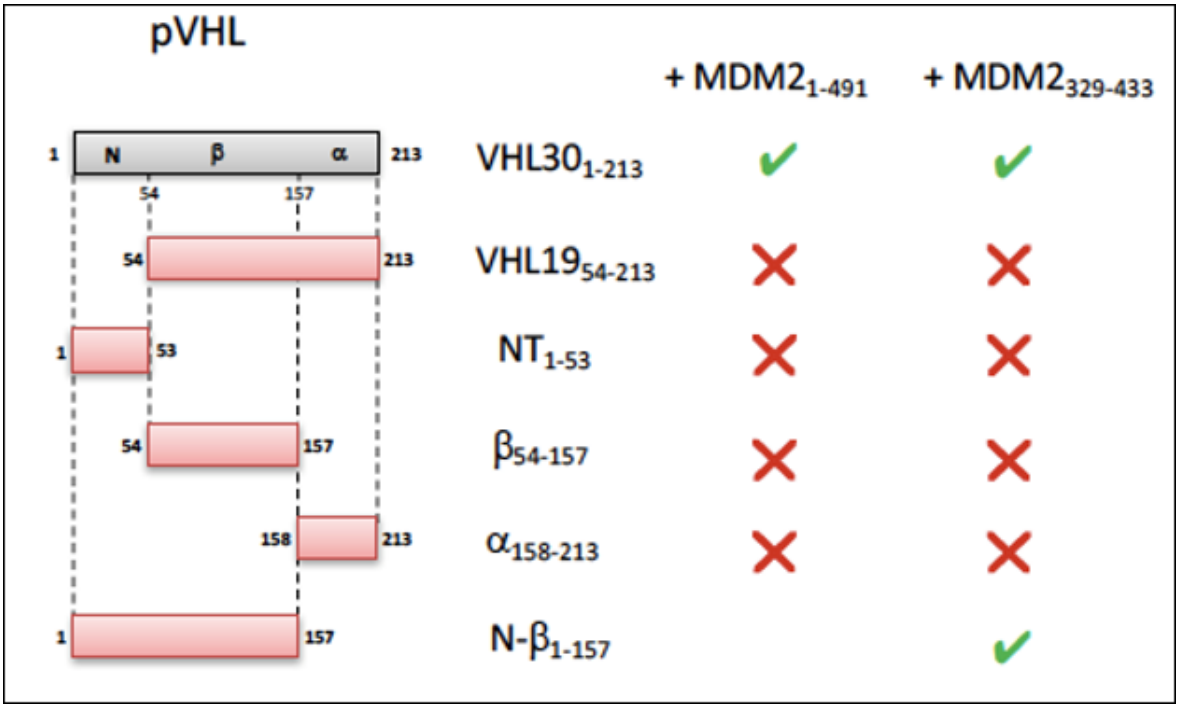


Figure 20B Schematic view of the pVHL30 fragments tested for interaction by Y2H.

Ab initio structure prediction and intrinsically disorder flavors characterization of MDM2 fragments

The MDM2 region involved in pVHL30 binding is poorly characterized. Intrinsically disordered regions, lacking of a fixed tertiary structure, are often associated to protein cellular functions depending on chemical property coded in their sequence[14,15]. Thus, I decided to investigate whether structural elements in MDM2 can justify the different binding behaviors showed in Y2H assays. To this aim, an *ab initio* structure prediction and disorder flavors characterization were performed for all the four MDM2 fragments localizing within the disordered portion. For simplicity, I will rename the MDM2 portions 351-451, 351-433, 329-433 and 329-451 respectively fragments 1, 2, 3 and 4 (figure 21). Multiple predictions for each MDM2 fragments (~ 5000 decoys) show these region to maintain a disordered behavior, however relevant differences in secondary structure content were also predicted. In particular, the MDM2 fragment 1 corresponding to the minimal MDM2 portion sustaining association with pVHL30, is characterized by an high content of secondary structure. On the other hand, all of MDM2 fragments are predicted to share a small beta hairpin followed by a variable number of alpha bulges, with the only notable exception for fragment 3 which is predicted to lack the beta hairpin, whereas maintaining the alpha bulges. Fragment 4 contains few secondary structure elements and in

general it shows a full disordered conformation. Contrarily, fragment 2 which is unable to associate pVHL30, is predicted as the most compact. This analysis indicated that alpha bulges are conserved both in interacting and not interacting MDM2 fragments. These bulges localize in the stretch 351-433 with flanking region more frequently predicted as coil (fig 21). A detailed characterization of residue-residue interactions for each model was also performed with RING[185]. The MDM2 fragment 2, unable to bind pVHL30 showed the highest hydrogen bonds, Van der Waals and ionic interactions.

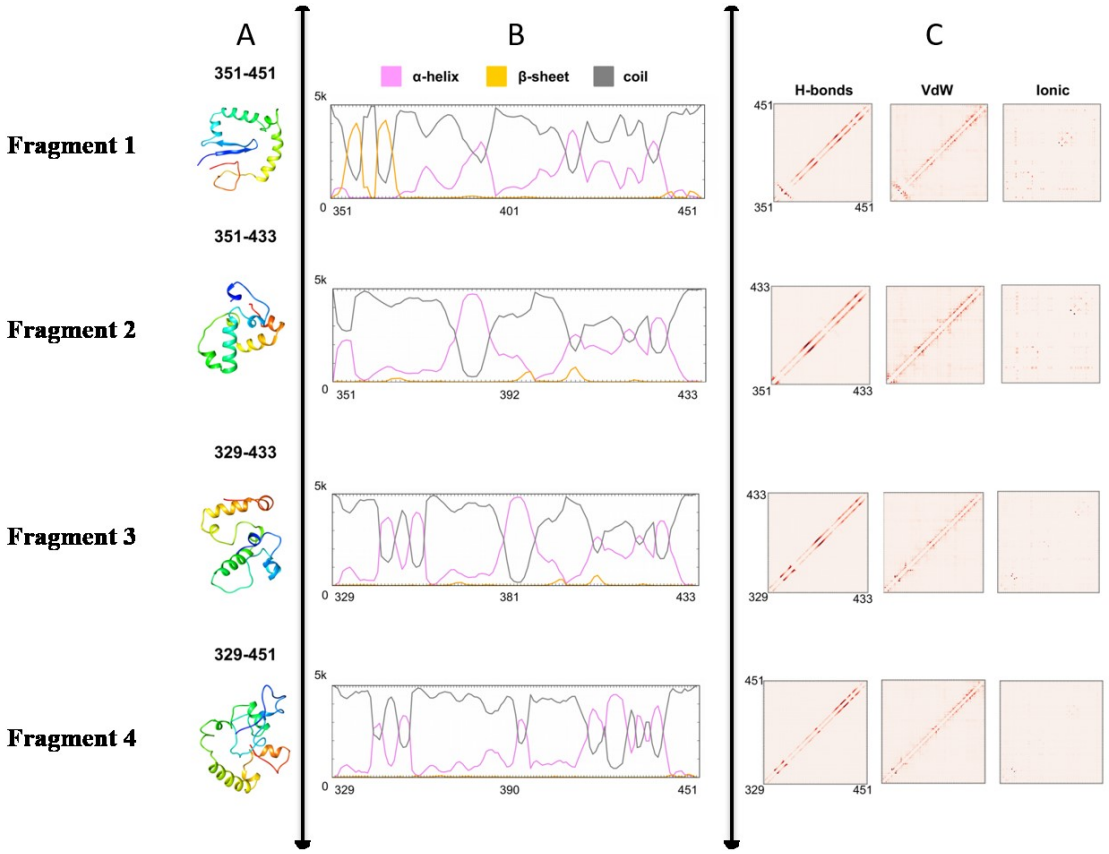


Figura 21 Characterization in silico of structural features of MDM2 fragments The most probable 3D structures are presented as cartoon and colored rainbow with blue and red marking the N- and C-terminus respectively. Secondary structure content calculated analyzing 5000 predictions is presented as plot. Intra-chain interactions describing network of interacting residues and type of interactions are presented as boxplots. (A) Ab inizio structure prediction – (B) secondary structure content – (C) conserved intra-chains interactions

This finding suggested that this MDM2 fragment, when isolated from flanking regions, is prone to assume a globular-like conformation. On the contrary, the others MDM2 pieces are characterized, on average, by comparable hydrogen bonds and a lower number of Van der Waals and ionic interactions. These characteristics lead to fragments more prone to

assume an extended conformation albeit being rich in secondary structure element. Taken together, all the observations indicate that shortening MDM2 fragment 4 removing the flanking regions, stabilizes the polypeptide reducing the intrinsically disordered behavior and putatively triggering a conformational switch to order.

Considering the relevance of the MDM2 fragment 4 in pVHL30 binding, an evolutionary study was also performed. In particular sequences of MDM2 fragments 4 from several organisms were aligned. As showed in Supplementary materials (fig S6) the region positively charged (aa 329-351) is maintained throughout evolution and mostly conserved in Artiodactyla and Primates. Since the number of acidic repeats in pVHL30 N-terminal tail is known to increase from Rodents to Primates[39], our data suggest that the isoform-specific interaction pVHL30-MDM2 could be specifically evolved in upper mammals. Notably, the MDM2 portion between residues 351 and 433 contains six ATM (ataxia-telangiectasia mutated kinase) phosphorylation sites (i.e. S386, S395, S407, T419, S425, S429). Interestingly, this evidence suggests that post translational modifications could operate as further regulator of MDM2 and pVHL30 association.

Discussion

The decision to investigate a possible physical interaction between pVHL30 and MDM2 started from the observation of p14ARF association with pVHL30 isoform that first showed a functional asymmetry in pVHL isoforms specialization[39,196], as well as representing another link between pVHL30 and p53 pathways. Functional interplay between oxygen sensing and cell cycle regulation was proposed by different authors[197,198]. In particular, both pVHL and HIF-1α act inducing p53, while hypoxia induced apoptosis requires a functional copy of both p53 and HIF-1α.[199]. Furthermore, it has been reported that MDM2 level decreases[200] in hypoxia conditions. In this chapter, I described a novel direct interaction between pVHL and MDM2. This association is isoform-specific, with experiments *in vivo* in both yeast cells and mammalian cells indicating that only pVHL30 is able to bind MDM2. We also found that the intrinsically disordered pVHL30 N-terminal tail has a key role in binding MDM2. Further findings showed that other domains of pVHL also mediate this interaction, as the N-terminal tail was found to be necessary but not sufficient alone. In particular, Y2H data highlighted that MDM2 binding requires the contribution of the pVHL beta domain, a well-known structural element mediating multiple protein-protein interactions[32], e.g. association with HIF-1α transcription factor. From literature, it is known that HIF-1α associate with MDM2[201] yielding hypoxia-induced

p53 stabilization and activation. Considering the novel association here presented, it can be speculated that proteins involved in cell cycle regulation and hypoxia sensing may form transient macromolecular complexes which are constantly formed and destroyed in different cellular compartments. Ideally, under mild hypoxia the amount of free pVHL increase allowing the pVHL30 functional re-localization with consequent MDM2 association. Referring to figure 17, the MDM2 region involved in pVHL30 binding is downstream from two overlapping binding sites for p14ARF (aa 210-304) and USP7 (aa 242-331), a deubiquitinating enzyme involved in p53 stabilization[202]. The MDM2 region (aa 351-433) contains multiple phosphorylation sites for the ATM Kinase known to be relevant for triggering the DNA damage response[203]. Data indicate that the pVHL30 binding is mediated by at least two linear interacting peptides (LIPs)[204] present on MDM2 and probably modulated by either indirect (e.g. variation of oxygen concentration) and direct (e.g. site-specific phosphorylation) mechanisms. All the evidences suggest that MDM2 could be part of transient macromolecular complexes together with pVHL30, p14ARF, HIF-1α and p53. In particular, it was proposed that the interaction of pVHL30 with p14ARF leads to ARF nucleolar translocation causing pVHL dissociation from pVHL-elonginC-elonginB (VCB) complex[196]. The pVHL30 fraction unbound can interact and stabilize p53 through its α-domain, inhibiting MDM2-mediated p53 degradation and its nuclear export[142]. A similar mechanism is also involved in p14ARF-dependent functional inhibition of MDM2[192]. In this study, I demonstrated that both the N-terminal tail and β-domain are necessary for MDM2 association suggesting that different regions of the same protein participate to MDM2 binding as already observed among linear motif-mediated protein-protein interactions[180]. The binding of pVHL30 to MDM2 may cover ATM-specific phosphorylation sites representing another way to regulate MDM2 activity. Indeed, it is known that ATM induces p53 accumulation and activation by regulating MDM2 oligomerization[203], while the pVHL30 association with ATM impairs MDM2-mediated ubiquitination and p53 nuclear export[142]. Considering that pVHL30 is a protein-hub able to interact with more than 500 partners[32], it is conceivable that it can participates as component of several ternary complexes formed under different conditions. Another interesting point is the presence in the same region of an USP7 binding site. USP7 is an hydrolase involved in de-ubiquitination of several target protein such as MDM2 and p53 to promote their stabilization[202]. Taken in account all these observations, the identification of pVHL30-MDM2 association suggest a novel HIF-1α independent function indicating that pVHL30 may play a role in MDM2 stabilization and mutual regulation between oxygen

sensing and cell cycle control. Further studies are however required to define the functional meaning of this association. The MDM2-pVHL30 complex could reflect an enzyme-substrate relationship since MDM2 is characterized by an E3-ligase catalytic activity. On the contrary, pVHL30 could act as substrate recognition particle for MDM2, similarly to its function in VCB complex.

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

Hypoxia is a common feature shared by the most active tumors, characterized by unregulated development and malignant progression[237]. The complex events forming the hypoxia response are mediated by hypoxia-inducible factor HIF-1α, a transcription factor involved in the regulation of numerous genes encoding proteins linked to the oxidative metabolism, energy production, cell cycle regulation, as well as red blood cell and vascular endothelial growth factor (VEGF) homeostasis[238–240]. The hypoxia response, as suggested by the name, is triggered by oxygen deprivation. Under hypoxia, HIF-1α is stabilized and translocated to the nucleus where it activates hypoxia-response element (HRE) promoting hypoxia-regulated genes expression[191]. Also pVHL has a key role in this system, indeed at physiological oxygen levels, it recognizes and ubiquitinates hydroxylated HIF-1α for proteosomal degradation[241]. Deregulation of this network leads to cancer predisposition, such as in the von Hippel-Lindau syndrome, an inherited cancer disease characterized by the formation of tumors and cysts in different tissues[242]. Hypoxia is also important in regulating cell cycle progression and senescence, inducing growth arrest in cells exposed to prolonged oxygen deprivation[243]. Hypoxia-dependent cellular senescence has a critical role in normal tumor suppressor response[243], modulating early malignant transformation[244] and drug-resistance[245]. In particular, hypoxic induction of cell cycle arrest is linked to HIF-1α dependent transcription of the cyclin-dependent kinase inhibitors p21 (CDKN1A)[246] and p27 (CDKN1B)[247]. Together with p57 (CDKN1C), these form a small family of kinase inhibitors playing important roles in negative regulation of the cell cycle[248]. Considering these evidences, I decided to investigate whether also pVHL interacts with these proteins. Bioinformatics analysis, yeast two-hybrid screening and co-immunoprecipitation assays were used to predict, dissect and validate the interactions. I found that the CDKN1 proteins interact with pVHL and share a conserved region mimicking the HIF-1α motif responsible for pVHL binding. Intriguingly, a p27 site-specific mutation associated to cancer is shown to modulate this novel interaction. Results obtained suggest a new connection between the pathways regulating hypoxia and cell cycle progression. More details about this study are collected into a paper published on Scientific Reports (Sci. Rep.7, 46562; 2017) and annexed in the Appendix of this thesis.

3.3 High-throughput cDNA Y2H library screening

In human, the VHL protein is ubiquitously expressed and high levels were found in the central nervous system, kidneys, testis and lung[34]. It is a protein interaction hub[205] involved in several cellular processes. The most characterized function is its role as substrate recognition element in the VCB complex involved in HIF-1α degradation. Inactivation of VHL yields stabilization of the HIF-1α transcription factor and consequent activation of its target genes implicated in angiogenesis, cell growth and metabolism[206]. Mutations affecting pVHL confer different site-specific risks of cancer, suggesting pVHL to have multiple tissue-specific tumor suppressor functions[207]. Experimental investigations, like Y2H library screenings[208,209], provided a huge amount of data about pVHL partners useful to characterize the non-canonical pVHL functions. For this reason, I decided to use this approach to identify novel pVHL30 tissue-specific partners. To perform the experiment, I chose a human testis library for different reasons. First, pVHL is highly expressed in this tissue and second, the same library was previously used in our laboratory, to identify novel pVHL paralog (VLP) interactors. Thus, the comparison of results from library screenings could identify shared interactors suggesting a possible functional overlap among these two proteins. Furthermore, pVHL and HIF-1α are components of the oxygen-sensing mechanism involved in the regulation of oxygen levels. Oxygen levels are very important in this tissue as it is well-known that exposure to hypoxia may lead to male infertility[210], so it could be interesting investigate the role of pVHL in this context.

The Y187 haploid (MATa) yeast strain expressing a human testis cDNA library (available in the Clontech Mate & Plate™ system) was mated to the Y2HGold haploid strain (MATα) expressing pVHL30 as prey protein. The high quality library pre-transformed in Y187 haploid strain was provided by Clontech while the coding sequence for pVHL30 was cloned in the bait vector pGBKT7 using the In-fusion® procedure as described in Materials and Methods section. After DNA sequencing, the recombinant plasmid was used to transform Y2HGold yeast strain and the expression of the recombinant fusion protein Gal4-BD-pVHL30 was checked by Western blot. As reported in Supplementary materials (fig S7) the chimera protein (Gal4-BD-pVHL30) was correctly detected at the calculated molecular weight of 41 kDa.

At this point, I checked whether the pVHL30 expression could have some effects on the yeast growth rate, an important parameter to consider for the setup of the next step,

consisting in the mating procedure with the Y187 strain expressing the cDNA library. The bait protein did not alter both the yeast colonies growth-rate and their dimension, suggesting that pVHL30 expression is not toxic for yeast cells. Another important point was to verify that the expression of the bait protein alone was unable to activate the transcription of the different Y2H reporter genes (details in Material and Methods section) evaluated in the diploid strain obtained after mating.

Thus, Y2HGold expressing pGBKT7 VHL30 were plated on selective agar media (SD - Trp, SD - Trp + X-α-Gal and SD - Trp + X-α-Gal + Aba). As expected, in presence of Aba drug no colonies were observed on agar plates, while white colonies were present on SD - Trp and SD – Trp - X-α-Gal (data not shown). Before the mating experiment, a pilot ones was performed. The overnight mated culture was monitored for about 40 hours to estimate the mating duration. After 16 hours of incubation and then every 4 hours, a drop of the co-culture was observed under a phase contrast microscope (40X). As indicated in the manufacturer protocol, when the mating occurs 3-lobed structure (also known as “Mickey Mouse” face-shape) are visible. This peculiar yeast shape indicates the formation of budding diploid cells. After 20 hours, the 3-lobed shape structures were detected indicating the mating was going on. The presence of the 3-lobed structures does not indicate whether the mating is just started or it is finished. To address this problem, I decided to take an aliquot of mated culture after 24h and 30h of incubation and spread a dilution on different media i.e. YPGA, SD -Trp, SD -Leu and SD –Trp - Leu. Plates were incubated for 3 days at 28 °C. As reported in figure 22, the mated culture has grown healthy on the rich media (YPGA). Considering the selection nutritional markers of Y2H vectors, the growth-rate of the bait culture was evaluated on SD-Trp, while the one of the library culture on SD-Leu. The number of colonies on SD –Leu is lower than on SD –Trp. This was expected because only a small amount of yeast library aliquot was employed in the pilot experiment. Some colonies were present on SD –Trp –Leu confirming that the mating correctly occurred. Of note, the number of diploids was higher after 30h of incubation. Considering collectively the pilot experiment results and the library datasheet indications, I chose 30h as time for mating culture incubation.

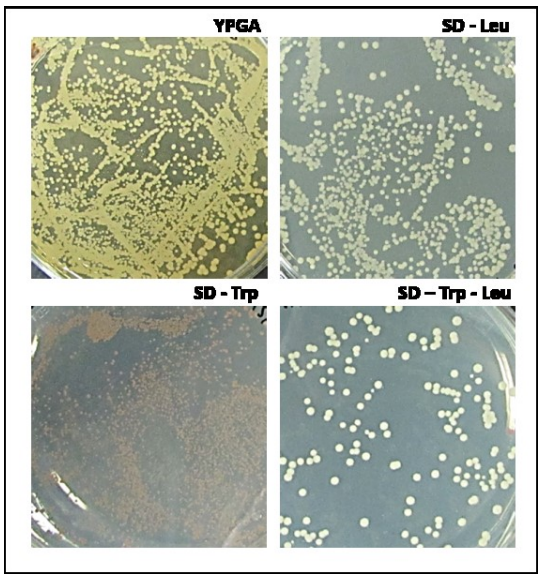


Figure 22 Small-scale mating controls. A mated culture aliquot was spread on permissive and selective agar plates. The culture was plated after 30h of mating incubation. Dilution showed 1:1000

At this point, large-scale mating was performed as reported in Materials and Methods section. To define the library titre, 100 µl of 1:10, 1:100, 1:1000 and 1:10000 dilutions were spread on SD –Leu agar plates (fig 23). Considering the 1:10000 dilution, $9x10^7$ cells/ml were counted consistent with protocol expectation (i.e. > $2x10^7$ cells/ml).

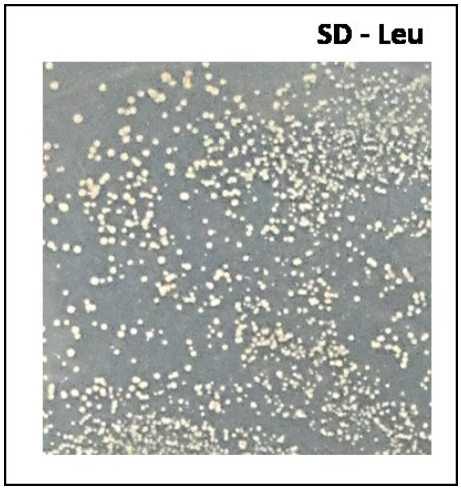


Figure 23 Titration of the library aliquot. Different dilutions were spread on SD -Leu medium in order to titre the library aliquot. Dilution reported 1:10000

To determine the mating efficiency, I calculated the number of total screened clones. An aliquot of 30 hours-mated culture was spread on SD –Trp, SD –Leu, SD –Trp - Leu and colonies were manually counted.

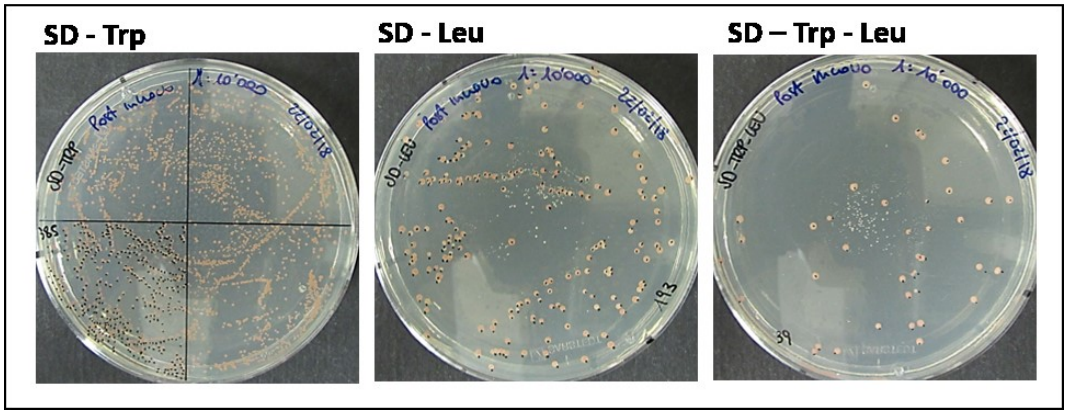


Figure 24 Diploids count. After 30 hours of mating, a co-culture aliquot was spread on selective agar plates in order to calculate the number of screened clones. Dilution reported 1:10000

Data are reported in the Table 15

media	N° clones	viability	Resusp. vol	Plating vol	Dilution	Screened clones
SD - Leu	193	Library	11,8 ml	100 µl	1:10000	$2{,}3 \times 10^{8}$
SD - Trp	1540	bait	11,8 ml	100 µl	1:10000	$1{,}82 \times 10^{9}$
SD – Trp -Leu	39	diploids	11,8 ml	100 µl	1:10000	$3{,}9 \times 10^{6}$

Table 15 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 considering the dilution.

These values were used to calculate the mating efficiency, as reported below.

$$\frac{\text{No of cfu/ml of diploids}}{\text{No of cfu/ml of limiting partner}} \times 100 = \%\ \text{diploids}$$

The limiting partner in this screening was the prey library. I calculated the percentage of diploids resulting in 20%; this value is very high compared with the estimated range 2-5%. During the experiment, the plates containing the mated culture were incubated at 28°C for 5 days and the presence of colonies was monitored from the third day, each day. Colonies were marked with progressive numbers.

	Day 3	Day 4	Day 5	
plates	**N° clones**	**N° clones**	**N° clones**	**N° clones**
1-10	54	70	12	136
11-20	77	90	3	170
21-30	76	66	11	153
31-40	81	101	3	185
41-50	81	70	2	153
51-58	53	75	8	136
tot/day	**422**	**472**	**39**	**933**

Table 16 Clones monitoring. Blue colonies was counted for 3 days. Number of clones per day/plates and the final total number are reported

At this point to avoid false positive, all clones were patched out on more stringent agar plates (i.e. SD-Leu-Trp-His-Ade named QDO) as shown in figure 25. After this step, the number of positive clones was reduced to 607. The 326 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 medium albeit retaining blue color. The number of false positive was very high but was expected considering the total number of blue colonies obtained.

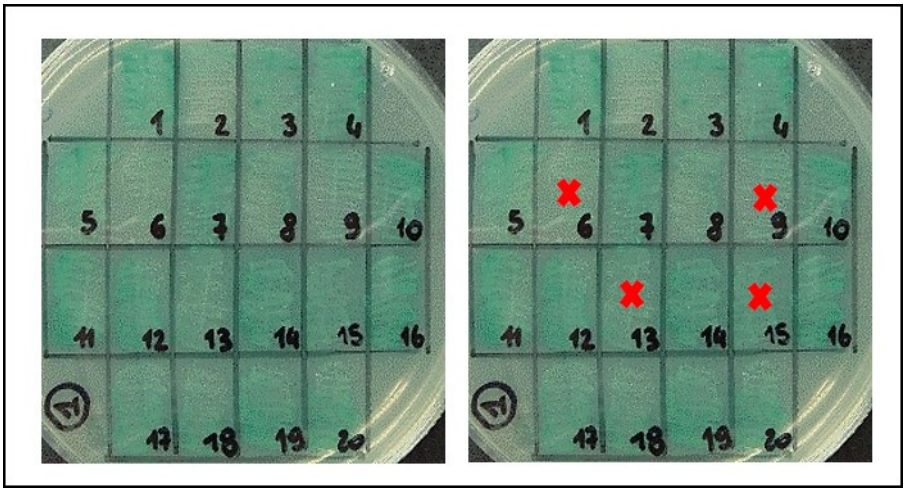


Figure 25 (previous page) Colony patches. Blue colonies were patched out on higher stringency agar plates (QDO+X-α-gal+Aba) in order to avoid false positive interactions. Here, an exemplificative picture representing clone 1 to 20 patches (left) is reported. Red crosses indicate four false positive clones that were not further analyzed (right).

After this step, positive clones identification process started. Total DNA (i.e. genome, bait and prey plasmids) was extracted from 607 yeast clones selected on QDO and analyzed by PCR. The PCR products were obtained using Gal-AD For and Gal4-AD Rev primers able to pair the vector sequences upstream and downstream the multiple cloning site of the pGADT7 plasmid. PCR samples were run on 0.8% agarose gels. Figure 26 shows one of these gels. (The entire subset of data generated for this analysis is reported as Supplementary materials figure S8). According to library datasheet, PCR products were included in 0.5 kb to 4.0 kb range.

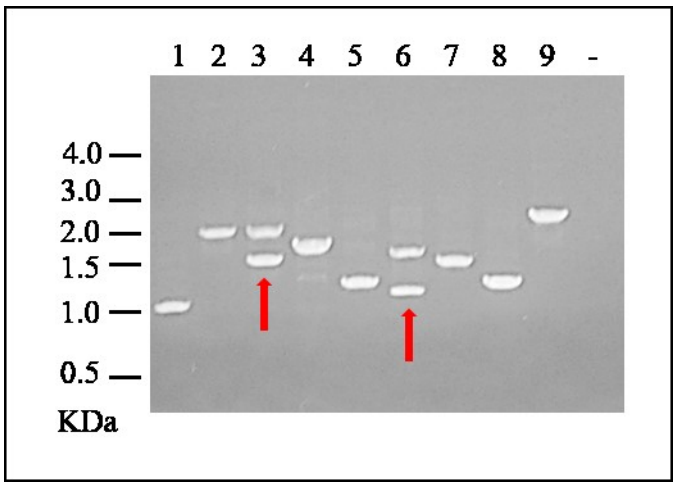


Figure 26 PCR fragments visualization on 0.8% agarose gel. Each line is marked with a number corresponding to the positive clone. The last column corresponds to the negative control. Red arrows indicates clones with more than one insert, i.e. 3 and 6.

This preliminary analysis identified yeast clones characterized by the presence of multiple recombinant plasmids carrying different inserts (i.e. clone 3 and 6). Ideally, each mixed population of DNA fragments should be further analyzed to define whether both the DNA fragments interact with pVHL30 or not. However, considered the relatively low amount of these mixed clones (88/607), I decided to exclude these problematic samples from the identification process.

On the contrary, all the other yeast DNAs (i.e. genome, bait and prey plasmids) were transformed in *E.coli* in order to amplify and isolate DNAs of prey plasmids. From each

transformation, five clones of each yeast-derived DNA were inoculated together in LBA and used to extract the DNA.

In parallel, to further validate the putative novel pVHL30 interactors also in a different yeast genetic background, each DNA was used to transform the Y190 yeast reporter strain, largely used in the previous results (chapter 3.1) to investigate Y2H binary interaction between VHL30 and MDM2. This strain was co-transformed with pGBKT7-VHL30 or pGBKT7-empty (Ø) vector (as prey auto-activation control) and plated on permissive (i.e.SD-Trp-Leu) and selective medium (i.e. SD –Trp-Leu-His supplemented with 30 mM or 60 mM 3-AT). For each transformation, two clones were patched out on both media. As reported in figure 27, I observed different yeast behaviors, with clones strongly binding pVHL30 survived resembling the positive control C+ (i.e. cl2), whereas other clones unable to grow on selective conditions as the negative control C- (i.e. cl1,cl5). A group of clones that auto-activate (i.e. cl4) was also observed. For this specific group, growth-rate of yeast co-expressing both prey-plasmid and empty vector was higher than growth-rate from yeast co-expressing the same plasmid and pVHL30. The entire subset of data generated for the experiment is reported as Supplementary materials (fig. S9).

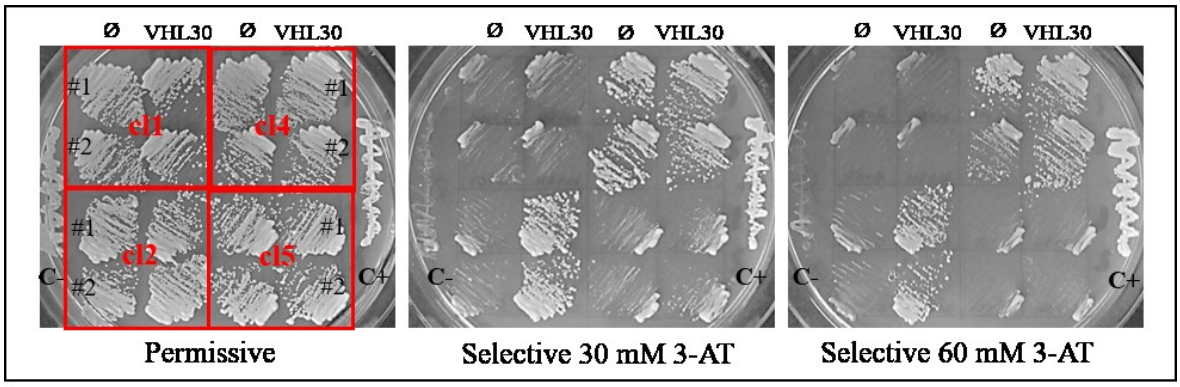


Figure 27 Yeast patches. On permissive and selective media, two clones (#1,#2) for each transformation were tested. Numbers indicate the analyzed clones (i.e. cl 1,2,4,5) co-transformed with pGBKT7VHL30 or pGBKT7 empty (Ø) vectors. C+ and Ø are for positive and negative Y2H controls, respectively. The scheme reported for the permissive plate is the same also used for the selective medium.

This second analysis further reduced the number of positive clones to 260. The plasmidic DNA isolated from bacteria corresponding to each positive clone was sequenced by the Sanger method using Gal4-AD For as primer (GATC service). Resulting cDNA sequences were translated with ExPASy Translate and the amino acidic sequences used to perform

BLAST search in Uniprot. This step found matches for 142 sequences (fig. S10). Each identified sequence was mapped on the corresponding full-length protein in order to define the real interacting fragment. Findings resulting from this analysis are presented in Table 17.

clone	protein name	protein size	binding fragment	function
cl 535	Jouberin	1196	978-1136	ciliogenesis
cl 338	Ankyrin repeat and EF-hand domain-containing protein 1	776	326-624	n.d.
cl 784	Ankyrin repeat and EF-hand domain-containing protein 1	776	325-624	n.d.
cl121	Ankyrin repeat domain-containing protein 11	2663	26-249	Chromatin regulator
cl 237	Ankyrin repeat domain-containing protein 11	2663	26-249	chromatin regulator
cl 491	Rho guanine nucleotide exchange factor 7	782	50-140	GEF apoptosis
cl 5	Protein BEX2	160	6-40	cell cycle regulator
cl 252	Protein BEX4	120	1-120	microtubule Dacetil.
cl 135	Protein BEX4	120	1-120	microtubule Dacetil.
cl 812	Breast cancer type 1 susceptibility protein	1863	1740-1863	E3-ub lig/DNA repair
cl 582	T-complex protein 1 subunit epsilon	541	1-230	actin/tubulin folding
cl 64	T-complex protein 1 subunit eta	543	145-368	actin/tubulin folding
cl 341	T-complex protein 1 subunit eta	543	146-474	actin/tubulin folding
cl 373	T-complex protein 1 subunit eta	543	203-536	actin/tubulin folding
cl 852	T-complex protein 1 subunit eta	543	145-472	actin/tubulin folding
cl 854	CDK5 regulatory subunit-associated protein 3	110	4-110	n.d.
cl 353	COMM domain-containing protein 1	190	1-190	protein ub regulator
cl 119	Copine-5	290	11-239	n.d.
cl 881	Cathepsin D (CTSD)	412	291-412	protease/prot degrad
cl 82	Death-inducer obliterator 1	2220	363-529	tumor suppressor
cl 95	Death-inducer obliterator 1	2220	363-529	tumor suppressor

cl 201	Death-inducer obliterator 1	2220	364-529	tumor suppressor
cl 131	Destrin	165	49-165	actin depolimerization
cl 75	Elongation factor 1-alpha 1	462	233-420	protein biosynthesis
cl 219	Elongation factor 1-alpha 1	462	295-462	protein biosynthesis
cl 175	Elongation factor 1-alpha 1	462	221-462	protein biosynthesis
cl 185	Elongation factor 1-alpha 1	462	295-462	protein biosynthesis
cl 239	Elongation factor 1-alpha 1	462	249-462	protein biosynthesis
cl 308	Elongation factor 1-alpha 1	462	297-462	protein biosynthesis
cl 332	Elongation factor 1-alpha 1	462	297-462	protein biosynthesis
cl 372	Elongation factor 1-alpha 1	462	213-462	protein biosynthesis
cl 900	Elongation factor 1-alpha 1	462	215-324	protein biosynthesis
cl 574	Elongation factor 1-alpha 1	462	165-434	protein biosynthesis
cl 138	Elongation factor 1-alpha 1	462	220-462	protein biosynthesis
cl 154	Elongation factor 1-alpha 1	462	197-434	protein biosynthesis
cl 476	Elongation factor 1-alpha 1	462	297-462	protein biosynthesis
cl 424	Elongation factor 1-alpha 1	462	231-462	protein biosynthesis
cl 324	Elongation factor 1-alpha 1	462	249-462	protein biosynthesis
cl 295	Elongation factor 1-alpha 1	462	288-462	protein biosynthesis
cl 655	Elongation factor 1-alpha 1	462	197-434	protein biosynthesis
cl 39	Elongin-C	112	1-112	protein degradation
cl 590	Elongin-C	112	1-112	protein degradation
cl 216	Elongin-C	112	1-112	protein degradation
cl 217	Elongin-C	112	1-112	protein degradation
cl 454	Elongin-C	112	1-112	protein degradation
cl 936	Elongin-C	112	1-112	protein degradation
cl 450	3'-5' exoribonuclease 1	349	35-209	histone mRNA degrad
cl 633	Electron transfer flavoprotein subunit alpha, mitochondrial	333	201-333	electron transport
cl 799	Electron transfer flavoprotein subunit alpha, mitochondrial	333	203-333	electron transport
cl 89	Electron transfer flavoprotein subunit alpha, mitochondrial	333	200-333	electron transport
cl 312	F-box only protein 28	368	1-126	ub/ prot degradation

cl 807	F-box only protein 34	711	1-53	SRP of E3-ub complex
cl 388	Flotillin-1	427	120-249	caveolae formation
cl 701	G patch domain and ankyrin repeat-containing protein 1	356	1-154	n.d.
cl 251	General transcription factor 3C polypeptide 2	911	1-290	DNA transcription
cl 348	Guanylate kinase (Fragment)	272	72-203	phosphorylation
cl 205	Guanylate kinase	197	21-197	phosphorylation
cl 313	Intraflagellar transport protein 88 homolog	833	39-282	ciliogenesis
cl 286	Inositol-trisphosphate 3-kinase	683	495-617	phosphorylation
cl 736	Kelch-like protein 10	608	472-608	ub/ prot degradation
cl 144	Microtubule-associated protein 1B	2468	2168-2468	microtub stabilization
cl 194	Microtubule-associated protein 1S	1059	778-1053	apoptosis
cl 287	Microtubule-associated protein 1S	1059	863-1059	apoptosis
cl 473	Microtubule-associated protein 1S	1059	677-932	apoptosis
cl 156	Microtubule-associated protein 1S	1059	815-1059	apoptosis
cl 630	28S ribosomal protein S9, mitochondrial	396	45-336	n.d.
cl 309	E3 ubiquitin-protein ligase MSL2	577	375-577	ub/ prot degradation
cl 102	Cytochrome c oxidase subunit 2	227	169-227	oxygen reduction
cl 235	Cytochrome c oxidase subunit 2	227	169-227	oxygen reduction
cl 561	Nischarin	1504	1386-1504	cell survival/migration
cl 236	Probable 28S rRNA (cytosine-C(5))-methyltransferase	429	291-428	methylation
cl 37	Putative methyltransferase NSUN7	718	510-696	methylation
cl 46	Putative methyltransferase NSUN7	718	510-523	methylation
cl 61	Putative methyltransferase NSUN7	718	510-718	methylation
cl 111	Putative methyltransferase NSUN7	718	510-718	methylation
cl 145	Putative methyltransferase NSUN7	718	510-701	methylation
cl 148	Putative methyltransferase NSUN7	718	510-687	methylation
cl 150	Putative methyltransferase NSUN7	718	510-718	methylation
cl 207	Putative methyltransferase NSUN7	718	510-630	methylation

cl 247	Putative methyltransferase NSUN7	718	510-718	methylation
cl 155	Putative methyltransferase NSUN7	718	510-718	methylation
cl 57	Putative methyltransferase NSUN7	718	510-718	methylation
cl 151	Putative methyltransferase NSUN7	718	510-718	methylation
cl 81	Protein disulfide-isomerase A3	505	343-500	protein folding
cl 124	Protein disulfide-isomerase A3	505	345-500	protein folding
cl 187	Protein disulfide-isomerase A3	505	345-500	protein folding
cl 198	Protein disulfide-isomerase A3	505	343-500	protein folding
cl 732	ATP-dependent 6-phosphofructokinase, platelet type	784	562-616	glycolysis
cl 381	Prohibitin-2	299	42-299	transcription inhibitor
cl 439	Prohibitin-2	299	42-290	transcription inhibitor
cl 218	Piwi-like protein 4	852	776-852	tumor enhancer
cl 815	Piwi-like protein 4	852	777-852	tumor enhancer
cl 600	Plakophilin-2	504	472-504	cell-cell adhesion
cl 104	cGMP-dependent protein kinase 1	376	320-376	phosphorylation
cl 226	cGMP-dependent protein kinase 1	376	320-376	phosphorylation
cl 261	26S proteasome regulatory subunit 4	440	1-250	protein degradation
cl 318	26S proteasome regulatory subunit 4	440	1-294	protein degradation
cl 387	26S proteasome regulatory subunit 4	440	1-180	protein degradation
cl 911	Arginine-glutamic acid dipeptide repeats protein	1566	59-134	cell survival control
cl 378	Arginine-glutamic acid dipeptide repeats protein	1566	59-134	cell survival control
cl 182	Histone deacetylase complex subunit SAP30	220	61-210	deacetylation
cl 892	Histone deacetylase complex subunit SAP30	220	77-220	deacetylation
cl 637	Histone deacetylase complex subunit SAP30	220	64-220	deacetylation
cl 475	Selenoprotein P	382	119-299	selenium transport
cl 256	Structural maintenance of chromosomes protein 5	1101	1-113	DNA repair
cl 400	Structural maintenance of chromosomes protein 5	1101	1-114	DNA repair

cl 68	U5 small nuclear ribonucleoprotein 200 kDa helicase	2136	1229-1334	RNA splicing
cl 72	SPARC	303	133-303	cell growth
cl 96	Spermatogenesis-associated protein 22	363	1-249	germ cell division
cl 100	Spermatogenesis-associated protein 22	363	27-269	germ cell division
cl 137	Spermatogenesis-associated protein 22	363	58-324	germ cell division
cl 215	Spermatogenesis-associated protein 22	363	22-189	germ cell division
cl 231	Spermatogenesis-associated protein 22	363	20-326	germ cell division
cl 830	Spermatogenesis-associated protein 22	363	22-286	germ cell division
cl 97	Spermatogenesis-associated protein 22	363	62-363	germ cell division
cl 158	Spermatogenesis-associated protein 22	363	20-356	germ cell division
cl 915	Spermatogenesis-associated protein 22	363	20-115	germ cell division
cl 393	Spermatogenesis-associated protein 22	363	24-294	germ cell division
cl 58	Spermatogenic leucine zipper protein 1	430	300-430	germ cell prolif and diff
cl 642	STAM-binding protein	424	151-424	protein degradation
cl 523	STAM-binding protein	424	151-370	protein degradation
cl 383	Histone-lysine N-methyltransferase SUV39H2	410	293-410	chromatin regulator
cl 880	Tudor domain-containing protein 7	1098	749-913	post-transcr regulator
cl 65	Testis-expressed protein 35	233	1-170	n.d.
cl 418	Testis-expressed protein 35	233	1-196	n.d.
cl 165	E3 ubiquitin-protein ligase TTC3	2025	998-1185	ub/ prot degradation
cl 166	E3 ubiquitin-protein ligase TTC3	2025	998-1185	ub/ prot degradation
cl 249	E3 ubiquitin-protein ligase TTC3	2025	998-1247	ub/ prot degradation
cl 691	Ubiquitin-conjugating enzyme E2 D2	147	8-147	protein ubiquitination
cl 533	Ubiquitin-conjugating enzyme E2 D2	147	8-112	protein ubiquitination
cl 48	Zinc finger and BTB domain-containing protein 17	803	459-662	cell cycle regulator
cl 56	Zinc finger and BTB domain-containing protein 17	803	306-450	cell cycle regulator
cl 74	Zinc finger and BTB domain-containing protein 17	803	459-723	cell cycle regulator
cl 524	Zinc finger and BTB domain-	803	459-711	cell cycle regulator

	containing protein 17			
cl 78	Zinc finger and BTB domain-containing protein 17	803	459-711	cell cycle regulator
cl 426	Zinc finger and BTB domain-containing protein 17	803	460-741	cell cycle regulator
cl 757	Zinc finger and BTB domain-containing protein 17	803	461-793	cell cycle regulator
cl 754	Zinc finger and BTB domain-containing protein 17	803	487-757	cell cycle regulator
cl 270	Zinc finger and BTB domain-containing protein 17	803	461-789	cell cycle regulator
cl 141	Zinc finger protein 200	395	200-359	spermatogenesis
cl 22	Zinc finger protein 668	619	319-365	transcription regulator
cl 768	Zinc finger protein 827	1081	1013-1076	transcription regulator

Table 17 List of proteins interacting with pVHL30.

To continue my investigation, I focused the next analysis on proteins presenting more than 2 hits. In particular, putative pVHL30 interactors can be subdivided into three main groups: i) highly represented (i.e. EEF1A1, NSUSN7, SPATA22 and ZBTB17) characterized by a number of hits within 17 and 9, ii) average represented with proteins showing a number of hits between 6 and 4 each (i.e. CCT7, MAP1S ,PDIA3 and EloC) and iii) normal represented including proteins with 3 hits each (i.e. DIDO1, PMSC1, SAP30, ETFA and TTC3).

Hits (i.e. independent DNA sequences encoding for the same protein) were aligned using ClustalX2 and visualized with Jalview 2.11.0 in order to identify shared region (if any) among the different clones. The Elongin-C was not included in this analysis as all interacting clones were found to encode for full-length protein (Table 17). As an example of the analysis work flow, I reported the alignment generated using fragment belonging from EEF1A1, which is the most represented protein in the dataset. In particular, I found 17 different hits corresponding to 12 different amino acidic sequences. These were aligned and compared with the EEF1A1 reference sequence retrieved by UniProt (ID P68104). As reported in figure 28, the alignment identified a fragment of 27 residues (aa 297- 324) shared among all clones, suggesting this region to be the pVHL30 binding motif.

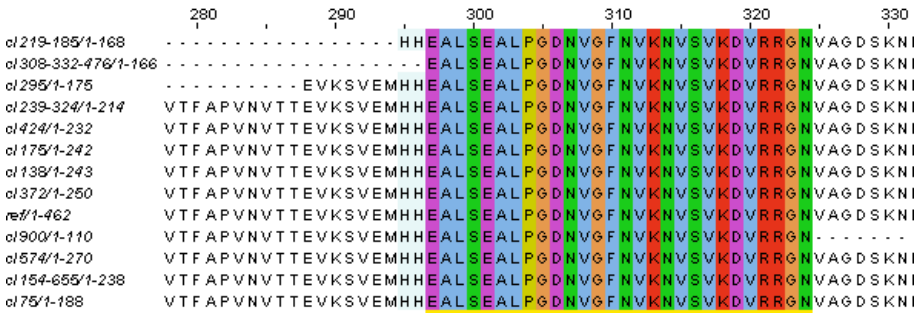


Figure 28 EEF1A1 multiple alignment. Library derived sequences and the reference (P68104) were aligned with ClustalX2 and visualized with Jalview.

The same protocol was used to investigate all the proteins with more than 2 hits. A total of 12 different putative pVHL30 binding fragments were identified (Table 18).

protein	n° hits	region shared
EEF1A1	17	297-324
NSUN7	12	510-523
SPATA22	10	62-115
ZBTB17	9*	487-662
CCT7	4	203-368
PDIA3	4	345-500
MAP1S	4*	863-1053
DIDO1	3	363-529
PSMC1	3	1-180
SAP30	3	77-210
ETFA	3	203-333
TTC3	3	998-1185

Table 18 Regions shared among different clones of a same protein (* one clone is excluded from this mapping analysis)

Multiple sequence alignment of the 12 fragments using MAFFT[211] showed no shared similarities among these putative binding sites, suggesting that each of them possess unique features. To identify the key-features conferring their ability to bind pVHL30, I decided to further characterize these fragments using different bioinformatics tools. In particular, investigation about secondary structure content, presence of functional domains,

linear motifs, post-translational modification sites and pathogenic mutations were taken into the account.

Interestingly, search with ELM, a database collecting eukaryotic linear motifs, predicted the presence of the USP7 binding motif in NSUN7, SPATA22 and PDIA3. The USP7 is a deubiquitinating enzyme involved in p53 stabilization[202], with high affinity for MDM2, the E3 ubiquitin-ligase normally involved in p53 degradation. As mentioned in chapter 3.1 of this manuscript, the pVHL30 binding region of MDM2 localizes up-stream to the USP7 interacting motif. Moreover, pVHL30 is known to interact with both USP33 and USP20, two enzyme belonging to the so-called USP family, whose USP7 is a further member[212]. Data from the literature indicate that pVHL30 mediates the ubiquitination and degradation of both USP33 and USP20, in turn down-regulating the pathways controlled by these two proteins[213].

On the other hand, the 12 sequence fragments were also analyzed with DiLiMOT[214] which identified 10 different linear motifs listed in figure 29. According to the significance threshold, the first six predicted motifs were considered as good candidates. The most represented are VGxxxK and PxxxVxxN, each shared among 4 different proteins and both found in EEF1A1 protein, a well-known pVHL30 interactor[215]. This evidence strongly supports their possible real involvement in pVHL30 binding. The remaining 4 motifs identified are rich in lysines (i.e. GxxKxxK, KKKxK, KxKxKxK, KxxxPK). Lysines are positively charged amino acids and possible sites of post-translational modifications. According to their electrostatic charge, these motifs could likely mediate binding with the pVHL30 N-terminal portion, which is characterized by a strong concentration of negatively charged residues. These 4 motifs were predicted with a P-value below the significance threshold, however this doesn't exclude completely their reliability, as real binder motifs can still occur also with a less significant values.

Motif	S_{cons} (help)	Proteins with motif (in unfiltered regions) (help)	Proteins total (non-redundant) (help)	P (binomial test) (help)
□ VGxxxK	2.22e-29	4	12	2.84e-05
□ PxxxVxxN	3.12e-24	4	12	2.20e-05
□ GxKxxK	1.42e-22	4	12	2.48e-04
□ KKKxK	1.13e-20	4	12	9.00e-06
□ KxKxKxK	1.39e-18	4	12	3.81e-06
□ KxxxPK	3.10e-18	5	12	1.29e-05
□ KxxKxxxP	1.20e-17	4	12	2.37e-04
□ KNxxxK	1.63e-16	4	12	3.15e-04
□ AxxVP	3.45e-16	4	12	2.18e-04
□ KKK	4.05e-16	5	12	2.76e-04

Figure 29 (previous page) List of linear motifs identified by DiLiMOT

Search on protein sequences with COSMIC[190] identified a number of mutations localizing in the fragments belonging from SPATA22, ZBTB17, CCT7, MAP1S and PDIA3 which correlate with different types of cancer, such like adenocarcinoma, squamous cell carcinoma, malignant melanoma and clear cell renal cell carcinoma. In particular, missense mutations R89I in SPATA22, R562C and R625W in ZBTB17, P235S and E316K in CCT7, R863Q and G891S in MAP1S as well as the substitutions S497S in ZBTB17 and L361L in PDIA3 correlate with pathogenic conditions. Starting from these findings, it can be speculated that mutations in these positions can alter the pVHL30 binding affinity thus impairing its tumor suppressor activity. Further investigations, however, are required to address this open question.

Taken into the account the presence of different clones encoding for a same protein, a total of 61 different pVHL30 interactors were identified from the library dataset. In details, 6 proteins revealed to be already known pVHL interactors, whereas remaining 55 candidates correspond to new interactors, which were never associated so far to pVHL30. According to VHLdb[32] the 6 known interactors are EEF1A1, CCT5, CCT7, Elongin-C, SNRNP200 and UBE2D2.

EEF1A1 is a GTP-binding protein with an important role in the translational machinery. Upon GTP binding, it delivers the aminoacylated-tRNA to the A site of the ribosome for mRNA decoding by codon-anticodon interaction[216]. It is also a pleiotropic protein expressed in human tumors including breast cancer, ovarian cancer and lung cancer. EEF1A1 plays a critical role in human diseases through enhancement of oncogenesis, blockade of apoptosis and increased cellular transformation[217]. Another very well characterized function of this protein is its involvement in the mammalian nuclear export machinery. In particular it is involved in the nuclear export of protein containing the transcription-dependent nuclear export motif (TD-NEM), such as VHL and the poly(A)-binding protein (PABP1)[215]. Nuclear export through TD-NEM , which is encoded by the consensus sequence DxGx2Dx2L, requires ongoing RNA Polymerase II-mediated transcription and operates independently of the classical CRM1/NES-mediated nuclear export pathway. EEF1A1 interacts specifically with TD-NEM of VHL and PABP1 and disrupting this interaction, by point mutations of key TD-NEM residues or treatment with

actinomycin D, an inhibitor of RNA PolII-dependent transcription, prevents assembly and nuclear export[215].

Another well-characterized pVHL interactor found among library dataset is the Elongin C. Elongin C is a regulatory subunit of the transcription factor B (SIII) complex. It activates RNA polymerase II mediated elongation by suppressing transient pausing of the polymerase at many sites within transcription units[218]. It is also a component of the VCB multiprotein complex which acts as an ubiquitin-ligase (E3) inducing proteasome-dependent degradation of targeted proteins. The hypoxia-inducible factor 1 alpha (HIF-1α) is the best characterized substrate of this complex[219]. Association of pVHL to Elongin B and C is important for pVHL tumor suppressor activity as pVHL mutation of Elongin B/C binding region was found in VHL-associated tumors[220].

Collectively, these known interactors can be considered as positive controls of our screening procedure, supporting the reliability of our library dataset. Undoubtedly, the most valuable part of our dataset is represented by the novel 55 pVHL30 interactors, so I decided to start the investigation of their known sub-cellular localization.

As reported on figure 30, new VHL30 binding partners localize preferentially into the nucleus and in the cytoskeleton (respectively, 29 and 7 proteins in these compartments). Remaining interactors are almost equally distributed among mitochondrion, endosome, membrane, chromosome and endoplasmic reticulum. Only one protein was found to participate the lysosome, whereas any protein was found to reside into the Golgi. Interestingly, two secreted proteins, i.e. cathepsin D (CTSD) and Selenoprotein P (SELENOP) were also identified.

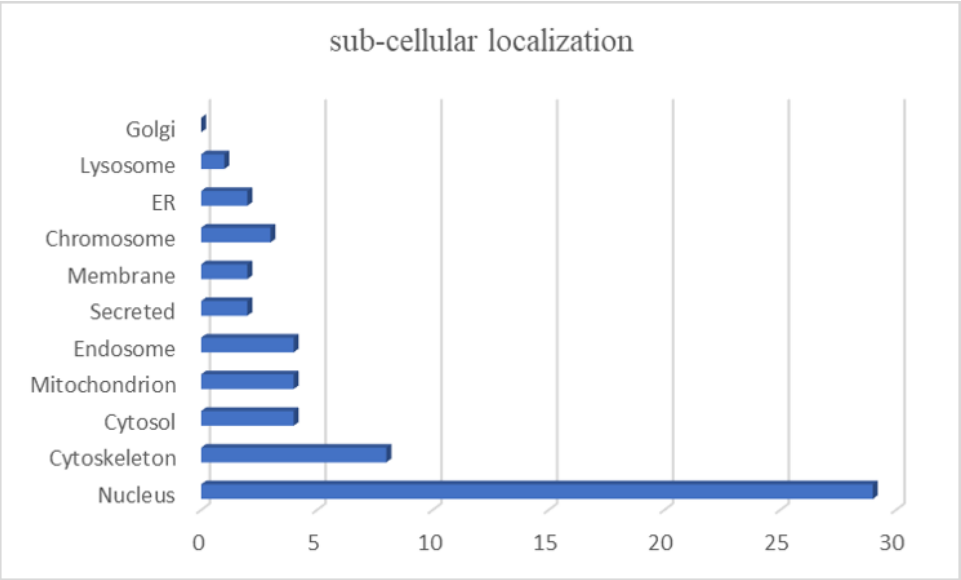


Figure 30 Sub-cellular localization of pVHL30 new interactors

The predominance of nuclear proteins is not surprising, as this compartment represents, with cytoplasm, the sub-cellular compartments in which pVHL is normally expressed at physiological conditions. In particular, the pVHL sub-localization is thought to be regulated in a cell cycle-dependent manner, that influences the protein shuttling between the nucleus and the cytoplasm[221].

To investigate about the pathways in which the 61 pVHL30 interactors (screening-derived) operate, a protein interacting network centered around pVHL30 was created using STRING[171]. Resulting network shows no trivial interaction among the 61 proteins but 5 entries. Indeed, only 9 interactions among pVHL30 interactors were identified. For this reason, the preliminary interacting network was enriched including the "interactors of interactors" retrieving no more than 50 proteins for each interactor.

The cluster analysis of the resulting extended network shows pVHL30 partners to be associated to multiple relevant cellular pathways (figure 31). In particular, a major central cluster composed by 43 proteins highly interconnected was identified. Among them, PSMC1, GUK1 and STAMBP1, EEF1A1, CCT5 and CCT7 are pVHL30 partners identified by the Y2H screening. This cluster seems to have a bi-modal distribution with a most populated part containing HIF-1α and several proteins involved in protein degradation. This finding is expected as it is confirmatory of the known pVHL30 involvement in this process. This part of the cluster includes PSMC1 and multiple proteins forming the proteasome subunits. The proteasome plays a key role in the maintenance of

protein homeostasis by removing misfolded or damaged proteins, which could impair cellular functions, and by removing proteins whose functions are no longer required. It participates in numerous cellular processes, including cell cycle progression, apoptosis, or DNA damage repair. Processes in which pVHL30 is also involved. On the other hand, in the same cluster were found several proteins belonging from the chaperonin family (i.e. CCT5,CCT7) known to play a key role in the correct folding of cytoskeleton proteins. This evidence can be related to the pVHL30 involvement in the regulation of microtubule dynamics[88], an example of HIF-1A independent pVHL function. As reported in figure 31, others 4 pVHL30 interactors (i.e. DIDO1, GTF3C2, SNRNP200 and PFKP) are member of a cluster composed by 11 nodes. This cluster is centered around CDC5L, a protein with a key role in cell-cycle control. Also this pathway was proposed to be linked to pVHL[193] function. Collectively, these findings highlighted already known pVHL30 functions. On the other hand, most of the 61 interactors identified so far, showed no trivial connections in the preliminary network suggesting that pVHL30 may directly link these pathways, putatively also exerting novel not yet characterized functions.

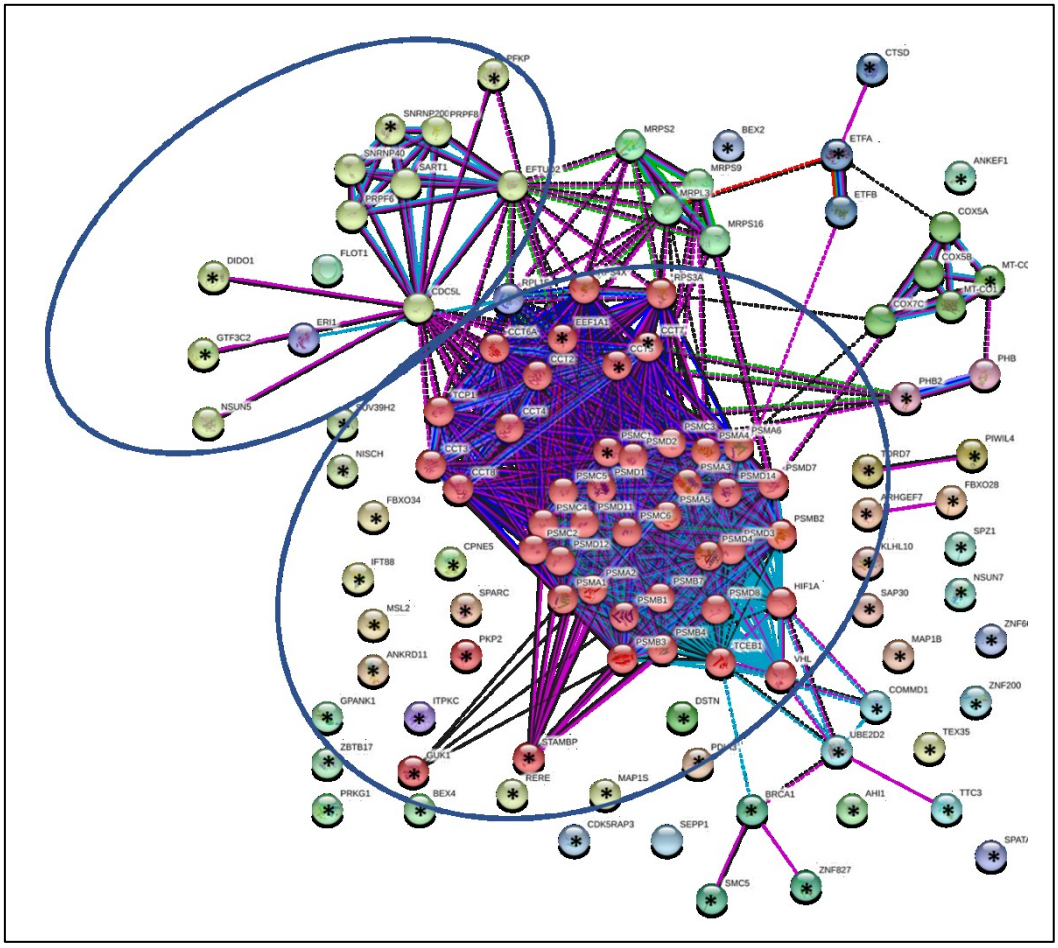


Figure 31 (previous page) STRING functional network. Two clusters are highlighted corresponding to proteins involved in protein degradation, cytoskeleton and cell-cycle regulation. Black stars mark the VHL30 interactors screening-derived.

All of the data described so far consider only 142 entries of 260 that were correctly associated to a corresponding human protein. Instead, for other 118 entries the identification process failed. In this case cDNA fragments were deeply analyzed with tBLASTx as described in materials and methods section. These entries were found to correspond to peptides containing non-coding region i.e. 3'- or 5'-untraslated regions (UTR) rather than normally translated protein regions. Together, the Y2H library screening produced two different datasets, one accounting for 142 entries corresponding to 61 different proteins and a second dataset collecting other 118 pVHL interacting amino acidic sequences not attributable to individual coding regions. From literature it is known that Y2H screening results in identification of many out-of-frame (OOF) clones that encode peptides with no sequence homology to known proteins. These OOF clones generate during library construction from cloning restriction-digested cDNAs fused to the AD. Nevertheless, deeper analysis of these peptides could still reveal common short linear motifs (SLiMs) responsible for their selection[222]. In other words, they can represent a further reservoir of putative new linear motifs able to mediate association with pVHL30.

Discussion

The Y2H screening is a useful technique vastly used to construct extensive protein-protein interaction maps for humans and several model organisms[223]. In this study, I took advantage from this approach to investigate the human pVHL30 interactome in testis tissue. The analysis identified 260 positive clones further classified into 2 groups: 61 human proteins and 118 non coding peptides. Taken into account the high number of interactors identified, the attention was firstly focused on human proteins presenting at least twice in the dataset. However, the generated dataset contains more information that will be analyzed and discussed in the future. The screening identified 6 known pVHL30 interactors confirming the system reliability. Indeed, pVHL30 was detected as partner of Elongin C, EEF1A1 and other 4 known interactors suggesting the "good ability" of the technique to find out *in vivo* real associations outside the physiological context in which they occur. On the other hand, this approach allows to obtain a huge amount of information about proteins interacting with the subject of interest. In this case 55 new interactors were found, including proteins involved in relevant cellular pathways, such as the cell-cycle

regulation, DNA damage repair, apoptosis and cytoskeleton regulation. Very interestingly, in most cases a specific fragment of each interactor was detected, suggesting its involvement in pVHL30 binding. However, information available at this step of the analysis are not sufficient to conclude whether each fragment alone is able to bind pVHL30 or not. The preliminary interacting network created using these interactors showed limited connections among them at the pathway level. These findings may suggest a broader involvement of pVHL30 in other pathways not yet identified. Another interesting point to discuss, is the ability of pVHL30 to recognize and interact with the UTR regions, as showed by the interaction with 118 entries codifying for out-of-frames clones. The UTRs regions are known to have a key role in post-transcriptional gene regulation for the maintenance of cellular homeostasis. They contain different regulatory elements involved in pre-mRNA processing, mRNA stability and translation initiation[224]. Thus, it can be speculated that pVHL30 acts as tumor suppressor also regulating these processes. Further studies will be performed on these non-coding sequences in order to detect possible new linear motifs involved in this particular pVHL30 binding. In addition, sequences mapping provided interesting information from a structural point of view defining for each interactor the region involved in pVHL30 binding. It should be reminded that the preliminary bioinformatics characterization of these interacting fragments was performed so far only for the most represented in the library dataset. It is expected that other valuable information will be obtained extending the analysis to all of the pVHL30 interacting fragments. Collectively, the screening showed that pVHL30 is able to interact with multiple new proteins suggesting novel putative testis-specific VHL functions. Further investigations for each new pVHL30 interactor will be necessary in order to understand the functional meaning beyond these associations.

3.4 The functional interaction between pVHL and AR

Mutations of the von-Hippel Lindau tumor suppressor (pVHL) are causative of VHL syndrome, a hereditary condition associated with tumors arising in multiple organs[225]. The pVHL loss of function mutations (LOF) lead to an increased risk to develop pheochromocytoma, hemangioblastoma and clear cell renal cell carcinoma (ccRCC)[226]. On the other hand, androgen receptor (AR) is a transcription factor normally activated by its natural ligands (testosterone or dihydrotestosterone) and involved in the regulation and maintenance of the male sexual phenotype. In its inactive state, AR localizes to cytosol in association with heat shock proteins (HSPs)[227]. Androgen binding results in dissociation from HSPs, protein stabilization due to decreased degradation by the ubiquitin-proteasome system (UPS), dimerization, translocation to nucleus and binding to androgen-responsive elements (ARE) to regulate genes expression[228]. Structurally, AR is composed of three domains: the amino-terminal domain, the DNA-binding domain (DBD) and the ligand-binding domain (LBD)[152]. DBD and LBD are linked by a hinge region. AR mutations can cause prostate cancer and other types of malignancies, as well as the so-called androgen insensitivity syndromes[229]. These evidence suggest that both pVHL and AR are key factors linked to urological tumors (i.e. kidney, bladder, testis and prostate cancer). Data on cancer incidence show a sex discrepancy with men having a 2-to-4-fold higher risk than women to develop cancer, an incidence that is independent of geographical regions and socio-economic level[230]. The genetic basis for this sex discrepancy in cancer is poorly understood. In 2013, it was demonstrated that pVHL associates with DBD and hinge domains of AR and induces AR ubiquitination. The pVHL–AR association stimulated AR transcriptional activity upon dihydrotestosterone (DHT) treatment[159]. On the other hand, in 2014 pVHL was shown to inhibit AR transcriptional activity without altering AR turnover. In this case the AR region involved in pVHL binding corresponded to LBD[158]. Taken together, information related to the structural and functional interaction between pVHL and AR are conflicting and lack mechanistic insights. For this reason, I decided to understand which functional relationship occurs between these two proteins. Here, I tested the hypothesis that the interaction between pVHL and AR contributes to sex discrepancy in urological malignancies.

First of all, I verified whether pVHL co-localizes in the same subcellular compartment with AR. To address this question, HeLa cells were co-transfected with vectors to

transiently overexpress pVHL30-GFP and AR wild-type (i.e. AR24Q) or the pathogenic poly-Q expanded (i.e. AR65Q). Then, transfected cells were treated with either DHT 0.5 nM or vehicle (EtOH) for 24 h and processed the cells for immunofluorescence analysis. Results are reported in figure 32. As expected DHT 0.5 nM treatment triggered AR translocation from the cytoplasm into the nucleus. Conversely, no differences were observed in pVHL30-GFP localization corresponding to nucleus and cytoplasm in presence and absence of androgens, indicating that these hormones do not affect pVHL subcellular localization. This behavior is shared among AR wild-type and the poly-Q expanded form (AR65Q). The same analysis was also performed using the plasmid expressing GFP only as control (data not shown). Taken together these observations indicate that pVHL30-GFP and AR co-localize in cancer cells.

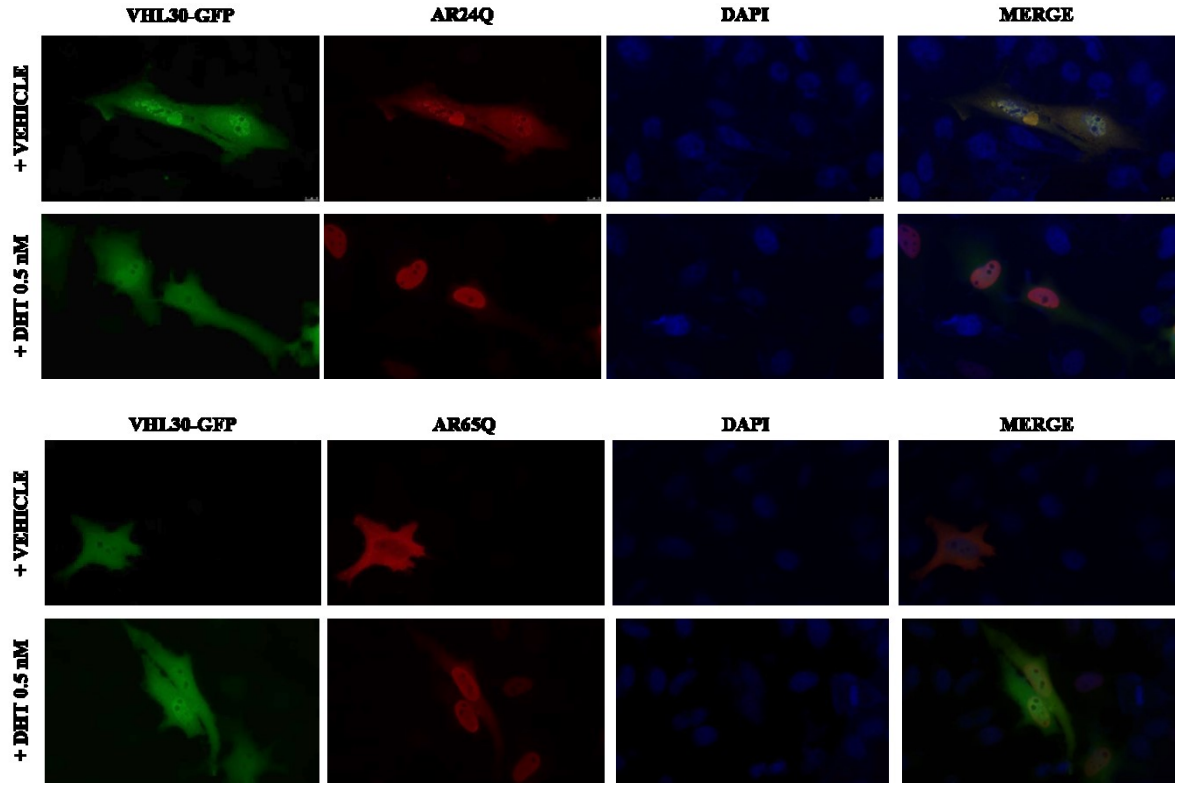


Figure 32 pVHL30 and AR co-localization in cancer cells. HeLa cells were co-transfected with pVHL30-GFP, AR24Q or AR65Q, labeled with anti-AR antibody and visualized by fluorescence microscope. Analysis was performed in the presence and absence of DHT (0.5 nM). Green corresponds to VHL30-GFP, while red is for AR. Nuclei are marked with DAPI. On the right single channels overlapped.

To verify whether pVHL30 forms a complex with AR24Q, I performed a Co-immunoprecipitation in HEK293T cells transiently expressing HA-pVHL30 and FLAG-AR24Q. Cells were treated with DHT 10 nM for 24 hours. As reported in figure 33, AR24Q forms a complex with pVHL30. No signal was detected in sample overexpressing

pVHL30 alone, confirming that pVHL30 co-immunoprecipitation was specifically driven by AR24Q.

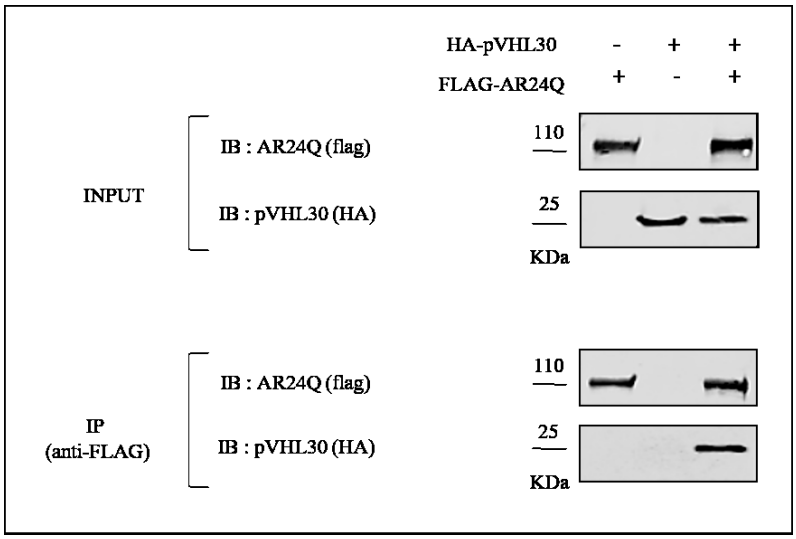


Figure 33 Validation of pVHL30-AR24Q interaction in HEK293T. Cells were transfected with plasmids overexpressing HA-pVHL30 and/or flag-AR24Q proteins, as indicated on the top row. Cells were treated with DHT 10 nM for 24h. Upon AR24 immunoprecipitation with anti-FLAG antibody, presence of pVHL30 in the immunoprecipitates (IP) was detected using anti-HA antibody. n=3

Taking advantage for the Y2H system, I tried to identify which AR24Q portion is responsible for pVHL30 binding. To this purpose, I transformed yeast strain Y190 with recombinant plasmids expressing pVHL30, AR24Q-LBD or AR24Q-NT. The AR24Q-LBD corresponds to the receptor portion involved in ligand binding, located into the C-terminal part of the protein, and corresponding to residues 670-920. Instead, the AR24Q-NT is the N-terminal portion of the receptor corresponding to residues 1-555. Recombinant plasmids used were previously generated and tested in the laboratory. Yeast cells were serially diluted and spotted on permissive and selective medium. Of note, in this case selective medium, which composition is described in materials and methods section, was complemented with DHT 0.5 nM. Indeed, previous data in Y2H indicated that the AR24Q-LBD interacting property is ligand-dependent. To evaluate yeasts growth, in each experiment an internal positive control corresponding to yeast co-expressing AR24Q-LBD and AR24Q-NT was added to the positive (C+) and the negative (C-) ones usually used as reference. Results are summarized in figure 34.

The AR24Q-LBD and the AR24Q-NT did not bind pVHL30, as indicated by the yeast inability to survive on selective medium. Indeed, the yeast strains used here are unable to grow as well as the negative control. These results are preliminary and will be completed

performing the Y2H assay expressing AR24Q full-length (aa 1-920) and the AR24Q hinge region (aa 556-669). Furthermore, as described for pVHL30-MDM2 interaction (Results section 3.1) flanking regions can influence the binding properties of a fragment. Thus, after the validation in Y2H of AR full-length binding to pVHL30 , it could be interesting to generate other AR fragments in order to identify the exact AR portion involved in pVHL30 binding.

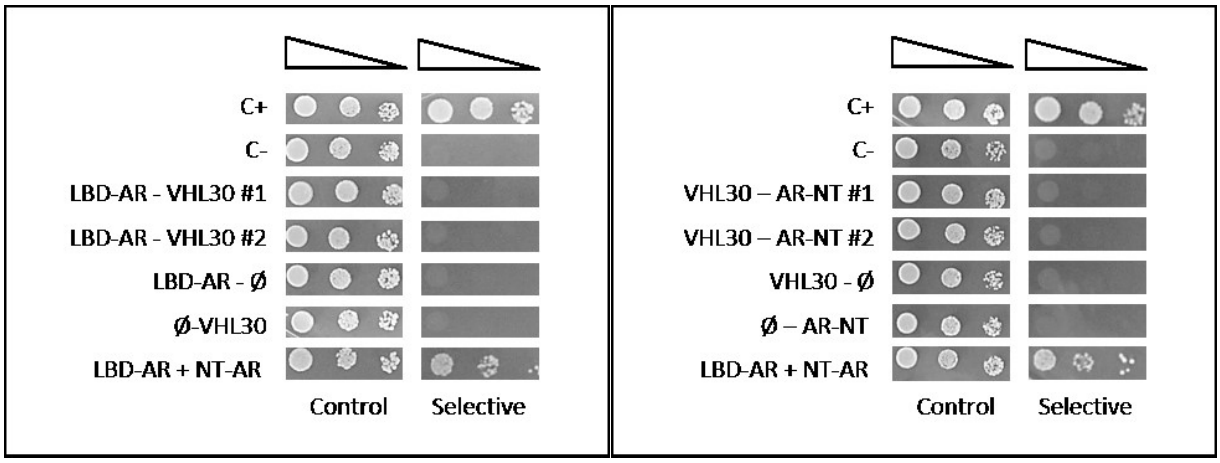


Figure 34 Dissection of the AR binding region. Panels summarize Y2H results of the interaction analysis between pVHL30 and LBD-AR or AR-NT. Serial dilutions of yeast cells were spotted on both permissive (left) and selective (right) medium. C+ and C- corresponds to positive and negative controls. Ø indicated an empty vector (i.e. auto-activation control).

The well-characterized pVHL–HIF-1α interaction is regulated by post-translational modifications. In particular, hydroxylation of proline residues within the oxygen-dependent degradation domain of HIF-1 α (ODD) is required for pVHL binding[44]. Prolines are targeted by prolyl hydroxylase domain containing proteins (PHD)[106]. Considering this aspect, I sought to address whether the pVHL30 – AR interaction is similarly dependent on proline hydroxylation. First of all, a bioinformatics analysis of AR sequence was performed to identify potential proline residues than can be modified by hydroxylation. As shown in figure 35, proline hydroxylation prediction identified multiple potential sites mostly lying within the AR N-terminal portion and the DNA-binding domain. The analysis identified high-confidence predictions for a poly-proline stretch, spanning residues 373-381 (NM_000044). The region immediately upstream this insertion contains a LxLAxP (L is leucine, A alanine, P proline, and X any amino acid) motif, which resembles the canonical pVHL LxxLAP binding motif. Intriguingly, in the same area there is a second and partially overlapping motif, YxxxxLϕ (Y is tyrosine, ϕ any hydrophobic amino acid), which regulates order-to-disorder transitions[231]. Also AR cancer-associated proline

mutations derived from COSMIC were considered. I found the poly-proline stretch to be enriched with mutations found in cancers affecting different tissues.

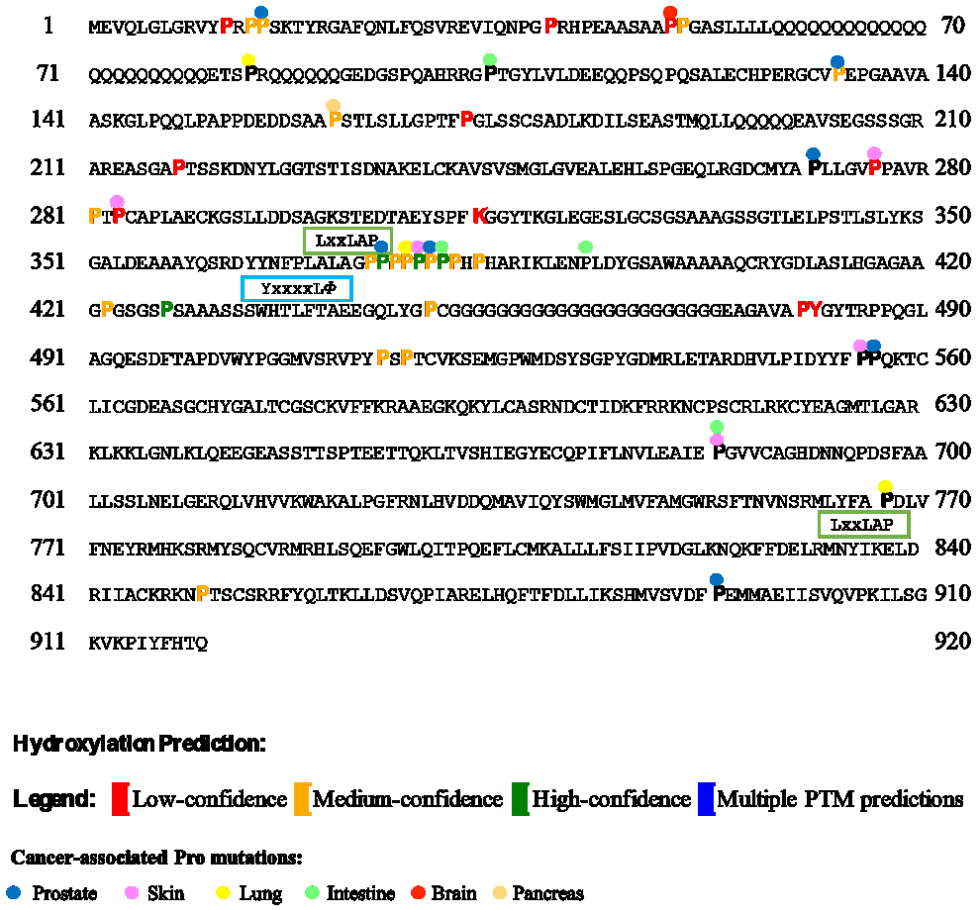


Figure 35 AR sequence with predicted hydroxylation sites and mutations in cancer. Prolines are marked with different colors corresponding to the prediction confidence level. Colored points over the P are for cancer-associated proline mutations and correspondent tissues involved. Rectangles indicate putative binding motifs.

Taken into the account the bioinformatics prediction, I verified whether the pVHL30-AR interaction is regulated by prolines hydroxylation using the three-hybrid system (Y3H). The Y3H was previously employed in the laboratory to visualize the pVHL association with HIF-1α. This system corresponds to a modified Y2H in which the bait vector allows the expression of two proteins simultaneously. Thus, in this case I took advantage from this system to co-express the PHD3, an enzyme involved in prolines hydroxylation, together with AR24Q-NT and pVHL30. Results are summarized in picture 36. Yeast co-expressing pVHL30, ODD (HIF-1α) and PHD3 were used as internal positive control, they survived on selective medium as expected. Instead, the yeast strain of interest failed to grow in the same conditions, indicating that pVHL30 was not able to associate with AR24Q-NT, also

in presence of PHD3. This observation excluded the involvement of PHD3 in AR proline hydroxylation. However, further investigations are required to address whether the prolines hydroxylation is required or not for AR-pVHL30 interaction. Indeed, PHD1 or PHD2, other two member of the same enzyme family, could be involved in AR prolines hydroxylation. On the contrary, the AR – pVHL30 association could be detected in yeast only expressing AR full-length.

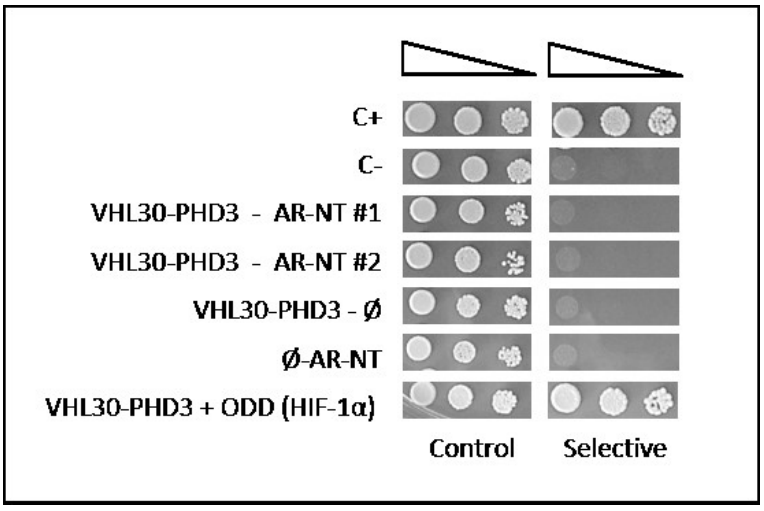


Figure 36 Analysis of pVHL30 interaction with AR24Q-NT in presence of PHD3. Serial dilutions of yeast cells were spotted on both permissive (left) and selective (right) medium. C+ and C- corresponds to positive and negative controls. Ø indicates an empty vector (i.e. auto-activation control).

Considering that also the AR65Q co-localizes with pVHL30 when overexpressed (figure 32), I investigated the relationship between pVHL30 and the AR poly-Q expanded. Currently, no information is available about this AR variant. First of all, I verified whether AR65Q is able to form a complex with pVHL30. Data reported in figure 37, show that AR65Q specifically immunoprecipitates pVHL30 in presence of DHT, similarly to what observed for AR24Q.

From literature it is known that androgens binding regulates multiple aspects of AR function, including subcellular localization (as shown in figure 32), interactions with transcriptional regulators and protein degradation. Ligand binding regulates these activities by altering AR conformation, chaperone composition, and protein interactions[232]. Thus, I addressed whether the interaction between pVHL30 and AR (both wild type and poly-Q expanded) is influenced by DHT. To verify this aspect, cells used to perform the Co-ip experiments were treated with DHT 10 nM or vehicle for 24 hours before the lysis. Results

are summarized in figure 37. As expected, ligand addition induced AR stabilization. Indeed samples treated with DHT show higher AR levels compared to the untreated ones. Despite these differences in AR levels, anti-AR antibody (441-Santacruz) seems to immunoprecipitate AR with the same affinity in all samples. Conversely, the amount of pVHL30 immunoprecipitated changes. In particular, pVHL30 was immunoprecipitated better when the AR is ligand unbounded. These observations suggest that AR have a higher affinity for pVHL30 when is unbounded. On the contrary, affinity decreases when the ligand is bounded to the receptor.

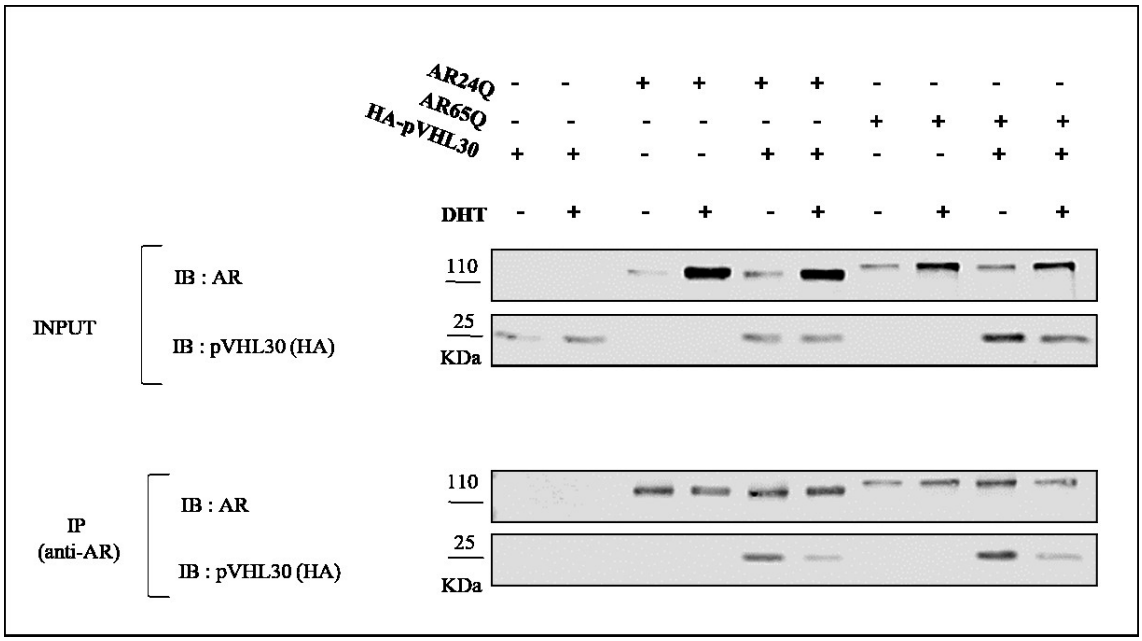


Figure 37 Analysis of AR-pVHL30 association with/without DHT. Cells were transfected with plasmids overexpressing HA-pVHL30 and/or flag-AR24Q or flag-AR65Q proteins, as indicated on the top row. Cells were treated with DHT 10 nM or vehicle for 24h. Upon AR immunoprecipitation with anti-AR antibody, presence of pVHL30 in the immunoprecipitates (IP) was detected using anti-HA antibody. n=3

At this point, I decided to investigate about the functional meaning of these interactions. I asked whether pVHL30 influences AR transcriptional activity. To address this question, *in vitro* luciferase assays was used. I transfected HEK293T cells with non-expanded or expanded polyQ-AR and pVHL30 together with a reporter construct, in which transcription of the firefly luciferase reporter gene is controlled by a promoter recognized by AR (androgen-responsive element, ARE). In addition, I transfected cells with a Renilla-expressing vector, in order to normalize the firefly luciferase signal on transfection efficiency. Transfected cells were treated with DHT 0.5 nM for 24 hours. The treatment triggers AR translocation into the nucleus and AR binding to specific regions of chromatin to induce AR transactivation. As shown in figure 38, in presence of pVHL30 the

transcriptional activity of both non-expanded and expanded polyQ-AR was reduced. This observation was also confirmed by a preliminary RT-PCR on p21, which is an AR target gene (data not shown).

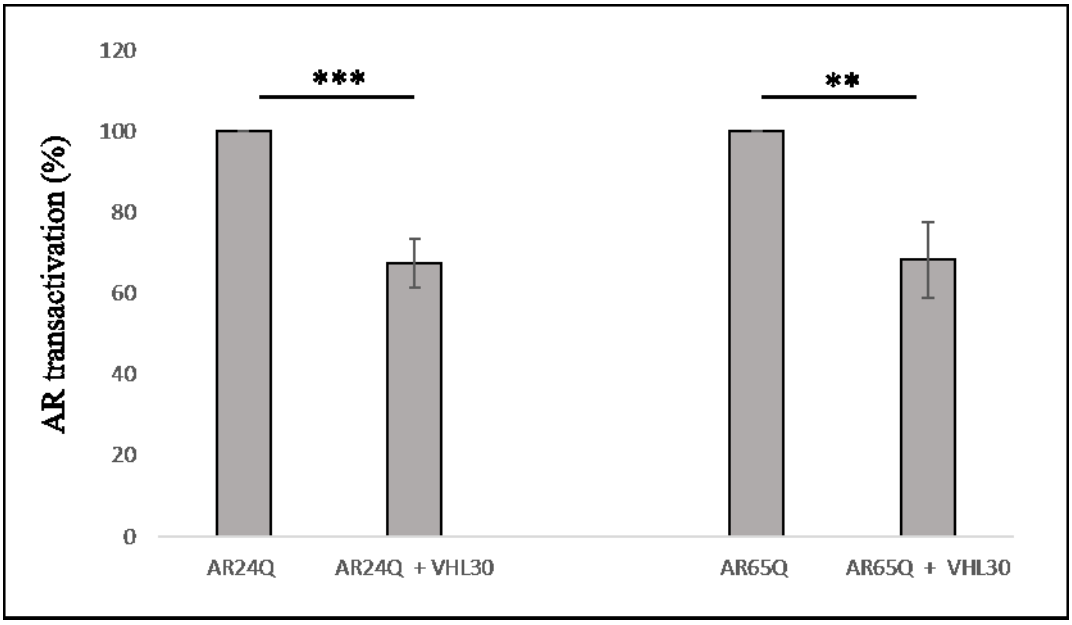


Figure 38 Analysis of pVHL30 effects on AR transactivation activity. AR transactivation was measured by luciferase assay in HEK293T cells transfected as indicated. Graph, mean +/- sem, n = 4, *=p<0.05, **=p<0.01, ***=p<0.001 ,Student t-test.

The same analysis was also performed using the AR24Q and AR65Q ubiquitination defective variant K846AK848A. As reported in picture 39, in this case pVHL30 did not modify the AR transcriptional activity. This result suggests pVHL30 ubiquitinates AR at these sites.

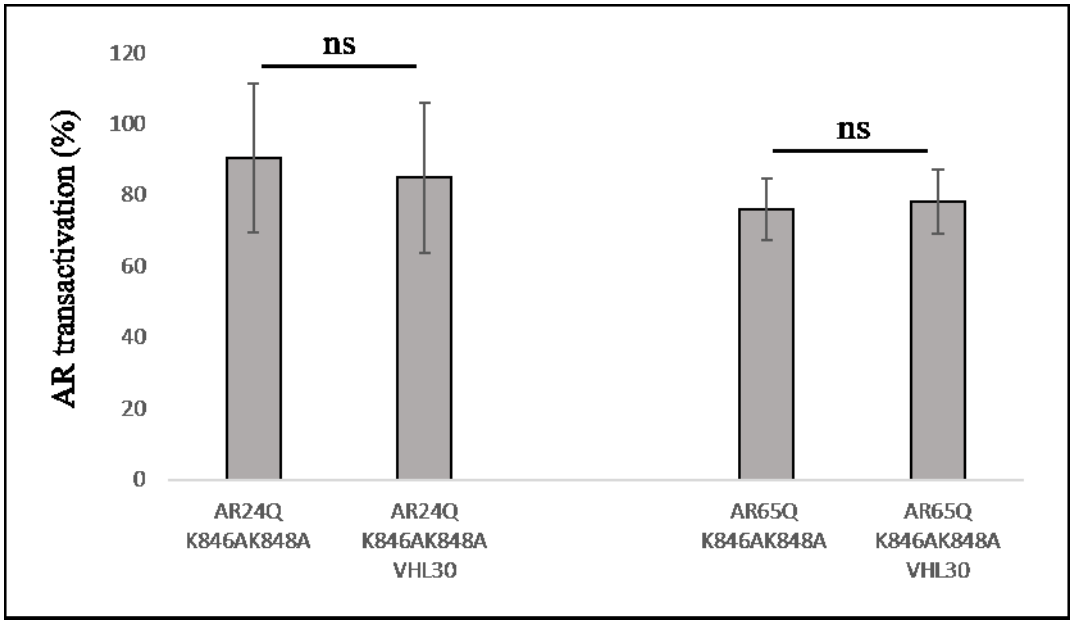
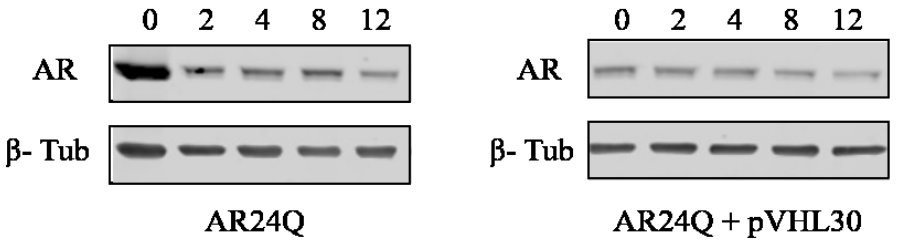


Figure 39 (previous page) Analysis of pVHL30 effects on AR mutant K846AK848A. AR transactivation was measured by luciferase assay in HEK293T cells transfected as indicated. Graph, mean +/- sem, n = 4, p=0.05, ns= non-significant, Student t-test.

Transcriptional assay data suggest the pVHL30 involvement in AR protein degradation. Thus, I asked whether pVHL30 alters AR turnover. I transfected HEK293T cells with AR24Q alone or together with pVHL30, blocked de novo protein synthesis with Cycloheximide 50 μM (CHX) and collected samples at different time points. Then, samples were visualized by western blot and protein levels were quantified with ImageJ. Results are reported in figure 40. The co-expression of pVHL30 accelerates the AR24Q turnover, indeed samples co-expressing the two proteins showed lower AR levels compared to those expressing AR24Q alone. Of note, according to its E3 ubiquitin ligase activity, pVHL30 reduced AR24Q levels also at time point 0 when de novo synthesis is not blocked. Thus, differences in AR24Q levels were visible also at the starting point.



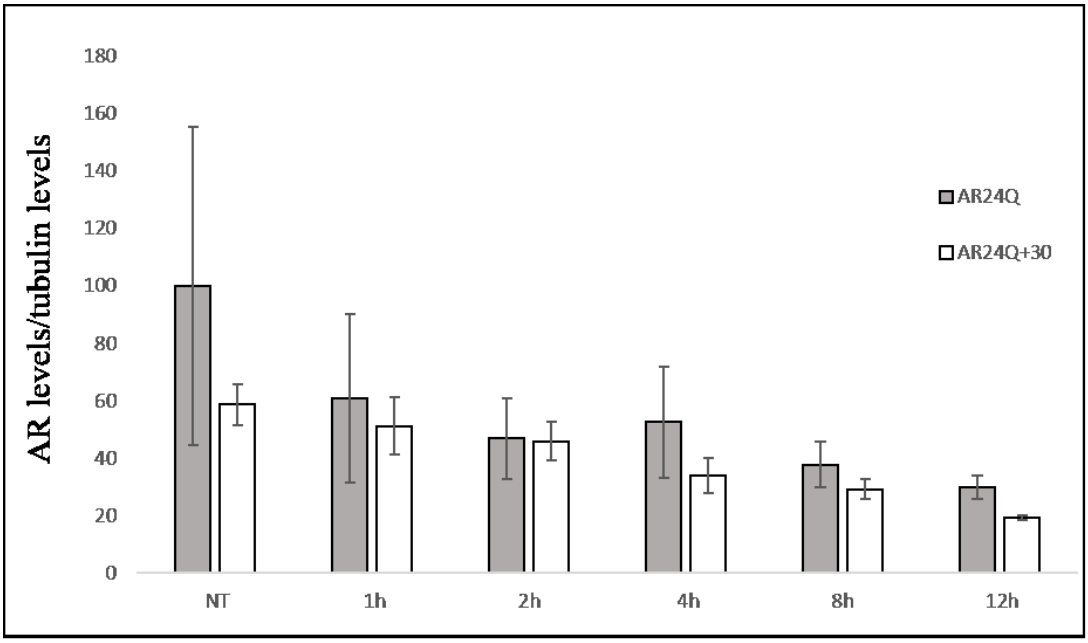


Figure 40 (previous page) Analysis of AR protein turnover with/without pVHL30. Cells were transfected as indicated, treated with CHX 50 μM and collected at different time points. On the top a representative WB. On the bottom the protein quantification is shown in the histogram. Graph, mean +/- sem, n =4.

The same preliminary experiment was also performed for AR65Q (data non shown). The result obtained is the same observed for AR24Q, indeed pVHL30 when co-expressed increases AR65Q protein turnover. On the other hand, pVHL30 did not alter the protein turnover of the ubiquitination defective mutant AR24Q K846AK848A (data not shown). These last experiments were performed only one times and will be repeated to confirm this observation. Collectively, our data indicate that pVHL30 increases both AR24Q and AR65Q turnover. The effect disappears with the AR ubiquitination-defective mutant confirming that probably pVHL30 ubiquitinates AR at residues K846 and K848. To conclude, all evidence demonstrates that pVHL30 co-localizes and interacts with AR24Q and AR65Q. The association is negatively regulated by AR ligand, which indeed stabilizes AR. Data also show that pVHL30 increases AR turnover with consequent reduction in AR levels. Of note, all experiments were performed working in overexpression. Thus, to verify the physiological relevance of this protein-protein interaction further analyses will be performed considering endogenous AR and pVHL.

Discussion

Sex is a key factor affecting the etiology, pathogenesis and prognosis of specific types of cancer. As result several kind of tumors like urological malignances (i.e. kidney, bladder, testis and prostate cancer) show a higher incidence in males compared to females. This sex discrepancy is not due to the environment. It is linked to unknown genetic biological factors. In order to understand the molecular mechanism yielding the sex gap, I focused on the relationship between the androgen receptor and pVHL. Data obtained in this study indicate that pVHL30 overexpression (GOF) leads to decreased AR levels suggesting pVHL to be an E3-ubiquitin ligase involved in AR degradation. The physical association between pVHL30 and AR was negatively regulated by ligand, suggesting that the DHT-dependent conformational changes induce pVHL30 release. This hypothesis could be reasonable considering that ligand binding to AR triggers AR stabilization and the activation of its transcriptional activity. The study presented here is at the beginning and data produced represent a starting point for further investigations. However, it can be speculated that the relationship between AR and the tumor suppressor pVHL30 connect

AR, that is not an oncogene *per se*, to cancer insurgence and development. The androgen receptor has a key role in male reproductive system development and regulation. Thus, it could be proposed that alteration in AR degradation by pVHL30 can contribute to the pathogenesis of urogenital malignancies in males. The identification of the AR-pVHL30 axis, whether confirmed, could represent a viable molecular mechanism explaining the sex discrepancy in cancer onset. On the other hand, according to its role in the VCB complex, pVHL30 could act as substrate recognition particle (SRP) of MDM2, the E3 ubiquitin ligase known to mediate AR degradation. Indeed, it is known that MDM2 together with the kinase Akt form a complex with AR and promote phosphorylation-dependent AR ubiquitylation, resulting in AR degradation by the proteasome[233]. This mechanism could be regulated by oxygen availability considering that pVHL suppresses the kinase activity of Akt in a proline hydroxylation-dependent fashion[145]. Furthermore, as described in results section 3.1, I demonstrated that pVHL30 and MDM2 physically interact. Taken together all of these considerations, suggest that AR, pVHL30, MDM2 and Akt could be part of transient macromolecular complexes intimately involved in the development of multiple human cancers. Further investigations are required to verify which hypothesis is the most promising and to define the functional role of AR and pVHL30 interaction.

4 – Discussion and Conclusions

Data described in this work identified novel pVHL interactors pointing out new alternative functions linking this tumor suppressor to the regulation of cell-cycle. In particular, our findings distinguished a novel isoform-specific interaction with MDM2 which underline the role of pVHL30 N-terminal tail in mediating this association. This evidence, taken together with the previous observation that p14ARF forms a binary association with pVHL30[39] suggests the two pVHL isoforms to have evolved different biological functions. On the other hand, we described a novel association between the pVHL and CDKN1 family members. This interaction is sustained by the pVHL β domain, which recognizes a binding motif conserved among CDKN1s. We proposed all these interactions as further link between cell-cycle progression and the oxygen-sensing mechanism involved in the hypoxia adaptive response. Thus, it can be speculated that transient macromolecular complexes composed of proteins participating in the regulation of both these cellular processes may constantly be formed and destroyed in different cell compartments. This idea of transient complexes is further sustained by the limited number of pVHL binding surfaces compatible with its small dimension. Indeed, more than 500 different partners are known to bind pVHL[32] suggesting that each interaction probably occurs in a particular time and space as well as under different cell conditions. According to this hypothesis, a large-scale screening in testis-tissue identified new 55 tissue-specific interactors suggesting pVHL to participate relevant cellular pathways already associated with this protein (i.e. cell-cycle regulation, DNA damage repair, apoptosis and cytoskeleton regulation) but also others cellular processes not yet characterized, possibly specific of this tissue. An interesting point to take into account is that all of these novel interactions were detected in a hydroxylation-independent manner considering that yeast cells lack enzymes involved in this post-translational modification. It could be assumed that in absence of oxygen, when pVHL is not associated to HIF-1α, it is free to mediate alternative interactions relevant for other cellular pathways. Thus, oxygen could influence the pVHL "fate" to acting as an E3-ubiquitin ligase involved in HIF-1α degradation or as multifunctional adaptor in other cellular processes. Furthermore, data presented in chapter 3.4 showed a functional relationship between pVHL30 and the androgen receptor. In particular pVHL30 overexpression (GOF) yields decreased AR levels suggesting pVHL to be an E3-ubiquitin ligase involved in AR degradation. This evidence is very interesting from the therapeutic point of view. It's known that AR dysfunction leads to several pathological conditions like

prostate cancer, thus targeting the altered AR for degradation could be a useful approach. At the moment a promising strategy named PROTAC[160] based on the ability of small molecules to bring together an E3-ubiquitin ligase and a protein to degrade is in phase I clinical trial. This system was already employed using MDM2 to degrade AR[161]. In particular, a non-steroidal androgen receptor ligand (SARM) was connected to the MDM2 ligand known as nutlin. The SARM-nutlin PROTAC recruits AR to MDM2, which functions as an E3 ubiquitin ligase. This leads to the ubiquitination of AR and its subsequent degradation by the proteasome. Upon treatment of HeLa cells with 10 μM PROTAC for 7 h, authors observed a decrease in AR levels[161]. Thus, according to our data pVHL could be used as substitute of MDM2 in this system. Another interesting point is that a number of studies[234–236] suggest that there is a crosstalk between the AR and HIF-1α pathways. Indeed, it was found that dihydrotestosterone enhances expression of the glucose transporter GLUT-1, one of the HIF-1α target genes. Similarly, it was demonstrated that hypoxia enhances the expression of prostate-specific antigen (PSA) that is one of the AR target genes[236]. Furthermore, chromatin immunoprecipitation assay indicated HIF-1 to interact AR on the human PSA gene promoter[236]. In the context of prostate cancer, these results point HIF-1 to cooperate with AR to activate the expression of several genes related to tumor angiogenesis, invasion, and progression (e.g. glycolytic pathway enzymes, VEGF). Taken into account the ability of pVHL to associate and mark for degradation both the AR and HIF-1α transcription factors, it can be speculated that pVHL orchestrates the "switch" among this two pathways that mutually act to sustain the tumor growth. Collectively taken, findings presented in this work, shed new light to the plethora of pVHL non-canonical functions adding relevant information about pVHL-isoforms specialization, as well as suggesting novel putative testis-specific and sex-linked pVHL functions.

5 – References

1. Naghavi, M. *et al.* 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).

2. Michor, F., Iwasa, Y. & Nowak, M. A. Dynamics of cancer progression. *Nat. Rev. Cancer* **4**, 197–205 (2004).

3. Kinzler, K. W. & Vogelstein, B. Gatekeepers and caretakers. *Nature* **386**, 761–763 (1997).

4. Bissell, M. J. & Radisky, D. Putting tumours in context. *Nat. Rev. Cancer* **1**, 46–54 (2001).

5. Gomy, I. & Diz, M. D. P. E. Hereditary cancer risk assessment: insights and perspectives for the Next-Generation Sequencing era. *Genet. Mol. Biol.* **39**, 184–8 (2016).

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

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

8. Garber, J. E. & Offit, K. Hereditary Cancer Predisposition Syndromes. *J. Clin. Oncol.* **23**, 276–292 (2005).

9. Malkin, D. Li-Fraumeni Syndrome. *Genes Cancer* **2**, 475–484 (2011).

10. Shuin, T. *et al.* Von Hippel–Lindau Disease: Molecular Pathological Basis, Clinical Criteria, Genetic Testing, Clinical Features of Tumors and Treatment. *Jpn. J. Clin. Oncol.* **36**, 337–343 (2006).

11. von Hippel E. Die anatomische Grundlage der von mir beschriebenen‘sehr seltenen Erkrankung der Netzhaut’. *Græfes Arch. Ophthalmol* **79**, 350–377 (1911).

12. Lindau A. Studien Über Kleinhirncysten. Bau, Pathogenese and Beziehungen zur Angiomatosis retinae. *Acta Pathol Microbiol Scand* **1**, 1–128 (1926).

13. Nordstrom-O’Brien, M. *et al.* Genetic analysis of von Hippel-Lindau disease. *Hum. Mutat.* **31**, n/a-n/a (2010).

14. 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).

15. Kim, E. & Zschiedrich, S. Renal Cell Carcinoma in von Hippel-Lindau Disease-From Tumor Genetics to Novel Therapeutic Strategies. *Front. Pediatr.* **6**, 16 (2018).

16. Opocher, G., Conton, P., Schiavi, F., Macino, B. & Mantero, F. Pheochromocytoma in von Hippel–Lindau disease and neurofibromatosis type 1. *Fam. Cancer* **4**, 13–16 (2005).

17. Gaal, J. *et al.* Parasympathetic Paragangliomas Are Part of the Von Hippel-Lindau Syndrome. *J. Clin. Endocrinol. Metab.* **94**, 4367–4371 (2009).

18. van Asselt, S. J. *et al.* Pancreatic cyst development: insights from von Hippel-Lindau disease. *Cilia* **2**, 3 (2013).

19. Choo, D. *et al.* Endolymphatic sac tumors in von Hippel—Lindau disease. *J. Neurosurg.* **100**, 480–487 (2004).

20. Choyke, P. L. *et al.* Epididymal cystadenomas in von Hippel-Lindau disease. *Urology* **49**, 926–31 (1997).

21. Werness, B. A. & Guccion, J. G. Tumor of the broad ligament in von Hippel-Lindau disease of probable mullerian origin. *Int. J. Gynecol. Pathol.* **16**, 282–5 (1997).

22. Ben-Skowronek, I. & Kozaczuk, S. Von Hippel-Lindau Syndrome. *Horm. Res. Paediatr.* **84**, 145–152 (2015).

23. Crespigio, J. *et al.* Von Hippel–Lindau disease: a single gene, several hereditary tumors. *J. Endocrinol. Invest.* **41**, 21–31 (2018).

24. Latif, F. *et al.* Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science (80-.).* **260**, 1317–1320 (1993).

25. Prowse, A. H. *et al.* Somatic inactivation of the VHL gene in Von Hippel-Lindau disease tumors. *Am. J. Hum. Genet.* **60**, 765–71 (1997).

26. Nielsen, S. M. *et al.* Von Hippel-Lindau Disease: Genetics and Role of Genetic Counseling in a Multiple Neoplasia Syndrome. *J. Clin. Oncol.* **34**, 2172–2181 (2016).

27. Woodward, E. R. *et al.* Comparative Sequence Analysis of the VHL Tumor Suppressor Gene. *Genomics* **65**, 253–265 (2000).

28. Zatyka, M. *et al.* Genetic and functional analysis of the von Hippel-Lindau (VHL) tumour suppressor gene promoter. *J. Med. Genet.* **39**, 463–472 (2002).

29. 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.* **95**, 11661–11666 (1998).

30. Stebbins, C. E., Kaelin, W. G. & Pavletich, N. P. Structure of the VHL-ElonginC-ElonginB Complex: Implications for VHL Tumor Suppressor Function. *Science (80-.).* **284**, 455–461 (1999).

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

32. Tabaro, F. *et al.* VHLdb: A database of von Hippel-Lindau protein interactors and mutations. *Sci. Rep.* **6**, 31128 (2016).

33. 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).

34. Richards, F., 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).

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

36. 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).

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

38. Iwai, K. *et al.* Identification of the von Hippel-lindau tumor-suppressor protein as part of an active E3 ubiquitin ligase complex. *Proc. Natl. Acad. Sci. U. S. A.* **96**,

12436–41 (1999).

39. Minervini, G. *et al.* Isoform-specific interactions of the von Hippel-Lindau tumor suppressor protein. *Sci. Rep.* **5**, 12605 (2015).

40. Chesnel, F. *et al.* The von Hippel–Lindau tumour suppressor gene: uncovering the expression of the pVHL172 isoform. *Br. J. Cancer* 1–9 (2015). doi:10.1038/bjc.2015.189

41. Martella, M. *et al.* Molecular analysis of two uncharacterized sequence variants of the VHL gene. *J. Hum. Genet.* **51**, 964–968 (2006).

42. Hascoet, P. *et al.* The pVHL172 isoform is not a tumor suppressor and up-regulates a subset of pro-tumorigenic genes including TGFB1 and MMP13. *Oncotarget* **8**, 75989–76002 (2017).

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

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

45. Kim, W. Y. & Kaelin, W. G. Role of VHL gene mutation in human cancer. *J. Clin. Oncol.* **22**, 4991–5004 (2004).

46. Semenza, G. L. Hypoxia-Inducible Factors in Physiology and Medicine. *Cell* **148**, 399–408 (2012).

47. Schito, L. & Semenza, G. L. Hypoxia-Inducible Factors: Master Regulators of Cancer Progression. *Trends in Cancer* **2**, 758–770 (2016).

48. McKeown, S. R. Defining normoxia, physoxia and hypoxia in tumours—implications for treatment response. *Br. J. Radiol.* **87**, (2014).

49. Jungermann, K. & Kietzmann, T. Oxygen: Modulator of metabolic zonation and disease of the liver. *Hepatology* **31**, 255–260 (2000).

50. Vaupel, P. & Mayer, A. Hypoxia in cancer: significance and impact on clinical outcome. *Cancer Metastasis Rev.* **26**, 225–239 (2007).

51. WENGER, R. H. Cellular adaptation to hypoxia: O_2 -sensing protein hydroxylases, hypoxia-inducible transcription factors, and O_2 -regulated gene expression. *FASEB*

J. **16**, 1151–1162 (2002).

52. Wang, G. L., Jiang, B. H., Rue, E. A. & Semenza, G. L. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proc. Natl. Acad. Sci.* **92**, 5510–5514 (1995).

53. Wu, D., Potluri, N., Lu, J., Kim, Y. & Rastinejad, F. Structural integration in hypoxia-inducible factors. *Nature* **524**, 303–308 (2015).

54. Huang, L. E., Gu, J., Schau, M. & Bunn, H. F. Regulation of hypoxia-inducible factor 1 is mediated by an O2-dependent degradation domain via the ubiquitin-proteasome pathway. *Proc. Natl. Acad. Sci.* **95**, 7987–7992 (1998).

55. Smythies, J. A. *et al.* Inherent DNA -binding specificities of the HIF -1α and HIF -2α transcription factors in chromatin. *EMBO Rep.* **20**, (2019).

56. Gu, Y. Z., Moran, S. M., Hogenesch, J. B., Wartman, L. & Bradfield, C. A. Molecular characterization and chromosomal localization of a third alpha-class hypoxia inducible factor subunit, HIF3alpha. *Gene Expr.* **7**, 205–13 (1998).

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

58. Koh, M. Y. & Powis, G. Passing the baton: the HIF switch. *Trends Biochem. Sci.* **37**, 364–372 (2012).

59. Loenarz, C. & Schofield, C. J. Expanding chemical biology of 2-oxoglutarate oxygenases. *Nat. Chem. Biol.* **4**, 152–156 (2008).

60. Yang, M., Su, H., Soga, T., Kranc, K. R. & Pollard, P. J. Prolyl hydroxylase domain enzymes: important regulators of cancer metabolism. *Hypoxia (Auckland, N.Z.)* **2**, 127–142 (2014).

61. Mahon, P. C., Hirota, K. & Semenza, G. L. FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity. *Genes Dev.* **15**, 2675–2686 (2001).

62. Bruick, R. K. & McKnight, S. L. A Conserved Family of Prolyl-4-Hydroxylases That Modify HIF. *Science (80-.).* **294**, 1337–1340 (2001).

63. McNeill, L. a *et al.* Hypoxia-inducible factor asparaginyl hydroxylase (FIH-1) catalyses hydroxylation at the beta-carbon of asparagine-803. *Biochem. J.* **367**, 571–575 (2002).

64. Hirsilä, M., Koivunen, P., Günzler, V., Kivirikko, K. I. & Myllyharju, J. Characterization of the Human Prolyl 4-Hydroxylases That Modify the Hypoxia-inducible Factor. *J. Biol. Chem.* **278**, 30772–30780 (2003).

65. Pereira, T., Zheng, X., Ruas, J. L., Tanimoto, K. & Poellinger, L. Identification of Residues Critical for Regulation of Protein Stability and the Transactivation Function of the Hypoxia-inducible Factor-1α by the von Hippel-Lindau Tumor Suppressor Gene Product. *J. Biol. Chem.* **278**, 6816–6823 (2003).

66. Rabinowitz, M. H. Inhibition of Hypoxia-Inducible Factor Prolyl Hydroxylase Domain Oxygen Sensors: Tricking the Body into Mounting Orchestrated Survival and Repair Responses. *J. Med. Chem.* **56**, 9369–9402 (2013).

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

68. Metzen, E. *et al.* Intracellular localisation of human HIF-1alpha hydroxylases: implications for oxygen sensing. *J. Cell Sci.* **116**, 1319–1326 (2003).

69. Appelhoff, R. J. *et al.* Differential Function of the Prolyl Hydroxylases PHD1, PHD2, and PHD3 in the Regulation of Hypoxia-inducible Factor. *J. Biol. Chem.* **279**, 38458–38465 (2004).

70. Kishida, T., Stackhouse, T. M., Chen, F., Lerman, M. I. & Zbar, B. Cellular proteins that bind the von Hippel-Lindau disease gene product: mapping of binding domains and the effect of missense mutations. *Cancer Res.* **55**, 4544–8 (1995).

71. Kamura, T. *et al.* Rbx1, a Component of the VHL Tumor Suppressor Complex and SCF Ubiquitin Ligase. *Science (80-.).* **284**, 657–661 (1999).

72. 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**, 732–41 (1998).

73. Lyapina, S. A., Correll, C. C., Kipreos, E. T. & Deshaies, R. J. Human CUL1 forms

an evolutionarily conserved ubiquitin ligase complex (SCF) with SKP1 and an F-box protein. *Proc. Natl. Acad. Sci.* **95**, 7451–7456 (1998).

74. Bai, C. *et al.* SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* **86**, 263–74 (1996).

75. Garrett, K. P. *et al.* Positive regulation of general transcription factor SIII by a tailed ubiquitin homolog. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 7172–6 (1995).

76. Pause, A. *et al.* The von Hippel-Lindau tumor-suppressor gene product forms a stable complex with human CUL-2, a member of the Cdc53 family of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 2156–61 (1997).

77. Nakayama, K. I. & Nakayama, K. Ubiquitin ligases: cell-cycle control and cancer. *Nat. Rev. Cancer* **6**, 369–381 (2006).

78. Wenger, R. H. & Gassmann, M. Oxygen(es) and the hypoxia-inducible factor-1. *Biol. Chem.* **378**, 609–16 (1997).

79. Jośko, J., Gwóźdź, B., Jedrzejowska-Szypułka, H. & Hendryk, S. Vascular endothelial growth factor (VEGF) and its effect on angiogenesis. *Med. Sci. Monit.* **6**, 1047–52

80. Richard, D. E., Berra, E. & Pouyssegur, J. Non-hypoxic pathway mediates the induction of hypoxia inducible factor 1 alpha (HIF-1 alpha) in vascular smooth muscle cells. *J. Biol. Chem.* **275**, 26765–71 (2000).

81. Feldser, D. *et al.* Reciprocal positive regulation of hypoxia-inducible factor 1alpha and insulin-like growth factor 2. *Cancer Res.* **59**, 3915–8 (1999).

82. Krishnamachary, B. *et al.* Regulation of colon carcinoma cell invasion by hypoxia-inducible factor 1. *Cancer Res.* **63**, 1138–43 (2003).

83. Semenza, G. L. Targeting HIF-1 for cancer therapy. *Nat. Rev. Cancer* **3**, 721–732 (2003).

84. Dang, C. V & Semenza, G. L. Oncogenic alterations of metabolism. *Trends Biochem. Sci.* **24**, 68–72 (1999).

85. Chen, C., Pore, N., Behrooz, A., Ismail-Beigi, F. & Maity, A. Regulation of *glut1* mRNA by Hypoxia-inducible Factor-1. *J. Biol. Chem.* **276**, 9519–9525 (2001).

86. Devarajan, P. *et al.* The von Hippel-Lindau Gene Product Inhibits Renal Cell Apoptosis via Bcl-2-dependent Pathways. *J. Biol. Chem.* **276**, 40599–40605 (2001).

87. 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).

88. Hergovich, A., Lisztwan, J., Barry, R., Ballschmieter, P. & Krek, W. Regulation of microtubule stability by the von Hippel-Lindau tumour suppressor protein pVHL. *Nat. Cell Biol.* **5**, 64–70 (2003).

89. Lutz, M. S. & Burk, R. D. Primary Cilium Formation Requires von Hippel-Lindau Gene Function in Renal-Derived Cells. *Cancer Res.* **66**, 6903–6907 (2006).

90. Stickle, N. H. *et al.* pVHL modification by NEDD8 is required for fibronectin matrix assembly and suppression of tumor development. *Mol. Cell. Biol.* **24**, 3251–61 (2004).

91. Russell, R. C. & Ohh, M. The Role of VHL in the Regulation of E-Cadherin: A New Connection in an Old Pathway. *Cell Cycle* **6**, 56–59 (2007).

92. Bader, H. L. & Hsu, T. Systemic VHL gene functions and the VHL disease. *FEBS Lett.* **586**, 1562–1569 (2012).

93. Frew, I. J. & Krek, W. pVHL: A Multipurpose Adaptor Protein. *Sci. Signal.* **1**, pe30–pe30 (2008).

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

95. Uversky, V. N. & Dunker, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta* **1804**, 1231–64 (2010).

96. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. & Jones, D. T. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *J. Mol. Biol.* **337**, 635–645 (2004).

97. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M. & Obradović, Z. Intrinsic disorder and protein function. *Biochemistry* **41**, 6573–82 (2002).

98. 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).

99. Iakoucheva, L. M. *et al.* The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**, 1037–1049 (2004).

100. Collins, M. O., Yu, L., Campuzano, I., Grant, S. G. N. & Choudhary, J. S. Phosphoproteomic Analysis of the Mouse Brain Cytosol Reveals a Predominance of Protein Phosphorylation in Regions of Intrinsic Sequence Disorder. *Mol. Cell. Proteomics* **7**, 1331–1348 (2008).

101. Davey, N. E. *et al.* Attributes of short linear motifs. *Mol. Biosyst.* **8**, 268–81 (2012).

102. Diella, F. *et al.* Understanding eukaryotic linear motifs and their role in cell signaling and regulation. *Front. Biosci.* **13**, 6580–603 (2008).

103. Van Roey, K. *et al.* Short Linear Motifs: Ubiquitous and Functionally Diverse Protein Interaction Modules Directing Cell Regulation. *Chem. Rev.* (2014). doi:10.1021/cr400585q

104. Mészáros, B., Kumar, M., Gibson, T. J., Uyar, B. & Dosztányi, Z. Degrons in cancer. *Sci. Signal.* **10**, eaak9982 (2017).

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

106. McDonough, M. A. *et al.* Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc. Natl. Acad. Sci.* **103**, 9814–9819 (2006).

107. Masoumi-Moghaddam, S., Amini, A. & Morris, D. L. The developing story of Sprouty and cancer. *Cancer Metastasis Rev.* **33**, 695–720 (2014).

108. Peier, M., Walpen, T., Christofori, G., Battegay, E. & Humar, R. Sprouty2 expression controls endothelial monolayer integrity and quiescence. *Angiogenesis* **16**, 455–68 (2013).

109. Frank, M. J. *et al.* Expression of sprouty2 inhibits B-cell proliferation and is epigenetically silenced in mouse and human B-cell lymphomas. *Blood* **113**, 2478–2487 (2009).

110. Anderson, K. *et al.* Regulation of Cellular Levels of Sprouty2 Protein by Prolyl

Hydroxylase Domain and von Hippel-Lindau Proteins. *J. Biol. Chem.* **286**, 42027–42036 (2011).

111. Hacohen, N., Kramer, S., Sutherland, D., Hiromi, Y. & Krasnow, M. A. sprouty encodes a novel antagonist of FGF signaling that patterns apical branching of the Drosophila airways. *Cell* **92**, 253–63 (1998).

112. Impagnatiello, M.-A. *et al.* Mammalian Sprouty-1 and -2 Are Membrane-Anchored Phosphoprotein Inhibitors of Growth Factor Signaling in Endothelial Cells. *J. Cell Biol.* **152**, 1087–1098 (2001).

113. Heir, P. *et al.* Oxygen-dependent Regulation of Erythropoietin Receptor Turnover and Signaling. *J. Biol. Chem.* **291**, 7357–7372 (2016).

114. Kapitsinou, P. P. *et al.* Hepatic HIF-2 regulates erythropoietic responses to hypoxia in renal anemia. *Blood* **116**, 3039–3048 (2010).

115. Shuai, K. & Liu, B. Regulation of JAK–STAT signalling in the immune system. *Nat. Rev. Immunol.* **3**, 900–911 (2003).

116. Jelkmann, W. Regulation of erythropoietin production. *J. Physiol.* **589**, 1251–1258 (2011).

117. Russell, R. C. *et al.* Loss of JAK2 regulation via a heterodimeric VHL-SOCS1 E3 ubiquitin ligase underlies Chuvash polycythemia. *Nat. Med.* **17**, 845–853 (2011).

118. Le Couedic, J. *et al.* Missense mutation of the erythropoietin receptor is a rare event in human erythroid malignancies. *Blood* **87**, (1996).

119. Gardie, B. *et al.* The role of PHD2 mutations in the pathogenesis of erythrocytosis. *Hypoxia (Auckland, N.Z.)* **2**, 71–90 (2014).

120. Xie, L. *et al.* Oxygen-regulated beta(2)-adrenergic receptor hydroxylation by EGLN3 and ubiquitylation by pVHL. *Sci. Signal.* **2**, ra33 (2009).

121. Cheong, H. I. *et al.* Hypoxia sensing through β-adrenergic receptors. *JCI Insight* **1**, (2016).

122. Jones, L. C., Okino, S. T., Gonda, T. J. & Whitlock, J. P. Myb-binding Protein 1a Augments AhR-dependent Gene Expression. *J. Biol. Chem.* **277**, 22515–22519 (2002).

123. Fan, M. *et al.* Suppression of mitochondrial respiration through recruitment of p160 myb binding protein to PGC-1 : modulation by p38 MAPK. *Genes Dev.* **18**, 278–289 (2004).

124. Hara, Y. *et al.* Molecular characterization of Mybbp1a as a co-repressor on the Period2 promoter. *Nucleic Acids Res.* **37**, 1115–26 (2009).

125. Lai, Y., Qiao, M., Song, M., Weintraub, S. T. & Shiio, Y. Quantitative Proteomics Identifies the Myb-Binding Protein p160 as a Novel Target of the von Hippel-Lindau Tumor Suppressor. *PLoS One* **6**, e16975 (2011).

126. Akaogi, K., Ono, W., Hayashi, Y., Kishimoto, H. & Yanagisawa, J. MYBBP1A suppresses breast cancer tumorigenesis by enhancing the p53 dependent anoikis. *BMC Cancer* **13**, 65 (2013).

127. Kurban, G. *et al.* Collagen matrix assembly is driven by the interaction of von Hippel–Lindau tumor suppressor protein with hydroxylated collagen IV alpha 2. *Oncogene* **27**, 1004–1012 (2008).

128. Schafer, K. A. The Cell Cycle: A Review. *Vet. Pathol.* **35**, 461–478 (1998).

129. Morgan, D. O. Principles of CDK regulation. *Nature* **374**, 131–134 (1995).

130. Pines J. Cyclins: wheels within wheels. *Cell Growth Differ* **2**, 305–10 (1991).

131. Vermeulen, K., Van Bockstaele, D. R. & Berneman, Z. N. The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Prolif.* **36**, 131–149 (2003).

132. Besson, A., Dowdy, S. F. & Roberts, J. M. CDK Inhibitors: Cell Cycle Regulators and Beyond. *Dev. Cell* **14**, 159–169 (2008).

133. Shaw, P. H. The Role of p53 in Cell Cycle Regulation. *Pathol. - Res. Pract.* **192**, 669–675 (1996).

134. Levine, A. J. p53, the cellular gatekeeper for growth and division. *Cell* **88**, 323–31 (1997).

135. Oren, M. Regulation of the p53 Tumor Suppressor Protein. *J. Biol. Chem.* **274**, 36031–36034 (1999).

136. Zhang, Y., Xiong, Y. & Yarbrough, W. G. ARF promotes MDM2 degradation and

stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. *Cell* **92**, 725–34 (1998).

137. Gartel, A. L. & Shchors, K. Mechanisms of c-myc-mediated transcriptional repression of growth arrest genes. *Exp. Cell Res.* **283**, 17–21 (2003).

138. Zhou, B. P. *et al.* Cytoplasmic localization of p21Cip1/WAF1 by Akt-induced phosphorylation in HER-2/neu-overexpressing cells. *Nat. Cell Biol.* **3**, 245–252 (2001).

139. Liang, J. *et al.* PKB/Akt phosphorylates p27, impairs nuclear import of p27 and opposes p27-mediated G1 arrest. *Nat. Med.* **8**, 1153–1160 (2002).

140. Lin, H.-K. *et al.* Phosphorylation-dependent regulation of cytosolic localization and oncogenic function of Skp2 by Akt/PKB. *Nat. Cell Biol.* **11**, 420–432 (2009).

141. Ashcroft, M. *et al.* Phosphorylation of HDM2 by Akt. *Oncogene* **21**, 1955–1962 (2002).

142. Roe, J.-S. *et al.* p53 stabilization and transactivation by a von Hippel-Lindau protein. *Mol. Cell* **22**, 395–405 (2006).

143. Essers, P. B. *et al.* The von Hippel–Lindau tumor suppressor regulates programmed cell death 5-mediated degradation of Mdm2. *Oncogene* **34**, 771–779 (2015).

144. 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).

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

146. Ampofo, E. *et al.* Phosphorylation of the von Hippel–Lindau protein (VHL) by protein kinase CK2 reduces its protein stability and affects p53 and HIF-1α mediated transcription. *Int. J. Biochem. Cell Biol.* **42**, 1729–1735 (2010).

147. Roe, J.-S. *et al.* Phosphorylation of von Hippel-Lindau protein by checkpoint kinase 2 regulates p53 transactivation. *Cell Cycle* **10**, 3920–3928 (2011).

148. Cook, M. B., McGlynn, K. A., Devesa, S. S., Freedman, N. D. & Anderson, W. F. Sex disparities in cancer mortality and survival. *Cancer Epidemiol. Biomarkers*

Prev. **20**, 1629–1637 (2011).

149. Edgren, G., Liang, L., Adami, H.-O. & Chang, E. T. Enigmatic sex disparities in cancer incidence. *Eur. J. Epidemiol.* **27**, 187–196 (2012).

150. Gulía, C. *et al.* Androgen insensitivity syndrome. *Eur. Rev. Med. Pharmacol. Sci.* **22**, 3873–3887 (2018).

151. Gandhi, J. *et al.* The molecular biology of prostate cancer: current understanding and clinical implications. *Prostate Cancer Prostatic Dis.* **21**, 22–36 (2018).

152. Tan, M. E., Li, J., Xu, H. E., Melcher, K. & Yong, E. Androgen receptor: structure, role in prostate cancer and drug discovery. *Acta Pharmacol. Sin.* **36**, 3–23 (2015).

153. Giovannucci, E. *et al.* The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc. Natl. Acad. Sci.* **94**, 3320–3323 (1997).

154. Stanford, J. L. *et al.* Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer Res.* **57**, 1194–8 (1997).

155. Tut, T. G., Ghadessy, F. J., Trifiro, M. A., Pinsky, L. & Yong, E. L. Long Polyglutamine Tracts in the Androgen Receptor Are Associated with Reduced *Trans* -Activation, Impaired Sperm Production, and Male Infertility [1]. *J. Clin. Endocrinol. Metab.* **82**, 3777–3782 (1997).

156. Spada, A. R. La, Wilson, E. M., Lubahn, D. B., Harding, A. E. & Fischbeck, K. H. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**, 77–79 (1991).

157. Koryakina, Y., Ta, H. Q. & Gioeli, D. Androgen receptor phosphorylation: biological context and functional consequences. *Endocr. Relat. Cancer* **21**, T131-45 (2014).

158. Wang, J. *et al.* The Von Hippel-Lindau Protein Suppresses Androgen Receptor Activity. *Mol. Endocrinol.* **28**, 239–248 (2014).

159. Chen, S., Chen, K., Zhang, Q., Cheng, H. & Zhou, R. Regulation of the transcriptional activation of the androgen receptor by the UXT-binding protein VHL. *Biochem. J.* **456**, 55–66 (2013).

160. Scheepstra, M., Hekking, K. F. W., van Hijfte, L. & Folmer, R. H. A. Bivalent

Ligands for Protein Degradation in Drug Discovery. *Comput. Struct. Biotechnol. J.* **17**, 160–176 (2019).

161. Schneekloth, A. R., Pucheault, M., Tae, H. S. & Crews, C. M. Targeted intracellular protein degradation induced by a small molecule: En route to chemical proteomics. *Bioorg. Med. Chem. Lett.* **18**, 5904–5908 (2008).

162. Brachmann, R. K. & Boeke, J. D. Tag games in yeast: the two-hybrid system and beyond. *Curr. Opin. Biotechnol.* **8**, 561–8 (1997).

163. 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**, 22995–22999 (1997).

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

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

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

167. 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).

168. Bradford, M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **72**, 248–254 (1976).

169. Hakhverdyan, Z. *et al.* Rapid, optimized interactomic screening. *Nat. Methods* **12**, 553–560 (2015).

170. 10X Running buffer. *Cold Spring Harb. Protoc.* **2006**, pdb.rec10475 (2006).

171. Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).

172. Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models of

Biomolecular Interaction Networks. *Genome Res.* **13**, 2498–2504 (2003).

173. UniProt Consortium, T. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **46**, 2699–2699 (2018).

174. 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).

175. Train, C.-M., Glover, N. M., Gonnet, G. H., Altenhoff, A. M. & Dessimoz, C. Orthologous Matrix (OMA) algorithm 2.0: more robust to asymmetric evolutionary rates and more scalable hierarchical orthologous group inference. *Bioinformatics* **33**, i75–i82 (2017).

176. 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).

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

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

179. Finn, R. D. *et al.* The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016).

180. Dinkel, H. *et al.* ELM 2016—data update and new functionality of the eukaryotic linear motif resource. *Nucleic Acids Res.* **44**, D294–D300 (2016).

181. 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).

182. Leaver-Fay, A. *et al.* Rosetta3. in *Methods in enzymology* **487**, 545–574 (2011).

183. Wang, R. Y.-R. *et al.* Modeling Disordered Regions in Proteins Using Rosetta. *PLoS One* **6**, e22060 (2011).

184. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–

2637 (1983).

185. 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).

186. Gasteiger, E. *et al.* ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784–3788 (2003).

187. Qiu, W.-R., Sun, B.-Q., Xiao, X., Xu, Z.-C. & Chou, K.-C. iHyd-PseCp: Identify hydroxyproline and hydroxylysine in proteins by incorporating sequence-coupled effects into general PseAAC. *Oncotarget* **7**, 44310–44321 (2016).

188. Jia, C.-Z., He, W.-Y. & Yao, Y.-H. OH-PRED: prediction of protein hydroxylation sites by incorporating adapted normal distribution bi-profile Bayes feature extraction and physicochemical properties of amino acids. *J. Biomol. Struct. Dyn.* **35**, 1–7 (2016).

189. Ismail, H. D., Newman, R. H. & Kc, D. B. RF-Hydroxysite: a random forest based predictor for hydroxylation sites. *Mol. Biosyst.* **12**, 2427–35 (2016).

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

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

192. Sherr, C. J. Divorcing ARF and p53: an unsettled case. *Nat. Rev. Cancer* **6**, 663–73 (2006).

193. Minervini, G. *et al.* Novel interactions of the von Hippel-Lindau (pVHL) tumor suppressor with the CDKN1 family of cell cycle inhibitors. *Sci. Rep.* **7**, 46562 (2017).

194. Riley, M. F. & Lozano, G. The Many Faces of MDM2 Binding Partners. *Genes Cancer* **3**, 226–39 (2012).

195. Challen, C., Anderson, J. J., Chrzanowska-Lightowlers, Z. M. A., Lightowlers, R. N. & Lunec, J. Recombinant human MDM2 oncoprotein shows sequence composition selectivity for binding to both RNA and DNA. *Int. J. Oncol.* **40**, 851–9 (2012).

196. 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).

197. Semenza, G. L. VHL and p53: Tumor Suppressors Team Up to Prevent Cancer. *Mol. Cell* **22**, 437–439 (2006).

198. Sutton, T. A. *et al.* p53 regulates renal expression of HIF-1{alpha} and pVHL under physiological conditions and after ischemia-reperfusion injury. *Am. J. Physiol. Renal Physiol.* **295**, F1666-77 (2008).

199. Hammond, E. M. & Giaccia, A. J. The role of p53 in hypoxia-induced apoptosis. *Biochem. Biophys. Res. Commun.* **331**, 718–725 (2005).

200. Alarcón, R., Koumenis, C., Geyer, R. K., Maki, C. G. & Giaccia, A. J. Hypoxia induces p53 accumulation through MDM2 down-regulation and inhibition of E6-mediated degradation. *Cancer Res.* **59**, 6046–51 (1999).

201. Chen, D., Li, M., Luo, J. & Gu, W. Direct Interactions between HIF-1α and Mdm2 Modulate p53 Function. *J. Biol. Chem.* **278**, 13595–13598 (2003).

202. Li, M. *et al.* Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* **416**, 648–653 (2002).

203. Cheng, Q., Chen, L., Li, Z., Lane, W. S. & Chen, J. ATM activates p53 by regulating MDM2 oligomerization and E3 processivity. *EMBO J.* **28**, 3857–3867 (2009).

204. Piovesan, D. & Tosatto, S. C. E. Mobi 2.0: an improved method to define intrinsic disorder, mobility and linear binding regions in protein structures. *Bioinformatics* **34**, 122–123 (2018).

205. 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).

206. Kaelin Jr, W. G. The von Hippel–Lindau tumour suppressor protein: O2 sensing and cancer. *Nat. Rev. Cancer* **8**, 865–873 (2008).

207. Ong, K. R. *et al.* Genotype-phenotype correlations in von Hippel-Lindau disease. *Hum. Mutat.* **28**, 143–149 (2007).

208. Li, Z. *et al.* Ubiquitination of a novel deubiquitinating enzyme requires direct binding to von Hippel-Lindau tumor suppressor protein. *J. Biol. Chem.* **277**, 4656–62 (2002).

209. Chen, K. *et al.* ZBRK1, a novel tumor suppressor, activates VHL gene transcription through formation of a complex with VHL and p300 in renal cancer. *Oncotarget* **6**, 6959–6976 (2015).

210. Reyes, J. G. *et al.* The Hypoxic Testicle: Physiology and Pathophysiology. *Oxid. Med. Cell. Longev.* **2012**, 1–15 (2012).

211. Rozewicki, J., Li, S., Amada, K. M., Standley, D. M. & Katoh, K. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res.* **47**, W5–W10 (2019).

212. Li, Z., Wang, D., Messing, E. M. & Wu, G. VHL protein-interacting deubiquitinating enzyme 2 deubiquitinates and stabilizes HIF-1α. *EMBO Rep.* **6**, 373–378 (2005).

213. Li, Z. *et al.* Identification of a deubiquitinating enzyme subfamily as substrates of the von Hippel–Lindau tumor suppressor. *Biochem. Biophys. Res. Commun.* **294**, 700–709 (2002).

214. Neduva, V. & Russell, R. B. DILIMOT: discovery of linear motifs in proteins. *Nucleic Acids Res.* **34**, W350-5 (2006).

215. Khacho, M. *et al.* eEF1A Is a Novel Component of the Mammalian Nuclear Protein Export Machinery. *Mol. Biol. Cell* **19**, 5296–5308 (2008).

216. Browne, G. J. & Proud, C. G. Regulation of peptide-chain elongation in mammalian cells. *Eur. J. Biochem.* **269**, 5360–5368 (2002).

217. Abbas, W., Kumar, A. & Herbein, G. The eEF1A Proteins: At the Crossroads of Oncogenesis, Apoptosis, and Viral Infections. *Front. Oncol.* **5**, 75 (2015).

218. Bradsher, J. N., Jackson, K. W., Conaway, R. C. & Conaway, J. W. RNA polymerase II transcription factor SIII. I. Identification, purification, and properties. *J. Biol. Chem.* **268**, 25587–93 (1993).

219. Kamura, T. *et al.* Activation of HIF1alpha ubiquitination by a reconstituted von Hippel-Lindau (VHL) tumor suppressor complex. *Proc. Natl. Acad. Sci. U. S. A.* **97**,

10430–5 (2000).

220. Gnarra, J. R. *et al.* Molecular cloning of the von Hippel-Lindau tumor suppressor gene and its role in renal carcinoma. *Biochim. Biophys. Acta* **1242**, 201–10 (1996).

221. Ye, Y. *et al.* Subcellular localization of the von Hippel-Lindau disease gene product is cell cycle-dependent. *Int. J. cancer* **78**, 62–9 (1998).

222. Liu, Y. *et al.* Yeast two-hybrid junk sequences contain selected linear motifs. *Nucleic Acids Res.* **39**, e128 (2011).

223. Roberts, G. G., Parrish, J. R., Mangiola, B. A. & Finley, R. L. High-Throughput Yeast Two-Hybrid Screening. in *Methods in molecular biology (Clifton, N.J.)* **812**, 39–61 (2012).

224. Schuster, S. L. & Hsieh, A. C. The Untranslated Regions of mRNAs in Cancer. *Trends in Cancer* **5**, 245–262 (2019).

225. Vortmeyer, A. O. *et al.* Somatic point mutation of the wild-type allele detected in tumors of patients with VHL germline deletion. *Oncogene* **21**, 1167–70 (2002).

226. Maher, E. R. & Kaelin, W. G. von Hippel-Lindau disease. *Medicine (Baltimore).* **76**, 381–91 (1997).

227. Pratt, W. B. & Toft, D. O. Steroid Receptor Interactions with Heat Shock Protein and Immunophilin Chaperones*. *Endocr. Rev.* **18**, 306–360 (1997).

228. Li, J. & Al-Azzawi, F. Mechanism of androgen receptor action. *Maturitas* **63**, 142–148 (2009).

229. QUIGLEY, C. A. *et al.* Androgen Receptor Defects: Historical, Clinical, and Molecular Perspectives*. *Endocr. Rev.* **16**, 271–321 (1995).

230. Cook, M. B. *et al.* Sex Disparities in Cancer Incidence by Period and Age. *Cancer Epidemiol. Biomarkers Prev.* **18**, 1174–1182 (2009).

231. Bah, A. *et al.* Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519**, 106–109 (2015).

232. Yang, C.-S. *et al.* Ligand binding to the androgen receptor induces conformational changes that regulate phosphatase interactions. *Mol. Cell. Biol.* **27**, 3390–404 (2007).

233. Lin, H.-K., Wang, L., Hu, Y.-C., Altuwaijri, S. & Chang, C. Phosphorylation-dependent ubiquitylation and degradation of androgen receptor by Akt require Mdm2 E3 ligase. *EMBO J.* **21**, 4037–48 (2002).

234. Mitani, T. *et al.* Hypoxia enhances transcriptional activity of androgen receptor through hypoxia-inducible factor-1α in a low androgen environment. *J. Steroid Biochem. Mol. Biol.* **123**, 58–64 (2011).

235. Mabjeesh, N. J., Willard, M. T., Frederickson, C. E., Zhong, H. & Simons, J. W. Androgens stimulate hypoxia-inducible factor 1 activation via autocrine loop of tyrosine kinase receptor/phosphatidylinositol 3'-kinase/protein kinase B in prostate cancer cells. *Clin. Cancer Res.* **9**, 2416–25 (2003).

236. Horii, K. *et al.* Androgen-Dependent Gene Expression of Prostate-Specific Antigen Is Enhanced Synergistically by Hypoxia in Human Prostate Cancer Cells. *Mol. Cancer Res.* **5**, 383–391 (2007).

237. Wilson, W. R. & Hay, M. P. Targeting hypoxia in cancer therapy. *Nat. Rev. Cancer* **11**, 393–410 (2011).

238. Bárdos, J. I. & Ashcroft, M. Negative and positive regulation of HIF-1: A complex network. *Biochim. Biophys. Acta - Rev. Cancer* **1755**, 107–120 (2005).

239. Kim, J., Tchernyshyov, I., Semenza, G. L. & Dang, C. V. HIF-1-mediated expression of pyruvate dehydrogenase kinase: A metabolic switch required for cellular adaptation to hypoxia. *Cell Metab.* **3**, 177–185 (2006).

240. Pugh, C. W. & Ratcliffe, P. J. Regulation of angiogenesis by hypoxia: role of the HIF system. *Nat. Med.* **9**, 677–684 (2003).

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

242. Iliopoulos, O., Kibel, A., Gray, S. & Kaelin, W. G. Tumour suppression by the human von Hippel-Lindau gene product. *Nat. Med.* **1**, 822–6 (1995).

243. Welford, S. M. & Giaccia, A. J. Hypoxia and senescence: the impact of oxygenation on tumor suppression. *Mol. Cancer Res.* **9**, 538–44 (2011).

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

245. 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).

246. Koshiji, M. *et al.* HIF-1α induces cell cycle arrest by functionally counteracting Myc. *EMBO J.* **23**, 1949–1956 (2004).

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

248. Nakayama, K. & Nakayama, K. Cip/Kip cyclin-dependent kinase inhibitors: brakes of the cell cycle engine during development. *BioEssays* **20**, 1020–1029 (1999).

6 – Supplementary materials

Results section 3.1

Figure S1

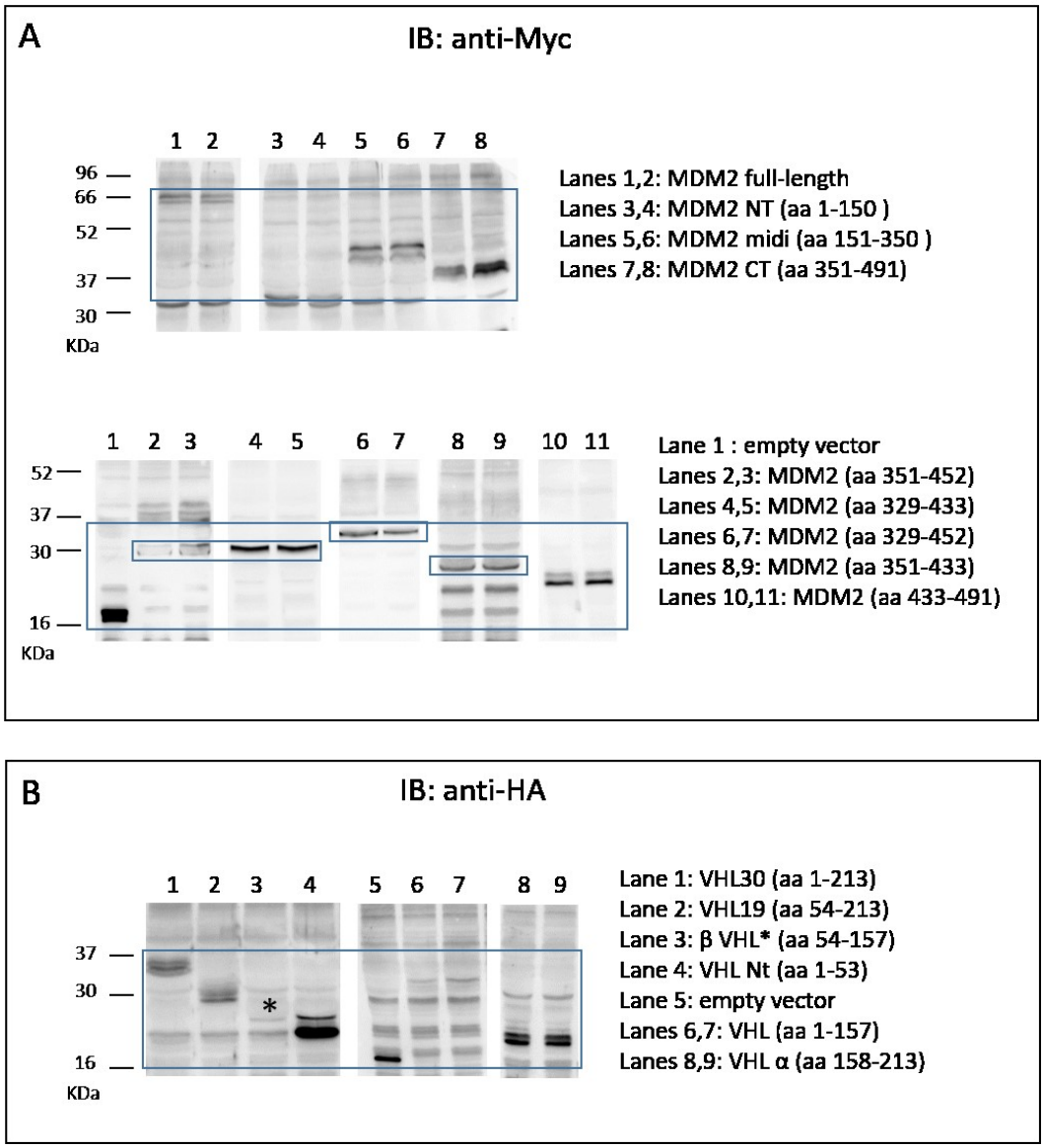


Fig. S1 Western blots to confirm expression of Gal4-fusion proteins in yeast

Proteins were extracted as described in details in materials and methods section and visualized in Western Blot. The bait proteins (i.e. MDM2 full-length and its fragments) are detected by anti-Myc antibody (panel A); whereas the prey proteins (i.e. p VHL30 and its fragments) are identified using the anti-HA (panel B). Each line correspond to the protein or the protein fragment as indicated in the legend.

Figure S2

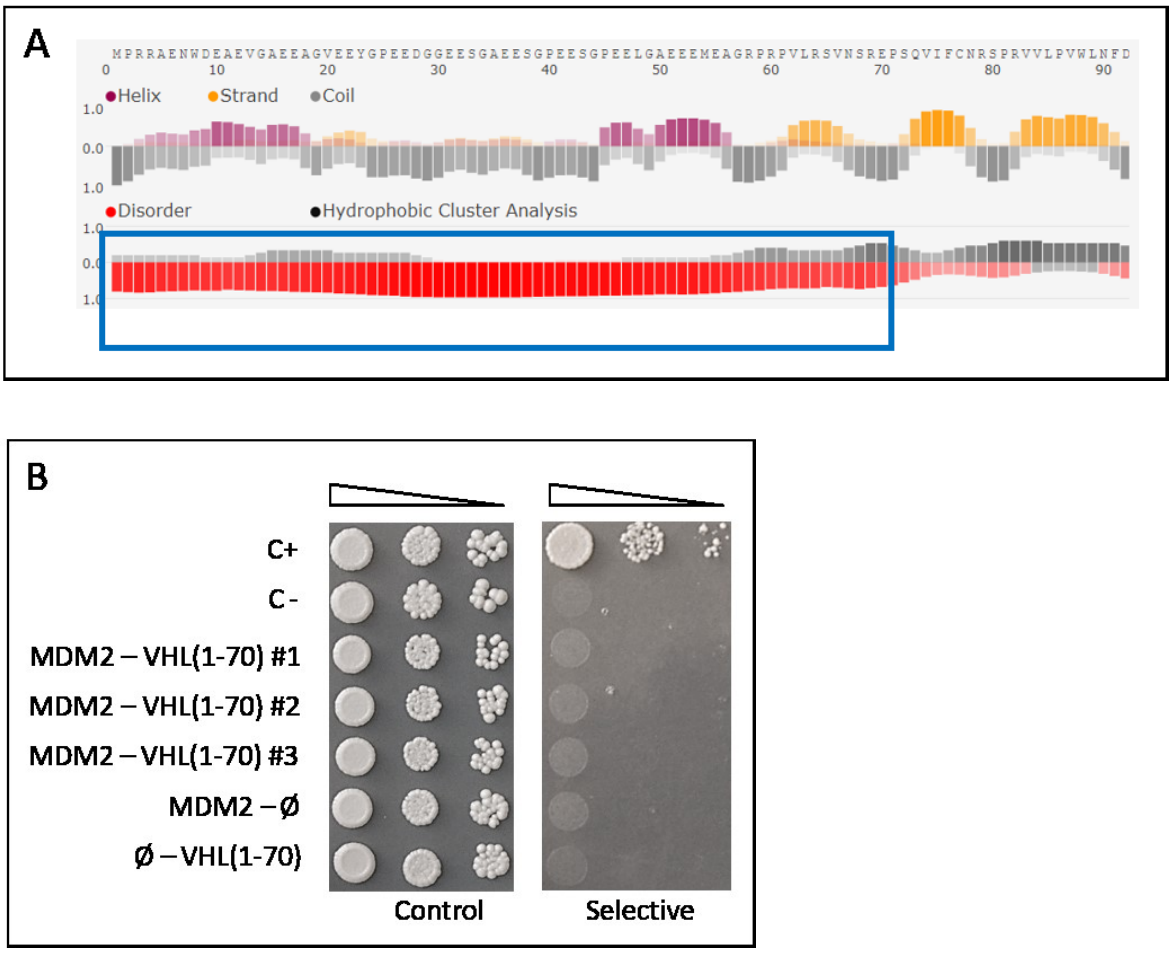


Fig. S2 Bioinformatic characterization of pVHL30 N-terminal tail and validation in Y2H

As reported in panel A the bioinformatics characterization of pVHL30 sequence with FELLS shows that the disorder part of the protein comprise residues from 1 to 70. This portion was then tested in Y2H assay as reported in panel B. Yeast co-expressing MDM2 and pVHL (1-70) were serially diluted and spotted on permissive (left) and selective (right) media and incubated for 4-6 days at 30°C. C+ and C- are positive and negative internal controls, respectively, while the pairings with empty plasmids are controls for autoactivation. Image is representative of 3 independent experiments.

Figure S3

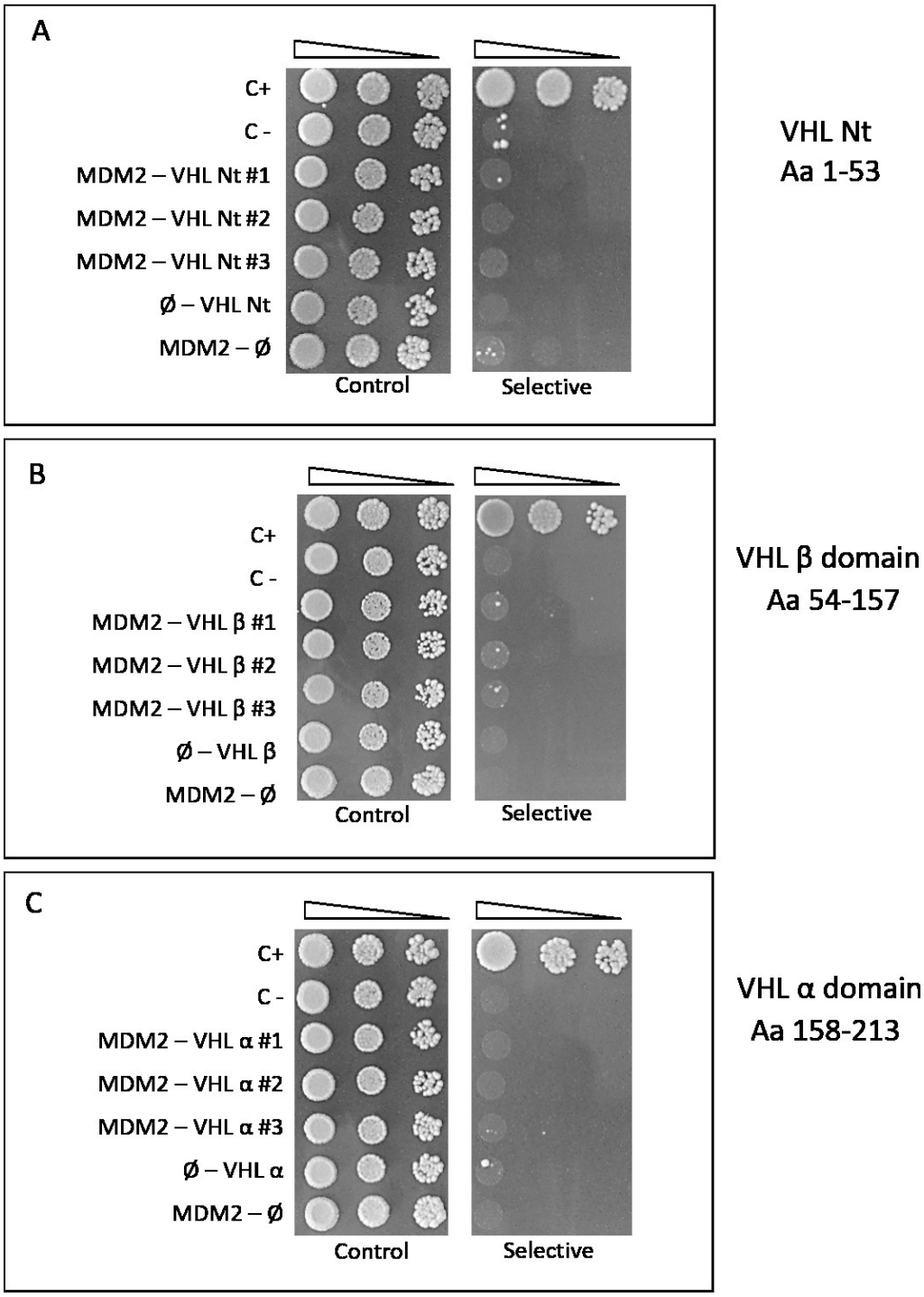


Figure S3. Molecular dissection of pVHL binding region

Yeast cells co-expressing MDM2 together with truncated mutants of pVHL (3 clones), or co-transformed with the empty vector (Ø), were serially diluted, and spotted on permissive (left) and selective (right) media, and incubated for 4-6 days at 30°C. Panel A: pVHL-NT (aa 1-53); Panel B: pVHL-beta (aa 54-157); Panel C: pVHL-α (aa 158-213). C+ and C- are positive and negative internal controls, respectively, while the pairings with empty plasmids are controls for autoactivation. Images are representative of 3 independent experiments.

Figure S4

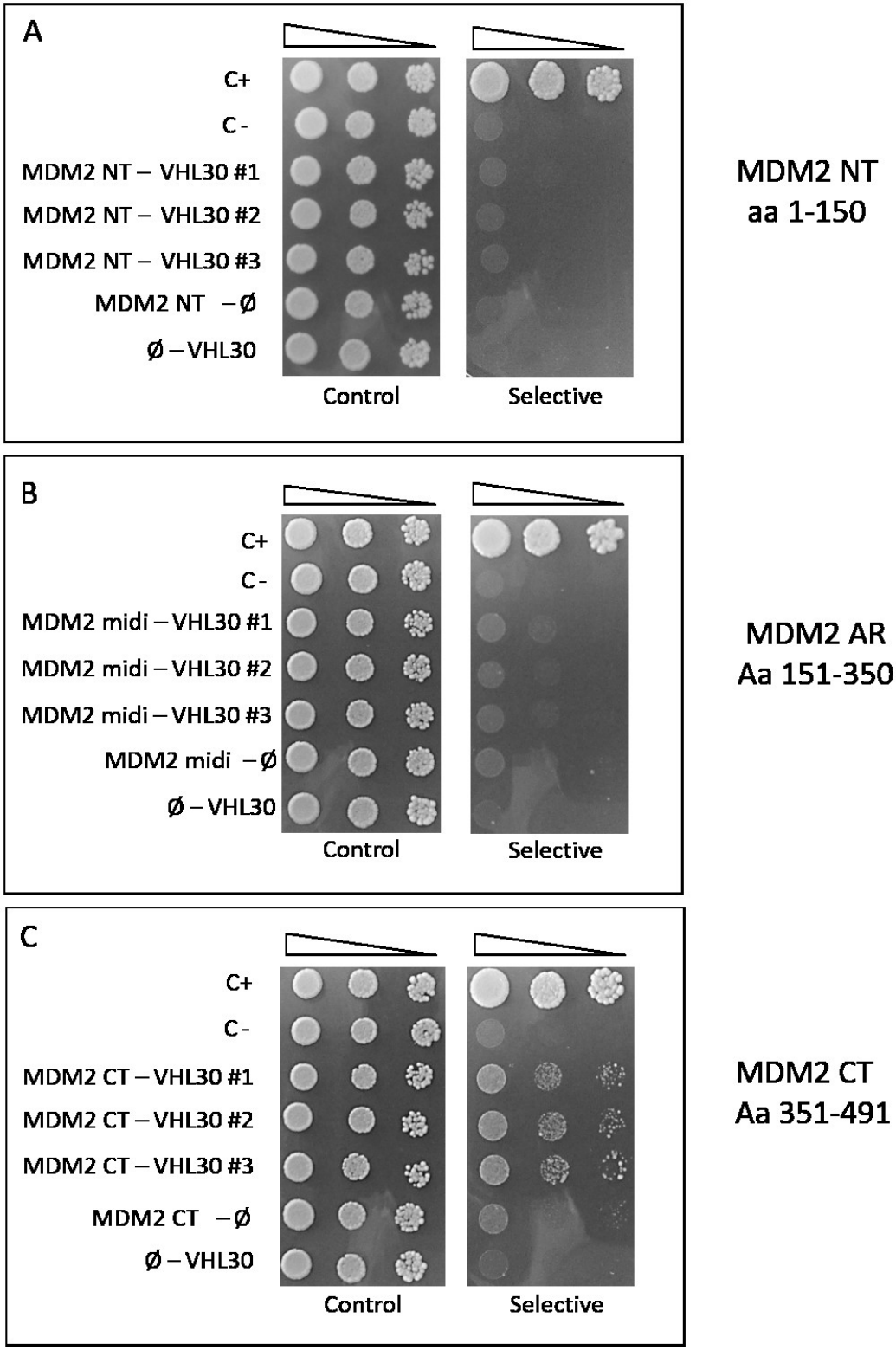


Figure S4. Molecular dissection of MDM2 binding region

Yeast cells co-expressing VHL30 together with truncated mutants of MDM2 (3 clones), or co-transformed with the empty vector (Ø), were serially diluted, and spotted on permissive (left) and selective (right) media, and incubated for 4-6 days at 30°C. Panel A: MDM2-NT (aa 1-150); Panel B: MDM2 AR (aa 151-350); Panel C: MDM2-CT (aa 351-491). C+ and C- are positive and negative internal controls, respectively, while the pairings with empty plasmids are controls for autoactivation. Images are representative of 3 independent experiments.

Figure S5

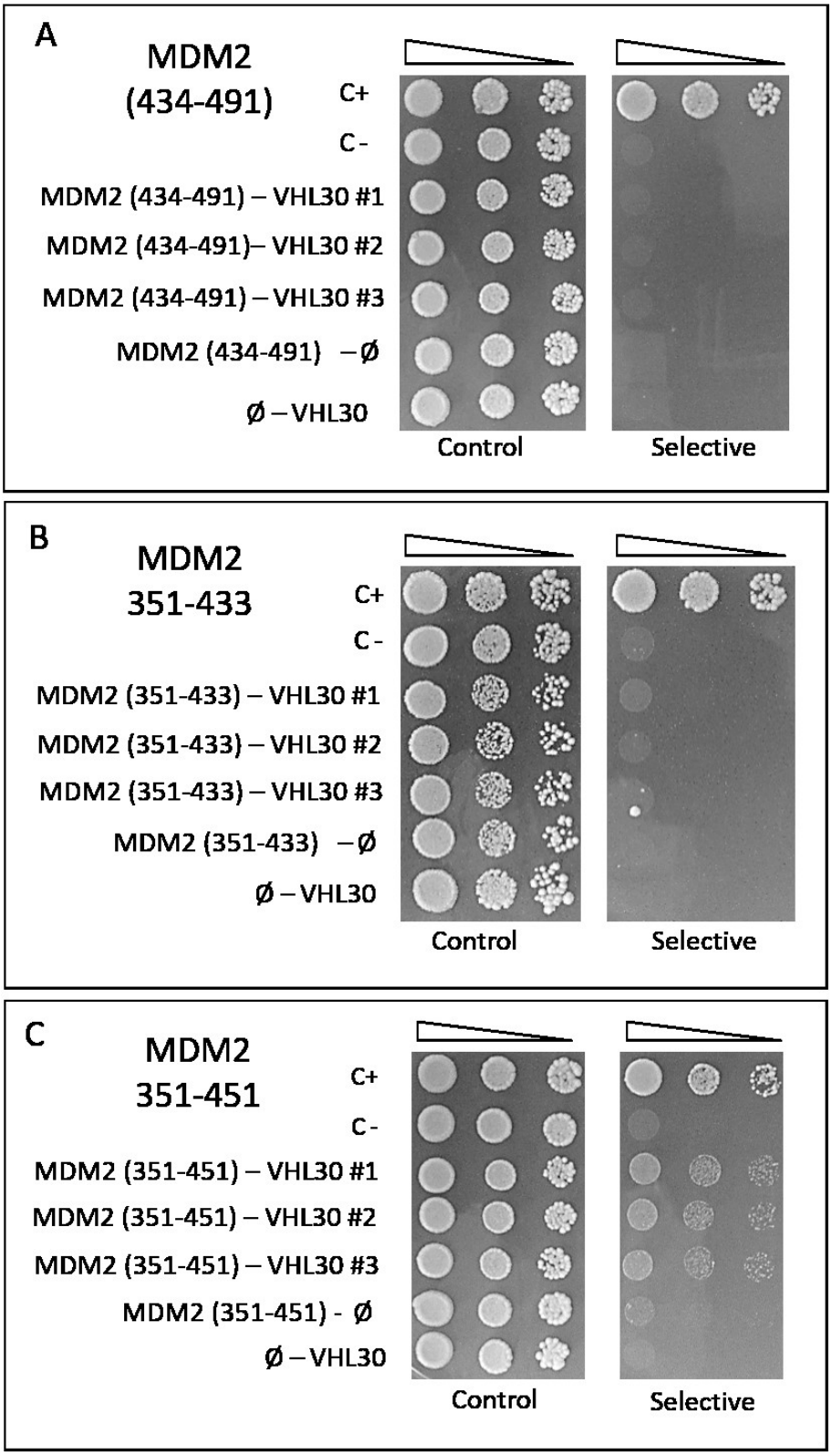


Figure S5. Molecular dissection and binding properties of regions within the MDM2-CT.

Yeast 2H experiments analyzing the interactions between the three MDM2 fragments 434-491, 351-433, 351-451 and pVHL30. Assays has been performed has previously described and the growth of three independent clones for each MDM2 fragment on the same plate is shown. Images are representative of 3 independent experiments.

Figure S6

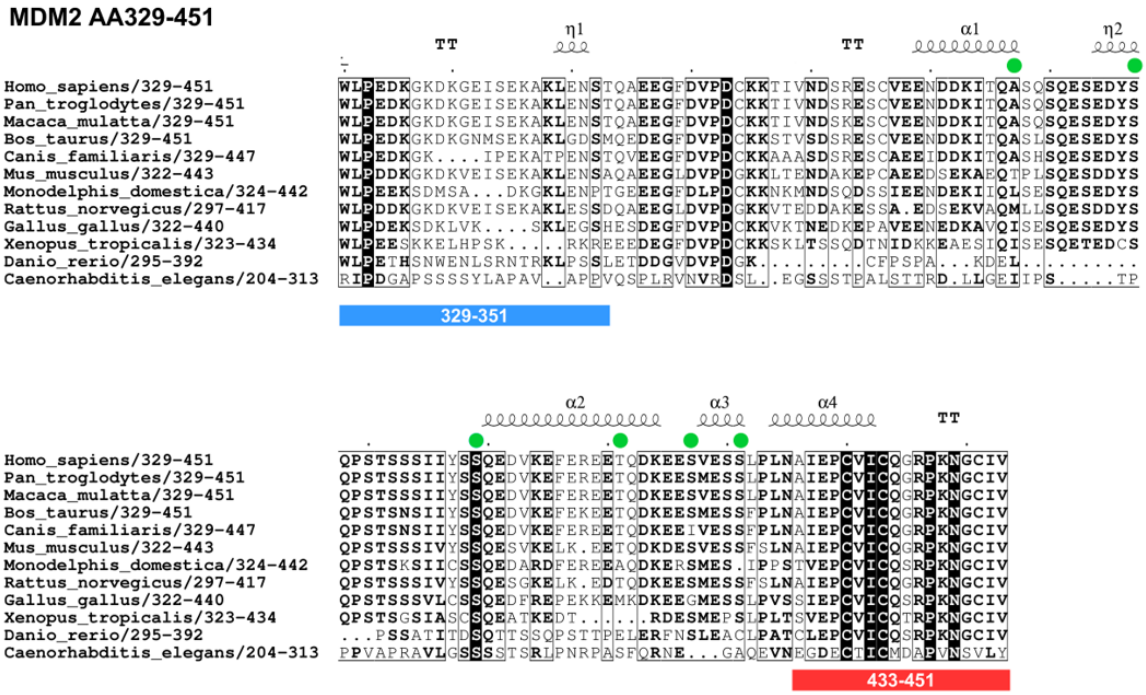


Figure S6. Multiple sequence alignment of MDM2 329-451 fragment.

Colored boxes represent the position of the two flanking regions presumed to drive MDM2/pVHL30 association. Green sphere highlight position of phosphorylation sites.

Results section 3.3

Figure S7

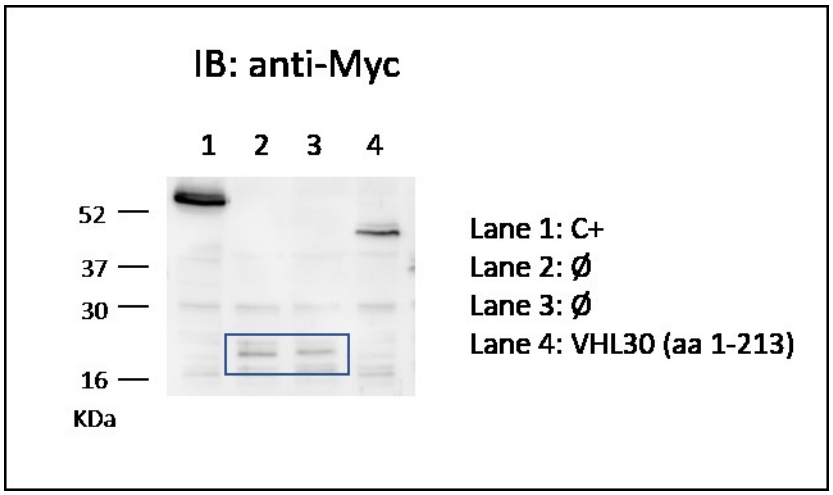
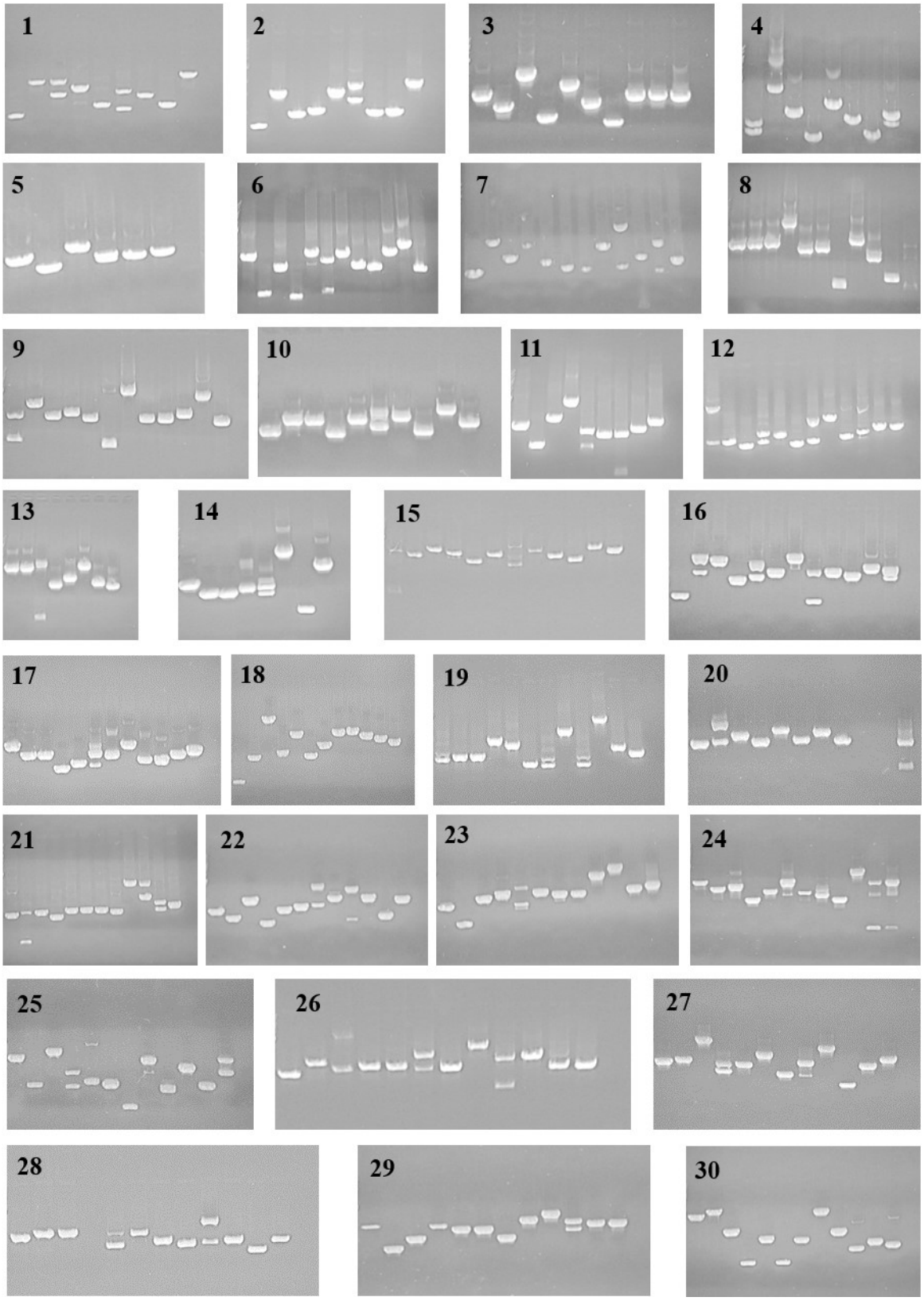


Figure S7. Western blot to confirm the expression of pVHL30 as bait in Y2H

Proteins were extracted as described in details in materials and methods section and visualized in Western Blot. The bait protein (i.e. VHL30) IS detected by anti-Myc (line 4). C+ and Ø correspond to positive (p53) and negative control respectively.

Figure S8


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

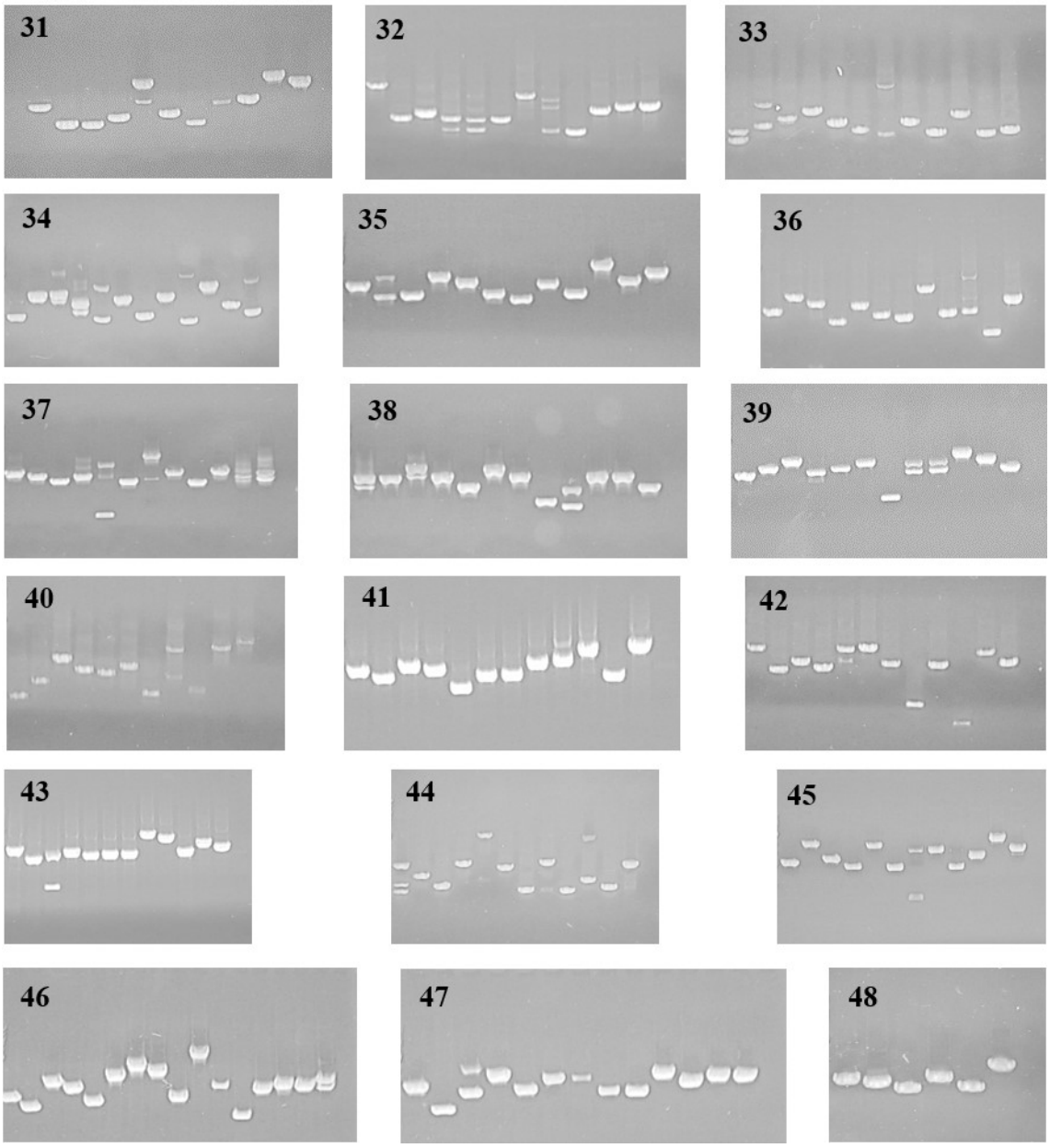
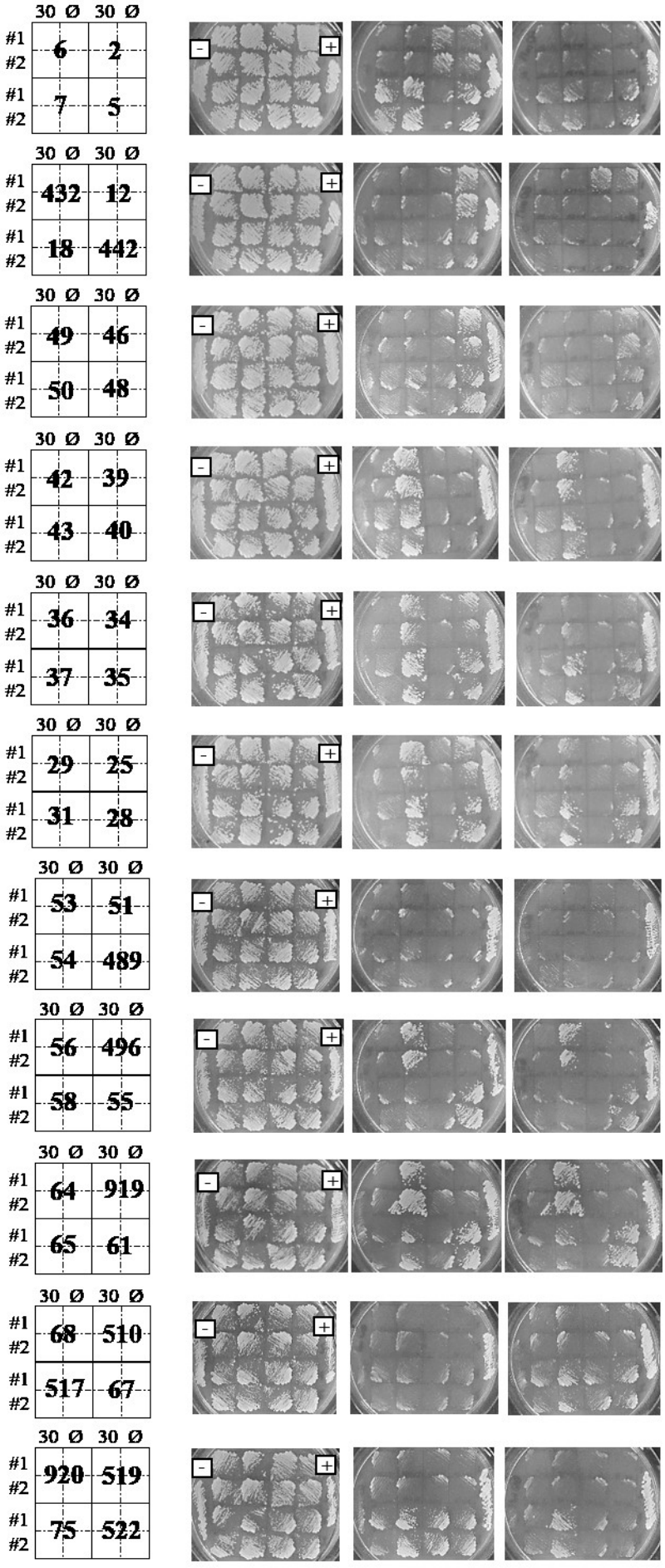


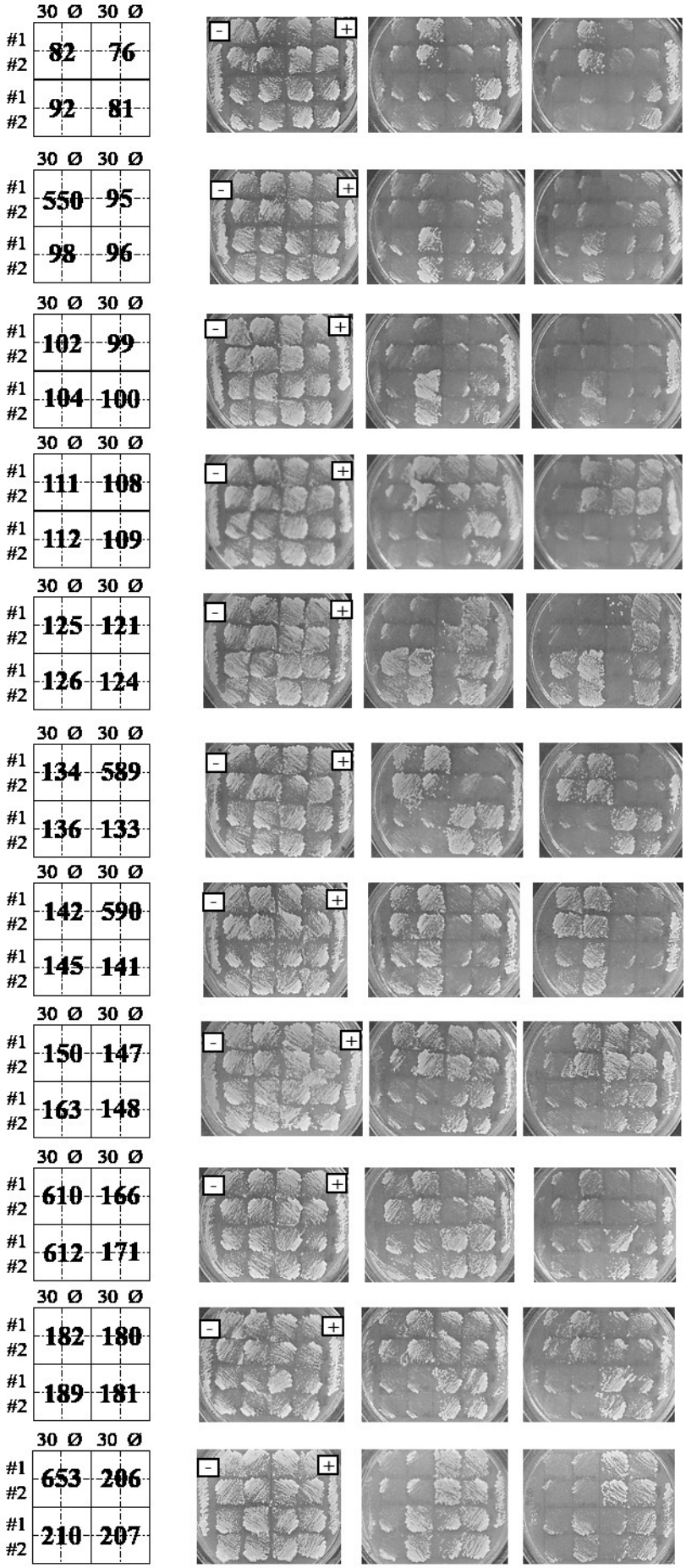
Figure S8. Library plasmids inserts amplifications.

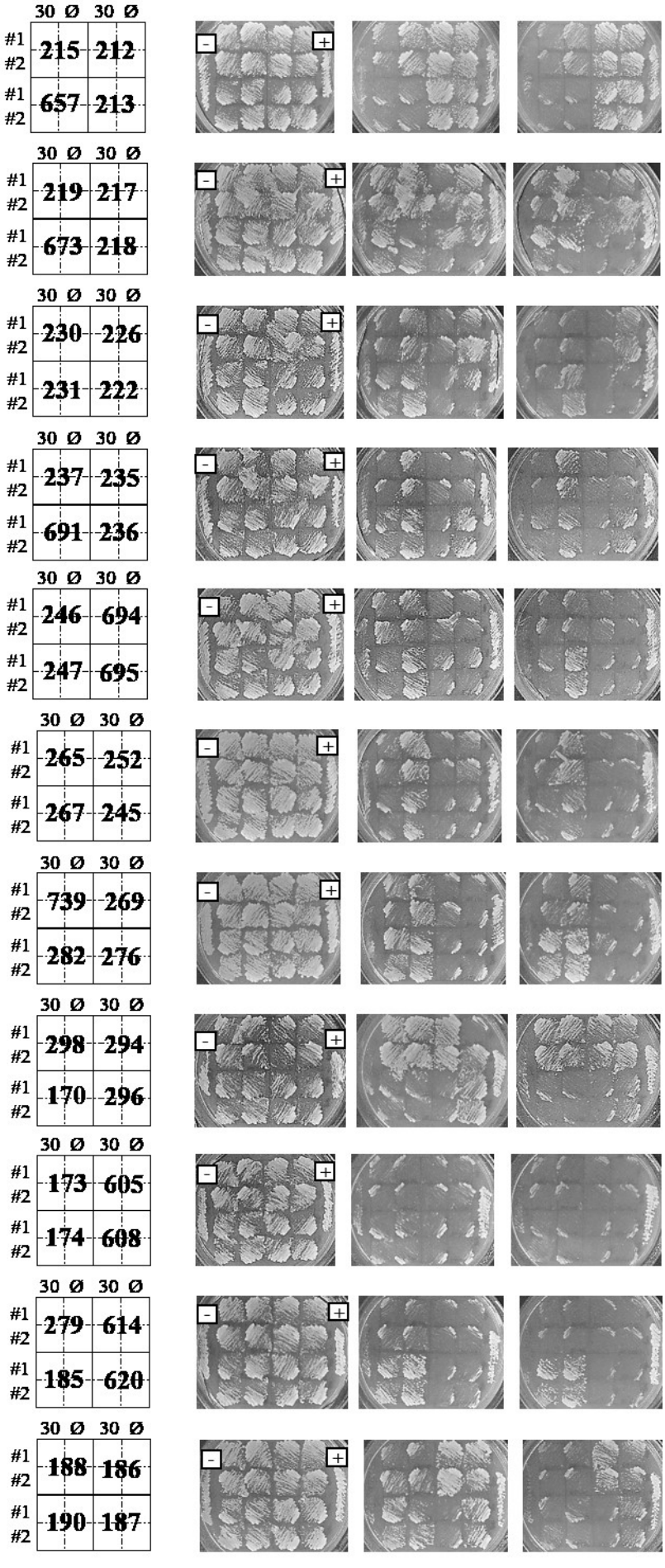
PCR products were differently separated on 0.8% agarose gel. Each line corresponds to a library clone identified by a number (n) reported in the following table. – is the PCR negative control (i.e. reaction mix without any plasmid). Red numbers correspond to clones with more than one insert.

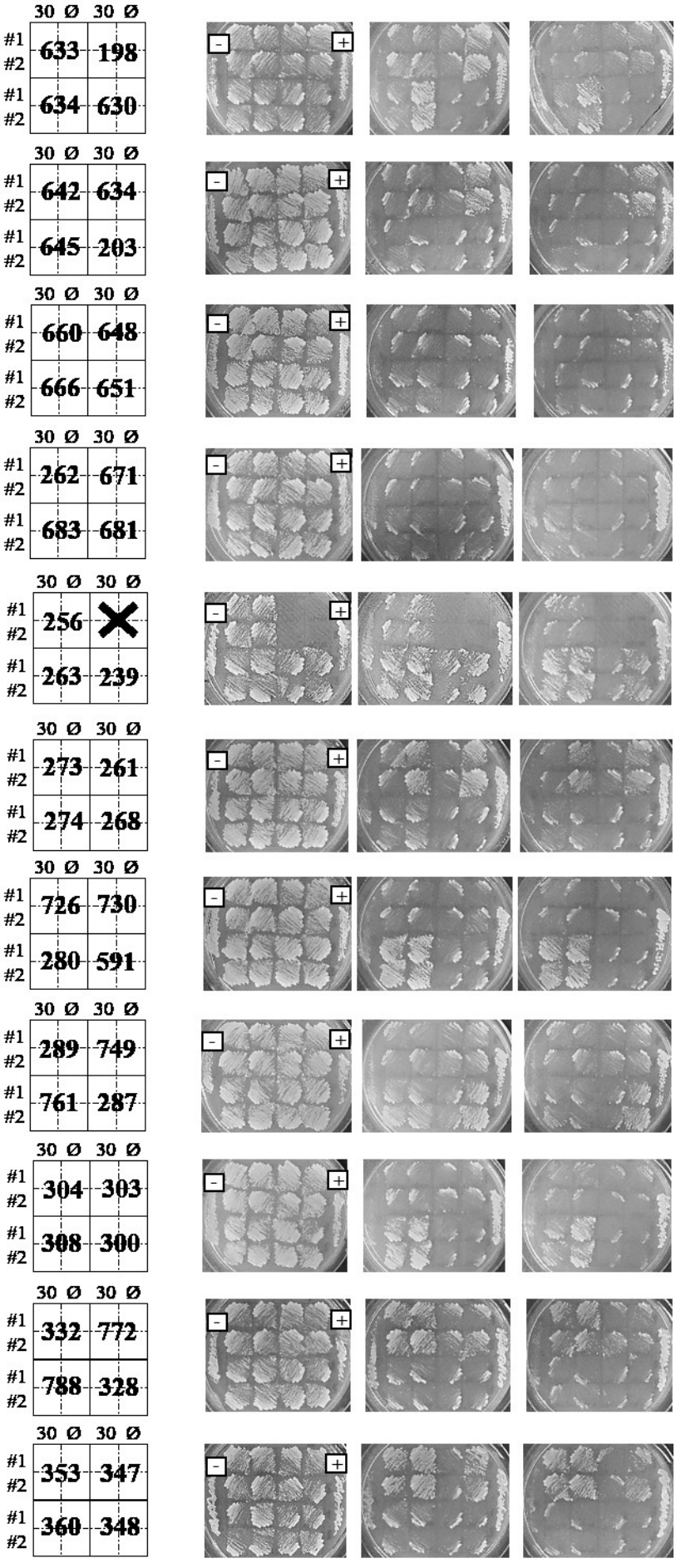
Gel number	Loading (clones number)
1	919;61;62;64;65;507;510;67;68; - (figure in the main text)
2	2;5;6;7;12;13;429;432;18; -
3	20;22;23;442;443;24;25;28;29;31; -
4	33;34;35;36;37;39;40;470; -
5	42;43;46;47;59;499; -
6	48;49;50;51;489;52;53;54;496;55;56;58; -
7	517;519;522;920;75;76;81;82;84;85;532;92; -
8	95;96;550;98;99;100;102;104;108;109; -
9	110;111;112;116;119;120;121;124;125;126;589;133; -
10	134;136;137;139;590;140;141;142;145;147; -
11	201;206; 207;653;210;212;213;215;657; -
12	253;256;263;264;257;721;260;261;268;271;273;274; -
13	148;150; 159;163;164;166;171;610; -
14	612; 180;181;182;183;189;625;165; -
15	113;572;117;574;580;127;130;131;582;583;135;138; -
16	265;266;267;269;728;276;739;281;282;286;288;292; -
17	662;216;217;218;219;220;673;226;222;230;231;233; -
18	235;236;237;691;694;695;246;247;249;251;252;245; -
19	658;660;666;669;224;671;674;681;682;262;683;239; -
20	170;603;605;608;173;174;175;176;-;x;310
21	614;178;620;298;185;186;187;188;190;628;192;194; -
22	197;198;630;633;634;203;204;642;643;645;648;651; -
23	813;822;372;373;376;830;832;833;835;836;847;848; -
24	730;938;726;280;740;744;749;284;287;289;293;759; -
25	761;303;300;306;304;308;317;318;772;328;332;333; -
26	570;571;578;581;129;588;156;151;152;600;595;602; -
27	154;158;144;592;601;155;167;596;599;911;439;444; -
28	525;526;78;x;86;87;88;89;90;533;535;540; -
29	880;888;889;897;419;900;950;523;72;73;74;524; -
30	309;773;777;319;322;323;324;327;785;311;312;767; -
31	x;97;544;545;920;294;101;103;106;561;562;564; -
32	494;506;514;69;77;529;543;551;555;560;552;565; -
33	342;795;788;347;348;353;355;360;361;362;815;367; -
34	227;646;232;240;241;699;697;701;696;706;711;703; -
35	768;770;313;315;316;775;939;330;779;784;336;338; -
36	849;852;943;865;944;871;890;876;886;884;883;881; -
37	162;923;619;624;191;626;627;924;637;199;927;928; -
38	709;707;714;270;729;731;732;736;752;754;757;295; -
39	205;229;932;208;654;655;656;934;661;664;936;672; -
40	391;392;393;398;399;400;401;410;411;412;415;891; -
41	377;831;381;383;378;837;841;844;384;387;388;854; -
42	892;418;421;423;424;426;427;1;8;9;10; -
43	790;350;351;354;810;358;359;364;365;370;823;374; -
44	368;369;340;341;793;940;346;349;799;802;807;812; -
45	449;450;457;458;475;467;465;468;463;480;481;482; -
46	906;907;16;17;430;435;910;21;436;438;26;27;448;451;454;456; -
47	459;41;461;462;464;473;466;476;484;485;915;491;493; -
48	63,503,60,918,57, -

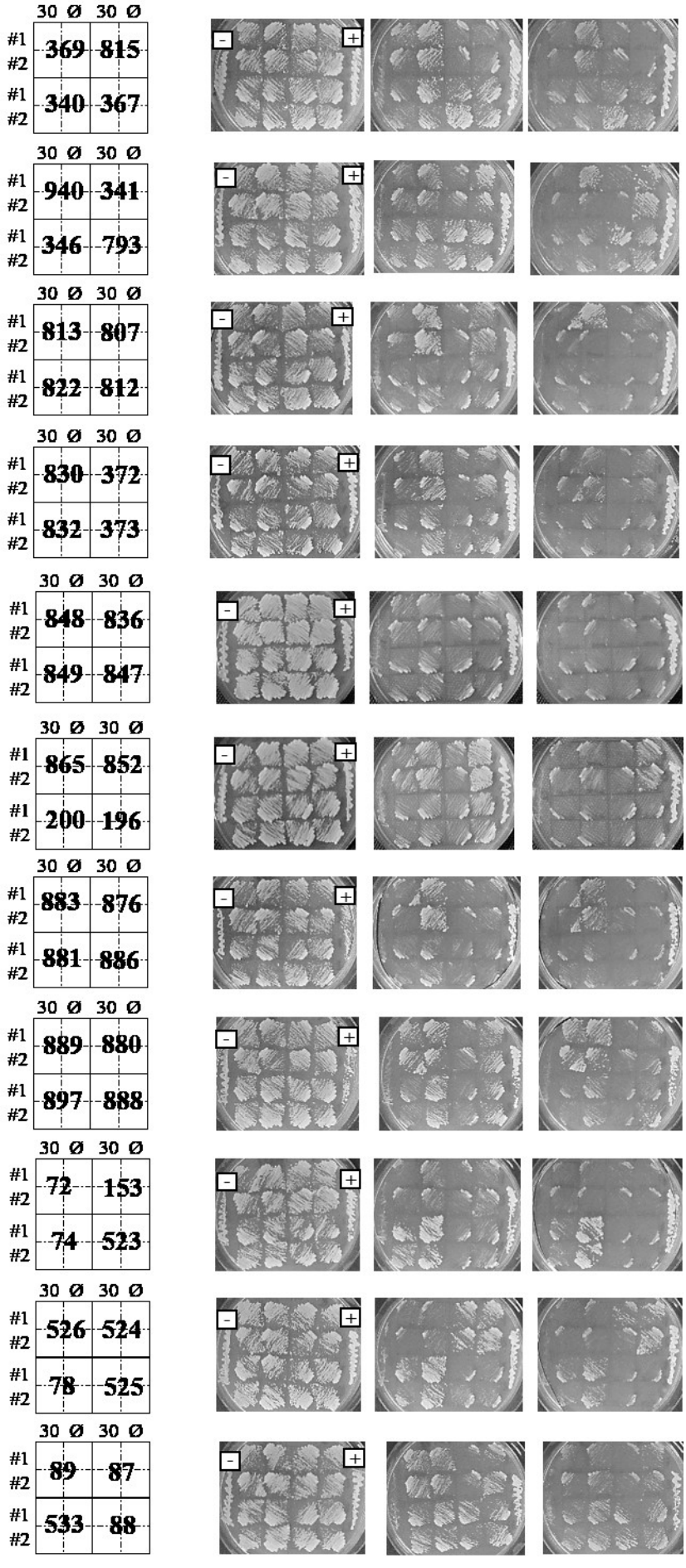
Figure S9

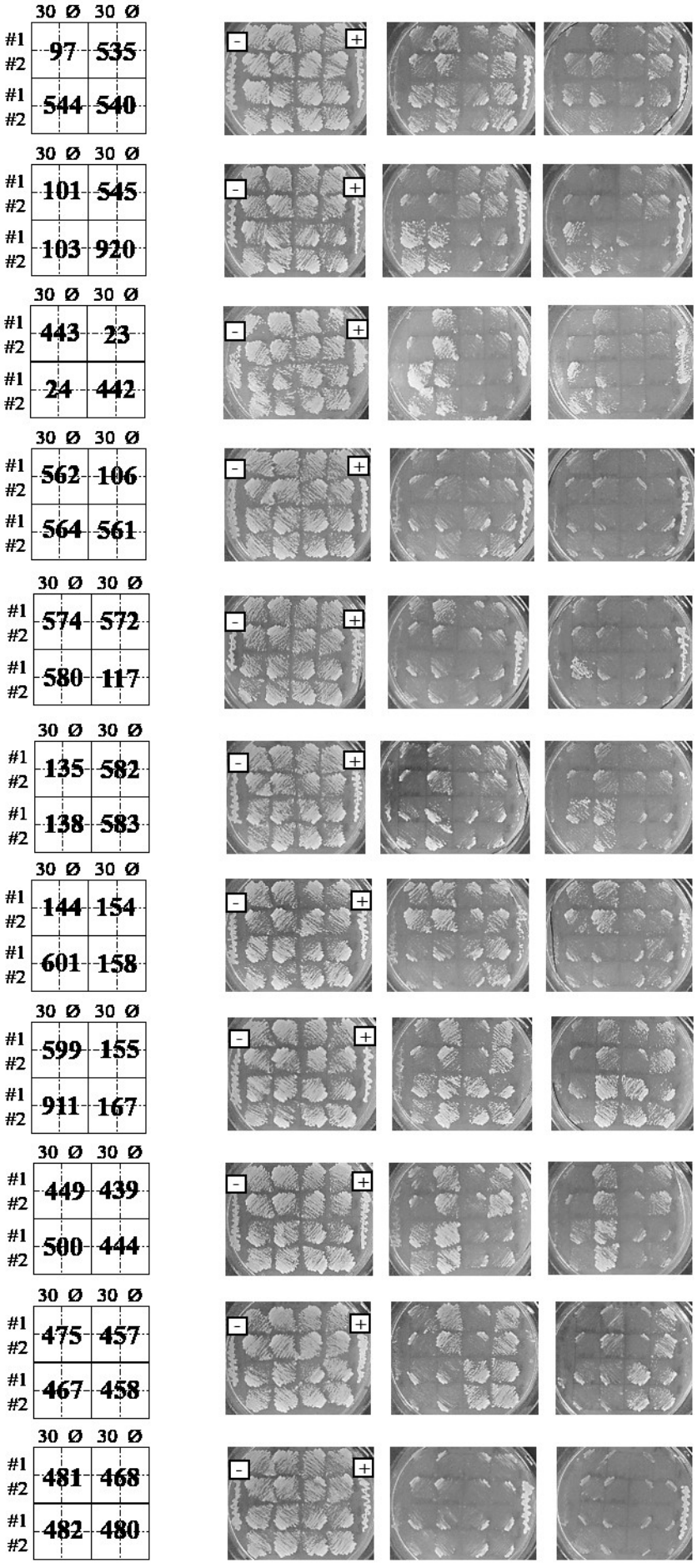

30 Ø 30 Ø
#1 #2 6 2
#1 #2 7 5
30 Ø 30 Ø
#1 #2 432 12
#1 #2 18 442
30 Ø 30 Ø
#1 #2 49 46
#1 #2 50 48
30 Ø 30 Ø
#1 #2 42 39
#1 #2 43 40
30 Ø 30 Ø
#1 #2 36 34
#1 #2 37 35
30 Ø 30 Ø
#1 #2 29 25
#1 #2 31 28
30 Ø 30 Ø
#1 #2 53 51
#1 #2 54 489
30 Ø 30 Ø
#1 #2 56 496
#1 #2 58 55
30 Ø 30 Ø
#1 #2 64 919
#1 #2 65 61
30 Ø 30 Ø
#1 #2 68 510
#1 #2 517 67
30 Ø 30 Ø
#1 #2 920 519
#1 #2 75 522

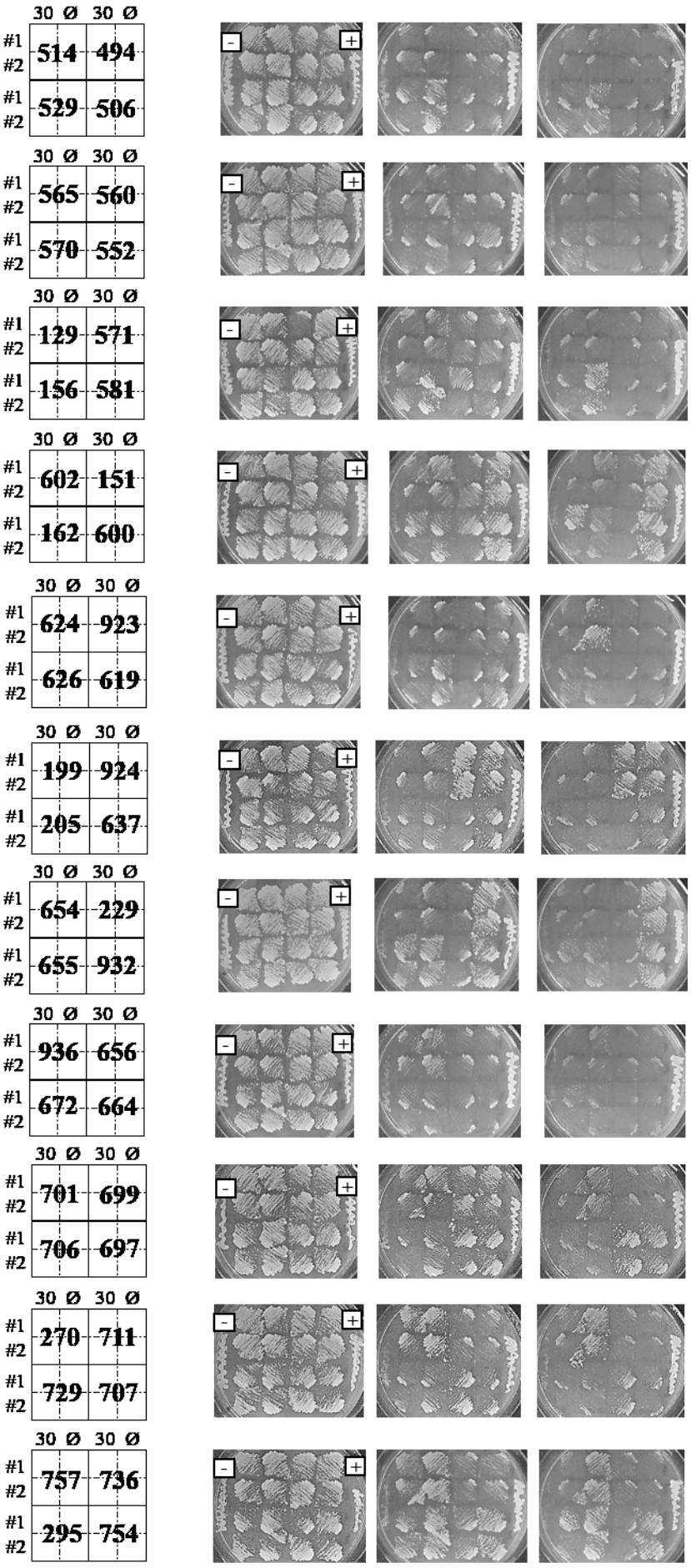

30 Ø 30 Ø
#1 82 76
#2
#1 92 81
#2
30 Ø 30 Ø
#1 550 95
#2
#1 98 96
#2
30 Ø 30 Ø
#1 102 99
#2
#1 104 100
#2
30 Ø 30 Ø
#1 111 108
#2
#1 112 109
#2
30 Ø 30 Ø
#1 125 121
#2
#1 126 124
#2
30 Ø 30 Ø
#1 134 589
#2
#1 136 133
#2
30 Ø 30 Ø
#1 142 590
#2
#1 145 141
#2
30 Ø 30 Ø
#1 150 147
#2
#1 163 148
#2
30 Ø 30 Ø
#1 610 166
#2
#1 612 171
#2
30 Ø 30 Ø
#1 182 180
#2
#1 189 181
#2
30 Ø 30 Ø
#1 653 206
#2
#1 210 207
#2


30 Ø 30 Ø
#1 215 212
#2
#1 657 213
#2
30 Ø 30 Ø
#1 219 217
#2
#1 673 218
#2
30 Ø 30 Ø
#1 230 226
#2
#1 231 222
#2
30 Ø 30 Ø
#1 237 235
#2
#1 691 236
#2
30 Ø 30 Ø
#1 246 694
#2
#1 247 695
#2
30 Ø 30 Ø
#1 265 252
#2
#1 267 245
#2
30 Ø 30 Ø
#1 739 269
#2
#1 282 276
#2
30 Ø 30 Ø
#1 298 294
#2
#1 170 296
#2
30 Ø 30 Ø
#1 173 605
#2
#1 174 608
#2
30 Ø 30 Ø
#1 279 614
#2
#1 185 620
#2
30 Ø 30 Ø
#1 188 186
#2
#1 190 187
#2

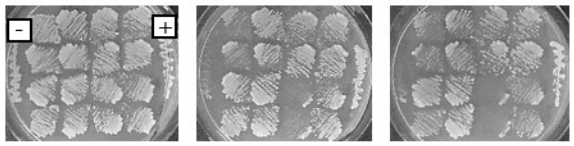

30 Ø 30 Ø
#1 633 198
#2
#1 634 630
#2
30 Ø 30 Ø
#1 642 634
#2
#1 645 203
#2
30 Ø 30 Ø
#1 660 648
#2
#1 666 651
#2
30 Ø 30 Ø
#1 262 671
#2
#1 683 681
#2
30 Ø 30 Ø
#1 256
#2
#1 263 239
#2
30 Ø 30 Ø
#1 273 261
#2
#1 274 268
#2
30 Ø 30 Ø
#1 726 730
#2
#1 280 591
#2
30 Ø 30 Ø
#1 289 749
#2
#1 761 287
#2
30 Ø 30 Ø
#1 304 303
#2
#1 308 300
#2
30 Ø 30 Ø
#1 332 772
#2
#1 788 328
#2
30 Ø 30 Ø
#1 353 347
#2
#1 360 348
#2


30 Ø 30 Ø
#1 369 815
#2
#1 340 367
#2
30 Ø 30 Ø
#1 940 341
#2
#1 346 793
#2
30 Ø 30 Ø
#1 813 807
#2
#1 822 812
#2
30 Ø 30 Ø
#1 830 372
#2
#1 832 373
#2
30 Ø 30 Ø
#1 848 836
#2
#1 849 847
#2
30 Ø 30 Ø
#1 865 852
#2
#1 200 196
#2
30 Ø 30 Ø
#1 883 876
#2
#1 881 886
#2
30 Ø 30 Ø
#1 889 880
#2
#1 897 888
#2
30 Ø 30 Ø
#1 72 153
#2
#1 74 523
#2
30 Ø 30 Ø
#1 526 524
#2
#1 78 525
#2
30 Ø 30 Ø
#1 89 87
#2
#1 533 88
#2

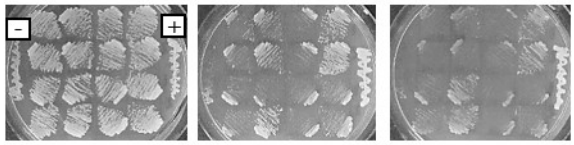

30 Ø 30 Ø
#1 97 535
#2
#1 544 540
#2
30 Ø 30 Ø
#1 101 545
#2
#1 103 920
#2
30 Ø 30 Ø
#1 443 23
#2
#1 24 442
#2
30 Ø 30 Ø
#1 562 106
#2
#1 564 561
#2
30 Ø 30 Ø
#1 574 572
#2
#1 580 117
#2
30 Ø 30 Ø
#1 135 582
#2
#1 138 583
#2
30 Ø 30 Ø
#1 144 154
#2
#1 601 158
#2
30 Ø 30 Ø
#1 599 155
#2
#1 911 167
#2
30 Ø 30 Ø
#1 449 439
#2
#1 500 444
#2
30 Ø 30 Ø
#1 475 457
#2
#1 467 458
#2
30 Ø 30 Ø
#1 481 468
#2
#1 482 480
#2


30 Ø 30 Ø
#1 514 494
#2
#1 529 506
#2
30 Ø 30 Ø
#1 565 560
#2
#1 570 552
#2
30 Ø 30 Ø
#1 129 571
#2
#1 156 581
#2
30 Ø 30 Ø
#1 602 151
#2
#1 162 600
#2
30 Ø 30 Ø
#1 624 923
#2
#1 626 619
#2
30 Ø 30 Ø
#1 199 924
#2
#1 205 637
#2
30 Ø 30 Ø
#1 654 229
#2
#1 655 932
#2
30 Ø 30 Ø
#1 936 656
#2
#1 672 664
#2
30 Ø 30 Ø
#1 701 699
#2
#1 706 697
#2
30 Ø 30 Ø
#1 270 711
#2
#1 729 707
#2
30 Ø 30 Ø
#1 757 736
#2
#1 295 754
#2

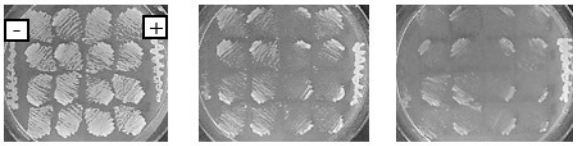
	30 Ø	30 Ø
#1 #2	324	322
#1 #2	327	326



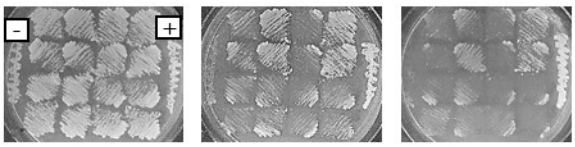
	30 Ø	30 Ø
#1 #2	768	785
#1 #2	313	312



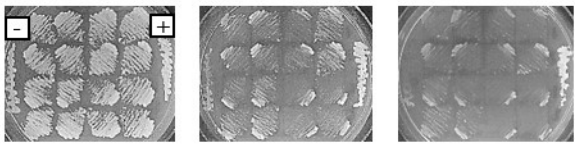
	30 Ø	30 Ø
#1 #2	775	315
#1 #2	613	316



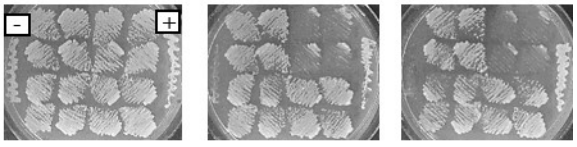
	30 Ø	30 Ø
#1 #2	784	330
#1 #2	336	779



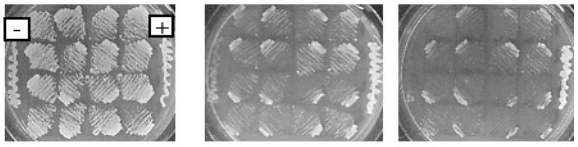
	30 Ø	30 Ø
#1 #2	350	338
#1 #2	354	790



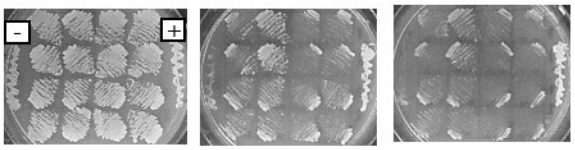
	30 Ø	30 Ø
#1 #2	359	810
#1 #2	364	358



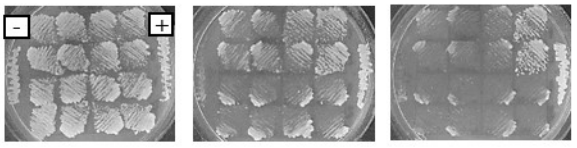
	30 Ø	30 Ø
#1 #2	823	365
#1 #2	374	370



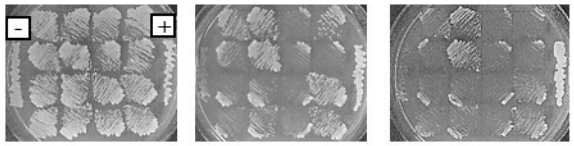
	30 Ø	30 Ø
#1 #2	381	377
#1 #2	383	831



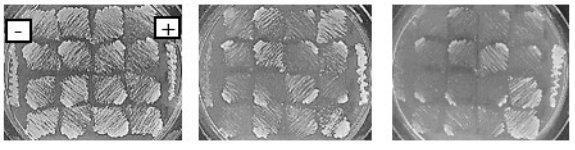
	30 Ø	30 Ø
#1 #2	841	378
#1 #2	844	837

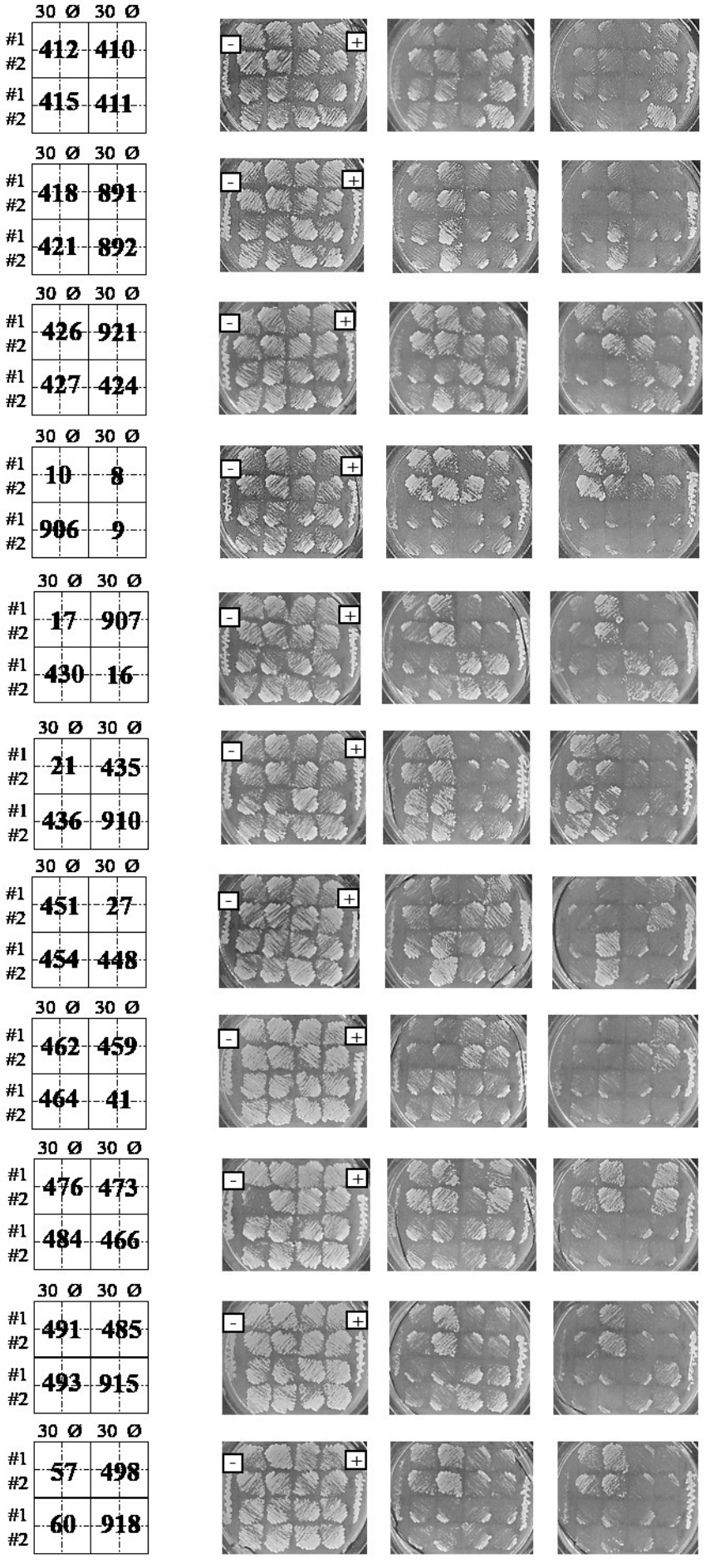


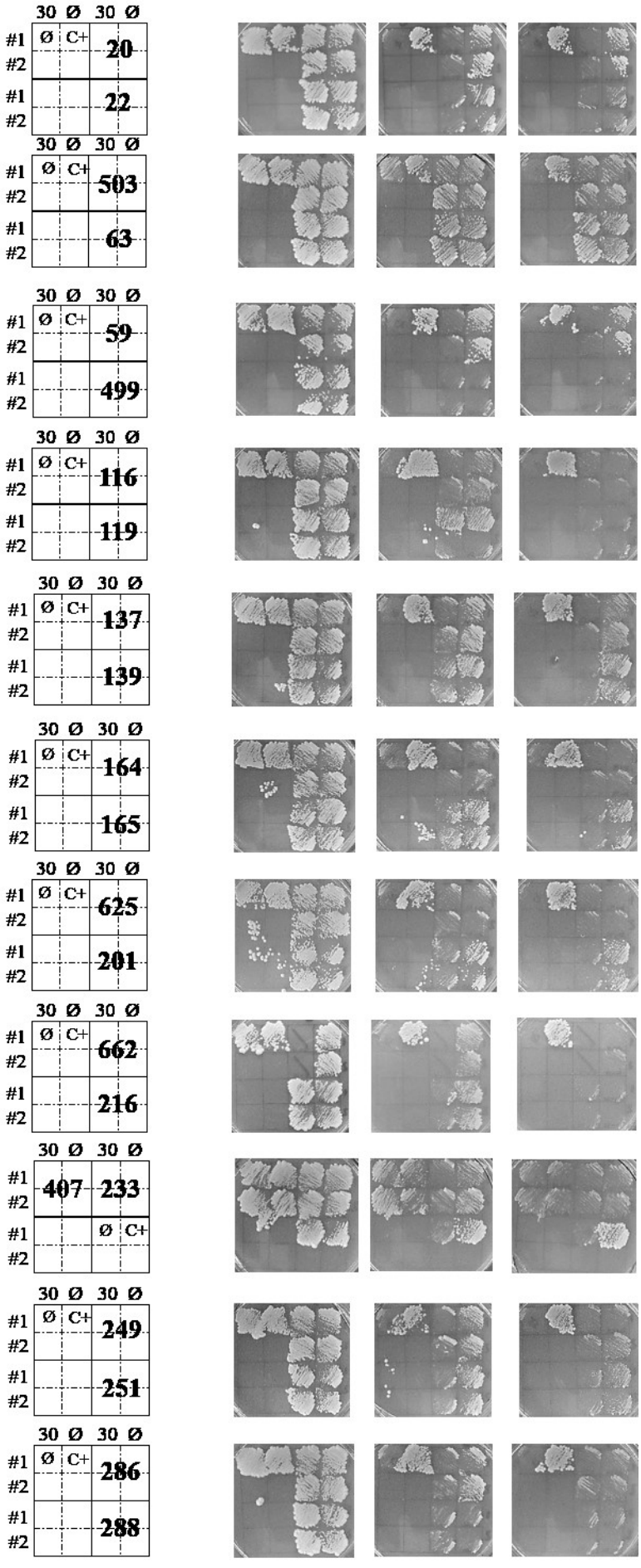
	30 Ø	30 Ø
#1 #2	388	384
#1 #2	854	387

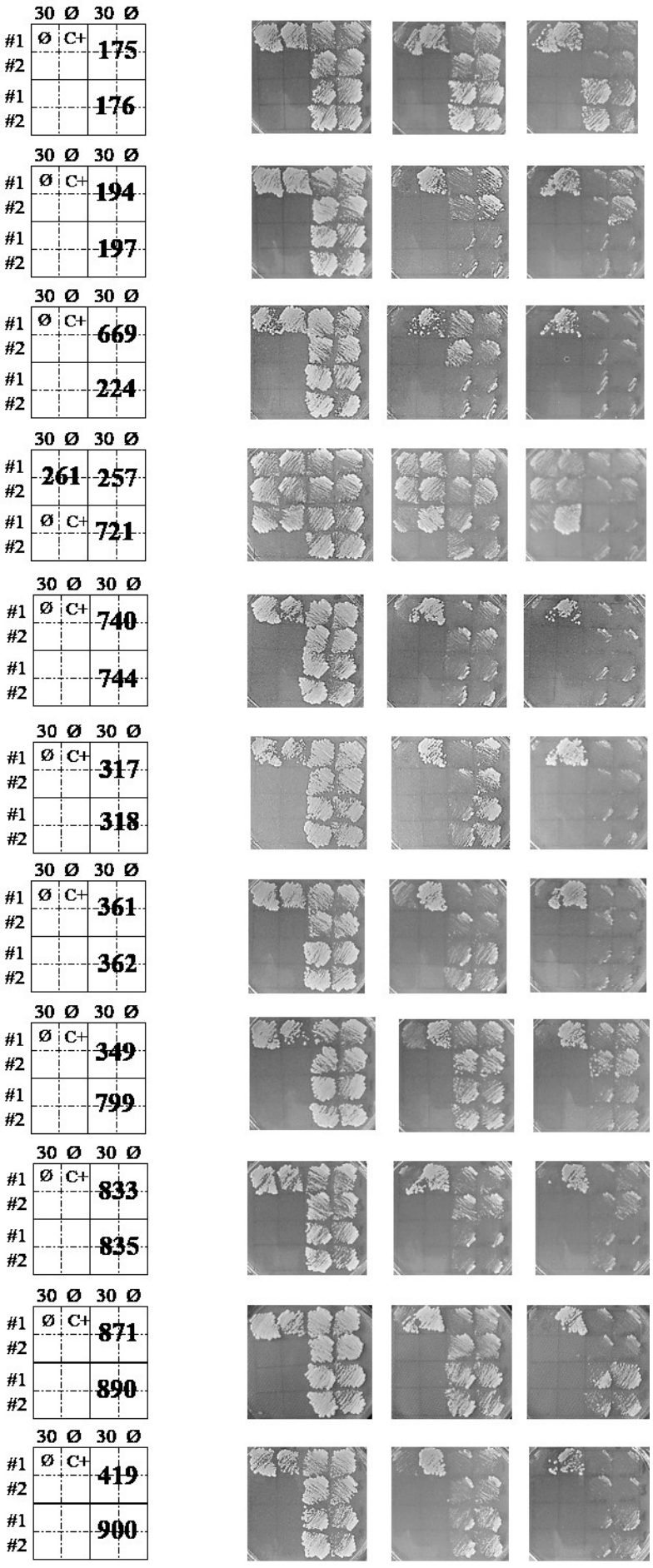


	30 Ø	30 Ø
#1 #2	400	393
#1 #2	401	398




30 Ø 30 Ø
#1 412 410
#2
#1 415 411
#2
30 Ø 30 Ø
#1 418 891
#2
#1 421 892
#2
30 Ø 30 Ø
#1 426 921
#2
#1 427 424
#2
30 Ø 30 Ø
#1 10 8
#2
#1 906 9
#2
30 Ø 30 Ø
#1 17 907
#2
#1 430 16
#2
30 Ø 30 Ø
#1 21 435
#2
#1 436 910
#2
30 Ø 30 Ø
#1 451 27
#2
#1 454 448
#2
30 Ø 30 Ø
#1 462 459
#2
#1 464 41
#2
30 Ø 30 Ø
#1 476 473
#2
#1 484 466
#2
30 Ø 30 Ø
#1 491 485
#2
#1 493 915
#2
30 Ø 30 Ø
#1 57 498
#2
#1 60 918
#2


30 Ø 30 Ø
#1 Ø C+ 20
#2
#1 22
#2
30 Ø 30 Ø
#1 Ø C+ 503
#2
#1 63
#2
30 Ø 30 Ø
#1 Ø C+ 59
#2
#1 499
#2
30 Ø 30 Ø
#1 Ø C+ 116
#2
#1 119
#2
30 Ø 30 Ø
#1 Ø C+ 137
#2
#1 139
#2
30 Ø 30 Ø
#1 Ø C+ 164
#2
#1 165
#2
30 Ø 30 Ø
#1 Ø C+ 625
#2
#1 201
#2
30 Ø 30 Ø
#1 Ø C+ 662
#2
#1 216
#2
30 Ø 30 Ø
#1 407 233
#2
#1 Ø C+
#2
30 Ø 30 Ø
#1 Ø C+ 249
#2
#1 251
#2
30 Ø 30 Ø
#1 Ø C+ 286
#2
#1 288
#2


30 Ø 30 Ø
#1 Ø C+ 175
#2
#1 176
#2
30 Ø 30 Ø
#1 Ø C+ 194
#2
#1 197
#2
30 Ø 30 Ø
#1 Ø C+ 669
#2
#1 224
#2
30 Ø 30 Ø
#1 261 257
#2
#1 Ø C+ 721
#2
30 Ø 30 Ø
#1 Ø C+ 740
#2
#1 744
#2
30 Ø 30 Ø
#1 Ø C+ 317
#2
#1 318
#2
30 Ø 30 Ø
#1 Ø C+ 361
#2
#1 362
#2
30 Ø 30 Ø
#1 Ø C+ 349
#2
#1 799
#2
30 Ø 30 Ø
#1 Ø C+ 833
#2
#1 835
#2
30 Ø 30 Ø
#1 Ø C+ 871
#2
#1 890
#2
30 Ø 30 Ø
#1 Ø C+ 419
#2
#1 900
#2

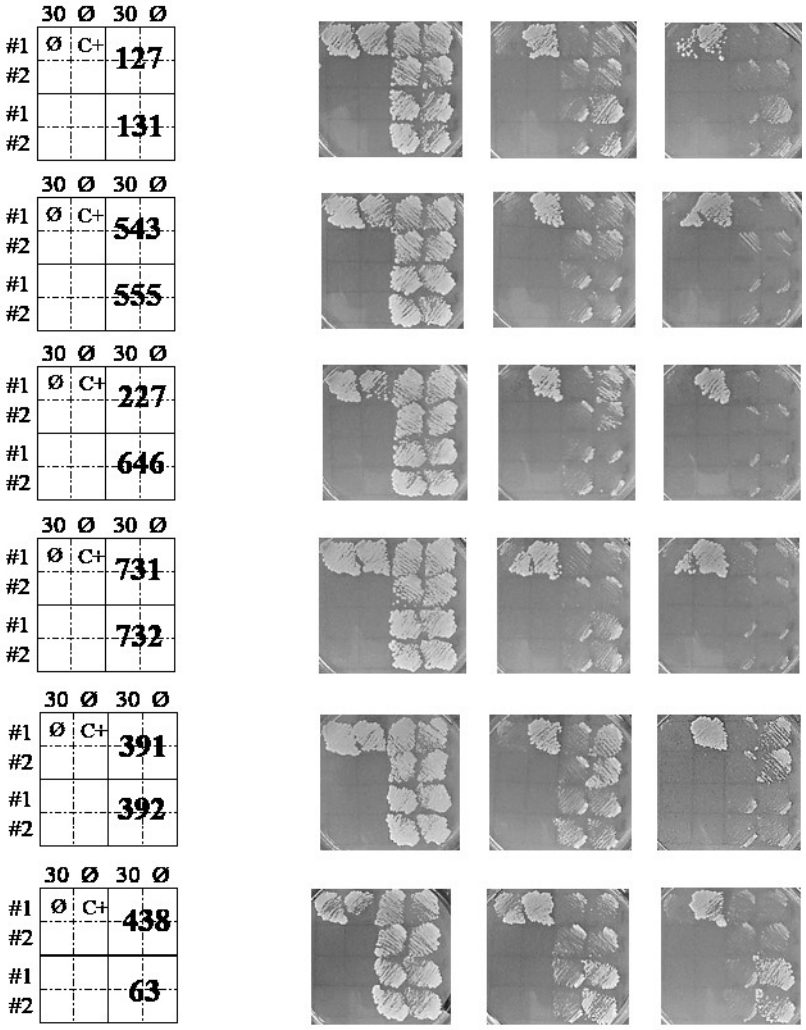


Figure S9 Yeast patches

Each library clone was co-transformed with pGBKT7 empty (Ø) or pGBKT7VHL30 (30) as indicated on the top row of the scheme. Each clone is identified by a number (n). Two independent colonies (#1,#2) of each transformation were tested on permissive (right) and selective media (left). + or C+ correspond to the Y2H positive control, whereas – or C- correspond to the negative one.

Clone n	interaction	
	yes	no
732	■	
391	■	
392		■
438		■
26	■	
46	■	
48	■	
49		■
50	■	
697		■

Clone n	interaction		Clone n	interaction		Clone n	interaction	
	yes	no		yes	no		yes	no
2			96			230		
5			550			231		
6			98			235		
7			99			236		
12			100			237		
429			102			691		
432			104			694		
18			108			695		
23			109			246		
442			111			247		
443			112			252		
24			121			245		
25			124			265		
28			125			267		
29			126			269		
31			589			276		
34			133			739		
35			134			282		
36			136			294		
37			590			296		
39			141			298		
40			142			170		
42			145			605		
43			147			608		
51			148			173		
489			150			174		
53			163			614		
54			166			620		
496			171			279		
55			610			185		
56			612			186		
58			180			187		
919			181			188		
61			182			190		
64			189			198		
65			206			630		
510			207			633		
67			653			634		
68			210			203		
517			212			642		
519			213			645		
522			215			648		
920			657			651		
75			217			660		
76			218			666		
81			219			671		
82			673			681		
92			226			262		
95			222			683		

Clone n	interaction		Clone n	interaction		Clone n	interaction	
	yes	no		yes	no		yes	no
239	■		865	■		167		■
256	■		200	■		599	■	
263		■	876		■	911	■	
261			886		■	439	■	
268		■	883	■		444		■
273	■		881	■		449	■	
274	■		880	■		500	■	
730		■	888	■		457	■	
591		■	889		■	458		■
726		■	897	■		475	■	
280		■	153	■		467		■
749	■		523	■		468		■
287	■		72	■		480		■
289		■	74	■		481		■
761	■		524	■		482		■
303		■	525		■	494	■	
300		■	526		■	506		■
304		■	78	■		514		■
308	■		87	■		529	■	
772	■		88		■	560		■
328		■	89	■		552	■	
332	■		533	■		565	■	
788		■	535	■		570		■
347	■		540	■		571		■
348	■		97	■		581		■
353	■		544		■	129	■	
360		■	545	■		156	■	
815	■		920		■	151	■	
367		■	101		■	600	■	
369	■		103		■	602	■	
340	■		503		■	162		■
341	■		63		■	923		■
793		■	106		■	619	■	
940	■		561	■		624		■
346	■		562		■	626	■	
807	■		564	■		924		■
812	■		572	■		637	■	
813	■		117		■	199	■	
822	■		574	■		205	■	
372	■		580		■	229	■	
373	■		582	■		932	■	
830	■		583		■	654		■
832	■		135	■		655	■	
836		■	138	■		656		■
847		■	154	■		664		■
848	■		158	■		936	■	
849	■		144	■		672		■
852	■		601		■	699	■	
196		■	155	■		701	■	

Clone n	interaction		Clone n	interaction		Clone n	interaction	
	yes	no		yes	no		yes	no
706	■		393	■		22	■	
711		■	398	■		59	■	
707		■	400	■		499		■
270	■		401		■	116	■	
729	■		410	■		119	■	
736	■		411	■		137	■	
754	■		412		■	139	■	
757	■		415		■	164		■
295	■		891		■	165	■	
322		■	892	■		625		■
326	■		418	■		201	■	
324	■		421	■		662		■
327		■	921		■	216	■	
785	■		424	■		233	■	
312	■		426	■		407		
768	■		427	■		249	■	
313	■		8		■	251	■	
315	■		9		■	286	■	
316		■	10		■	288		■
775		■	906		■	175	■	
613		■	907		■	176		■
330	■		16	■		194	■	
779		■	17	■		197		■
784	■		430			669		■
336		■	435			224		■
338	■		910	■		257	■	
790		■	21			721		■
350		■	436			261	■	
354	■		27	■		740	■	
810		■	448		■	744		■
358		■	451		■	317	■	
359		■	454	■		318	■	
364		■	459	■		361		■
365	■		41		■	362	■	
370		■	462		■	349		■
823	■		464	■		799	■	
374		■	473	■		833	■	
377	■		466		■	835		■
831		■	476	■		871		■
381	■		484		■	890		■
383	■		485		■	419	■	
378	■		915	■		900	■	
837		■	491	■		127	■	
841		■	493		■	131	■	
844		■	498		■	543	■	
384	■		918		■	555		■
387	■		57	■		227	■	
388	■		60		■	646		■
854	■		20	■		731		■

Figure S10

clone	protein name	gene	UniProt ID	isoforms	sequences
ID07HJ15 cl 535	Jouberin	**AHI1**	Q8N157 AHI1_HUMAN	canonical sequence	**LVKQRLETVTEVIRSCAAKVNKNLSFTSPPAVSSQQSKLKQSNMLTAQEILHQFGFTQTGIISIERKPCN HQVDTAPTVVALYDYTANRSDELTIHRGDIIRVFFKDNEDWWYGSIGKGQEGYFPANHVASETLYQELP PEIKERSPPLSPEEKTKIEK**
ID07HJ94 cl 338	Ankyrin repeat and EF-hand domain-containing protein 1	**ANKEF1**	Q9NU02 ANKE1_HUMAN	canonical sequence	**ALRLHDWSVEREAFLREAFAVLDRGDGSISKNDFVMVLEERQDYASSEQLAAIAHLHEKTRGGGVNIN EFFKGTRYLNKSFVLGSYGPKKKEKGMGKKGKKGKFVLPLPICVIPEYAFPRRQDGGPPYYMIETYKN VTDSSRFNRDHPPEHPIQDDSVWYIDDSEKVFSNINIITKAGDLASLKKAFESGIPVDMKDNYYKTPLM TACASGNIDVVKFLLEKGANVNATDNFLWTPLHFACHAGQQDIVELLVESGALIDAASINNSTPLNRAIE SCRLDTVKYLLDIGAKFQLENRKG**
ID07HJ93 cl 784	Ankyrin repeat and EF-hand domain-containing protein 1	**ANKEF1**	Q9NU02 ANKE1_HUMAN	canonical sequence	**WALRLHDWSVEREAFLREAFAVLDRGDGSISKNDFVMVLEERQDYASSEQLAAIAHLHEKTRGGGVN INEFFKGTRYLNKSFVLGSYGPKKKEKGMGKKGKKGKFVLPLPICVIPEYAFPRRQDGGPPYYMIETY KNVTDSSRFNRDHPPEHPIQDDSVWYIDDSEKVFSNINIITKAGDLASLKKAFESGIPVDMKDNYYKTP LMTACASGNIDVVKFLLEKGANVNATDNFLWTPLHFACHAGQQDIVELLVESGALIDAASINNSTPLNR AIESCRLDTVKYLLDIGAKFQLENRKG**
ID_66HJ10 _7H _cl121	Ankyrin repeat domain-containing protein 11	**ANKRD11**	Q6UB99 ANR11_HUMAN	canonical sequence	**TGKKDKDKVSLTKTPKLERGDGGKEVRERASKRKLPFTAGANGEQKDSDTEKQGPERKRIKKEPVT RKAGLLFGMGLSGIRAGYPLSERQQVALLMQMTAEESANSPVDTTPKHPSQSTVCQKGTPNSASKT KDKVNKRNERGETRLHRAAIRGDARRIKELISEGADVNVKDFAGWTALHEACNRGYYDVAKQLLAAG AEVNTKGLDDDTPLHDAANNGHYKV**
ID66HJ41 Clone 237	Ankyrin repeat domain-containing protein 11	**ANKRD11**	Q6UB99 ANR11_HUMAN	canonical sequence	**TGKKDKDKVSLTKTPKLERGDGGKEVRERASKRKLPFTAGANGEQKDSDTEKQGPERKRIKKEPVT RKAGLLFGMGLSGIRAGYPLSERQQVALLMQMTAEESANSPVDTTPKHPSQSTVCQKGTPNSASKT KDKVNKRNERGETRLHRAAIRGDARRIKELISEGADVNVKDFAGWTALHEACNRGYYDVAKQLLAAG AEVNTKGLDDDTPLHDAANNGHYKV**
ID 07IA36 cl 491	Rho guanine nucleotide exchange factor 7	**ARHGEF7**	Q14155 ARHG7_HUMAN	Isoform 3	**PGTIEKVYPEPRSESECLSNIREFLRGCGASLRLETFDANDLYQGQNFNKVLSSLVTLNKVTADIGLGS DSVCARPSSHRIKSFDSLGSQS**
ID_66HI64_2A_cl5	Protein BEX2	**BEX2**	Q9BXY8-2 BEX2_HUMAN	isoform 2	**VCGAKCCGDAPHVENREEETARIGPGVMESKEERA**
ID66HJ48 Clone 252	Protein BEX4	**BEX4**	Q9NWD9 BEX4_HUMAN	canonical sequence	**MESKEELAANNLNGENAQQENEGGEQAPTQNEEESRHLGGGEGQKPGGNIRRGRVRRLVPNFRWA IPNRHIEHNEARDDVERFVGQMMEIKRKTREQQMRHYMRFQTPEPDNHYDFCLIP**
ID07HJ26 cl135	Protein BEX4	**BEX4**	Q9NWD9 BEX4_HUMAN	canonical sequence	**MESKEELAANNLNGENAQQENEGGEQAPTQNEEESRHLGGGEGQKPGGNIRRGRVRRLVPNFRWA IPNRHIEHNEARDDVERFVGQMMEIKRKTREQQMRHYMRFQTPEPDNHYDFCLIP**
ID_66IA15 cl812	Breast cancer type 1 susceptibility protein	**BRCA1**	P38398 BRCA1_HUMAN	canonical + isoform 5,3,8,4,7	**VVNGRNHQGPKRARESQDRKIFRGLEICCYGPFTNMPTDQLEWMVQLCGASVVKELSSFTLGTGVH PIVVVQPDAWTEDNGFHAIGQMCEAPVVTREWVLDSVALYQCQELDTYLIPQIPHSHY**
ID07HJ25 cl 582	T-complex protein 1 subunit epsilon	**CCT5**	P48643 TCPE_HUMAN	canonical sequence	**MASMGTLAFDEYGRPFLIIKDQDRKSRLMGLEALKSHIMAAKAVANTMRTSLGPNGLDKMMVDKDG DVTVTNDGATILSMMDVDHQIAKLMVELSKSQDDEIGDGTTGVVVLAGALLEEAEQLLDRGIHPIRIAD GYEQAARVAIEHLDKISDSVLVDIKDTEPLIQTAKTTLGSKVVNSCHRQMAEIAVNAVLTVADMERRDVD FELIKVEGKVGGRLEDTKLIKGVIV**
ID_66HI91_6E_cl64	T-complex protein 1 subunit eta	**CCT7**	Q99832 TCPH_HUMAN	canonical sequence + isoform 3, 4, 5, 8	**KADKVEQRKLLEKCAMTALSSKLISQQKAFFAKMVVDAVMMLDDLLQLKMIGIKKVQGGALEDSQLV AGVAFKKTFSYAGFEMQPKKYHNPKIALLNVELELKAEKDNAEIRVHTVEDYQAIVDAEWNILYDKLEK IHHSGAKVVLSKLPIGDVATQYFADRDMFCAGRVPEEDLKRTMMACGGSIQTSVNALSADVLGRCQV FEETQIGGERYNFFTGCPKAK**
ID_66IA10 cl341	T-complex protein 1 subunit eta	**CCT7**	Q99832 TCPH_HUMAN	canonical + isoform 3, 4	**ADKVEQRKLLEKCAMTALSSKLISQQKAFFAKMVVDAVMMLDDLLQLKMIGIKKVQGGALEDSQLVA GVAFKKTFSYAGFEMQPKKYHNPKIALLNVELELKAEKDNAEIRVHTVEDYQAIVDAEWNILYDKLEKI HHSGAKVVLSKLPIGDVATQYFADRDMFCAGRVPEEDLKRTMMACGGSIQTSVNALSADVLGRCQVF EETQIGGERYNFFTGCPKAKTCTFILRGGAEQFMEETERSLHDAIMIVRRAIKNDSVVAGGGAIEMELS KYLRDYSRTIPGKQQLLIGAYAKALEIIPRQLCDNAGFDATNILNKLRARHAQGGTW**
ID_66IA19 cl373	T-complex protein 1 subunit eta	**CCT7**	Q99832 TCPH_HUMAN	canonical + isoform 3, 4	**GALEDSQLVAGVAFKKTFSYAGFEMQPKKYHNPKIALLNVELELKAEKDNAEIRVHTVEDYQAIVDAE WNILYDKLEKIHHSGAKVVLSKLPIGDVATQYFADRDMFCAGRVPEEDLKRTMMACGGSIQTSVNALS ADVLGRCQVFEETQIGGERYNFFTGCPKAKTCTFILRGGAEQFMEETERSLHDAIMIVRRAIKNDSVVA GGGAIEMELSKYLRDYSRTIPGKQQLLIGAYAKALEIIPRQLCDNAGFDATNILNKLRARHAQGGTWYG VDINNEDIADNFEAFVWEPAMVRINALTAASEAACLIVSVDETIKNPRSTVDAPTAAGRG**
ID_66IA25 cl852	T-complex protein 1 subunit eta	**CCT7**	Q99832 TCPH_HUMAN	canonical + isoform 3, 4	**KADKVEQRKLLEKCAMTALSSKLISQQKAFFAKMVVDAVMMLDDLLQLKMIGIKKVQGGALEDSQLV AGVAFKKTFSYAGFEMQPKKYHNPKIALLNVELELKAEKDNAEIRVHTVEDYQAIVDAEWNILYDKLEK IHHSGAKVVLSKLPIGDVATQYFADRDMFCAGRVPEEDLKRTMMACGGSIQTSVNALSADVLGRCQV FEETQIGGERYNFFTGCPKAKTCTFILRGGAEQFMEETERSLHDAIMIVRRAIKNDSVVAGGGAIEMEL SKYLRDYSRTIPGKQQLLIGAYAKALEIIPRQLCDNAGFDATNILNKLRARHAQGG**

ID07IA05 cl 854	CDK5 regulatory subunit-associated protein 3	**CDK5RAP3**	J3QRX0 J3QRX0_HUMAN	canonical sequence	**WGQELGPSARAHRHPDQQAARLAGGQKALQPEMAESGADDPREDQCCHPGHARERRDRPAAVW VLHSLLSLPKNPGPSQRHRGLHEEYFWPILFTADEGLAGDYSSV**

ID_66IA05 cl353	COMM domain-containing protein 1	**COMMD1**	Q8N668 COMD1_HUMAN	canonical + isoform 2	**MAAGELEGGKPLSGLLNALAQDTFHGYPGITEELLRSQLYPEVPPEEFRPFLAKMRGILKSIASADMD FNQLEAFLTAQTKKQGGITSDQAAVISKFWKSHKTKIRESLMNQSRWNSGLRGLSWRVDGKSQSRH SAQIHTPVAIIELELGKYGQESEFLCLEFDEVKVNQILKTLSEVEESISTLISQPN**
ID_66HJ09 _5H _cl119	Copine-5	**CPNE5**	A0A0J9YWA1 A0A0J9YWA1_HUMAN	canonical sequence	**ITVSCRNLLDKDMFSKSDPLCVMYTQGMENKQWREFGRTEVIDNTLNPDFVRKFIVDYFFEEKQNLRF DLYDVDSKSPDLSKHDFLGQAFCTLGEIVGSPGSRLEKPLTIGAFSLNSRTGKPMPAVSNGSGLWMES LRTTGLEASGGVPGKKCGTIILSAEELSNCRDVATMQFCANKLDKKDFFGKSDPFLVFYRSNEDGTFTI CHKTEVMKNTLNPVWQTFSIPVRA**
ID66IA48 cl 881	Homo sapiens cathepsin D (CTSD), RefSeqGene on chromosome 11	**CTSD**	P07339 CATD_HUMAN	canonical sequence	**EAIVDTGTSLMVGPVDEVRELQKAIGAVPLIQGEYMIPCEKVSTLPAITLKLGGKGYKLSPEDYTLKVS QAGKTLCLSGFMGMDIPPPSGPLWILGDVFIGRYYTVFDRDNNRVGFAEAARL**
ID_66HI99 _8F _cl82	Death-inducer obliterator 1	**DIDO1**	Q9BTC0 DIDO1_HUMAN	canonical sequence + isoform 1, 2 3	**DQGIKGRIEKAANPSGKKKLKIFQPVIEAPGASKCIGPGCCHVAQPDSVYCSNDCILKHAAATMKFLSS GKEQKPKPKEKMKMKPEKPSLPKCGAQAGIKISSVHKRPAPEKKETTVKKAVVVPARSEALGKEAAC ESSTPSWASDHNYNAVKPEKTAAPSPSLLYK**
ID_66HJ01 _3G _cl95	Death-inducer obliterator 1	**DIDO1**	Q9BTC0 DIDO1_HUMAN	canonical sequence + isoform 1, 2 3	**DQGIKGRIEKAANPSGKKKLKIFQPVIEAPGASKCIGPGCCHVAQPDSVYCSNDCILKHAAATMKFLSS GKEQKPKPKEKMKMKPEKPSLPKCGAQAGIKISSVHKRPAPEKKETTVKKAVVVPARSEALGKEAAC ESSTPSWASDHNYNAVKPEKTAAPSPSLLYK**
ID66HJ28 Clone 201	Death-inducer obliterator 1	**DIDO1**	Q9BTC0 DIDO1_HUMAN	canonical sequence + isoform 1, 2 3	**QGIKGRIEKAANPSGKKKLKIFQPVIEAPGASKCIGPGCCHVAQPDSVYCSNDCILKHAAATMKFLSSG KEQKPKPKEKMKMKPEKPSLPKCGAQAGIKISSVHKRPAPEKKETTVKKAVVVPARSEALGKEAACE SSTPSWASDHNYNAVKPEKTAAPSPSLLYK**
ID07HJ24 cl 131	Destrin	**DSTN**	P60981 DEST_HUMAN	canonical + isoform 2	**VEEGKEILVGDVGVTITDPFKHFVGMLPEKDCRYALYDASFETKESRKEELMFFLWAPELAPLKSKMIY ASSKDAIKKKFQGIKHECQANGPEDLNRACIAEKLGGSLIVAFEGCPV**
ID_66HI96 _5F _cl75	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**DCILPPTRPTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPVNVTTEVKSVEMHHEALSE ALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIA CKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPP**
ID66HJ34 Clone 219	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**HHEALSEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVL DCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRFAVRD MRQTVAVGVIKAVDKKAAGAGKVTKSAQKAQKAK**
ID66HJ54 Clone 175	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**GNASGTTLLEALDCILPPTRPTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPVNVTTEV KSVEMHHEALSEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAG YAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRF AVRDMRQTVAVGVIKAVDKKAAGAGKVTKSAQKAQKAK**
ID66HJ55 Clone 185	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**HHEALSEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVL DCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRFAVRD MRQTVAVGVIKAVDKKAAGAGKVTKSAQKAQKAK**
ID_66HJ88 cl239	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**PLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPVNVTTEVKSVEMHHEALSEALPGDNVGFNVKNV SVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSG KKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRFAVRDMRQTVAVGVIKAVDKKAAGAGKV TKSAQKAQKAK**
ID_66HJ98 cl308	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**EALSEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDC HTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRFAVRDMR QTVAVGVIKAVDKKAAGAGKVTKSAQKAQKAK**
ID_66IA02 cl332	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**EALSEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDC HTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRFAVRDMR QTVAVGVIKAVDKKAAGAGKVTKSAQKAQKAK**
ID_66IA18 cl372	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**GWKVTRKDGNASGTTLLEALDCILPPTRPTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTF APVNVTTEVKSVEMHHEALSEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIIL NHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESF SDYPPLGRFAVRDMRQTVAVGVIKAVDKKAAGAGKVTKSAQKAQKAK**
ID66IA53 cl 900	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**KVTRKDGNASGTTLLEALDCILPPTRPTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPV NVTTEVKSVEMHHEALSEALPGDNVGFNVKNVSVKDVRRGN**
ID07HJ22 cl 574	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**KRYEEIVKEVSTYIKKIGYNPDTVAFVPISGWNGDNMLEPSANMPWFKGWKVTRKDGNASGTTLLEAL DCILPPTRPTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPVNVTTEVKSVEMHHEALSE ALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIA CKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRFAVRDMRQTVA**

ID07HJ27 cl 138	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**DGNASGTTLLEALDCILPPTRPTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPVNVTTE VKSVEMHHEALSEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISA GYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGR FAVRDMRQTVAVGVIKAVDKKAAGAGKVTKSAQKAQKAK**

ID07HJ28 cl 154	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**NGDNMLEPSANMPWFKGWKVTRKDGNASGTTLLEALDCILPPTRPTDKPLRLPLQDVYKIGGIGTVP VGRVETGVLKPGMVVTFAPVNVTTEVKSVEMHHEALSEALPGDNVGFNVKNVSVKDVRRGNVAGDS KNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGD AAIVDMVPGKPMCVESFSDYPPLGRFAVRDMRQTVA**
ID 07IA34 cl 476	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**EALSEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDC HTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRFAVRDMR QTVAVGVIKAVDKKAAGAGKVTKSAQK**
ID07IA14 cl 424	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**ALDCILPPTRPTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPVNVTTEVKSVEMHHEAL SEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTA HIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRFAVRDMRQTVA VGVIKAVDKKAAGAGKVTKSAQKAQKAK**
ID07HJ86 cl 324	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**PLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPVNVTTEVKSVEMHHEALSEALPGDNVGFNVKNV SVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSG KKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRFAVRDMRQTVAVGVIKAVDKKAAGAGKV TKSAQKAQKAK**
ID07HJ82 cl 295	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**EVKSVEMHHEALSEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAGFTAQVIILNHPGQIS AGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGDAAIVDMVPGKPMCVESFSDYPPLG RFAVRDMRQTVAVGVIKAVDKKAAGAGKVTKSAQKAQKAK**
ID07HJ56 cl 655	Elongation factor 1-alpha 1	**EEF1A1**	P68104 EF1A1_HUMAN	canonical + isoform 2	**NGDNMLEPSANMPWFKGWKVTRKDGNASGTTLLEALDCILPPTRPTDKPLRLPLQDVYKIGGIGTVP VGRVETGVLKPGMVVTFAPVNVTTEVKSVEMHHEALSEALPGDNVGFNVKNVSVKDVRRGNVAGDS KNDPPMEAAGFTAQVIILNHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLEDGPKFLKSGD AAIVDMVPGKPMCVESFSDYPPLGRFAVRDMRQTVAVGVIKAVDKKAAGAGKVTKSAQKAQKAK**
ID_66HI86_4C_cl39	Elongin-C	**ELOC**	Q15369 ELOC_HUMAN	canonical isoform	**MDGEEKTYGGCEGPDAMYVKLISSDGHEFIVKREHALTSGTIKAMLSGPGQFAENETNEVNFREIPSH VLSKVCMYFTYKVRYTNSSTEIPEFPIAPEIALELLMAANFLDC**
ID66HJ19 Clone 590	Elongin-C	**ELOC**	Q15369 ELOC_HUMAN	canonical isoform + isoform 2	**MDGEEKTYGGCEGPDAMYVKLISSDGHEFIVKREHALTSGTIKAMLSGPGQFAENETNEVNFREIPSH VLSKVCMYFTYKVRYTNSSTEIPEFPIAPEIALELLMAANFLDC**
ID66HJ31 Clone 216	Elongin-C	**ELOC**	Q15369 ELOC_HUMAN	canonical isoform + isoform 2	**MDGEEKTYGGCEGPDAMYVKLISSDGHEFIVKREHALTSGTIKAMLSGPGQFAENETNEVNFREIPSH VLSKVCMYFTYKVRYTNSSTEIPEFPIAPEIALELLMAANFLDC**
ID66HJ32 Clone 217	Elongin-C	**ELOC**	Q15369 ELOC_HUMAN	canonical isoform + isoform 2	**MDGEEKTYGGCEGPDAMYVKLISSDGHEFIVKREHALTSGTIKAMLSGPGQFAENETNEVNFREIPSH VLSKVCMYFTYKVRYTNSSTEIPEFPIAPEIALELLMAANFLDC**
ID 07IA30 cl 454	Elongin-C	**ELOC**	Q15369 ELOC_HUMAN	canonical isoform + isoform 2	**MDGEEKTYGGCEGPDAMYVKLISSDGHEFIVKREHALTSGTIKAMLSGPGQFAENETNEVNFREIPSH VLSKVCMYFTYKVRYTNSSTEIPEFPIAPEIALELLMAANFLDC**
ID07HJ57 cl 936	Elongin-C	**ELOC**	Q15369 ELOC_HUMAN	canonical isoform + isoform 2	**MDGEEKTYGGCEGPDAMYVKLISSDGHEFIVKREHALTSGTIKAMLSGPGQFAENETNEVNFREIPSH VLSKVCMYFTYKVRYTNSSTEIPEFPIAPEIALELLMAANFLDC**
ID07HJ36 cl 450	3'-5' exoribonuclease 1	**ERI1**	Q8IV48 ERI1_HUMAN	canonical sequence	**PEETQQCKFDGQETKGSKFITSSASDFSDPVYKEIAITNGCINRMSKEELRAKLSEFKLETRGVKDVLK KRLKNYYKKQKLMLKESNFADSYYDYICIIDFEATCEEGNPPEFVHEIIEFPVVLLNTHTLEIEDTFQQYV RPEINTQLSDFCISLTGITQDQVDRADTFPQVLKK**
ID_66HJ85 cl633	Electron transfer flavoprotein subunit alpha, mitochondrial	**ETFA**	P13804 ETFA_HUMAN	canonical + isoform 2	**DQKLTKSDRPELTGAKVVVSGGRGLKSGENFKLLYDLADQLHAAVGASRAAVDAGFVPNDMQVGQT GKIVAPELYIAVGISGAIQHLAGMKDSKTIVAINKDPEAPIFQVADYGIVADLFKVVPEMTEILKKK**
ID_66IA13 cl799	Electron transfer flavoprotein subunit alpha, mitochondrial	**ETFA**	P13804 ETFA_HUMAN	canonical + isoform 2	**KLTKSDRPELTGAKVVVSGGRGLKSGENFKLLYDLADQLHAAVGASRAAVDAGFVPNDMQVGQTGK IVAPELYIAVGISGAIQHLAGMKDSKTIVAINKDPEAPIFQVADYGIVADLFKVVPEMTEILKKK**
ID07HJ13 cl 89	Electron transfer flavoprotein subunit alpha, mitochondrial	**ETFA**	P13804 ETFA_HUMAN	canonical + isoform 2	**LDQKLTKSDRPELTGAKVVVSGGRGLKSGENFKLLYDLADQLHAAVGASRAAVDAGFVPNDMQVGQ TGKIVAPELYIAVGISGAIQHLAGMKDSKTIVAINKDPEAPIFQVADYGIVADLFKVVPEMTEILKKK**
ID07HJ88 cl 312	F-box only protein 28	**FBXO28**	Q9NVF7 FBX28_HUMAN	canonical + isoform 2	**MAAAAEERMAEEGGGGQGDGGSSLASGSTQRQPPPPAPQHPQPGSQALPAPALAPDQLPQNNTLV ALPIVAIENILSFMSYDEISQLRLVCKRMDLVCQRMLNQGFLKVERYHNLCQKQVKAQLPR**
ID_66IA14 cl807	F-box only protein 34	**FBXO34**	Q9NWN3 FBX34_HUMAN	canonical sequence	**MHLKPYWKLQKKEHPPEVSRETQRTPMNHQKAVNDETCKASHITSSVFPSASL**
ID07IA04 cl 388	Flotillin-1	**FLOT1**	O75955 FLOT1_HUMAN	canonical sequence	**IYKDRQKFSEQVFKVASSDLVNMGISVVSYTLKDIHDDQDYLHSLGKARTAQVQKDARIGEAEAKRDA GIREAKAKQEKVSAQYLSEIEMAKAQRDYELKKAAYDIEVNTRRAQADLAYQLQVAKTKQQI**

ID07HJ74 cl 701	G patch domain and ankyrin repeat-containing protein 1	**GPANK1**	O95872 GPAN1_HUMAN	canonical sequence	**MSRPLLITFTPATDPSDLWKDGQQQPQPEKPESTLDGAAARAFYEALIGDESSAPDSQRSQTEPARE RKRKKRRIMKAPAAEAVAEGASGRHGQGRSLEAEDKMTHRILRAAQEGDLPELRRLLEPHEAGGAG GNINARDAFWWTPLMCAARAG**

ID66HJ47 Clone 251	General transcription factor 3C polypeptide 2	**GTF3C2**	Q8WUA4 TF3C2_HUMAN	canonical sequence + isoform 2	**MDTCGVGYVALGEAGPVGNMTVVDSPGQEVLNQLDVKTSSEMTSAEASVEMSLPTPLPGFEDSPDQ RRLPPEQESLSRLEQPDLSSEMSKVSKPRASKPGRKRGGRTRKGPKRPQQPNPPSAPLVPGLLDQS NPLSTPMPKKRGRKSKAELLLLKLSKDLDRPESQSPKRPPEDFETPSGERPRRRAAQVALLYLQELA EELSTALPAPVSCPEGPKVSSPTKPKKIRQPAACPGGEEVDGAPRDEDFFLQVEAEDVEESEGPSESS SEPEPVVPRSTPRGSTSGKQKPH**
ID_66IA04 cl348	Guanylate kinase (Fragment)	**GUK1**	B1ANH3 B1ANH3_HUMAN	canonical sequence	**RHVGPQACGAERAFGSWEEHPAEEAAPGAQRHLWLQRVPYHEEPEARRGERQREVMQRDIAAGDF IEHAEFSGNLYGTSKVAVQAVQAMNRICVLDVDLQGVRNIKATDLRPIYISVQPPSLHVLEQRLRQR**
ID07HJ53 cl 205	Guanylate kinase	**GUK1**	Q16774 KGUA_HUMAN	canonical + isoform 2	**LKRLLQEHSGIFGFSVSHTTRNPRPGEENGKDYYFVTREVMQRDIAAGDFIEHAEFSGNLYGTSKVAV QAVQAMNRICVLDVDLQGVRNIKATDLRPIYISVQPPSLHVLEQRLRQRNTETEESLVKRLAAAQADM ESSKEPGLFDVVIINDSLDQAYAELKEALSEEIKKAQRTGA**
ID07HJ90 cl 313	Intraflagellar transport protein 88 homolog	**IFT88**	Q13099 IFT88_HUMAN	canonical + isoform 2	**EELENDAAFQQAVRTSHGRRPPITAKISSTAVTRPIATGYGSKTSLASSIGRPMTGAIQDGVTRPMTAV RAAGFTKAALRGSAFDPLSQSRGPASPLEAKKKDSPEEKIKQLEKEVNELVEESCIANSCGDLKLALE KAKDAGRKERVLVRQREQVTTPENINLDLTYSVLFNLASQYSVNEMYAEALNTYQVIVKNKMFSNAGI LKMNMGNIYLKQRNYSKAIKFYRMALDQVPSVNKQMRIK**
ID66HJ52 Clone 286	Inositol-trisphosphate 3-kinase	**ITPKC**	Q96DU7 IP3KC_HUMAN	canonical sequence	**LVKARERPRPRKDMYEKMVAVDPGAPTPEEHAQGAVTKPRYMQWRETMSSTSTLGFRIEGIKKADGT CNTNFKKTQALEQVTKVLEDFVDGDHVILQKYVACLEELREALEISPFFKTHEV**
ID07HJ79 cl 736	Kelch-like protein 10	**KLHL10**	Q6JEL2 KLH10_HUMAN	canonical sequence	**GIGVIAYGEHVYAVGGFDGANRLRSAEAYSPVANTWRTIPTMFNPRSNFGIEVVDDLLFVVGGFNGFTT TFNVECYDEKTDEWYDAHDMSIYRSALSCCVVPGLANVEEYAARRDNFPGLALRDEVKYSASTSTLP V**
ID07HJ30 cl 144	Microtubule-associated protein 1B	**MAP1B**	P46821 MAP1B_HUMAN	canonical sequence	**ECPSITADANIDSEDESETIPTDKTVTYKHMDPPPAPVQDRSPSPRHPDVSMVDPEALAIEQNLGKALK KDLKEKTKTKKPGTKTKSSSPVKKSDGKSKPLAASPKPAGLKESSDKVSRVASPKKKESVEKAAKPT TTPEVKAARGEEKDKETKNAANASASKSAKTATAGPGTTKTTKSSAVPPGLPVYLDLCYIPNHSNSKN VDVEFFKRVRSSYYVVSGNDPAAEEPSRAVLDALLEGKAQWGSNMQVTLIPTHDSEVMREWYQETH EKQQDLNIMVLASSSTVVMQDESFPACKIEL**
ID66HJ57 Clone 194	Microtubule-associated protein 1S	**MAP1S**	Q66K74 MAP1S_HUMAN	canonical + isoform 2	**GAEETPPTSVSESLPTLSDSDPVPLAPGAADSDEDTEGFGVPRHDPLPDPLKVPPPLPDPSSICMVDP EMLPPKTARQTENVSRTRKPLARPNSRAAAPKATPVAAAKTKGLAGGDRASRPLSARSEPSEKGGRR APLSRKSSTPKTATRGPSGSASSRPGVSATPPKSPVYLDLAYLPSGSSAHLVDEEFFQRVRALCYVIS GQDQRKEEGMRAVLDALLASKQHWDRDLQVTLIPTFDSVAMHTWYAETHARHQALGITVLGSNSMV SMQDDAFP**
ID_66HJ96 cl287	Microtubule-associated protein 1S	**MAP1S**	Q66K74 MAP1S_HUMAN	canonical + isoform 2	**RKPLARPNSRAAAPKATPVAAAKTKGLAGGDRASRPLSARSEPSEKGGRAPLSRKSSTPKTATRGP SGSASSRPGVSATPPKSPVYLDLAYLPSGSSAHLVDEEFFQRVRALCYVISGQDQRKEEGMRAVLDA LLASKQHWDRDLQVTLIPTFDSVAMHTWYAETHARHQALGITVLGSNSMVSMQDDAFPACKVEF**
ID 07IA33 cl 473	Microtubule-associated protein 1S	**MAP1S**	Q66K74 MAP1S_HUMAN	canonical + isoform 2	**VTTPSLPAEVGSPHSTEVDESLSVSFEQVLPPSAPTSEAGLSLPLRGPRARRSASPHDVDLCLVSPCE FEHRKAVPMAPAPASPGSSNDSSARSQERAGGLGAEETPPTSVSESLPTLSDSDPVPLAPGAADSDE DTEGFGVPRHDPLPDPLKVPPPLPDPSSICMVDPEMLPPKTARQTENVSRTRKPLARPNSRAAAPKA TPVAAAKTKGLAGGDRASRPLSARSEPSEKGGRAPLSRKSSTPKTATRGPSGSA**
ID07HJ45 cl 156	Microtubule-associated protein 1S	**MAP1S**	Q66K74 MAP1S_HUMAN	canonical + isoform 2	**KTARQTENVSRTRKPLARPNSRAAAPKATPVAAAKTKGLAGGDRASRPLSARSEPSEKGGRAPLSR KSSTPKTATRGPSGSASSRPGVSATPPKSPVYLDLAYLPSGSSAHLVDEEFFQRVRALCYVISGQDQR KEEGMRAVLDALLASKQHWDRDLQVTLIPTFDSVAMHTWYAETHARHQALGITVLGSNSMVSMQDD AFPACKVEF**
ID_66HJ84 cl630	28S ribosomal protein S9, mitochondrial	**MRPS9**	P82933 RT09_HUMAN	canonical sequence	**QILRLRHTAFVIPKKNVPTSKRETYTEDFIKKQIEEFNIGKRHLANMMGEDPETFTQEDIDRAIAYLFPSG LFEKRARPVMKHPEQIFPRQRAIQWGEDGRPFHYLFYTGKQSYYSLMHDVYGMLLNLEKHQSHLQA KSLLPEKTVTRDVIGSRWLIKEELEEMLVEKLSDLDYMQFIRLLEKLLTSQCGAAEEEFVQRFRRSVTL ESKKQLIEPVQYDEQGMAFSKSEGKRKTAKAEAIVYKHGSGRIKVNGIDYQLYFPITQDREQLMFPFHF VDRLGKHDVTCTVSGGG**
ID07HJ83 cl 309	E3 ubiquitin-protein ligase MSL2	**MSL2**	Q9HCI7 MSL2_HUMAN	canonical + isoform 2	**KRESKISLQPIATVPNGGTTPKISKTVLLSTKSMKKSHEHGSKKSHSKTKPGILKKDKAVKEKIPSHHF MPGSPTKTVYKKPQEKKGCKCGRATQNPSVLTCRGQRCPCYSNRKACLDCICRGCQNSYMANGEK KLEAFAVPEKALEQTRLTLGINVTSIAVRNASTSTSVINVTGSPVTTFLAASTHDDKSLDEAIDMRFDC**
ID66HJ14 Clone 102	Cytochrome c oxidase subunit 2	**MT-CO2**	P00403 COX2_HUMAN	canonical sequence	**GLKTDAIPGRLNQTTFTATRPGVYYGQCSEICGANHSFMPIVLELIPLKIFEMGPVFTL**
ID66HJ39 Clone 235	Cytochrome c oxidase subunit 2	**MT-CO2**	P00403 COX2_HUMAN	canonical sequence	**GLKTDAIPGRLNQTTFTATRPGVYYGQCSEICGANHSFMPIVLELIPLKIFEMGPVFTL**

ID07HJ19 cl 561	Nischarin	**NISCH**	Q9Y2I1 NISCH_HUMAN	canonical + isoform 2	**DCVHYPLPEFAKEPPQRDRYRLDDGRRVRDLDRVLMGYQTYPQALTLVFDDVQGHDLMGSVTLDHF GEVPGGPARASQGREVQWQVFVPSAESREKLISLLARQWEALCGRELPVELTG**
ID66HJ40 Clone 236	Probable 28S rRNA (cytosine-C(5))-methyltransferase	**NSUN5**	Q96P11 NSUN5_HUMAN	canonical + isoform 2, 5	**PSDPRYHEVHYILLDPSCSGSGMPSRQLEEPGAGTPSPVRLHALAGFQQRALCHALTFPSLQRLVYS TCSLCQEENEDVVRDALQQNPGAFRLAPALPAWPHRGLSTFPGAEHCLRASPETTLSSGFFVAVIERV EVP**

ID_66HI77_2C_cl37	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETKAS ANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRERQTHFLRPRPEDRMV ALKPIKIVLPPVFMPFSSPQGIRSRMPTQHLYCRWVAPKALVPTCLPTHSLS**
ID_66HI80_11C_cl46	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLAR**
ID_66HI90_4E_cl61	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETKAS ANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRERQTHFLRPRPEDRMV ALKPIKIVLPPVFMPFSSPQGIRSRMPTQHLYCRWVAPKALVPTCLPTHSLSRKEEKPKDDTPSSLLRP PRRWL**
ID_66HJ07 _2H _cl111	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETKAS ANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRERQTHFLRPRPEDRMV ALKPIKIVLPPVFMPFSSPQGIRSRMPTQHLYCRWVAPKALVPTCLPTHSLSRKEEKPKDDTPSSLLRP PRRWL**
ID66HJ21 Clone 145	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETKAS ANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRERQTHFLRPRPEDRMV ALKPIKIVLPPVFMPFSSPQGIRSRMPTQHLYCRWVAPKALVPTCLPTHSLSRKEEK**
ID66HJ22 Clone 148	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETKAS ANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRERQTHFLRPRPEDRMV ALKPIKIVLPPVFMPFSSPQGIRSRMPTQHLYCRWVAPKALVP**
ID66HJ23 Clone 150	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETKAS ANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRERQTHFLRPRPEDRMV ALKPIKIVLPPVFMPFSSPQGIRSRMPTQHLYCRWVAPKALVPTCLPTHSLSRKEEKPKDDTPSSLLRP PRRWL**
ID66HJ29 Clone 207	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETKAS ANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRER**
ID66HJ45 Clone 247	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETKAS ANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRERQTHFLRPRPEDRMV ALKPIKIVLPPVFMPFSSPQGIRSRMPTQHLYCRWVAPKALVPTCLPTHSLSRKEEKPKDDTPSSLLRP PRRWL**
ID07HJ31 cl 155	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETKAS ANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRERQTHFLRPRPEDRMV ALKPIKIVLPPVFMPFSSPQGIRSRMPTQHLYCRWVAPKALVPTCLPTHSLSRKEEKPKDDTPSSLLRP PRRWL**
ID 07IA37 cl 57	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**PGDPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETK ASANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRERQTHFLRPRPEDR MVALKPIKIVLPPVFMPFSSPQGIRSRMPTQHLYCRWVAPKALVPTCLPTHSLSRKEEKPKDDTPSSLL RPPRRWL**
ID07HJ46 cl 151	Putative methyltransferase NSUN7	**NSUN7**	Q8NE18 NSUN7_HUMAN	canonical sequence	**DPSETVSVNDVLARAAAKGLLDGIELGKSSKREKKKKKSKTSLTKGATTDNGIQMKIAEFLNRETKAS ANLSETVTKPPLPQKNTAQVGASSQTRKPNKLAPHPAVPAFVKNTCPSRPRERQTHFLRPRPEDRMV ALKPIKIVLPPVFMPFSSPQGIRSRMPTQHLYCRWVAPKALVPTCLPTHSLSRKEEKPKDDTPSSLLRP PRRWL**
ID_66HI98 _7F _cl81	Protein disulfide-isomerase A3	**PDIA3**	P30101 PDIA3_HUMAN	canonical sequence	**SRDGKALERFLQDYFDGNLKRYLKSEPIPESNDGPVKVVVAENFDEIVNNENKDVLIEFYAPWCGHCK NLEPKYKELGEKLSKDPNIVIAKMDATANDVPSPYEVRGFPTIYFSPANKKLNPKKYEGGRELSDFISY LQREATNPPVIQEEKPKKKKK**
ID_66HJ11 _8H _cl124	Protein disulfide-isomerase A3	**PDIA3**	P30101 PDIA3_HUMAN	canonical sequence	**DGKALERFLQDYFDGNLKRYLKSEPIPESNDGPVKVVVAENFDEIVNNENKDVLIEFYAPWCGHCKNL EPKYKELGEKLSKDPNIVIAKMDATANDVPSPYEVRGFPTIYFSPANKKLNPKKYEGGRELSDFISYLQ REATNPPVIQEEKPKKKKK**
ID66HJ56 Clone 187	Protein disulfide-isomerase A3	**PDIA3**	P30101 PDIA3_HUMAN	canonical sequence	**DGKALERFLQDYFDGNLKRYLKSEPIPESNDGPVKVVVAENFDEIVNNENKDVLIEFYAPWCGHCKNL EPKYKELGEKLSKDPNIVIAKMDATANDVPSPYEVRGFPTIYFSPANKKLNPKKYEGGRELSDFISYLQ REATNPPVIQEEKPKKKKK**
ID66HJ58 Clone 198	Protein disulfide-isomerase A3	**PDIA3**	P30101 PDIA3_HUMAN	canonical sequence	**SRDGKALERFLQDYFDGNLKRYLKSEPIPESNDGPVKVVVAENFDEIVNNENKDVLIEFYAPWCGHCK NLEPKYKELGEKLSKDPNIVIAKMDATANDVPSPYEVRGFPTIYFSPANKKLNPKKYEGGRELSDFISY LQREATNPPVIQEEKPKKKKK**

ID07HJ78 cl 732	ATP-dependent 6-phosphofructokinase, platelet type	**PFKP**	Q01813 PFKAP_HUMAN	canonical + isoform 2	**TCDRIKQSASGTKRRVFIIETMGGYCGYLANMGGLAAGADAAYIFEEPFDIRDL**
ID07HJ99 cl 381	Prohibitin-2	**PHB2**	Q99623 PHB2_HUMAN	canonical sequence	**TVEGGHRAIFFNRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRARPRKISSPTGSKDLQMVNISLRVLSRP NAQELPSMYQRLGLDYEERVLPSIVNEVLKSVVAKFNASQLITQRAQVSLLIRRELTERAKDFSLILDDV AITELSFSREYTAAVEAKQVAQQEAQRAQFLVEKAKQEQRQKIVQAEGEAEAAKMLGEALSKNPGYI KLRKIRAAQNISKTIATSQNRIYLTADNLVLNLQDESFTRGSDSLIKGKK**

ID07HJ34 cl 439	Prohibitin-2	**PHB2**	Q99623 PHB2_HUMAN	canonical sequence	**TVEGGHRAIFFNRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRARPRKISSPTGSKDLQMVNISLRVLSRP NAQELPSMYQRLGLDYEERVLPSIVNEVLKSVVAKFNASQLITQRAQVSLLIRRELTERAKDFSLILDDV AITELSFSREYTAAVEAKQVAQQEAQRAQFLVEKAKQEQRQKIVQAEGEAEAAKMLGEALSKNPGYI KLRKIRAAQNISKTIATSQNRIYLTADNLVLNLQDESFTRG**
ID66HJ33 Clone 218	Piwi-like protein 4	**PIWIL4**	Q7Z3Z4 PIWL4_HUMAN	canonical sequence	**CRGTVSPTYYNVIYDDNGLKPDHMQRLTFKLCHLYYNWPGIVSVPAPCQYAHKLTFLVAQSIHKEPSLE LANHLFYL**
ID_66IA07 cl815	Piwi-like protein 4	**PIWIL4**	Q7Z3Z4 PIWL4_HUMAN	canonical sequence	**RGTVSPTYYNVIYDDNGLKPDHMQRLTFKLCHLYYNWPGIVSVPAPCQYAHKLTFLVAQSIHKEPSLEL ANHLFYL**
ID07HJ47 cl 600	Plakophilin-2	**PKP2**	Q99959 PKP2_HUMAN	canonical sequence	**PGAVAHACNPSTLGGQGGRITRSGVRDQPDQHG**
ID_66HJ05 _10G _cl104	cGMP-dependent protein kinase 1	PRKG1	Q13976 KGP1_HUMAN	canonical + isoform beta and 3	**SFKHLIGGLDDVSNKAYEDAEAKAKYEAEAAFFANLKLSDFNIIDTLGVGGFGRVEL**
ID66HJ35 Clone 226	cGMP-dependent protein kinase 1	**PRKG1**	Q13976 KGP1_HUMAN	canonical + isoform beta and 3	**SFKHLIGGLDDVSNKAYEDAEAKAKYEAEAAFFANLKLSDFNIIDTLGVGGFGRVEL**
ID_66HJ91 cl261	26S proteasome regulatory subunit 4	**PSMC1**	P62191 PRS4_HUMAN	canonical + isoform 2	**MGQSQSGGHGPGGGKKDDKDKKKKYEPPVPTRVGKKKKKTKGPDAASKLPLVTPHTQCRLKLLKL ERIKDYLLMEEEFIRNQEQMKPLEEKQEEERSKVDDLRGTPMSVGTLEEIIDDNHAIVSTSVGSEHYVSI LSFVDKDLLEPGCSVLLNHKVHAVIGVLMDDTDPLVTVMKVEKAPQETYADIGGLDNQIQEIKESVELP LTHPEYYEEMGIKPPKGVILYGPPGTGKTLLAKAVANQTSATFLRVVGSELIQKYLGDGPKLVRELFRV AEEHAPSIVFIDEIDAIGTKR**
ID_66IA00 cl318	26S proteasome regulatory subunit 4	**PSMC1**	P62191 PRS4_HUMAN	canonical + isoform 2	**MGQSQSGGHGPGGGKKDDKDKKKKYEPPVPTRVGKKKKKTKGPDAASKLPLVTPHTQCRLKLLKL ERIKDYLLMEEEFIRNQEQMKPLEEKQEEERSKVDDLRGTPMSVGTLEEIIDDNHAIVSTSVGSEHYVSI LSFVDKDLLEPGCSVLLNHKVHAVIGVLMDDTDPLVTVMKVEKAPQETYADIGGLDNQIQEIKESVELP LTHPEYYEEMGIKPPKGVILYGPPGTGKTLLAKAVANQTSATFLRV**
ID07IA03 cl 387	26S proteasome regulatory subunit 4	**PSMC1**	P62191 PRS4_HUMAN	canonical + isoform 2	**MGQSQSGGHGPGGGKKDDKDKKKKYEPPVPTRVGKKKKKTKGPDAASKLPLVTPHTQCRLKLLKL ERIKDYLLMEEEFIRNQEQMKPLEEKQEEERSKVDDLRGTPMSVGTLEEIIDDNHAIVSTSVGSEHYVSI LSFVDKDLLEPGCSVLLNHKVHAVIGVLMDDTDPLVTVMKVEKAP**
ID07HJ33 cl 911	Arginine-glutamic acid dipeptide repeats protein	**RERE**	Q9P2R6 RERE_HUMAN	canonical sequence	**EDNDNNSATAEESTKKNKKKPPKKKSRYERTDTGEITSYITEDDVVYRPGDCVYIESRRPNTPYFICSIQ DFKLVH**
ID07IA01 cl 378	Arginine-glutamic acid dipeptide repeats protein	**RERE**	Q9P2R6 RERE_HUMAN	canonical sequence	**EDNDNNSATAEESTKKNKKKPPKKKSRYERTDTGEITSYITEDDVVYRPGDCVYIESRRPNTPYFICSIQ DFKLVH**
ID66HJ27 Clone 182	Histone deacetylase complex subunit SAP30	**SAP30**	O75446 SAP30_HUMAN	canonical sequence	**PGPGQLCCLREDGERCGRAAGNASFSKRIQKSISQKKVKIELDKSARHLYICDYHKNLIQSVRNRRKR KGSDDDGGDSPVQDIDTPEVDLYQLQVNTLRRYKRHFKLPTRPGLNKAQLVEIVGCHFRSIPVNEKDT LTYFIYSVKNDKNK**
ID07IA11 cl 892	Histone deacetylase complex subunit SAP30	**SAP30**	O75446 SAP30_HUMAN	canonical sequence	**GRAAGNASFSKRIQKSISQKKVKIELDKSARHLYICDYHKNLIQSVRNRRKRKGSDDDGGDSPVQDID TPEVDLYQLQVNTLRRYKRHFKLPTRPGLNKAQLVEIVGCHFRSIPVNEKDTLTYFIYSVKNDKNKSDL KVDSGVH**
ID07HJ51 cl 637	Histone deacetylase complex subunit SAP30	**SAP30**	O75446 SAP30_HUMAN	canonical sequence	**GQLCCLREDGERCGRAAGNASFSKRIQKSISQKKVKIELDKSARHLYICDYHKNLIQSVRNRRKRKGS DDDGGDSPVQDIDTPEVDLYQLQVNTLRRYKRHFKLPTRPGLNKAQLVEIVGCHFRSIPVNEKDTLTYF IYSVKNDKNKSDLKVDSGVH**
ID07HJ38 cl 475	Selenoprotein P	**SELENOP**	P49908 SEPP1_HUMAN	canonical sequence	**NQTDVWTLLNGSKDDFLIYDRCGRLVYHLGLPFSFLTFPYVEEAIKIAYCEKKCGNCSLTTLKDEDFCK RVSLATVDKTVETPSPHYHHEHHHNHGHQHLGSSELSENQQPGAPNAPTHPAPPGLHHHHKHKGQH RQGHPENRDMPASEDLQDLQKKLCRKRCINQLLCKLPTDSELAPRS**
ID_66HJ89 cl256	Structural maintenance of chromosomes protein 5	**SMC5**	Q8IY18 SMC5_HUMAN	canonical sequence	**MATPSKKTSTPSPQPSKRALPRDPSSEVPSKRKNSAPQLPLLQSSGPFVEGSIVRISMENFLTYDICEV SPGPHLNMIVGANGTGKSSIVCAICLGLAGKPAFMGRADKVGFF**
ID07IA08 cl 400	Structural maintenance of chromosomes protein 5	**SMC5**	Q8IY18 SMC5_HUMAN	canonical sequence	**MATPSKKTSTPSPQPSKRALPRDPSSEVPSKRKNSAPQLPLLQSSGPFVEGSIVRISMENFLTYDICEV SPGPHLNMIVGANGTGKSSIVCAICLGLAGKPAFMGRADKVGFFV**
ID_66HI94 _12E _cl68	U5 small nuclear ribonucleoprotein 200 kDa helicase	**SNRNP200**	O75643 U520_HUMAN	canonical sequence	**DSEVILHHEYFLLKAKYAQDEHLITFFVPVFEPLPPQYFIRVVSDRWLSCETQLPVSFRHLILPEKYPPP TELLDLQPLPVSALRNSAFESLYQDKFPFFNPIQTQ**

ID66IA56 cl 72	SPARC	**SPARC**	P09486 SPRC_HUMAN	canonical sequence	**EGTKKGHKLHLDYIGPCKYIPPCLDSELTEFPLRMRDWLKNVLVTLYERDEDNNLLTEKQKLRVKKIHE NEKRLEAGDHPVELLARDFEKNYNMYIFPVHWQFGQLDQHPIDGYLSHTELAPLRAPLIPMEHCTTRF FETCDLDNDKYIALDEWAGCFGIKQKDIDKDLVI**
ID_66HJ02 _4G _cl96	Spermatogenesis-associated protein 22	**SPATA22**	Q8NHS9 SPT22_HUMAN	canonical + isoform 3	**MKRSLNENSARSTAGCLPVPLFNQKKRNRQPLTSNPLKDDSGISTPSDNYDFPPLPTDWAWEAVNPE LAPVMKTVDTGQIPHSVSRPLRSQDSVFNSIQSNTGRSQGGWSYRDGNKNTSLKTWNKNDFKPQCK RTNLVANDGKNSCPVSSGAQQQKQLRIPEPPNLSRNKETELLRQTHSSKISGCTMRGLDKNSALQTL KPNFQQNQYKKQMLDDIPEDNTLKETSLYQLQFKEKASSLRIISAVIES**
ID_66HJ04 _8G _cl100	Spermatogenesis-associated protein 22	**SPATA22**	Q8NHS9 SPT22_HUMAN	canonical + isoform 3	**RNRQPLTSNPLKDDSGISTPSDNYDFPPLPTDWAWEAVNPELAPVMKTVDTGQIPHSVSRPLRSQDS VFNSIQSNTGRSQGGWSYRDGNKNTSLKTWNKNDFKPQCKRTNLVANDGKNSCPVSSGAQQQKQL RIPEPPNLSRNKETELLRQTHSSKISGCTMRGLDKNSALQTLKPNFQQNQYKKQMLDDIPEDNTLKET SLYQLQFKEKASSLRIISAVIESMKYWREHAQKTVLLFEVLAV**
ID66HJ17 Clone 137	Spermatogenesis-associated protein 22	**SPATA22**	Q8NHS9 SPT22_HUMAN	canonical + isoform 2	**DWAWEAVNPELAPVMKTVDTGQIPHSVSRPLRSQDSVFNSIQSNTGRSQGGWSYRDGNKNTSLKTW NKNDFKPQCKRTNLVANDGKNSCPVSSGAQQQKQLRIPEPPNLSRNKETELLRQTHSSKISGCTMRG LDKNSALQTLKPNFQQNQYKKQMLDDIPEDNTLKETSLYQLQFKEKASSLRIISAVIESMKYWREHAQ KTVLLFEVLAVLDSAVTPGPYYSKTFLMRDGKNTLPCVFYEIDRELPRLIRGRVHRCVGNYDQKKN**
ID66HJ30 Clone 215	Spermatogenesis-associated protein 22	**SPATA22**	Q8NHS9 SPT22_HUMAN	canonical + isoform 3	**FNQKKRNRQPLTSNPLKDDSGISTPSDNYDFPPLPTDWAWEAVNPELAPVMKTVDTGQIPHSVSRPL RSQDSVFNSIQSNTGRSQGGWSYRDGNKNTSLKTWNKNDFKPQCKRTNLVANDGKNSCPVSSGAQ QQKQLRIPEPPNLSRNKETELLRQTHSSKISGCTMR**
ID66HJ37 Clone 231	Spermatogenesis-associated protein 22	**SPATA22**	Q8NHS9 SPT22_HUMAN	canonical sequence	**PLFNQKKRNRQPLTSNPLKDDSGISTPSDNYDFPPLPTDWAWEAVNPELAPVMKTVDTGQIPHSVSR PLRSQDSVFNSIQSNTGRSQGGWSYRDGNKNTSLKTWNKNDFKPQCKRTNLVANDGKNSCPVSSG AQQQKQLRIPEPPNLSRNKETELLRQTHSSKISGCTMRGLDKNSALQTLKPNFQQNQYKKQMLDDIP EDNTLKETSLYQLQFKEKASSLRIISAVIESMKYWREHAQKTVLLFEVLAVLDSAVTPGPYYSKTFLMR DGKNTLPCVFYEIDRELPRLIRGRVHRCVGNYDQKKNIF**
ID_66IA20 cl830	Spermatogenesis-associated protein 22	**SPATA22**	Q8NHS9 SPT22_HUMAN	canonical sequence	**FNQKKRNRQPLTSNPLKDDSGISTPSDNYDFPPLPTDWAWEAVNPELAPVMKTVDTGQIPHSVSRPL RSQDSVFNSIQSNTGRSQGGWSYRDGNKNTSLKTWNKNDFKPQCKRTNLVANDGKNSCPVSSGAQ QQKQLRIPEPPNLSRNKETELLRQTHSSKISGCTMRGLDKNSALQTLKPNFQQNQYKKQMLDDIPED NTLKETSLYQLQFKEKASSLRIISAVIESMKYWREHAQKTVLLFEVLAVLDSAVTPGPYYSKTFLM**
ID07HJ17 cl 97	Spermatogenesis-associated protein 22	**SPATA22**	Q8NHS9 SPT22_HUMAN	canonical + isoform 2	**EAVNPELAPVMKTVDTGQIPHSVSRPLRSQDSVFNSIQSNTGRSQGGWSYRDGNKNTSLKTWNKND FKPQCKRTNLVANDGKNSCPVSSGAQQQKQLRIPEPPNLSRNKETELLRQTHSSKISGCTMRGLDKN SALQTLKPNFQQNQYKKQMLDDIPEDNTLKETSLYQLQFKEKASSLRIISAVIESMKYWREHAQKTVL LFEVLAVLDSAVTPGPYYSKTFLMRDGKNTLPCVFYEIDRELPRLIRGRVHRCVGNYDQKKNIFQCVSV RPASVSEQKTFQAFVKIADVEMQYYINVMNET**
ID07HJ29 cl 158	Spermatogenesis-associated protein 22	**SPATA22**	Q8NHS9 SPT22_HUMAN	canonical + isoform 2	**PLFNQKKRNRQPLTSNPLKDDSGISTPSDNYDFPPLPTDWAWEAVNPELAPVMKTVDTGQIPHSVSR PLRSQDSVFNSIQSNTGRSQGGWSYRDGNKNTSLKTWNKNDFKPQCKRTNLVANDGKNSCPVSSG AQQQKQLRIPEPPNLSRNKETELLRQTHSSKISGCTMRGLDKNSALQTLKPNFQQNQYKKQMLDDIP EDNTLKETSLYQLQFKEKASSLRIISAVIESMKYWREHAQKTVLLFEVLAVLDSAVTPGPYYSKTFLMR DGKNTLPCVFYEIDRELPRLIRGRVHRCVGNYDQKKNIFQCVSVRPASVSEQKTFQAFVKIADVEMQY Y**
ID 07IA35 cl 915	Spermatogenesis-associated protein 22	**SPATA22**	Q8NHS9 SPT22_HUMAN	canonical + isoform 3	**PLFNQKKRNRQPLTSNPLKDDSGISTPSDNYDFPPLPTDWAWEAVNPELAPVMKTVDTGQIPHSVSR PLRSQDSVFNSIQSNTGRSQGGWSYRDGN**
ID07IA06 cl 393	Spermatogenesis-associated protein 22	**SPATA22**	Q8NHS9 SPT22_HUMAN	canonical sequence	**QKKRNRQPLTSNPLKDDSGISTPSDNYDFPPLPTDWAWEAVNPELAPVMKTVDTGQIPHSVSRPLRS QDSVFNSIQSNTGRSQGGWSYRDGNKNTSLKTWNKNDFKPQCKRTNLVANDGKNSCPVSSGAQQQ KQLRIPEPPNLSRNKETELLRQTHSSKISGCTMRGLDKNSALQTLKPNFQQNQYKKQMLDDIPEDNTL KETSLYQLQFKEKASSLRIISAVIESMKYWREHAQKTVLLFEVLAVLDSAVTPGPYYSKTFLMRDGKNT LP**
ID_66HI89_12D_cl58	Spermatogenic leucine zipper protein 1	**SPZ1**	Q9BXG8	canonical isoform	**KHELEEQVKKLSHDTYSLQLMAALLENECQILQQRVEILKELHHQKQGTLQEKPIQINYKQDKKNQKP SEAKKVEMYKQNKQAMKGTFWKKDRSCRSLDVCLNKKACNTQFNIHVARKALRGKMRSASSLR**
ID66HJ60 Clone 642	STAM-binding protein	**STAMBP**	O95630 STABP_HUMAN	canonical sequence	**QQKQQQLEQEQFHAFEEMIRNQELEKERLKIVQEFGKVDPGLGGPLVPDLEKPSLDVFPTLTVSSIQP SDCHTTVRPAKPPVVDRSLKPGALSNSESIPTIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILC GKLMRNEFTITHVLIPKQSAGSDYCNTENEEELFLIQDQQGLITLGWIHTHPTQTAFLSSVDLHTHCSY QMMLPESVAIVCSPKFQETGFFKLTDHGLEEISSCRQKGFHPHSKDPPLFCSCSHVTVVDRAVTITDLR**

ID66IA55 cl 523	STAM-binding protein	**STAMBP**	O95630 STABP_HUMAN	canonical sequence	**QQKQQQLEQEQFHAFEEMIRNQELEKERLKIVQEFGKVDPGLGGPLVPDLEKPSLDVFPTLTVSSIQP SDCHTTVRPAKPPVVDRSLKPGALSNSESIPTIDGLRHVVVPGRLCPQFLQLASANTARGVETCGILC GKLMRNEFTITHVLIPKQSAGSDYCNTENEEELFLIQDQQGLITLGWIHTHPTQTAFLSSVDLHTHCSY QMMLPESVAIVCSPK**
ID07IA00 cl 383	Histone-lysine N-methyltransferase SUV39H2	**SUV39H2**	Q9H5I1 SUV92_HUMAN	canonical + isoform 1,2	**RGQFYDNKGITYLFDLDYESDEFTVDAARYGNVSHFVNHSCDPNLQVFNVFIDNLDTRLPRIALFSTRTI NAGEELTFDYQMKGSGDISSDSIDHSPAKKRVRTVCKCGAVTCRGYLN**
ID66IA49 cl 880	Tudor domain-containing protein 7	**TDRD7**	Q8NHU6 TDRD7_HUMAN	canonical + isoform 2, 3	**ELREIPPRFLQEMIAIPPQAIKCCLADLPQSIGMWTPDAVLWLRDSVLNCSDCSIKVTKVDETRGIAHVY LFTPKNFPDPHRSINRQITNADLWKHQKDVFLSAISSGADSPNSKNGNMPMSGNTGENFRKNLTDVIK KSMVDHTSAFSTEELPPPVHLSKPGEH**
ID_66HI92_7E_cl65	Testis-expressed protein 35	**TEX35**	Q5T0J7 TEX35_HUMAN	canonical + isoform 2, 3, 5	**MSAKRAELKKTHLSKNYKAVCLELKPEPTKTFDYKAVKQEGRFTKAGVTQDLKNELREVREELKEKM EEIKQIKDLMDKDFDKLHEFVEIMKEMQKDMDEKMDILINTQKNYKLPLRRAPKEQQELRLMGKTHRE PQLRPKKMDGASGVNGAPCALHKKTMAPQKTKQG**

ID07IA12 cl 418	Testis-expressed protein 35	**TEX35**	Q5T0J7 TEX35_HUMAN	canonical + isoform 2, 3, 5	**MSAKRAELKKTHLSKNYKAVCLELKPEPTKTFDYKAVKQEGRFTKAGVTQDLKNELREVREELKEKM EEIKQIKDLMDKDFDKLHEFVEIMKEMQKDMDEKMDILINTQKNYKLPLRRAPKEQQELRLMGKTHRE PQLRPKKMDGASGVNGAPCALHKKTMAPQKTKQGSLDPLHHCGTCCEKCLLCALKNNYNRG**
ID66HJ24 Clone 165	E3 ubiquitin-protein ligase TTC3	**TTC3**	P53804 TTC3_HUMAN	canonical + isoforms TPRDII and TPRDIII	**DSRCTVLRKQDSGEAPFSSTKVKNKSKKKKPKDSKPMLVGSGTTSVTSNNEIITSSEDHSNRNSDSA GPFAVPDHLRQDVEEFEALYDQHSNEYVVRNKKLWDMNPKQKCSTLYDYFSQFLEEHGPLDMSNKM FSAEYEFFPEETRQILEKAGGLKPFLLGCPRFVVIDNCIALKKVASRLKKKRKKK**
ID66HJ25 Clone 166	E3 ubiquitin-protein ligase TTC3	**TTC3**	P53804 TTC3_HUMAN	canonical + isoforms TPRDII and TPRDIII	**DSRCTVLRKQDSGEAPFSSTKVKNKSKKKKPKDSKPMLVGSGTTSVTSNNEIITSSEDHSNRNSDSA GPFAVPDHLRQDVEEFEALYDQHSNEYVVRNKKLWDMNPKQKCSTLYDYFSQFLEEHGPLDMSNKM FSAEYEFFPEETRQILEKAGGLKPFLLGCPRFVVIDNCIALKKVASRLKKKRKKK**
ID66HJ46 Clone 249	E3 ubiquitin-protein ligase TTC3	**TTC3**	P53804 TTC3_HUMAN	canonical + isoforms TPRDII and TPRDIII	**DSRCTVLRKQDSGEAPFSSTKVKNKSKKKKPKDSKPMLVGSGTTSVTSNNEIITSSEDHSNRNSDSA GPFAVPDHLRQDVEEFEALYDQHSNEYVVRNKKLWDMNPKQKCSTLYDYFSQFLEEHGPLDMSNKM FSAEYEFFPEETRQILEKAGGLKPFLLGCPRFVVIDNCIALKKVASRLKKKRKKKNIKTKVEEISKAGEY VRVKLQLNPAAREFKPDVKSKPVSDSSSAPAFENVKPKPVSANSPKP**
ID66HJ42 Clone 691	Ubiquitin-conjugating enzyme E2 D2	**UBE2D2**	P62837 UB2D2_HUMAN	canonical sequence	**KELNDLARDPPAQCSAGPVGDDMFHWQATIMGPNDSPYQGGVFFLTIHFPTDYPFKPPKVAFTTRIYH PNINSNGSICLDILRSQWSPALTISKVLLSICSLLCDPNPDDPLVPEIARIYKTDREKYNRIAREWTQKYA M**
ID07HJ14 cl 533	Ubiquitin-conjugating enzyme E2 D2	**UBE2D2**	P62837 UB2D2_HUMAN	canonical sequence	**KELNDLARDPPAQCSAGPVGDDMFHWQATIMGPNDSPYQGGVFFLTIHFPTDYPFKPPKVAFTTRIYH PNINSNGSICLDILRSQWSPALTISKVLLSICSLLCD**
ID_66HI81_1D_cl48	Zinc finger and BTB domain-containing protein 17	**ZBTB17**	Q13105 ZBT17_HUMAN	canonical + isoforms 2,3	**GNLKAHLKIHIADGPLKCRECGKQFTTSGNLKRHLRIHSGEKPYVCIHCQRQFADPGALQRHVRIHTG EKPCQCVMCGKAFTQASSLIAHVRQHTGEKPYVCERCGKRFVQSSQLANHIRHHDNIRPHKCSVCSK AFVNVGDLSKHIIIHTGEKPYLCDKCGRGFNRVDNLRSHVKTVHQGKAGIKILEPEEGSEVSVVTVDDM**
ID_66HI83_11D_cl56	Zinc finger and BTB domain-containing protein 17	**ZBTB17**	Q13105 ZBT17_HUMAN	canonical + isoforms 2,3	**HKCEDCGKEFTHTGNFKRHIRIHTGEKPFSCRECSKAFSDPAACKAHEKTHSPLKPYGCEECGKSYR LISLLNLHKKRHSGEARYRCEDCGKLFTTSGNLKRHQLVHSGEKPYQCDYCGRSFSDPTSKMRHLET HDTDKEHKCPH**
ID66IA57 cl 74	Zinc finger and BTB domain-containing protein 17	**ZBTB17**	Q13105 ZBT17_HUMAN	canonical + isoforms 2,3	**GNLKAHLKIHIADGPLKCRECGKQFTTSGNLKRHLRIHSGEKPYVCIHCQRQFADPGALQRHVRIHTG EKPCQCVMCGKAFTQASSLIAHVRQHTGEKPYVCERCGKRFVQSSQLANHIRHHDNIRPHKCSVCSK AFVNVGDLSKHIIIHTGEKPYLCDKCGRGFNRVDNLRSHVKTVHQGKAGIKILEPEEGSEVSVVTVDDM VTLATEALAATAVTQLTVVPVGAAVTADETEVLKAEISKAVKQVQEEDPNTHILYACDSCG**
ID66IA58 cl 524	Zinc finger and BTB domain-containing protein 17	**ZBTB17**	Q13105 ZBT17_HUMAN	canonical + isoforms 2,3	**GNLKAHLKIHIADGPLKCRECGKQFTTSGNLKRHLRIHSGEKPYVCIHCQRQFADPGALQRHVRIHTG EKPCQCVMCGKAFTQASSLIAHVRQHTGEKPYVCERCGKRFVQSSQLANHIRHHDNIRPHKCSVCSK AFVNVGDLSKHIIIHTGEKPYLCDKCGRGFNRVDNLRSHVKTVHQGKAGIKILEPEEGSEVSVVTVDDM VTLATEALAATAVTQLTVVPVGAAVTADETEVLKAEISKAVKQVQEEDP**
ID07HJ11 cl 78	Zinc finger and BTB domain-containing protein 17	**ZBTB17**	Q13105 ZBT17_HUMAN	canonical + isoforms 2,3	**GNLKAHLKIHIADGPLKCRECGKQFTTSGNLKRHLRIHSGEKPYVCIHCQRQFADPGALQRHVRIHTG EKPCQCVMCGKAFTQASSLIAHVRQHTGEKPYVCERCGKRFVQSSQLANHIRHHDNIRPHKCSVCSK AFVNVGDLSKHIIIHTGEKPYLCDKCGRGFNRVDNLRSHVKTVHQGKAGIKILEPEEGSEVSVVTVDDM VTLATEALAATAVTQLTVVPVGAAVTADETEVLKAEISKAVKQVQEEDP**
ID07IA15 cl 426	Zinc finger and BTB domain-containing protein 17	**ZBTB17**	Q13105 ZBT17_HUMAN	canonical + isoforms 3	**NLKAHLKIHIADGPLKCRECGKQFTTSGNLKRHLRIHSGEKPYVCIHCQRQFADPGALQRHVRIHTGE KPCQCVMCGKAFTQASSLIAHVRQHTGEKPYVCERCGKRFVQSSQLANHIRHHDNIRPHKCSVCSKA FVNVGDLSKHIIIHTGEKPYLCDKCGRGFNRVDNLRSHVKTVHQGKAGIKILEPEEGSEVSVVTVDDMV TLATEALAATAVTQLTVVPVGAAVTADETEVLKAEISKAVKQVQEEDPNTHILYACDSCGDKFLDANSL AQHVRIHTA**

ID07HJ81 cl 757	Zinc finger and BTB domain-containing protein 17	**ZBTB17**	Q13105 ZBT17_HUMAN	canonical + isoforms 3	**LKAHLKIHIADGPLKCRECGKQFTTSGNLKRHLRIHSGEKPYVCIHCQRQFADPGALQRHVRIHTGEK PCQCVMCGKAFTQASSLIAHVRQHTGEKPYVCERCGKRFVQSSQLANHIRHHDNIRPHKCSVCSKAF VNVGDLSKHIIIHTGEKPYLCDKCGRGFNRVDNLRSHVKTVHQGKAGIKILEPEEGSEVSVVTVDDMVT LATEALAATAVTQLTVVPVGAAVTADETEVLKAEISKAVKQVQEEDPNTHILYACDSCGDKFLDANSLA QHVRIHTAQALVMFQTDADFYQQYGPGGTWPAGQVLQAGELVFRPRDGAEGQPALAETSP**
ID07HJ80 cl 754	Zinc finger and BTB domain-containing protein 17	**ZBTB17**	Q13105 ZBT17_HUMAN	canonical + isoforms 3	**GNLKRHLRIHSGEKPYVCIHCQRQFADPGALQRHVRIHTGEKPCQCVMCGKAFTQASSLIAHVRQHT GEKPYVCERCGKRFVQSSQLANHIRHHDNIRPHKCSVCSKAFVNVGDLSKHIIIHTGEKPYLCDKCGR GFNRVDNLRSHVKTVHQGKAGIKILEPEEGSEVSVVTVDDMVTLATEALAATAVTQLTVVPVGAAVTAD ETEVLKAEISKAVKQVQEEDPNTHILYACDSCGDKFLDANSLAQHVRIHTAQALVMFQTDADFYQQY**
ID07HJ76 cl 270	Zinc finger and BTB domain-containing protein 17	**ZBTB17**	Q13105 ZBT17_HUMAN	canonical + isoforms 3	**LKAHLKIHIADGPLKCRECGKQFTTSGNLKRHLRIHSGEKPYVCIHCQRQFADPGALQRHVRIHTGEK PCQCVMCGKAFTQASSLIAHVRQHTGEKPYVCERCGKRFVQSSQLANHIRHHDNIRPHKCSVCSKAF VNVGDLSKHIIIHTGEKPYLCDKCGRGFNRVDNLRSHVKTVHQGKAGIKILEPEEGSEVSVVTVDDMVT LATEALAATAVTQLTVVPVGAAVTADETEVLKAEISKAVKQVQEEDPNTHILYACDSCGDKFLDANSLA QHVRIHTAQALVMFQTDADFYQQYGPGGTWPAGQVLQAGELVFRPRDGAEGQPALA**

ID66HJ20 Clone 141	Zinc finger protein 200	**ZNF200**	P98182 ZN200_HUMAN	canonical isoform + isoform 2, 3	**QEKERLNTSIPQKRKMRNLLVTIENDTPLEELSKYVDISIIALTRNRRTRRWYTCPLCGKQFNESSYLIS HQRTHTGEKPYDCNHCGKSFNHKTNLNKHERIHTGEKPYSCSQCGKNFRQNSHRSRHEGIHIREKIFK CPECGKTFPKNEEFVLHLQSHE**
ID_66HI68_12A_cl22	Zinc finger protein 668	**ZNF668**	Q96K58 ZN668_HUMAN	canonical + isoform 2	**QPADLAMHRRVHTGDRPFKCLQCDKTFVASWDLKRHALVHSGQRPFR**
ID07HJ89 cl 768	Zinc finger protein 827	**ZNF827**	Q17R98 ZN827_HUMAN	canonical + isoform 2	**EKPEKGFECVFCNFVCKTKNMFERHLQIHLITRMFECDVCHKFMKTPEQLLEHKKCHTVPTGGL**

Figure S10: List of pVHL30 interacting protein library-derived

Information about the 142 interacting clones encoding human proteins are reported in the table above. Details in the follow order: number of clone; name of the protein; name of the corresponding gene; Uniprot ID of the protein; isoforms; sequence of the interacting fragment

Supplementary Table 1: Primers list

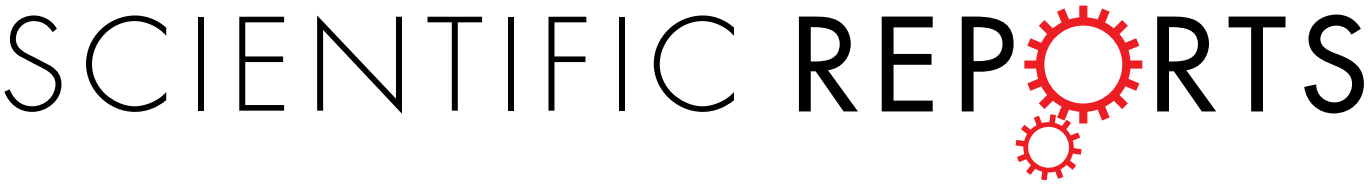
Plasmid Name	Primer Name	n°	sequence (5'-3')	Protein expressed
pGAD –pVHL30	VHL30 For	1	ggaggccagtgaattcATGCCCCGGAGGGCGGAGAA	Gal4AD-pVHL30 (1-213)
	VHL30Rev	2	cacccgggtggaattgTCAATCTCCCATCCGTTGAT	
pGAD –pVHL19	VHL19 For	3	ggaggccagtgaattcATGGAGGCCGGGCGGCCGCG	Gal4AD-pVHL19 (54-213)
	VHL30Rev	2		
pGAD –pVHL β	VHL19 For	3		Gal4AD-pVHLβ (54-157)
	VHL-β Rev	4	cacccgggtggaattgCTAAGTATACACTGGCAGTGTGATATTGGC	
pGAD –pVHL α	VHL-α For	5	ggaggccagtgaattcCTGAAAGAGCGATGCCTCC	Gal4AD-pVHLα (158-213)
	VHL-α Rev	6	cacccgggtggaattgTCAATCTCCCATCCGTTGATGTGC	
pGAD –pVHL-Nt	VHL30 For	1		Gal4AD-pVHL-Nt (1-53)
	VHL-Nt Rev	7	cacccgggtggaattgTCACTCCTCCTCGGCGCCCA	
pGAD –pVHl(1-157)	VHL30 For	1		Gal4AD-pVHL (1-157)
	VHL157Rev	8	cacccgggtggaattg TTAAGTATACACTGGCAGTGTGATATTGG	

pGBK -MDM2 (1-491)	MDM2 For	9	catggaggccgaattcATGGTGAGGAGCAGGCAAATGTGCAATACC	Gal4BD-MDM2 (1-491)
	MDM2 Rev	10	ggatccccgggaattgttaGGGGAAATAAGTTAGCACAATCATTTGAATTGG	
pGBK –MDM2 (1-150)	MDM2 For	9		Gal4BD-MDM2 (1-150)
	MDM2 (1-150) Rev	11	ggatccccgggaattgTTATGAAGATGAAGGTTTCTCTTCCTGAAGC	
pGBK –MDM2 (151-350)	MDM2 (151-350) For	12	catggaggccgaattcCATTTGGTTTCTAGACCATCTACCTC	Gal4BD-MDM2 (151-350)
	MDM2 (151-350) Rev	13	ggatccccgggaattgTTAGTTTTCCAGTTTGGCTTTCTCAGAGATTTCC	
pGBK–MDM2 (351-491)	MDM2 (351-491) For	14	catggaggccgaattcTCAACACAAGCTGAAGAGGGCTTTGATGTTCC	Gal4BD-MDM2(351-491)
	MDM2 Rev	10		
pGBK –MDM2 (434-491)	MDM2 (434-491) For	15	catggaggccgaattcGCCATTGAACCTTGTGTGATTTGTCAAGG	Gal4BD-MDM2 (434-491)
	MDM2 Rev	10		
pGBK –MDM2 (351-433)	MDM2 (351-491) For	14		Gal4BD-MDM2 (351-433)
	MDM2 (351-433) Rev	15	ggatccccgggaattgTTAATTAAGGGGCAAACTAGATTCCACACTC	

pGBK –MDM2 (351-451)	MDM2 (351-491) For	14		Gal4BD-MDM2 (351-451)
	MDM2 (351-451) Rev	16	ggatccccgggaattgTTAATGGACAATGCAACCATTTTTAGGTCG	
pGBK – MDM2 (329-433)	MDM2 (329-433) For	17	catggaggccgaattcTGGGCCCTTCGTGAGAATTGGCTTCC	Gal4BD-MDM2 (329-433)
	MDM2 (351-433) Rev	15		
pGBK – MDM2 (329-451)	MDM2 (329-433) For	17		Gal4BD-MDM2 (329-451)
	MDM2 (351-452) Rev	16		
pCDNA3.1 flag-MDM2			Purchased by Genescript cat n: OHu28568	Flag-MDM2 (1-491)
pCDNA3.1 flag-VHL30			Purchased by Genescript cat n: OHu23297	Flag-VHL30 (1-213)
pCDNA3.1 HA-VHL30	HA-VHL For	18	taccgagctcggatcATGGAGTACCCATACGACGTACCAGATTACG	HA-VHL30 (1-213)
	HA-VHL Rev	19	gatatctgcagaattTCAATCTCCCATCCGTTGATGTGCAATGCGC	
pCDNA3.1 HA-VHL19	HA-VHL For	18		

	HA-VHL Rev	19		HA-VHL19 (54-213)
pBRIDGE MCSI: pVHL30 MCSII: PHD3	pBRIDGE VHL30 For	20	tgtatcgccggaattcATGCCCCGGAGGGCGGAGAACTGGG	Myc-pVHL30 and PHD3 simultaneoulsy
	pBRIDGE VHL30 Rev	21	ggatccccgggaattTCAATCTCCCATCCGTTGAT	
pBRIDGE MCSI: Ø MCSII: PHD3	Received from Bex's Lab			
pBRIDGE MCSI:ODD MCSII: PHD3	pBRIDGE ODD For	22	tgtatcgccggaattcGACACAGATTTAGACTTGG	Myc-pVHL30 and ODD(HIF-1α) simultaneoulsy
	pBRIDGE ODD Rev	23	ggatccccgggaattAACGTAACTGGAAGTCATCATCC	
sequencing	Gal4 AD For	24	TACCACTACAATGGATG	----
sequencing	Gal4 AD Rev	25	GGTTTTTCAGTATCTACG	----
sequencing	T7	26	TAATACGACTCACTATAGGG	----

7- Appendix



OPEN

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

Received: 11 November 2016
Accepted: 17 March 2017
Published: 20 April 2017

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

Germline inactivation of the von Hippel-Lindau (VHL) tumor suppressor predisposes patients to develop different highly vascularized cancers. pVHL targets the hypoxia-inducible transcription factor (HIF-1α) for degradation, modulating the activation of various genes involved in hypoxia response. Hypoxia plays a relevant role in regulating cell cycle progression, inducing growth arrest in cells exposed to prolonged oxygen deprivation. However, the exact molecular details driving this transition are far from understood. Here, we present novel interactions between pVHL and the cyclin-dependent kinase inhibitor family CDKN1 (p21, p27 and p57). Bioinformatics analysis, yeast two-hybrid screening and co-immunoprecipitation assays were used to predict, dissect and validate the interactions. We found that the CDKN1 proteins share a conserved region mimicking the HIF-1α motif responsible for pVHL binding. Intriguingly, a p27 site-specific mutation associated to cancer is shown to modulate this novel interaction. Our findings suggest a new connection between the pathways regulating hypoxia and cell cycle progression.

Hypoxia is a common feature shared by the most active tumors, characterized by unregulated development and malignant progression[1]. The complex reactions forming the hypoxia response are mediated by the hypoxia-inducible factor HIF-1α, a transcription factor regulating numerous genes encoding proteins involved in the oxidative metabolism, energy production, cell cycle regulation, as well as red blood cell and vascular endothelial growth factor (VEGF) homeostasis[2–4]. At normal oxygen concentrations, the PHD (prolyl-4-hydroxylase domain) enzymes[5,6] catalyze hydroxylation of two specific HIF-1α prolines P402 and P564 in the N- and C-terminal oxygen-dependent domains (NODD and CODD). Hydroxylated HIF-1α is rapidly targeted for proteosomal degradation by the von Hippel-Lindau tumor suppressor protein (pVHL), an E3 ubiquitin ligase complex substrate recognition element[7]. Hypoxia inhibits PHD activity, impairing pVHL recognition and thereby promoting HIF-1α stabilization[8]. Once stabilized, HIF-1α is translocated to the nucleus, where it activates hypoxia response elements (HRE) promoting hypoxia-regulated gene expression[9]. Deregulation of this network is known to predispose to cancer onset, e.g. in von Hippel-Lindau syndrome, an inherited cancer syndrome characterized by the formation of tumors and cysts in different tissues[10]. Hypoxia is also important in regulating senescence[11]. Programmed cellular senescence is a physiological response evolved to limit the proliferation rate of normal mammalian cells[12]. A cell reaching the so-called Hayflick limit ceases or strongly reduces proliferation, while its metabolism is preserved. Under physiologic conditions, cellular senescence is regulated by many stimuli[13], including oncogene activity, telomere shortening, oxidative stress and DNA damage. Hypoxia-dependent cellular senescence is also thought to have a critical role in normal tumor suppressor response[11,14], modulating early malignant transformation[15] and drug-resistance[14]. In particular, hypoxic induction of cell cycle arrest is linked to HIF-1α dependent transcription of the cyclin-dependent kinase inhibitors p21 (CDKN1A)[16] and p27 (CDKN1B)[17]. Together with p57 (CDKN1C), these form a small family of kinase inhibitors playing important roles in negative regulation of the cell cycle[18]. It is well known that p21 mediates G1 growth arrest[19] and its transcription is mainly regulated by the tumor suppressor p53 in response to DNA damage[20]. HIF-1α dependent

[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. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to S.C.E.T. (email: silvio.tosatto@unipd.it)

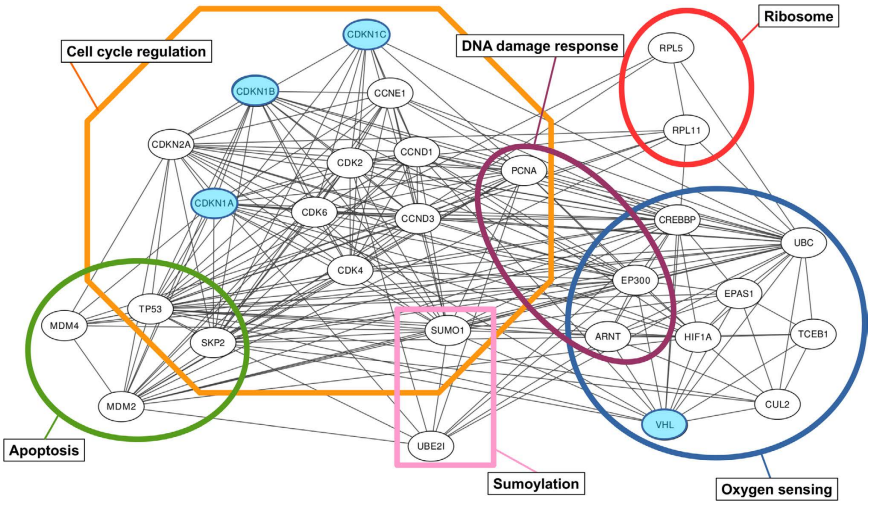


Figure 1. Prediction of pVHL/CDKN1 interactions. Protein-protein interaction network generated with STRING[26]. Functional connections between the pVHL and proteins involved in cell cycle regulation. Connections between nodes represent experimental evidence for interaction. Colored boxes group proteins participating in the same pathway or sharing similar function, e.g. ribosomal proteins. pVHL and CDKN1 are marked in blue.

transcription seems to regulate the activation of a specific genetic program designed to slow down the cell cycle in a p53-independent fashion, with marked progression into S phase during hypoxia rather than apoptosis[17]. 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[21]. However it is also thought that a ternary complex between p53, HIF-1α and Mdm2 may promote pVHL-independent degradation of HIF-1α and modulation of p53 activity[22]. Functional connections between hypoxia response and cell cycle regulation are reinforced by recent evidence linking pVHL and p14ARF[23,24], a modulator of Mdm2 function[25] arising from an alternative reading frame product of the CDKN2A locus encoding the p16ink4a1 tumor suppressor[26]. Here, we describe a novel interaction between pVHL and the CDKN1 inhibitor family. *In silico* sequence, structure and interaction analyses have been complemented with yeast two-hybrid and mammalian cell assays to define the molecular details driving this novel interaction. A conserved region shared among CDKN1 members was found to be responsible for the interaction, with at least one cancer-related mutations able to affect binding.

Results

***In silico* investigation of the hypoxia response and apoptotic pathways suggests that pVHL and the CDKN1 protein family could be physically linked.** At the pathway level, pVHL and p53 are thought to associate[27]. To shed light on the functional connection between hypoxia response and cell cycle regulation, an interaction network centered around the proteins pVHL, p53 and CDKN1s was generated with STRING[28] (Fig. 1). Two almost functionally overlapping pVHL isoforms are reported[29–31] (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 group. The network can be divided into six different clusters representing different biological processes (Fig. 1). 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[18]. Less obvious is their connection with the second largest cluster (9 nodes), which accounts for proteins involved in oxygen sensing and DNA damage response. Functional enrichment in GO[32] 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[33,34]. In general, the functional connection with RNA polymerase II subunits seems to have multiple functional roles. pVHL is thought to modulate Rpb1[33] expression and is necessary for the oxidative-stress-induced interaction of Rpb1 with DNA[33], as well as to suppress hsRPB7-induced VEGF promoter transactivation[34]. Considering network topology and connectivity, the data supports a strong functional link between oxygen sensing and cell cycle regulation. Further, pVHL was also proposed to mediate a HIF-1α-independent senescence program[35], 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[36], we decided to address the possible connection between pVHL and the CDKN1 family experimentally.

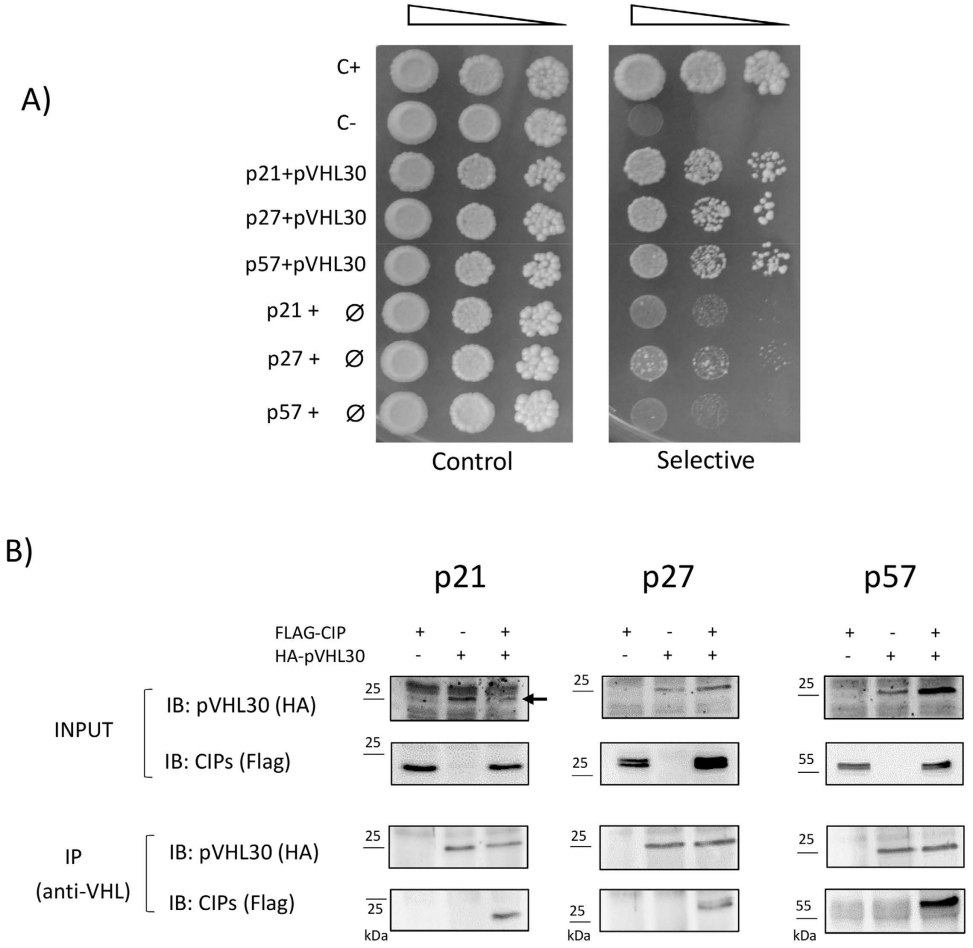


Figure 2. Validation of pVHL/CDKN1 interactions. (**A**) Yeast two hybrid (Y2H) assay of pVHL binding to the CDKN1 proteins. Serial dilutions of yeast cells were spotted on both permissive (*left*) and selective (*right*) media, and incubated for several days at 30 °C. The assayed interaction is shown on the left, with ∅ 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. (**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.

pVHL is able to interact with all CDKN1 proteins in Y2H and co-IP experiments. The protein-protein interaction between pVHL and members of the CDKN1 family (p21, p27 and p57) has been investigated by a genetic two-hybrid system in yeast cells (Y2H). As shown in Fig. 2A, yeast cells expressing either p21, p27, or p57 alone in selective medium almost failed to grow, as well as with pVHL30 only (Supplementary Figures S1–S3). Cell growth was markedly improved in presence of pVHL30, indicating that the proteins were able to associate *in vivo*. Our data further suggested a common pVHL interaction site with CDKN1 protein sequences. Co-immunoprecipitation experiments in human cells were performed to confirm the interactions in a more physiological context. A series of plasmids able to over-express either HA-tagged pVHL30 protein, or the Flag-tagged CDKN1 proteins in mammalian cells were constructed. Recombinant plasmids were used to transiently transfect HEK293T cells and perform Co-IP assays from total cell lysates using a specific anti-VHL antibody[37,38]. Our data (Fig. 2B) indicates that all three CDKN1 proteins were able to interact with pVHL30, as demonstrated by their presence in the immunoprecipitate revealed with the anti-Flag antibody (bottom panels). Taken together, the experiments show that these proteins can form at least binary complexes in human, and notably kidney cells.

A conserved motif in the CDKN1 domain resembles HIF-1α, and MD simulations suggest the pVHL β-domain to drive the interaction. The CDKN1 proteins share a conserved 48-residue domain located at the N-terminus identified in Pfam as the CDI protein family (pfam02234). Multiple sequence alignment showed relevant (64%) sequence similarity with the CODD[7] 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

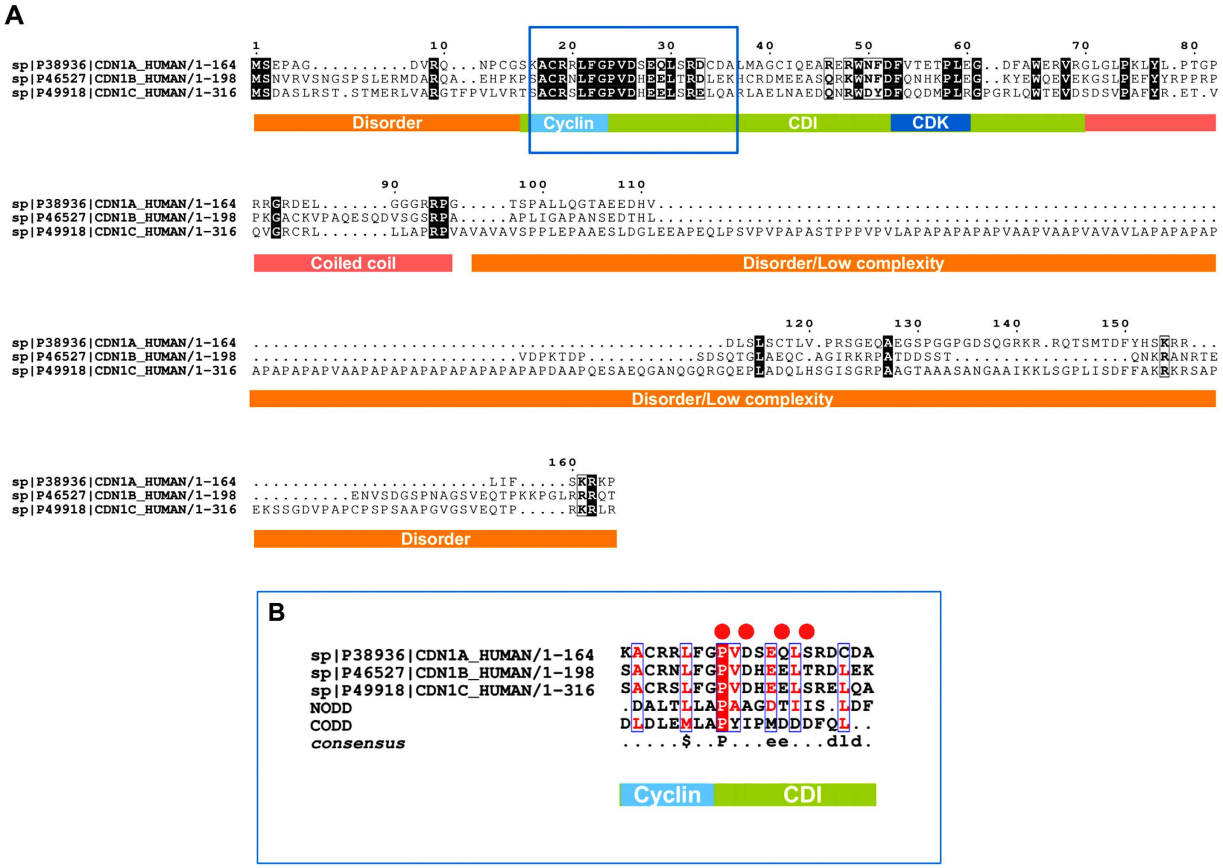


Figure 3. Overview of the CDKN1 sequence features. (**A**) Multiple sequence alignment and feature analysis of the CDKN1 protein family. The functional motif organization is presented as a colored bar below. (**B**) Close-up of sequence conservation between the CDI domain and NODD/CODD motifs of HIF-1α. 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.

amino acids 556–574 of HIF-1α (Fig. 3), a functional region responsible for the interaction with pVHL[7]. Data in the literature[39–41] as well as secondary structure prediction of CDI performed with FELLS[42] suggested this short segment to adopt random coil structure when not in complex to cyclin A. Combining these findings, we asked whether CDI may sustain the interaction with pVHL in a HIF-1α like fashion. Inspection of the p27 crystal structure (PDB identifier: 1JSU) showed the region partially overlapping the putative CODD motif, forming a short coil (residues 25–37) binding cyclin A. This finding reinforced the idea that this segment may adopt an extended conformation able to interact with pVHL under physiological conditions. Intriguingly, four cancer-related p27 mutations were found to affect the CODD-like region[43]. Whereas multiple experimental evidences describe the pVHL/CODD interaction[7,44], no structural data is available for the interactions investigated here. The corresponding p27 region (p27-CODD-like, residues 27–51) was modeled, using the HIF-1α CODD crystal structure (PDB code: 1LM8, chain H) as 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. The p27-NT backbone was predicted to assume a β-sheet-like conformation binding the fourth pVHL β-strand through 2 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[45] showed that p27-E40 is able to form interactions with other residues on the same interface of the pVHL β-domain (Supplementary Figure S4). 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.

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 (158–204). 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

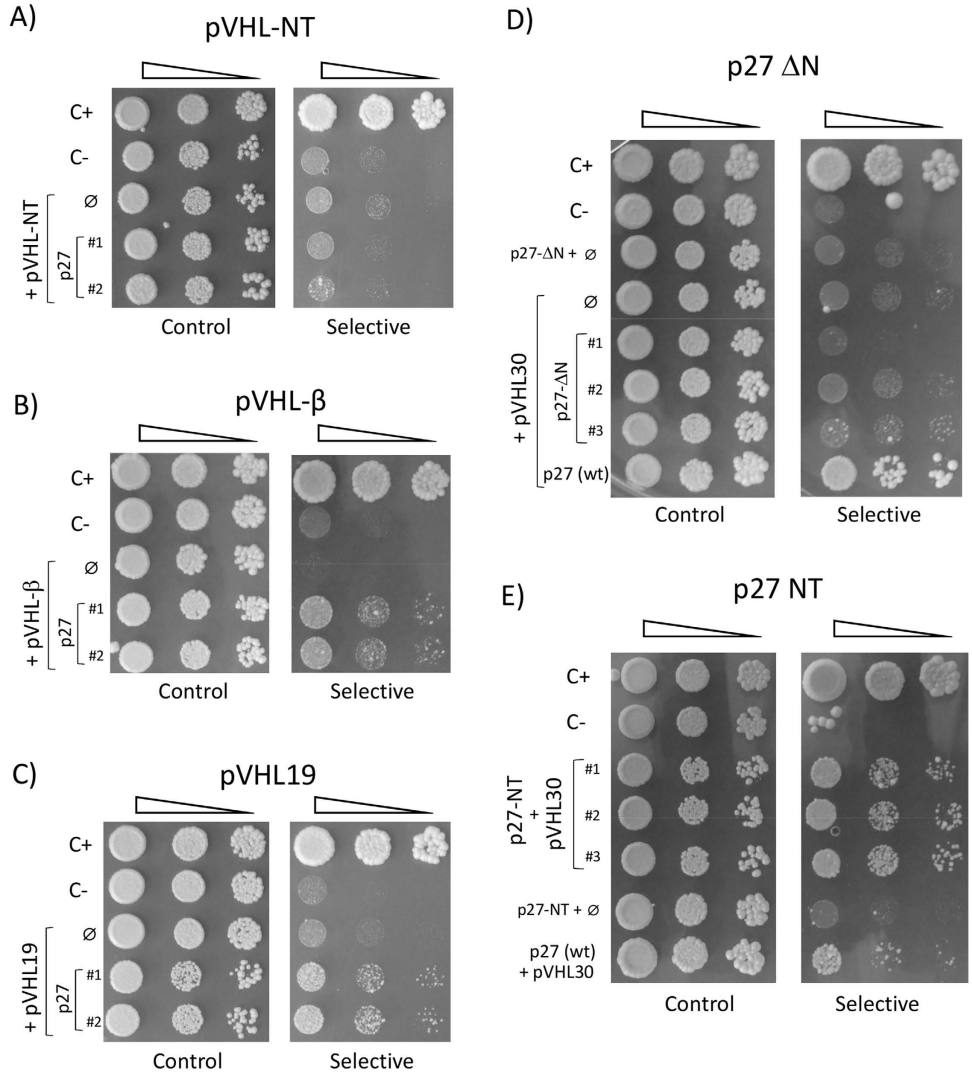


Figure 4. Y2H dissection map of pVHL binding to p27. Yeast two hybrid (Y2H) assays are shown of pVHL binding to p27, including fragments of either protein. Serial dilutions of yeast cells were spotted on both permissive (*left*) and selective (*right*) media, and incubated for several days at 30 °C. The assayed interaction is shown on the left, with ∅ 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. (**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. Plates were incubated at 30 °C for longer time (8 days), 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. Plates were incubated at 30 °C for 4 days.

with any CDKN1 protein are unable to grow in selective medium (Fig. 4A, Supplementary Figures S5 and S6). The pVHL β-domain is able to bind any member of the CDKN1 family, as indicated by yeast cell growth (Fig. 4B, Supplementary Figures S5 and S7). The results also confirmed that the pVHL19 isoform, lacking the N-terminus is able to interact with all CDKN1 proteins like the pVHL30 protein (Fig. 4C, Supplementary Figures S5 and S8). 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.

The CDKN1 region involved in pVHL binding was mapped based on the conservation shared by the three CDKN1 sequences. Yeast plasmids expressing either their N-terminal tail containing the CDI domain (p27-NT residues: 1–60; p21-NT: 1–49; p57-NT: 1–61) or the corresponding C-terminal moiety lacking the N-terminus (p27-ΔN residues: 61–198; p21-ΔN: 50–164; p57-ΔN: 62–316) were generated to test the effects of CDKN1 binding to pVHL30. As shown in Fig. 4D, loss of the p27 N-terminus clearly 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

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 1. *In silico* prediction of variations between wild-type and mutant pVHL/p27. Free binding energy (Δ_{bind}G, electrostatic component) is calculated with APBS[48], while stability prediction and electrostatic solvation free energy of mutations affecting p27 (CDKN1b) were calculated with NeEMO[46] and Bluues[47].

data have been also obtained for both p21 and p57 (Supplementary Figures S9 and S10), strongly supporting the notion that the CDI domain is responsible for CDKN1 binding to pVHL30.

The final demonstration came by directly assaying the interaction between pVHL30 and the N-terminus of each CDKN1 protein. As shown in Fig. 4E (and Supplementary Figure S11), the N-terminus of p27 was able to sustain the growth of pVHL30-expressing yeast cells even more efficiently than full-length p27. Similar results have been also observed for the N-terminus of p57 (Supplementary Figures S5 and S11). Although data on the p21-NT fragment cannot be considered due to unspecific activation of the reporter gene (i.e. auto-activation, Supplementary Figure S9) our 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.

Pathological p27 mutations may influence pVHL binding. The predicted binding site between pVHL and the CODD-like p27 motif, shared by all CDKN1 proteins, has been further tested by inserting several N-terminal p27 missense mutations. The substitutions P35L, D37N, E40K and T42A (Fig. 3B) have been selected considering the conservation of the CODD motif, their position on the interaction surface as well as their pathological relevance in cancer[43], i.e. hematological malignancies compatible with the deregulation of the pVHL/HIF-1α axis. None of these had been characterized with respect to the normal function of p27 as CDK inhibitor. *In silico* replacements of both D37N and T42A were predicted by NeEMO[46] to stabilize (or have a negligible impact) on the pVHL/p27 binding moiety (Table 1), while the contrary is predicted for E40K and partially for T42A. 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, we characterized the electrostatic properties of the putative pVHL/p27 complex wild-type in respect to the complexes formed with mutants p27 using Bluues[47] (Fig. 5). 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). Aspartic acid 27 is exposed to solvent and a change to asparagine is predicted to stabilize the complex little. Conversely, E40K is predicted to markedly destabilize the interaction. During MD simulations, we observed the formation of a salt bridge between p27 E40 and pVHL R107. A change in lysine clearly abolishes this interaction, suggesting a repulsive effect. We predict a difference in electrostatic solvation energy of 183.33 kJ/mol for this mutant, a value high enough to destabilize the complex. A slightly positive variation (3.58 kJ/mol) is predicted instead for T42A. As T42 is located in a partially hydrophobic pocket at the pVHL/p27 interface, we believe it should have a modest impact on complex formation. APBS[48] was used to quantify the binding free energy electrostatic component, required for pVHL/p27 complex formation (Table 1). The binding free energy predicted for the wild-type complex is Δ_{bind}G −22.07 kcal/mol, a favorable value for complex formation. Similar values are predicted for mutants D37N and T42A (Δ_{bind}G −22.56 and −20.65 kcal/mol, respectively) and Δ_{bind}G −28.05 kcal/mol is predicted for P35L. Instead, a positive value is predicted for E40K (Δ_{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.

The predictions have been then experimentally tested by Y2H assay (Fig. 6). Both p27 T42A and D37N mutants retained the ability to associate pVHL30 almost completely, as indicated by similar yeast cell growth in selective medium. The data further confirm that P35L does not perturb binding, as its introduction did not affect the growth rate of mutant yeast cells. On the contrary, E40K reduced p27 binding to pVHL30, as yeast cells expressing the p27 almost failed to grow in selective medium. The negative effect of the mutation has been confirmed by testing additional clones (Supplementary Figures S12 and S13). Collectively, our results on p27 mutants further support that p27/pVHL binding may be described by the *in silico* model.

Discussion

The starting point for our 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, involved in its binding with the pVHL protein. We 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. Our data point to 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[49] which is thought to drive cell cycle entry at the G1/S transition by decreasing p27

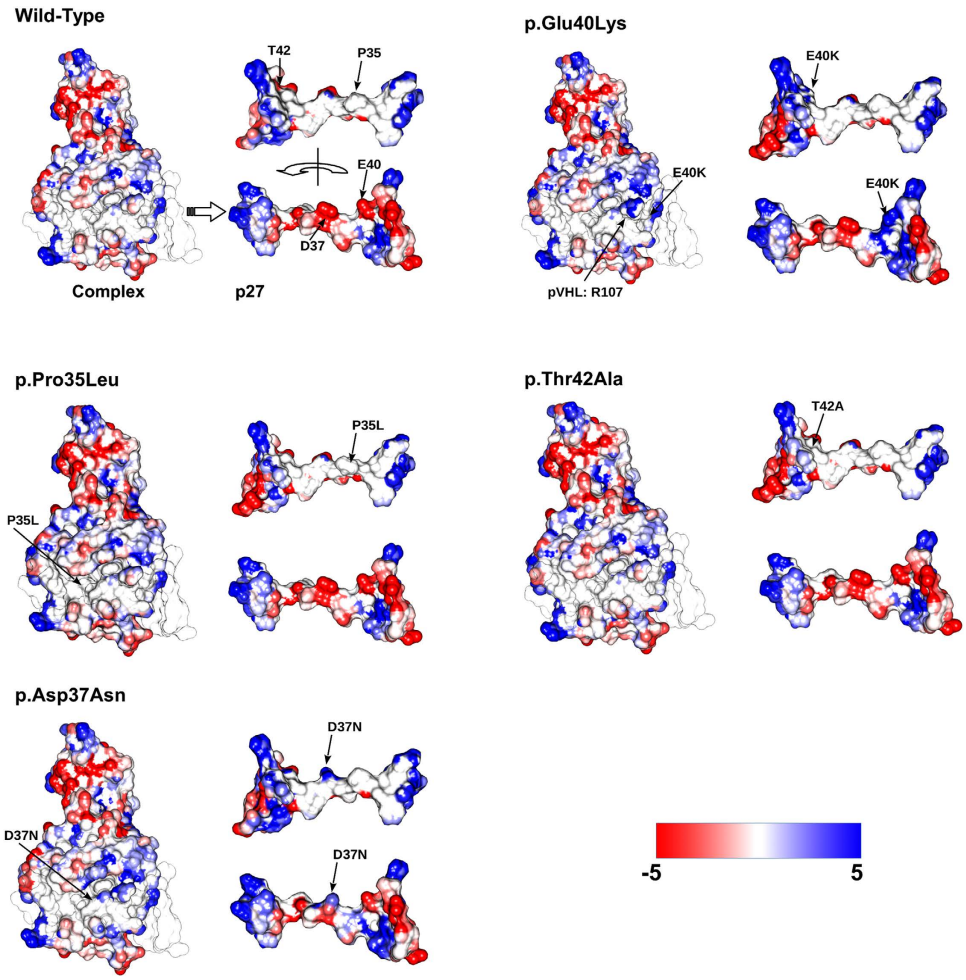


Figure 5. ***In silico*** **prediction of mutations putatively affecting pVHL/p27 binding.** Solvent accessible surface representation of a predicted pVHL/p27-NT complex colored by electrostatic potential. Blue represents positively charged areas, red negative. For each mutation, the full complex is 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[47].

stability. In this scenario, interaction with pVHL could serve as further modulator of p27 half-life. Our data highlights the common features between the interactions of the pVHL-β domain with both the CDKN1-CDI and HIF1-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[43] 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 potentially able to exert their 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[50]. 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[33,34,51] in a CODD-like fashion, however all these interactions require 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[5,6], reducing competition for interface B. Very recently, a novel hydroxylation-independent interaction between pVHL and Aurora kinase A (AURKA)[52], 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,

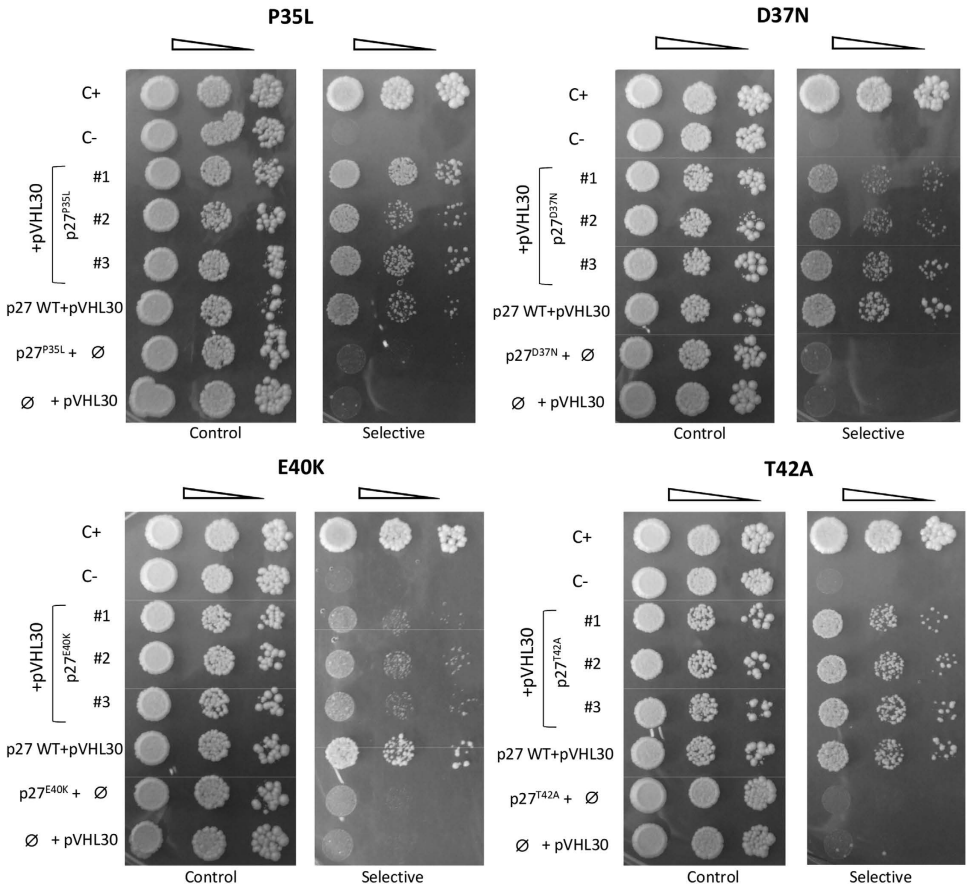


Figure 6. Effects p27 mutagenesis on pVHL binding. *Y2H analysis of the p27 mutations P35L, D37N, E40K and T42A*. Yeast cells co-expressing pVHL30 together with the indicated mutant p27 isoforms were assayed in Y2H. Three independent mutant clones are shown. Yeast cells have been grown in selective medium (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.

the same pVHL region[53,54] mediates its association with over 40 proteins beyond HIF-1/2α[54], 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[55–58].

Although preliminary correlations between hypoxia response and the CDKN1 proteins have been reported[49,59,60], 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[61]. 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[62]. Increased p27 levels have been observed in several pVHL-mutant cell lines[63], further pointing to their functional association. Future work is needed to address if (and how) binding with pVHL/VCB could impact both 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. Our results on site-specific mutations, 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.

Materials and Methods

Interaction network and sequence feature analysis. Amino acid sequences (with UniProt accession numbers in parentheses) for pVHL (P40337), p21 (P38936), p27 (P46527) and p57 (P49918) were retrieved

from UniProt[64] selecting the canonical sequence and visualized with Jalview[65]. Alignment was performed with T-Coffee[66] using default parameters. Disorder was assessed with MobiDB[67] and DisProt[41], while functional domains were retrieved from Pfam[68] and InterPro[69]. A protein-protein interaction network centered around pVHL and CDKNs was derived from STRING[28]. To maximize data reliability, text-mining and neighborhood interactions were excluded. The first shell of interactors was populated with pVHL, 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 confidence parameter (0.400) was used and the resulting network analyzed with Cytoscape[70].

Molecular dynamics simulations. The 1.8 Å crystal structure of pVHL (PDB identifier: 1LM8)[7] and the 2.3 Å structure of p27 (PDB identifier: 1JSU)[71] were used as starting models. The pVHL/p27 complex was constructed by homology modeling using Modeller[72] through superimposition to the pVHL/HIF-1α complex. All simulations were carried out with GROMACS[73] 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[74] using a 2 fs time step. Trajectories were compared in terms of RMSD and root-mean-square fluctuation (RMSF). RING 2.0[45] was used to estimate variation in residue interaction network with strict distance thresholds and the "one interaction" option.

Selection and interpretation of mutations putatively affection pVHL/p27 interaction. Considering both sequence conservation with CODD motif and their specific association with cancer, the following cancer related p27 mutations putatively affecting the pVHL/p27 interaction were retrieved from COSMIC[43]: P35L, D37N, E40K and T42A. The first three localize in or are immediately close to conserved positions, while p.T42A has been chosen because it was found in patients with hematological malignancies[75]. Both driver and passenger mutations affecting the CODD-like region were included in this study due to the novelty of the pVHL/p27 interaction and a lack of specific bibliographic data. Mutations were placed with Bluues[47] 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 was predicted for each mutant with NeEMO[46]. Mutations on the pVHL/p27 interaction were investigated predicting the electrostatic $\Delta_{bind}G$ values with APBS[48], a method which evaluates the energy difference between the solvated unbound interactors and the solvated complex.

Plasmid constructs. The pcDNA3.1-derived plasmids carrying synthetic full-length cDNA sequences coding for the human pVHL30, p21, p27 and p57 proteins were purchased from GenScript (GenEZ plasmids OHu23297, OHu27895, OHu26670 and OHu27234 respectively). Plasmids were used as starting point to transfer the cDNA sequences (full-length or corresponding to the different protein regions) in the pGADT7 and pGBKT7 plasmids (Clontech), used to perform yeast two-hybrid assays. Untagged cDNA sequences were amplified by PCR using specific primers (Supplementary Table S1) carrying 15 nucleotides long 5′ ends corresponding to specific regions surrounding the EcoRI site in the MCS of pGADT7 and pGBKT7 vectors. PCR products were directly cloned in two vectors linearized by digestion with EcoRI using the In-Fusion® HD Cloning Kit (Clontech) following the manufacturer protocols. All recombinant plasmids used in the yeast two-hybrid assays (Supplementary Table S1) are able to express the proteins of interest in fusion with either the DNA binding domain (DBD) or the activation domain (AD) of the Gal4 transcription factor. The cMyc and HA epitopes present in the fusion proteins with the Gal4 DBD and AD, respectively, allow to readily verify the expression of the chimeric proteins in yeast cells. GenScript plasmids have been directly used to overexpress the different CDKN1 proteins as N-terminal FLAG-tagged fusion polypeptides in mammalian cells. The 5′-end of each cDNA was joined in-frame with the sequence coding for the FLAG epitope. The In-Fusion® HD Cloning Kit (Clontech) was used to clone the pVHL30 cDNA in the BamHI/EcoRI sites of pcDNA3.1 vector, with the addition of the HA epitope sequence at the N-terminus. The resulting pVHL30-pcDNA3.1 plasmid was used in mammalian cells to overexpress pVHL30 as N-terminal HA-tagged protein. Yeast plasmids expressing single-residue mutagenized p27 (P35L, D37N, E40K and T42A) were obtained using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies), starting from the corresponding wild-type plasmid in a single-round PCR step, following the manufacturer procedures. All sequences cloned in the recombinant plasmids described in this paper have been verified by Sanger sequencing prior to use in the experiments.

Yeast two-hybrid (Y2H) assays. Interactions between the three CDKN1 family members and pVHL have been investigated by the two-hybrid assay, using the Matchmaker® Gold Two Hybrid System (Clontech) using standard Yeast growth conditions, media and transformation protocols. Experimental conditions to evaluate the positivity of the assay have been first established by following the growth of the Y190 reporter strain in selective medium lacking histidine supplemented with 30 mM 3-AT. In the assays, the same negative and positive internal controls have been always added, represented respectively by the Y190 yeast strain carrying both empty pGBKT7 and pGADT7 plasmids (no binding) and Y190 cells co-transformed with the pGADT7-T (Gal4 AD-SV40 large T-antigen) and pGBKT7–53 (Gal4 DBD-murine p53, fragment 72–390) plasmids (strong binding) provided by the manufacturer. Yeast Y190 cells co-expressing the fusion protein with the Gal4 DBD and AD domains alone, or *vice-versa*, have also been tested to exclude false positivity of the assay due to the auto-activation by Gal4-fusion proteins. Expression in yeast cells of the Gal4-fusion proteins has been checked by Western blot analysis (Supplementary Figure S15), while the functional status of pVHL was verified testing the well-known pVHL/HIF-1α interaction with and without co-expressing PHD3 (Supplementary Figure S16). Multiple (3–5) transformations were generally performed, where (at least) two independent colonies for each experiment were picked, inoculated in liquid medium, and grown to exponential phase. Yeast cells were then serially diluted 10-fold, and spotted on either solid selective medium lacking histidine and containing 30 mM 3-AT, or permissive medium

(i.e., supplemented by histidine), to check both cell viability and number. Growth of yeast strains was constantly monitored for 3 to 8 days at 30 °C. When necessary, interactions have been further tested in selective medium supplemented with 60 mM 3-AT, to increase the stringency of the assay by doubling the concentration of the competitive inhibitor. Total yeast proteins have been obtained by TCA-based solubilization of yeast cells, as described in ref. 76, followed by standard Western Blot analysis[77], by using either anti-HA (Abcam, ab16918), or anti-Myc (Abcam, ab127421) antibodies, to reveal the HA-tagged Gal4AD-pVHL, or the Myc-tagged Gal4BD-CDKN1 proteins, respectively.

Transfection and co-immunoprecipitation from HEK293T cells. Human kidney HEK293T cells were used for co-immunoprecipitation experiments. Cells grown at confluence 70–80% in a 1 ml well microplate were plated for transfection using the Lipofectamine 2000 DNA transfection protocol (Invitrogen). For each transfection, approximatively 5 μg of total DNA, i.e. pcDNA3.1-derived plasmids (empty, and/or expressing either HA-pVHL30 or FLAG-CDKN1) were used. After 24 hours, transfected cells were re-suspended in lysis buffer (20 mM HEPES-Na pH7.4, 150 mM NaCl, 5 mM CHAPS) supplemented with 1X PIC (Protease Inhibitors Cocktail, Sigma). Cell lysates have been centrifuged (10′ at 600 rpm, 4 °C), and the resulting PNS (post-nuclear supernatant) quantified by Bradford assay. For co-immunopreCDKN1itations, 5 μl of protein A magnetic beads (Pierce Thermoscientific) were pre-incubated (1 hour at RT) with 2 μg of anti-VHL antibody (Santa Cruz, sc-5575). About 0.2 mg of PNS were added and incubated for 4 hours at 4 °C. Beads were finally washed 3 times using lysis buffer and eluted by incubating the beads 5 min at 70 °C in 25 μl in 1X NuPAGE LDS sample buffer (Invitrogen) supplemented with 0.1 M DTT. Both PNS and immunoprecipitated samples were subjected to standard SDS-PAGE and Western blot. Membranes were probed with either anti-HA (Abcam, ab16918), or anti-FLAG (Abcam, ab45766) antibodies, to reveal the HA-tagged pVHL30 or FLAG-CDKN1 proteins, respectively.

References

1. Wilson, W. R. & Hay, M. P. Targeting hypoxia in cancer therapy. *Nature reviews. Cancer* **11,** 393–410 (2011).
2. Bárdos, J. I. & Ashcroft, M. Negative and positive regulation of HIF-1: a complex network. *Biochimica et biophysica acta* **1755,** 107–20 (2005).
3. Kim, J., Tchernyshyov, I., Semenza, G. L. & Dang, C. V. HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia. *Cell Metab.* **3,** 177–185 (2006).
4. Pugh, C. W. & Ratcliffe, P. J. Regulation of angiogenesis by hypoxia: role of the HIF system. *Nat. Med.* **9,** 677–684 (2003).
5. Jaakkola, P. *et al.* Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science* (*New York, N.Y.*) **292,** 468–72 (2001).
6. Ivan, M. *et al.* HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science* **292,** 464–468 (2001).
7. Hon, W.-C. *et al.* Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. *Nature* **417,** 975–8 (2002).
8. McDonough, M. A. *et al.* Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc. Natl. Acad. Sci. USA* **103,** 9814–9819 (2006).
9. Maxwell, P. H. *et al.* The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* **399,** 271–275 (1999).
10. 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).
11. Welford, S. M. & Giaccia, A. J. Hypoxia and Senescence: The impact of oxygenation on tumor suppression. *Mol Cancer Res* **9,** 538–544 (2011).
12. Hayflick, L. & Moorhead, P. S. The serial cultivation of human diploid cell strains. *Experimental Cell Research* **25,** 585–621 (1961).
13. Ben-Porath, I. & Weinberg, R. A. The signals and pathways activating cellular senescence. *Int. J. Biochem. Cell Biol.* **37,** 961–976 (2005).
14. Sullivan, R., Paré, 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).
15. Ohtani, N., Mann, D. J. & Hara, E. Cellular senescence: its role in tumor suppression and aging. *Cancer Sci.* **100,** 792–797 (2009).
16. Koshiji, M. *et al.* HIF-1alpha induces cell cycle arrest by functionally counteracting Myc. *EMBO J.* **23,** 1949–1956 (2004).
17. Goda, N. *et al.* Hypoxia-inducible factor 1alpha is essential for cell cycle arrest during hypoxia. *Mol. Cell. Biol.* **23,** 359–369 (2003).
18. Nakayama, K. & Nakayama, K. Cip/Kip cyclin-dependent kinase inhibitors: brakes of the cell cycle engine during development. *Bioessays* **20,** 1020–1029 (1998).
19. Brugarolas, J. *et al.* Radiation-induced cell cycle arrest compromised by p21 deficiency. *Nature* **377,** 552–557 (1995).
20. 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).
21. Ravi, R. *et al.* Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1α. *Genes Dev* **14,** 34–44 (2000).
22. Chen, D., Li, M., Luo, J. & Gu, W. Direct Interactions between HIF-1α and Mdm2 Modulate p53 Function. *J. Biol. Chem.* **278,** 13595–13598 (2003).
23. 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).
24. Minervini, G. *et al.* Isoform-specific interactions of the von Hippel-Lindau tumor suppressor protein. *Sci Rep* **5,** 12605 (2015).
25. 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).
26. 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).
27. Roe, J.-S. *et al.* p53 Stabilization and Transactivation by a von Hippel-Lindau Protein. *Molecular Cell* **22,** 395–405 (2006).
28. Franceschini, A. *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41,** D808–815 (2013).
29. 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. USA* **95,** 11661–11666 (1998).
30. Schraml, P. *et al.* Relevance of Nuclear and Cytoplasmic von Hippel Lindau Protein Expression for Renal Carcinoma Progression. *Am J Pathol* **163,** 1013–1020 (2003).
31. Ratcliffe, P. J. New insights into an enigmatic tumour suppressor. *Nat Cell Biol* **5,** 7–8 (2003).

32. The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Research*, doi: 10.1093/nar/gkw1108 (2016).
33. 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,** 2701–2717 (2008).
34. 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).
35. 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).
36. 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,** 710 (2014).
37. 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).
38. 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).
39. 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).
40. Yoon, M.-K., Mitrea, D. M., Ou, L. & Kriwacki, R. W. Cell cycle regulation by the intrinsically disordered proteins p21 and p27. *Biochemical Society Transactions* **40,** 981–988 (2012).
41. Piovesan, D. *et al.* DisProt 7.0: a major update of the database of disordered proteins. *Nucleic Acids Res.* **45,** D1123–D1124 (2017).
42. Piovesan, D., Minervini, G. & Tosatto, S. C. E. FELLS: Fast Estimator of Latent Local Structure. *Submitted* http://protein.bio.unipd.it/fells/ (2016).
43. Forbes, S. A. *et al.* COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **43,** D805–811 (2015).
44. Min, J.-H. *et al.* Structure of an HIF-1alpha -pVHL complex: hydroxyproline recognition in signaling. *Science* **296,** 1886–1889 (2002).
45. Piovesan, D., Minervini, G. & Tosatto, S. C. E. The RING 2.0 web server for high quality residue interaction networks. *Nucleic Acids Res.*, doi: 10.1093/nar/gkw315 (2016).
46. 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** Suppl 4, S7 (2014).
47. Walsh, I. *et al.* Bluues server: electrostatic properties of wild-type and mutated protein structures. *Bioinformatics* **28,** 2189–2190 (2012).
48. 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. USA* **98,** 10037–10041 (2001).
49. 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* **14,** 143 (2015).
50. Gordan, J. D., Bertout, J. A., Hu, C.-J., Diehl, J. A. & Simon, M. C. HIF-2alpha promotes hypoxic cell proliferation by enhancing c-myc transcriptional activity. *Cancer Cell* **11,** 335–347 (2007).
51. Guo, J. *et al.* pVHL suppresses kinase activity of Akt in a proline-hydroxylation-dependent manner. *Science* **353,** 929–932 (2016).
52. Hasanov, E. *et al.* Ubiquitination and regulation of AURKA identifies a hypoxia-independent E3 ligase activity of VHL. *Oncogene*, doi: 10.1038/onc.2016.495 (2017).
53. Leonardi, E., Murgia, a. & Tosatto, S. C. E. Adding structural information to the von Hippel-Lindau (VHL) tumor suppressor interaction network. *FEBS letters* **583,** 3704–10 (2009).
54. Tabaro, F. *et al.* VHLdb: A database of von Hippel-Lindau protein interactors and mutations. *Sci Rep* **6,** 31128 (2016).
55. 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).
56. 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).
57. Child, E. S. & Mann, D. J. The intricacies of p21 phosphorylation: protein/protein interactions, subcellular localization and stability. *Cell Cycle* **5,** 1313–1319 (2006).
58. Chu, I. *et al.* p27 phosphorylation by Src regulates inhibition of Cyclin E-Cdk2 and p27 proteolysis. *Cell* **128,** 281–294 (2007).
59. 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).
60. 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,** 4565–4578 (2005).
61. Starostina, N. G. & Kipreos, E. T. Multiple degradation pathways regulate versatile CIP/KIP CDK inhibitors. *Trends Cell Biol.* **22,** 33–41 (2012).
62. 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).
63. Sgambato, A. *et al.* Deregulated expression of p27(Kip1) in human breast cancers. *Clin Cancer Res* **3,** 1879–1887 (1997).
64. UniProt: a hub for protein information. *Nucleic Acids Res* **43,** D204–D212 (2015).
65. 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).
66. Notredame, C., Higgins, D. G. & Heringa, J. T-coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* **302,** 205–217 (2000).
67. Potenza, E., Di Domenico, T., Walsh, I. & Tosatto, S. C. E. MobiDB 2.0: an improved database of intrinsically disordered and mobile proteins. *Nucleic Acids Res.* **43,** D315–320 (2015).
68. Punta, M. *et al.* The Pfam protein families database. *Nucleic Acids Research* **40,** D290–D301 (2012).
69. Finn, R. D. *et al.* InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Res.* **45,** D190–D199 (2017).
70. Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **13,** 2498–2504 (2003).
71. Russo, A. A., Jeffrey, P. D., Patten, A. K., Massagué, J. & Pavletich, N. P. Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. *Nature* **382,** 325–331 (1996).
72. Eswar, N. *et al.* Comparative protein structure modeling using MODELLER. *Curr Protoc Protein Sci* Chapter 2, Unit 2.9 (2007).
73. Hess, B., Kutzner, C., Biophysics, C., Spoel, D. V. D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable *Molecular Simulation.* 435–447 (2008).
74. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J Comput Chem* **26,** 1781–1802 (2005).
75. Papaemmanuil, E. *et al.* Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122,** 3616–3627; quiz 3699 (2013).
76. Wright, A. P., Bruns, M. & Hartley, B. S. Extraction and rapid inactivation of proteins from Saccharomyces cerevisiae by trichloroacetic acid precipitation. *Yeast* **5,** 51–53 (1989).
77. Sambrook, J. & Russell, D. *Molecular cloning: a laboratory manual* (Cold Spring Harbor Laboratory Press, 2001).

Acknowledgements
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 S.T. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions
S.T., G.M. and R.L. conceived the experiments. G.M.,. R.B. and A.F. performed the experiments. G.M., R.L. and G.S. analyzed the data. G.M., R.L. and S.T. wrote the manuscript.

Additional Information
Supplementary information accompanies this paper at http://www.nature.com/srep

Competing Interests: The authors declare that have no significant competing financial, professional or personal interests that might have influenced the performance or presentation of the work described in this manuscript.

How to cite this article: Minervini, G. *et al.* Novel interactions of the von Hippel-Lindau (pVHL) tumor suppressor with the CDKN1 family of cell cycle inhibitors. *Sci. Rep.* **7**, 46562; doi: 10.1038/srep46562 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.