

An investigation of the role of ASPP2 in HIF signaling and cancer



Hok Fung Chan
Balliol College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Hilary 2020

Declaration

This thesis is submitted to the University of Oxford in support of my application for the degree of Doctor of Philosophy in Clinical Medicine. The work described herein was performed while I was a graduate student at the Ludwig Institute for Cancer Research, under the supervision of Professor Xin Lu and Professor Sir Peter Ratcliffe.

All the work presented in this thesis is the result of my own work unless stated here and in the corresponding sections and has not been submitted in any previous application for any degree. Interactomic study of ASPP2 by IP-MS was performed jointly with former lab graduate student Dr. Thomas Leissing (University of Oxford). In the phosphoproteomic study of ASPP2-regulated phosphosites, phosphopeptide enrichment and liquid chromatography electrospray ionization tandem mass spectrometric measurements were performed by Dr. Kundan Sharma, a former lab post-doc and phosphoproteomic data analysis was jointly performed with Dr. Sharma. The results from pan-cancer gene co-expression analysis were based upon data generated by the TCGA Research Network: <https://www.cancer.gov/tcga>. ImageJ macro for quantification of nuclear/non-nuclear signals was developed by Dr. Mark Shipman (former microscope manager, Ludwig Oxford).

This thesis does not exceed 50,000 words, exclusive of bibliography, as required by the Nuffield Department of Medicine.

Acknowledgements

I am very grateful to many people who have been involved in various ways during my study here in Ludwig.

First and foremost, I would like to thank my supervisors Professor Xin Lu and Professor Sir Peter Ratcliffe for giving me the opportunity to work on this exciting and challenging project and their guidance throughout my DPhil study. I would also like to express my gratitude to Nuffield Department of Medicine, Ludwig Institute for Cancer Research, and Balliol College for their financial and administrative support, as well as my examiners Professor Margaret Ashcroft and Professor Francesca Buffa, my transfer and confirmation examiners Professor David Mole, Professor Chris Schofield, and Professor Benedict Kessler for their time and effort.

I have been very lucky to have a great bunch of people in the Lu lab and Ratcliffe lab even since I started my DPhil. I would like to especially thank Xiao Qin, Felix Zhou, Joseph Kaplinsky, Kundan Sharma, Ludo Buti, Stefanie Zentelis, Carlos Ruiz Puig, Dennis Sangberg, Thomas Carroll, Artem Smirnov, Khatoun Al Moussawi, Jie Yin, Thomas Leissing, Imogen Roth, Jingyi Ma, Veronique Lafleur, Olivia Lombardi, Norma Masson, and Ya-Min Tian. They have been great mentors and friends who supported and inspired me at various stages during my DPhil.

I am very grateful to have the opportunity to get to know and share my time here with many enthusiastic and inspiring colleagues in the Ludwig and I have learned and benefitted enormously through collaborating and discussing with them. In particular I would like to acknowledge Pakavarin Louphrasitthiphol, Carol Leung, Xiao Qin, Kundan Sharma, Felix Zhou, Joseph Kaplinsky, Biao Ma, Thomas Leissing, Stefanie Zentelis, Norbert Volkmar, Jie Yin, and Matthew Cockman, for their support, advice, and stimulating conversations in/outside of research.

I would also like to thank the colleagues who have supported me in different ways throughout my study. In particular, Felix Zhou and Joseph Kaplinsky as my bioinformatics mentors giving invaluable guidance and feedbacks in the construction and improvement of the bioinformatic pipeline and also patiently teaching me some relevant concepts; Peter Chan from the Big Data Institute and George Qian from for their constructive feedbacks on bioinformatics; Mark Shipman and Richard Lisle for their invaluable support in microscopy imaging and screening; Veronique Lafleur who patiently mentored me to learn the chromatin immunoprecipitation protocol,

as well as troubleshooting and scientific exchange on HIF together with Olivia Lombardi and Norma Masson; Mary Muers, Professor Sarah de Val, and Xiao Qin for sparing the time to read my thesis and give me invaluable feedbacks; as well as our fantastic support and administration staff: Veronica Counsell, Alexandra Ward, Jo Wainwright, Stan Ng, Jan Twitchen, James Hall, Carly Morris, and Stephen Fischer.

Importantly I would like to thank my parents, grandparents, and in-laws for giving me the freedom to do what I am passionate about, and for their unconditional love, support, and encouragement, especially my mother who believed in me even when I was a little kid and encouraged me to dream and pursue something meaningful in life.

Last but not least, I am forever indebted to my wife, my best friend, and my soul mate, Shuyan Lin, who stayed by my side throughout this challenging, and at times frustrating journey with unwavering love and support. I am also thankful to Tang Tang, Pa Pa, and Pom Pom, our three guinea pigs, for their innocence and company through the ups and downs.

Thank you very much.

Abstract

The ability to sense and respond to hypoxia is essential for the survival of living organisms and is frequently co-opted by tumor cells to adapt to hypoxic microenvironment. As the pivotal transcription factor for hypoxic responses, the hypoxia inducible factors (HIFs) transactivate a wide repertoire of genes that can mediate divergent functional outcomes. The regulation of HIF target gene selectivity is complex, involving the canonical regulation by oxygen sensors prolyl hydroxylases and factor inhibiting HIF (FIH) and the regulated recruitment of transcriptional cofactors, such as p300. However, how HIF target gene selectivity is regulated in cancers has yet to be fully understood.

Apoptosis stimulating protein of p53 2 (ASPP2) belongs to the ASPP protein family whose members are established regulators of p53-family proteins. Recently, it was found that ASPP2 is one of the alternative substrates of FIH and may be a potential competitor of HIF α for FIH activity. However, whether ASPP2 plays a role in the regulation of HIF activity in cancer remains unclear.

In this work we showed that ASPP2 regulates the expression of a subset of HIF target genes irrespective of FIH status. Using HIF target gene *EGLN3* for further investigation, we further found that ASPP2 does not affect either the protein expression, subcellular localization, or the level of DNA-binding of HIFs to the HRE-containing *EGLN3* loci.

To explore whether ASPP2 modulates the expression of *EGLN3* and potentially other ASPP2-sensitive HIF targets through regulating another signaling pathway that crosstalks with HIFs at the transcriptional level, we constructed a bioinformatics pipeline that combines gene coexpression analysis and gene set enrichment analysis to successfully identify both known and hitherto unknown ASPP2-coexpressed pathways at a pan-cancer level.

Among these ASPP2-coexpressed pathways, we focused on the hippo pathway in which transcriptional coactivator YAP has previously been reported to act as a transcriptional cofactor of HIF1 α . We found that the loss of ASPP2 reduces the nuclear localization of YAP under hypoxia. It also attenuates the nuclear interaction between YAP and HIF1 α , as well as the level of DNA-binding of YAP at HRE-containing loci of *EGLN3*.

Overall, this study demonstrates for the first time that ASPP2 modulates a subset of HIF target gene expression, potentially by regulating the interaction between YAP and HIF1 α . Our work also identifies novel pan-cancer ASPP2-associated pathways that would assist in unlocking hypothesis to expand the functions of ASPP2 in human cancers and beyond.

Contents

List of Figures	xi
List of Tables	xiv
Abbreviations	xvi
1 Introduction	1
1.1 Cancer and its microenvironment	1
1.2 Oxygen sensing and responses by HIF	4
1.2.1 Oxygen	4
1.2.2 Structure of HIF heterodimer	5
1.2.3 Regulation of HIF α stability by prolyl hydroxylation	8
1.2.4 Regulation of HIF α transactivational activity by asparaginyl hydroxylation	10
1.2.5 HIF-mediated adaptive responses	11
1.2.6 Regulation of HIF target gene selectivity	12
1.2.6.1 Histone lysine demethylases	14
1.2.6.2 p300/CBP	15
1.2.6.3 CDK8	16
1.2.6.4 YAP	17
1.3 HIF in cancer	19
1.3.1 Oxygen-independent regulation of HIF α activity in cancer	20
1.3.1.1 Crosstalk with tumor suppressive and oncogenic pathways	20
1.3.1.2 Loss of function of pVHL	21
1.3.1.3 Inactivation of the PHDs by TCA intermediates	22
1.3.1.4 The influence of FIH on HIF in cancer	23
1.3.1.5 ARD-containing proteins	24
1.4 ASPP2	26
1.4.1 ASPP protein family	26
1.4.2 Expression and subcellular localization of ASPP2	27
1.4.3 ASPP2 as a regulator of p53 family proteins	28

1.4.4	ASPP2 in development and tumorigenesis	29
1.4.5	ASPP2 p53-independent interactions	31
1.4.5.1	ASPP2 as a regulator of cell polarity	31
1.4.5.2	ASPP2 as a regulator of RAS activity	32
1.4.5.3	ASPP2's interaction with the Hippo signaling pathway	34
1.4.5.4	ASPP2's interaction with HIF regulator FIH	36
1.4.5.5	ASPP2 and hypoxia response	37
1.5	Aims of the study	38
2	Materials and Methods	41
2.1	General	41
2.2	Molecular biology techniques	41
2.2.1	Plasmid preparation	41
2.2.2	Agarose gel electrophoresis	42
2.2.3	Molecular subcloning	42
2.2.4	Bacterial culture	43
2.2.5	Bacterial transformation	43
2.2.6	Glycerol stocks	43
2.3	Tissue culture and cell-based techniques	44
2.3.1	Maintenance and passage of cell lines	44
2.3.2	Freezing and thawing cells	44
2.3.3	Cell transfection	45
2.3.4	RNA extraction and reverse transcription	45
2.3.5	RNA samples preparation for RNA-sequencing	46
2.3.6	Quantitative reverse transcription PCR	46
2.3.7	Small interfering-RNA-induced gene silencing	47
2.3.8	CRISPR-Cas9 gene-editing	48
2.3.9	Generation of stable cell line	48
2.3.10	PCR and Sanger sequencing	50
2.3.11	Immunofluorescence microscopy	50
2.3.12	Western blotting	52
2.3.13	Protein concentration determination	53
2.3.14	Immunoprecipitation	54
2.3.15	Chromatin immunoprecipitation	55
2.3.15.1	Cell fixation and harvest	55
2.3.15.2	Sonication of DNA	56
2.3.15.3	Samples pre-clearing, IP and washes	56
2.3.15.4	Elution, reversal of cross-linking, and proteinase K digestion	57

2.3.15.5	PCR purification	58
2.3.15.6	Quality control of DNA fragment size	58
2.3.16	<i>In situ</i> proximity ligation assay (PLA)	58
2.4	Proteomics sample preparation	60
2.4.1	Glycine elution	60
2.4.2	Methanol/Chloroform precipitation and in-solution digest . .	61
2.4.3	SEP-PAK C18 purifications	61
2.5	Phosphoproteomics sample preparation	62
2.6	Antibodies and primers	63
2.7	Bioinformatics methods	71
2.7.1	Bioinformatics for interactomics	71
2.7.2	Bioinformatics for phosphoproteomics	71
2.7.3	Bioinformatics for RNA-seq	72
2.7.4	Defining HIF target genes in U2OS cells	74
2.7.5	Pan-Cancer Gene Coexpression Analysis	75
3	Generation of ASPP2-KO cell lines using CRISPR-Cas9	78
3.1	Introduction	78
3.1.1	The CRISPR-Cas9 system for gene-editing	78
3.1.2	Results	80
3.1.2.1	Designing the CRISPR-Cas9 system for knocking out ASPP2	80
3.1.2.2	Validation of ASPP2-KO and DKO clones	82
3.1.3	Summary and Brief Discussion	84
4	Investigation of the effects of ASPP2 on HIF transcriptional ac- tivity	87
4.1	Introduction	87
4.2	Results	89
4.2.1	Investigating the effects of ASPP2 on HIF target gene expres- sion using a small panel of HIF targets	89
4.2.2	ASPP2 modulates the expression of a subset of HIF target genes irrespective of FIH status	93
4.2.3	ASPP2 modulates PHD3 expression at the protein level under hypoxia	100
4.2.4	Generation of stable ASPP2-KO U2OS cell line with inducible ASPP2 expression	102
4.2.5	Restoration of ASPP2 expression in ASPP2-KO cells rescues PHD3 protein expression in hypoxia	106

4.2.6	ASPP2 does not affect the protein expression of HIF and p300 under hypoxia	107
4.2.7	ASPP2 does not affect the subcellular localization of HIF1 α under hypoxia	110
4.2.8	The loss of ASPP2 does not affect HIF binding to HRE-containing regions within/near <i>EGLN3</i> gene locus under hypoxia	112
4.2.8.1	Optimization of ChIP protocol to assess the HRE-binding activity of HIF	112
4.2.8.2	The loss of ASPP2 does not affect HIF binding to HRE-containing regions within/near <i>EGLN3</i> gene locus under hypoxia	116
4.2.9	The loss of ASPP2 does not affect p300 or p53 binding to the HRE-containing regions within/near PHD3 gene locus under hypoxia	118
4.2.9.1	Examining p53 response in U2OS cells upon nutlin-3a activation	119
4.2.9.2	The loss of ASPP2 does not affect p300 or p53 binding to the HRE-containing regions within/near <i>EGLN3</i> gene locus under hypoxia	121
4.2.10	Echinomycin abolishes ASPP2's regulation of <i>EGLN3</i> expression	123
4.2.11	ASPP2 and HIF1 α do not seem to directly interact with each other under hypoxia	124
4.2.11.1	ASPP2 and HIF1 α have different subcellular localization under hypoxia	124
4.2.11.2	ASPP2 and HIF1 α cannot be co-immunoprecipitated under hypoxia	126
4.2.12	The regulation of PHD3 expression by ASPP2 is not sensitive to HIF1 α knockdown in hypoxia	127
4.3	Summary and Brief Discussion	128
5	Pan-cancer <i>TP53BP2</i> co-expression analysis	136
5.1	Introduction	136
5.2	Results	140
5.2.1	Overview, assembly, and preprocessing of input data	140
5.2.2	Confounder identification	143
5.2.3	Beta coefficient-based coexpression analysis	146
5.2.4	Evaluation of Beta coefficient-based coexpression analysis	148

5.2.5	Functional interpretation of <i>TP53BP2</i> coexpressed genes . . .	152
5.2.6	Validation of <i>TP53BP2</i> gene coexpression results	157
5.2.6.1	Integrative analysis with ASPP2 transcriptomic data in U2OS cells	158
5.2.6.2	Integrative analysis with ASPP2 interactomic data in U2OS cells	162
5.2.6.3	Integrative analysis with ASPP2 phosphoproteomic data in U2OS cells	164
5.2.6.4	Wet-lab validation of selected <i>TP53BP2</i> coexpressed pathways	168
5.3	Summary and Brief Discussion	171
6	Studying the effects of ASPP2 on the potential HIF1α-YAP crosstalk	175
6.1	Introduction	175
6.2	Results	178
6.2.1	ASPP2 physically interacts with YAP under hypoxia	178
6.2.2	ASPP2-KO affects the nuclear availability of YAP under hypoxia	179
6.2.3	Knocking out ASPP2 attenuated the interaction between YAP and HIF1 α under hypoxia	181
6.2.4	Knocking out ASPP2 compromised the binding of YAP to HRE-containing loci of <i>EGLN3</i> and <i>VEGFA</i>	185
6.3	Summary and Brief Discussion	188
7	Final discussion and perspectives	191
7.1	The role of ASPP2 in HIF signaling	191
7.2	The role of ASPP2 in the regulation of YAP	194
7.3	The role of ASPP2 in cancer	196
7.4	Conclusion	200
	References	201
	Appendices	

List of Figures

1.1	Overview of HIF signaling pathway	6
1.2	Genetic alterations events of <i>HIF1AN</i> in human cancers	23
1.3	Major structural domains of ASPP2	27
1.4	ASPP2 protein expression in human tissues	28
1.5	Genetic alterations events of <i>TP53BP2</i> in human cancers	31
1.6	Multiple hypoxia response signatures display differential expression patterns between wild-type and ASPP2 ^{Δ3/Δ3} mouse intestinal organoids	38
1.7	Major points of regulation that influence HIF signaling activity . .	39
3.1	Transcript variants of <i>TP53BP2</i>	81
3.2	Experimental workflow of CRISPR/Cas9.	83
3.3	Validation of ASPP2-KO and DKO candidates.	85
4.1	The loss of ASPP2 selectively affected the mRNA expression of HIF target genes under physiological hypoxia and hypoxia-mimicking condition.	91
4.2	The loss of ASPP2 selectively affected the mRNA expression of HIF target genes under both normoxia and hypoxia in the absence of FIH.	94
4.3	ASPP2 affects the gene expression of a subset of HIF target genes irrespective of FIH status	98
4.4	The status of FIH affects how ASPP2 affects the gene expression of a subset of HIF target genes	99
4.5	The loss of ASPP2 affects the protein expression of PHD3 under physiological in multiple types of cells.	101
4.6	Design of the piggyBac Tet-On system to create inducible ASPP2 cell lines	102
4.7	Validation of the stable cell lines expressing dox-inducible ASPP2. .	105
4.8	Re-introducing ASPP2 rescues the reduction of PHD3 expression under hypoxia	107
4.9	Major points of regulation of the HIF pathway	108
4.10	The loss of ASPP2 did not influence the protein expression of HIF and p300 under physiological hypoxia	109

4.11	The loss of ASPP2 does not affect the subcellular localization of HIF1 α under hypoxia	111
4.12	Optimization of chromatin fragment size for ChIP	113
4.13	Quality control of the ChIP-qPCR protocol.	115
4.14	The loss of ASPP2 does not affect HIF binding to the HRE-containing loci of <i>EGLN3</i> under hypoxia	117
4.15	The U2OS cell line shows functional p53 transcriptional response upon nutlin-3a activation	120
4.16	The loss of ASPP2 did not affect the recruitment of p300 or p53 to the HIF-binding loci of <i>EGLN3</i> under hypoxia	122
4.17	Echinomycin abrogated the modulatory effect of ASPP2 on <i>EGLN3</i> expression in hypoxia-mimicking condition	124
4.18	ASPP2 and HIF1 α did not co-localize under hypoxia	125
4.19	Physical interaction was not detected between ASPP2 and HIF1 α or HIF1 β under normoxia and hypoxia	127
4.20	HIF1 α knockdown did not affect the regulation of PHD3 expression by ASPP2 under hypoxia	129
5.1	Overview of the gene coexpression analysis pipeline	141
5.2	Overview of the TCGA pan-cancer transcriptomic dataset	142
5.3	Somatic copy number as a potential confounder in gene coexpression analysis.	145
5.4	Beta Coefficient-based coexpression Analysis (BCCA) effectively adjusted for confounding by SCN and facilitated the identification of functionally-related coexpressed genes.	150
5.5	Pan-cancer enrichment patterns of REACTOME pathways coexpressed with <i>TP53BP2</i>	155
5.6	Integrative analysis with ASPP2 transcriptomics data in U2OS cells	159
5.7	Identification of potential ASPP2 interactors using IP-MS.	163
5.8	Integration of ASPP2 interactomic analysis with <i>TP53BP2</i> pan-cancer gene coexpression analysis	165
5.9	Integration of ASPP2 phosphoproteomic analysis with <i>TP53BP2</i> pan-cancer gene coexpression analysis.	167
5.10	Investigating the effects of ASPP2 knockout on the rhythmic expression of PER1 and CRY1 under normoxia and hypoxia	171
6.1	ASPP2 physically interacts with YAP under hypoxia	179
6.2	Knocking out ASPP2 affects YAP localization under hypoxia	180
6.3	Knocking out ASPP2 attenuates the interaction between nuclear HIF1 α and YAP under hypoxia.	183

6.4	Knocking out <i>ASPP2</i> compromises the binding of YAP to the HRE-containing region within the gene loci of <i>EGLN3</i> and <i>VEGFA</i>	188
7.1	Controlling YAP subcellular localization represents one of the potential mechanisms that may explain the regulation of ASPP2 HIF target gene expression	192
7.2	ASPP2 expression positively correlates with more aggressive breast cancer subtypes and p53 mutations	198
7.3	High ASPP2 gene expression is associated with poorer relapse-free survival in breast cancer patients with the basal-like subtype	199
S1	Validation of IP-MS hits by WB	lxiv
S2	Detection of ASPP2-PP1 α interaction using PLA	lxv
S3	Controls for PLA for the detection of YAP-HIF1 α interaction . . .	lxvi

List of Tables

2.1	PCR protocol for plasmid cloning	42
2.2	qPCR preparation protocol	46
2.3	qPCR thermal cycling condition	47
2.4	List of primary antibodies	63
2.5	List of Secondary Antibodies	65
2.6	List of Primers used for Molecular cloning	66
2.7	List of Primers used for sequencing	66
2.8	List of Primers used for qPCR	67
2.9	List of Primers used for ChIP-qPCR	69
2.10	List of plasmids used in this study	70
3.1	Results of gRNA prediction using CRISPOR	82
4.1	Top 5 enriched GO terms in genes downregulated in ASPP2-KO U2OS relative to WT	95
4.2	Top 5 enriched GO terms in genes upregulated in ASPP2-KO U2OS relative to WT	95
4.3	Enriched GO terms in ASPP2-sensitive HIF target genes	96
5.1	Genomic locations of the top 20 <i>TP53BP2</i> coexpressed genes	153
5.2	Enriched REACTOME pathways in ASPP2-KO compared to WT U2OS cells	160
S1	Top 20 potential off-target sites	iii
S2	Top 500 <i>TP53BP2</i> pan-cancer co-expressed genes	vi
S3	Bottom 500 <i>TP53BP2</i> pan-cancer co-expressed genes	xxi
S4	REACTOME pathways coexpressed with <i>TP53BP2</i> at the pan- cancer level	xxxvi
S5	Differentially expressed genes up-regulated in ASPP2-KO U2OS cells compared to WT U2OS cells	xlvi
S6	Differentially expressed genes down-regulated in ASPP2-KO U2OS cells compared to WT U2OS cells	xlix
S7	List of ASPP2-downregulated translation-related genes	liv

S8	Enriched KEGG pathways for candidate ASPP2 interactors identified by IP-MS	lv
S9	List of candidate ASPP2-regulated phosphosites identified by phosphoproteomic analysis	lvii

Abbreviations

$\Delta 3/\Delta 3$	exon 3 deletion
2-OG	2-oxoglutarate
4E-BP	Eukaryotic translation initiation factor 4E-binding protein 1
Ac-BSA	Acetylated BSA
ACC	Adrenocortical carcinoma
ACTRT1	Actin Related Protein T1
AhR	Aryl hydrocarbon receptor
AJ	Adherens junction
ANKRD37	Ankyrin Repeat Domain-Containing Protein 37
ARD	Ankyrin Repeat Domain
ARNT	Aryl hydrocarbon receptor nuclear translocator
ARNT2	Aryl hydrocarbon receptor nuclear translocator 2
Asn	Asparagine
ASPP1	Apoptosis-stimulating protein of p53 1
ASPP2	Apoptosis-stimulating protein of p53 2
ATP	Adenosine triphosphate
BAX	BCL2 Associated X, Apoptosis Regulator
BCCA	Beta coefficient-based co-expression analysis
bHLH-PAS	basic helix-loop-helix-PER-ARNT-SIM
BLCA	Bladder urothelial carcinoma
BMAL1	Brain and muscle ARNT-like 1
BNIP3	BCL2-Interacting Protein 3
bp	Base pair
BRCA	Breast invasive carcinoma
BSA	Bovine serum albumin
C-TAD	C-terminal-transactivation domain
CA9	Carbonic anhydrase 9
CAG	CMV early enhancer/chicken β actin
Cas9	CRISPR associated protein 9
CBP	CREB-binding protein
CCN2	Cellular communication network factor 2
ccRCC	Clear cell renal cell carcinoma

CDK8	Cyclin-dependent kinase 8
CDK9	Cyclin-dependent kinase 9
CDKN1A	Cyclin Dependent Kinase Inhibitor 1A
cDNA	Complementary DNA
CESC	Endocervical adenocarcinoma
CH1	Cysteine-rich domain 1
CH2	Cysteine-rich domain 2
CH3	Cysteine-rich domain 3
ChIP	Chromatin immunoprecipitation
ChIP-seq	ChIP-sequencing
CHOL	Cholangioma
CLOCK	Clock circadian regulator
CNS	Central nervous system
COAD	Colon adenocarcinoma
CRC	Colorectal adenocarcinoma
CREB	CRE-response element binding protein
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
CRY1	Cryptochrome circadian regulator 1
CTGF	Connective tissue growth factor
DAPI	4,6-diamidino-2-phenylindole
DEGs	Differentially expressed genes
DLBC	Diffuse large B cell lymphoma
DMEM	Dulbecco's modified eagle medium
DMOG	Dimethyloxalylglycine
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
Dnase	Deoxyribonuclease
Dox	Doxycycline
DSB	Double-stranded break
dsDNA	double-stranded DNA
DTT	Dithiothreitol
E2F	E2 factor family of transcription factors
ECL	Enhanced chemiluminescent
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EGLN1	Egl nine homolog 1
EGLN2	Egl nine homolog 2
EGLN3	Egl nine homolog 3

eIF-4E	Eukaryotic translation initiation factor 4E
EIF2A	Eukaryotic Translation Initiation Factor 2A
EPAS1	Endothelial PAS Domain Protein 1
EPO	Glycoprotein hormone erythropoietin
ERK2	Mitogen-activated protein kinase 1
ESCA	Esophageal carcinoma
ESI	Electrospray ionization
et al.	Et alia
etc.	Et cetera
FACS	Fluorescence-activated cell sorting
FDR	False discovery rate
FH	Fumarate hydratase
FIH	Factor inhibiting HIF
FRMD6	FERM domain-containing protein 6
GBM	Glioblastoma multiforme
GDC	Genomic Data Commons
GFP	Green fluorescent protein
GOF	Gain-of-function
GRACE	Genomic Regression Analysis of Coordinated Expression
gRNA	guide-RNA
GSEA	Gene set enrichment analysis
GTFs	General transcription factors
H3K27ac	Histone 3 lysine 27 acetylation
H3K9	Histone 3 lysine 9
HAT	Acetyl-transferase domain
HER2	Epidermal growth factor receptor 2
HIF	Hypoxia-inducible factor
HIF1AN	Hypoxia Inducible Factor 1 Subunit Alpha Inhibitor
hiHTGs	hypoxia-inducible HIF target genes
HK2	Hexokinase 2
HNSC	Head and neck squamous cell carcinoma
HRE	Hypoxia response element
HRP	Horseradish peroxidase
i.e.	id est
iASPP	Inhibitor of Apoptosis-Stimulating Protein of p53
IBiD	Interferon response factor-binding domain
IF	Immunofluorescence
IgG	Immunoglobulin G
IκB	NFκB inhibitor

Indel	Insertion/deletion
IP	Immunoprecipitation
ITRs	Inverted terminal repeats
JmjC	Jumonji-C domain
kDa	Kilodalton
KDMs	Histone demethylases
KEGG	Kyoto Encyclopedia of Genes and Genomes
KICH	Kidney chromophobe
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney renal papillary cell carcinoma
KIX	CREB and c-Myb binding domain
KO	Knockout
KRAS	Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
L-2HG	L-2-hydroxyglutarate
LAML	Acute myeloid leukemia
LATS1	Large tumour suppressor homologue 1
LATS2	Large tumour suppressor homologue 2
LB	Lysogen broth
LC	Liquid chromatography
LDHA	Lactate dehydrogenase A
LFQ	Label-free quantification
LGG	Brain lower-grade glioma
LIHC	Liver hepatocellular carcinoma
LOF	Loss-of-function
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
mAb	monoclonal antibody
MAPK	Mitogen-activated protein kinase
MDM2	Mouse double minute 2 homolog
MEM	Minimal essential medium
MESO	Mesothelioma
MOB1A	MOB domain kinase activator 1A
MOB1B	MOB domain kinase activator 1B
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MST1	Mammalian STE20-like protein kinase 1
MST2	Mammalian STE20-like protein kinase 2
mTORC1	Mammalian target of rapamycin complex 1
MXI1	MAX interactor 1

MYC	Myelocytomatosis proto-oncogene
MYPT1	Myosin phosphatase target subunit 1
N-TAD	N-terminal-transactivation domain
NAD	Nicotinamide adenine dinucleotide
NES	Normalized enrichment score
NFkB	Nuclear factor kappa-light-chain-enhancer
NHEJ	Non-homologous end joining
Notch	Neurogenic locus notch homolog protein
ODD	Oxygen-dependent degradation domain
OV	Ovarian carcinoma
OXPHOS	Oxidative phosphorylation
PAAD	Pancreatic adenocarcinoma
PAGE	Polyacrylamide gel electrophoresis
PAM	Protospacer adjacent motif
Par3	Partitioning-defective protein 3
PARD3	Partitioning-defective 3 homolog
PARP2	Poly(ADP-Ribose) Polymerase 2
PB	PiggyBac
PBS	Phosphate-buffered saline
PCPG	Pheochromocytoma and paraganglioma
PCR	Polymerase chain reaction
PDK1	Pyruvate dehydrogenase kinase 1
PER1	Period circadian regulator 1
PFA	Paraformaldehyde
PHDs	Prolyl hydroxylases
PIC	Transcription Pre-initiation complex
PIG3	p53-induced gene 3 protein
PKM2	Pyruvate kinase isoform 2
PLA	Proximity ligation assay
PP1	Protein phosphatase 1
PRAD	Prostate adenocarcinoma
Pro	Proline
PTEN	Phosphatase and tensin homolog
PUMA	p53 upregulated modulator of apoptosis
pVHL	von Hippel-Lindau tumor-suppressor protein
QC	Quality control
qPCR	Quantitative polymerase chain reaction
qRT-PCR	Quantitative reverse transcription PCR
R	R programming language and software environment

RASSF8	Ras-association domain-containing protein 8
READ	Rectal adenocarcinoma
RID	Nuclear receptor interaction domain
RNA	Ribonucleic acid
RNA-seq	RNA-sequencing
RNAPII	RNA polymerase II
ROCK1	Rho-associated protein kinase 1
ROCK2	Rho-associated protein kinase 2
ROS	Reactive oxygen species
RSEM	RNASeq by Expectation Maximization
rtTA	Reverse tetracycline transactivator
s0	Artificial within-group variance
SARC	Sarcoma
SAV1	Salvador homologue 1
SCN	Somatic copy number
SCNA	Somatic copy number alteration
SDH	Succinate dehydrogenase
SDS-PAGE	Sodium dodecyl sulfate
Siah2	Absentia homolog 2
SIM	Single-minded homolog
siRNA	Small-interfering RNA
SKCM	Skin cutaneous melanoma
SLC2A1	Solute carrier family 2 member 1
SMAD	Mothers against decapentaplegic homolog
STAD	Stomach carcinoma
STAT3	Signal transducer and activation of transcription 3
STES	Stomach and esophageal carcinoma
STY	Serine, threonine, tyrosine
TALENs	Transcription activator-like effector nucleases
TAZ	Transcriptional coactivator with PDZ-binding motif
TBS	Tris-buffered saline
TBST	TBS with Tween-20
TCA	Tricarboxylic acid cycle
TCGA	The Cancer Genome Atlas
TEADs	TEA domain family members
Tet	Tetracycline
TGCT	Testicular germ cell tumors
THCA	Thyroid carcinoma
THYM	Thymoma

TP53	Tumor protein p53
TP53BP2	TP53-binding protein 2
tracrRNA	Trans-activating crRNA
TRE	Tetracycline-response element
Trp53bp2	Transformation related protein 53 binding protein 2
UCEC	Uterine corpus endometrial carcinoma
UCS	Uterine carcinosarcoma
UV	Ultraviolet
UVM	Uveal melanoma
VEGFA	Vascular endothelial growth factor
WB	Western blot
WT	Wild-type
YAP	Yes-associated protein
ZFNs	Zinc finger nucleases

1

Introduction

1.1 Cancer and its microenvironment

Cancer is a global leading cause of morbidity and mortality. It afflicts an estimated 18.1 million of people worldwide and causing an estimated 9.6 million of deaths in 2018 (Bray et al. 2018). These figures are projected to increase further by over 20 percent over the next decade (Bray et al. 2012). The socioeconomical implications of cancer are arguably even more far-reaching than these numbers suggest, considering the physical, emotional, and financial burdens cancer inflicts on the patients, their family, and the society.

Driving these mind-boggling statistics and enormous burdens is a disease with profoundly heterogeneous traits, behaviors, and causal mechanisms. At its core, cancer is a result of genetic alterations, through which a subpopulation of normal cells become defective in the molecular machineries regulating their proliferation, differentiation, migration, and death, and consequently, transform into tumor cells with uncontrolled growth and invasiveness (Hanahan et al. 2000; Meyerson

et al. 2010). Such transformation has been proposed, based on multiple lines of evidence, to be a multi-step process that involves a succession of mutations at the gene or chromosome level (Cho et al. 1992; Fearon et al. 1990; Hahn et al. 1999; Kinzler et al. 1996; Renan 1993). Although cancer can arise due to different combinations of genetic alterations, there are genes that are more frequently mutated or altered than the others. In a simplistic view, one group of such genes are called proto-oncogenes, for example, *KRAS* (Simanshu et al. 2017), which often function on one side of cellular homeostasis - promoting cell survival, growth, and migration, etc. Conversely, another group of frequently altered genes are termed tumor suppressor genes, for example, *TP53* (Muller et al. 2014). These genes often function on the other side of the homeostatic processes - slowing down growth, repairing DNA damage, instructing suicidal program, to name a few examples. During tumorigenesis and cancer progression, mutations and alterations of these proto-oncogenes and tumor suppressor genes often result in gain-of-function effects for the former and loss-of-function effects for the latter. Consequently, these effects disrupt different cellular homeostatic processes and confer survival and growth advantages to the host cells, promoting their malignant growth.

Although the consequences of genetic alterations in different tumors can be manifested in different ways, there are still some common hallmarks that are shared among different types of tumors. These were famously described by Hanahan and Weinberg in 2000 and further expanded upon a decade later (Hanahan et al. 2000, 2011). These hallmarks summarize the aberrations of tumor cells in growth and growth-prohibitive signaling, responses to various intra- and extra-cellular

stresses, metabolic changes, invasion and metastasis, and their interactions with the immune system.

Importantly, it has become clear that these hallmarks of cancer are not solely attributed to cell-autonomous properties of tumor cells, but also a result of the dynamic, reciprocal interplay between the tumor cells and different factors and components in the tumor microenvironment, such as the infiltrating immune cells, endothelial cells, and extracellular matrix.

Among the many different factors within the tumor microenvironment, oxygen, or lack thereof, is an essential factor that constantly shapes the interactions between tumor cells and their microenvironment. Indeed, hypoxia is a prevalent feature in most solid tumors as a result of uncontrolled cell proliferation and defective vasculature (Helmlinger et al. 1997; Vaupel et al. 2004). Rapid proliferation of tumor cells create an oxygen demand that exceeds their vascular supply. This unmet demand is further coupled with the structural and functional abnormality of tumor vasculature as a result of uncontrolled angiogenesis, leading to heterogeneous perfusions in tumors (Gillies et al. 1999). This results in necrotic cell death in the core of the tumor mass due to severe depletion of oxygen, i.e. anoxia, and the formation of hypoxic and peri-necrotic regions. Therefore, it is important to develop a better understanding of the interactions of oxygen sensing and response pathways with key tumor signaling pathways.

1.2 Oxygen sensing and responses by HIF

1.2.1 Oxygen

A continuous supply of oxygen is necessary for the survival and proper functioning of all human cells, which are largely dependent on the aerobic energy production process called oxidative phosphorylation (OXPHOS). Oxygen sits at the very end of OXPHOS as the final electron acceptor, enabling the establishment of an electrochemical gradient in the electron transport chain that leads to the synthesis of adenosine triphosphate (ATP), which is the energy currency of human cells to fuel various cellular activities.

This reliance on oxygen makes human cells susceptible to abrupt and prolonged changes in oxygen level or metabolic demands, which can have undesirable or lethal consequences, such as cytotoxicity due to a heightened production of reactive oxygen species (Waypa et al. 2016). To minimize these potential damages, cellular oxygen levels are maintained close to certain set points *in vivo*, which can vary from near 21% oxygen in our upper respiratory tract to around 1% oxygen at the renal corticomedullary junction (Gregg L. Semenza 2010). These physiological oxygen levels are termed normoxia.

Over the course of their lives, human cells may encounter various extra- and intracellular stresses that challenge the delivery and utilization of oxygen, leading to a reduction of local oxygen tension, which is conventionally termed hypoxia. When confronted with hypoxic challenges, human cells are equipped with a robust oxygen sensing machinery that enables them to sense the fluctuations of oxygen availability

and respond accordingly in an attempt to restore oxygen homeostasis, or avoid more extensive damage by committing suicide if the former option is not feasible.

A major component of this oxygen sensing machinery is a family of transcription factors called hypoxia-inducible factors (HIFs), which, together with their associated upstream regulators, constitute a pivotal signaling pathway that is conserved evolutionarily among metazoans to regulate cellular oxygen homeostasis (Kaelin et al. 2008; Gregg L Semenza 2012).

1.2.2 Structure of HIF heterodimer

HIFs are an obligate protein heterodimer consisting of an alpha subunit and a beta subunit (Semenza et al. 1992; Gu et al. 1998; H. Tian et al. 1997; G. L. Wang et al. 1995). Both subunits belong the basic helix-loop-helix-PER-ARNT-SIM (bHLH-PAS) transcription factor family, of which members can be categorized into class I and class II. Class I bHLH-PAS members, which include the alpha subunit of HIF, are mainly regulated by various environmental signals. Class II bHLH-PAS members, such as the beta subunit of HIF, on the other hand, are usually constitutively expressed. Members from this transcription factor family function by forming class I: class II-heterodimers, which usually bind to E-Box elements (5'-CACGTG-3') or E-Box-like elements in the DNA and are involved in initiating various developmental and stress-induced transcriptional programs (Kewley et al. 2004).

In humans, HIF alpha subunit can be encoded by three different genes, namely, *HIF1A*, *EPAS1*, and *HIF3A*, resulting in three different isoforms, HIF1 α , HIF2 α , and HIF3 α (Ema et al. 1997; Gu et al. 1998; Iyer, Leung, et al. 1998), respectively. All three isoforms contain a bHLH domain for DNA binding and two PAS domains

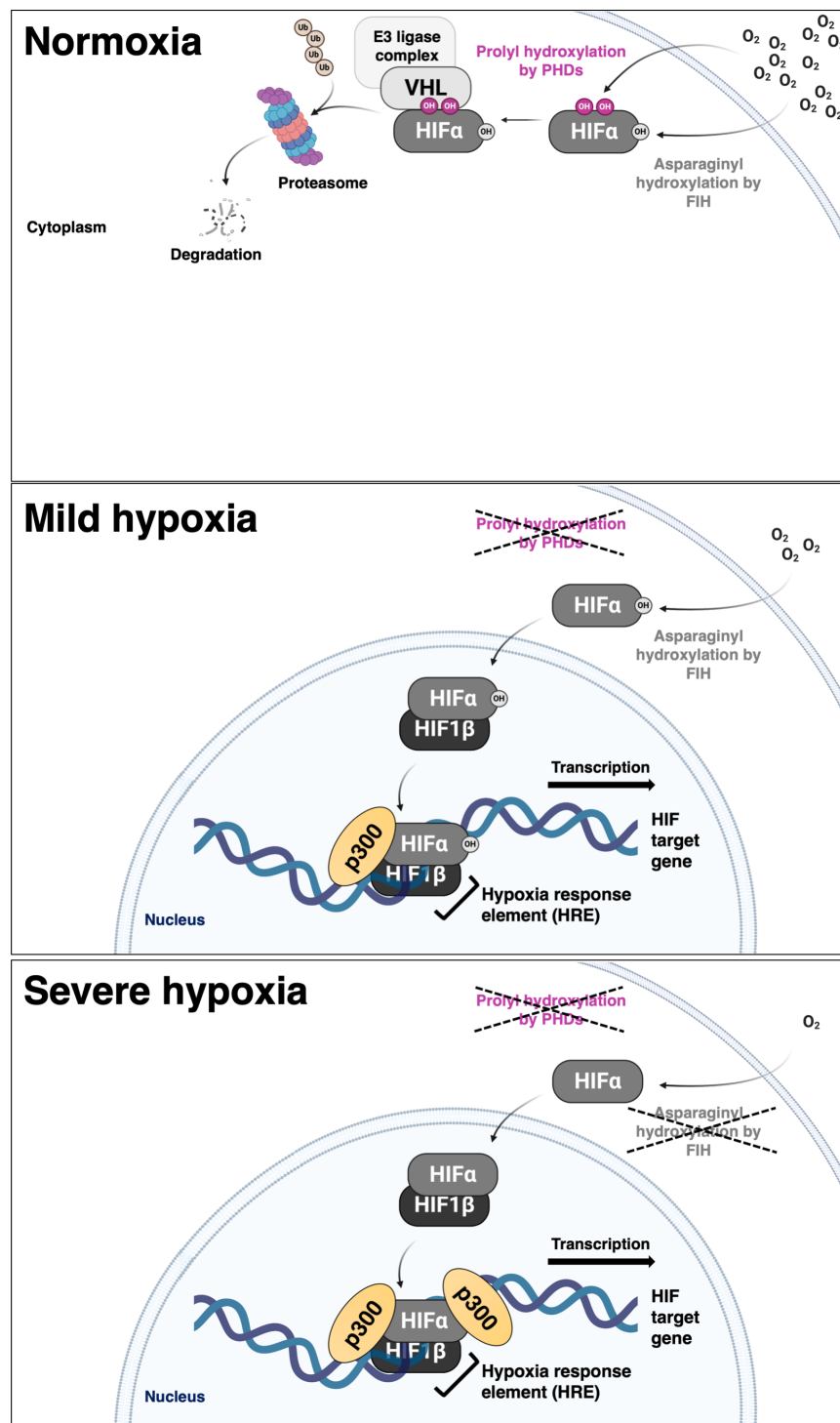


Figure 1.1: Overview of HIF signaling pathway. Schematic of a brief overview of the regulation of hypoxia-inducible factors (HIFs) by oxygen-dependent prolyl and asparaginyl hydroxylation under normoxia, mild hypoxia, and severe hypoxia. PHDs: prolyl hydroxylases; FIH: Factor-inhibiting HIF; VHL: Von Hippel-Lindau tumor suppressor.

for heterodimerization in the N-terminus, and an N-terminal-transactivation domain (N-TAD) in the C-terminus, which confers binding specificity and transactivation functions, likely by interacting with different transcriptional factors and co-factors (Akiri et al. 2003; Blau et al. 1996; Cheng-Jun Hu et al. 2007). HIF1 α and HIF2 α , but not HIF3 α , also have a C-terminal transactivation domain (C-TAD) in their C-terminus, which has also been shown to interact with transcriptional cofactors and synergize with N-TAD in transactivating HIF target genes (Yan et al. 2007).

HIF1 α and HIF2 α , in particular, have a high degree of sequence homology in their bHLH and PAS domains and share a common DNA-binding motif - 5'-(A/G)CGTG-3' - which is an E-box-like element termed the hypoxia response element (HRE). Despite their similarities, HIF1 α and HIF2 α have been shown to display differential tissue-level and temporal expression patterns under hypoxia and more recent reports have suggested that they have overlapping but distinct DNA-binding patterns and transcriptional output (Holmquist-Mengelbier et al. 2006; Smythies et al. 2019; H. Tian et al. 1997; Wiesener et al. 2003).

HIF3 α , on the other hand, has multiple splice variants and is more divergent compared to HIF1 α and HIF2 α in terms of sequence and structural homology (Maynard et al. 2003). Functionally, HIF3 α is also the least similar among the three isoforms as some HIF3 α variants have been reported to play an opposing role to those of HIF1 α and HIF2 α (Makino et al. 2001; Maynard et al. 2005). This study mainly focuses on HIF1 α and HIF2 α and will use the designation HIF α when these two isoforms are referred to generically.

The heterodimeric partner of HIF α , HIF's β subunit, is also encoded by different genes, namely, aryl hydrocarbon receptor nuclear translocator (*ARNT* and *ARNT2*),

producing two paralogs - ARNT (also known as HIF1 β) and ARNT2 - with high sequence similarity to each other (Hirose et al. 1996). However, while HIF1 β is ubiquitously expressed in most cell types, the expression of ARNT2 is restricted to neuronal and renal cells (Drutel et al. 1996; Hirose et al. 1996), and therefore only HIF1 β will be focused on in this study. HIF1 β is mainly localized in the nucleus and can heterodimerize with other class I members of the bHLH-PAS family in addition to HIF α , including aryl hydrocarbon receptor (AhR) and single-minded homolog (SIM) proteins. AhR heterodimerizes with ARNT upon xenobiotic activation and AhR:ARNT heterodimer recognizes and binds to another E-Box-like element called the xenobiotic response element, while SIM:ARNT heterodimers are more specialized in neuronal functions (Denison et al. 2003; Kewley et al. 2004).

1.2.3 Regulation of HIF α stability by prolyl hydroxylation

Unlike the expression of HIF1 β , which is constitutively expressed, the stability of HIF α is highly sensitive to oxygen and represents one of the major points of regulation in the HIF pathway.

The protein expression of HIF α is tightly regulated by post-translational prolyl hydroxylation at its oxygen-dependent degradation domain (ODD) embedded in the N-TAD. The two highly conserved proline residues in the ODD (Pro-402 and Pro-564 in HIF1 α , Pro-405 and Pro-531 in HIF2 α) can be catalyzed by a group of three iron- and 2-oxoglutarate (2-OG)-dependent prolyl hydroxylases (PHDs) 1-3, which are encoded by *EGLN*1,2, and 3, respectively (Bruick et al. 2001; Epstein et al. 2001; Ivan et al. 2002; Masson et al. 2001). They belong to the extended superfamily of 2-OG-dependent dioxygenases (Schofield et al. 1999) and require molecular oxygen

and 2-OG as co-substrates, as well as iron-binding at the catalytic site for their hydroxylase activity. In the hydroxylation reaction, one oxygen atom is inserted into the proline residue, forming a hydroxyproline, while the other oxygen atom is inserted into 2-OG, resulting in the formation of carbon dioxide and succinate. Importantly, PHDs have relatively low affinities for oxygen as suggested by *in vitro* biochemical studies (Hirsilä et al. 2003). This feature enables them to sense oxygen over the entire physiological range and connects the availability of oxygen to the downstream components of the hypoxia response pathway.

In normoxia, the prolyl-hydroxylated form of HIF α is recognized and bound by the von Hippel-Lindau tumor-suppressor protein (pVHL) (Huang et al. 1998; F. Yu et al. 2001; Maxwell et al. 1999). pVHL is a subunit of a multi-protein E3-ubiquitin ligase complex responsible for substrate recognition. Consequently, the binding of pVHL to HIF α leads to polyubiquitination and proteasomal degradation of the latter. In hypoxic conditions, however, the activity of PHDs diminishes as the availability of oxygen declines. More HIF α thus escapes degradation and become stabilized. Upon stabilization, HIF α is imported into the nucleus, where it heterodimerizes with HIF1 β . This heterodimer then binds to the HRE-containing loci located within the promoter, intronic, or enhancer regions of its target genes, and recruits transcriptional co-factors, such as histone lysine acetyltransferases p300 and its paralog CRE-response element binding protein (CREB)-binding protein (CBP), *en route* to the formation of the transcription initiation complex and transactivation of target genes (Arany et al. 1996; Gregg L Semenza 2002) (discussed further below).

1.2.4 Regulation of HIF α transactivational activity by asparaginyl hydroxylation

In addition to its stability, the transactivational activity of HIF α constitutes another point of regulation of the HIF pathway.

One such regulation can be attributed to the asparaginyl hydroxylation mediated by another iron- and 2-OG-dependent dioxygenase called factor inhibiting HIF (FIH), encoded by *HIF1AN*. Under normoxic condition, FIH hydroxylates an asparagine residue within the C-terminal transactivation domain (C-TAD) of HIF α (Asn-803 in HIF1 α , Asn-847 in HIF2 α) (Mahon et al. 2001; D. Lando et al. 2002; David Lando et al. 2002). This hydroxylation disrupts the physical interaction between C-TAD of HIF α and the CH1 domain of p300/CBP. This disruption impairs the transactivation function of HIF α C-TAD and leads to the repression of a subset of HIF target genes (M. C. Chan et al. 2016; Kasper et al. 2005). Similar to the PHDs, FIH also requires oxygen as a co-factor for its hydroxylation activity, therefore its negative regulation of HIF transcriptional activity is also controlled by the availability of oxygen. Interestingly, the affinity of FIH for oxygen is higher than that of the PHDs. Consistent with this observation, it has been shown that FIH still retains partial activity in severely hypoxic cells, in which the PHDs are largely inhibited (Stolze et al. 2004; Tarhonskaya et al. 2015; Y.-M. Tian et al. 2011a), suggesting that FIH hydroxylation may serve as a fine-tuning mechanism to modulate the transactivational activity of HIF α when it is stabilized under hypoxic condition.

1.2.5 HIF-mediated adaptive responses

Studies using chromatin immunoprecipitation assays have identified over hundreds of HIF target genes (Schödel et al. 2011; Smythies et al. 2019; Xia et al. 2009). Many of them encode proteins that are implicated in different functions to maintain oxygen homeostasis, such as increasing oxygen supply to the cells. Example HIF target genes that function in oxygen delivery include glycoprotein hormone erythropoietin (EPO) and vascular endothelial growth factor (VEGFA) (Kimura et al. 2001; Semenza et al. 1992). The transactivation of *EPO* by HIF leads to an increased circulating levels of EPO, which in turn stimulates the increased production of red cell mass and oxygen delivery to the cells (Scholz et al. 1990). VEGFA is a key proangiogenic factor with pleiotropic functions related to the formation of new blood vessels from endothelial precursor cells (vasculogenesis) and pre-existing vasculature (angiogenesis). By binding to the VEGF receptors, VEGFA activates multiple receptor tyrosine kinase signaling cascades to orchestrate angiogenic responses, such as the proliferation and migration of endothelial cells and the induction of vascular permeability (Koch et al. 2012).

Another major HIF-mediated adaptive response to hypoxia is to reduce oxygen consumption by the cells through reprogramming cellular metabolism. As mentioned in the earlier subsection, most human cells primarily rely on OXPHOS to provide energy for various metabolic and mechanical cellular processes as the complete oxidation of glucose into carbon dioxide and water provides a much greater yield of ATP than glycolysis alone, in which glucose is incompletely oxidized into lactate. However, the downside of OXPHOS is that it also produces reactive oxygen species

(ROS) as a by-product due to the premature reaction between electrons and oxygen at the electron transport chain complex I or complex III (Murphy 2009). ROS are important second messengers that function in various signaling pathways but their levels need to be kept at a tightly regulated balance by the cellular antioxidant system as excessive ROS have deleterious effects by damaging various cellular components due to their oxidative properties. Hypoxia results in a heightened production of ROS, and therefore a major rationale underlying the metabolic reprogramming mediated by HIF is to mitigate potential damages caused by ROS while maintaining ATP production, exemplified by the lethality of the excessive ROS in *HIF1A*-knockout mouse embryonic fibroblasts (J. W. Kim et al. 2006; Zhang et al. 2008). Metabolic adaptation to hypoxia involves HIF activating the transcription of various glycolytic enzymes and glucose transporters, such as hexokinase 2 (HK2), solute carrier family 2 member 1 (SLC2A1), and lactate dehydrogenase A (LDHA) to increase glycolytic flux (C. Chen et al. 2001; Hayashi et al. 2004; C.-J. Hu et al. 2003; Iyer, Kotch, et al. 1998). HIF also transactivates multiple target genes, such as pyruvate dehydrogenase kinase 1 (PDK1), MAX interactor 1 (MXI1), and BCL2-Interacting Protein 3 (BNIP3) to suppress mitochondrial metabolism and OXPHOS (Bellot et al. 2009; Corn et al. 2005; Zhang et al. 2007; J. W. Kim et al. 2006; Zhang et al. 2008).

1.2.6 Regulation of HIF target gene selectivity

Transactivation of HIF target genes requires HIF heterodimer binding to the DNA at its consensus motif. However, the human genome contains over a million HIF consensus motifs (HRE, 5'-(A/G)CGTG-3') and only a small subset of them has

been reported to be bound by HIF (Schödel et al. 2011). Over the past decade, increasing efforts have been made to understand the DNA-binding selectivity of HIF, aided by the advances in next-generation sequencing methodologies. Studies using chromatin immunoprecipitation (ChIP) in combination with pan-genomic technologies such as ChIP-on-chip and ChIP-sequencing (ChIP-seq) have shed light on the HIF1 α -, HIF2 α -, and HIF1 β -binding sites, which are in the order of hundreds using higher stringency criteria, such as using the existence of HRE in the detected loci (Schödel et al. 2011) or the overlapping binding peaks between HIF α and HIF1 β (Xia et al. 2009). Coupling with epigenetic analyses using the above pan-genomic techniques, it was further revealed that HIF binding is highly correlated with gene loci that are within accessible chromatin regions and already transcriptionally active in normoxia, indicated by DNase I hypersensitivity and histone H3 lysine 4 methylation, respectively (Schödel et al. 2011). This feature suggests that normoxic and hypoxic epigenetic regulation, such as histone lysine methylation described below, helps set the tone for HIF binding and thereby defining HIF-mediated transactivation activity in response to hypoxia.

The integration of these pan-genomic analyses with transcriptomic analyses based on microarray or RNA-sequencing (RNA-seq) experiments performed in normoxia and hypoxia further allows researchers to differentiate direct transcriptional targets of HIF from indirect targets. Such integrative analyses reveal a significant association of HIF-bound genes preferentially with genes up-regulated in response to hypoxia, suggesting that HIF acts in *cis* primarily as a transcriptional activator, while the majority of its transcriptionally repressive activities is mediated through

indirect mechanisms, for instance, by transactivating transcriptional repressors (Yun et al. 2002) and microRNAs (Kulshreshtha et al. 2007).

1.2.6.1 Histone lysine demethylases

Moreover, several Jumonji-C (JmjC) domain-containing histone lysine demethylases, including KDM3A, KDM4B, KDM4C, and KDM5B, which belong to the same superfamily of 2-OG-dependent dioxygenases as the PHDs and FIH, are identified among the direct HIF target genes (Pollard et al. 2008; Schödel et al. 2011; Xia et al. 2009). These KDMs are epigenetic regulators that can either enhance or suppress gene expression by reversing specific histone methylation marks and some of them have been reported to serve as a transcriptional coactivator of HIF1 α , suggesting a potential positive feedback regulation. For instance, KDM3A has been shown to act as a coactivator of HIF1 α in urinary bladder cancer cells and promotes HIF1 α -mediated glycolytic switch and cell proliferation (Wan et al. 2017), whereas KDM4C interacts with HIF1 α in breast cancer cells to stimulate its transactivation of target genes involved in glycolysis and metastasis (Luo et al. 2012). Interestingly, two JmjC domain-containing demethylases, KDM5A and KDM6A, are recently identified as direct oxygen sensors (Batie et al. 2019; Chakraborty et al. 2019). Due to their low affinities for oxygen, hypoxia directly regulates the demethylation activities of these histone modifiers independent of HIF and other 2-OG related metabolites, resulting in global histone methylation changes that precede cellular transcriptional response (Batie et al. 2019; Chakraborty et al. 2019). How these oxygen-sensitive KDMs interact with HIF to coordinate cellular transcriptional response to physiological

and tumor hypoxia and whether there are other unidentified oxygen-sensitive JmjC domain-containing proteins remain important areas to be elucidated.

1.2.6.2 p300/CBP

The transactivation activity of HIF is not solely determined by HIF binding to accessible HRE loci alone. The interaction of HIF with transcriptional co-factors also provides a major layer of regulation to enable differential transactivation of HIF target genes in a context-dependent manner, as the recruitment of different cofactors to a given HRE gene locus by HIF regulates the assembly and function of the pre-initiation complex (PIC) which - formed by RNA polymerase II (RNAPII) and the general transcription factors (GTFs) TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH - is responsible for the transcription initiation and elongation of a given HIF target gene (Luse 2013).

Until recently, p300/CBP was the only transcriptional coactivator of HIF that had been extensively studied and cited. p300/CBP are two highly structurally-related lysine acetyltransferases with highly conserved functional domains, both functioning as transcriptional coactivators in a number of different signaling pathways and are required for normal embryonic development (H. M. Chan et al. 2001; Yao et al. 1998). They exert transcriptional regulation, in part, by coupling a wide variety of motif-specific transcription factors with components of the general transcriptional machinery (H. M. Chan et al. 2001; Vo et al. 2001).

Early studies demonstrated that p300/CBP interact with the C-terminal transactivation domain (C-TAD) of HIF α through its CH1 domain, and this interaction stimulates the transcription of several well-established HIF target genes, including

VEGFA and *EPO* (Arany et al. 1996; Dames et al. 2002; Steven J. Freedman et al. 2003; Kallio et al. 1998). Interestingly, although the homozygous deletion of CH1 domain in either p300 or CBP largely suppresses the transactivation activity of HIF1 α C-TAD in mouse embryonic fibroblasts (Kasper et al. 2005), gene expression profiling studies by Kasper *et al.* showed that disruption of the interaction between p300/CBP and HIF1 α did not have the same consequence for the expression of each HIF target gene. In fact, according to their sensitivity to the deletion of p300/CBP CH1 domain, Kasper *et al.* suggested that HIF target genes could be broadly categorized into three subgroups of which their gene expression were largely dependent, partially dependent, and mostly independent of the recruitment of p300/CBP by HIF, suggesting that rather than acting as a binary on/off switch of HIF transactivation activity, p300/CBP functions as a modulator to selectively control HIF target gene expression (Kasper et al. 2005; Kasper et al. 2006), thereby providing another regulatory layer to fine-tune HIF transcriptional output. The findings above were subsequently corroborated by Chan *et al.* using a different approach, in which they compared the transcriptional responses from cultured breast cancer cell line MCF-7 treated with a small molecule inhibitor that selectively inhibits PHDs activity and a broad-spectrum 2-OG-dependent dioxygenase inhibitor that inhibits both PHDs and FIH (M. C. Chan et al. 2016).

1.2.6.3 CDK8

Another transcriptional coactivator that connects HIF and the formation of PIC is cyclin-dependent kinase 8 (CDK8), which is a subunit of the Mediator complex. Mediator is an evolutionarily conserved multiprotein complex in metazoans,

functioning as a critical regulator of the transcriptional machinery, including the initiation, pausing, and elongation of RNAPII (Allen et al. 2015; Malik et al. 2010, 2000). Over 30 subunits have been identified for the Mediator complex and its variable subunit composition permits differential and precise regulation of RNAPII activity in a stimulus- and cell type-specific manner.

Being a subunit that associates with the Mediator complex in a facultative manner (Y. Liu et al. 2001; Taatjes et al. 2002), CDK8 has been shown to co-activate a large number of HIF1 α target genes in response to hypoxia (Galbraith et al. 2013). Mechanistically, using HIF target gene *ANKRD37* as an example, Galbraith *et al.* found that CDK8 affects neither HIF1 α binding nor histone acetylation at the *ANKRD37* locus. Instead, CDK8 is recruited to the HIF target gene locus through interacting with the transactivation domain of HIF1 α (Galbraith et al. 2013). This recruitment is necessary for the further recruitment of the super elongation complex, leading to the release of paused RNAPII and transcription elongation of *ANKRD37* (Galbraith et al. 2013). Dysregulation of CDK8 has been found in multiple cancers and implicated in both oncogenic and tumor suppressive signalling (Donner et al. 2007; Firestein et al. 2008; Kapoor et al. 2010; McClelland et al. 2015; McDermott et al. 2017). Given also that CDK8 can both stimulate and suppress RNAPII activity (Hengartner et al. 1998; Knuesel et al. 2009), the functional consequence of HIF1 α -CDK8 interaction in cancers is likely to be context-specific.

1.2.6.4 YAP

Another transcriptional coactivator that has been reported to interface with HIF and is frequently dysregulated in cancer is the Yes-associated protein (YAP),

which is a major effector of the Hippo signaling pathway alongside its paralog, transcriptional coactivator with PDZ-binding motif (TAZ). Originally identified in *Drosophila melanogaster*, Hippo signaling is an evolutionarily conserved pathway with a crucial role in sensing and controlling organ size and tissue homeostasis through coordinated regulation of cell growth and cell death (F. X. Yu et al. 2015). Central to this pathway is a kinase cascade, consisting of two core serine/threonine kinases - MST1/2 and LATS1/2 - and two scaffold proteins - SAV1 and MOB1A/1B. When this kinase cascade is turned on in response to various upstream signals - especially those related to cellular and tissue architecture - such as cell-cell contacts (Zhao et al. 2007) and cell polarity (Genevet et al. 2011), the activated MST1/2-SAV1 complex phosphorylates and thereby activates the LATS1/2-MOB1A/B complex, which in turn phosphorylates transcriptional coactivators YAP/TAZ. Such phosphorylation serves as an inhibitory signal, promoting YAP/TAZ nuclear export, cytoplasmic retention by 14-3-3 proteins and subsequent proteasomal degradation. When this kinase cascade is switched off, YAP/TAZ accumulates in the nucleus, where it can interact with various transcription factors, particularly the TEA domain family members (TEADs), to co-activate a broad transcriptional program for tissue homeostasis. In addition to hippo-dependent mechanisms, YAP/TAZ activity can also be regulated by hippo-independent mechanisms, such as through Rho signaling in response to extracellular matrix cues (Nardone et al. 2017).

Being two important cellular regulators in response to changes in cellular oxygen availability and structural features respectively, crosstalk has been reported between HIF and YAP/TAZ in tumor cells under hypoxia through different mechanisms. For example, using a triple-negative breast cancer cell line (MDA-MB-231), it was

found that the upstream regulator of YAP/TAZ, LATS2, is marked for proteasomal degradation by E3 ubiquitin ligase SIAH2, whose activity is enhanced under hypoxia (Ma et al. 2015). This increased destabilization of LATS2 promotes the nuclear translocation of YAP, which was found to physically interact with HIF1 α and co-localize with HIF1 α at a HRE-containing locus within the promoter region of HIF target gene *VEGFA* (Ma et al. 2015). The interaction between HIF and YAP will be further discussed in Chapter six.

1.3 HIF in cancer

Being the master regulator of cellular adaptive responses to hypoxia, HIF has important roles in cancer development, in which oxygen homeostasis is often disrupted by tumor hypoxia and various aberrant genetic and epigenetic changes, resulting in the activation of the HIF pathway. Indeed, compared to normal tissues, elevated levels of HIF1 α and/or HIF2 α have been observed (Zhong et al. 1999; Talks et al. 2000) and are correlated with adverse prognosis in many types of human cancers (Giatromanolaki et al. 2001; Griffiths et al. 2007; Koukourakis et al. 2002; Nanni et al. 2009; Schmitz et al. 2009; Sun et al. 2007; Theodoropoulos et al. 2005; Yamamoto et al. 2008). Such correlation may be underlay by the fact that a number of HIF target genes have major roles in multiple hallmark features of cancer, including cancer stem cell maintenance, epithelial-mesenchymal transition, invasion and metastasis, evasion of anti-tumor immune response, resistance to chemo- and radiotherapy (J. Chen et al. 2014; Moeller et al. 2007; Xiong et al. 2018; Xu et al. 2019), as well as the metabolic reprogramming and neovascularization mentioned earlier in subsection 1.2.5.

However, despite these apparent oncogenic associations, whether the activation of HIF activity has an overall tumor-promoting, neutral, or even tumor suppressive outcome seems to be highly context-dependent. For example, in clear cell renal cell carcinoma (ccRCC) where HIF2 α is identified as a major oncogenic driver, HIF1 α plays an antitumorigenic role (Kondo et al. 2003; Raval et al. 2005; Salama et al. 2015; Shen et al. 2011), whereas in colon carcinoma, the roles of HIF1 α and HIF2 α are reversed (Imamura et al. 2009). The divergent functional consequences of HIF α isoforms in different cancer types may be due to cell-type- and HIF-isoform-specific differences in the transactivated HIF target genes (C.-J. Hu et al. 2003; Raval et al. 2005; Salama et al. 2015; Shen et al. 2011; Smythies et al. 2019). Given the importance of HIF target gene selectivity in cancer development, a better understanding of how this selectivity is regulated is needed and will be further explored in this study.

1.3.1 Oxygen-independent regulation of HIF α activity in cancer

As well as oxygen-dependent regulation of HIF activity via the PHDs and FIH due to the depletion of oxygen inside tumors, the activity of HIF α can also be regulated in an oxygen-independent manner.

1.3.1.1 Crosstalk with tumor suppressive and oncogenic pathways

HIF activity can be modulated as a result of its crosstalk with some tumor suppressive and oncogenic pathways that are frequently altered in cancers. p53, for example, is one of the most important proteins involved in tumor suppression and its encoding gene, *TP53*, is the most frequently mutated gene in human

cancers (Kandoth et al. 2013). In response to various stress signals, such as hyperproliferative stimuli and genotoxic insults, p53 can be induced and activated to orchestrate a wide range of antiproliferative transcriptional programs including cell cycle arrest, senescence, and apoptosis, to stop the proliferation of aberrant cells (Biegging et al. 2014). Hypoxia has also been demonstrated to induce p53 in a HIF1 α -dependent manner (An et al. 1998), although conflicting reports suggest that p53 stabilization under hypoxia may also depend on other factors such as the severity and duration of hypoxia (Hammond et al. 2002), the co-existence of other stress signals (Kaluzová et al. 2004), and the phosphorylation status of HIF1 α (Suzuki et al. 2001). The functional consequence of HIF-p53 crosstalk remains unclear: some reports have shown that p53 can act as a negative regulator of HIF1 α by promoting its degradation (Choy et al. 2009; Ravi et al. 2000) or competing for common transcriptional coactivator p300 (Blagosklonny et al. 1998; T. Schmid et al. 2004), whereas a more recent report has demonstrated that p53 can promote the DNA binding of HIF1 α at the HRE loci of several HIF target genes and facilitate their transcription under hypoxia (Madan et al. 2019). Given the complexity of these two pathways and the differences in the experimental setup in these studies, whether p53 has an overall promoting or inhibitory effect on HIF activity in human cancers is likely to be context-dependent.

1.3.1.2 Loss of function of pVHL

In addition to signaling crosstalk, HIF can be activated in cancers due to the inactivation of other components in the HIF signaling pathway. A prime example is pVHL loss of function (LOF), which is found in the majority of ccRCC as a result

of biallelic inactivation of *VHL* due to somatic mutation, hypermethylation, or loss of heterozygosity. pVHL LOF abrogates the proteolytic regulation of HIF α and results in its constitutive stabilization and enhanced transactivation activity. This is supported by the observation that pVHL loss in two human renal carcinoma cell lines, RCC4 and RCC10, led to aerobic glycolysis, reduced mitochondrial biogenesis and cellular respiration that could be rescued by the re-introduction of pVHL or knocking down the expression of HIF1 α and HIF2 α (Zhang et al. 2007).

1.3.1.3 Inactivation of the PHDs by TCA intermediates

PHDs represent another key component of the HIF pathway that becomes dysregulated in certain types of cancers. Unlike pVHL, inactivating mutations of PHDs' encoding genes - *EGLN1*, *EGLN2*, *EGLN3* - have rarely been reported in cancers (Astuti et al. 2011). However, since 2-OG is an obligatory substrate for the negative regulation mediated by PHDs on HIF α (Jaakkola et al. 2001), these dioxygenases are therefore sensitive to aberrant signals from the TCA cycle, in which 2-OG is an intermediate product. Examples are inactivating mutations of succinate dehydrogenase (SDH) in paraganglioma and pheochromocytoma and fumarate hydratase (FH) in leiomyomatosis and renal cell carcinoma (Pollard et al. 2003), both functioning in the TCA cycle by catalyzing successive steps to convert succinate into fumarate and the hydration of fumarate to malate. Succinate and fumarate accumulate in *SDH*- and *FH*-mutant tumors, resulting in the inhibition of PHDs activity by out-competing 2-OG for PHDs binding, leading to constitutive activation of HIF activity (Koivunen et al. 2007; Selak et al. 2005).

1.3.1.4 The influence of FIH on HIF in cancer

Like PHDs, genetic alterations of FIH's encoding gene, *HIF1AN*, are not frequently observed in cancers [Figure 1.2] and in contrast to PHDs, FIH is not susceptible to SDH and FH deficiencies as it is not inhibited by fumarate and succinate (Koivunen et al. 2007).

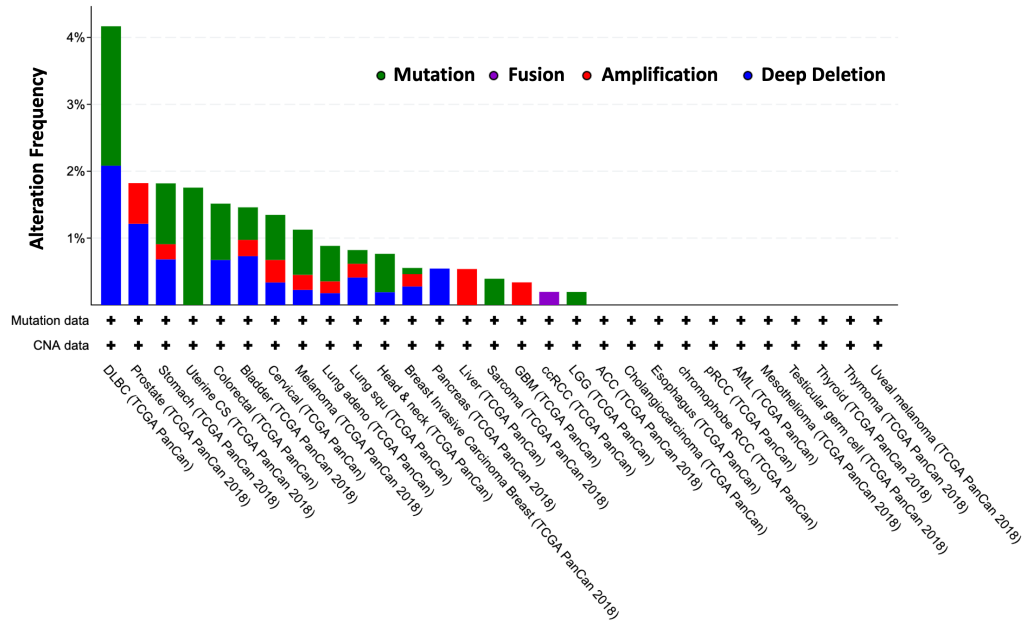


Figure 1.2: Genetic alterations events of *HIF1AN* in human cancers. Bar graph showing the frequency of genetic alterations of *HIF1AN* in 9661 patients in the TCGA PanCancer cohort. Information and bar graph were obtained from cBioPortal (<http://cbioportal.org>) (Cerami et al. 2012).

However, FIH can still be inhibited by TCA intermediates oxaloacetate and citrate, therefore any dysregulation of the TCA cycle that causes the accumulation of oxaloacetate or citrate (Peng et al. 2019) in cancers still has the potential to influence FIH activity, thereby modulating HIF transcriptional activity.

Apart from the TCA cycle, oxidative stress, a common phenomenon that tumor cells encounter due to prolonged hypoxia (Guzy et al. 2005), reoxygenation (Moeller et al. 2004), and mitochondrial dysfunction (Vyas et al., n.d.), may also influence

HIF activity through FIH. Compared to PHDs, FIH has also been found to be much more sensitive to ROS hydrogen peroxide, which inhibits the catalytic activity of FIH on HIF1 α Asn-803 and potentiates the expression of a subset of HIF target genes under hypoxia in a FIH-dependent manner (Norma Masson et al. 2012).

1.3.1.5 ARD-containing proteins

HIF α , however, are not the only substrates for FIH. Studies using different approaches to study the interactome of FIH, including yeast two-hybrid screen (Cockman et al. 2006), affinity purification-mass spectrometry following the use of small-molecule FIH inhibitor to stabilize FIH-substrate interactions (Cockman, Webb, Kramer, et al. 2009; Coleman et al. 2007), and *in silico* prediction based on sequence homology of pre-existing FIH substrates (Schmierer et al. 2010), have consistently identified a group of alternative FIH substrates sharing the ankyrin repeats (AR) domain (ARD). AR is a widely existing motif of 33 amino acids, named after the cytoskeletal protein ankyrin, which contains 24 copies of this repeat (J. Li et al. 2006; Voronin et al. 2008). Structurally, an AR has a helix-turn-helix conformation, where two α -helices run in an anti-parallel fashion and are connected by a hairpin-like β -sheet (J. Li et al. 2006; Voronin et al. 2008). ARD-containing proteins can contain a variable number of ARs, which stack in tandem to form a stably-folded domain involved in protein-protein interactions (J. Li et al. 2006; Voronin et al. 2008).

Reports to date have identified over 20 ARD-containing proteins to be hydroxylated by FIH (Cockman, Webb, and Ratcliffe 2009). They include a number of proteins functioning in various important signaling pathways in normal development

and tumorigenesis, such as Notch1-3 in Notch signaling, p105 in nuclear factor kappa-light-chain-enhancer of activated B cells (NFkB) signaling, and Apoptosis-stimulating protein of p53 2 (ASPP2) in p53 signaling (Cockman et al. 2006; Coleman et al. 2007; Janke et al. 2013; Leissing 2017; Wilkins et al. 2009; Zheng et al. 2008). However, the functional consequences of FIH hydroxylation on these ARD-containing proteins have remained largely unclear. One postulation is that some ARD-containing FIH substrates may be capable of fine-tuning HIF-mediated transcriptional responses by sequestering FIH away from HIF. This hypothesis was supported by several lines of evidence. First, it was reported that endogenous FIH hydroxylation in several of its substrates, including Notch1, MYPT1, and I κ B, are incomplete (Cockman et al. 2006; Coleman et al. 2007; James D Webb et al. 2009b), suggesting that FIH activity may be limiting *in vivo*. Second, FIH substrates bind to FIH with variable affinities and some ARD-containing substrates, as shown by studies on the Notch family members (Coleman et al. 2007; Koivunen et al. 2004; Wilkins et al. 2009; Zheng et al. 2008), may have a higher binding affinities for FIH than that of HIF α . Third and most importantly, ectopic expression of several ARD-containing FIH substrates have been shown to enhance HIF transcriptional activity by competing with HIF α transactivation domain for FIH hydroxylation (Shin et al. 2009; James D Webb et al. 2009b; Zheng et al. 2008). Given also that FIH binds to its substrates more strongly in their unhydroxylated form compared to their hydroxylated form (Cockman, Webb, Kramer, et al. 2009; Cockman, Webb, and Ratcliffe 2009; Schmierer et al. 2010), these above findings suggest that unhydroxylated ARD-containing FIH substrates may be able to modulate HIF activity by serving as a "sink" for FIH binding and hydroxylation. However,

ARD-containing proteins tend to have long half-lives (X. Cai et al. 2006) and reversal mechanisms for asparagine hydroxylation have not yet been identified. Moreover, not all ARD-containing FIH substrates are able to compete with HIF α transactivation domain for FIH binding (James D Webb et al. 2009b). This thus calls into question the relevance of ARD-containing FIH substrates as a regulator of HIF activity and it is likely that only some of them have the potential to modulate HIF activity in cancer in a context-dependent manner. One such potential candidate is ASPP2, which is involved in several important signalling pathways that are often dysregulated in cancer.

1.4 ASPP2

1.4.1 ASPP protein family

ASPP2 belongs to the ASPP protein family that consists of three members - ASPP1, ASPP2 and inhibitor of ASPP (iASPP) (Daniele Bergamaschi et al. 2003b; Samuels-Lev et al. 2001) - that are encoded by three different genes on separate chromosomes. The ASPP family is conserved in higher animals and all three members share common structural features of ARD, SH3 domain and proline-rich region at their C termini (Sullivan et al. 2007), but differ in their N termini, where ASPP1 and ASPP2 have an evolutionarily conserved RAS-association domain that is not found in iASPP (Tidow et al. 2007a; Y. Wang et al. 2013) [Figure 1.3]. These structural domains allow ASPP2 to bind to a number of oncoproteins and tumor suppressors (discussed further below) to regulate diverse signaling pathways in cancer development with functional consequences that are yet to be fully explored and understood.

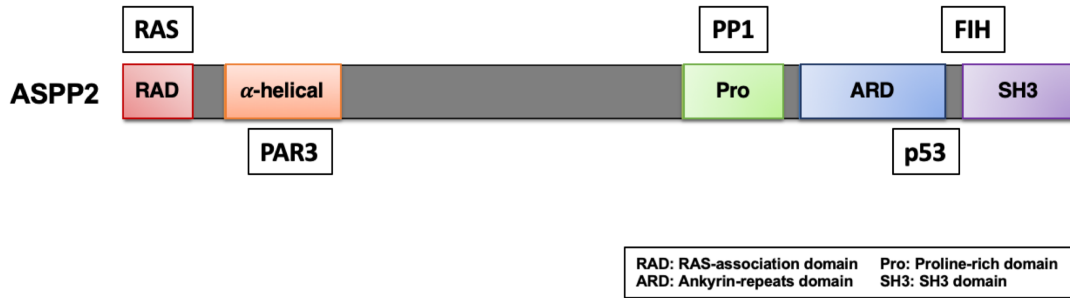


Figure 1.3: Major structural domains of ASPP2. Schematic showing the major structural domains of ASPP2. Some of the major interactors with reported binding sites on ASPP2 are shown in brackets.

All three ASPP proteins have been identified as important regulators of the p53 family proteins. While iASPP is a negative regulator of p53 by selectively inhibiting the transactivation of p53's pro-apoptotic target genes (Ahn et al. 2009; Daniele Bergamaschi et al. 2003a; Daniele Bergamaschi et al. 2006; S. Chen et al. 2019), ASPP1 and ASPP2 are activators of p53 and selectively promote its proapoptotic activity (discussed further below) (D. Bergamaschi et al. 2004).

1.4.2 Expression and subcellular localization of ASPP2

ASPP2 is encoded by *TP53BP2* mapped to chromosome 1q42.1 and consists of 1128 amino acids. It is ubiquitously expressed in most human tissues [Figure 1.4] and generally has a cytoplasmic and junctional localization (W. Cong et al. 2010; Kuniyoshi Iwabuchi et al. 1998; Royer et al. 2014; Sottocornola et al. 2010). However, the ARD domain of ASPP2 contains a nuclear localization signal (Sachdev et al. 1998) and the nuclear localization of ASPP2 has also been observed in first trimester placentas (Mak et al. 2013) and LPS-induced murine macrophages (Turn-

quist et al. 2014), suggesting that the localization of ASPP2 is context-dependent.

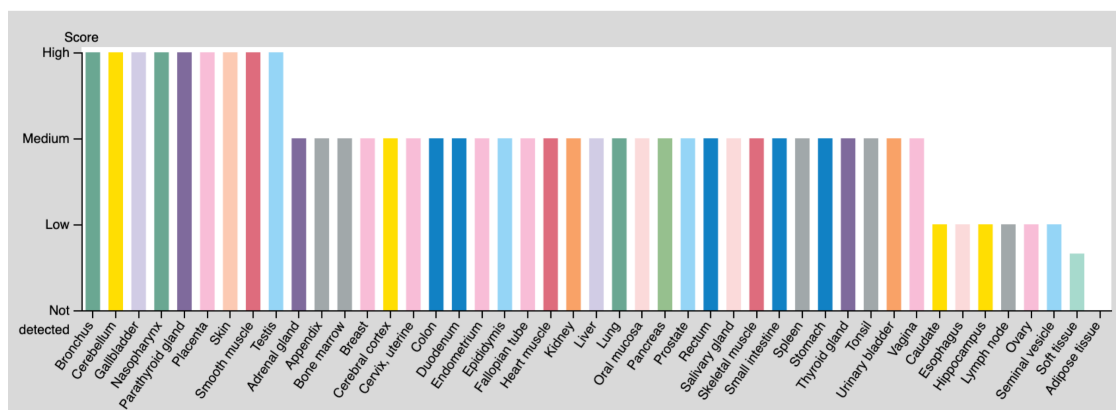


Figure 1.4: ASPP2 protein expression in human tissues. Bar graph showing the protein expression of ASPP2 in 44 human tissues. Data is obtained from Human Protein Atlas (Uhlen et al. 2015) (<https://www.proteinatlas.org/ENSG00000143514-TP53BP2/tissue>). Image credit: Human Protein Atlas.

1.4.3 ASPP2 as a regulator of p53 family proteins

Initially, a shorter form of the protein called p53-binding protein 2 (53BP2) was identified as an interactor of tumor suppressor p53 in a yeast two-hybrid assay (Iwabuchi et al. 1994). Crystallographic studies of the interactions of ASPP2-p53 complex have shown that the ARD and SH3 domain of ASPP2 bind to the DNA-binding domain of p53 (Gorina et al. 1996). Using luciferase assays as a measure of the transactivation activity of p53, it was found that ASPP2 selectively enhanced p53 transactivation function at the promoters of pro-apoptotic target genes *BAX*, *TP53I3*, and *BBC3*, but not cell cycle arrest-related target gene *CDKN1A* or negative regulator *MDM2* (Samuels-Lev et al. 2001). ASPP2 was subsequently found to interact also with p53 family members p63 and p73. Using luciferase-based techniques in a similar manner to the above p53 studies, it was found that ASPP2

also promotes the transactivation activity of p63 and p73 at the promoters of BAX, PIG3, and PUMA, but not p21 or MDM2 (D. Bergamaschi et al. 2004). These studies support the observation that ASPP2 enhanced damage-induced apoptosis in various cultured cell lines (Lopez et al. 2000; Samuels-Lev et al. 2001; Zhu et al. 2005) and suggest that ASPP2 can play a tumor suppressive role through promoting the pro-apoptotic activity of p53, p63, and p73, although how ASPP2 mechanistically mediates these selective effects and how ASPP2 may affect other transcriptional targets of p53, p63, and p73 have remained unclear.

1.4.4 ASPP2 in development and tumorigenesis

Early studies using *in vivo* mouse models shed more light on the role of ASPP2 in normal and tumor development. Two major approaches of mouse models targeting ASPP2 have been reported. Vives *et al.* targeted ASPP2 expression by deleting the exon 3 of its encoding gene *Trp53bp2*, resulting in a frame-shift mutation in the open reading frame and loss of protein expression (Vives et al. 2006). Using this approach on mice in a 129SvJxC57BL/6 genetic background, homozygous deletion of ASPP2 exon 3 resulted in postnatal lethality around weaning, which was likely caused by hydrocephalus and cardiac abnormalities upon histological examinations (Vives et al. 2006). Kampa *et al.*, on the other hand, depleted ASPP2 by replacing *Trp53bp2* exon 10-17 with neomycin resistant gene and the resulting ASPP2-null mice were inviable due to embryonic lethality (Kampa et al. 2009). Despite the differences in the timing of the lethality, which could be due to the different deletion strategies, these findings suggest that ASPP2 has important roles during normal embryonic development.

While ASPP2 homozygous deletion results in embryonic/postnatal lethality, ASPP2 heterozygosity in aging mice led to accelerated spontaneous and γ -irradiation-induced tumor development (Kampa et al. 2009; Vives et al. 2006). Since ASPP2 loss of heterozygosity was not observed in the resulting tumors, this suggests that ASPP2 has a tumor suppressive role in a haploinsufficient manner during the tumor initiation stage. Potential genetic cooperation between ASPP2 and p53 in tumor development was also studied by Vives et al. and Kampa et al. by crossing p53 heterozygous mice with ASPP2 heterozygous mice (Kampa et al. 2009; Vives et al. 2006). Of note, combined heterozygosity of ASPP2 and p53 was only found to accelerate the development of lymphoma at 42 weeks but not the development of sarcoma and lymphoma at 72 weeks (Kampa et al. 2009; Vives et al. 2006), suggesting that ASPP2 is capable of mediating p53-dependent and p53-independent tumor suppressive functions in a context-specific manner.

Interestingly, while deletion of ASPP2's encoding gene, *TP53BP2*, is rare in human cancers, amplification events of *TP53BP2* are observed in a subpopulation of several cancer types, notably in breast invasive carcinoma, where *TP53BP2* is amplified in 93 out of 1084 patients in the TCGA breast cancer cohort [Figure 1.5]. The functional implication of such amplification events, however, remains unclear. Mutation of ASPP2 is also observed in a small subpopulation of melanoma, stomach, and bladder cancer patients but effects of these mutational events on the function of ASPP2 also remain unknown [Figure 1.5].

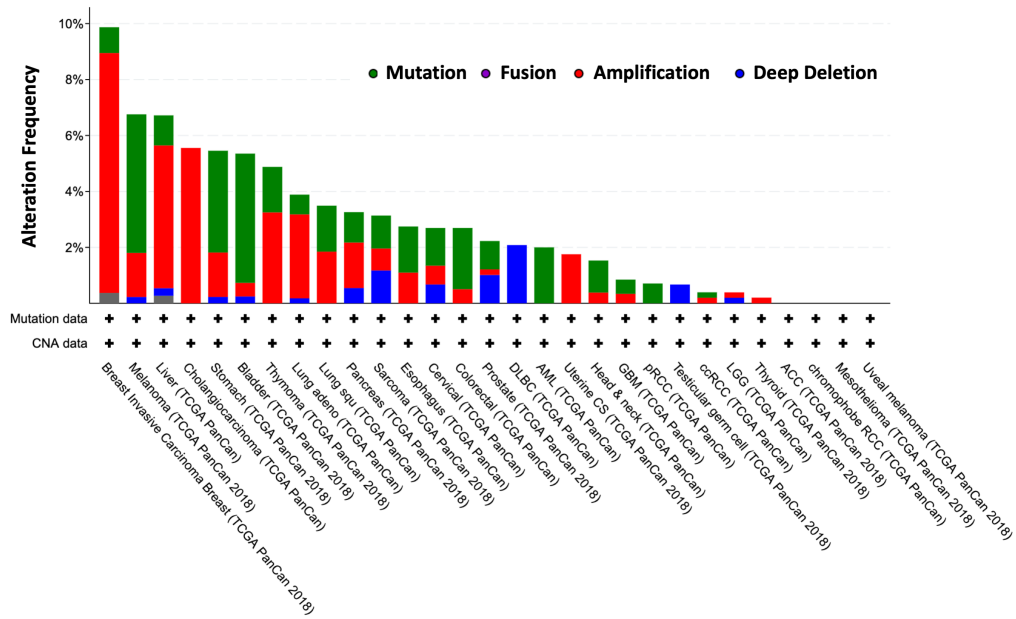


Figure 1.5: Genetic alterations events of *TP53BP2* in human cancers. Bar graph showing the frequency of genetic alterations of *TP53BP2* in 9661 patients in the TCGA PanCancer cohort. Information and bar graph were obtained from cBioPortal (<http://cbioportal.org>) (Cerami et al. 2012).

1.4.5 ASPP2 p53-independent interactions

In addition to p53, ASPP2 has been reported to interact with several other important regulatory proteins, such as PAR3, RAS, YAP/TAZ, and FIH.

1.4.5.1 ASPP2 as a regulator of cell polarity

Studying ASPP2 exon3-deleted mouse embryos revealed that the loss of ASPP2 led to a disorganized neuroepithelium, among a few other gross CNS abnormalities (Sottocornola et al. 2010). This led to the hypothesis that ASPP2 may be involved in the regulation of cell polarity and it was subsequently found that, using canine MDCK cells, ASPP2 interacts and colocalizes with PAR3 at the cell-cell junction in an interdependent manner to regulate epithelial cell polarity (W. Cong et al. 2010; Sottocornola et al. 2010). Such findings are consistent with the role of *Drosophila* ASPP, which is homologous to human ASPP1/2, as

a non-core component of the adherens junctions (AJs) and a regulator of AJs remodelling during *Drosophila* retinal development (Langton et al. 2007; Langton et al. 2009). The above phenotypic and mechanistic studies raise an important, yet unanswered question about the precise role(s) and relevance of ASPP2 in the regulation of cell polarity in solid tumors, where dysregulated cell polarity is one of the hallmark features. Interestingly, ASPP2 has also been found to be one of the major cellular targets of viral proteins produced by certain cancer-causing pathogens, including CagA from *Helicobacter pylori* and the core protein from Hepatitis C virus, to suppress the apoptotic function of p53 (Cao et al. 2004; Buti et al. 2011). These viral proteins are also known to target cellular polarity regulators to disrupt host cell polarity (Awad et al. 2013; Saadat et al. 2007), which partially contributes to their oncogenic potential. Whether these viral proteins subverts host cell polarity through their interaction with ASPP2 is another important question that awaits further elucidation.

1.4.5.2 ASPP2 as a regulator of RAS activity

Apart from cell polarity, ASPP2 is involved in the regulation of RAS signalling, which is often dysregulated in the tumorigenic process as a result of activation mutations or overexpression of the RAS proto-oncogenes, as well as constitutive activation of upstream activators, such as the epidermal growth factor receptor family proteins (Downward 2003; Buday et al. 2008). Normal cells respond to such aberrant activation of RAS signalling by activating the senescence program, which refers to a permanent state of cell cycle arrest and represents one of the major barriers to suppress or slow the development of cancer (Bartkova et al. 2006;

Ferbeyre et al. 2002; A. C. Lee et al. 1999; Serrano et al. 1997). It was found that RAS-induced senescence was in part mediated by ASPP2, which interacts with activated RAS at the plasma membrane through its RAS-association domain at the N-terminal (Y. Wang et al. 2013; Z. Wang et al. 2013; Tidow et al. 2007b). This interaction potentiates downstream Raf-ERK activity, but instead of promoting more proliferation, it results in the induction of senescence (Z. Wang et al. 2013). RAS activation also promotes the interaction between ASPP2 and p53, leading to an increased p53 transcriptional and pro-apoptotic activity (Y. Wang et al. 2013). Of note, ASPP2 has also been shown to mediate RAS-induced senescence through other mechanisms in a p53-independent manner, such as by suppressing RAS-induced autophagy (Yihua Wang et al. 2012) or RAS-induced SUMOylation on cyclin D1 (X.D. et al. 2011). These findings shed further light on the tumor suppressive functions of ASPP2, in part by mediating multiple mechanisms that connect aberrant growth signals to senescence induction in order to terminate unrestrained proliferation of pre-neoplastic cells. However, it has been shown that tumor cells may be able to overcome senescence through different mechanisms, such as by inactivating p53 or upregulating the activity of H3K9 demethylases, and that tumor cells that have escaped senescence possesses higher tumor-initiation potential compared to their never-senesced counterparts. It is therefore tempting to speculate whether the functions of ASPP2 in the regulation of senescence can be a double-edged sword by promoting RAS-induced senescence in pre-neoplastic cells and stimulating RAS activity to promote expansion of senescence-escaped tumor cells.

1.4.5.3 ASPP2's interaction with the Hippo signaling pathway

In addition to RAS signalling, ASPP2 is linked to the Hippo pathway, which, as discussed in the previous section, is heavily implicated in both normal physiological development and tumorigenesis. Central to the Hippo pathway is a serine/threonine kinase cascade that regulates the activity of its downstream effects, YAP/TAZ, by relaying phosphorylation signals. Phosphorylation is a dynamic post-translational modification that serves as one of the most important regulatory mechanisms and its robustness lies in the readily reversible nature of phosphate groups, mediated by the antagonistic actions of protein kinases and phosphatases, acting as on-off switches to regulate signalling pathways in response to different stimuli. While the mechanisms of how the activation of the core Hippo serine/threonine kinase cascade eventually phosphorylates, and thereby inhibits, the activity of YAP/TAZ have been relatively well-studied, less have been known about the potential phosphatase signalling mechanisms that regulates the Hippo pathway through dephosphorylation. It was found that ASPP2 can interact with YAP and TAZ, and, importantly, induces their dephosphorylation by recruiting protein phosphatase 1 (PP1), which is a serine/threonine phosphatase responsible for the majority of the dephosphorylation activity in eukaryotic cells (C.-Y. Liu et al. 2011; Royer et al. 2014). Human genome contains 3 genes that encode 4 closely related PP1 catalytic subunits – alpha, beta, and 2 gamma splice variants. PP1 is evolutionarily conserved and all its isoforms show similar enzymatic properties and display no substrate specificity *in vitro*. However, PP1 activity *in vivo* is highly specific and regulated by over 200 regulatory proteins that interact with PP1, forming stable complexes with distinct subcellular

localization and substrate specificities (Bollen et al. 2010; Heroes et al. 2013).

Most PP1 regulatory proteins contain a consensus sequence called the RVxF motif that is required for PP1 binding (Bollen et al. 2010). Of note, the binding of RVxF motif alone to PP1 affects neither the structure nor the substrate specificity of PP1, suggesting that additional motifs on PP1 regulatory proteins are responsible for regulating PP1 substrate specificity (Bollen et al. 2010; Heroes et al. 2013). ASPP2 is one of the PP1 regulatory proteins and it has been shown both by our lab and others to form a stable complex with PP1 (Bertran et al. 2019; Egloff et al. 1997; Nicholas R. Helps et al. 1995; Royer et al. 2014). Subsequent crystallographic studies of ASPP2-PP1 complex showed that ASPP2 binds to PP1 through its RVxF motif and SH3 domain (Bertran et al. 2019). The RVxF motif is necessary for ASPP2-PP1 binding, as mutating Val-922 and Phe-924 residues within this motif to alanine residues abolished the binding between ASPP2 and PP1 (Llanos et al. 2011). Functionally, a C-terminal fragment of ASPP2 containing the RVxF motif and SH3 domain was reported to inhibit PP1 dephosphorylation on glycogen phosphorylase but not on myosin P-light chains (Nicholas R Helps et al. 2000), suggesting that ASPP2 and PP1 may together form a holoenzyme with distinct substrate specificity. Mice bearing a homozygous knock-in mutation in the ASPP2 RVxF motif, i.e. ASPP2(V922A, F924A), are embryonically lethal. These mice started to manifest signs of deformation on embryonic day (E) 9.5 and none of them were viable past E13.5 (Dr. Elizabeth Slee, Lu lab unpublished data). This suggests that ASPP2-PP1 complex has indispensable functions during embryonic developmental process. One such function may be through the dephosphorylation of YAP. ASPP2-PP1 complex was shown to mediate the dephosphorylation of Serine-127 on YAP (Royer

et al. 2014). Knocking down ASPP2 in colon adenocarcinoma cell line Caco-2 cells and the depletion of ASPP2 in mouse colonic crypts reduced nuclear localization of YAP, and in the latter case ASPP2 depletion also reduced the mRNA expression of *CCN2*, which is a YAP/TAZ target (Royer et al. 2014). Intriguingly, only the co-overexpression of YAP and ASPP2, but not the overexpression of ASPP2 alone, increased the luciferase activity of TEAD reporter (Royer et al. 2014). In a separate study, ectopically expressed ASPP2 was also found to interact with TAZ, increasing TAZ nuclear localization, while knocking down ASPP2 similarly led to a reduction in *CCN2* mRNA expression (C.-Y. Liu et al. 2011). However, the contexts and consequences of ASPP2-YAP/TAZ interactions in cancers are not yet clear.

1.4.5.4 ASPP2's interaction with HIF regulator FIH

As mentioned in subsubsection 1.3.1.5, ASPP2 is one of the recently identified ARD-containing substrates of FIH (Janke et al. 2013). Using peptides and protein constructs, Janke *et al.* showed that ASPP2 can be hydroxylated at asparagine-986 (N986) in a FIH-dependent manner (Janke et al. 2013). They further showed that knocking down FIH expression led to a mild reduction of the interaction between ASPP2 and cell polarity regulator PAR3 (Janke et al. 2013). However, whether this effect is due to a direct effect of FIH hydroxylation on ASPP2 and whether there are other biological consequences of FIH-ASPP2 interaction remain unclear.

Using recombinant protein binding and co-immunoprecipitation with endogenous proteins, unpublished work from our former lab member, Dr. Thomas Leissing, have shown that FIH can indeed interact with ASPP2 at the endogenous level and that ASPP2 contains two potential FIH-hydroxylation sites, N986 and asparagine-984

(N984) (Leissing 2017): hydroxylated N986 was detected in wild-type U2OS cells where FIH was expressed at an endogenous level, while hydroxylated N984 was detected only when FIH was overexpressed. Interestingly, Dr. Leissing also showed that ASPP2 can successfully compete with HIF α -CAD for FIH binding and reduce the asparaginyl hydroxylation of HIF α -CAD *in vitro* (Leissing 2017), thus raising the question of whether ASPP2 could be a potential modulator of HIF activity.

1.4.5.5 ASPP2 and hypoxia response

The expression of ASPP2 has been recently reported to be subject to hypoxia-specific regulation. It was found that hypoxia promotes the activity of E3 ubiquitin ligase SIAH2, which in turn targets ASPP2 for proteasomal degradation, thereby reducing ASPP2 expression under hypoxia (H. Kim et al. 2014). It was also observed that hypoxia can upregulate the expression of microRNA-205, which targets ASPP2 and suppresses its expression (X. Wang et al. 2016).

Although the aforementioned reports revealed that hypoxia regulates ASPP2 expression, whether ASPP2 has a role to play in hypoxia, and specifically if and how ASPP2 regulates HIF signaling activity remains unknown.

Interestingly, when analyzing our group's in-house RNA-seq data from intestinal organoids that were derived from wild-type (WT) and ASPP2-exon3-deleted (ASPP2-deficient) mice, we found in an unbiased gene set variation analysis (GSVA) (Hänzelmann et al. 2013) that multiple hypoxia response gene signatures displayed differential expression patterns between WT and ASPP2-deficient conditions [Figure 1.6]. Since many of the genes that constitute these hypoxia response signatures are HIF target genes and HIF is a major regulator that orchestrates

hypoxia response, this observation prompted us to investigate the potential role of ASPP2 in HIF signaling activity.

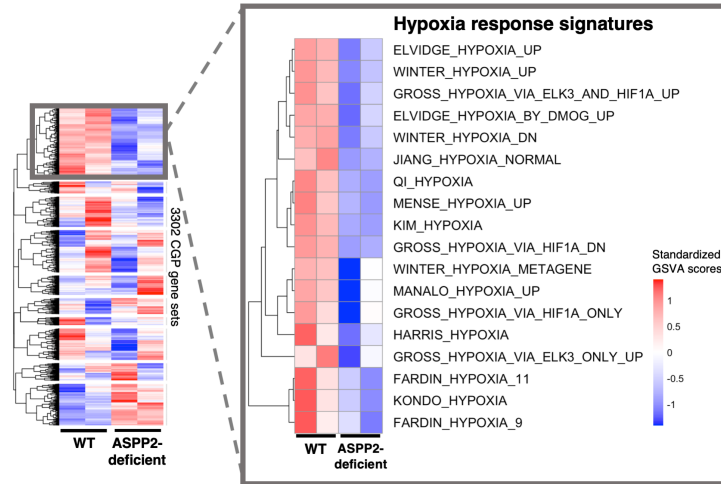


Figure 1.6: Multiple hypoxia response signatures display differential expression patterns between wild-type and ASPP2 Δ^3/Δ^3 mouse intestinal organoids. Heatmaps showing results of GSEA for indicated gene sets for biological duplicate wild-type and ASPP2 Δ^3/Δ^3 mouse intestinal organoids grown in normoxia. All 3297 gene sets from the Molecular Signature Database chemical and genetic perturbation collection were used for GSEA analysis.

1.5 Aims of the study

The goal of the work described in this thesis is to explore the role of ASPP2 in HIF signalling in cancer with a view to providing insights of how ASPP2, which interacts with and regulates multiple oncoproteins and tumor suppressors implicated in cancer development, may contribute to the context-dependent regulation of HIF activity in cancers.

Based on the observation that (i) ASPP2 can compete with HIF α -CAD for FIH hydroxylation in *in vitro* competition assay and (ii) the differential expression patterns of multiple hypoxia response gene signatures between wild-type and

ASPP2 Δ^3/Δ^3 mouse intestinal organoids, we hypothesized that ASPP2 could regulate HIF signaling activity.

This hypothesis is addressed in this work, highlighted by a transcriptomic-based approach to first understand the effects, if any, of ASPP2 on HIF transcriptional activity. This is followed by mechanistic studies that employ a combination of cell-based molecular biology techniques, *in silico* gene coexpression study, and multi-omics analyses to identify potential underlying regulatory mechanisms, focusing on four major points of regulation that influence HIF activity [Figure 1.7], namely, the expression of HIFs, the DNA-binding activity of HIFs, the recruitment of HIF-associated transcriptional cofactors, and the crosstalk of HIFs with other signaling pathways.

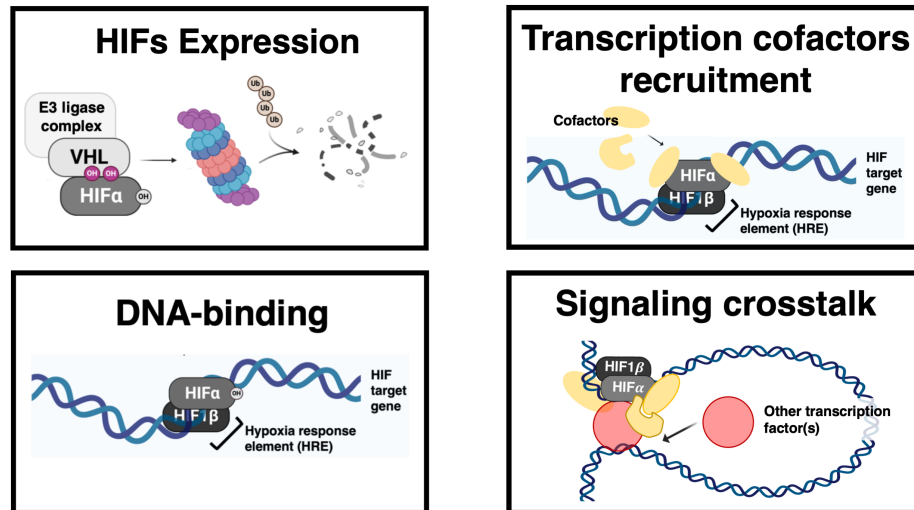


Figure 1.7: Major points of regulation that influence HIF signaling activity. Schematic showing four major points of regulation that can influence HIF signaling activity, including the expression of HIFs, DNA-binding activity of HIFs, recruitment of transcription cofactors by HIFs, and crosstalk between HIFs and other signaling pathways.

Whether and how ASPP2 could regulate HIF signaling activity are addressed in

this work by a combination of cell-based molecular biology techniques, as well as *in silico* gene coexpression study and multi-omics analyses, with an aim to provide insights of how signaling crosstalk with ASPP2-regulated pathways may contribute to the context-dependent regulation of HIF activity in cancers.

2

Materials and Methods

2.1 General

This chapter provided an overview of the methods and materials used in this study. If not otherwise stated, all chemicals and reagents were ordered from Sigma-Aldrich, oligonucleotides were purchased from Eurofins Scientific, and enzymes and buffers used in microbiological techniques were from New England Biolabs. All plasmids listed in Table 2.10 were sequenced before use.

2.2 Molecular biology techniques

2.2.1 Plasmid preparation

Purification of plasmid DNA was performed using QIAprep Spin Miniprep Kit (Qiagen) or GeneElute HP Plasmid Maxiprep Kit (Sigma-Aldrich) according to the manufacturers' protocols. Purified plasmids were eluted using ultra-pure water and stored at -20°C.

Step	Temp	Time
Initial denaturation	98°C	30 sec
30 cycles	98°C	5 sec
	60°C	30 sec
	72°C	20 sec
Final extension	72°C	2 min
Final hold	4°C	∞

Table 2.1: PCR protocol for plasmid cloning

2.2.2 Agarose gel electrophoresis

Nucleic acids were purified using agarose gel electrophoresis. 0.8-1% agarose (molecular grade, Bioline) was heated in TAE buffer (40 mM Tris-Acetate pH = 8.0, 1 mM EDTA) containing ethidium bromide (1:50,000 dilution) or SYBR Safe DNA Gel Stain (1:2000 dilution, Sigma-Aldrich). DNA samples were mixed in 4x loading buffer (Fermentas) before loading on the gel. Electrophoresis was performed at 100 V for 45-60 min in TAE buffer. GeneRuler 1 kb DNA ladder (Fermentas) was used as a marker and DNA bands were visualised with Ultraviolet (UV) light (254 nm).

2.2.3 Molecular subcloning

PCR was performed to amplify ASPP2 cDNA from V5-pcDNA3.1 plasmid following the protocol outlined in Table 2.1 using Q5 High Fidelity DNA polymerase and primers listed in Table 2.6. The primers were designed to introduce XhoI and ClaI restriction enzyme sites flanking the insert.

Amplified insert was purified using agarose gel electrophoresis. The desired band was excised and DNA extracted using the gel extraction kit (QIAGEN) according to the manufacturers' protocol. Purified insert and target piggyBac vector, pPB-MCS-

FLAG-Puro, were digested using restriction enzymes XhoI and ClaI according to the manufacturer's protocol. Digested insert and target vector were used for agarose gel electrophoresis and gel-purified using Qiagen Gel Extraction Kit (Qiagen).

Ligation was performed using T4 DNA ligase according to the manufacturer's protocol and 1 μ L of the ligation reaction was used for bacterial transformation.

Sequencing primers for the resulting piggyBac plasmid were listed in Table 2.7.

2.2.4 Bacterial culture

Bacterial cultures were grown in pre-autoclaved lysogen broth medium (LB, 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) with ampicillin (1:1000 dilution). Antibiotic stock solutions (1000x) of ampicillin (50 mg/mL in water) were filter-sterilized and stored at -20 °C until use.

2.2.5 Bacterial transformation

The plasmid DNA of interest was transformed into α -select electro-competent cells, bronze efficiency (Bioline) according to the manufacturers' protocol and plated on LB agar (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, 20 g/L agar) supplemented with appropriate antibiotics. Bacterial colonies were cultured overnight at 37 °C and subsequently at 4 °C for two weeks if not immediately used.

2.2.6 Glycerol stocks

For long term storage, bacterial cultures were grown to stationary phase and mixed with 1 volume of 2x glycerol buffer (65% glycerol, 0.1 M MgSO₄, 45 mM Tris-HCl pH = 7.5) and snap-freezed and stored at -80°C.

2.3 Tissue culture and cell-based techniques

2.3.1 Maintenance and passage of cell lines

All cell culture were carried out using standard sterile techniques for human and mouse tissue culture. ASPP2-Knockout (KO) AGS cells were generated by Dr. Ludovico Buti, a former lab post-doc. FIH-KO U2OS cells were kindly provided by Dr. Ya-Min Tian (Ratcliffe group, University of Oxford). Primary astrocytes from wild-type (WT) and ASPP2-KO mice as a result of ASPP2 exon 3 deletion ($\Delta 3/\Delta 3$) were kindly provided by fellow lab DPhil student Mr. Jie Yin.

Immortalized cell lines and primary astrocytes were cultured in DMEM supplemented with 10% (v/v) FBS, L-glutamine, and penicillin/streptomycin. Cells were incubated at 37 °C in the presence of 5% CO₂. Depending on the confluence, medium was changed every 2–4 days. Once confluent, cells were first washed with PBS once and incubated in Trypsin-EDTA at 37 °C until detachment. Trypsin was subsequently inactivated with culture medium containing FBS, and cells were seeded into new flasks/dishes at the desired cell density.

2.3.2 Freezing and thawing cells

For freezing cells, cells were collected by trypsinisation and centrifuged at 200 g for 5 minutes at room temperature. The supernatant was aspirated and the cell pellet was resuspended in freezing medium and transferred to a labelled cryovial (1 ml per vial). The cryovial was kept in a Mr. FrostyTM freezing container at 80°C for at least 24 hours, before being transferred to liquid nitrogen for long-term storage.

For cell thawing, cryovials taken from liquid nitrogen were incubated at 37°C until the medium became liquid. The cells were transferred to a 15 ml centrifuge

tube (Corning) and centrifuged at 200 g for 5 minutes at room temperature. The supernatant was then discarded, and the cell pellet was resuspended in culture medium and seeded into new flasks/dishes at the desired density.

2.3.3 Cell transfection

Fugene 6 transfection reagent (Promega) was used for cell transfection unless otherwise stated. Transfection was performed according to the manufacturer's protocol. Downstream treatments or assays were performed 24 hours after transfection unless otherwise stated.

2.3.4 RNA extraction and reverse transcription

Cells were first lysed by directly scraping in RLT buffer (Qiagen) supplemented with freshly added 1% 2-Mercaptoethanol. Same lysis protocol was performed for cells incubated inside hypoxia workstation. RNA was then extracted using the RNeasy Mini Kit (Qiagen). The concentration and quality of RNA were measured by NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific) and samples with poor purity ratios were discarded or cleaned up with QIAquick PCR Purification Kit (Qiagen).

Reverse transcription was performed using QuantiTect Reverse Transcription Kit (Qiagen) per 1 µg of RNA samples. cDNA samples were eluted with 120 µL of nuclease-free water (Ambion) and were either proceeded with qPCR immediately or stored at 4°C within one week or long-term at -20°C.

Reagent	Per reaction (μ L)
PowerUp SYBR Master Mix	7.5
Forward primer (10 μ M)	2
Reverse primer (10 μ M)	2
cDNA	5.5

Table 2.2: qPCR preparation protocol

2.3.5 RNA samples preparation for RNA-sequencing

RNA was extracted as described in subsection 2.3.4, and QC on the TapeStation System using the High-Sensitivity RNA ScreenTape Assay according to the manufacturer's protocol. Evaluation of RNA quality was measured by the RNA Integrity Number (RIN^e). Only samples with RIN^e ≥ 9.5 were submitted for RNA-sequencing. ERCC ExFold RNA Spike-In Mixes (Ambion) were added prior to Library prep using QuantSeq Forward kit (LEXOGEN), using 500 ng of starting material to minimize the PCR amplification step. Samples were sequenced on HiSeq 4000 (Illumina) carried out using the Wellcome Trust genomic service, Oxford.

2.3.6 Quantitative reverse transcription PCR

PowerUp SYBR Green Master Mix (Thermo Fisher Scientific) was used to perform Quantitative reverse transcription PCR (qRT-PCR) using the real-time PCR cycler, the Rotor-Gene Q (Qiagen). Reactions were performed in duplicate with a final volume of 15 μ L. Reaction constituents were illustrated in Table 2.2 and the thermal cycling condition was outlined in Table 2.3.

Relative mRNA abundances were normalised to the mRNA level of endogenous *beta*-Actin. Shown qPCR results were analysed by the $\delta\delta C_t$ -method as previously described (Schmittgen et al. 2008). The threshold cycles were determined by the

Step	Temp (°C)	Time	Cycles
UDG activation	50	2 min	Hold
Polymerase activation	95	2 min	Hold
Denaturation	95	15 sec	40
Annealing/extension	60	1 min	

Table 2.3: qPCR thermal cycling condition

Qiagen Rotor-Gene Software. Analyses and visualization of the qPCR results were performed by using Prism 8 (GraphPad Software Inc.) in which the multiple t-tests function was used for determining statistical significance.

2.3.7 Small interfering-RNA-induced gene silencing

Small interfering-RNA (siRNA) oligo from Sigma were used to silence the mRNA expression of HIF1 α . U2OS cells seeded in 24-well plates at about 60% confluence were transiently transfected using DharmaFECT I Transfection Reagent (Horizon Discovery Group plc). Briefly, 25 nM of anti-HIF1 α siRNA or ON-TARGETplus Non-targeting control siRNAs (Dharmacon) were added to a total of 100 μ L of Opti-MEM Reduced Serum Medium (Gibco), and 2.5 μ L DharmaFECT I reagent to 100 μ L of Opti-MEM, in separate tubes. After incubation for 5 minutes with gentle mixing, the contents of the two tubes were combined and incubated for a further 20 minutes. This mixture was added to 400 μ L of complete cell culture medium (DMEM supplemented with 10% (v/v) FBS, L-glutamine, and penicillin/streptomycin) and the resulting transfection medium was added to each 24-well after removal of the original cell culture medium.

The siRNA targeting HIF1 α was kindly provided by Dr. Veronique Lafleur from the Ratcliffe group (University of Oxford). Its efficiency was previously assessed by the Ratcliffe lab and re-assessed by qPCR analysis as described in chapter 4.

2.3.8 CRISPR-Cas9 gene-editing

ASPP2-KO and ASPP2-FIH-Double-KO U2OS cell lines were created using WT U2OS and FIH-KO U2OS, respectively, following the protocol in Zhang *et al.* (Ran et al. 2013). Briefly, the single-guide-RNA (sgRNA) target ASPP2 exon 6 (5'- GAG CAC TTA AAG GCC ACG -3') was cloned into the CRISPR-Cas9 plasmid carrying the Cas9 nuclease (PX459) using BbsI restriction enzyme sites. The plasmid was then sequenced using hU6-F primer to confirm the incorporation of the sgRNA.

WT-U2OS and FIH-KO U2OS cells were then either co-transfected with PX459 and a GFP expression-only plasmid (pCMV-GFP), or transfected with pCMV-GFP plasmid alone. 36 hours post-transfection, GFP-positive cells were FACS-sorted and plated at low density on 15cm dish. Cell culture medium was supplemented with 8 µg/mL of puromycin to select resistant single cell clones. Single cell clones were then picked and expanded and analyzed by western blot.

The CRISPR-Cas9 plasmid, pSpCas9(BB)-2A-Puro (PX459) V2.0, was a gift from Feng Zhang (Addgene plasmid ID: 62988). The GFP-expression plasmid, pCMV-GFP, was a gift from Connie Cepko (Addgene plasmid ID: 11153). FACS sorting was performed by the Flow Sorting Facility Manager, Mr. Andrew Worth (Jenner Institute, Oxford).

2.3.9 Generation of stable cell line

To generate stable ASPP2-KO cell line expressing doxycycline-inducible ASPP2, ASPP2-KO cells were co-transfected with the piggyBac transposon vector carrying ASPP2 cDNA downstream of the Tet-ON promoter, the piggyBac transposase vector (Transposagen) and the piggyBac vector expressing reverse tetracycline

transactivator (rtTA) from a CAG constitutive promoter and neomycin resistance gene [Table 2.10]. The transfection was performed using Lipofectamine LTX with PLUS Reagent (Thermo Fisher Scientific).

24 hours after transfection, the cells were selected with neomycin (Gibco) at a concentration of 700 $\mu\text{g/ml}$ for one week and the cells that survived were single-cell sorted into 96-well plates. FACS sorting was performed by Mr. Andrew Worth (Jenner Institute, Oxford). Single cell clones were then expanded and splitted in a 1:2:2 ratio with 20% seeded on a normal 96-well plate and 80% seeded on two separate NuncTM MicroWellTM 96-Well Optical-Bottom Plates with Polymer Base (Thermo Fisher Scientific) for immunofluorescence screening. The two optical-bottom 96-well plates were treated with or without doxycycline hydrochloride solution (Sigma) at 50 ng/mL for 16 hours. Immunofluorescence was then performed as described in subsection 2.3.11 using anti-FLAG antibody (Sigma) as the primary antibody. The 96-well Optical-Bottom Plates were then screened using under a confocal microscope system (LSM 710; Carl Zeiss) using a macro developed by the Microscope Manager, Dr. Richard Lisle (Ludwig Institute, Oxford). Clones that showed positive FLAG-tag staining were selected for further characterization by western blot as described in chapter 4 and sequencing as described in subsection 2.3.10.

The original piggyBac vectors were from Dr. Kazuhiro Murakami (RIKEN, Kobe, Japan) and were kindly provided by Dr. Pakavarin Louphrasitthiphol from the Goding lab (Ludwig Institute, Oxford).

2.3.10 PCR and Sanger sequencing

Genomic DNA was isolated from clonally expanded CRISPR-Cas9-modified candidate ASPP2-KO and ASPP2-FIH-DKO U2OS cells using GenElute Mammalian Genomic DNA Purification Kit according to the manufacturer's protocol. A 270-bp segment containing the ASPP2 sgRNA target site was amplified by PCR using Phusion High-Fidelity DNA polymerase using primers ASPP2-Fw1 and ASPP2-Rev1 [Table 2.6].

Same procedures were performed for candidate clones of stable ASPP2-KO cell lines transfected with the piggyBac vector, pPB-MCS-FLAG-Puro, with a Tet-On system and ASPP2 cDNA incorporated. A 779-bp segment containing the ASPP2 RVxF motif and the C-terminally-tagged 3X-FLAG sequence was amplified by PCR using primers ASPP2-RVxF-F and ASPP2-RVxF-R [Table 2.6].

The PCR amplicon was then purified using QIAquick PCR Purification Kit (Qiagen) and subcloned into a TOPO TA vector using TOPO TA Cloning Kit (Invitrogen). Random colonies were picked and submitted for Sanger Sequencing using T7 promoter and M13 Reverse sequencing primers [Table 2.7].

2.3.11 Immunofluorescence microscopy

For immunofluorescence experiments, coverslips (VWR, round 13 mm) were placed in flat-bottom multi-well cell culture plates and the plates were sterilized by UV irradiation in UV Crosslinker (Hoefer Scientific) for 2 minutes. Cells were seeded onto these pre-sterilized coverslips in multi-well plates. Downstream treatments or assays were performed 12 hours after seeding unless stated otherwise.

At designated timepoints, cells on the coverslips were washed with PBS and fixed in 4% paraformaldehyde in PBS for 15 minutes at room temperature. For cells incubated in hypoxia workstation, they were washed with PBS pre-equilibrated inside the hypoxia workstation for at least 24 hours at the designated oxygen level and fixed in 4% paraformaldehyde in PBS for 10 minutes inside the hypoxia workstation. The samples were then washed 3 times with PBS and either stored at 4°C for use within one month or proceed immediately to permeabilization by 0.2% Triton X-100 in PBS for 10 minutes at room temperature, then blocked with Image-iT FX Signal Enhancer (Thermo Fisher Scientific) for 1 hour and stained overnight at 4°C with IF antibodies diluted in Antibody Diluent, Background Reducing (Dako), at the recommended dilution level [Table 2.6]. After primary antibody incubation, the coverslips were washed 4 times with IF washing buffer (2% bovine serum albumin, 0.2% Triton X-100, 0.1% Tween-20 in PBS), before incubating with the appropriate Alexa Fluor secondary antibodies diluted in Antibody Diluent (Dako) at the recommended dilution level [Table 2.5] for 30 minutes. DAPI (1:2000 dilution) was used to stain nucleic acids. Coverslips were then washed 4 times with IF buffer before mounting on PolysineTM Microscope Adhesion Slides (Thermo Fisher Scientific) using ProLong Gold Antifade Mountant (Life Technologies Ltd). Slides were observed under a confocal microscope system (LSM 710; Carl Zeiss) and were stored for 1 week at 4°C or long-term at -20°C.

Images were acquired using Zen imaging software (Carl Zeiss) according to the manufacturer's guidelines. Quantification of nuclear and non-nuclear intensity of the microscopy images was calculated by ImageJ 1.42q software (National Institutes

of Health) using an ImageJ macro developed by former Microscopy Manager, Dr. Mark Shipman (Ludwig Institute, Oxford).

2.3.12 Western blotting

For protein expression analysis, cells seeded at similar cell densities were first washed with PBS and lysed directly on the plate by adding either ice-cold 1% NETN buffer (50 mM Tris-HCl pH = 7.4, 1% NP40, 150 mM sodium chloride, 20 mM magnesium chloride supplemented with protease and phosphatase inhibitors) or 4x SDS-PAGE loading buffer (200mM Tris-HCl pH 6.8, 8% SDS, 0.4% bromophenol blue, 40% glycerol freshly supplemented with 5% 2-Mercaptoethanol before use). For cells incubated in hypoxia workstation, cell lysis was performed inside the chamber.

Apart from samples lysed by 4x SDS-PAGE loading buffer, all lysates were put on immediately ice before further processing after being taken out of the hypoxia workstation. Apart from samples lysed by 4x SDS-PAGE loading buffer, the protein concentration of the lysates were measured as described in subsection 2.3.13, topped up to the same concentration using the corresponding lysis buffer, and boiled.

Boiled samples were run on 1.5mm NuPAGE Novex 4-12% Bis-Tris Protein Gels (Thermo Fisher Scientific) for proteins with under 200 kDa molecular weight, or 1.0mm NuPAGE 3-8% Tris-Acetate Protein Gels (Thermo Fisher Scientific) for proteins with molecular weight higher than 200 kDa. Blue Prestained Protein Standard, Broad Range (NEB) was used as the protein ladder. Electrophoresis was run in XCell4 SureLock™ Mini- or Midi-Cell Electrophoresis System (Thermo Fisher Scientific) for about 1 hour and 10 minutes using 1X NuPAGE MOPS SDS Running Buffer (Thermo Fisher Scientific).

Gels were transferred at 80V, 400A, for 3 hours at 4°C or at 20V, 400A, overnight at room temperature with a -20°C freezer block inside the transfer tank. Nitrocellulose or PVDF membranes (GE Healthcare) were blocked with 5% semi-skimmed milk in TBST buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5% Tween-20) before incubating with primary antibody in TBST buffer at the recommended dilution level [Table 2.6] overnight at 4°C or for 1 hour at room temperature. Membranes were washed three times with TBST buffer with rocking for 10 minutes each and incubated with horseradish peroxidase (HRP)-conjugated secondary antibody diluted in TBST buffer at the recommended dilution level [Table 2.5] and incubated for 40 minutes on a rocking platform at room temperature.

Membranes were washed three more times with TBST buffer with rocking for 10 minutes per wash and ECL working solution was prepared by mixing Amersham ECL Reagent 1 and Reagent 2 (GE Healthcare) at a 1:1 ratio. Membranes were incubated with ECL working solution for 2 minutes before exposing in the dark room. X-ray films scanned from western blots were only adjusted for size, orientation, brightness, or contrast. No signal was otherwise manipulated in relation to the raw image.

2.3.13 Protein concentration determination

Bio-Rad Protein Assay Dye Reagent Concentrate (Bio-Rad) was used to determine the protein concentration of cell lysates. Six serial dilutions of 2 mg/mL BSA were used as a standard curve. Dilutions of BSA or 1 µL of cell lysate were pipetted into a 96-well plate and mixed with 200 µL of 1x Bio-Rad assay reagent. The plate was read using a spectrophotometer (Anthus Labtech) at an absorbance of

595 nm. The concentration of the lysates were calculated by the Magellan Data Analysis Software (Tecan) based on the standard curve.

2.3.14 Immunoprecipitation

For immunoprecipitation (IP) experiments using whole cell lysates, cells were lysed using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Fisher Scientific) according to the manufacturer's instruction or ice-cold 1% NETN buffer (50 mM Tris-HCl pH = 7.4, 1% NP40, 150 mM sodium chloride, 20 mM magnesium chloride supplemented with protease and phosphatase inhibitors) for 45 minutes with occasional pipetting. Cell debris was pelleted at 16,000g for 15 minutes at 4 °C and the supernatant was transferred into a pre-chilled Eppendorf tube. The protein concentration of the supernatant was determined using the Bradford assay (Bio-RAD) according to the manufacturers' protocol. 5% of the supernatant was kept as input and boiled in 2X SDS-PAGE loading buffer (100mM Tris-HCl pH 6.8, 4% SDS, 0.2% bromophenol blue, 20% glycerol with 5% 2-Mercaptoethanol added freshly before use) for 5 minutes and stored at -20°C until immunoblotting. Protein lysates were incubated with IP antibodies at 4°C on an end-over-end rotator overnight, followed by 1 hour of magnetic beads capture using Dynabeads Protein G (Invitrogen). The beads were then washed three times with 1 mL ice cold 0.5% NETN buffer and boiled in 1X SDS-PAGE loading buffer for 10 minutes. The boiled samples were used for western blot analysis immediately or freezed in -20°C.

2.3.15 Chromatin immunoprecipitation

2.3.15.1 Cell fixation and harvest

For chromatin immunoprecipitation (ChIP) experiments, cells were grown to about 80% confluent in 15 cm dishes. Cross-linking was done by directly adding 540 μ L of 37 % paraformaldehyde (Molecular Biology grade) directly to the 15 cm dishes with 20 mL of medium to achieve a final concentration of 1%. For cells incubated in the hypoxia workstation, the cross-linking step was performed in the same way inside the hypoxia workstation and the 15 cm dishes were immediately taken out of the hypoxia workstation after paraformaldehyde addition. The 15 cm dishes were put immediately on ice with gentle rocking for 10 minutes.

This was then followed by quenching with the addition of 2.5 mL of 1M glycine in PBS solution to achieve a final concentration of 125 mM. The 15 cm dishes were further incubated at room temperature with gentle rocking for 10 minutes, followed by twice washing with 20 mL of ice-cold PBS solution. The adherent cells were scraped in 5 mL of ice-cold PBS and transferred into a falcon tube (Corning).

After centrifugation at 200 g at 4°C for 10 minutes, the supernatant was carefully aspirated and the cell pellet was resuspended in 500 μ L ChIP SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris pH 8.1 supplemented with 2x cOmplete Protease Inhibitor Cocktail (Roche)) per 15 cm dish, followed by 10 minutes of incubation on ice. The lysates were then diluted 1:1 with ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris pH 8.1, 167 mM NaCl).

2.3.15.2 Sonication of DNA

The diluted lysates were divided into 1 mL portions and these 1 mL portions were transferred to 15 mL TPX tubes (Diagenode) for sonication using a Bioruptor water bath sonicator (Diagenode). The water bath sonicator was pre-cooled to 4°C before use. The sonicator was set to cycles of 15-second on, 15-second off, for 20 minutes in total, on high amplitude. During sonication, Dynabeads protein G (Invitrogen) were prepared by vortexing and washing with ChIP dilution buffer twice before being put on ice.

Sonicated samples were then transferred to 1.5 mL eppendorf tubes and centrifuged at 13,000 rpm for 10 mins at 4°C. Supernatants from same conditions were then carefully transferred to new falcon tubes and 20 µL of the supernatants for each condition were transferred to 1.5 mL Safe-Lock eppendorf tubes (Eppendorf Ltd) as 2% input for downstream fragment size quality control and to process as an input sample control for IP. 2% input samples were stored at -20°C before further use.

2.3.15.3 Samples pre-clearing, IP and washes

The remaining samples were diluted 1:6 with ChIP dilution buffer to reduce the concentration of SDS to less than 1% so that it would be suitable for subsequent IP. Diluted samples were then pre-cleared with 50 µL of the washed Dynabeads protein G (Invitrogen) per 15 cm dish on an end-over-end rotator at 4°C. Pre-cleared samples were then collected by using a DynaTM-2 Magnet (Thermo Fisher Scientific). An appropriate amount of ChIP antibody was added to each sample per IP in a 15 mL falcon tube, and the mixture was incubated at 4°C for 16 hours on an end-over-end

rotator. This was followed by the addition of 20 μ L washed Dynabeads Protein G (Invitrogen) and a further incubation of 1 hour on an end-over-end rotator at 4°C.

After the 1-hour antibody capture, the supernatant was aspirated by using the DynaTM-2 Magnet (Thermo Fisher Scientific), and the beads were resuspended in 800 μ L of low salt washing buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.1, 150 mM NaCl), transferred to a new 1.5 mL eppendorf tube with care using Axygen Maximum Recovery pipette tips (Fisher Scientific), and incubated for 5 minutes at 4°C on an end-over-end rotator. This washing step was repeated for two more times, and the beads were then resuspended and washed in high salt washing buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.1, 500 mM NaCl) for three times, and then in lithium chloride washing buffer (1% Igepal, 1 mM EDTA, 10 mM Tris pH 8.1, 250 mM LiCl, 1% sodium deoxycholate) for three times. This was followed by two more washes with TE washing buffer (10 mM Tris pH 8.0, 1 mM EDTA). All the washes were done on an end-over-end rotator for 5 minutes at 4°C.

2.3.15.4 Elution, reversal of cross-linking, and proteinase K digestion

After all the washes, the beads were resuspended in 240 μ L of ChIP elution buffer (0.1 M NaHCO₃, 1% SDS in autoclaved ultra-pure water) and transferred to Safe-Lock eppendorf tubes (Eppendorf Ltd). 2% input were thawed and its volume were also topped up to 240 μ L with ChIP elution buffer. Both ChIP and input samples were then added with 10 μ L of 5M NaCl, 2 μ L of Proteinase K (20 mg/mL) and 1 μ L of DNase-free RNase A (20 mg/mL). The resulting mixtures were then incubated on ThermoMixer with ThermoTop (Fisher Scientific) at 1,400 rpm for 16 hours at 65°C.

2.3.15.5 PCR purification

After the incubation, ChIP and input samples were then added with 1 mL of PB buffer (Qiagen) and 10 μ L of 3M sodium acetate pH 5.2. This was followed by PCR purification using Qiagen MiniElute PCR Purification columns (Qiagen) according to the manufacturer's protocol. Elution was then performed with 200 μ L of Nuclease-free water (Invitrogen) in Safe-Lock eppendorf tubes (Eppendorf Ltd). Eluted samples were stored at 4°C for short-term use and -20°C for long-term storage.

2.3.15.6 Quality control of DNA fragment size

To quality control (QC) the size of the fragmented chromatin, purified DNA were analyzed by D1000 ScreenTape Assay (Agilent) using the TapeStation system (Agilent) according to the manufacturer's protocol. The size and quality of the DNA samples were confirmed by size verification in the resulting gel image and electrophoregram using the TapeStation Analysis Software (Agilent).

2.3.16 *In situ* proximity ligation assay (PLA)

Cells were seeded in a 2-well insert placed on a -slide 8 well (Ibidi), with boundary encircled with PAP pen (Sigma). At 80-100% confluency, cells were washed once with PBS and fixed with 4% Paraformaldehyde/PBS at room temperature for 15 minutes and washed once with PBS. Fixed cells were permeabilized with 0.3% Triton-X-100 in PBS for 10 minutes at RT and blocked with 1% bovine serum albumin/PBS (0.22m-filtered) for 1 hour at RT.

In situ PLA detection was carried out using the appropriate Duolink kit components (Sigma-Aldrich) according to the protocol of the manufacturer. Briefly, mouse anti-ASPP2 antibody and rabbit anti-PP1 α antibody were added in 40 μ L

PLA antibody diluent and incubated in a humidified chamber at 4°C overnight. The cells were then washed with wash buffer A (0.01M Tris, 0.15M NaCl, 0.05% Tween 20 in ultra pure water) twice for 5 minutes each, then secondary antibodies, Duolink anti-mouse PLA-PLUS probe and Duolink anti-rabbit PLA-MINUS probe (Sigma-Aldrich), were added and incubated at 37°C in humidified chamber for 1 hour. Three washes with wash buffer A were followed by addition of the ligation mix and incubation at 37°C for 30 minutes, followed by another two washes. Thereafter, the amplification reaction was carried out at 37°C for 100 minutes. Afterwards, the cells were washed twice with wash buffer B (0.2M Tris, 0.1M NaCl) and once with 0.1X wash buffer B. Mounting was done with Duolink *In Situ* Mounting Medium with DAPI (Sigma-Aldrich).

Samples were observed using a confocal microscope system (LSM 710, Carl Zeiss). Images were acquired using Zen imaging software (Carl Zeiss) according to the manufacturer's guidelines. For each experiment, 5 images per sample were acquired using identical microscope settings.

Image analysis was performed in ImageJ with the assistance of former Microscope Manager, Dr. Mark Shipman, and Dr. Richard Lisle, current Microscope Manager (Ludwig Institute, Oxford). Total PLA signals in the image were counted and divided by the number of nuclei to obtain average signals per cell in each image. Nuclear interactions per cell were defined as the number PLA signals in the nucleus divided by the number of nuclei per image. Counting was performed using an automated ImageJ macro developed by Dr. Mark Shipman. Based on the counting data, means and standard deviations, in the form of bar graphs, or boxplots of the independently acquired images were presented. Statistical analyses of the

quantification results were carried out using multiple t-test function in Prism 8 (GraphPad Software Inc.). Adjusted p value of 0.05 was used as the significance threshold unless otherwise stated.

To improve visibility, contrast and brightness of the images were slightly enhanced in the thesis. For this, all images to be compared were treated the same and this had no effect on the quantification, which was performed using the unmodified images.

2.4 Proteomics sample preparation

For ASPP2 interactome study by IP-liquid chromatography tandem mass spectrometry (LC-MS/MS), biological triplicates were prepared. For each replicate, cells were seeded in a 15 cm dish and grown to 60-80% confluent. They were then transfected with designated plasmids and processed as described in the IP section [subsection 2.3.14] using anti-V5 antibody and the downstream procedures were outlined below.

2.4.1 Glycine elution

For elution, glycine elution buffer was added and the sample incubated under gentle rotation for 30 minutes at 4°C. The beads were separated from the elution buffer and the supernatant was transferred into a new tube. The supernatant was neutralized using 1M Tris-HCl buffer at pH8. Dithiothreitol (DTT) was added to the sample and incubated at room temperature for 30-60 minutes prior to the addition of iodoacetamide alkylation reagent. The samples were allowed to sit for another 30-60 minutes at room temperature.

2.4.2 Methanol/Chloroform precipitation and in-solution digest

Methanol/Chloroform precipitation and in-solution digest were performed according to established protocols in Professor Benedikt Kessler Group (University of Oxford). Briefly, glycine-eluted samples were first diluted in reaction buffer to 200 μ L. 600 μ L of Methanol and 150 μ L of chloroform were then added to the sample, followed by vortexing, addition of 450 μ L of water and further vortexing. After centrifugation at 16,000 g for 1 minute at room temperature, the upper phase was removed without disturbing the protein precipitate in the interphase. Further 450 μ L of methanol was added and vortexing was performed again prior to centrifugation. The supernatant was removed, yielding the protein pellet. The pellet was re-suspended in urea buffer by vortexing and sonication. The concentration was diluted and trypsin added and the digestion mixture incubated at 37°C for 15-18 hours. The resulting digested peptides were de-salted following the SEP-PAK C18 purification protocol.

2.4.3 SEP-PAK C18 purifications

SEP-PAK C18 columns (Waters) were equilibrated with 5 mL of solution B (65% CH₃CN, 35% H₂O, 0.1% HCOOH) and subsequently 10 mL of solution A (98% H₂O, 2% CH₃CN, 0.1% HCOOH). The samples from the in-solution digest were diluted in 10 mL solution A and slowly applied on the columns. The columns were washed with an additional 10 mL of solution A before the peptides were eluted with 0.6 mL of solution B twice. Desalted peptides were dried under reduced pressure and the resulting pellet re-suspended in 20 μ L of solution A by vortexing and sonication and analysed using LC MS/MS.

2.5 Phosphoproteomics sample preparation

For ASPP2 phosphoproteomics study using IP-liquid chromatography electrospray ionization tandem mass spectrometry (IP-LC-ESI-MS/MS), sample preparation was performed as described in section 2.4. In-solution digestion and SEP-PAK C18 desalting were performed as described in subsection 2.4.2 and subsection 2.4.3, respectively.

Dried pellet after desalting was resuspended, and 5% of each sample were used for proteomics and injected directly into LC-MS. The rest of the sample were injected into Ultimate 300 UHPLC (Thermo Scientific) for High pH reverse phase pre-fractionation of the peptides, using XBridge C-18 column (150 x 1.0 mm, Waters UK) and Buffer A (double-distilled water, pH 10) and Buffer B (90% Acetonitrile, pH 10). Samples were injected (in 200 μ l volume) on to the column over a 100-minute gradient of increasing amount of Buffer B to collect 50 fractions of 400 μ l each. The fractions 12-43 are concatenated and pooled into 8 fractions, which were then dried and used for phosphor-peptide enrichment using titanium dioxide as described in Sharma *et al.* 2015 (Sharma et al. 2015).

2.6 Antibodies and primers

Primary antibodies

Table 2.4: List of primary antibodies

Antigen	Clone	Host/Type	Source	Catalog No.	Applications
ASPP2	DX54.10	Mouse ascites	Lu lab	–	WB
ASPP2	19	Mouse mAb	Santa Cruz	sc-135818	WB/IF/IP/PLA
β -Actin	C-2	Mouse mAb	Santa Cruz	803	WB
β -Tubulin	–	Rabbit pAb	Abcam	6046	WB
FIH	162c	Mouse mAb	Ratcliffe lab	–	WB
FLAG-tag	M2	Mouse mAb	Sigma	F1804	WB/IF/IP
HIF1 α	PM14	Rabbit sera	Ratcliffe lab	–	ChIP
HIF1 α	54	Mouse mAb	BD	610959	WB/IP/IF/PLA
HIF1 β	–	Rabbit pAb	Novus Biologicals	100110	ChIP
HIF2 α	PM9	Rabbit sera	Ratcliffe lab	–	ChIP
HIF2 α	190b	Rabbit sera	Ratcliffe lab	–	WB
Myc-tag	9E10	Mouse mAb	Lu lab	–	IP
p300	20	Rabbit pAb	Santa Cruz	585	WB/IP

Continued on next page

Table 2.4 – continued from previous page

Antigen	Clone	Host/Type	Source	Catalog No.	Applications
p300	–	Rabbit pAb	Abcam	10485	ChIP
p53	DO-1	Mouse mAb	Santa Cruz	126	WB/ChIP
p53	DO-1	Mouse ascites	Lu lab	–	WB/ChIP
p53	–	Rabbit pAb	Cell Signaling	9282	WB/ChIP
PHD3	–	Rabbit pAb	Abcam	30782	WB
PP1 α	E-9	Mouse mAb	Santa Cruz	7482	WB
PP1 α	C-19	goat pAb	Santa Cruz	6104	WB
PP1 α	–	rabbit pAb	Sigma	0154	PLA
RASSF8	4B1	mouse mAb	Santa Cruz	81934	WB
V5-tag	–	Mouse mAb	Serotec	MCA1360	WB/IP
V5-tag	–	Rabbit pAb	abcam	ab15828	WB/IP
YAP	D8H1X	Rabbit mAb	Cell Signaling	14074	WB/IF/IP/ChIP
YAP	63.7	Mouse mAb	Santa Cruz	101199	WB/IF
ZO-1	D6L1E	Rabbit mAb	Cell Signaling	13663	IF

Abbreviations: ChIP - Chromatin Immunoprecipitation; mAb - Monoclonal antibody; pAb - Polyclonal antibody; IF - Immunofluorescence; IP - Immunoprecipitation; PLA - Proximity ligation assay; WB - Western blotting.

Secondary antibodies

Table 2.5: List of Secondary Antibodies

Type	Host	Supplier	Catalog No.	Application
Alexa Fluor®488 Anti-Rabbit IgG	Donkey	Invitrogen	A21206	IF
Alexa Fluor®488 Anti-Rabbit IgG	Goat	Invitrogen	A11008	IF
Alexa Fluor®647 Anti-Rabbit IgG	Goat	Invitrogen	A27040	IF
Alexa Fluor®488 Anti-Mouse IgG	Donkey	Invitrogen	A21202	IF
Alexa Fluor®546 Anti-Mouse IgG	Donkey	Invitrogen	A10036	IF
Alexa Fluor®647 Anti-Mouse IgG	Donkey	Invitrogen	A31571	IF
Alexa Fluor®647 Anti-Mouse IgG	Goat	Invitrogen	A21235	IF
Anti-Goat Immunoglobulins/HRP	Rabbit	Dako	P0449	WB
Anti-Mouse Immunoglobulins/HRP	Rabbit	Dako	P0161	WB
Anti-Rabbit Immunoglobulins/HRP	Pig	Dako	P0399	WB

Abbreviations: HRP - Horseradish peroxidase; IF - Immunofluorescence; WB - Western blotting.

Primers

Table 2.6: List of Primers used for Molecular cloning

Construct	Name	Sequence
ASPP2	PB-ASPP2-Fw	TAA GCA CTC GAG CAC CAT GAT GCC GAT GTT TCT T
	PB-ASPP2-Rv	TAA GCT ATC GAT GAG GCC AAG CTC CTT TGT CT
	ASPP2-Fw1	ATT CTA GGG TGG GCT CAG GAG TAT
	ASPP2-Rev1	CCT TGC CTT CAT CAA ATT CTC AGT CAA
	ASPP2-RVxF-F	ACA AAC TTG CGT AAA ACT GGC
	ASPP2-RVxF-R	TCG TCA TCC TTG TAG TCG ATG T

Table 2.7: List of Primers used for sequencing

Target/Vector	Name	Sequence
PX459	hU6-F	GAG GGC CTA TTT CCC ATG ATT
	Cas9-F	TTT ATG GCG AGG CGG CGG
	Cas9-R	GTG GGC TTG TAC TCG GTC AT
pPB-MCS-FLAG-Puro	XhoI-F	CTC GAG CAC CAT GAT GCC

Continued on next page

Table 2.7 – continued from previous page

Target/Vector	Name	Sequence
ASPP2	XhoI-R	TTT CTT TTT AAA CTT GGA TCC TG
	Teton-F	TCT AGA CTG GAC AAG AGC AAA GT
	Teton-R	TCC CCG GGG AGC ATG TCA AG
	RVxF-F	ACA AAC TTG CGT AAA ACT GGC
	RVxF-R	TCG TCA TCC TTG TAG TCG ATG T
Miscellaneous	T7-F	TAA TAC GAC TCA CTA TAG GG
	T7-R	GCT AGT TAT TGC TCA GCG G
	BGH-R	TAG AAG GCA CAG TCG AGG
	pCMV-F	GAG CTC GTT TAG TGA ACC GTC
	M13-R	TGT AAA ACG ACG GCC AGT

Table 2.8: List of Primers used for qPCR

Gene	Name	Sequence
<i>ACTB</i>	ACTB-F	CAT CCA CGA AAC TAC CTT CAA CTC C
	ACTB-R	GAG CCG CCG ATC CAC ACG
Continued on next page		

Table 2.8 – continued from previous page

Gene	Name	Sequence
<i>ANKRD37</i>	ANKRD37-F	TGT GTT GCC GTG CTC AGA CAG A
	ANKRD37-R	ACC CAC GTG ACA TCA GCA CTT C
<i>BNIP3</i>	BNIP3-F	GCC ATC GGA TTG GGG ATC TA
	BNIP3-R	AGC TCA GAA GTA ATC CAC TAA CG
<i>CA9</i>	CA9-F	AGC ACA GAA GGG GAA CCA AAG
	CA9-R	ATG AGC AGG ACA GGA CAG TTA C
<i>HIF1A</i>	HIF1 α -F	TAT GAG CCA GAA GAA CTT TTA GGC
	HIF1 α -R	CAC CTC TTT TGG CAA GCA TCC TG
<i>EGLN1</i>	PHD2-F	GCA GCA TGG ACG ACC TGA TA
	PHD2-R	CCA TTG CCC GGA TAA CAA GC
<i>EGLN3</i>	PHD3-F	TCC TGC GGA TAT TTC CAG AGG
	PHD3-R	GGT TCC TAC GAT CTG ACC AGA A
	PHD3-F2	CTG GTC CTC TAC TGC GGG A
	PHD3-R2	AGC CAC CAT TGC CTT AGA CCT C
<i>VEGFA</i>	VEGFA-F	GGC AGA ATC ATC ACG AAG TGG T
	VEGFA-R	GGC ACA CAG GAT GGC TTG AA

Table 2.9: List of Primers used for ChIP-qPCR

Gene	Name	Sequence
<i>CDKN1A</i>	CHIP-CDKN1A-FW	AGC CTT CCT CAC ATC CTC CT
	CHIP-CDKN1A-RV	GGA ATG GTG AAA GGT GGA AA
<i>BBC3</i>	CHIP-BBC3-FW	CCT GCT CTG GTT TGG TGA GT
	CHIP-BBC3-RV	TTG GGT GAG ACC CAG TAA GG
<i>BAX</i>	CHIP-BAX-FW	AAC CAG GGG ATC TCG GAA G
	CHIP-BAX-RV	AGT GCC AGA GGC AGG AAG T
<i>CCN2</i>	CHIP-CTGF-FW	GCC AAT GAG CTG AAT GGA GT
	CHIP-CTGF-RV	CAA TCC GGT GTG AGT TGA TG
<i>EGLN3</i>	CHIP-EGLN3-FW1	AGT GTC CGT TCC CAG CTC AG
	CHIP-EGLN3-RV1	TAG GCA CAG TAA ACA GGC CC
<i>EGLN3</i>	CHIP-EGLN3-FW2	TGT CTT GAT GTC TGC TTC TTG
	CHIP-EGLN3-RV2	GAT GCT TTT GTT GTT CAC CC
<i>MDM2</i>	CHIP-MDM2-FW	GTT CAG TGG GCA GGT TGA CT
	CHIP-MDM2-RV	CGG AAC GTG TCT GAA CTT GA
<i>VEGFA</i>	CHIP-VEGFA-FW	TGG GTG AAT GGA GCG AGC AG
	CHIP-VEGFA-RV	TAT TGA AGG GGG CAG GGG AAG G

Table 2.10: List of plasmids used in this study

Construct	Vector	Source	Additional information
Empty vector	pEF6	Ratcliffe lab	HA-tag
	pcDNA 3.1	Lu lab	V5-tag
EGFP	pCMV-GFP	Gyrd-Hansen lab	–
PiggyBac transposase	SPB00125	Goding lab	–
pSpCas9	PX459	Zhang lab	Puromycin resistance
rtTA	pB-rtTA	Coding lab	Neomycin resistance
Wild-type ASPP2	pcDNA 3.1	Lu lab	V5-tag
	pPB-MCS-FLAG-Puro	Lu lab	3xFLAG-tag, Tet-ON

2.7 Bioinformatics methods

2.7.1 Bioinformatics for interactomics

For the ASPP2 interactome study by IP-LC-MS/MS, peptide annotation, quantification and normalization were performed by Ms Georgina Berridge from the Ludwig Institute, Oxford. Briefly, the identified peptides were mapped against the human proteome using the UniProt18 Database and PEAKS Software. Progenesis (Nonlinear Dynamics) was used for label-free quantification. Normalization of different IPs in the same experiment was done by normalizing each IP against all proteins identified in the IP.

Normalized abundance of the identified proteins were analyzed using Perseus (Tyanova et al. 2016). To minimize false positives, only proteins with at least two unique peptides and those that were detected in all three biological replicates of the same experimental condition were used for further analysis. Proteins were log transformed and were inspected for log-normal distribution. Differential expression analysis between samples from V5-ASPP2-IP and empty-vector-IP was performed using two-sample t-test. 5000 randomizations were performed and a threshold cutoff using a combination of artificial within-group variance (s_0) = 0.3 and false discovery rate = 0.5 was used to determine candidate ASPP2 interactors. s_0 was a built-in parameter of Perseus to control the relative importance of statistical significance and effect size (Tusher et al. 2001).

2.7.2 Bioinformatics for phosphoproteomics

MaxQuant (v1.5.2.8, MPI for Biochemistry) was used for processing raw files. MS/MS spectra were then searched against the UniProtKB/Swiss-Prot human

database. Precursor and fragment ion mass tolerances were set to 6 and 20 ppm after initial recalibration, respectively. STY phosphorylation, protein N-terminal acetylation, and methionine oxidation were allowed as variable modifications. Cysteine carbamidomethylation was defined as a fixed modification. Minimal peptide length was set to seven amino acids, with a maximum of two missed cleavages. The false discovery rate (FDR) was set to 1% on both the peptide and the protein level using a forward-and-reverse concatenated decoy database approach.

Data analysis was conducted with Perseus software (v1.5.2.4, MPI for Biochemistry). After removing decoy and contaminant entries, identified phosphosites with a localization probability lower than 0.75 and those whose protein levels were differentially expressed between WT and ASPP2-KO conditions were filtered out. Differentially expressed proteins were defined as having an absolute log2 fold change of LFQ intensity greater than 2 and false discovery rate less than 0.05.

The resulting list of phosphosites were then analyzed using R using the criteria described in subsubsection 5.2.6.3.

2.7.3 Bioinformatics for RNA-seq

For the RNA-seq experiment conducted in U2OS cell lines, preprocessing of raw fastq files and read alignment were performed by the Oxford Genomics Centre in Oxford. Briefly, FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) was performed on raw fastq files to assess read quality and PCR duplication events, and then mapped to human genome build hg38 (GRCh38, 2015) using STAR aligner (Dobin et al. 2013). Counts data generated by featureCounts from the Rsubread package in R (Liao et al. 2019) was used as an input for

differential gene expression analysis. Of note, one ASPP2-KO replicate was excluded from downstream analysis due to abnormally low read counts (A. Green, personal communication, 18 June, 2019) and therefore technical quadruplicate of WT, FIH-KO, ASPP2-FIH-Double-KO (DKO), and technical triplicate of ASPP2-KO U2OS samples were used for further analysis.

Before proceeding to downstream analysis, multidimensional scaling by using the plotMDS function in the edgeR package (Robinson et al. 2010) to check that replicates from the same experimental groups are close to each other and far from those that belong to different experimental groups.

Differential expression analyses were performed in R using the edgeR package. Briefly, read counts for each sample set were first filtered for genes whose expression is < 1 count per million. The remaining read counts were modelled using a negative binomial model by edgeR and differential expression analysis was performed using the quasi-likelihood F-test in the edgeR package to identify differentially expressed genes between experimental groups. Genes with a false discovery rate of less than 0.1 were considered as differentially expressed and used for further analysis. As controls, normalized expression values of ASPP2 and FIH were compared between experimental groups to validate the knockout status of the samples. Heatmaps were generated from the edgeR-library normalized reads of genes with standardization and hierarchical clustering using the Pheatmap package in R (<https://cran.r-project.org/web/packages/pheatmap/index.html>).

For the RNA-seq analysis of mouse intestinal organoids, counts data were obtained from Dr. Xiao Qin from Xin Lu lab. Sequencing was performed by the Oxford Genomics Centre in Oxford and the preprocessing steps were as described

above for the U2OS cell lines. Trimmed means of M values (TMM) from the edgeR package was used for counts normalization. For gene set variation analysis (GSVA) (Hänzelmann et al. 2013), normalized counts were used as input and the C2 chemical and genetic perturbation collection from the Molecular Signature Database (Subramanian et al. 2005). Heatmaps of the GSVA results were generated using Pheatmap package in R.

2.7.4 Defining HIF target genes in U2OS cells

HIF target genes in U2OS cells were defined by using HIF1 β ChIP-seq data in WT U2OS cells treated with hypoxia and the RNA-seq data in WT U2OS cells treated with normoxia and hypoxia. The HIF1 β hypoxia ChIP-seq data was used as this was the only existing in-house ChIP-seq data conducted in the U2OS cell line, and HIF1 β binding under hypoxia represents a reliable indicator of HIFs DNA binding activity under hypoxia (D. Mole, Personal communication, 20 Feb, 2020). This ChIP-seq data contained one sample and no input sample and its corresponding ChIP-seq peaks in bed-file format was preprocessed by Dr. James Platt and Dr. Rafik Salama (D. Mole, Personal communication, 20 Feb, 2020). Peak annotation was carried out by using the Homer package with peaks being assigned to the nearest gene (Heinz et al. 2010).

To determine HIF target genes in U2OS cells, we first used the RNA-seq data as mentioned in (subsection 2.7.3) and performed differential expression analysis was performed between WT U2OS cells treated with normoxia and hypoxia to identify hypoxia response genes. Using a false discovery rate (FDR) of less than 0.1 as a threshold, we identified 3975 hypoxia response genes, which were then intersected

with the HIF1 β hypoxia ChIP-seq peaks with the top 2000 peak scores. The top 2000 selection was an arbitrary threshold based on the lack of input sample and the numerical range of identified HIF1 β -binding sites in 4 other cell lines (from 970 sites to 2081 sites (V. Schmid et al. 2019)). From this intersection, 666 overlapped genes were identified and considered as the HIF target genes in U2OS cells.

2.7.5 Pan-Cancer Gene Coexpression Analysis

The bioinformatics pipeline for performing pan-cancer gene coexpression analysis was written in R. The input of the pipeline included the TCGA pan-cancer RNA-seq data as the expression dataset and its corresponding metadata which included clinical information, gene-level somatic copy number estimates, and binary gene-level non-silent mutation calls as potential sources of confounding. These input data were publicly accessible from the Genomic Data Commons (GDC) Data Portal.

The TCGA pan-cancer RNA-seq data comprised 11070 patient tumor samples from 31 cancer types [Figure 5.2] and the gene expression of 20532 genes. Transcript quantification and normalization were performed by the TCGA. Briefly, RNA-Seq by Expectation Maximization (RSEM) (B. Li et al. 2011) was used for quantification, followed by upper-quartile normalization (G. F. Gao et al. 2019). Batch correction was performed by Hoadley *et al.* to adjust for batch effects from different sequencing platforms, sequencing centers, and sequencing plates using an empirical Bayes/ComBat-based method (Bullard et al. 2010; Hoadley et al. 2018; Johnson et al. 2007).

Missing data in the TCGA pan-cancer RNA-seq data was identified by using R packages "funModeling" (Pablo 2019), which was used to compute the number and

percentage of missing gene expression values, and "Amelia" (Honaker et al. 2011), which was used to visualize the distribution of missing gene expression values. In total, 3045 genes were listwise deleted from the dataset as it was difficult to accurately impute the expression values of these genes since they were missing from the entire acute myeloid leukaemia cohort, and the majority of stomach, esophageal, ovarian, and uterine corpus endometrial carcinoma cohorts.

Further preprocessing was performed to remove uninformative genes that were not expressed or very lowly expressed in most samples, as well as those whose expression displayed little variation across samples. This filtering was done by calculating the median absolute deviation of the expression of each gene in the dataset. As a heuristic cutoff, only genes with MAD greater than the median MAD were included for downstream analysis.

Data normalization was inspected by using boxplots to examine the distribution of mRNA expression values across samples of each TCGA cancer types, as well as across all cancer types in the entire pan-cancer dataset as a whole.

Potential confounders were identified by testing the association between *ASPP2* expression and each of the variables in the metadata. This was done by running univariate linear regression models for each individual metadata variables, which were the independent variables, with *ASPP2* expression being the dependent variable. A metadata variable would be considered as a potential confounder if the adjusted coefficient of determination (R^2) was greater than 0.1 and the p-value for the metadata variable was less than 0.05.

Pan-cancer gene coexpression analysis was performed by running multivariate linear regression models for each of the 9602 genes that remained after the

aforementioned preprocessing step in each of the 31 TCGA cohorts Equation 2.1.

$$y_i = \beta_0 + \beta_{xi}x + \sum_{k=1}^n \beta_k COV_k + \epsilon \quad (2.1)$$

In these models, y represented each of the 9602 genes, x represented the given gene of interest, and COV represented a confounding variable. For a given gene of interest x , the beta coefficient β_{xi} was used as the confounder(s)-adjusted coexpression score for each gene y_i in each of the 31 TCGA cohorts, as long as its corresponding p-value, after being corrected for multiple comparisons, was less than 0.05, otherwise the gene coexpression score would be zero. The pan-cancer coexpression score for each gene y_i was obtained by calculating the mean of the 31 coexpression scores. A ranked gene list was then obtained by ranking the 9602 genes based on their respective pan-cancer coexpression scores.

Gene sets used for overrepresentation analysis and pre-ranked gene set enrichment analysis (GSEA) were downloaded from the Molecular Signature Database (Liberzon et al. 2011). Overrepresentation analysis was performed by using the enricher function in "clusterProfiler" package (G. Yu et al. 2012). Pre-ranked gene set enrichment analysis was performed by using the "fgsea" package in R (Sergushichev 2016). For ASPP2 coexpressed pathways obtained by pre-ranked GSEA, only the REACTOME pathways with a FDR less than 0.01 were considered as enriched. Clustering of the enriched coexpressed pathways into functional groups was performed manually based on the description of the pathways and the leading edge subset of genes. Overrepresentation analyses in subsection 5.2.6 were performed using STRING (<http://www.string-db.org>) (Szklarczyk et al. 2015).

3

Generation of ASPP2-KO cell lines using CRISPR-Cas9

3.1 Introduction

3.1.1 The CRISPR-Cas9 system for gene-editing

To assess the influence of ASPP2 on HIF activity and its relation to FIH, CRISPR-Cas9 was used in WT U2OS and FIH-Knockout (FIH-KO) U2OS cells to create ASPP2-Knockout (ASPP2-KO) and ASPP2-FIH-Double-Knockout (DKO) cell lines. U2OS cell line was used as the major experimental model for CRISPR-Cas9 and the subsequent investigations introduced in the later chapters as it is one of the more commonly used cancer cell line models that are relatively easy to culture and manipulate. Most importantly, using this cell line also allows us to leverage existing in-house lab data that were generated using this cell line.

CRISPR-Cas system is a RNA-guided gene-editing tool originated from bacteria and archaea as a form of their adaptive immune response against viral infection (Barrangou et al. 2007). Several types of CRISPR-Cas system have been identified

and the one used in this study is the type II CRISPR-Cas9 system from *Streptococcus pyogenes* (Jinek et al. 2012; Koonin et al. 2017). As its name suggests, CRISPR-Cas9 has two important components - CRISPR and Cas9. CRISPR consists of the targeting crRNA that engages targeted dsDNA based on Watson-Crick base pairing and the activating tracrRNA that binds to the repeat sequences in the crRNA (Deltcheva et al. 2011). Together they form the guide (gRNA) that directs Cas9, a DNA endonuclease, to the site of cleavage. To activate the cleavage by Cas9, it needs to interact also with a protospacer adjacent motif (PAM) sequence in the form of 5'-NGG-3' contiguous to the target dsDNA (Bolotin et al. 2005; Garneau et al. 2010; Gasiunas et al. 2012; Jinek et al. 2012). The cleavage by Cas9 results in DSBs 3 base pair (bp) upstream of the PAM. Importantly, these DSBs are predominantly corrected by the error-prone NHEJ mechanism. This frequently results in random indels at and near the DSB site, and may therefore knockout the expression of the target gene as a result of a loss-of-function mutation.

CRISPR-Cas9 offers significant advantages over its predecessors, Zinc finger nucleases (ZFNs) and Transcription activator-like effector nucleases (TALENs). Whereas ZFNs and TALENs use DNA-binding motifs to direct site-specific cleavages and thus requires designing and cloning of new proteins whenever a new site is targeted (Carroll 2011; Joung et al. 2013), CRISPR-Cas9 is simpler and more flexible as Cas9 can be engineered to target a new site by modifying the gRNA sequence.

3.1.2 Results

3.1.2.1 Designing the CRISPR-Cas9 system for knocking out ASPP2

To target the gene locus of human *TP53BP2*, the gene that encodes ASPP2, using the CRISPR-Cas9 system, gRNAs were designed using three different gRNA design tools, namely MIT web sever (Hsu et al. 2013), gRNA Designer (Doench et al. 2016) and CRISPOR (Haeussler et al. 2016), which employ both overlapping and different prediction methods based on machine learning models to predict on-target and off-target activities of gRNAs.

The following criteria were followed during gRNA design in order to maximize on-target efficiency while minimizing off-target effects:

- (i) Target specificity: each candidate gRNA must contain at least 2 mismatches to each off-target site. Importantly, any gRNA that has off-targets with no mismatches in the 12 bp adjacent to the PAM, i.e. the seed sequence (L. Cong et al. 2013; Jinek et al. 2012; Semenova et al. 2011), would not be selected;
- (ii) Target positioning: the earliest exon that is present in all splice variants of *TP53BP2* should be targeted;
- (iii) Sequence criteria: each candidate gRNA must have less than 4-nucleotide homopolymers and a %GC between 20% to 80% (T. Wang et al. 2014);

Since *TP53BP2* has another N-terminally truncated variant [Figure3.1] (Kuniyoshi Iwabuchi et al. 1998; Naumovski et al. 1996; Samuels-Lev et al. 2001), exon 6 was selected as the target region for gRNA design as it was the earliest exon that was present in both splice variants.

Similar results of gRNA design were obtained from the 3 prediction tools and the representative results from CRISPOR were shown [Table 3.1]. The MIT

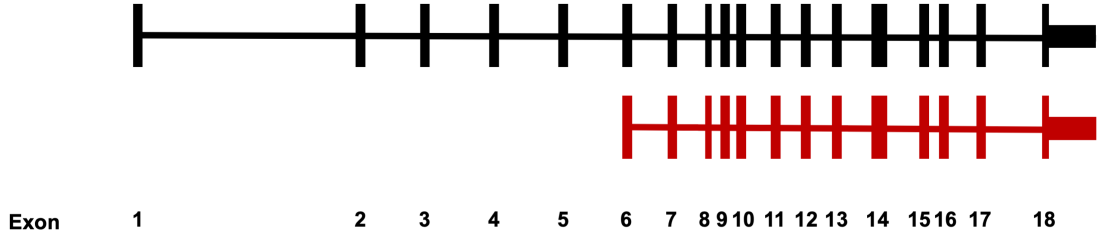


Figure 3.1: Transcript variants of *TP53BP2*. Schematic showing the two transcripts for *TP53BP2* due to alternative splicing. The full-length transcript (black) encodes ASPP2. The shorter transcript (red), starting from exon 6, gives rise to a N-terminally truncated isoform termed BBP.

and CFD specificity scores indicate potential off-target effects and gRNAs with fewer predicted off-targets would have higher MIT and CFD scores (Doench et al. 2016; Hsu et al. 2013). Doench '16 predicted the on-target activity of the gRNAs and Out-of-Frame- and Lindel-Score estimated the likelihood of a frameshift mutation. The most important feature was the number of off-targets indicated by each number of mismatches. Shown in parentheses were the off-targets with no mismatches in the 12 bp juxtaposing the PAM, i.e. the seed sequence, and they represented the most likely off-targets.

gRNA#1 was chosen as it is the only gRNA that fulfilled the aforementioned criteria and had the lowest off-target potential and the highest predicted on-target activity. Its top twenty highest-scoring potential off-targets were mostly in the intergenic regions and all of them contained 4 mismatches in which at least one was in the seed sequence, suggesting that gRNA#1 had a minimal potential off-target effect [Table S1].

	gRNA#1	gRNA#2	gRNA#3	gRNA#4	gRNA#5
MIT Specificity Score	82	78	68	65	62
CFD Specificity Score	91	88	81	81	67
Total Off-targets	96	120	198	255	423
Doench-16 Score	73	71	52	36	41
Out-of-frame Score	67	64	51	62	65
Lindel Score	86	82	73	76	85
Off-targets for 0,1,2,3,4 Mismatches Next to PAM	0-0-0-0-0	0-0-0-1-3	0-0-1-1-5	0-0-0-3-9	0-0-1-2-2
Off-targets for 0,1,2,3,4 Mismatches	0-0-1-9-86	0-0-0-11-109	0-0-2-18-178	0-0-3-29-223	0-0-13-28-382

Table 3.1: Results of gRNA prediction using CRISPOR. Table showing the top 5 predicted gRNAs with their respective on-target and off-target scores.

The length of gRNA#1 was consistent with those of the conventional gRNAs, which was 20-nucleotide long (Hsu et al. 2013; Jinek et al. 2012). However, as truncated gRNAs with lengths of 17- and 18-nucleotide were demonstrated to have an improved specificity and a comparable on-target activity when compared with the conventional 20-nucleotide gRNAs (Fu et al. 2014), the first two nucleotides from the 5'end of gRNA#1 were removed from the design.

The truncated gRNA#1 was then cloned into CRISPR-Cas9 plasmid PX459 [Figure 3.2A] and was transiently transfected into WT and FIH-KO U2OS cells. Candidate single-cell KO clones were generated as illustrated in the workflow [Figure 3.2B].

3.1.2.2 Validation of ASPP2-KO and DKO clones

To validate the KO of ASPP2 expression, protein lysates from each clonally expanded single-cell clone were analyzed by western blot [Figure 3.3A] with lysates from WT and FIH-KO U2OS cells as positive controls. Compared to WT and FIH-KO U2OS

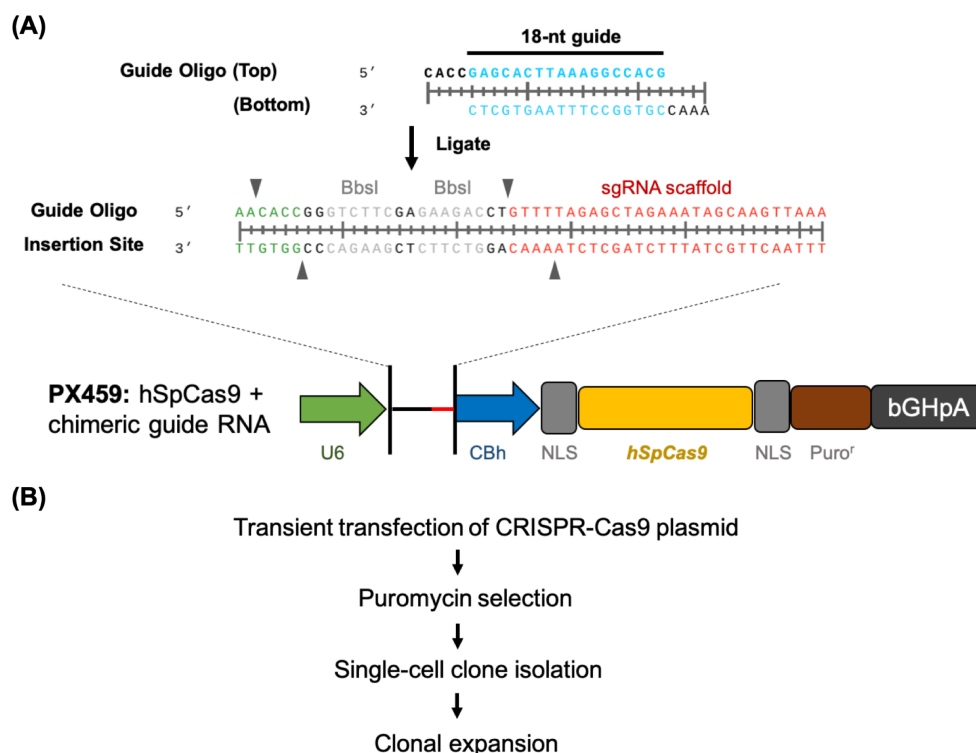


Figure 3.2: Experimental workflow of CRISPR/Cas9. (A) Schematic illustrating the cloning of the gRNA into the CRISPR-Cas9 plasmid PX459. (B) WT and FIH-KO cells were respectively transfected with PX459 containing the gRNA oligo. Cells were then selected with 1 μ g/mL puromycin for one week to enrich for cells with gene insertion. Cell that survived puromycin selection were then seeded as single cells on 96-well plates, expanded and subsequently collected for validation analyses.

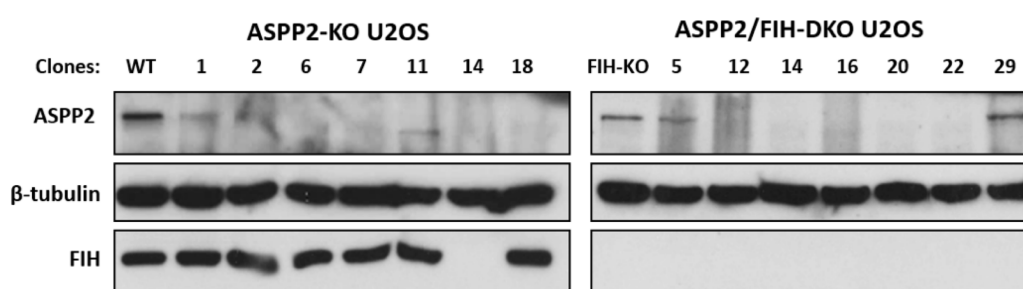
cells, ASPP2 expression was not detected in KO clones #6, #14 and #18, and in DKO clones #14, #20 and #22. The absence of FIH expression in KO-clones #14 suggested that it should be a DKO clone instead.

To confirm the western blot results were ASPP2-specific, immunofluorescence was performed using a different α -ASPP2 antibody from the one used for western blot [Figure 3.3B]. Consistent with the western blot results, ASPP2 expression was not detected in both the ASPP2-KO and DKO clones.

Selected KO clones were also evaluated at the nucleotide level by Sanger-sequencing multiple colonies of *E. coli* transformed with the plasmid into which the amplicons of the CRISPR-Cas9-targeted region was cloned. The Sanger sequencing

results showed that the selected clones either had a single nucleotide deletion at the predicted Cas9 cleavage site or multiple indels at and/or near the predicted Cas9 cleavage site, thus confirming the gene-edit [Figure 3.3C].

(A)



3.1.3 Summary and Brief Discussion

ASPP2-KO and DKO U2OS cells were generated by using the CRISPR-Cas9 gene-editing system. The resulting single-cell clones were then evaluated by western blot, immunofluorescence and Sanger sequencing to confirm the knockout of ASPP2 expression.

Although different prediction tools were used to select the gRNA with the least potential off-target effects and truncation of the gRNA was made to further improve targeting specificity, the extent of off-target effects in the selected ASPP2-KO and DKO clones had not been elucidated. Based on the results of *in silico* off-target prediction, the potential off-target with the fewest mismatches was intergenic and close to 20kb away from its closest gene *ACTRT1*, while other potential off-targets

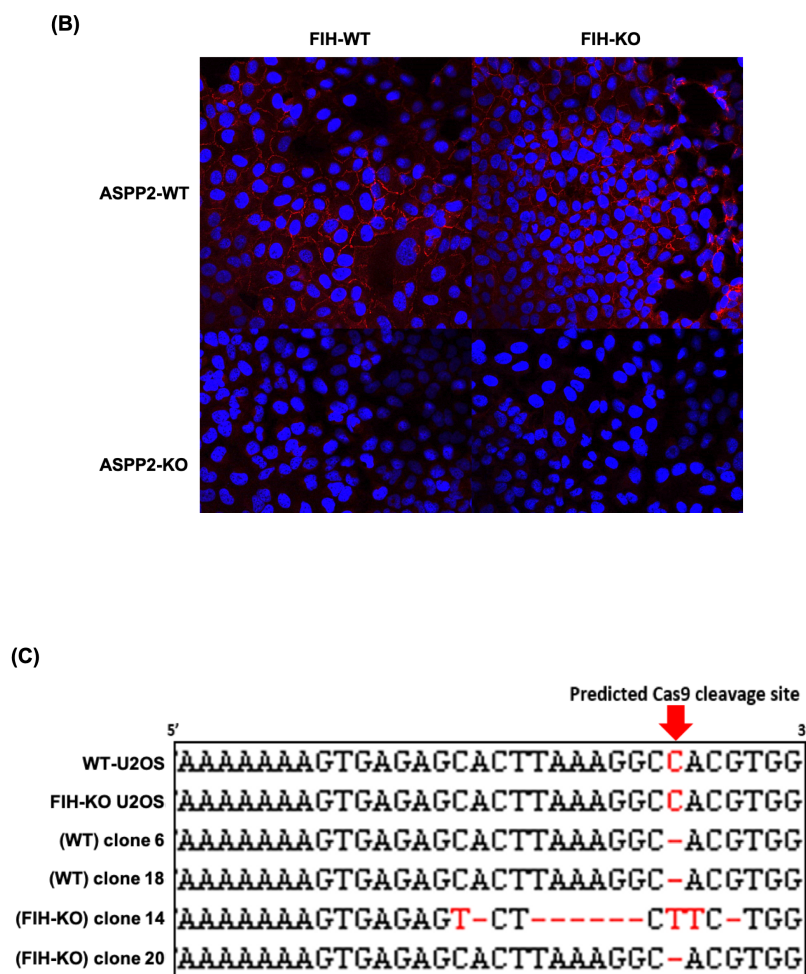


Figure 3.3: Validation of *ASPP2-KO* candidates. (A) Western blot analysis of whole cell lysates from candidate *ASPP2-KO* and DKO clones. WT and FIH-KO U2OS cells served as positive controls. (B) Immunofluorescence images showing the loss of *ASPP2* in the selected clones of *ASPP2-KO* (#18) and DKO (#14). WT and FIH-KO U2OS cells served as positive controls. Images are representative examples of two independent replicates of this experiment with at least three images taken per sample. (C) Sanger sequencing analysis of the nucleotide sequences near the predicted Cas9 cleavage site (red arrow) on *TP53BP2* locus obtained from selected *ASPP2-KO* and DKO clones. WT and FIH-KO U2OS cells were used as positive controls.

with the highest prediction scores were either intergenic and/or having at least one mismatch close to the PAM, making them unlikely off-targets. Therefore it was difficult to select any of them, nor was it practical to clone all of them and sequence for potential mutations. Considering the cost of whole-genome sequencing at the

time of generating these cell lines, sequencing option was also not considered.

However, the *in silico* prediction tools used in this study share some limitations. They do not take into consideration parameters such as potential off-target sites with alternative PAM sequences or with more than four mismatches (Cradick et al. 2013; Fu et al. 2013; Hsu et al. 2013; Pattanayak et al. 2013). Given that the cost of sequencing has become more affordable and new methods have been developed to assess off-target effects genome-wide in an unbiased manner based on CRISPR-Cas9-induced DSBs or recruitment of endogenous DNA repair machinery (Tsai et al. 2015; Wienert et al. 2019), it may be worth applying one of these methods to the ASPP2-KO and DKO clones for a more stringent evaluation of potential confounding off-target effects caused by CRISPR-Cas9. Future attempts of gene-editing using CRISPR-Cas9 may also consider using alternative methods to deliver the gRNA-Cas9 complex, such as the RNP delivery, to further improve targeting efficiency and specificity (S. Kim et al. 2014).

4

Investigation of the effects of ASPP2 on HIF transcriptional activity

4.1 Introduction

Over the past few decades of research, it has become clear that both oxygen limitation and oxygen sensing are inevitable parts of tumorigenesis. As the key pathway in oxygen sensing and homeostasis, tumor hypoxia inevitably activates the HIF pathway. The HIF pathway involves a complex transcriptional cascade and its activation results in the transactivation of a large number of genes mediated by its nuclear effectors, the HIF1/2 α in complex with ARNT. It is known that a number of HIF target genes are conducive to the tumorigenic process, such as pro-angiogenic genes that instruct more blood vessels to be created and glycolytic genes that underpin the metabolic switch fueling further tumor growth. In contrast, some HIF target genes are detrimental to tumorigenesis, such as proapoptotic genes that activates cell death signaling pathways. Natural selection implies that tumor cells that outcompete the others would have altered the transcriptional output

of HIF pathway to their growth advantage, through genetic and/or epigenetic means, or at least be able to compensate for the negative aspects of HIF activation. Elucidating the mechanisms underpinning the dysregulation of the HIF pathway is thus important in order to stop tumor cell growth.

As discussed in the introduction subsection 1.2.2, the transcriptional activity of HIF is regulated in both an oxygen-dependent and -independent manner. FIH represents one of the oxygen-dependent regulators of HIF. Its Asn hydroxylation in the HIF α -CAD prevents the association between HIF α -CAD and p300, thereby impairing the expression of p300-dependent HIF target genes. This hydroxylation has been shown to be influenced by the overexpression of certain ARD-containing substrates for FIH, such as Notch1 and MYPT1, as it is observed that they could compete with HIF α -CAD for FIH hydroxylation (Coleman et al. 2007; James D. Webb et al. 2009a). However, since these observations are based on *in vitro* competition assays and reporter assays, their *in vivo* physiological relevance remains to be elucidated.

Interestingly, ASPP2 is an ARD-containing substrate for FIH (Janke et al. 2013) and it has been shown to be capable of competing with HIF α -CAD for FIH hydroxylation *in vitro* (Leissing 2017). Additionally, the expression of ASPP2 is also subject to hypoxia-specific regulations. Hypoxia upregulates the expression of ubiquitin ligase SIAH2 and its increased expression promotes the proteasomal degradation of ASPP2 (H. Kim et al. 2014). Besides SIAH2, it has also been found that ASPP2 could be targeted by hypoxia-induced microRNA MiR-205 for degradation (X. Wang et al. 2016). Furthermore, we also observed multiple hypoxia response signature displayed differential expression patterns between WT

and ASPP2-deficient mouse intestinal organoids. Taken together, it is reasonable to hypothesize that ASPP2 regulates HIF transcriptional activity, potentially through a FIH-dependent mechanism. This hypothesis was addressed in this chapter by first looking at the effects of ASPP2 on HIF target gene expression, followed by the investigation of potential underlying mechanisms.

4.2 Results

4.2.1 Investigating the effects of ASPP2 on HIF target gene expression using a small panel of HIF targets

To investigate whether ASPP2 influences the transcriptional activity of HIF, the expression of a panel of HIF target genes were first assessed in WT and ASPP2-KO U2OS cells under normoxia and physiological hypoxia.

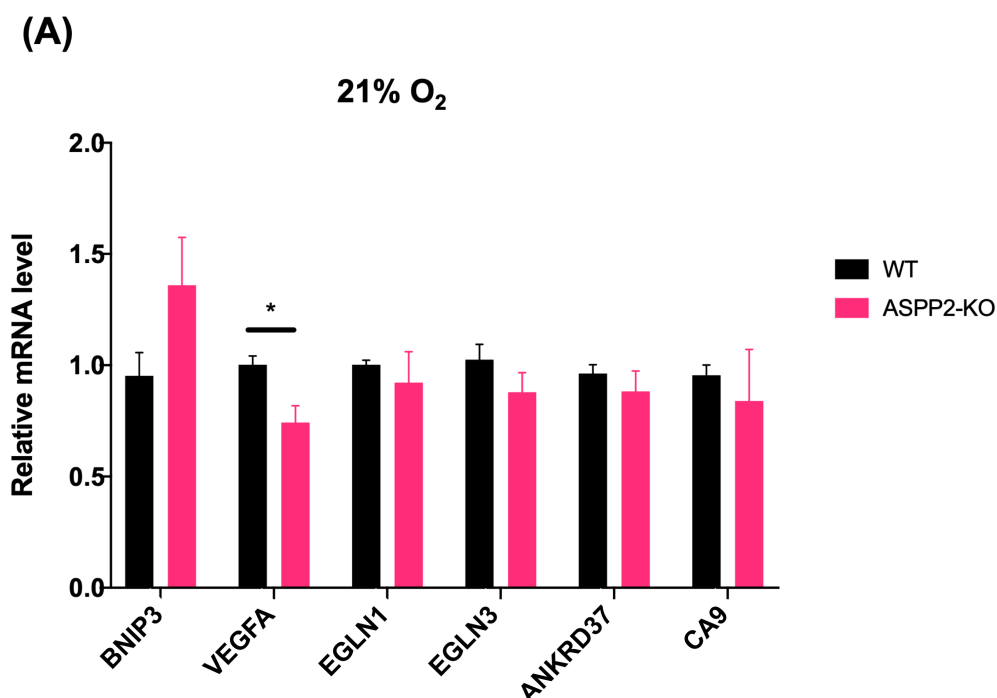
The genes in this panel were selected with an aim to create a small, yet representative HIF target gene set that included genes reported to be sensitive to FIH inhibition (*EGLN3*, *ANKRD37*, *CA9*) and those that are unaffected by FIH inhibition (*BNIP3*, *VEGFA*, *EGLN1*) (M. C. Chan et al. 2016). All of the selected genes have been shown to be hypoxia-inducible in a HIF-dependent manner and their gene loci bound by HIF (HIF1 α /HIF2 α /ARNT) under hypoxia in different cell lines.

To minimize potential reoxygenation of the samples when harvesting the cells, they were lysed inside the hypoxia workstation and the resulting lysates were placed immediately on ice before further processing. The mRNA expression of these HIF targets were measured by qPCR analysis.

Under normoxia, the loss of ASPP2 did not have a significant effect on the expression of HIF targets apart from *VEGFA*, which displayed a mild reduction

($p = 0.04$) compared to that in WT cells [Figure 4.1A].

In line with our expectation, hypoxia treatment induced the expression of all six HIF target genes to different extents - from about 2-fold in *VEGFA* to around 80% induction in *CA9* [Figure 4.1B]. Interestingly, the loss of ASPP2 under hypoxia resulted in a significant reduction of mRNA expression of *EGLN3*, *VEGFA*, and to a lesser extent, *EGLN1* ($p = 0.008$, $p = 0.006$ and $p = 0.03$ for *EGLN3*, *VEGFA* and *EGLN1*, respectively), whereas the gene expression of *BNIP3*, *ANKRD37* and *CA9* were not noticeably affected [Figure 4.1B]. Such observation suggested that ASPP2 might selectively modulate the gene expression of certain HIF targets, including *EGLN3*, *VEGFA* and *EGLN1*.



One major question heading into this chapter was whether ASPP2 could modulate HIF transcriptional activity through FIH. Since FIH was less sensitive to oxygen level compared to PHDs and it had been reported to retain partial asparaginyl hydroxylase activity even at 0.1% of oxygen (Y.-M. Tian et al. 2011b), treatment of

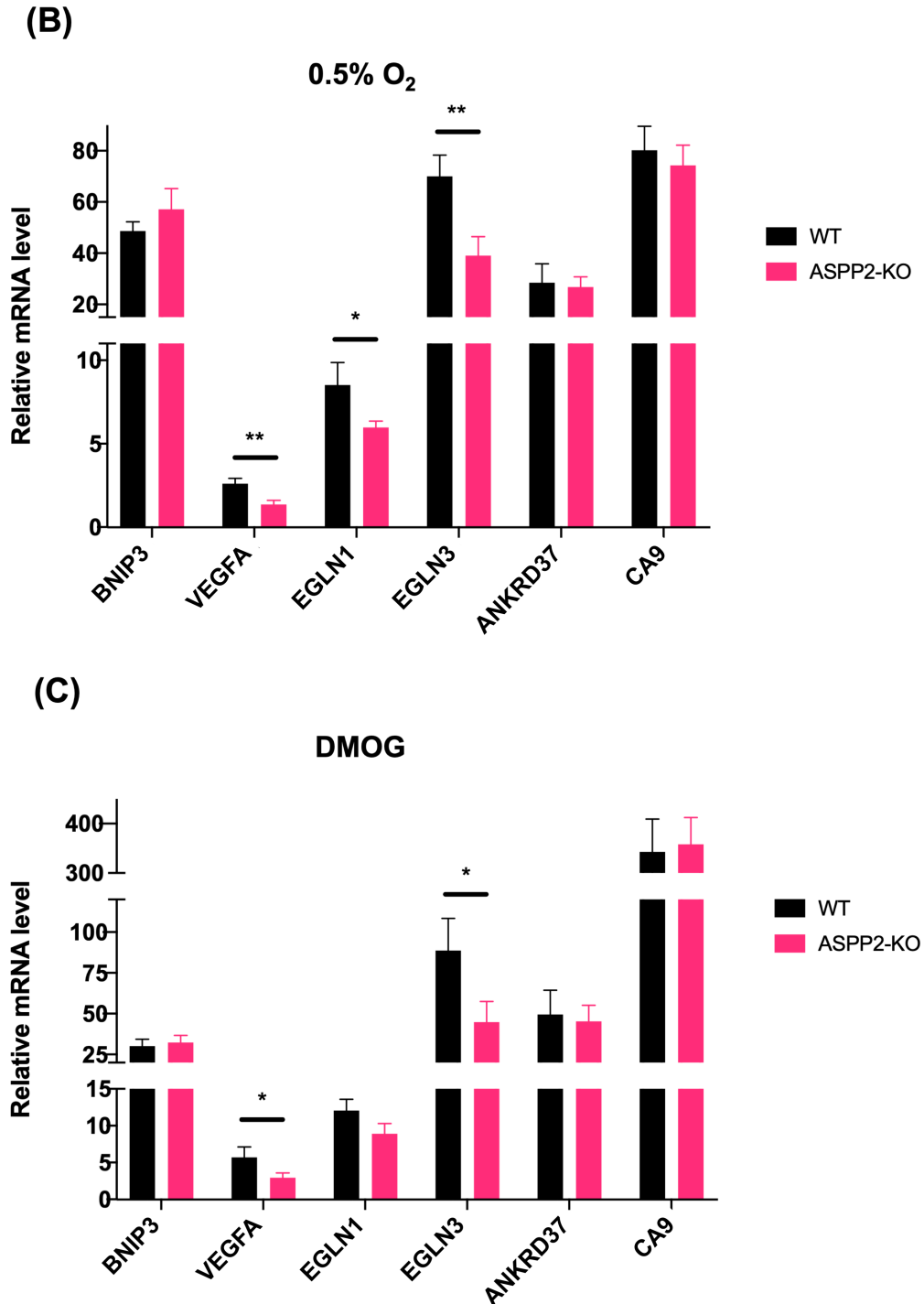


Figure 4.1: The loss of ASPP2 selectively affected the mRNA expression of HIF target genes under physiological hypoxia and hypoxia-mimicking condition. qPCR analysis of the gene expression of the indicated HIF target genes in WT and ASPP2-KO U2OS cells. Cells were incubated at (A) 21% or (B) 0.5% O₂ for 16 hours, or (C) treated with 1 mM of DMOG for 16 hours before RNA extraction. Values represented mean $\pm$ standard deviation; Experiment was performed in biological duplicate and technical triplicate; * $p < 0.05$, ** $p < 0.01$.

physiological hypoxia could not effectively address whether the modulatory effect of ASPP2 on HIF targets was dependent on FIH or not. For this, WT and ASPP2-KO cells were also treated with dimethyloxalylglycine (DMOG) for 16 hours. DMOG was a pan-hydroxylase inhibitor that was reported to inhibit both PHDs and FIH activity by 2OG competition and thereby inducing a hypoxia-mimicking condition (Jaakkola et al. 2001; David Lando et al. 2002; Yeh et al. 2017). As expected, DMOG treatment up-regulated the gene expression of all six HIF target genes to varying degrees [Figure 4.1C]. Consistent with our observation under physiological hypoxia, the loss of ASPP2 did not affect the expression of *BNIP3*, *ANKRD37* and *CA9*, but it led to a significant decrease of *EGLN3* and *VEGFA* expression ($p = 0.03$ and $p = 0.04$ for *EGLN3* and *VEGFA*, respectively) [Figure 4.1C]. *EGLN1* expression, in this case, was not significantly affected by the status of ASPP2 [Figure 4.1C], which might imply that it was not a high-confidence target of ASPP2, or different HIF target genes could be sensitive to the status of ASPP2 under different conditions.

Given that DMOG might exert other effects than inhibiting FIH activity and to rule out the possibility that DMOG might not completely eliminate FIH activity, qPCR analysis was also performed to measure the expression of the same HIF target gene panel in FIH-KO and DKO U2OS cells subject to either normoxia or physiological hypoxia. Interestingly, in the absence of FIH, the expression of all HIF targets in the panel, except for *BNIP3*, was sensitive to the loss of ASPP2 even under normoxia ($p = 0.04$ and $p = 0.03$ for *VEGFA* and *EGLN1*, respectively; $p = 0.02$ for *PHD3*, *ANKRD37* and *CA9*) [Figure 4.2A]. As expected, all six HIF target genes were induced under hypoxia. In accordance with our observation using WT and ASPP2-KO cells, the gene expression of *EGLN3* and *VEGFA* was

significantly reduced in DKO compared to FIH-KO cells ($p = 0.005$ and $p = 0.008$ for *EGLN3* and *VEGFA*, respectively) [Figure 4.2B]. However, the gene expression of *ANKRD37* and *CA9*, which were not sensitive to the status of ASPP2 in the presence of FIH, was also down-regulated in DKO cells relative to that in FIH-KO cells ($p = 0.03$ and $p = 0.02$ for *ANKRD37* and *CA9*, respectively) [Figure 4.2B].

Taken together, these results suggested that ASPP2 could selectively modulate the expression of HIF target genes, such as *EGLN3* and *VEGFA* in the selected HIF target gene panel, under both physiological hypoxia and hypoxia-mimicking condition induced by DMOG. This modulatory effect of ASPP2 was observed regardless of the status of FIH - when it was present, when its expression was knocked out, and when its hydroxylase activity was inhibited - suggesting that such effects did not require FIH. It should be noted that this modulatory effect of ASPP2 did not seem to be completely independent of FIH as the status of ASPP2 had different but overlapping effects on the selected panel of HIF targets between the presence and absence of FIH under normoxia and hypoxia. Potential implications of these observation would be elaborated further in the subsequent discussion. However, since our data suggested that this modulatory effect of ASPP2 did not necessarily require FIH, subsequent efforts will first focus on ASPP2 as the sole independent variable in order to elucidate the potential relationship between ASPP2 and the HIF pathway.

4.2.2 ASPP2 modulates the expression of a subset of HIF target genes irrespective of FIH status

To have a global view of the effects of ASPP2 on HIF target genes, RNA-sequencing (RNA-seq) analysis was performed using the aforementioned four cell lines - WT,

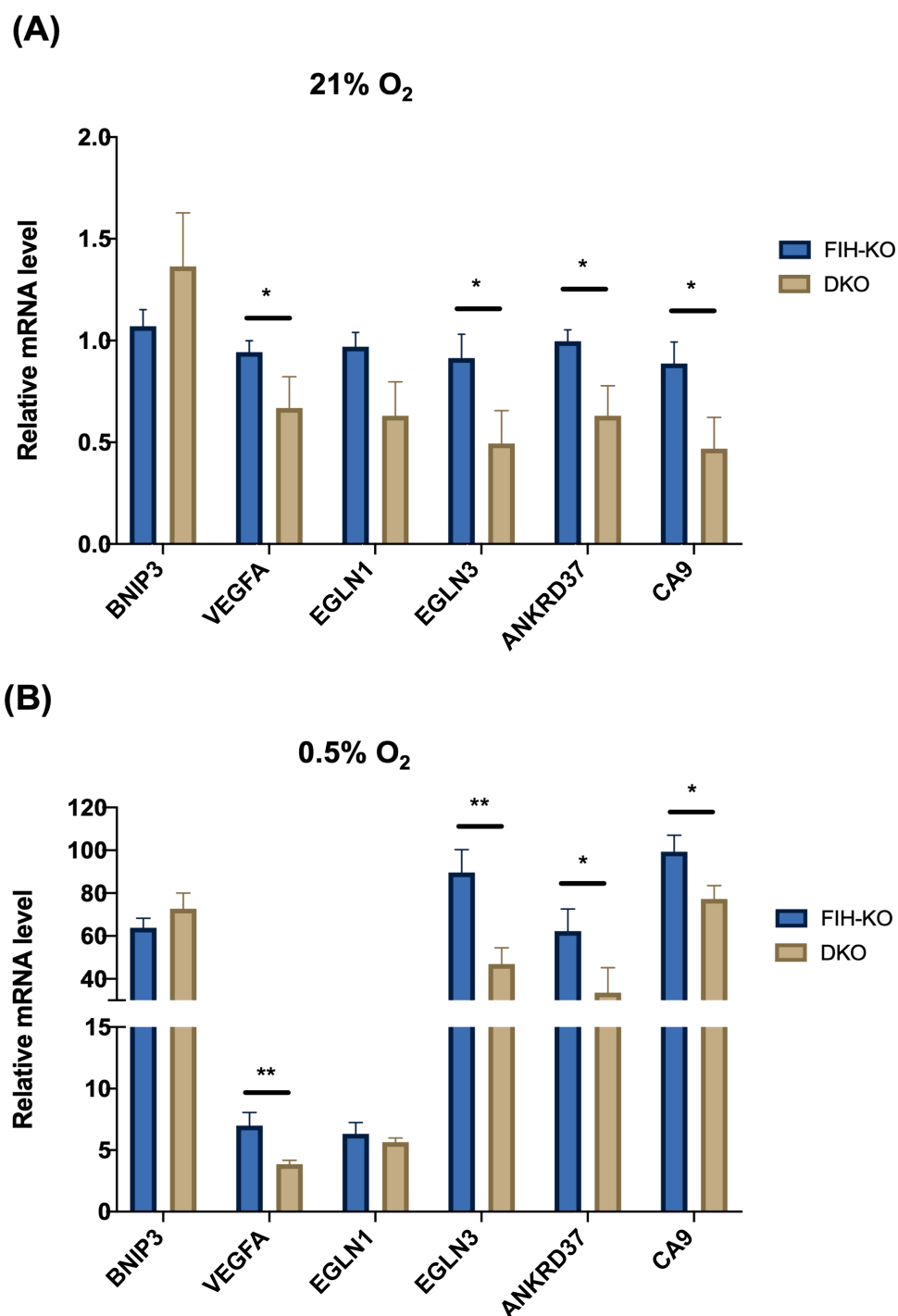


Figure 4.2: The loss of ASPP2 selectively affected the mRNA expression of HIF target genes under both normoxia and hypoxia in the absence of FIH. qPCR analysis of the gene expression of the indicated HIF target genes in FIH-KO and DKO U2OS cells. Cells were incubated at (A) 21% or (B) 0.5% O₂ for 16 hours before RNA extraction. Values represented mean $\pm$ standard deviation; Experiment was performed in biological duplicate and technical triplicate; *p < 0.05, **p < 0.01.

Gene Ontology (GO) terms (FDR < 0.05)
muscle organ development
striated muscle tissue development
muscle tissue development
skeletal muscle organ development
negative regulation of growth

Table 4.1: Top 5 enriched GO terms in genes downregulated in ASPP2-KO U2OS relative to WT

Gene Ontology (GO) terms (FDR < 0.05)
positive regulation of epithelial cell migration
respiratory tube development
lung development
small GTPase mediated signal transduction
regulation of epithelial cell migration

Table 4.2: Top 5 enriched GO terms in genes upregulated in ASPP2-KO U2OS relative to WT

ASPP2-KO, FIH-KO, and DKO U2OS cells in technical quadruplicate subject under both normoxia and hypoxia. To understand the effects of ASPP2 on HIF target gene expression, we performed differential gene expression analysis between WT and ASPP2-KO U2OS cells under hypoxia to identify ASPP2-sensitive genes under hypoxia. Using FDR less than 0.1 as the threshold, 4749 ASPP2-sensitive genes were identified and their enriched Gene Ontology terms were shown in Table 4.1 and Table 4.2.

By intersecting these 4749 ASPP2-sensitive genes under hypoxia with the 666 HIF target genes (for the detailed derivation of HIF target genes, please refer to subsection 2.7.4), we identified 297, out of the total 666, HIF target genes that were sensitive to ASPP2 status under hypoxia. This suggested that ASPP2 affects the transcript levels of a subset of HIF target genes, and their enriched

Gene Ontology (GO) terms (FDR < 0.05)
cellular metabolic process
cellular macromolecule metabolic process
organonitrogen compound metabolic process
nitrogen compound metabolic process
cellular protein metabolic process
primary metabolic process
organic substance metabolic process
macromolecule modification
metabolic process
cellular protein modification process
macromolecule metabolic process
protein metabolic process
glycogen metabolic process
cellular process

Table 4.3: Enriched GO terms in ASPP2-sensitive HIF target genes

Gene Ontology terms revealed that these genes are involved in various generic metabolic processes (Table 4.3).

Given the known interaction between ASPP2 and FIH, which is a negative regulator of HIF α , we further studied whether the effects of ASPP2 on HIF target genes are dependent on the status of FIH. To do that we performed differential expression analysis between FIH-KO and ASPP2-FIH-DoubleKO U2OS cells under hypoxia to identify the ASPP2-sensitive genes under hypoxia in the absence of FIH, then intersecting them with HIF target genes to identify the ASPP2-sensitive HIF target genes under hypoxia in the absence of FIH.

From this we identified 240 HIF targets that are sensitive to ASPP2 in the absence of FIH. By intersecting them with the ASPP2-sensitive HIF targets that we identified in the presence of FIH, we identified 69 HIF targets that are sensitive to ASPP2 irrespective of FIH status (Figure 4.3). No GO terms were found to

be statistically significantly enriched in these 69 genes.

On the other hand, we also identified 399 HIF targets that were sensitive to ASPP2 depending on FIH status (Figure 4.4). These genes can be further categorized into three major clusters: the first cluster contains HIF targets that were sensitive to ASPP2 only in the presence of FIH (Figure 4.4, top cluster); the second cluster contains HIF targets that were sensitive to ASPP2 only in the absence of FIH (Figure 4.4, middle cluster); and more interestingly, the third cluster contains ASPP2-sensitive HIF targets whose direction of expression regulation switches depending on FIH status (Figure 4.4, bottom cluster).

Such observation suggest that the status of FIH affects how ASPP2 modulates a subset of HIF target gene expression. However, this data alone could not tell us whether these are the direct effects due to the interaction between ASPP2 and FIH, or indirect effects, since FIH hydroxylates many other substrates. To circumvent this problem we considered using ASPP2 hydroxylation mutant, in which the two asparagine residues were mutated to alanine so that the mutant can no longer be hydroxylated by FIH. However, colleagues in Professor Chris Schofield group found that these two asparagine residues are structurally important as they serves as a crucial link for the SH3 domain of ASPP2 to fold back to the ankyrin repeat domain, so mutation to alanine potentially disrupts protein conformation (T. Leissing, Personal communication, 11 Mar, 2020). Due to this limitation in this thesis we would focus on the subset of HIF target genes that were sensitive to ASPP2 irrespective of FIH status.

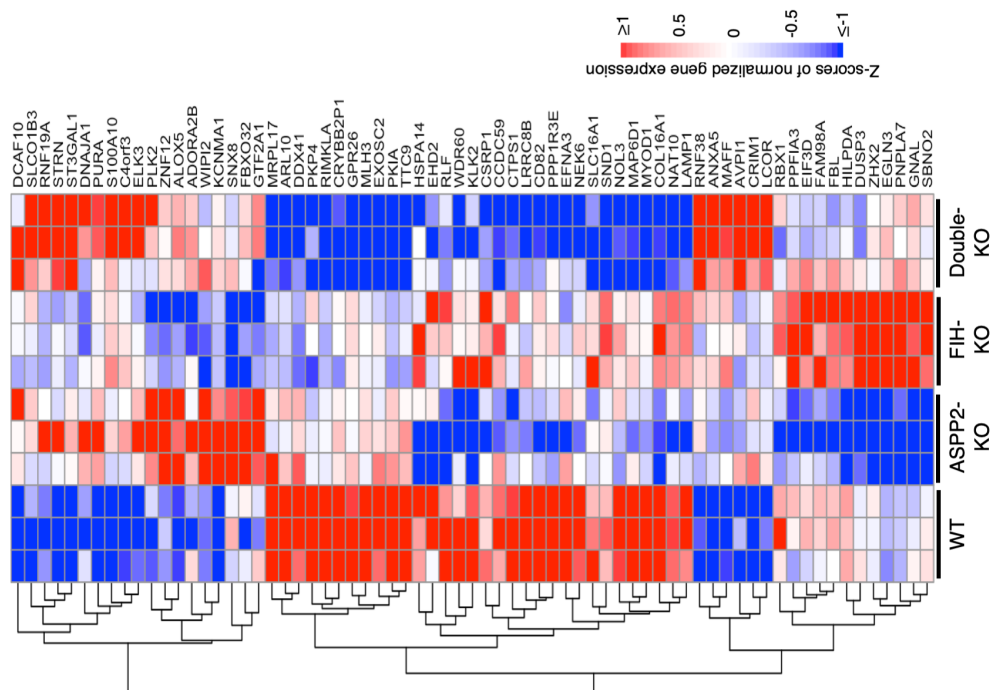


Figure 4.3: ASPP2 affects the gene expression of a subset of HIF target genes irrespective of FIH status. Clustered heatmap showing the expression of the HIF target genes that are sensitive to ASPP2 irrespective of FIH status in WT, ASPP2-KO, FIH-KO and Double-KO U2OS cells under hypoxia. Color key indicated the intensity associated with z-scores of normalized expression values.

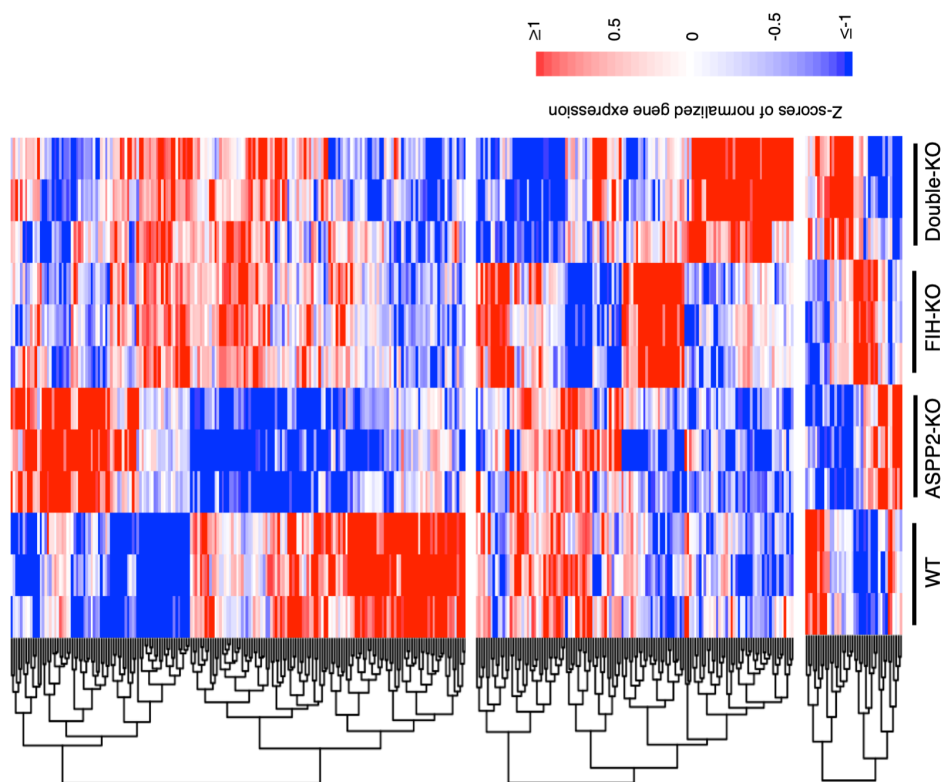


Figure 4.4: The status of FIH affects how ASPP2 affects the gene expression of a subset of HIF target genes. Clustered heatmap showing the expression of HIF target genes that are sensitive to ASPP2 depending on the status of FIH in WT, ASPP2-KO, FIH-KO and Double-KO U2OS cells under hypoxia. Color key indicated the intensity associated with z-scores of normalized expression values.

4.2.3 ASPP2 modulates PHD3 expression at the protein level under hypoxia

Our previous gene expression analyses using qPCR and RNA-seq suggested that the status of ASPP2 could affect the gene expression of 69 HIF target genes irrespective of FIH status (Figure 4.3). To assess whether this had any relevance at the protein level and whether it was only limited to one particular cell line, western blot analysis was performed using multiple types of cells subject to physiological hypoxia for 18 hours.

The types of cells used in this experiment included the WT and ASPP2-KO U2OS cells as described earlier [Figure 4.5A], WT and ASPP2-KO gastric adenocarcinoma AGS cells [Figure 4.5B], and primary astrocytes derived from WT and ASPP2-null littermate mice as a result of an exon 3 deletion in *TP53BP2* (ASPP2 Δ^3/Δ^3), which encodes ASPP2 [Figure 4.5C]. (ASPP2-KO AGS cells and primary astrocytes were kindly provided by Dr. Ludovico Buti and Mr. Jie Yin, respectively. For their cell culture maintenance, please kindly refer to subsection 2.3.1.)

Among these 69 ASPP2-sensitive HIF targets, *EGLN3*, which encodes the prolyl hydroxylase PHD3, is the better-known HIF target gene and is used by the Ratcliffe group as a positive control in HIFs ChIP experiments in different cell lines (O. Lombardi, Personal communication, 5 July, 2019). We therefore used it as the candidate ASPP2-sensitive HIF target for further investigation.

PHD3 protein expression was not detected in all three types of cells under normoxia, but was clearly induced after 18 hours of incubation at 0.5% O₂. Importantly, PHD3 expression was noticeably reduced in ASPP2-KO compared to wild-type of all three types of cells under hypoxia [Figure 4.5A-C], thus suggesting

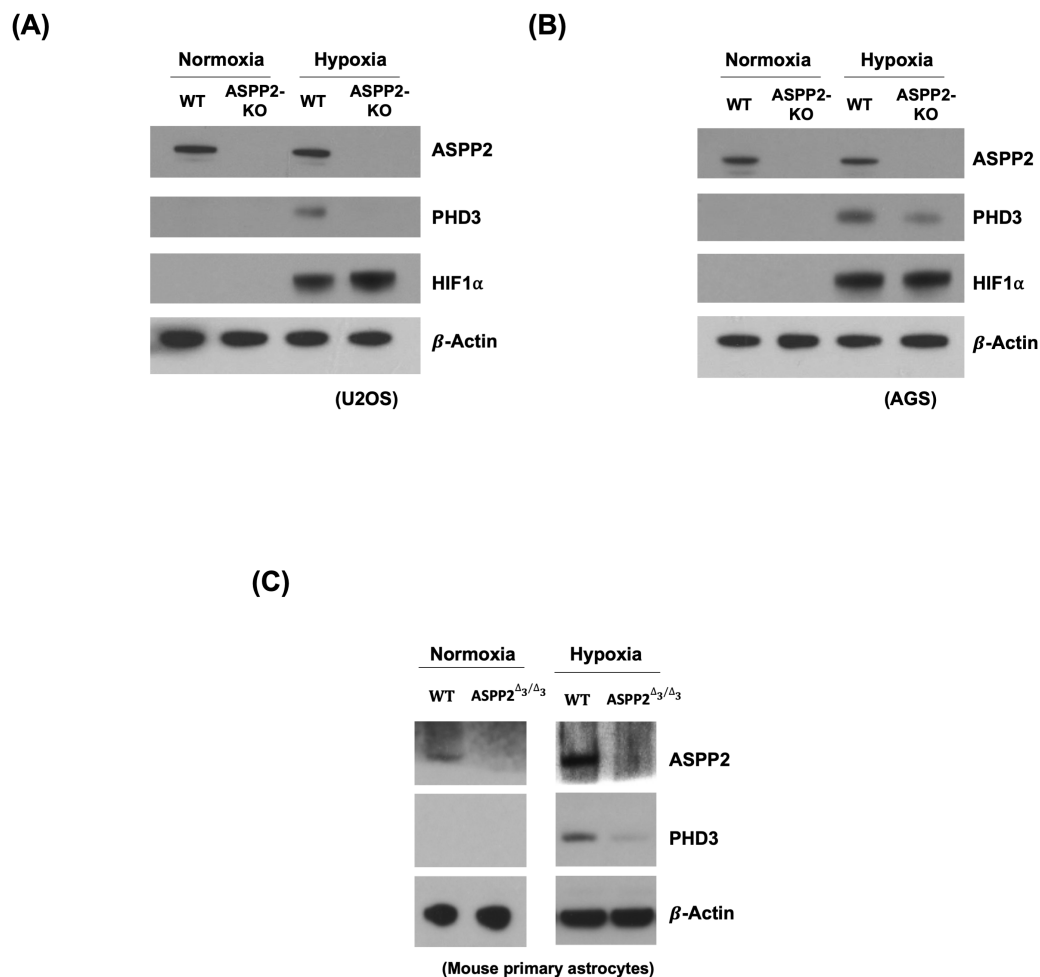


Figure 4.5: The loss of ASPP2 affects the protein expression of PHD3 under physiological in multiple types of cells. Western blot analysis of protein expression in cell lysates from WT and ASPP2-KO (KO) (A) U2OS and (B) AGS cells, and (C) ASPP2^{Δ3/Δ3} mouse primary astrocytes from littermates, using the indicated antibodies. Cells were incubated under 21% O₂ (normoxia) or 0.5% O₂ (hypoxia) for 18 hours before lysis. Results in (A) and (B) were representative of two independent experiments, and the result in (C) was representative of astrocytes cultured from two pairs of WT and ASPP2^{Δ3/Δ3} littermates.

that the modulatory effect of ASPP2 on PHD3 occurred at both the mRNA and protein levels.

4.2.4 Generation of stable ASPP2-KO U2OS cell line with inducible ASPP2 expression

To further confirm the reduction of PHD3 expression was due to the loss of ASPP2, an attempt was made to restore ASPP2 expression in ASPP2-KO cells to assess whether the phenotype could be rescued. Thus a stable ASPP2-KO U2OS cell line with doxycycline (dox)-inducible ASPP2 expression was generated using a piggyBac (PB) transposon-based system (Ding et al. 2005).

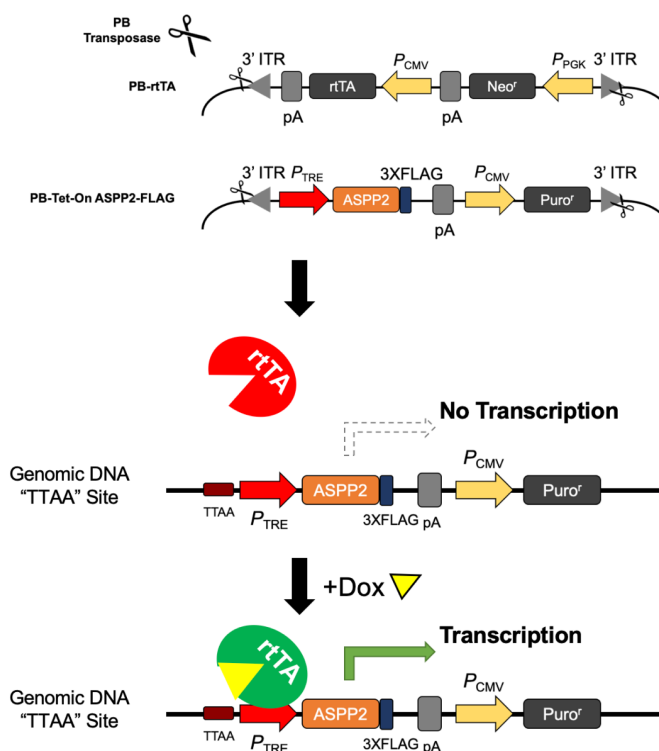


Figure 4.6: Design of the piggyBac Tet-On system to create inducible ASPP2 cell lines. Schematic depicting the three major components of the piggyBac Tet-On system: the piggyBac (PB) transposase recognized the inverted repeats flanking the cargo in the piggyBac vector, and transposed the cargo to "TTAA"-containing regions in the genome. The presence of the reverse tetracycline transactivator (rtTA) protein and doxycycline would activate the TRE promoter, leading to the transcription of *TP53BP2* encoding ASPP2.

The basis of the PB system could be understood as a "cut-and-paste" mechanism [Figure 4.6]. Our cargo-of-interest, i.e. ASPP2 cDNA, was cloned into a PB transposon vector, under the control of a dox-inducible promoter. It was fused at the C-terminal with a 3X-FLAG tag and flanked on both sides by a 13-bp sequence called inverted terminal repeats (ITRs) (Cary et al. 1989; Malcolm J. Fraser et al. 1995; M J Fraser et al. 1996). With the help of a helper plasmid that contained the PB transposase, which recognized the ITRs, it then excised our cargo of interest at the ITRs and ligated it into random sites in the target genome containing a "TTAA" tetranucleotide sequence (Malcolm J. Fraser et al. 1995; M J Fraser et al. 1996). This allowed for a more efficient transfer of our cargo of interest over traditional transfection due to the higher transposition rate than that of random insertion events (Ding et al. 2005). The inducible nature of this model was achieved by the Tet-On system, which required the co-transfection of a third plasmid that carried the reverse tetracycline transactivator (rtTA) protein (Gossen et al. 1995) and the neomycin resistance gene. The rtTA protein in the absence of dox is incapable of binding to the tetracycline response element (TRE) in the dox-inducible promoter. When bound by dox, however, rtTA binds to the TRE and transactivates the gene downstream. In this case, it enabled the controllable expression of FLAG-tagged ASPP2 by doxycycline treatment.

After neomycin selection and expansion of single-cell colonies, the single-cell clones were first analyzed by western blot to assess the expression of ASPP2. Selected single-cell clones B3 and 2B9 were incubated for 12 hours with cell culture medium containing the designated concentrations of doxycycline and cell lysates were subject to western blot analysis. In both clones, ASPP2 was detected by the anti-ASPP2

antibody at a molecular weight (175 kDa) similar to the endogenous ASPP2 in the WT U2OS, while no ASPP2 was detected in the ASPP2-KO cells, suggesting that the antibody was specific [Figure 4.7A]. The expression of inducible ASPP2 was restored to a level higher than the endogenous level in WT cells at the starting concentration of 50 ng/mL of dox and became saturated beyond 100 ng/mL of dox [Figure 4.7A]. However, basal leakage of ASPP2 expression was observed in clone 2B9 [Figure 4.7A], therefore clone B3 was selected to be used for subsequent experiments.

In order to assess whether the inducible ASPP2 had a subcellular localization similar to that of the endogenous ASPP2, immunofluorescence was performed using an anti-ASPP2 antibody different from the one used in the western blot analysis. As shown in the microscopy images, endogenous ASPP2 in the WT U2OS was predominantly junctional and it was barely detectable in the ASPP2-KO cells, suggesting that this staining of ASPP2 was specific [Figure 4.7B]. Dox treatment at 5 ng/mL induced ASPP2 in the inducible cell line (clone B3), with also a predominantly junctional localization similar to that of the endogenous ASPP2 [Figure 4.7B]. Increasing the concentration of dox to 50 ng/mL further increased the fluorescence intensity of ASPP2, suggesting that the expression level of inducible ASPP2 could be manipulated with varying levels of dox [Figure 4.7B]. The inducible ASPP2 at 50 ng/mL of Dox exhibited a clear junctional and cytoplasmic localization, which seemed logical since it was overexpressed at the given level of dox compared to the endogenous ASPP2 expression in the WT U2OS cells [Figure 4.7B].

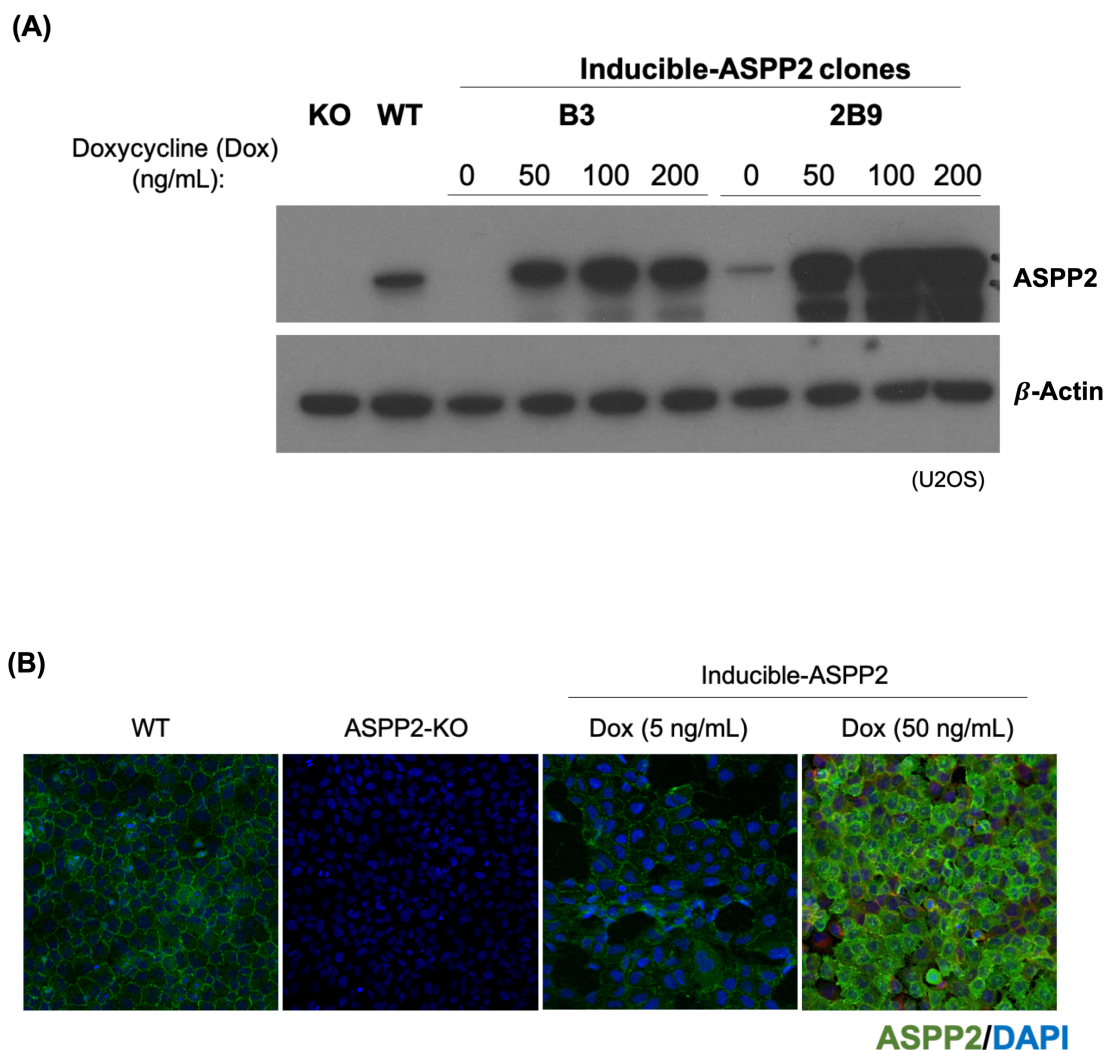


Figure 4.7: Validation of the stable cell lines expressing dox-inducible ASPP2. (A) Western blot analysis of the two clones of ASPP2-KO U2OS cells stably expressing doxycycline-inducible FLAG-tagged ASPP2, treated with increasing doses of doxycycline, using the indicated antibodies. WT and ASPP2-KO (KO) U2OS cells were used as positive and negative controls, respectively. β -Actin was used as the loading control. (B) Immunofluorescence staining of ASPP2 (green), with nuclei counterstained with DAPI (blue) in WT, ASPP2-KO, and ASPP2-KO U2OS cells stably expressing doxycycline-inducible FLAG-tagged ASPP2, treated with the indicated doses of doxycycline. Images are representative examples of two independent replicates of this experiment with at least three images taken per sample.

4.2.5 Restoration of ASPP2 expression in ASPP2-KO cells rescues PHD3 protein expression in hypoxia

After validating that the inducible ASPP2 could be recognized by anti-ASPP2 antibody by western blot and immunofluorescence analysis, and shared a similar molecular weight and subcellular localization to that of the endogenous ASPP2, ASPP2-KO U2OS cells stably expressing inducible FLAG-tagged ASPP2 (hereafter referred to as Tet-On-ASPP2 U2OS cells) was used to test whether PHD3 expression could be restored upon the reintroduction of ASPP2. It was first treated with or without 10 ng/mL of dox overnight to induce the expression of ASPP2. Together with WT and ASPP2-KO cells, they were then subject to normoxia or physiological hypoxia for 18 hours, followed by cell lysis and western blot analysis. Under normoxia, PHD3 expression was not detected in WT, ASPP2-KO or the Tet-On-ASPP2 U2OS cells, but it was clearly induced under hypoxia in all three cell lines [Figure 4.8]. Consistent with previous experiments, PHD3 expression was noticeably reduced in ASPP2-KO cells compared to that in the WT cells [Figure 4.8]. Importantly, PHD3 expression in the Tet-On-ASPP2 cell line without dox was similar to that in the ASPP2-KO cells, but upon induction of ASPP2 expression by dox treatment, PHD3 expression was restored to a level close to the endogenous level in the WT cells [Figure 4.8]. This result therefore indicated that the modulatory effect of ASPP2 on the expression of PHD3 was indeed due to ASPP2. However, it should be noted that this experiment lacked a negative control, i.e. ASPP2-KO U2OS cells stably expressing only the Tet-On system with an empty vector, to preclude potential artifacts due to doxycycline induction alone. Therefore it is critical that this cell line is generated and included in experiments alongside

the Tet-On-ASPP2 U2OS cell line in order to prevent any effects of doxycycline from confounding experimental results.

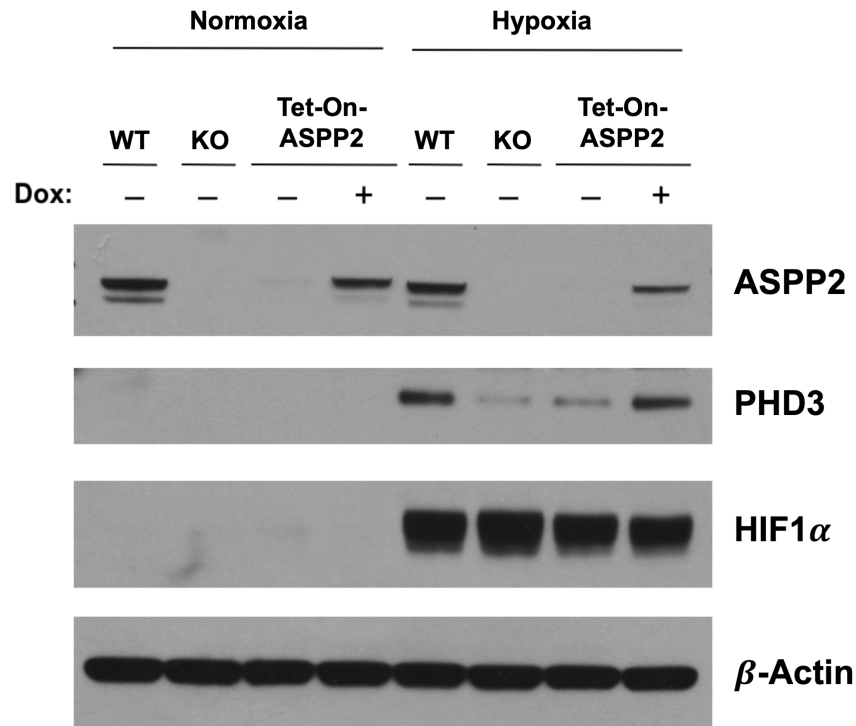


Figure 4.8: Re-introducing ASPP2 rescues the reduction of PHD3 expression under hypoxia. Representative western blot of protein expression using the indicated antibodies in cell lysates of WT, ASPP2-KO, and Tet-On-ASPP2 U2OS cells treated with or without doxycycline (dox). Cells were incubated under normoxia or hypoxia (0.5% O₂) for 16 hours before lysis; This result was representative of two independent experiments.

4.2.6 ASPP2 does not affect the protein expression of HIF and p300 under hypoxia

So far our data suggested that ASPP2 regulates the expression of PHD3 under hypoxia, but the underlying mechanism(s) has not been established. Since PHD3 is a well-established direct HIF target gene, it was hypothesized that ASPP2 can regulate PHD3 by directly affecting HIF.

The HIF pathway is directly regulated at several different stages [Figure

4.9]. Among them, arguably the most important and well-studied one is the oxygen-dependent regulation of HIF α 's protein expression by the PHDs prolyl hydroxylases, which formed the core of the mammalian oxygen sensing machinery (Kaelin et al. 2008).

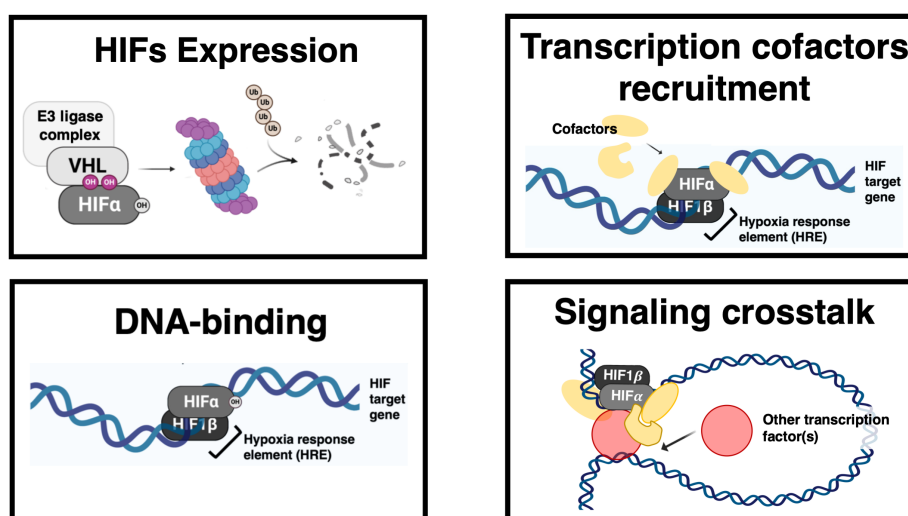


Figure 4.9: Major points of regulation of the HIF pathway. Schematic illustrating the major points of regulation of the HIF pathway.

Therefore, to examine our hypothesis, it was first investigated that whether ASPP2 can affect the protein expression of HIF α . WT and ASPP2-KO cells were subject to physiological hypoxia and their expression of HIF1 α and HIF2 α were tracked at the 0th, 4th, 8th, 16th and 24th hours of hypoxia and analyzed by western blot.

As expected, HIF1 α was not detected in normoxia (indicated in Figure 4.10 as 0-hour), but it was detected after 4 hours of hypoxia and its expression peaked at the eight hours of hypoxia [Figure 4.10]. This was followed by a decrease after

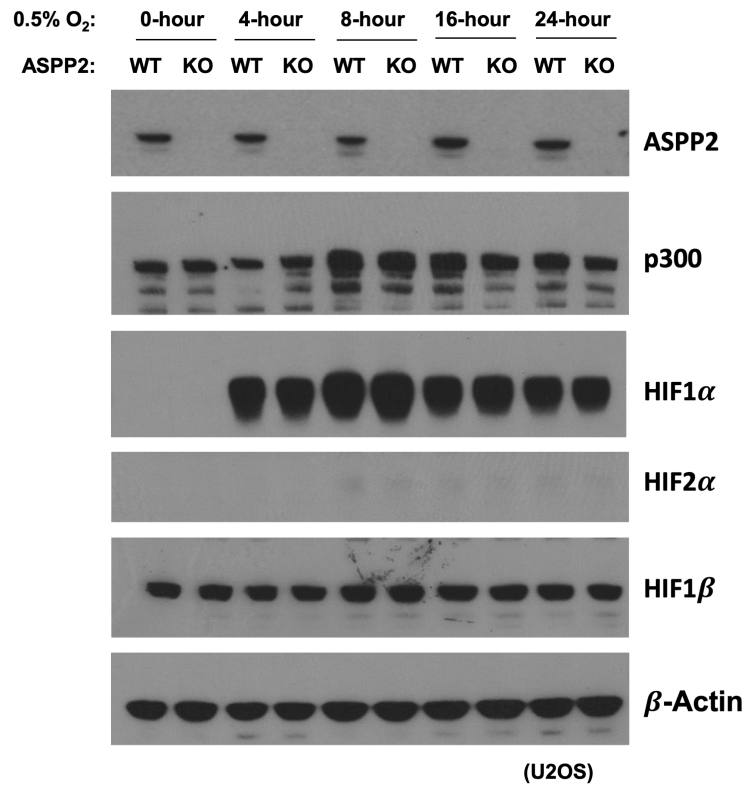


Figure 4.10: The loss of ASPP2 did not influence the protein expression of HIF and p300 under physiological hypoxia. Western blot of protein expression using the indicated antibodies in WT and ASPP2-KO U2OS cells lysed at the designated timepoints after incubation at 0.5% O₂.

16 hours of hypoxia and a further decline after 24 hours. This decline was to a level lower than that seen after 4 hours of hypoxia but was still detectable [Figure 4.10]. This result was in concordance with various reports that HIF α was subject to various multiple negative feedback mechanisms under prolonged hypoxia (Appelhoff et al. 2004; Cavadas et al. 2015; Horak et al. 2010; Marxsen et al. 2004; Stiehl et al. 2006; Uchida et al. 2004). However, there was no observable difference of HIF1 α expression between WT and ASPP2-KO cells at all the measured timepoints under hypoxia [Figure 4.10].

On the other hand, HIF2 α expression was much weaker than that of HIF1 α at all the time-points measured [Figure 4.10]. It was not expressed until eight

hours of hypoxia, when it was very faintly detected. This decreased after 16 and 24 hours of hypoxia [Figure 4.10]. Importantly, there seemed to be no difference of HIF2 α expression between WT and ASPP2-KO cells at all five timepoints measured under hypoxia [Figure 4.10].

As previously discussed, HIF α does not function alone but functions as a transcription factor by forming a heterodimer with ARNT (HIF1 β). It also recruits transcriptional coactivators including p300 as part of the transcriptional machinery. In contrast to the oxygen-dependent nature of HIF α , the expression of HIF1 β and p300 were constitutively expressed. Nevertheless, their expression were also measured in this hypoxia time-course experiment to test if ASPP2 affected their expression. Consistent with the literature and similar to the expression of housekeeping gene β -Actin, the expression of HIF1 β was stable over all five time-points measured and there was no observable difference between WT and ASPP2-KO cells [Figure 4.10]. The expression of p300 was also similar between WT and ASPP2-KO cells at all time-points, although it seemed to have an increased expression at the eighth hour under hypoxia in both WT and ASPP2-KO cells, coinciding with the peak of HIF1 α expression [Figure 4.10].

Taken together, these results suggested that ASPP2 did not influence the protein expression of HIF1 α , HIF2 α , ARNT and p300 under hypoxia, at least up to 24 hours at 0.5% O₂.

4.2.7 ASPP2 does not affect the subcellular localization of HIF1 α under hypoxia

In addition to the protein expression of HIF and p300 under hypoxia, the subcellular localization of HIF was the next point of regulation that was considered, as the

whereabouts of HIF is key to its function as a transcription factor. To test if ASPP2 can influence HIF subcellular localization, immunofluorescence analysis was performed in WT and ASPP2-KO U2OS cells before and after hypoxia. Based on the results of the previous time-course experiment, where HIF2 α was minimally expressed, only HIF1 α was analyzed in the immunofluorescence analysis.

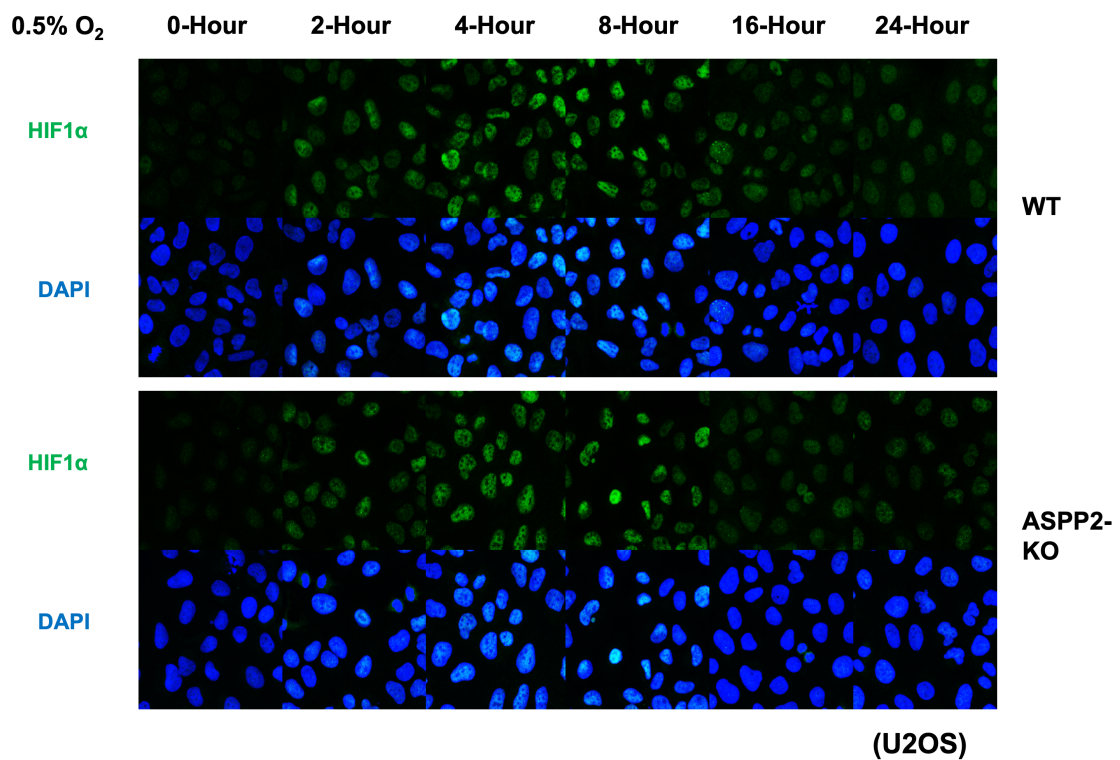


Figure 4.11: The loss of ASPP2 does not affect the subcellular localization of HIF1 α under hypoxia. Immunofluorescence images of HIF1 α (green), with nuclei counterstained with DAPI (blue). WT (top panel) and ASPP2-KO (bottom panel) U2OS cells were incubated at 0.5% O₂ and fixed at the designated timepoints. Images are representative examples of two independent replicates of this experiment with at least three images taken per sample.

HIF1 α was barely detectable in the nucleus of WT cells in normoxia, suggesting that HIF1 α was still expressed in normoxia, but at a very low level [Figure 4.11]. After two hours or more of hypoxia, HIF1 α was clearly induced, predominantly

in the nucleus [Figure 4.11]. In keeping with the previous time-course western blot analysis, the fluorescence intensity of nuclear HIF1 α was diminished after 16 and 24 hours of hypoxia [Figure 4.11]. Importantly, HIF1 α had a primarily nuclear subcellular localization in both WT and ASPP2-KO cells [Figure 4.11]. This suggested that ASPP2 did not affect the subcellular localization of HIF1 α .

4.2.8 The loss of ASPP2 does not affect HIF binding to HRE-containing regions within/near *EGLN3* gene locus under hypoxia

Although our data suggests that ASPP2 does not affect the protein expression and subcellular localization of HIF1 α , the transcriptional activity of HIF1 α may still be directly regulated independently of its protein expression and subcellular localization (Galbraith et al. 2013; J. S. Lee et al. 2011; Perez-Perri et al. 2016; Sena et al. 2013). The transcriptional activity of HIF1 α also depends on its ability to directly bind to sites containing HRE sequences in the genome as a heterodimer with ARNT. Therefore it was investigated whether ASPP2 could affect the level of HIF binding to the HREs within and near PHD3 gene locus. For this, ChIP coupled with qPCR analysis (ChIP-qPCR) was performed.

4.2.8.1 Optimization of ChIP protocol to assess the HRE-binding activity of HIF

The protocol of ChIP-qPCR was first optimized to ensure the chromatin was fragmented into optimal sizes (100-600 base pairs (bp) for ChIP-qPCR, 150-300 bp for ChIP-seq). To do this, a time-course pilot sonication experiment was performed on the cell lysates of hypoxia-treated and formaldehyde-fixed WT U2OS cells. The purified DNA from different sonicated time-points were then analyzed by

electrophoresis, using the Agilent TapeStation system. The size distribution of the DNA fragments was illustrated in the electropherograms [Figure 4.12]. 20 minutes of sonication produced DNA fragments with sizes ranged from ~100 to ~600 bp, with the majority of them centering at ~300 bp. Further increasing the sonication time did not lead to significant changes, therefore 20-minute was used for sonicating the chromatin for the ChIP-qPCR experiments in this study unless stated otherwise.

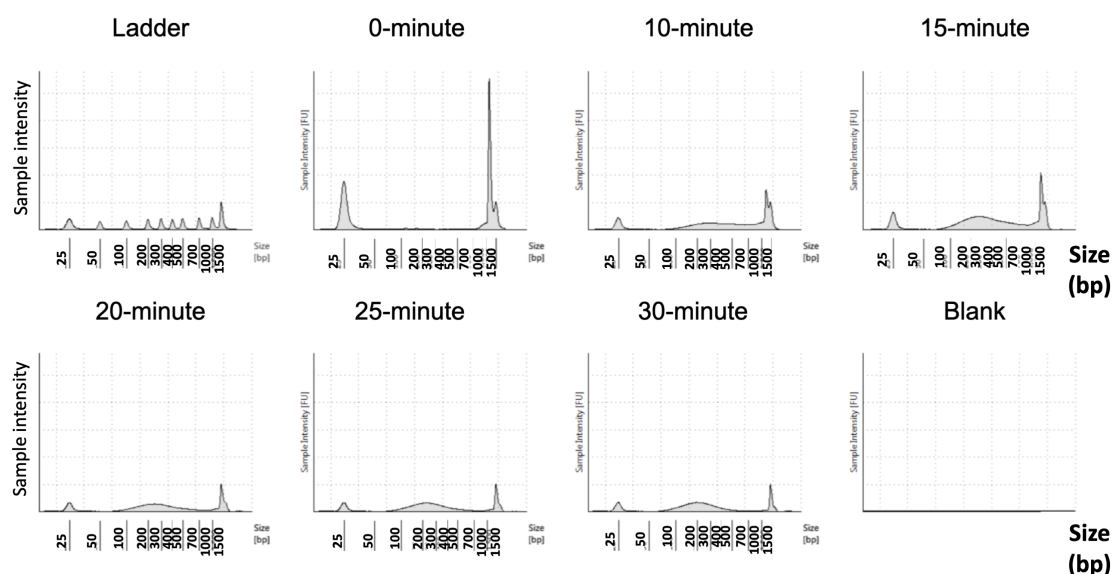


Figure 4.12: Optimization of chromatin fragment size for ChIP. Electropherograms of purified sheared DNA from WT U2OS cell lysates sonicated for the indicated length of time. D1000 ladder and blank sample were used as controls. The rightmost and leftmost peaks represented the upper and lower markers, respectively.

Further adjustments were made to optimize the ChIP protocol in order to raise the signal-to-noise ratio. They included

- (i) Magnetic beads coupled covalently with protein G was blocked with acetylated BSA (Ac-BSA) to cover unspecific sites on the beads. Ac-BSA was used instead

of BSA as acetylation inactivates any potential nuclease activity in the BSA that might interfere with the chromatin samples;

(ii) Fixed, sonicated cell lysates were pre-cleared with Ac-BSA blocked magnetic beads for four hours before performing the ChIP step to remove any non-specific proteins binding to the magnetic beads;

(iii) DNA-LoBind tubes were used to improve chromatin recovery during the washing steps and each sample was transferred to a new tube in each washing step to reduce potential unspecific proteins or DNA sticking to the tubes;

(iv) a higher volume (240 μ L) of nuclease-free water was used to elute the purified DNA instead of the recommended 20 μ L to reduce the undesirable effects due to pipetting errors.

HIF1 β and Immunoglobulin G (IgG) ChIP-qPCRs were then performed on hypoxia-treated WT U2OS cells using this modified protocol to assess whether the technique worked. The chromatin fragment size was again measured to ensure that it was within the optimal range [Figure 4.13A].

Percentage of input DNA was used to represent the enrichment of HIF1 β and IgG at the HRE-containing regions of HIF target genes *EGLN3* and *NDRG1* [Figure 4.13B]. A region in *ACTG1*, where no binding of HIFs (including HIF1 α , HIF2 α and HIF1 β) was detected based on multiple in-house ChIP-seq data of the Ratcliffe group (data not shown), was used as a negative control locus in addition to the IgG ChIP [Figure 4.13B]. HIF1 β showed significant enrichment at the HRE-containing regions of both *EGLN3* and *NDRG1* in comparison to the negative control locus of *ACTG1* ($p = 0.001$ for *EGLN3*, $p = 0.02$ for *NDRG1*) [Figure 4.13B]. The enrichment of HIF1 β at the HRE-containing regions of *EGLN3* and

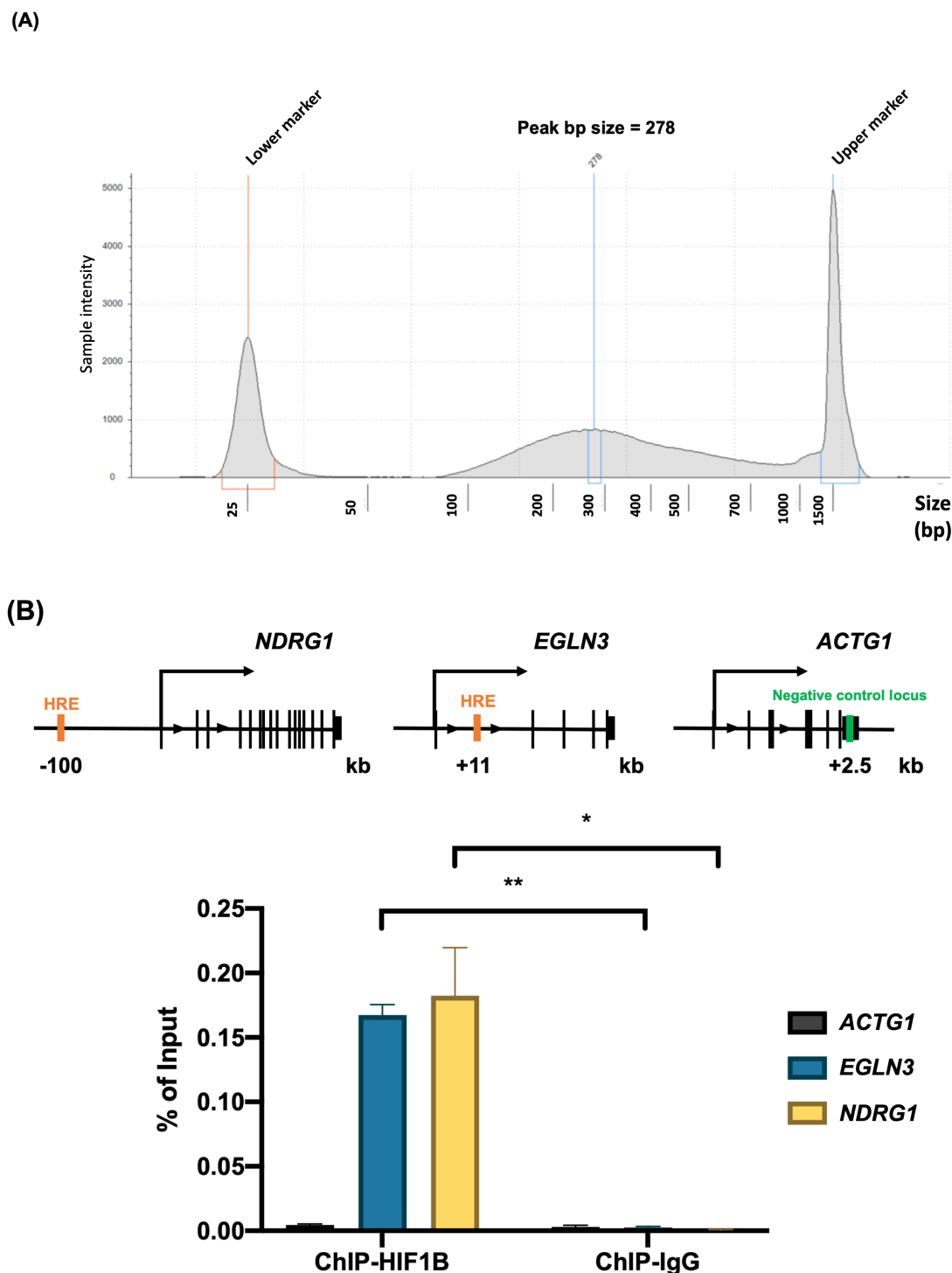


Figure 4.13: Quality control of the ChIP-qPCR protocol. (A) Electropherogram of purified sheared DNA from the 2% genomic DNA input of WT U2OS cell lysates. Cell lysates were sonicated for 20 minutes using Bioruptor as described in chapter 2. The rightmost and leftmost peaks represented the upper and lower markers, respectively; (B) Schematics of ChIP-qPCR amplicons and ChIP-qPCR analyses with the indicated antibodies at the enhancers of EGLN3 and NDRG1 or the non-specific region of G-Actin. Values were mean $\pm$ standard deviation; Biological duplicate and technical triplicate for each group; * $p < 0.05$, ** $p < 0.01$.

NDRG1, respectively, was also significantly greater than that of the IgG control ($p = 0.001$ for *EGLN3*, $p = 0.02$ for *NDRG1*), while the IgG control had insignificant enrichment at the HRE-containing regions of *EGLN3* and *NDRG1* compared to the negative control locus of *ACTG1* [Figure 4.13B].

Taken together, these data suggested that the ChIP protocol was reasonably optimized and could be used to assess the DNA-binding level of our transcription factors of interest.

4.2.8.2 The loss of ASPP2 does not affect HIF binding to HRE-containing regions within/near *EGLN3* gene locus under hypoxia

In order to investigate whether ASPP2 influenced HIF binding to the HRE-containing regions within/near *EGLN3* gene locus, ChIP-qPCR was performed using WT and ASPP2-KO U2OS cells subject to eight hours of physiological hypoxia. This duration of hypoxia incubation was decided based on our earlier finding that HIF1 α had the highest expression after eight hours of hypoxia [Figure 4.10]. ChIP was performed against all the major nuclear effectors of the HIF pathways, including HIF1 β , HIF1 α and HIF2 α and the enrichment of these transcription factors at the HRE-containing regions within or near *EGLN3* gene locus were measured by qPCR and represented as percentage of the genomic DNA input. Two HRE-containing regions for *EGLN3* - *EGLN3_A* and *EGLN3_B* - were analyzed as these two regions were found to contain HIF-ChIP peaks in multiple in-house ChIP-seq datasets from the Ratcliffe group under hypoxia in multiple cell lines, including U2OS cells [Figure 4.14].

Both HIF1 β and HIF1 α were significantly enriched at *EGLN3_A* and *EGLN3_B* compared to the negative control locus of *ACTG1* in both WT and ASPP2-KO cells

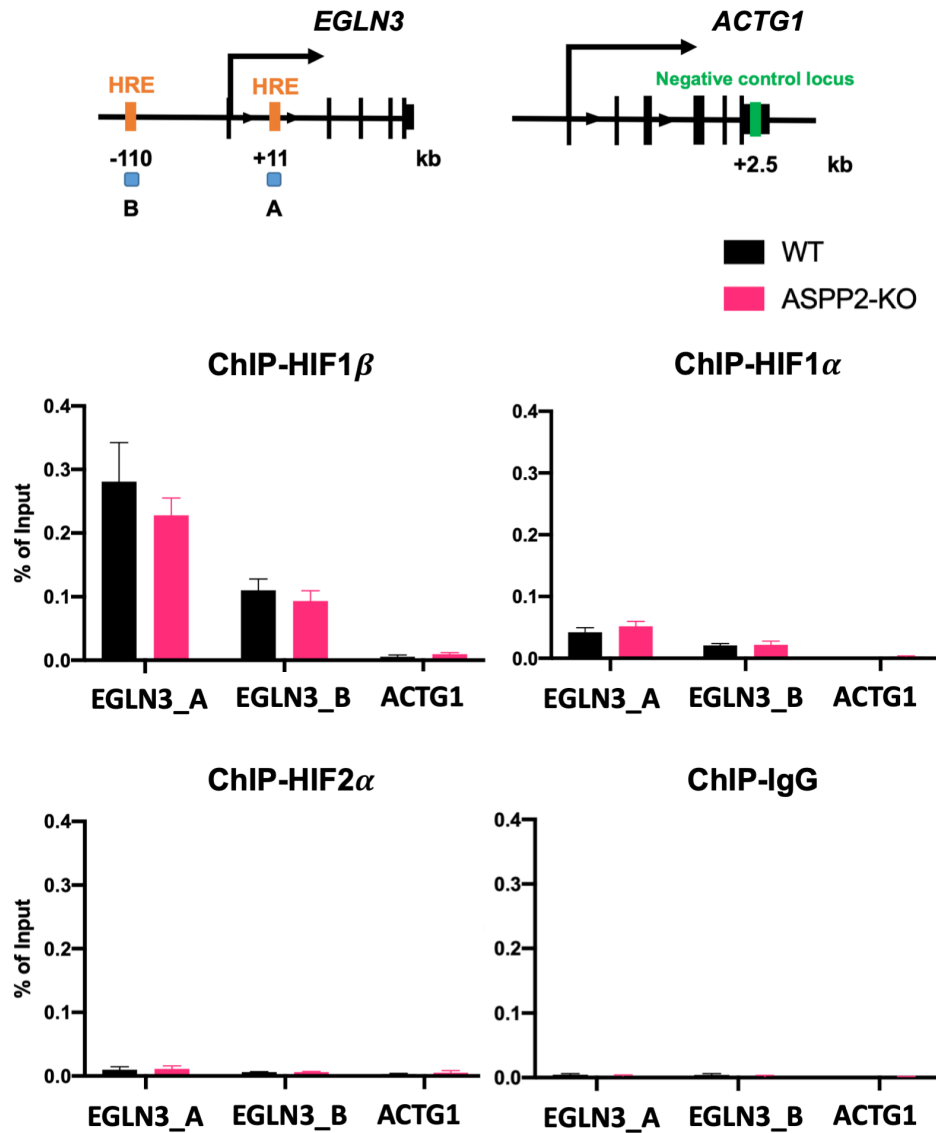


Figure 4.14: The loss of ASPP2 does not affect HIF binding to the HRE-containing loci of *EGLN3* under hypoxia. Schematics of ChIP-qPCR amplicons and ChIP-qPCR analyses with the indicated antibodies at the enhancers of *EGLN3* or non-specific region of *ACTG1*. ChIP experiment was performed using WT and ASPP2-KO U2OS cells incubated under 0.5% O₂ for 8 hours. Values were mean $\pm$ standard deviation; Biological duplicate and technical duplicate for each locus.

[Figure 4.14]. Both transcription factors also had significantly higher enrichment at these two *EGLN3* HRE-containing loci than that of the IgG control [Figure 4.14]. HIF2 α , on the other hand, was not enriched at *EGLN3_A* and *EGLN3_B* when compared internally to the negative control region of *ACTG1* or to that of the IgG control [Figure 4.14]. This was likely due to its very low expression at the eighth hour of hypoxia [Figure 4.10]. Importantly, there was little difference between WT and ASPP2-KO cells in terms of the enrichment of HIF1 β and HIF1 α at *EGLN3_A* and *EGLN3_B* [Figure 4.14], suggesting that ASPP2 did not affect HIF binding to the selected two HRE-containing regions within/near *EGLN3* gene locus under hypoxia.

4.2.9 The loss of ASPP2 does not affect p300 or p53 binding to the HRE-containing regions within/near PHD3 gene locus under hypoxia

In addition to protein expression, subcellular localization and HRE-binding, the recruitment of transcriptional cofactors is also an important point of regulation of HIF transcriptional activity. p300, one of the first HIF transcriptional cofactors to be identified, is known to be recruited to the HIF transcriptional complex to promote HIF transcriptional activity (Arany et al. 1996; S. J. Freedman et al. 2002). Notably, it was reported that ASPP2 can also directly interact with p300 to regulate the transcriptional activity of p53 (Gillot et al. 2011). ASPP2 has long been reported to regulate p53 and selectively promotes its apoptotic function. Interestingly, various interactions between HIF1 α and p53 have been reported and it has been suggested that HIF1 α could directly interact with p53 and modulate its transcriptional activity (An et al. 1998; Sánchez-Puig et al. 2005; Suzuki et al. 2001). However, it is not known whether p53 can also be recruited to the HIF transcriptional

complex and influence the transactivation of HIF target genes. Based on the reported relationships between ASPP2, p53 and p300, we hypothesized that ASPP2 may affect the recruitment of either or both of these two factors by HIF to the HIF-binding sites targeting PHD3.

4.2.9.1 Examining p53 response in U2OS cells upon nutlin-3a activation

Before testing this hypothesis, we first established whether our U2OS cell lines possessed a functional wild-type p53 response, and whether p53 could bind to some of its better known target genes when activated. For this, WT and ASPP2-KO cells were first treated with 5 μ M nutlin-3a for 24 hours. Nutlin-3a is a widely used p53 activator that stabilizes p53 by inhibiting its interaction with MDM2 (Vassilev et al. 2004). The expression of p53 target genes including *CDKN1A*, *MDM2* and *BAX*, were then analyzed by qPCR. Compared to DMSO treatment, 24 hours of nutlin-3a treatment induced significant expression of all three p53 targets in WT U2OS cells [Figure 4.15A].

The ability of p53 to bind to these p53 target genes in WT U2OS cells was also validated by ChIP-qPCR. Since the anti-p53 ChIP antibody used in our lab's previous p53 ChIP-seq experiments was discontinued, three different anti-p53 ChIP antibodies were used in the ChIP-qPCR and their efficiencies were compared with that of rabbit and mouse IgG. As shown in Figure 4.15B, eight hours of 5 μ M nutlin-3a treatment induced a significantly higher enrichment of p53 at the p53-binding sites of *BAX*, *MDM2* and *CDKN1A* than both the mouse and rabbit IgG. In comparison, little p53 enrichment was detected in the non-p53-binding region of *ACTG1* in both DMSO- and nutlin-3a-treated conditions [Figure 4.15B]. Among

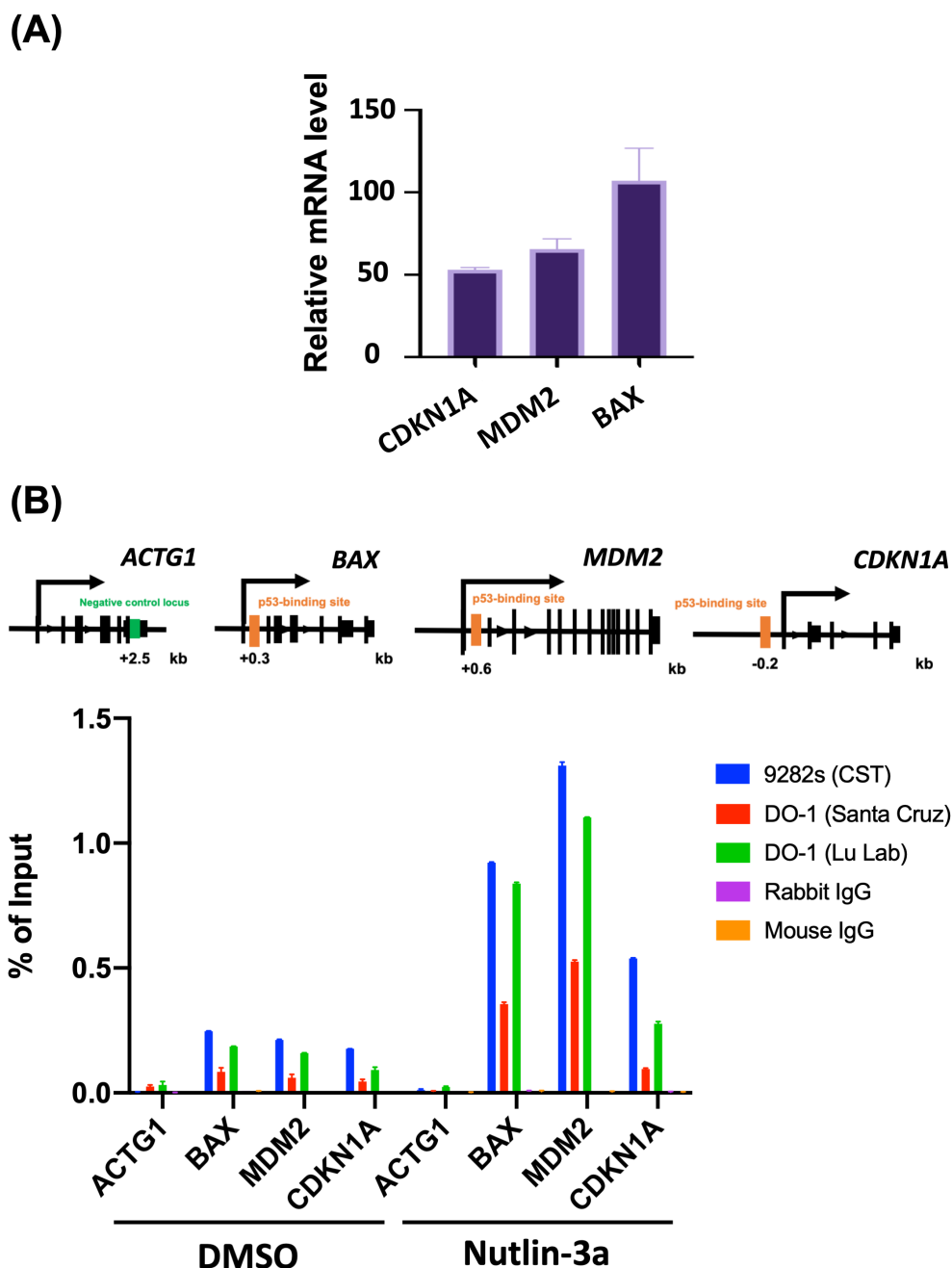


Figure 4.15: The U2OS cell line shows functional p53 transcriptional response upon nutlin-3a activation. (A) qPCR analyses of the expression of selected p53 target genes in nutlin-3a treated WT U2OS cells. Gene expression was normalized to that in the DMSO-treated WT U2OS cells. Values were mean $\pm$ standard deviation; Biological duplicate and technical triplicate for each group; (B) Schematics of ChIP-qPCR amplicons and ChIP-qPCR analyses with the indicated antibodies at the enhancers of *BAX*, *CDKN1A*, and *MDM2*, as well as at the non-specific region of *ACTG1*. ChIP experiment was performed using WT U2OS cells treated with DMSO or nutlin-3a. Values were mean $\pm$ standard deviation; One biological replicate and technical duplicate for each group;

the three anti-p53 ChIP antibodies tested, 9282S from Cell Signalling Technology gave the highest p53 enrichment at all the three p53 target gene loci [Figure 4.15B], therefore it was selected for subsequent p53 ChIP experiments. Overall, these results suggest that our WT U2OS cells possess a functional wild-type p53 response and we then examined whether ASPP2 could influence the recruitment of p53 or p300 to HIF-binding sites of *EGLN3*.

4.2.9.2 The loss of ASPP2 does not affect p300 or p53 binding to the HRE-containing regions within/near *EGLN3* gene locus under hypoxia

To test our hypothesis, p53- and p300-ChIPs were performed in parallel with HIF1 β -ChIP in hypoxia-treated WT and ASPP2-KO U2OS cells and the enrichment of these transcription factors at the two HRE-containing regions of *EGLN3* was analyzed by qPCR [Figure 4.16].

Consistent with previous results, HIF1 β was significantly enriched at the two HIF-binding sites of *EGLN3*, i.e. *EGLN3_A* and *EGLN3_B*, relative to the non-HIF-binding locus of *ACTG1* but there was no noticeable difference between WT and ASPP2-KO cells [Figure 4.15]. It was also not significantly enriched at the p53-binding locus of *CDKN1A* [Figure 4.16]. For p53-ChIP, while p53 enrichment at the p53-binding locus of *CDKN1A* was significantly higher than that of the non-p53-binding site of *ACTG1*, there was no noticeable enrichment of p53 at the two HRE-binding loci of *EGLN3* relative to the non-HIF-binding locus of *ACTG1* [Figure 4.16]. This suggests that p53 was not recruited to these two HIF-binding sites of PHD3 under hypoxia. For p300-ChIP, the enrichment of p300 at the two HRE-containing loci of *EGLN3* was at least over four times greater than

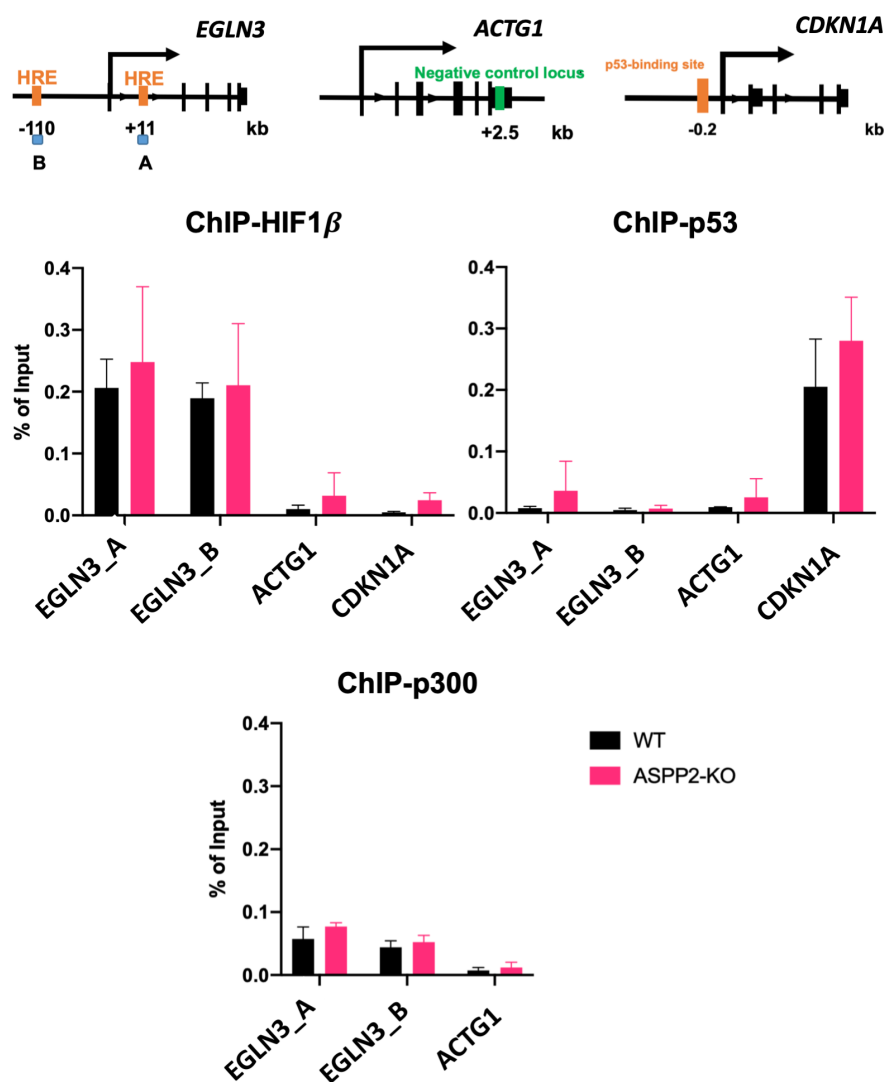


Figure 4.16: The loss of ASPP2 did not affect the recruitment of p300 or p53 to the HIF-binding loci of *EGLN3* under hypoxia. Schematics of ChIP-qPCR amplicons (top panel) and ChIP-qPCR analyses (bottom panel) with the indicated antibodies at the enhancer regions of *EGLN3* and *CDKN1A* and the non-specific region of *ACTG1*. Values were mean $\pm$ standard deviation; Biological duplicate and technical triplicate for each group.

that at the non-HIF-binding locus of *ACTG1*, however, there was no observable difference of p300 enrichment at the two HRE-containing loci of *EGLN3* between WT and ASPP2-KO cells [Figure 4.16].

Taken together, these results suggest that ASPP2 did not have an observable influence on the recruitment of either p300 or p53 to the HIF-binding loci of PHD3.

4.2.10 Echinomycin abolishes ASPP2's regulation of *EGLN3* expression

Since the HIF-binding loci analyzed by ChIP-qPCR was limited and genome-wide analysis of HIF-binding has not been conducted in this study, we next used an alternative approach to preliminarily assess whether the regulation of *EGLN3* by ASPP2 was dependent on HIF HRE-binding activity. For this, echinomycin, a small-molecule inhibitor of HIF DNA-binding activity, was used. Echinomycin recognizes and binds to HREs in the genomic DNA, thereby blocking HIF from binding to them and inhibiting its transcriptional activity (Kong et al. 2005).

In keeping with previous results, the loss of ASPP2 led to a reduction of the expression of *EGLN3* ($p = 0.03$), but not *ANKRD37*, under DMOG-treated condition [Figure 4.17B]. As expected, the introduction of echinomycin greatly reduced the induction of both *EGLN3* and *ANKRD37* expression (from as high as ~100-fold to under 4-fold for *EGLN3* and from above 70-fold to under 6-fold for *ANKRD37*) [Figure 4.17B]. Interestingly, the introduction of echinomycin abolished the difference of *EGLN3* expression between DMOG-treated WT and ASPP2-KO cells, suggesting that the regulation of *EGLN3* expression by ASPP2 might require the HRE-binding activity of HIF.

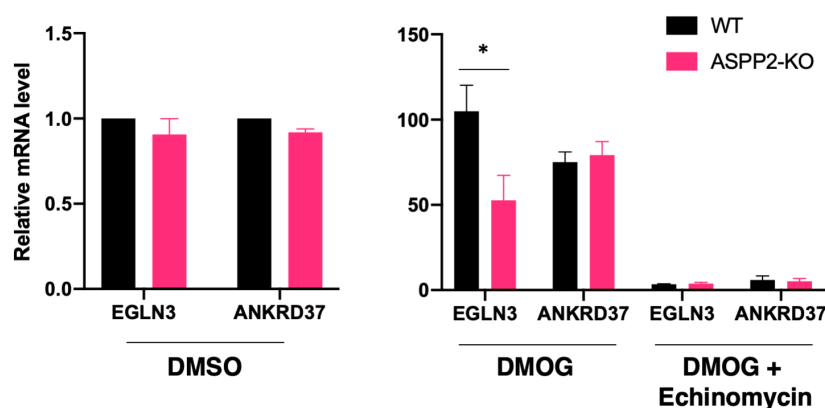


Figure 4.17: Echinomycin abrogated the modulatory effect of ASPP2 on *EGLN3* expression in hypoxia-mimicking condition. qPCR analysis of *EGLN3* and *ANKRD37* gene expression in WT and ASPP2-KO U2OS cells treated with DMSO, DMOG, or DMOG plus echinomycin for 16 hours. The expression level of each gene was normalized to that in the DMSO-treated WT U2OS cells. Values were mean $\pm$ standard deviation; Biological duplicate and technical triplicate for each group; * $p < 0.05$.

However, it should be acknowledged that the echinomycin used in this experiment was a kind gift from Dr. Pakavarin Louphrasitthiphol (Colin group, Ludwig, Oxford) and although it had been used by researchers in the Colin group, whether it can effectively and specifically inhibit HRE binding had not been tested in our lab. Therefore it would be critical that the effectiveness and specificity of this drug be tested by, e.g. measuring the luminescence of HRE luciferase reporter and non-HRE (e.g. E-box) luciferase reporter under normoxia and hypoxia with and without echinomycin.

4.2.11 ASPP2 and HIF1 α do not seem to directly interact with each other under hypoxia

4.2.11.1 ASPP2 and HIF1 α have different subcellular localization under hypoxia

So far, our results have suggested that ASPP2 regulates PHD3 expression independently of HIF1 α 's protein expression, subcellular localization, binding at

the selected HRE-containing loci as well as recruitment of p300 and potentially p53 to those HIF-binding loci of PHD3. We therefore next investigated that whether these two proteins co-localized or physically interact with each other under hypoxia: if no direct interaction between them was detected, this would suggest that ASPP2 regulates PHD3 expression independently of HIF. To test whether ASPP2 and HIF1 α co-localized, the subcellular localization of these two proteins under physiological hypoxia were analyzed by immunofluorescence at four different time-points. Throughout the hypoxia time-course experiment, ASPP2 was localized predominantly at the cell junction. Additionally, some ASPP2 was also detected in the cytoplasm after six hours of hypoxia [Figure 4.18]. HIF1 α displayed a primarily nuclear localization throughout the hypoxia time-course [Figure 4.18].

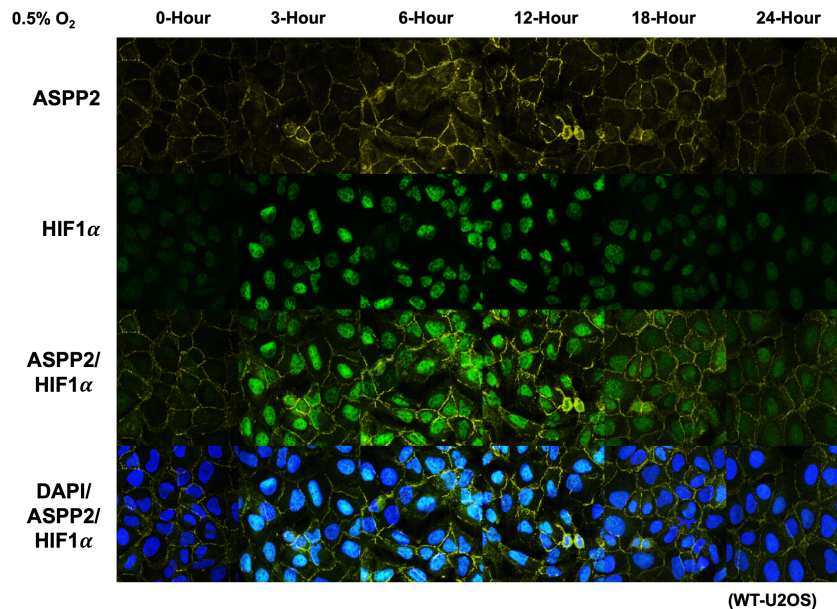


Figure 4.18: ASPP2 and HIF1 α did not co-localize under hypoxia. Representative immunofluorescence co-staining of ASPP2 (yellow) and HIF1 α (green), with nuclei counterstained with DAPI (blue). WT U2OS cells were incubated at 0.5% O₂ and fixed at the designated timepoints. Images are representative examples of two independent replicates of this experiment with at least three images taken per sample.

4.2.11.2 ASPP2 and HIF1 α cannot be co-immunoprecipitated under hypoxia

Although the results of this immunofluorescence analysis suggest that ASPP2 and HIF1 α did not co-localize with each other due to their different predominant subcellular localization, it is known that HIF1 α can be shuttled between the nucleus and the cytoplasm (Chachami et al. 2009; Mylonis et al. 2008). Given that ASPP2 could be detected in the cytoplasm after six hours of hypoxia, it is possible that it could interact with HIF1 α , potentially in the cytoplasm. To test this idea, co-immunoprecipitation (co-IP) was performed. Briefly, the Tet-On-ASPP2 U2OS cells were treated with dox to induce ASPP2, which was fused with a 3X FLAG tag at its C-terminal. These cells were then treated with either normoxia or physiological hypoxia and the whole cell lysates were subsequently subject to IP against FLAG-tag or the control IgG. The co-IPed proteins were analyzed by western blot.

More FLAG-tagged ASPP2 was detected in the FLAG-IP fraction than in the input and no FLAG-tagged ASPP2 was detected in the IgG-IP fraction under both normoxia and hypoxia, suggesting that FLAG-tagged ASPP2 was enriched in the IP fraction and this IP was specific [Figure 4.19]. PP1 α , a known interactor of ASPP2 (Bertran et al. 2019; Egloff et al. 1997; Nicholas R. Helps et al. 1995), was used as a positive control for this co-IP and it was co-immunoprecipitated by FLAG-ASPP2 but not the IgG. As expected, HIF1 α was detected in the input under hypoxia but not under normoxia. However, it was not detected in the co-IP fraction. HIF1 β (ARNT), on the other hand, was detected in both normoxia and hypoxia, but it was also not co-immunoprecipitated by FLAG-ASPP2 [Figure 4.19]. Therefore these

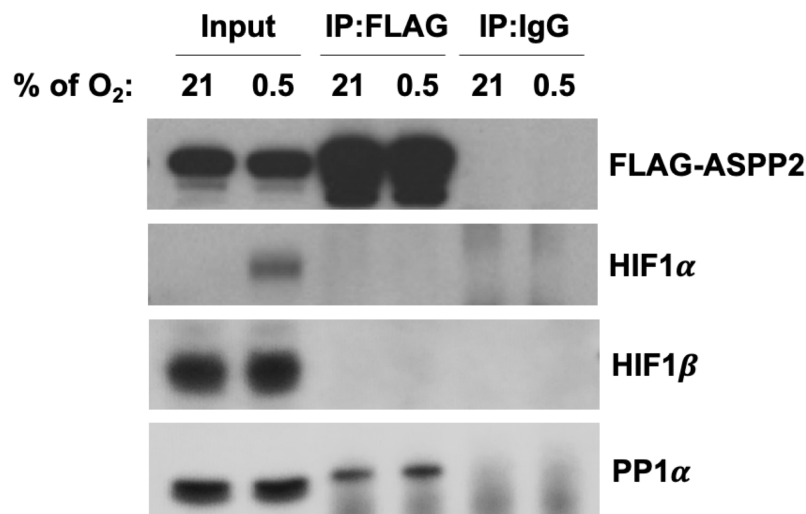


Figure 4.19: Physical interaction was not detected between ASPP2 and HIF1 α or HIF1 β under normoxia (21% O₂) and hypoxia (0.5% O₂). Co-IP experiment was performed using ASPP2-KO cells stably expressing doxycycline inducible FLAG-tagged ASPP2. Following doxycycline induction, FLAG-tagged ASPP2 was immunoprecipitated from normoxia- or hypoxia-treated cells using an anti-FLAG antibody and co-immunoprecipitates were analyzed by western blot using the indicated antibodies. This result was representative of two independent experiments.

results suggested that ASPP2 does not interact with HIF1 α or HIF1 β under both normoxia and hypoxia, at least at the time-point tested.

4.2.12 The regulation of PHD3 expression by ASPP2 is not sensitive to HIF1 α knockdown in hypoxia

To further confirm if ASPP2 regulates PHD3 expression independently of HIF1 α , WT and ASPP2-KO U2OS cells were treated with anti-*HIF1A* siRNA to knockdown *HIF1A* expression before subjecting them to either normoxia or hypoxia incubation. The mRNA expression of *EGLN3* was then analyzed by qPCR.

Compared to scrambled control, treatment of anti-*HIF1A* siRNA reduced *HIF1A* expression to a similar extent in WT and ASPP2-KO cells, by over 5-fold in normoxia and close to 10-fold in hypoxia [Figure 4.20]. As expected, *EGLN3* expression

was significantly induced in both scrambled control-treated WT and ASPP2-KO cells under hypoxia [Figure 4.20]. Treatment of anti-*HIF1A* siRNA resulted in a decrease in hypoxic *EGLN3* expression of both WT and ASPP2-KO cells (~2.7-fold decrease in WT and ~3.5-fold decrease in ASPP2-KO) [Figure 4.20], likely due to the knockdown of *HIF1A* expression in hypoxia.

On the other hand, the loss of ASPP2 did not have a noticeable effect on *EGLN3* expression regardless of the siRNA treatment under normoxia, whereas in hypoxia, it resulted in a comparable fold decrease in *EGLN3* expression regardless of the siRNA treatment (near 2-fold decrease when treated with scrambled control, ~2.5-fold when treated with anti-HIF1 α siRNA) [Figure 4.20].

Taken together, these data suggested that the regulation of PHD3 expression by ASPP2 is hypoxia-specific. However, it could not be concluded whether this regulation would require at least a certain level of HIF1 α to be expressed or not, as HIF1 α expression was not completely abolished after the knockdown, as indicated by the reduced induction of hypoxic PHD3 expression.

4.3 Summary and Brief Discussion

The work presented in this chapter revealed that ASPP2 modulates the gene expression of a subset of HIF targets irrespective of FIH status. Using *EGLN3* as the candidate ASPP2-sensitive HIF target to understand the underlying mechanism, we found that neither the protein expression and subcellular localization of HIFs, nor the DNA-binding of HIFs at the HRE-containing *EGLN3* loci was regulated by ASPP2. However, we found that the modulatory effect of ASPP2 on *EGLN3* expression may require HRE-binding. Taken together, these observation prompted us to make a

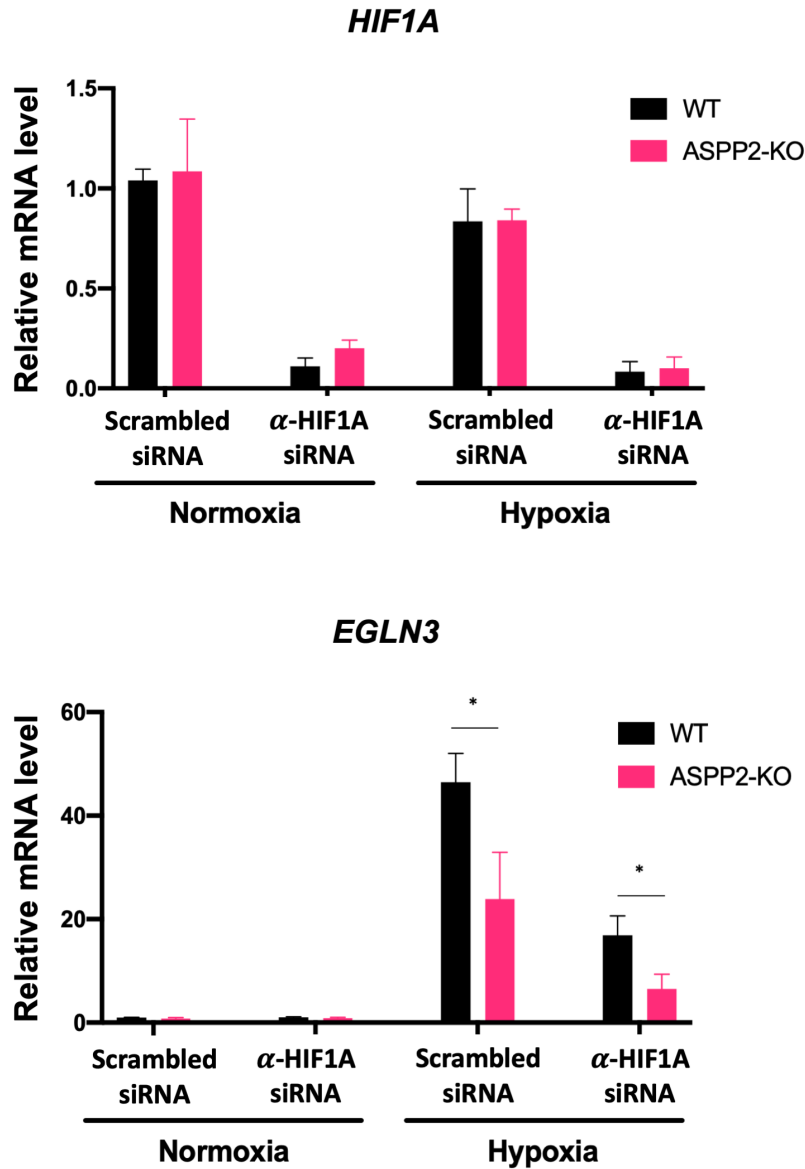


Figure 4.20: HIF1 α knockdown did not affect the regulation of PHD3 expression by ASPP2 under hypoxia. qPCR analyses of gene expression of *HIF1A* (top) and *EGLN3* (bottom) in WT and ASPP2-KO U2OS cells treated with either non-targeting siRNA (scrambled) or anti-*HIF1A* siRNA under normoxia or hypoxia (0.5% O₂ for 16 hours). Values were mean $\pm$ standard deviation; biological duplicate and technical duplicate for each group; *p < 0.05.

new hypothesis that ASPP2 may regulate HIF target gene expression by regulating another signaling pathway that crosstalks with HIFs at the transcriptional level.

This chapter set out to address the hypothesis that ASPP2 could modulate the transcriptional activity of HIF by competing with HIF α for FIH hydroxylation.

Starting with a small-scale wet lab experiment, a HIF target gene panel consisting of six well-established HIF targets, together with the ASPP2-KO U2OS cell lines generated in the previous chapter were used to get a preliminary understanding of the potential effects of ASPP2 on HIF transcriptional activity.

We observed that knocking out ASPP2 affected the expression of certain HIF target genes in the panel, including *EGLN3* and *VEGFA*, under both physiological hypoxia and hypoxia-mimicking condition. Such modulatory effects of ASPP2 were selective in the sense that the loss of ASPP2 alone did not have a significant effect on the other three HIF target genes in the panel.

This was consistent with our observation in the global transcriptomic analyses through which ASPP2's effects on the expression of HIF target genes were assessed. It was found that ASPP2 can alter the gene expression of a subset of HIF target genes under hypoxia irrespective of FIH status, although FIH status affects how ASPP2 modulates a different subset of HIF target gene expression. However, further investigation to elucidate the inter-relationship between FIH, ASPP2, and HIF activity, was hindered by the technical limitation posed by the ASPP2 double hydroxylation mutant and this direction was therefore not further pursued in this study.

Using PHD3 as our candidate for further validation, it was observed that the loss of ASPP2 resulted in the decrease of PHD3 expression at the protein level, and this effect was observable in cells of different origins, including the bone (U2OS),

the stomach (AGS) and the central nervous system (primary astrocytes). Such observation was important as they suggested that the observed effect of ASPP2 was not likely to be an artifact due to some idiosyncrasies of a particular cell line or knockout clone. This observation also implied that there might be potential generalized mechanism(s) that connected ASPP2 and HIF.

To further confirm that our observation could be explained by the loss of ASPP2 but not because of some random artifacts, the question was asked whether reintroduction of ASPP2 in ASPP2-KO cells could rescue the phenotype. For this, using a combination of piggyBac transposition and Tet-on systems, stable ASPP2-KO U2OS cell lines were generated to express inducible FLAG-tagged ASPP2 upon doxycycline treatment. After checking that the inducible ASPP2 could be detected by ASPP2-specific antibodies and shared similar molecular weight and subcellular localization with WT endogenous ASPP2, it was found that reintroduction of ASPP2 could rescue PHD3 expression in ASPP2-KO U2OS cells under hypoxia. These results thus strongly suggested that ASPP2 regulated PHD3 expression.

This chapter then moved on to investigate the potential mechanism(s) that could explain ASPP2-mediated regulation of *EGLN3* expression. Since *EGLN3* is a known HIF target and its expression was strongly induced by hypoxia and upon HIF α stabilization, it was hypothesized that ASPP2 regulates *EGLN3* expression directly through the HIF pathway. To address this hypothesis, several major points of regulations of the HIF pathway were considered, including protein expression, subcellular localization and HRE-binding activity.

Protein expression was assessed by performing a time-course study of protein expression that included HIF1 α , HIF2 α , ARNT and coactivator p300, it was

found that the loss of ASPP2 does not have a noticeable influence on any of their expression under hypoxia, suggesting that protein expression may not be a relevant point of regulation. The data also suggests that HIF1 α is the major HIF α isoform under hypoxia in these U2OS cell lines, at least under our assay conditions, based on the observation that HIF2 α was very weakly expressed throughout the 24-hour hypoxia time-course.

Subcellular localization was then investigated by a time-course immunofluorescence analysis of HIF1 α under hypoxia in WT and ASPP2-KO U2OS cells. It was found that HIF1 α had a mostly nuclear localization under hypoxia but it was not influenced by the loss of ASPP2.

ChIP-qPCR was then optimized and used to investigate the level of HIF binding to the HRE-containing loci of *EGLN3*. Two HRE-containing HIF-binding sites within/near the *EGLN3* gene locus were selected for ChIP-qPCR analysis based on multiple in-house ChIP-seq datasets from the Ratcliffe group. While HIF1 β and HIF1 α were found enriched at both mentioned HIF-binding loci by comparing with their enrichment at the negative control locus of *ACTG1* and also with the enrichment of IgG, their enrichment were not affected by the loss of ASPP2 under hypoxia. HIF2 α , on the other hand, was not enriched at the two HIF-binding loci of *EGLN3*.

It was also investigated by ChIP-qPCR whether ASPP2 influenced the recruitment of p300 and p53 to the two HIF-binding loci of *EGLN3*, given the reported interactions of these two proteins with ASPP2 and HIF1 α , respectively. However, the loss of ASPP2 did not seem to affect the enrichment of p300 at the two selected HIF-binding loci of *EGLN3*, while p53 was not enriched at those two loci.

These negative results, however, did not preclude the possibility that ASPP2 might influence HIF binding, or the recruitment of p300 or p53 to other HRE-containing loci in the genome that may interact with the promoter or enhancer(s) of *EGLN3*. A genome-wide analysis of HIF-binding sites in WT and ASPP2-KO cells under hypoxia, potentially together with Capture-C, would be required for a more thorough investigation.

As an alternative, echinomycin was used to assess whether the HRE-binding ability of HIF was required for the regulation of *EGLN3* expression by ASPP2. Although the data suggested that this regulation was inhibited by echinomycin under hypoxia-mimicking condition, attention should be paid to the potential confounding effects of echinomycin. Since the HRE sequence was a variant of the E-box sequence that was recognized by bHLH-PAS transcription factors, echinomycin was also reported to interfere with c-Myc/MAX transcriptional activity (Vlaminck et al. 2007), although its IC₅₀ was in the nM range, which was much higher than the dose (20 pM) that was used in this study. Additionally, echinomycin was also reported to interfere with topoisomerase II (Yoshinari et al. 1994) and DNA methylation (Adams et al. 1987), although the IC₅₀ was again much higher than the dose used in this study (IC₅₀ in the uM range for topoisomerase II, nM range for DNA methylation). With the potential confounding effects in mind, the data did bring forward a possibility that merited further investigation. This possibility was that ASPP2 might not affect the binding of HIF to the HRE-containing loci of *EGLN3* per se, but HIF-binding might be required for the regulation of PHD3 expression by ASPP2.

In addition to assessing several major points of regulation of the HIF pathway, it was also examined whether ASPP2 could directly interact with HIF. To address this question, the respective subcellular localization of ASPP2 and HIF1 α was analyzed by immunofluorescence under hypoxia to see if they co-localized with each other. There was no strong evidence supporting their co-localization as ASPP2 predominantly localized in the cell-cell contact while HIF1 α was mainly nuclear under hypoxia. However, the observation of some cytoplasmic ASPP2 at a certain time-point under hypoxia raised the question whether there was a physical interaction between the two proteins because HIF1 α was known to shuttle between the nucleus and the cytoplasm. However, neither HIF1 α nor ARNT was detected in the co-IP of FLAG-tagged ASPP2 under hypoxia, suggesting that ASPP2 did not physically interact with HIF, at least under the selected conditions.

It was also investigated whether the regulation of *EGLN3* expression by ASPP2 was sensitive to the level of HIF1 α under hypoxia. For this, siRNA was used to knockdown *HIF1A* expression in WT and ASPP2-KO cells under both normoxia and hypoxia. Although the expression of *EGLN3* in hypoxia was compromised by *HIF1A* knockdown, the relative difference of *EGLN3* expression between WT and ASPP2-KO cells was not significantly affected, suggesting that this regulation was not sensitive to a reduced level of *HIF1A* expression. The conclusion that could be made about this data was that this regulation seems to be hypoxia-specific, as the loss of ASPP2 does not have a noticeable effect on *EGLN3* expression in normoxia. However, whether this regulation is dependent on HIF1 α remains inconclusive, because HIF1 α knockdown did not completely eliminate HIF1 α expression under hypoxia and HIF was not the only pathway that was regulated by hypoxia. A more

conclusive picture would require knocking out HIF1 α in both WT and ASPP2-KO cell lines in future investigations.

So far, by the process of elimination, our data suggests that the regulation of *EGLN3* expression by ASPP2 was not due to any effects of ASPP2 on (i) the protein expression of HIF α and ARNT, (ii) the subcellular localization of HIF1 α , (iii) HIF-binding or the recruitment of p300 and p53 to the selected HIF-binding loci of *EGLN3*, or (iv) physical interactions between ASPP2 and HIF. Based on these observation, it was not likely that ASPP2 regulated *EGLN3* expression by directly acting on HIF. Interestingly, the echinomycin data suggested that the regulation of *EGLN3* expression by ASPP2 was abolished when the HRE-binding activity of HIF was inhibited. This raised the question of whether ASPP2 might instead regulate certain transcriptional factor(s) or co-factor(s) from another signalling pathway that interfaces with HIF to indirectly affect the transcriptional activity of HIF.

To examine this possibility, we decided to exploit the abundance of publicly accessible transcriptomic datasets from cancer patients, with the aim of identifying candidate signalling pathways or components that might link ASPP2 to the transcriptional activity of HIF and getting a bird's eye view of any potential association of ASPP2 with HIF signalling activity in a clinically relevant context.

5

Pan-cancer *TP53BP2* co-expression analysis

5.1 Introduction

In order to identify potential signalling pathways that might connect ASPP2 to the HIF pathway, we took a step back to first understand the signalling pathways in which ASPP2 functions.

Multiple ways can be used to characterize the functions of a gene. One common wet-lab-based approach is to perturb the gene of interest, for example by knocking it out at the DNA level or reducing its expression at the mRNA level, and investigate the resulting phenotypic changes that happen to the biological system of choice, such as a cancer cell line or a mouse model. Indeed some biological functions of ASPP2 have been identified in this way. For example, the significant CNS defects and the disorganized neuroepithelium displayed by ASPP2 Δ^3/Δ^3 mice led to the hypothesis and subsequent findings of ASPP2's involvement in cell polarity regulation (W. Cong et al. 2010; Sottocornola et al. 2010). The phenotypes of

artificial knockout models, however, may not always fully reflect the functions of the target gene or the roles played by the target gene in complex pathological conditions, such as cancer. Another widely-used wet-lab-based approach is to study the interactome of the protein of interest. This is based on the principle of guilt-by-association: functions of the target protein can be gauged by looking at its interacting partners with known functions. It is supported by multiple experimental datasets demonstrating that proteins interacting with each other are often involved in the same, or related biological functions (Fromont-Racine et al. 2000; Ito et al. 2001; Vazquez et al. 2003). In fact, one of the major pieces of evidence that led to the findings of ASPP2's role in p53-mediated apoptosis was from a yeast two-hybrid screen, in which the shorter isoform of ASPP2, 53BP2, was identified as a p53-binding protein (Iwabuchi et al. 1994).

The rapid accumulation of publicly accessible gene expression datasets as a result of high-throughput sequencing technologies has greatly facilitated the use of another complementary, guilt-by-association-based approach, i.e. gene coexpression analysis, to study gene functions. Similar to the concept of the protein-protein-interaction approach, gene coexpression analysis predicts the functions of the gene of interest from functions of genes whose expression correlate with that of the gene of interest. The underlying logic is based on the common observation reported by multiple studies in eukaryotes that functionally-related genes tend to share coordinated gene expression patterns, although such coexpression does not imply causality (Eisen et al. 1998; Hughes et al. 2000; Miki et al. 2001; Niehrs et al. 1999; L. F. Wu et al. 2002). For example, if we observe an uncharacterized gene correlated with

other genes that function in translation initiation, we might infer that the product of the uncharacterized gene is involved in translation initiation as well.

A major advantage of using gene coexpression analysis is that it represents a cost-effective way to predict gene functions in a clinically relevant context as large-scale genome-wide expression data from patients suffering from different types of cancers, such as those catalogued by The Cancer Genome Atlas (TCGA) (www.cancer.gov/tcga), are readily accessible.

To date, gene coexpression analysis has not been used to predict ASPP2's functions. To identify potential biological processes and signaling pathways in which ASPP2 functions in human cancers, we developed a bioinformatic pipeline to perform pan-cancer gene coexpression analysis of *TP53BP2* encoding ASPP2 using the pan-cancer TCGA RNA-sequencing (RNA-seq) dataset, which is profiled from over 10,000 cancer patients and includes more than 30 different types of cancers. The reasons for choosing this dataset are multifold. Firstly, this dataset has a much larger sample size than any other publicly accessible transcriptomic dataset. This large sample size is beneficial to our analysis as it allows us to have a higher level of confidence in the coexpression estimates and improves the statistical power to detect statistically significant associations that might otherwise be masked by the heterogeneity in our samples. Secondly, this dataset allows us to perform analyses over 30 different types of cancers, i.e. pan-cancer, instead of restricting our analyses to only a single cancer type. We consider that conservation across multiple cancer types would be a powerful criterion to identify coexpressed genes that are functionally related to ASPP2. A pan-cancer coexpression pattern of a certain gene pair may imply that such a coexpression relationship confers some sort of selective

advantage during the evolution of cancer cells, and it is more likely to reflect a functional relationship than a coexpression obtained from analyzing datasets limited to a single cell line or cell type. Thirdly, analyzing the TCGA dataset enables further integration of our results with other clinical, genomics, epigenomics, and proteomics data catalogued by TCGA as they are profiled from the same patients.

At the time when the above idea was conceived, several user-friendly platforms, such as Oncomine (Rhodes et al. 2004) and cBioPortal for Cancer Genomics (Cerami et al. 2012), already existed to allow researchers to search for genes coexpressed with their gene of interest in individual TCGA cancer datasets. However, their analysis did not take into consideration potential confounding variables that might be present in the dataset and obscure the relationships between our gene of interest and its coexpressed genes. Another platform with a confounder-adjustment feature called Genomic Regression Analysis of Coordinated Expression (GRACE) was also developed during the development our bioinformatic pipeline (L. Cai et al. 2017) but its adjustment is limited to potential confounding effects from somatic copy number only and coexpression analysis is limited to only 21 TCGA cancer types. Moreover, these platforms did not provide any downstream functional interpretation of any resulting coexpressed genes. We therefore developed a bioinformatic pipeline, written in R programming language, that integrated both the identification and functional interpretation of *TP53BP2* coexpressed genes with the capability of identifying potential confounders in the dataset and adjusting for them if necessary.

5.2 Results

5.2.1 Overview, assembly, and preprocessing of input data

Our coexpression analysis pipeline consists of four major steps [Figure 5.1]: (i) the assembly and preprocessing of the input data; (ii) the identification of potential confounders; (iii) the coexpression analysis; and (iv) the functional interpretation of coexpressed genes.

Our input data includes the TCGA pan-cancer RNA-seq dataset, curated clinical data, pan-cancer gene-level somatic copy number (SCN) estimates, and pan-cancer binary gene-level non-silent mutation calls. All four of them are part of the Pan-Cancer Atlas initiative driven by TCGA Research Network (Weinstein et al. 2013) and were used in a series of Pan-Cancer Atlas papers published in Cell journals (<https://www.cell.com/pb-assets/consortium/pancanceratlas/pancani3/index.html>). These input data are publicly accessible from the UCSC Xena platform (<https://xenabrowser.net/>) (Goldman et al. 2019).

The RNA-seq dataset comprised 11070 patient tumor samples from 32 cancer types, covering solid, hematologic and lymphatic malignancies originating from different organ systems [Figure 5.2]. Transcript quantification, normalization, and batch correction of this dataset were performed by TCGA (For more details, please refer to subsection 2.7.5).

The clinical data, which contains anonymized patients information (e.g. age and gender) and technical variables (e.g. sequencing platform, sequencing plates), SCN estimates, and non-silent mutation calls from the same patients were assembled

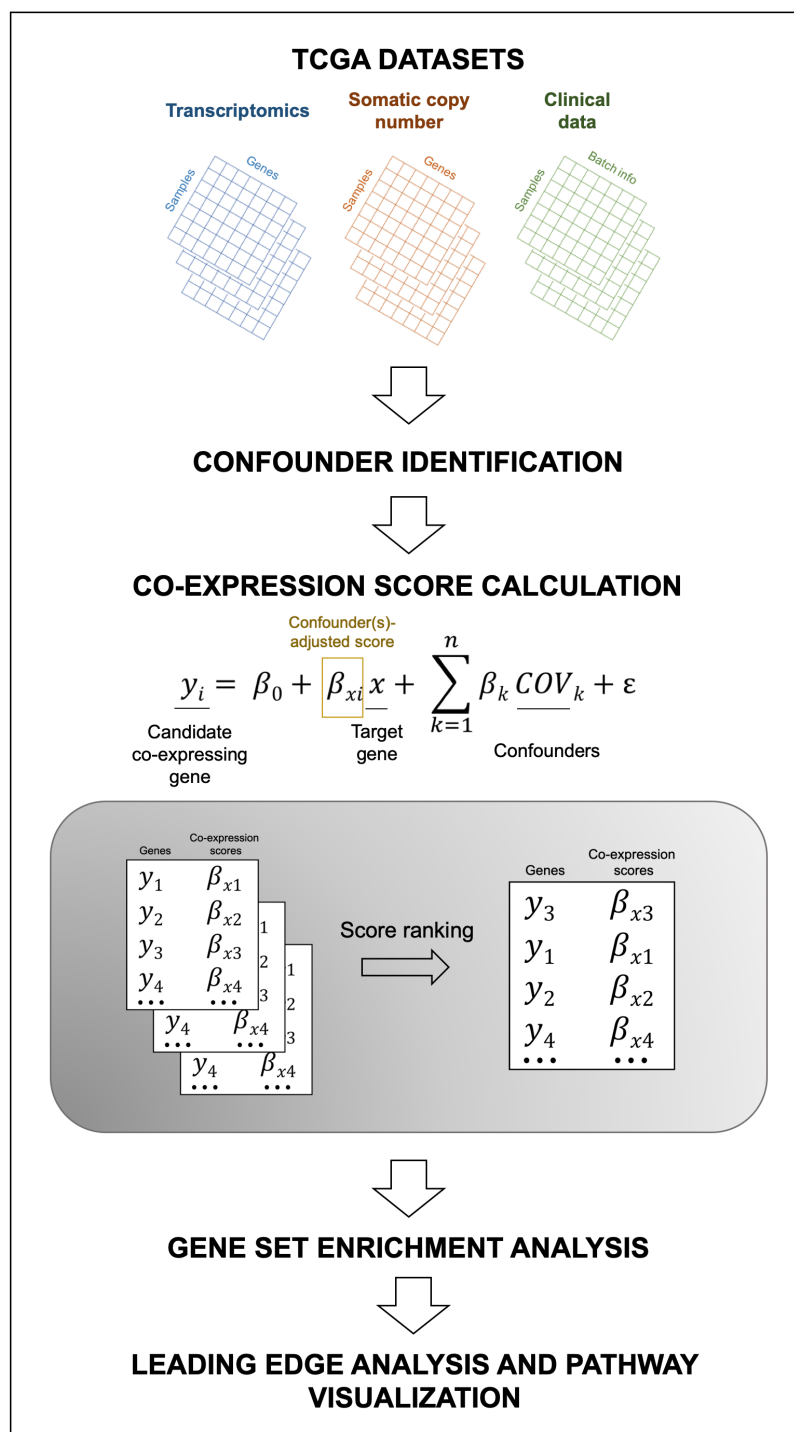


Figure 5.1: Overview of the gene coexpression analysis pipeline. This flowchart shows a brief overview of the major steps of our gene coexpression analysis pipeline, which includes a beta-coefficient-based method to calculate confounder(s)-adjusted gene coexpression scores.

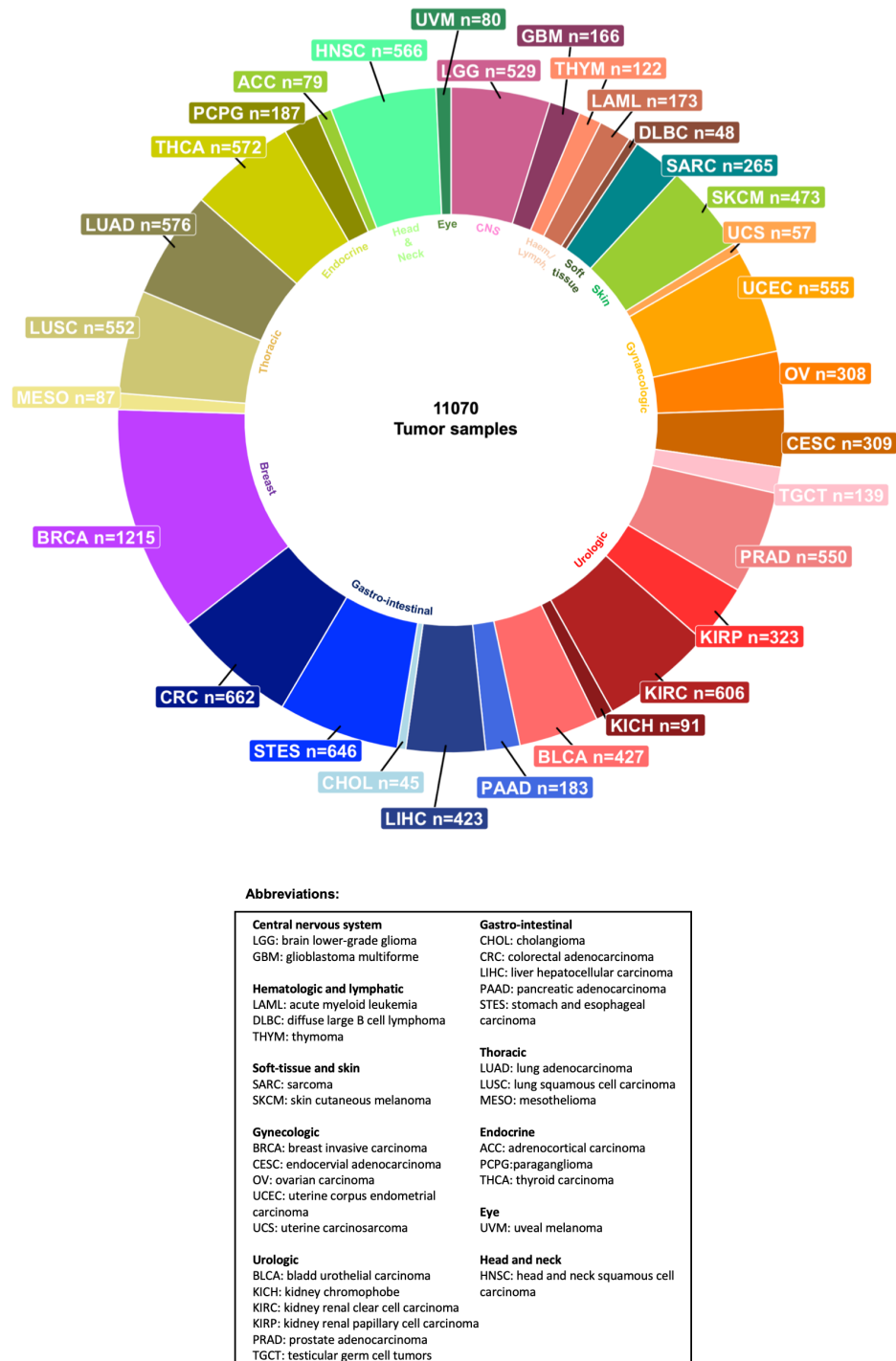


Figure 5.2: Overview of the TCGA pan-cancer transcriptomic dataset. This doughnut chart illustrates the distribution of the cancer types in the TCGA pan-cancer cohort.

together as the metadata to be used as an input to check for potential confounders, which will be elaborated in the next section.

The input data were then preprocessed to handle the missing data, filter lowly-expressed genes, and check if the RNA-seq data is normalized. Details of these preprocessing steps are included in subsection 2.7.5.

5.2.2 Confounder identification

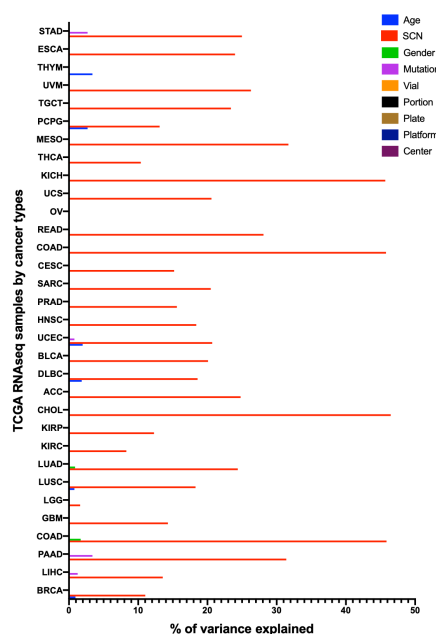
The next step is the identification of variables that may confound the association between *TP53BP2* and its coexpressed genes, so that their confounding effects could be adjusted for during the coexpression analysis. To be a confounder, a variable needs to satisfy three criteria. First, this variable needs to be associated with the independent variable, i.e. *TP53BP2* gene expression in our case. Second, this variable needs to be associated with the dependent variable, i.e. the expression of every other gene not named *TP53BP2* in the dataset. Third, this variable must not be an intermediary step in the causal pathway from the independent variable to the dependent variable.

To identify potential confounders, a simple linear regression [Equation 5.1] was used to measure the association between *TP53BP2* expression, and each of the following variables in the metadata: demographic variables (Age, Gender), technical variables (Sequencing centers, platforms, plates, portions, vials), and biological variables (*TP53BP2* mutation calls and *TP53BP2* SCN estimates), all of which are highly unlikely to represent a direct causal pathway from *TP53BP2* to its functionally-related coexpressed genes.

$$y = \beta_0 + \beta_x x + \epsilon \quad (5.1)$$

In this equation, y represents normalized *TP53BP2* expression, x represents the candidate confounder, β_0 is the predicted value of y when $x = 0$, and ϵ is the error term. The adjusted coefficient of determination (R^2) from this equation, which indicates how much variance of y that can be explained by x , is the square of the correlation coefficient. The p-value for x , on the other hand, indicates whether the association between x and y is statistically significant. We therefore used these two parameters: the adjusted R^2 to measure the association between normalized *TP53BP2* expression and each candidate confounder, which would be considered as a confounder if the two following conditions were met: (i) the corresponding adjusted R^2 is greater than 0.1, which indicates a mild association, and (ii) the p-value for x is smaller than 0.05, which indicates a statistically significant association.

(A)



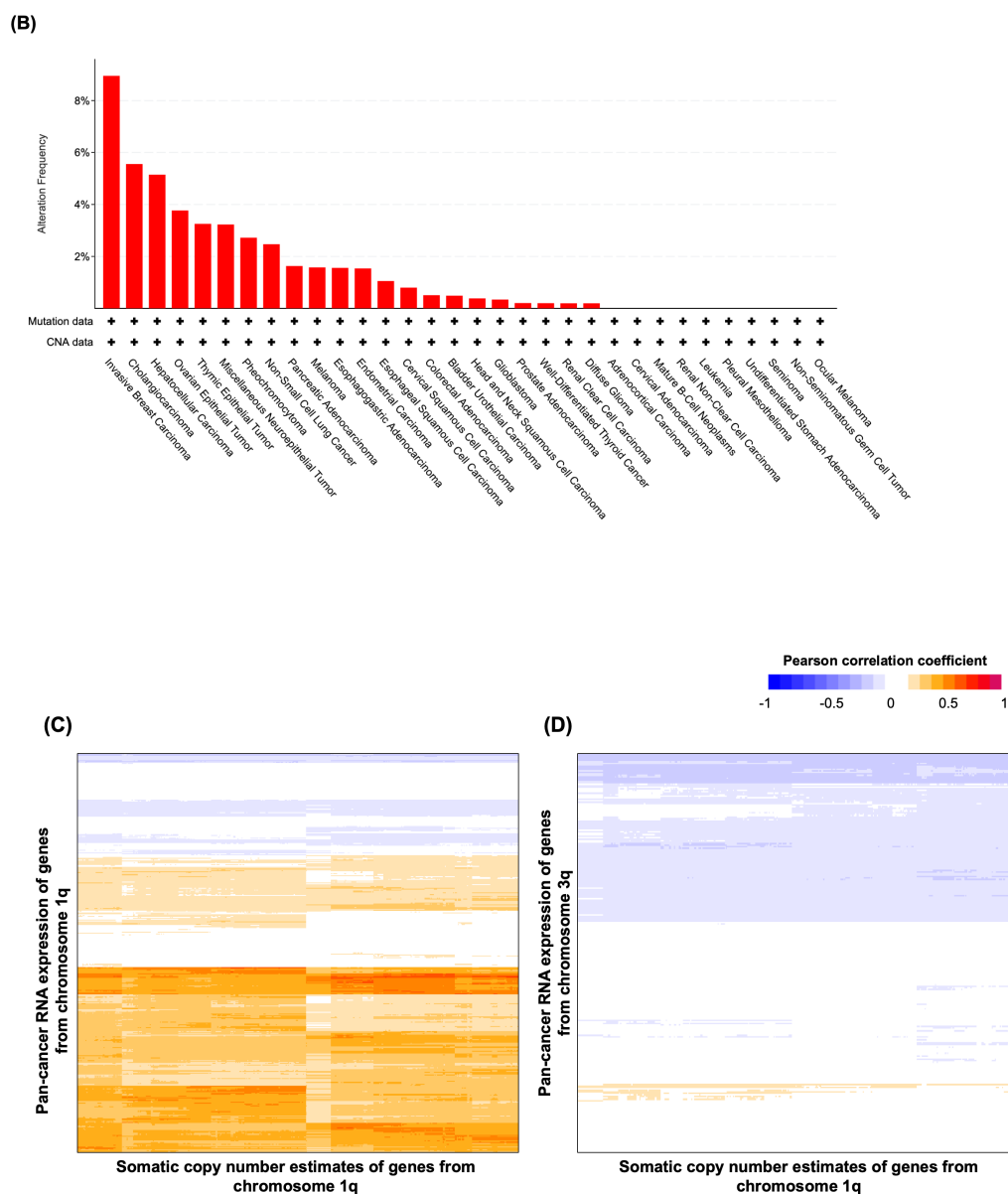


Figure 5.3: Somatic copy number as a potential confounder in gene coexpression analysis. (A) Bar plot showing the percentage of the variance of *TP53BP2* expression explained by each of the nine variables in each TCGA cancer as a measurement for correlation. SCN: somatic copy number estimates of *TP53BP2*; Mutation: Mutation calls of *TP53BP2*; (B) Bar plot showing the frequency of *TP53BP2*'s gene copy number amplification events in each TCGA cancer, obtained using the OncoPrint function on cBioPortal (www.cbioportal.org/) (J. Gao et al. 2013). (C and D) Correlation matrices showing the correlation between (C) copy number and RNA expression of genes from chromosome 1q and (D) between copy number of genes from chromosome 1q and RNA expression of genes from chromosome 3q. The color key on the top left represents Pearson correlation coefficient. Blue indicates highly negative correlation, red indicates highly positive correlation and white indicates no correlation.

As shown in Figure 5.3A, *TP53BP2* SCN estimates, but not the other input variables, satisfied the two aforementioned conditions. Focal and arm-level SCN alterations (SCNA) have been reported to be a frequent event in cancers (Beroukhi et al. 2010; Zack et al. 2013). Indeed, using the Oncoprint function of cBioPortal (cbioportal.org) to visualize genetic alteration events (J. Gao et al. 2013), we found that the copy number amplification of *TP53BP2*, which is located at chromosome 1q41, was observed in a subset of cancer populations in a number of TCGA cancers, such as invasive breast carcinoma, cholangiocarcinoma, and hepatocellular carcinoma [Figure 5.3B]. We also observed a widespread, strong positive correlation between SCN estimates and mRNA expression of genes on chromosome 1q [Figure 5.3C], likely due to arm-level SCNA events, whereas much less significant correlation was found between SCN estimates of genes on chromosome 1q and mRNA expression of genes on chromosome 3q, which was picked at random to represent another locus [Figure 5.3D]. A major premise of gene coexpression analysis is that coexpression may reflect regulatory gene-gene relationships or coordinated regulation of genes involved in the same, or related, biological functions and signaling pathways. However, this premise is obscured by SCN as the occurrence of SCNA in cancers causes coexpression of physically neighboring genes that may not be functionally related. Therefore, the confounding effects of SCN will be adjusted for in the gene coexpression analysis, which is the next step of our pipeline.

5.2.3 Beta coefficient-based coexpression analysis

To assess the correlation between two genes while taking into account of the confounding effects of SCN, we used a multivariate linear regression model [Equation 5.2],

in which there were four variables: (i) expression of each gene, excluding our gene of interest, as the dependent variable, denoted by y_i ; (ii) expression of our gene of interest as the independent variable, denoted by x ; (iii, iv) SCN estimates for the dependent and independent variables as the confounders, denoted by w_{yi} and w_x .

$$y_i = \beta_0 + \beta_{xi}x + \beta_{w_{yi}}w_{yi} + \beta_{w_x}w_x + \epsilon \quad (5.2)$$

A coexpression score was calculated for each gene (y_i) to estimate the strength of its association with our gene of interest (x) using the β coefficient of x (β_{xi}), which is an effect size measure that estimates the impact of the corresponding variable, x , on y_i , holding all other covariates (w_{yi} and w_x) constant.

To make the β_x for each gene comparable to each other, all the variables are standardized to z-scores. Thus in our model, the degree of coexpression between y_i and x , and that between y_j and x , are considered equal if a step change in x produces the same number of standard deviations of y_i as that of y_j .

The p-value of each β_{xi} (p_{xi}) is used to indicate the confidence we would have in the reliability of β_{xi} . p_{xi} tests the null hypothesis that β_{xi} has no correlation with the dependent variable y_i : only those β_{xi} whose p_{xi} is less than our significance level, which is 0.05 false discovery rate (FDR) to adjust for multiple comparisons. Conversely, any p_{xi} greater than 0.05 indicates that there is insufficient evidence in our data to support a non-zero correlation, in which case β_{xi} would be shrunk to zero.

In this way, a coexpression score was obtained for each gene in each TCGA cancer type, and the mean of the coexpression scores for each gene in all TCGA cancer types was computed as the pan-cancer coexpression score for each gene. This pan-cancer coexpression score was then ranked in descending order and the

positions on this ranked gene list could be interpreted as follows: (i) genes closer to the top of the list have higher positive pan-cancer coexpression scores which indicate greater positive associations with our gene of interest; (ii) genes closer to the bottom of the list have higher negative pan-cancer coexpression scores which indicate greater negative associations with our gene of interest; and (iii) genes close to the middle of the list have close to zero or zero pan-cancer coexpression scores, which indicate weak or no associations with our gene of interest.

Collectively, our method would hereafter be referred to as beta coefficient-based coexpression analysis (BCCA).

5.2.4 Evaluation of Beta coefficient-based coexpression analysis

If our proposed method for gene coexpression analysis, BCCA, was effective, we would expect it to be able to identify genes functionally related to our gene of interest while adjusting for confounding effects from SCN. We therefore evaluated BCCA by asking two questions: (i) whether it can adjust for confounding effects from SCN, and (ii) whether it can identify coexpressed genes involved in the same function/pathway as our gene of interest. The results of coexpression analysis using BCCA were then compared with two pre-existing methods for gene coexpression analysis: (i) cBioPortal (J. Gao et al. 2013), which uses Spearman’s correlation without adjusting for any confounding effects (hereafter referred to as cBioPortal), and (ii) Genomic Regression Analysis of Coordinated Expression (GRACE) (L. Cai et al. 2017), which uses a combination of univariate linear regression and Spearman’s correlation to estimate gene coexpression adjusting specifically for SCN’s effects.

To address these two questions, we initially selected *EIF2D* and *PARP2*, which are relatively well-characterized functionally and were also used by Cai *et al.* in the validation of GRACE (L. Cai et al. 2017), as our genes of interest for evaluation. However, since *EIF2D* was excluded from our dataset due to the presence of missing data (for the handling of missing data, please refer to subsection 2.7.5), *EIF2A* was used as an alternative gene of interest for evaluation as *EIF2A* and *EIF2D* perform closely related functions in translation (Kearse et al. 2017).

The first gene of interest for evaluation, *EIF2A*, is located at chromosome 3q25.1. *EIF2A* encodes Eukaryotic Translation Initiation Factor 2A (eIF2A), which functions in mRNA translation by directing the interaction between the initiator tRNA and 40S ribosomal subunit (Kearse et al. 2017). Gene coexpression analysis of *EIF2A* was performed using BCCA, GRACE, and cBioPortal in 10 TCGA cohorts. The resulting top 100 *EIF2A* coexpressed genes calculated by cBioPortal contained the largest number of genes from chromosome 3q in all 10 TCGA cohorts [Figure 5.4A], especially in LUAD, LUSC, BRCA, and KICH. Considering that the number of genes from chromosome 3q expected by chance, calculated by reshuffling the gene list 1000 times, is 2.88 ± 1.66 , the results from cBioPortal suggest a bias towards genes physically neighboring *EIF2A*, likely due to confounding by SCN. On the other hand, the application of GRACE significantly reduced this bias in all 10 TCGA cohorts tested [Figure 5.4A] and importantly, using BCCA, the number of top 100 *EIF2A* coexpressed genes from chromosome 3q was at least comparable to that estimated by GRACE in all 10 cohorts [Figure 5.4A].

We then assessed whether the top 100 *EIF2A* coexpressed genes obtained by these three different methods contain genes involved in mRNA translation,

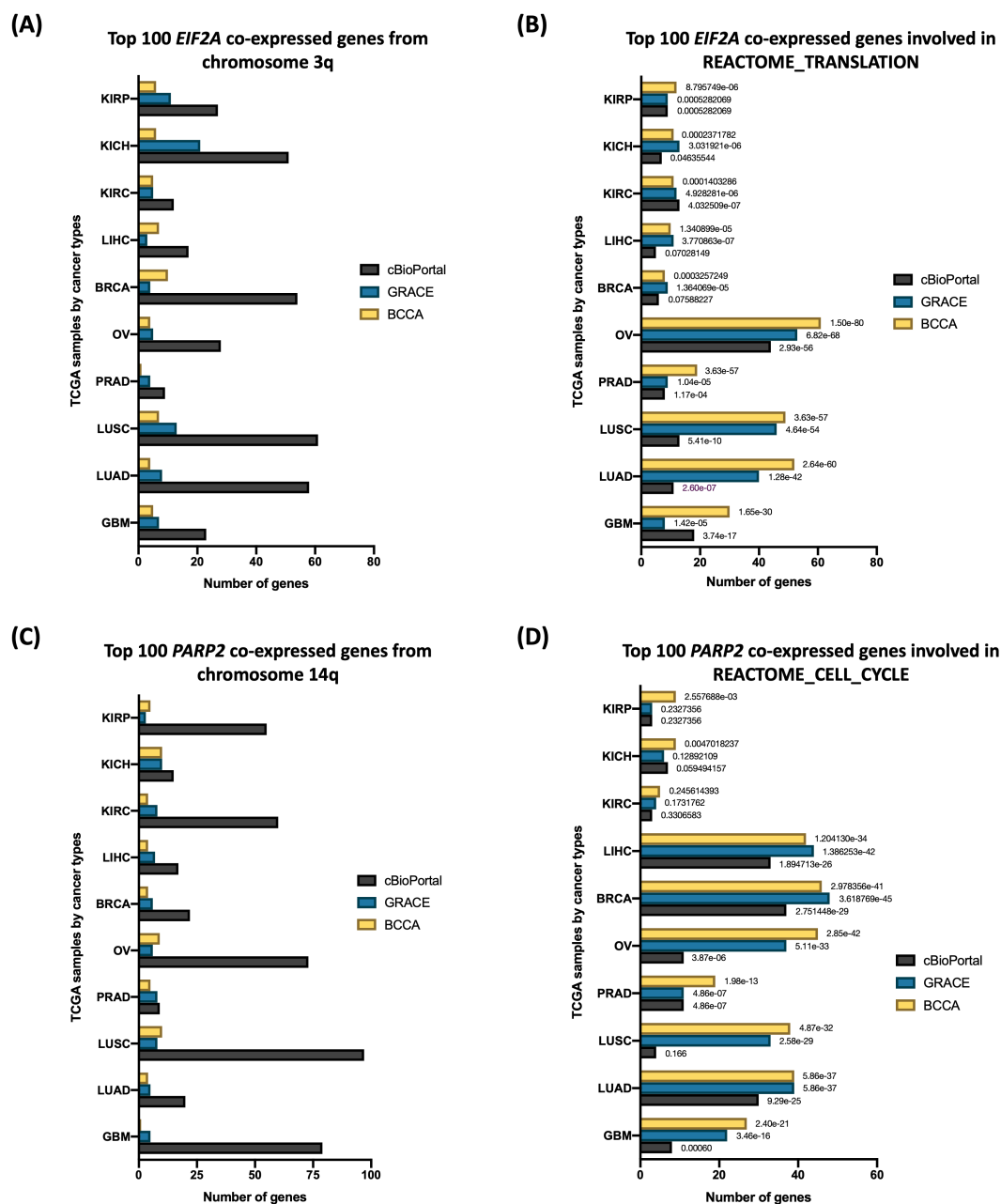


Figure 5.4: Beta Coefficient-based coexpression Analysis (BCCA) effectively adjusted for confounding by SCN and facilitated the identification of functionally-related coexpressed genes. Bar plots showing (A) the number of (A) top 100 *EIF2A* coexpressed genes from chromosome 3q, calculated respectively by cBioPortal, GRACE, and BCCA; (B) the number of top 100 *EIF2A* coexpressed genes involved in REACTOME Translation gene set. The number on the right of each bar represents the statistical significance (adjusted p-value) of the enrichment; (C) the number of top 100 *PARP2* coexpressed genes from chromosome 14q, calculated respectively by cBioPortal, GRACE, and BCCA, and (D) the number of top 100 *PARP2* coexpressed genes involved in REACTOME Cell Cycle gene set. The number on the right of each bar represents the p-value of the enrichment.

using the Translation gene set (R-HSA-72766) in the Reactome Pathway Database (<http://reactome.org>), which is a peer-reviewed database that catalogues genes organized into signaling pathways and biological processes (Croft et al. 2011). Consistent with the literature, the top 100 *EIF2A* coexpressed genes were enriched for mRNA translation process, and such enrichment was more significant for the coexpressed genes estimated by BCCA than those by GRACE or cBioPortal, in terms of both the number of Translation-enriched coexpressed genes identified and the corresponding statistical significance, in OV, PRAD, LUSC, LUAD, and GBM cohorts, while in KIRP, KICH, KIRC, LIHC, and BRCA cohorts all three methods showed comparable enrichment [Figure 5.4B].

The second gene for evaluation purposes is *PARP2* located at chromosome 14q11.2. This gene encodes Poly(ADP-Ribose) Polymerase 2 (PARP2), which is an important regulator of various mitotic functions in part through mediating base excision DNA repair and stabilization of replication fork (Ronson et al. 2018). Among the three coexpression methods, the top 100 *PARP2* coexpressed genes estimated by cBioPortal again displayed a bias towards physically neighboring genes in all 10 cohorts, considering the expected chance occurrence is 3.06 ± 1.61 [Figure 5.4C]. The top 100 *PARP2* coexpressed genes identified by both our method and GRACE contained significantly fewer number of genes located at chromosome 14q [Figure 5.4C], suggesting that both methods effectively minimized the confounding effects from *PARP2* copy number. On the other hand, the top 100 *PARP2* coexpressed genes identified by cBioPortal consistently had the lowest enrichment for Reactome Cell Cycle gene set (R-HSA-1640170) except in the KICH cohort [Figure 5.4D], whereas those identified by our method showed much higher enrichment for cell

cycle in all 10 TCGA cohorts and such enrichment were either comparable to or slightly higher than that estimated by GRACE [Figure 5.4D].

Taken together, the above observation suggested that our proposed gene coexpression method can effectively adjust for confounding effects from SCN and has the potential to identify coexpressed genes that are functionally related to the selected genes of interest. This gave us the confidence to use our method for *TP53BP2* to identify its potential functions at the pan-cancer level.

5.2.5 Functional interpretation of *TP53BP2* coexpressed genes

Using our proposed gene coexpression method BCCA and *TP53BP2* as our gene of interest, we produced (as described in subsection 5.2.3) a ranked gene list based on genes' pan-cancer coexpression scores. These scores indicate the strength of coexpression of each gene with *TP53BP2*, across the 31 TCGA cancer cohorts. Due to the length of this ranked gene list, only the top 500 and bottom 500 *TP53BP2* coexpressed genes were presented in Table S2 and Table S3, respectively, in the Appendix (Supporting information for chapter 5).

With this ranked gene list, we first checked whether our regression-based coexpression method effectively adjusted for the confounding effects from SCN by comparing the results with those obtained by cBioPortal without any confounder adjustment. All the top 20 *TP53BP2* coexpressed genes obtained by cBioPortal were located on chromosome 1q, neighboring *TP53BP2* gene locus [Table 5.1]. By contrast, the top 20 *TP53BP2* coexpressed genes obtained by our regression-based coexpression method BCCA were located across various chromosomes [Table 5.1]. This observation was consistent with our earlier evaluation results using *EIF2A*

and *PARP2* and it suggested that confounding effects of SCN were effectively adjusted for by our coexpression method.

Table 5.1: Genomic locations of the top 20 *TP53BP2* coexpressed genes

(a) cBioPortal			(b) BCCA		
Gene	Rank	Cytoband	Gene	Rank	Cytoband
WDR26	1	1q42.12	RNF169	1	11q13.4
C1orf55	2	1q42.12	RIF1	2	2q23.3
ACBD3	3	1q42.12	ROCK2	3	2p25.1
ASH1L	4	1q22	LATS1	4	6q25.1
BAT2L2	5	1q24.3	AFF4	5	5q31.1
ARID4B	6	1q24.3	PPP4R2	6	3p13
RAB3GAP2	7	1q41	STAG1	7	3q22.3
C1orf58	8	1q41	REST	8	4q12
TAF5L	9	1q41	JMJD1C	9	10q21.3
FBXO28	10	1q42.11	STRN	10	2p22.2
ZNF281	11	1q32.1	ASXL2	11	2p23.3
TPR	12	1q31.1	SCYL2	12	12q23.1
DHX9	13	1q25.3	TAOK1	13	17q11.2
GON4L	14	1q22	ROCK1	14	18q11.1
HEATR1	15	1q43	ASH1L	15	1q22
C1orf107	16	1q32.2	SBNO1	16	12q24.31
RBBP5	17	1q32.2	PIK3CA	17	3q26.32
NUP133	18	1q42.13	RSF1	18	11q14.1
CEP170	19	1q43	HIPK3	19	11p13
AHCTF1	20	1q44	NCOA2	20	8q13.3

Our next goal was to examine whether *TP53BP2* expression was highly correlated with certain groups of functionally related genes as a basis to infer its involvement in potential functions based on the guilt-by-association principle as mentioned in section 5.1.

To achieve this goal, pre-ranked Gene Set Enrichment Analysis (GSEA) (Ser-gushichev 2016) was performed to functionally interpret the ranked gene list at the biological/signaling pathway level. In this case, pre-ranked GSEA determines whether the members of a given gene set, which represents a collection of genes

functioning in a specific pathway or biological process, are randomly distributed in the ranked gene list. Such random distribution would suggest that there is a weak or insignificant correlation between this given gene set and *TP53BP2* expression. By contrast, if the members of this gene set are distributed near the top or the bottom of the ranked gene list, i.e. enriched, it would implicate a positive (if top) or negative (if bottom) correlation between this gene set and *TP53BP2* expression. Such correlation would in turn suggest the potential involvement of ASPP2 in the specific pathway that this enriched gene set represents.

To perform pre-ranked GSEA in an unbiased manner, the entire gene set collection from the REACTOME Pathway Database was used as the gene set and the ranked gene list of *TP53BP2* coexpressed genes was used as the pre-ranked gene list. Among the 1532 gene sets in this REACTOME gene set collection (Molecular Signatures Database version 7.0 (Subramanian et al. 2005)) , using a threshold of FDR less than 0.01, 135 gene sets were significantly enriched, out of which 43 gene sets were positively enriched and 92 gene sets were negatively enriched [Table S4].

From these 135 enriched gene sets, a commonly shared pan-cancer enrichment pattern emerged as indicated by the heatmap generated based on the enrichment scores of the enriched gene sets coexpressed with *TP53BP2* [Figure 5.5]. Hierarchical clustering of the TCGA cancer types identified five major clusters, with the 10 cancer types in cluster I showing the highest similarity of enriched *TP53BP2*-coexpressed pathways.

Based on the respective description and members of the 135 enriched gene sets, they were categorized into 10 groups [Figure 5.5]. The positively coexpressed groups

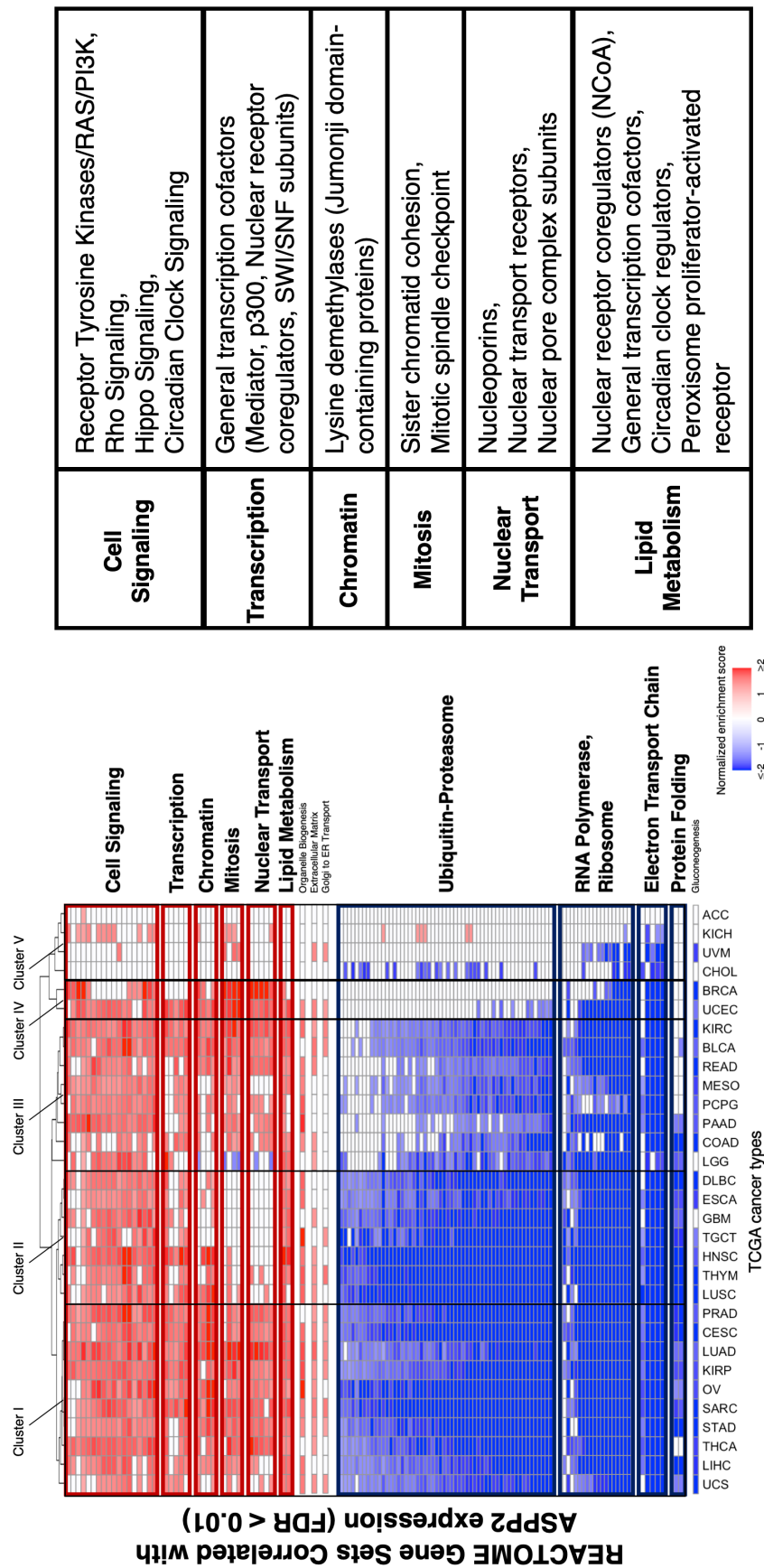


Figure 5.5: Pan-cancer enrichment patterns of REACTOME pathways coexpressed with *TP53BP2*. Heatmap (left) showing hierarchical clustering of TCGA cancer types based on the 135 coexpressed gene sets from REACTOME Database (Version 7.0) for *TP53BP2*. Each cell of the heatmap is colored by the normalized enrichment score of a gene set for a cancer type. White cells indicate a lack of enrichment. Red brackets on the heatmaps mark groups of functionally-similar gene sets that are positively coexpressed with *TP53BP2*, while blue brackets mark groups of functionally-similar gene sets that are negatively coexpressed with *TP53BP2*. Table (right) outlines the main gene sets found in each positively coexpressed group.

included cell signaling-related, transcription-related, chromatin-related, mitosis-related, nuclear-transport-related, lipid-metabolism-related pathways. The negatively coexpressed groups included ubiquitin-proteasome-related, RNA-polymerase-ribosome-related, electron-transport-chain-related, and protein-folding-related pathways. There were also four pathways that could not be categorized into any of the above groups and they were related to organelle biosynthesis, extracellular matrix, golgo-to-ER transport, and gluconeogenesis [Figure 5.5].

Each of the positively coexpressed groups was further divided into smaller subgroups, for example, the cell signaling group consists of *TP53BP2*-coexpressed pathways related to receptor tyrosine kinases/RAS/PI3K-related, rho-signaling-related, hippo-signaling-related, and circadian-clock-signaling-related gene sets [Figure 5.5]. The full details of the gene sets in these six groups can be found in Table S4.

Our gene coexpression analysis using BCCA identified two known signaling pathways that had previously been reported to be regulated by or associated with ASPP2. The first one was RAS GTPases Signaling [Figure 5.5], as ASPP2 had been shown to interact with activated RAS and stimulate downstream MAPK signaling (Yihua Wang et al. 2012; Z. Wang et al. 2013; Y. Wang et al. 2013). The second known association identified was Hippo Signaling [Figure 5.5], in which ASPP2 had been shown to be able to bring protein phosphatase 1 (PP1) to dephosphorylate YAP and TAZ (C.-Y. Liu et al. 2011; Royer et al. 2014).

The rest of these enriched pathways, on the other hand, represented potentially novel pathways regulated by ASPP2 as their associations with ASPP2 have not been previously reported.

5.2.6 Validation of *TP53BP2* gene coexpression results

So far, our bioinformatic pipeline had produced a ranked gene list which indicates the relative strength of association of each gene with *TP53BP2* and from this ranked gene list 135 *TP53BP2* coexpressed pathways were identified. However, these results could not inform us of causality or the type of relationships, such as transcriptional regulation, (post-)translational regulation, and protein-protein interactions, between *TP53BP2* and its coexpressed pathways.

To validate the pan-cancer gene coexpression results obtained from our bioinformatic pipeline, as well as to decipher the types of relationships between ASPP2 and the resulting *TP53BP2* coexpressed pathways, the pan-cancer gene coexpression results were integrated with the results from the following multi-omics experiments performed by using the ASPP2-Knockout (ASPP2-KO) and wild-type (WT) human osteosarcoma U2OS cancer cell lines as described in chapter 3: (i) 3'RNA-seq experiment using ASPP2-KO and WT U2OS cells in biological quadruplicate to assess whether *TP53BP2* coexpressed genes and enriched pathways identified by the coexpression analysis were regulated by ASPP2 at the transcriptional level; (ii) Interactome analysis performed using V5-tagged ASPP2 in ASPP2-KO U2OS cells to examine whether direct physical interactions exist between *TP53BP2* and its coexpressed genes; (iii) Phosphoproteomics analysis using ASPP2-KO and WT U2OS cells to assess whether *TP53BP2* coexpressed genes and enriched pathways identified by the coexpression analysis were regulated by ASPP2 at the post-translational (phosphorylation) level, given that ASPP2 is a regulatory subunit of protein phosphatase PP1 as discussed in Introduction subsection 1.4.5.3.

5.2.6.1 Integrative analysis with ASPP2 transcriptomic data in U2OS cells

Two different approaches: (i) a gene-level approach and (ii) a pathway-level approach were used for integrating gene coexpression results with 3'RNA-seq results.

In the gene-level approach, pre-ranked GSEA was performed using the genes differentially expressed (DEGs) between ASPP2-KO and WT U2OS cells from the 3'RNA-seq as the gene set, and the ranked gene list based on each gene's pan-cancer ASPP2 coexpression scores as the pre-ranked gene list. $\text{FDR} < 0.05$ and absolute log fold-changes > 2 were used for determining DEGs in the RNA-seq analysis (for details please refer to Materials and Methods subsection 2.7.3). The aim of this approach was to determine whether members of the gene set, i.e. the DEGs between ASPP2-KO and WT U2OS cells, tend to occur toward the top (or bottom) of our pre-ranked gene list, in which case the gene set would be correlated with the top- (or bottom-) *TP53BP2* coexpressed genes that are estimated by our pan-cancer coexpression analysis.

In total, 121 ASPP2-downregulated DEGs, i.e. genes overexpressed in ASPP2-KO U2OS cells [Table S5], and 158 ASPP2-upregulated DEGs, i.e. genes under-expressed in ASPP2-KO U2OS cells [Table S6], were identified by differential expression analysis.

Neither the genes downregulated in ASPP2-KO U2OS cells relative to wild-type [Figure 5.6A] nor the genes upregulated in ASPP2-KO U2OS cells relative to wild-type [Figure 5.6B] were enriched in the pre-ranked gene list ranked by *TP53BP2* pan-cancer coexpresion scores.

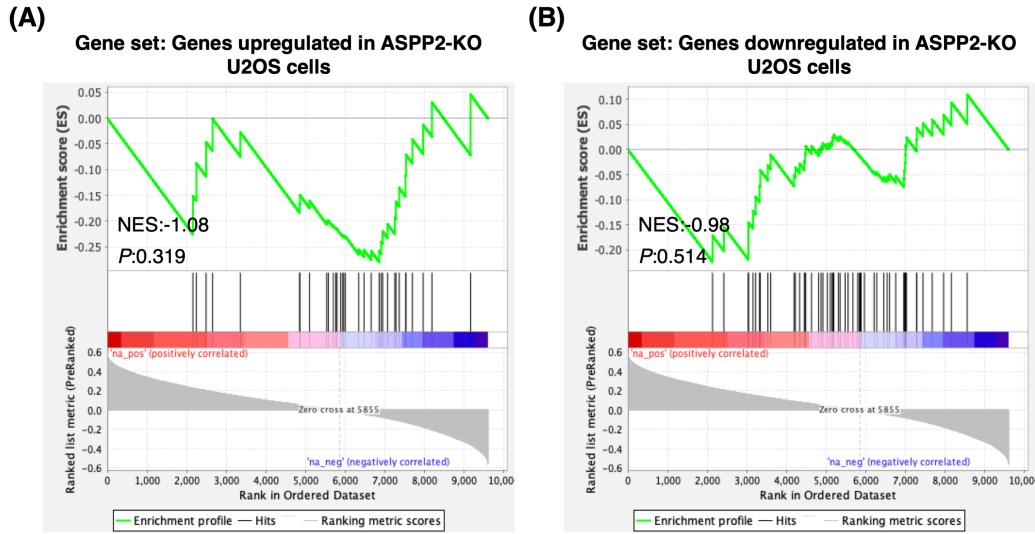


Figure 5.6: Integrative analysis with ASPP2 transcriptomics data in U2OS cells. Pre-ranked GSEA analyses showing the enrichment scores (ESs) of two gene sets, (A) Genes upregulated in ASPP2-KO U2OS cells and (B) Genes downregulated in ASPP2-KO U2OS cells, derived from U2OS ASPP2 transcriptomic data, in the ranked gene list derived from ASPP2 pan-cancer gene coexpression analysis. The ES is represented as a running-sum statistic colored in green. Starting from the top of the ranked gene list and considering one gene at a time, the ES increases when the encountered gene belongs to the gene set, and decreases when the encountered gene is not in the gene set. An enriched pathway, i.e. pathway member genes overrepresented close to the top or the bottom of the ranked gene list, tends to have an asymmetric distribution of the running sum. DEGs: differentially expressed genes; NES: normalized enrichment score; *P*: nominal p-value.

Although these results may suggest that transcriptional effects of ASPP2 in the U2OS cells under the experimental conditions used in our RNA-seq experiment might not explain the relationships between *TP53BP2* and its coexpressed genes captured by the pan-cancer coexpression analysis, we noticed that out of the total 279 differentially expressed genes between ASPP2-KO and wild-type U2OS cells identified in the RNA-seq experiment, only 87 of them were detected in the TCGA pan-cancer RNA-seq data. This low overlap may be due to various biological factors, such as the biological differences between cancer cell lines and patient biopsies, or other technical factors.

Since the discrepancy of the detected genes between the RNA-seq in the U2OS cells and the TCGA pan-cancer RNA-seq dataset may affect the outcome of the pre-ranked GSEA if the DEGs from the 3'RNA-seq were used as the gene set, we took a pathway-level approach to integrate the results of the pan-cancer gene coexpression analysis with the U2OS RNA-seq data.

To do this, we first performed pre-ranked GSEA directly on the results of the differential expression analysis from the 3'RNA-seq, using the entire gene set collection from the REACTOME Pathway Database (version 7.0) as the gene set. For the pre-ranked gene list, we used all the detected genes in the RNA-seq experiment ranked by their respective log fold change between ASPP2-KO and WT U2OS cells. Any enriched pathways from this pre-ranked GSEA were then compared with the 135 enriched pathways identified by the *TP53BP2* pan-cancer coexpression analysis to assess if overlapping enriched pathways existed.

Using nominal p-value < 0.01 as the threshold, 11 REACTOME pathways were found to be positively enriched in ASPP2-KO relative to WT U2OS cells [Table 5.2].

Table 5.2: Enriched REACTOME pathways in ASPP2-KO compared to WT U2OS cells

Pathway	NES	FDR
REACTOME KERATINIZATION	2.0465307	0
REACTOME FORMATION OF THE CORNIFIED ENVELOPE	1.9853158	0
REACTOME BILE ACID AND BILE SALT METABOLISM	1.7884121	0.003710575
REACTOME ION HOMEOSTASIS	1.742859	0.001872659
REACTOME ACTIVATION OF THE MRNA UPON BINDING OF THE CAP BINDING COMPLEX AND EIFS AND SUBSEQUENT BINDING TO 43S	1.6883214	0.002074689

Table 5.2 – continued from previous page

REACTOME Pathway	NES	FDR
REACTOME CARDIAC CONDUCTION	1.6629821	0
REACTOME GLUCONEOGENESIS	1.6592284	0.009578544
REACTOME MUSCLE CONTRACTION	1.6365694	0
REACTOME ACTIVATION OF NMDA RECEPTORS AND POSTSYNAPTIC EVENTS	1.6181237	0.005940594
REACTOME NEUROTRANSMITTER RECEPTORS AND POSTSYNAPTIC SIGNAL TRANSMISSION	1.4739096	0.001934236
REACTOME NEURONAL SYSTEM	1.4495864	0

Abbreviations: NES - Normalized enrichment score; FDR - False discovery rate.

Among these 11 enriched pathways, "REACTOME ACTIVATION OF THE MRNA UPON BINDING OF THE CAP BINDING COMPLEX AND EIFS AND SUBSEQUENT BINDING TO 43S" and "REACTOME GLUCONEOGENESIS" were also identified in the pan-cancer coexpression analysis. This overlap suggests that the coexpression relationships between *TP53BP2* and the leading edge subset of genes involved in these two pathways may likely be explained by ASPP2-mediated transcriptional suppression. Further investigations would be needed to elucidate the potential mechanisms underlying such suppression.

Taken together, these results indicated that some coexpression relationships identified in the pan-cancer coexpression analysis could be a result of ASPP2-mediated transcriptional suppression, supported by the overlap of the enriched translation-related and glucogenesis-related pathways - identified in both ASPP2 coexpression analysis using the TCGA pan-cancer dataset and the differential

expression analysis of the 3'RNA-seq from WT and ASPP2-KO U2OS cells.

5.2.6.2 Integrative analysis with ASPP2 interactomic data in U2OS cells

We next examined whether some of the *TP53BP2* coexpressed genes were physically interacting partners of ASPP2 by integrating the results of gene coexpression analysis with ASPP2 interactomic analysis.

To identify potential ASPP2 interacting partners, an immunoprecipitation-mass spectrometry (IP-MS) approach was used [Figure 5.7A]. Briefly, V5-tagged ASPP2 or empty vector was transiently transfected in ASPP2-KO U2OS cells and immunoprecipitated from the whole cell lysates using anti-V5 antibody. Empty vector was used as a negative control in this experiment. In cooperation with Dr. Thomas Leissing and Ms Georgina Berridge from the Ludwig Institute, 5% co-immunoprecipitates were saved for validation by western blot (WB) and the rest were further processed for in-solution tryptic digest and the resulting peptides were analyzed by liquid chromatography-mass spectrometry (LC-MS). Label-free quantification was then performed using Progenesis for protein quantification. The two IPs were normalised against all proteins identified in the IP. To minimize false positives, only proteins with at least two unique peptides and those that were detected in all three biological replicates of the same condition were used in subsequent analysis.

The abundance of ASPP2 and 33 other proteins were significantly enriched in V5-ASPP2-transfected U2OS cells in comparison to those transfected by empty vector control [Figure 5.7B], suggesting that these 33 proteins were potential ASPP2 interactors, which included known interactors including PP1 (Nicholas R. Helps

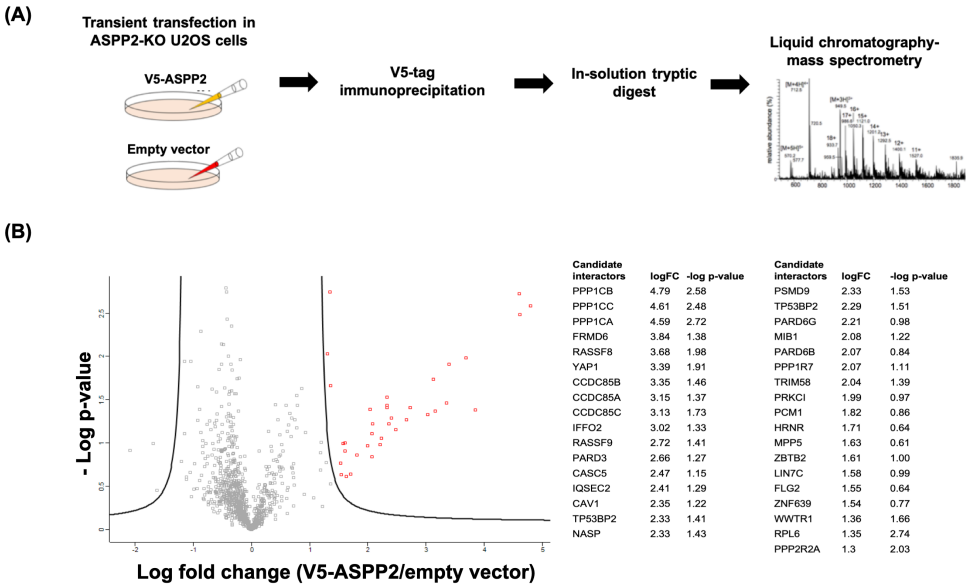


Figure 5.7: Identification of potential ASPP2 interactors using IP-MS. (A) Outline of the interactomic workflow using IP coupled with LC-MS to uncover potential ASPP2 interactors. (B) Left: volcano plot from the MS data showing the magnitudes and significance of the cellular proteins interacted with V5-ASPP2 compared with the control (empty vector). The x-axis represents the log2 fold change (V5-ASPP2/empty vector) and y axis represents log10 p-value showing statistical significance. The curved black line represents the threshold cutoff of $s_0 = 3$ and $FDR = 0.5$. Details for the parameters of threshold cutoff could be found in Materials and Methods. s_0 : artificial within-group variance (Tusher et al. 2001); FDR: false discovery rate. Right: list of candidate ASPP2 interacting proteins (colored in red in the volcano plot). LogFC: log fold change.

et al. 1995), PAR3 (W. Cong et al. 2010; Sottocornola et al. 2010), YAP (Royer et al. 2014), and TAZ (C.-Y. Liu et al. 2011). WB validation of V5-ASPP2, PP1 α , RASSF8, and YAP was performed by former Lu lab member Dr. Thomas Leissing and are shown in Figure S1.

To understand if these 33 candidate ASPP2 interactors were enriched for particular functions, they were used as an input for functional association and enrichment analyses by STRING and were found to be enriched in 27 KEGG pathways [Table S8], among which Hippo signaling pathway represented the top enriched pathway. Interestingly, PP1 family members (encoded by *PPP1CA*, *PPP1CB*, *PPP1CC*) contributed to the enrichment of 21 KEGG pathways and were

solely responsible for the enrichment of 13 KEGG pathways [Table S8], revealing a high degree of redundancy behind the enrichment of these KEGG pathways but also suggesting that PP1 may be a key component of ASPP2's mechanism of actions in various biological processes.

We then integrated the results from the interactomic analysis with the results from pan-cancer gene coexpression analysis. To do so, pre-ranked GSEA was performed using the 33 candidate ASPP2 interacting proteins as the gene set, and the ranked gene list containing *TP53BP2* coexpressed genes as the pre-ranked gene list.

The gene set containing 33 candidate ASPP2 interactors showed a statistically significant, positive enrichment [Figure 5.8A], suggesting that it was positively correlated with *TP53BP2* coexpressed genes. Enrichment analysis of the leading edges of this gene set identified three enriched pathways. They included Hippo signaling pathway, tight junction and human papillomavirus infection, with the latter two enriched pathways having overlapping leading edges with the Hippo signaling pathway [Figure 5.8B]. This finding suggests that physical interaction with ASPP2 may be one of the likely explanations for some of the *TP53BP2* coexpressed genes and the enriched Hippo pathway identified in the pan-cancer coexpression analysis.

5.2.6.3 Integrative analysis with ASPP2 phosphoproteomic data in U2OS cells

One likely function of ASPP2 in protein-protein interactions would be to mediate serine/threonine dephosphorylation of its binding partner(s) due to its ability to bind to PP1 as a holoenzyme, with known substrates including Hippo effectors YAP and TAZ. However, the impact of ASPP2:PP1 holoenzyme on the global phosphoproteome remains unexplored. To address this question, as well as to

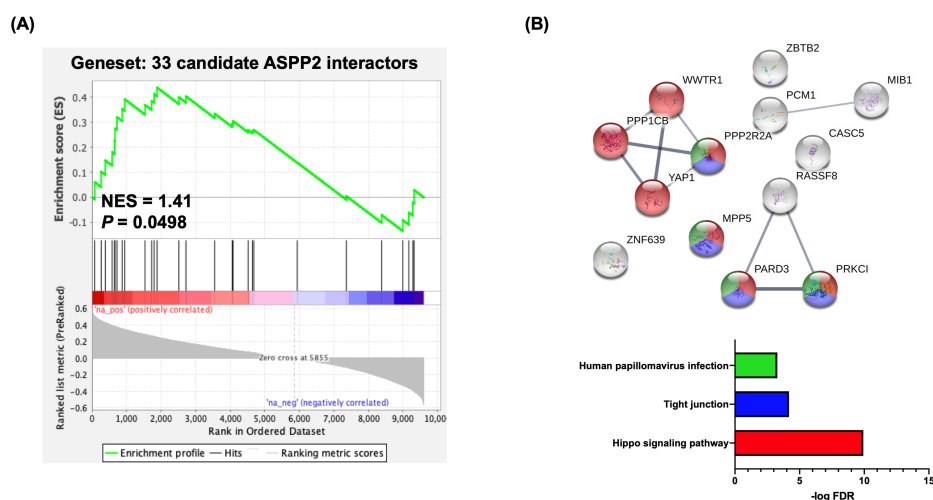


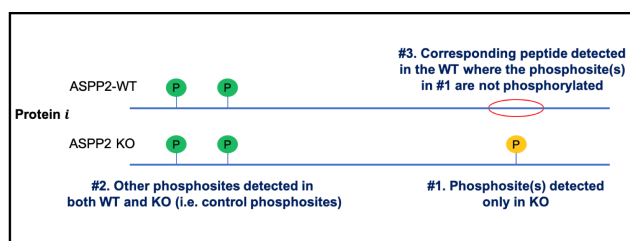
Figure 5.8: Integration of ASPP2 interactomic analyses with *TP53BP2* pan-cancer gene coexpression analysis. (A) Pre-ranked GSEA analyses showing the level of enrichment of the gene set consisting of 33 candidate ASPP2 interactors in the ranked gene list derived from *TP53BP2* pan-cancer coexpression analysis. NES: normalized enrichment score; *P*: nominal p-value. (B) STRING protein-protein interaction network of the leading edges that contributed to the enrichment of the 33 candidate ASPP2 interactors in (A). Leading edges involved in Hippo signaling pathway, tight junction, and human papillomavirus infection were highlighted in red, blue, and green, respectively. The statistical significance, in terms of $-\log$ FDR, of these three enriched pathways were indicated by the bar graph below. FDR: false discovery rate.

understand whether phosphorylation regulation could explain certain *TP53BP2* coexpressed genes and pathways, we first conducted a phosphoproteomic analysis using ASPP2-KO and WT U2OS cells in biological triplicate, before integrating the results from phosphoproteomic analysis and pan-cancer gene coexpression analysis. The phosphopeptide enrichment and liquid chromatography electrospray ionization tandem mass spectrometric measurements were performed by Dr. Kundan Sharma, a former post-doc of our lab and we jointly performed the phosphoproteomic data analysis.

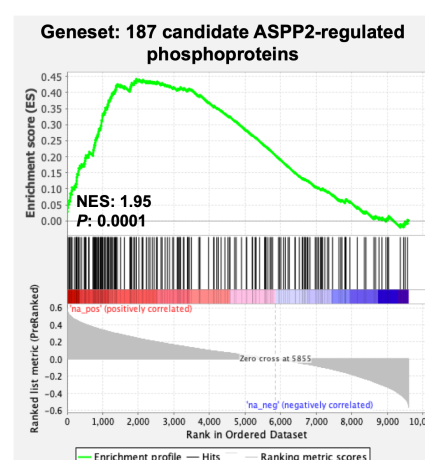
To identify potential ASPP2-regulated serine/threonine phosphosites, we used a highly stringent approach that consisted of 3 criteria [Figure 5.9A]: (i) the candidate

phosphosite must be binarily detected, i.e. the residue of interest is phosphorylated only in ASPP2-KO condition; (ii) Peptide(s) containing a corresponding non-phosphorylated site must also be detected in the WT condition; (iii) at least one control phosphosite, i.e. a phosphosite that is present in both WT and ASPP2-KO conditions, must be detected for the same protein.

(A)



(B)



Using this approach, we detected 268 phosphosites, corresponding to 187 proteins, that were binarily present in ASPP2-KO U2OS cells [Table S9]. These results were then integrated with the results from *TP53BP2* pan-cancer coexpression analysis by performing pre-ranked GSEA, in which the 187 candidate ASPP2-regulated phosphoproteins were used as the gene set, while the ranked gene list derived from the *TP53BP2* pan-cancer coexpression analysis was used as the pre-ranked gene list. This gene set showed a strong, positive enrichment [Figure 5.9B], suggesting that it was highly correlated with genes that were positively coexpressed with *TP53BP2*. Of note, the level of enrichment of this gene set, indicated by the normalized

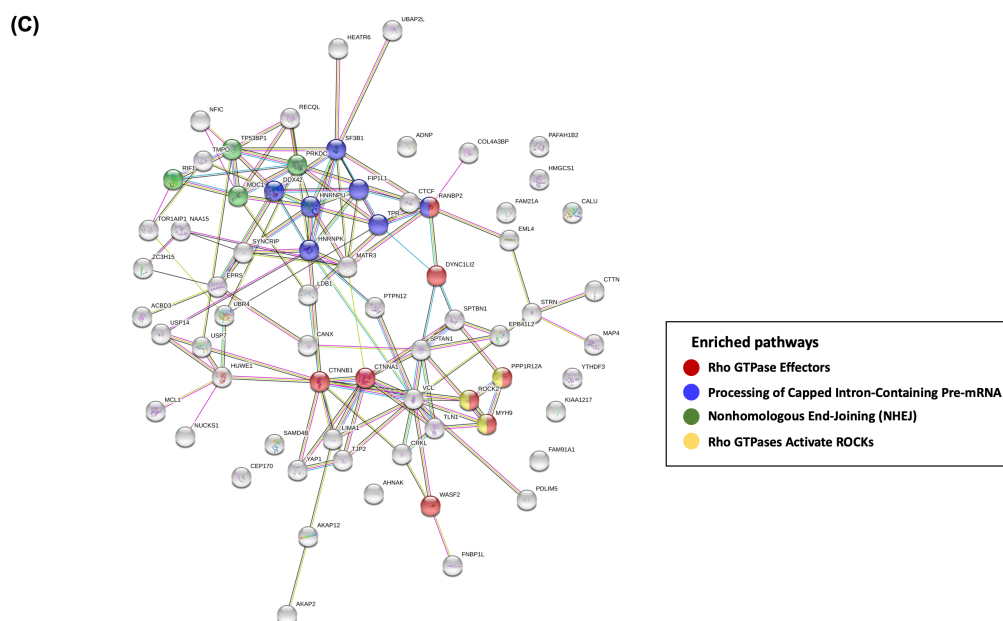


Figure 5.9: Integration of ASPP2 phosphoproteomic analysis with *TP53BP2* pan-cancer gene coexpression analysis. (A) Schematic showing the criteria for determining ASPP2-regulated serine/threonine phosphosites in the U2OS ASPP2 phosphoproteomics data. (B) Pre-ranked GSEA analyses showing the level of enrichment of the gene set consisting of 187 candidate ASPP2-regulated phosphoproteins in the ranked gene list derived from *TP53BP2* pan-cancer coexpression analysis. NES: normalized enrichment score; *P*: nominal p-value. (C) STRING protein-protein interaction network of the leading edges that contributed to the enrichment of the 187 candidate ASPP2-regulated phosphoproteins in (B). Four enriched REACTOME pathways with false discovery rate < 0.01 were highlighted.

enrichment score, was more significant than that from the gene sets (described in the previous sub-sections) derived from U2OS transcriptomic and proteomic datasets. Enrichment analysis of the leading edges identified 4 statistically significant pathways with false discovery rate (FDR) < 0.01: RHO GTPase effectors; RHO GTPases Activate ROCKs; Nonhomologous End-Joining (NHEJ); and Processing of Capped Intron-Containing Pre-mRNA [Figure 5.9C]. The first three pathways coincided with two of the 14 *TP53BP2* coexpressed pathways, i.e. Rho GTPases Signaling and DNA Repair, whereas Processing of Capped Intron-Containing Pre-mRNA contained genes involved in pre-mRNA splicing and were functionally related to the

translation pathway identified by the pan-cancer *TP53BP2* coexpression analysis. These overlaps suggested that ASPP2-mediated phosphorylation regulation was one likely reason underlying the coexpression relationships between *TP53BP2* and genes involved in Rho signaling pathway, DNA repair, and pre-mRNA splicing.

5.2.6.4 Wet-lab validation of selected *TP53BP2* coexpressed pathways

Apart from integrating the results from *TP53BP2* pan-cancer coexpression analysis with the aforementioned omics analyses conducted in ASPP2-KO and WT U2OS cells, we undertook wet-lab investigations to validate several *TP53BP2* coexpressed pathways, including Rho signaling, circadian clock, and Hippo signaling (discussed in chapter 6).

Rho signaling was one of the *TP53BP2* coexpressed pathways identified by our pan-cancer coexpression analysis [Figure 5.5] and had a significant overlap with the functional enrichment of the candidate ASPP2-regulated phosphoproteins in U2OS cells. We noticed that two major Rho effectors, ROCK1 and ROCK2, were among the top 15 *TP53BP2* coexpressed genes, and our phosphoproteomic analysis also suggested that ROCK2 is a potential substrate for ASPP2:PP1 holoenzyme as phosphorylation of ROCK2 serine 1132 was detected in ASPP2-KO but not wild-type U2OS cells [Table S9]. We therefore hypothesized that ASPP2 may regulate Rho/ROCK activity. This hypothesis was tested by our lab member, Mr. Jie Yin, who found that the depletion of ASPP2 in U2OS cells led to impaired ROCK activity upon Rho activation and largely abolished the phosphorylation of ROCK substrates MYPT1 and myosin light chain [Unpublished, data not shown]. Investigations are under way to further elucidate the underlying mechanism(s).

A major goal of our coexpression analysis was to identify potential signaling pathway(s) that might help us understand how ASPP2 may indirectly regulate HIF transcriptional activity. Among the six major groups of *TP53BP2* coexpressed pathways identified by the pan-cancer coexpression analysis [Figure 5.5], circadian clock signaling in the cell signaling group is one of the pathways that have been shown to crosstalk with the HIF pathway and regulate a subset of HIF target genes, including *EGLN3*, in a cell-type dependent manner (Adamovich et al. 2017; Peek et al. 2017; Y. Wu et al. 2017). A strong overlap was also reported between the DNA-binding loci of the core circadian clock effector brain and muscle ARNT-like 1 (BMAL1) and those of HIF1A, suggesting a strong co-regulation between these two pathways at the chromatin level (Y. Wu et al. 2017). In mammals, the clock circadian regulator (CLOCK) and BMAL1 heterodimerize to transactivate a number of E-box containing target genes, among them were the core genes, the period (*PER*) and cryptochrome (*CRY*), which in turn regulate CLOCK/BMAL1 activity in a feedback loop that contributes to the oscillating clock activity and consequentially the circadian rhythmicity (Buhr et al. 2013). Coincidentally, the U2OS cell line that we were using has been reported to display rhythmic transcription of core clock genes and is one of the commonly used cell lines to study circadian rhythm (Altman et al. 2015; Maier et al. 2009; Vollmers et al. 2009). We therefore preliminarily assessed the rhythmic transcriptional expression of *PER1* and *CRY1* in WT and ASPP2-KO U2OS cells under normoxia and hypoxia.

To initiate circadian rhythm under normoxia, these two cell lines were first synchronized with dexamethasone, and the mRNA expression of *PER1* and *CRY1* were measured every 6 hours in a 30-hour timecourse.

Consistent with the literature, in normoxic WT U2OS cells these two core clock components exhibited rhythmic expression profiles with peak expression at the 18-hour timepoint [Figure 5.10A and 5.10B]. However, in normoxic ASPP2-KO cells, such rhythmicity was noticeably dampened, suggesting that ASPP2 may be required to ensure rhythmic circadian clock signaling [Figure 5.10A and 5.10B].

For hypoxia treatment, ASPP2-KO and WT U2OS cells were incubated in a hypoxia chamber where they were synchronized with pre-equilibrated dexamethasone-containing media and the mRNA expression of *PER1* and *CRY1* were measured every 6 hours in a 30-hour timecourse. Interestingly, in sharp contrast to our observation under normoxia, *PER1* and *CRY1* were both suppressed after 6 hours of hypoxia in both cell lines and their rhythmic expression patterns were significantly dampened [Figure 5.10C and 5.10D], suggesting that the circadian clock activity in these U2OS cell lines might be suppressed under hypoxia, which could likely be caused by the acidic cell culture condition induced by hypoxia, as reported by Walton *et al.* (Walton et al. 2018).

Taken together, these preliminary findings suggested that ASPP2 may regulate the circadian clock activity under normoxia, at least for the two core clock target genes tested. However, since in our experimental model this pathway was likely to be suppressed under hypoxia, and due to the time constraint and the incompleteness of our current experimental model, the investigation of whether ASPP2 may regulate HIF transcriptional activity through its regulation of the circadian clock signaling was not pursued further in this study. Further investigations would be required to better understand the role of ASPP2 in the circadian clock pathway.

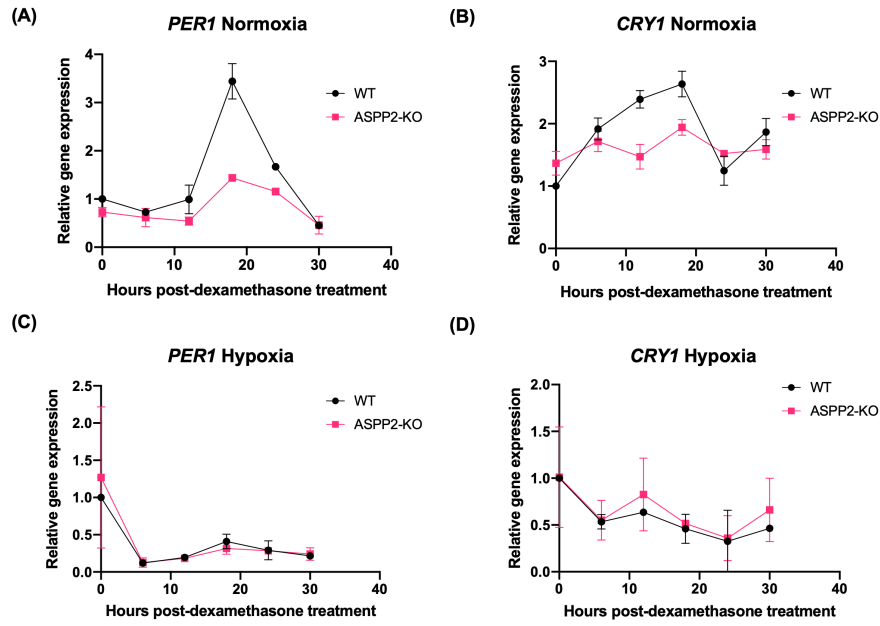


Figure 5.10: Investigating the effects of ASPP2 knockout on the rhythmic expression of *PER1* and *CRY1* under normoxia and hypoxia. Time-course qPCR analyses of gene expression of (A, C) *PER1* and (B, D) *CRY1* in dexamethasone-synchronized WT and ASPP2-KO U2OS cells under normoxia and hypoxia (0.5% O₂), respectively. Values were mean $\pm$ standard deviation; biological duplicate and technical triplicate for each group.

5.3 Summary and Brief Discussion

In this chapter we constructed a bioinformatics pipeline that employed gene coexpression analysis and gene set enrichment analysis to successfully identify, for the first time, a set of *TP53BP2*-coexpressed pathways with a commonly shared enrichment pattern at the pan-cancer level.

A serious problem that may occur when assessing gene-gene coexpression is the presence of confounders, which lead to spurious associations and false positives coexpressed genes. To alleviate the adverse effects brought about by confounders, our pipeline uses a regression-based approach to identify and adjust for potential confounders from given suspected variables. We identified SCN as a potential confounder as it is significantly associated with the pan-cancer *TP53BP2* expression,

likely as a result of focal- and arm-level SCNA events, which lead to coordinated changes in the expression of genes in proximity to each other. The association between SCN and gene expression had previously been reported in several studies and while the construction of our pipeline was in progress, Cai *et al.* published a similar but distinct regression-based coexpression method called GRACE that specifically controlled for SCN's effects (L. Cai et al. 2017). We therefore used their method as a benchmark to evaluate the effectiveness of our method using two genes, *EIF2A* and *PARP2*, that have relatively well characterized functions in translation and cell cycle, respectively. In 10 different TCGA cohorts, the results obtained by our method were comparable, if not better, than those from GRACE in terms of the removal of SCN's confounding effects and the number of identified functionally-related genes, thus supporting the effectiveness of our method.

Combining gene coexpression analysis and gene set enrichment analysis on the pan-cancer TCGA RNA-seq dataset, we identified 135 *TP53BP2*-coexpressed REACTOME pathways that exhibited a pan-cancer enrichment pattern and could be further classified into ten major functionally-related groups. These groups consisted of pathways that were related to RAS GTPases Signaling and Hippo Signaling - both known to be associated with ASPP2 - and for the first time our data showed that these two pathways are also linked to *TP53BP2* at the pan-cancer level, suggesting that ASPP2 may be able to regulate the expression or activities of these pathways or vice versa in multiple types of human cancers.

On the other hand, these groups also included pathways, for example, Rho signaling, that were hitherto unknown to be associated with ASPP2. The pan-cancer association of ASPP2 with these newly identified pathways would thus

provide new directions and unlock novel hypothesis to further explore new biological functions of ASPP2.

A limitation of coexpression analysis is that it tells us about associations but not causation, and it does not tell us exactly the kind of relationship that exists between our gene of interest and the coexpressed genes/pathways, even if that relationship is a regulatory one. To validate whether and what regulatory relationships existed between *TP53BP2* and at least some of its coexpressed genes/pathways, the results above were integrated with those from transcriptomic, interactomic, and phosphoproteomic analyses in ASPP2-KO and WT U2OS cells using GSEA analysis. These results suggested that at least three types of relationships are likely to underlie *TP53BP2* coexpression: transcriptional regulation (e.g. gluconeogenesis), physical interaction (e.g. Hippo signaling), and phosphorylation regulation (e.g. Rho signaling).

What about the other identified *TP53BP2* coexpressed pathways that did not overlap with those enriched in the three omics analyses we used? It is possible that the coexpression between *TP53BP2* and these pathways may reflect indirect correlation, i.e. confounded by identified variables, or it may reflect condition-specific regulatory relationships that do not take place in our current experimental model and condition. Illustrative examples of this latter point are the validation of ASPP2's involvement in Rho signaling and the circadian clock. Both pathways were not enriched in our omics analyses that were conducted in unperturbed ASPP2-KO and WT U2OS cell lines, but ASPP2's effects on ROCK activity and *PER1/CRY1* transcriptional rhythmicity were apparent upon Rho activation (by Mr. Jie Yin) and dexamethasone synchronization, respectively. It is therefore likely that the roles

of ASPP2 in the other yet-to-be-confirmed coexpressed pathways would become apparent in a condition-specific manner.

The second goal we had in mind when constructing this bioinformatic pipeline was to examine if any identified *TP53BP2*-coexpressed pathways may interface with HIFs at the transcriptional level, thus representing potential crosstalk mechanisms that may explain the selective transcriptional regulation of ASPP2 on HIF target genes. Among the identified *TP53BP2* coexpressed pathways, circadian clock was arguably the most closely connected to HIF transcriptional activity, as the core clock transcriptional effectors - bHLH-PAS transcription factors CLOCK and BMAL1 - have been shown to co-regulate a subset of HIF target genes at the transcriptional level. However, since our preliminary data suggested that circadian clock activity might be suppressed in our experimental system under hypoxia, and due to time limitation, this direction was unfortunately not pursued further in this study. Our wet-lab investigations thus focused on another identified *TP53BP2* coexpressed pathway - Hippo signaling - since the crosstalk between Hippo nuclear effector YAP and HIF signaling at the transcriptional level has also been reported (Ma et al. 2015). This will be discussed in the next chapter.

6

Studying the effects of ASPP2 on the potential HIF1 α -YAP crosstalk

6.1 Introduction

As the master regulator of oxygen homeostasis, HIF α is commonly stabilized in at least a subpopulation of tumor cells as a result of tumor hypoxia and aberrant genetic alterations (Linehan et al. 2009; Pouyssegur et al. 2006; Rajani Ravi et al. 2000). The major functional consequence of HIF α stabilization is the transactivation of its target genes that mediate a series of adaptive responses trying to restore oxygen homeostasis. Ironically, many of these genes are also involved in promoting the malignant features of tumors cells.

The transactivation program of HIF is much more complex than a simple on-off switch. Once stabilized, the oxygen-sensitive HIF α forms a heterodimer with the constitutively-expressing ARNT and this heterodimer can bind to the hypoxia-response element (HRE) of over 500 transcriptional targets (Schödel et al. 2011; Smythies et al. 2019; Xia et al. 2009). However, the DNA binding of HIF does

not necessarily lead to the transcription of its target genes, as different subsets of HIF targets that can be eventually transcribed depend on the context. Target selectivity is thus another major point of regulation in the HIF signalling pathway that can influence the outcome of HIF transcriptional response.

One major group of factors that may influence HIF target gene selectivity is transcriptional co-factors that associate with HIF. It is known that the activation of HIF target gene transcription involves a multiprotein complex that consists of the HIF heterodimer and other co-factors. One of the better known co-factors recruited by HIF is the general transcription coactivator p300/CBP, which modulates the transactivation of a subset of HIF targets (Kasper et al. 2005).

Apart from general transcriptional co-factors, HIF has also been shown to interact with nuclear effectors from other signalling pathways through signaling crosstalk. Of particular relevance in this study is the Hippo pathway effector YAP, given the observed coexpression between *ASPP2*'s encoding gene *TP53BP2* and the Hippo pathway identified by the pan-cancer coexpression analysis in the previous chapter.

Hippo signaling is an evolutionarily conserved pathway known for its key roles in organ size control and tissue homeostasis. Central to this pathway is a kinase cascade that senses and transduces various environmental cues, such as those from the cell junction and mechanical stress, into phosphorylation signals to regulate the nuclear availability of YAP and its mammalian homolog TAZ. Activation of Hippo signalling retains YAP/TAZ in the cytoplasm, whereas inhibition of this pathway allows YAP/TAZ to translocate into the nucleus, where they serve as transcriptional co-activators and activated different, and sometimes functionally opposing transcriptional programs depending on the transcription factor that it

binds to, for example, binding to TEAD transcription factors reportedly leads to cell survival and cell growth whereas binding to p73 reportedly activates a pro-apoptotic transcriptional program (Lapi et al. 2008; Matallanas et al. 2007; Tsutomu Oka et al. 2008; Varelas 2014).

Oxygen availability is among the environmental cues that regulate the Hippo pathway at different levels. It has been demonstrated that hypoxia can promote the transcriptional activity of YAP/TEAD through the upregulation of G-protein coupled receptor GPRC5A and its downstream RhoA activity to drive the expression of antiapoptotic regulator BCL2L1 (Greenhough et al. 2018). Hypoxia can also regulate key components of the Hippo pathway to modulate the nuclear availability of YAP. It has been shown that hypoxia induces ubiquitin E3 ligase SIAH2, which in turn inhibits the Hippo pathway by promoting the degradation of YAP's upstream kinase, LATS2 (Ma et al. 2015). In addition, hypoxia can also activate MAP kinase ERK2, and its stimulation reportedly results in the phosphorylation of the cytoplasmic retainer of YAP, 14-3-3 γ , leading to the release of YAP into the nucleus (Jia et al. 2019). Interestingly, in the latter two examples YAP was also shown to be recruited by HIF to promote the transcription of its target genes, including *VEGFA* and *PKM*.

Interestingly, it is known that *ASPP2* can physically interact with YAP and influence its subcellular localization (Royer et al. 2014). This reported interaction is corroborated by the findings of our pan-cancer coexpression analysis in chapter five, in which a strong coexpression relationship between *TP53BP2* and several key components of the Hippo pathway, including *YAP1*, was identified.

Since ASPP2 can affect YAP localization and YAP can be recruited by HIF as a transcriptional co-activator, we hypothesized that ASPP2 may modulate the interaction between HIF and YAP.

6.2 Results

6.2.1 ASPP2 physically interacts with YAP under hypoxia

Although the physical interaction between ASPP2 and YAP in the U2OS cells under normoxic condition was validated by western blot analysis [Figure S5.1], it was not known whether this interaction also takes place under hypoxia. To address this question, we first examined whether ASPP2 and YAP shared similar subcellular localization patterns under hypoxia. By immunofluorescence analysis of U2OS cells subject to hypoxia treatment, it was observed that ASPP2 predominantly had a junctional localization, whereas the majority of YAP was found in the nucleus [Figure 6.1A]. However, there also seemed to be a relatively small amount of YAP situated outside the nucleus and overlapped with ASPP2 staining, suggesting that a small subset of YAP might co-localize with ASPP2 under hypoxia at the condition tested [Figure 6.1A].

It was therefore examined further by co-IP to see if ASPP2 and YAP physically interact under hypoxia. Whole cell lysates from hypoxia-treated WT and ASPP2-KO U2OS cells were subject to IP using anti-YAP and anti-IgG antibody, and the resulting co-immunoprecipitates were analyzed by western blot. Both ASPP2-KO cells and IgG antibody were used as negative controls of the experiment. Analysis of the WB results showed that YAP was significantly enriched in the YAP-IP fraction but not in the IgG-IP fraction, suggesting that the IP was specific [Figure 6.1B].

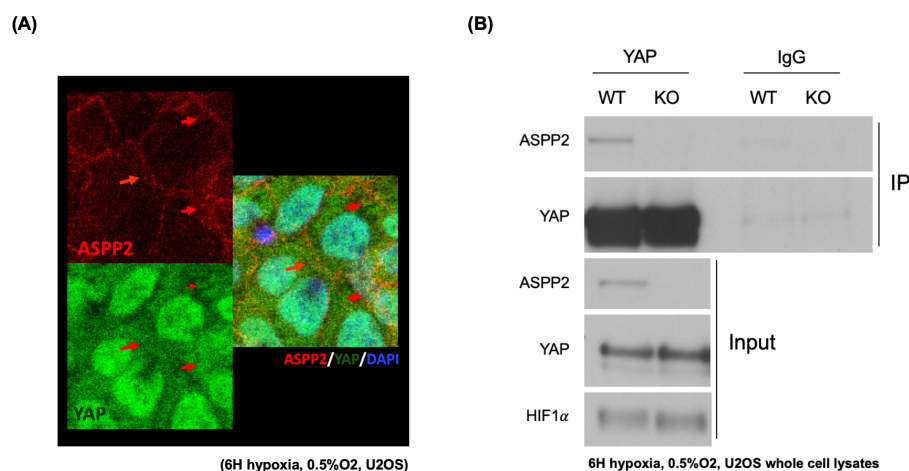


Figure 6.1: ASPP2 physically interacts with YAP under hypoxia. (A) Representative immunofluorescence staining of ASPP2 (red) and YAP (green), with nuclei counterstained with DAPI (blue) in WT U2OS cells incubated for 6 hours under 0.5% O₂. Red arrows indicated co-localization of ASPP2 and YAP. Images are representative examples of two independent replicates of this experiment with at least three images taken per sample. (B) Co-immunoprecipitation (Co-IP) experiment was performed using WT U2OS cells incubated for 6 hours under 0.5% O₂. Endogenous YAP was immunoprecipitated using an anti-YAP antibody or a control anti-IgG antibody. Co-immunoprecipitates and input (5% of the lysates) were analyzed by western blot using the indicated antibodies.

In agreement with the immunofluorescence analysis, a small but clearly observable amount of ASPP2 was detected in the WT U2OS cells in the YAP-IP fraction but not in the ASPP2-KO cells nor the IgG-IP fraction [Figure 6.1B]. The co-IP of ASPP2 by YAP suggested that they could physically interact under hypoxia.

6.2.2 ASPP2-KO affects the nuclear availability of YAP under hypoxia

It had previously been reported that ASPP2 could regulate the subcellular localization of YAP under normoxic condition (Royer et al. 2014). Since ASPP2-YAP interaction was also detected in hypoxia [Figure 6.1B], it was hypothesized that ASPP2 could also regulate the subcellular localization of YAP under hypoxia.

To test this hypothesis, immunofluorescence analysis was performed on WT and ASPP2-KO U2OS cells. It was observed that YAP had a more prominent nuclear localization in WT U2OS cells than in ASPP2-KO cells [Figure 6.2A].

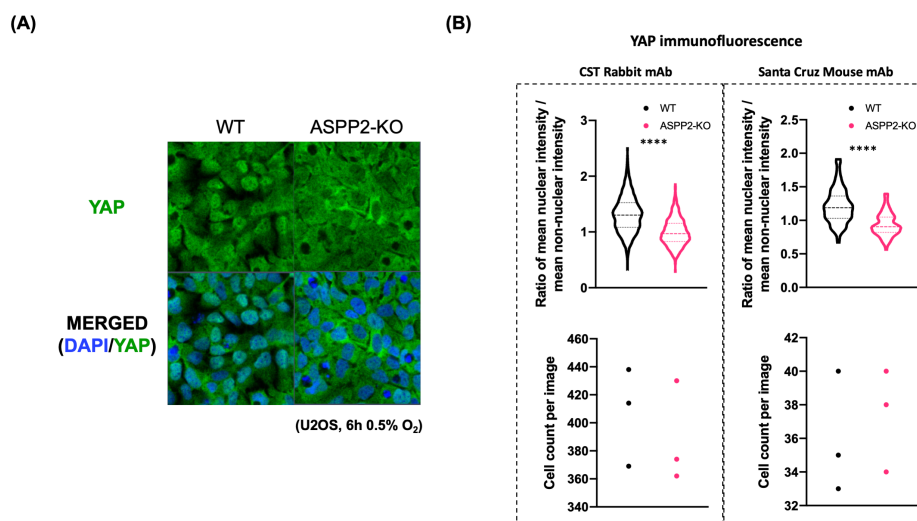


Figure 6.2: Knocking out ASPP2 affects YAP localization under hypoxia. (A) Immunofluorescence images of YAP (green), with nuclei counterstained with DAPI (blue) in WT and ASPP2-KO U2OS cells incubated for 6 hours under 0.5% O₂. Images are representative examples of three independent replicates of this experiment with at least three images taken per sample. (B) Top: Violin plots showing the distribution of the ratio of nuclear YAP mean intensity over non-nuclear YAP mean intensity of WT and ASPP2-KO U2OS cells after 6 hours of hypoxia (0.5% O₂). Each data point represents the said ratio in an individual cell from the analysis of three separate images taken from three independent replicates of each cell line stained by two different antiYAP antibodies: Cell Signaling Technology Rabbit mAb (#14047, left column) and Santa Cruz Mouse mAb (#101199, right column). Bottom: dot plots indicating the cell count per image. Data are mean $\pm$ standard deviation; ****p < 0.0001.

To obtain a more quantitative measurement, nuclear and non-nuclear immunofluorescence intensity of YAP were quantified by ImageJ and the ratio between these two intensities were then calculated to indicate the nuclear availability of YAP. Since cell density was known to influence the nuclear availability of YAP (Zhao et al. 2007), close attention was paid to keep the seeding densities of WT and

ASPP2-KO cells similar and only immunofluorescence images with similar cell counts between WT and ASPP2-KO cells were used for ImageJ analyses. As shown in the violin plots in Figure 6.2B, the majority of WT U2OS cells had an apparently higher nuclear availability of YAP in comparison to the majority of ASPP2-KO cells ($p < 0.0001$) [Figure 6.2B]. To rule out this observation was due to some artifacts of a particular immunofluorescence antibody (CST rabbit mAb), another widely used anti-YAP immunofluorescence antibody (Santa Cruz mouse mAb) was used for the analysis and similar observation was made [Figure 6.2B].

Taken together, these results indicated that ASPP2 could regulate the nuclear availability of YAP under hypoxia.

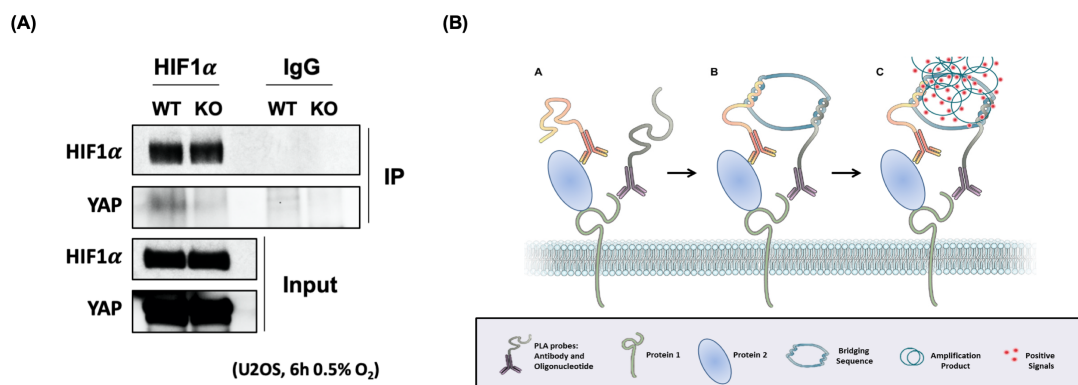
6.2.3 Knocking out ASPP2 attenuated the interaction between YAP and HIF1 α under hypoxia

Since HIF1 α had been reported to recruit YAP as its transcriptional co-factor under hypoxia and our results showed that ASPP2 could regulate the nuclear availability of YAP, it was then tested whether ASPP2 could influence the interactions between HIF and YAP under hypoxia.

For this, IP was used to detect the physical interaction between HIF1 α and YAP. Whole cell lysates from hypoxia-treated WT and ASPP2-KO U2OS cells were subject to IP using anti-HIF1 α antibody and the resulting co-immunoprecipitates were then washed and analyzed by WB.

As shown in Figure 6.3A, HIF1 α was clearly detected in the HIF1 α -IP fraction relative to the input and was not detectable in the IP-IgG fraction, suggesting that the IP was specific. In agreement with the literature that reported HIF-YAP interaction, YAP was detected in the IP-HIF1 α fraction but not in the IP-IgG

fraction, suggesting that it was co-immunoprecipitated by HIF1 α under hypoxia [Figure 6.3A]. Notably, there were visibly more YAP co-immunoprecipitated by HIF1 α in WT U2OS cells than in ASPP2-KO U2OS cells [Figure 6.3A]. On the other hand, the amount of YAP detected in the input fraction was similar in WT and ASPP2-KO U2OS cells [Figure 6.3A]. These results suggested that while knocking out ASPP2 did not affect the expression of YAP under hypoxia, it attenuated the interactions between HIF1 α and YAP, which could partially be due to less nuclear YAP in the ASPP2-KO U2OS cells in comparison to the WT U2OS cells.



To study the interactions between HIF1 α and YAP from another perspective, *in situ* proximity ligation assay (PLA) was used [Figure 6.3B]. The rationale was that PLA does not need to go through the disruptive protein extraction process inherent in co-IP as it can be applied directly on fixed cells, therefore it can facilitate the detection of weak or transient protein-protein interactions while also providing spatial information about the interacting proteins.

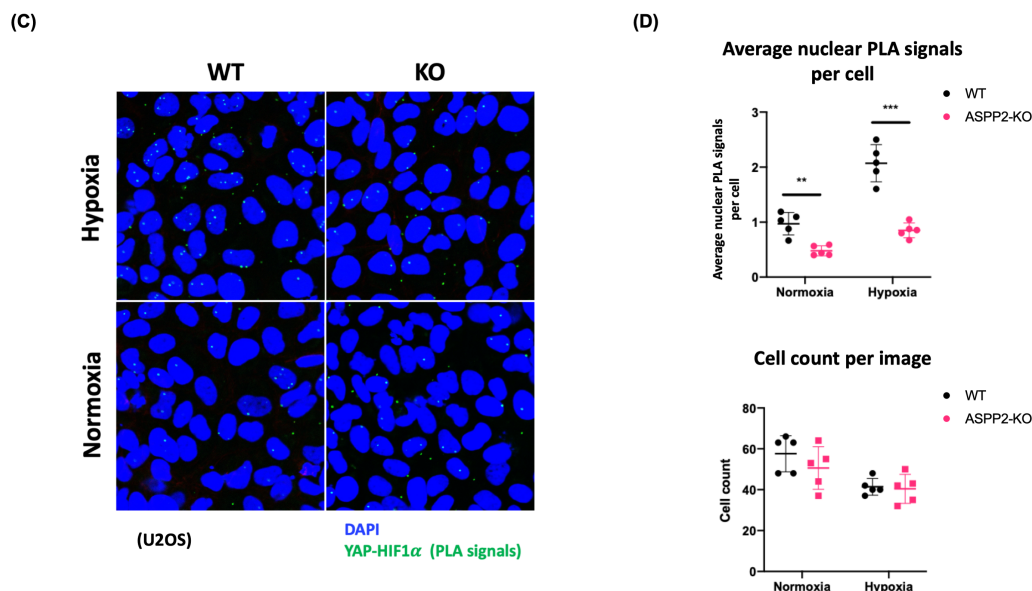


Figure 6.3: Knocking out ASPP2 attenuates the interaction between nuclear HIF1 α and YAP under hypoxia. (A) Co-IP experiment was performed using WT and ASPP2-KO U2OS cells incubated for 6 hours under 0.5% O₂. Endogenous HIF1 α was immunoprecipitated using an anti-HIF1 α antibody or a control anti-IgG antibody. Co-immunoprecipitates and input (5% of the lysates) were analyzed by western blot using the indicated antibodies. (B) Schematics showing the principle of *in situ* proximity ligation assay (PLA). PLA utilizes antibodies coupled to DNA oligonucleotides (oligo) that can be hybridized, amplified and visualized as fluorescent signals when in close proximity (at distances below 40 nm (Alam 2018; Bagchi et al. 2015; Zatloukal et al. 2014)). In the first step, two specific PLA probes, each carrying a unique oligo, respectively bind to their target proteins. In the second step, the oligos then guide the formation of circular DNA strands when bound in close proximity. In the next step the circular DNA would then serve as templates for localized amplification, allowing the interacting pair of proteins to be visualized through microscopy imaging. Adapted from Mereiter *et al.* 2016 with permission (Mereiter et al. 2016). (C) Immunofluorescence images of PLA signals (green), with nuclei counterstained with DAPI (blue) in WT and ASPP2-KO U2OS cells incubated for 6 hours under 21% O₂ (normoxia) or 0.5% O₂ (hypoxia). Images are representative examples of two independent replicates of this experiment with at least five images taken per sample. (D) Top: Dot plot showing the average nuclear PLA signals per cell. Each data point represents nuclear PLA signals in an individual cell from the analysis of five separate images taken from biological duplicates of WT or ASPP2-KO U2OS cells incubated for 6 hours under 21% O₂ (normoxia) or 0.5% O₂ (hypoxia). Bottom: dot plots indicating the cell count per image. Data are mean $\pm$ standard deviation; **p = 0.001, ***p < 0.001. Representative images are in Figure 6.3C

To set up this technique, the interaction between ASPP2 and PP1 was first studied in WT and ASPP2-KO U2OS cells, as this system allowed us to have ASPP2-KO condition as a negative control and there was also PLA-validated anti-PP1 antibody available, whereas the specificity of anti-ASPP2 antibody was previously tested in ASPP2-KO cells. Adopting and optimizing the protocol previously validated by Dr. Patricia Muller (CRUK Manchester), PLA signals were detected only in WT U2OS cells but not in ASPP2-KO U2OS cells (Figure S6.2A), indicating the specificity of the PLA signals. The PLA signals in the nucleus and the cytoplasm could also be successfully quantified using an ImageJ macro developed by our former microscope manager Dr. Mark Shipman (Figure S6.2B).

After setting up the protocol for *in situ* PLA, it was then used to assess the interaction between HIF1 α and YAP. For this, normoxia- and hypoxia-treated WT and ASPP2-KO cells were fixed and then proceeded with the *in situ* PLA protocol. Nuclear PLA signals in the resulting microscopy images were then quantified by the aforementioned ImageJ macro. As cell density was one of the environmental cues that could regulate YAP nuclear localization, only microscopy images with similar cell number were used for the analysis.

In normoxia, there were a higher number of average nuclear PLA signals detected in WT U2OS cells than in ASPP2-KO U2OS cells [Figure 6.3C and 6.3D]. The detection of PLA signals in normoxia indicated that there might be a basal level of HIF1 α expressed in normoxia, as observed in the previous chapter [Figure 4.9 and 4.16]. In hypoxia, the average nuclear PLA signals detected in the WT U2OS cells were significantly increased compared to that in normoxia [Figure 6.3C and 6.3D]. Notably, the average number of nuclear PLA signals in the WT U2OS cells

under hypoxia was significantly greater than that in the ASPP2-KO cells [Figure 6.3C and 6.3D]. On the other hand, PLA signals were barely detected in both WT and ASPP2-KO cells under normoxia and hypoxia when only either anti-YAP or anti-HIF1 α antibody was used for the *in situ* PLA assay [Figure S6.3], suggesting that the PLA signals detected were specific. These results suggested that on average there were more nuclear HIF1 α -YAP interactions in WT U2OS cells than in ASPP2-KO U2OS cells, and this difference was more significant in hypoxia, likely due to more stabilized HIF1 α under hypoxia to interact with YAP in the nucleus.

Taken together, the results from co-IP and *in situ* PLA confirmed the physical interaction between YAP and HIF1 α under hypoxia and knocking out ASPP2 could attenuate this interaction.

6.2.4 Knocking out ASPP2 compromised the binding of YAP to HRE-containing loci of *EGLN3* and *VEGFA*

Since it was observed that ASPP2 could influence the nuclear interaction between HIF1 α and YAP, it was then asked whether this influence would also affect HIF1 α 's recruitment of YAP as a transcriptional co-activator to promote HIF target gene transactivation.

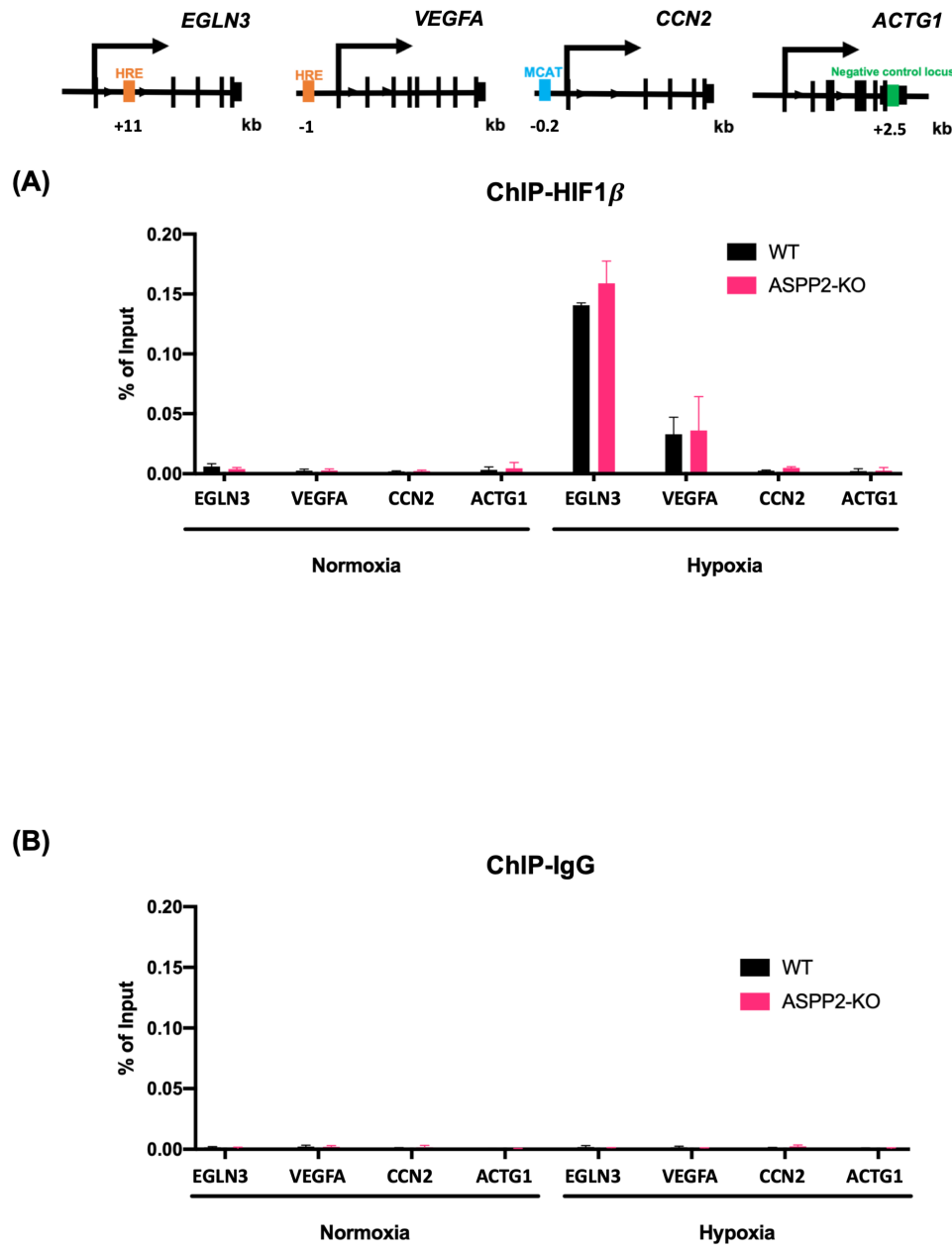
To test this idea, ChIP-qPCR was used to assess the enrichment of HIF and YAP at the HRE-containing loci of HIF targets *EGLN3* and *VEGFA*. Since HIF binds to HRE-containing loci of its target genes as a heterodimer of HIF1 α and HIF1 β , HIF1 β -ChIP was performed in parallel with YAP-ChIP in normoxia- and hypoxia-treated WT and ASPP2-KO U2OS cells. IgG-ChIP was also performed as a negative control. A non-HIF- and non-YAP-binding locus of *ACTG1* was used as a negative control and a previously reported YAP-binding site of its known target gene

CCN2 was used as a positive control for YAP-ChIP (Stein et al. 2015). It should be noted that YAP itself did not bind to the DNA directly, so the YAP-binding site above referred to the site where it was recruited by TEAD transcription factor.

As shown in Figure 6.4A, under normoxia, no significant enrichment of HIF1 β was detected in the HRE-containing regions of its target gene *EGLN3* and *VEGFA*, and the YAP-binding region of *CCN2*, in comparison to the enrichment of IgG [Figure 6.4B] or to the HIF1 β enrichment at the *ACTG1* negative control locus [Figure 6.4A]. The binding of HIF1 β to the 2 HRE-containing regions of *EGLN3* and *VEGFA* was significantly enriched in comparison to that of IgG [Figure 6.4A-B] and the negative control site of *ACTG1* after hypoxia treatment [Figure 6.4A], while there was still no observable enrichment of HIF1 β at the *CCN2* locus [Figure 6.4A]. Consistent with our earlier findings in the previous chapter, the different of HIF1 β enrichment at the two HIF target gene loci were not significantly different in the WT U2OD cells than in the ASPP2-KO cells [Figure 6.4A].

On the other hand, noticeable enrichment of YAP was detected in the YAP-binding locus of *CCN2* in comparison to the negative control region of *ACTG1* under both normoxia and hypoxia [Figure 6.4C], but there was no significant difference in enrichment between WT and ASPP2-KO U2OS cells, or between normoxia and hypoxia, suggesting that YAP's binding to *CCN2* gene locus, at least for the one tested in this assay, was unlikely to be affected by ASPP2 status and oxygen availability [Figure 6.4C].

Similar to the HIF1 β -ChIP, there were no significant enrichment of YAP at the two HRE-containing loci of *EGLN3* and *VEGFA* in both normoxia-treated WT and ASPP2-KO U2OS cells [Figure 6.4C]. Interestingly, there was a small but statistically



significant increase of YAP's enrichment at both of these HRE-containing loci in WT U2OS cells from normoxia to hypoxia ($p < 0.01$ for both *EGLN3* and *VEGFA*) [Figure 6.4C]. Such increase was also significant when compared with that of the IgG control, or YAP's enrichment at the negative control locus of *ACTG1* in hypoxia. Importantly, YAP's enrichment at the HRE-containing loci of *EGLN3* and *VEGFA* in WT U2OS cells were significantly higher than those in the ASPP2-KO U2OS cells

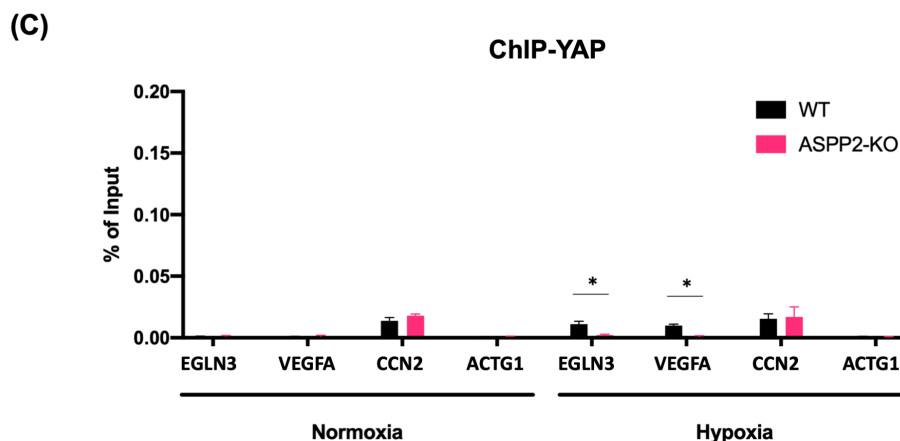


Figure 6.4: Knocking out *ASPP2* compromises the binding of YAP to the HRE-containing region within the gene loci of *EGLN3* and *VEGFA*. Schematics of ChIP-qPCR amplicons (top panel) and ChIP-qPCR analyses with ChIP (A) anti-HIF1 β (B) anti-IgG and (C) anti-YAP antibodies at the enhancer/promoter regions of *EGLN3*, *VEGFA*, and *CCN2*, as well as at the non-HIF/YAP-binding region of *ACTG1*. ChIP experiment was performed using WT and ASPP2-KO U2OS cells incubated under 21% O₂ (normoxia) or 0.5% O₂ (hypoxia) for 8 hours. Biological duplicate and technical duplicate for each locus. Values were mean $\pm$ standard deviation; *p < 0.05; MCAT: M-CAT motif (5'-CATTCCT-3')

(p = 0.04 and p = 0.01 for *EGLN3* and *VEGFA*, respectively) [Figure 6.4C]. These results suggest that YAP was recruited to the HRE-containing loci of *EGLN3* and *VEGFA* under hypoxia and knocking out ASPP2 compromised such recruitment.

6.3 Summary and Brief Discussion

Our work in this chapter revealed, for the first time, that ASPP2 could regulate the nuclear interaction between YAP and HIF1 α , as well as the recruitment of YAP to HRE-containing gene loci of HIF targets *EGLN3* and *VEGFA* under hypoxia.

This chapter set out to examine the hypothesis that ASPP2 could modulate the interaction between YAP and HIF based on the connections between HIF-YAP and ASPP2-YAP.

Using co-IP and immunofluorescence techniques, it was found that ASPP2-YAP

interaction could be observed under hypoxia, and knocking out ASPP2 reduced the nuclear localization of YAP under hypoxia.

We then asked whether this reduction would have any impact on HIF transcriptional activity, given that HIF1 α was reported to interact with YAP and recruit it as a transcriptional co-factor. For this, an *in situ* PLA-based approach was developed and together with co-IP, the physical interaction between YAP and HIF1 α was detected under hypoxia and this interaction was attenuated when ASPP2 was knocked out. Since in chapter four little evidence was found to indicate ASPP2's effect on the stability and subcellular localization of HIF1 α , the attenuation of HIF1 α -YAP interaction upon ASPP2-KO was likely due to the effect of ASPP2 exerted on YAP.

In agreement with previous reports, the enrichment of YAP at the HRE-containing locus of HIF target gene *VEGFA* was observed, and it was found that YAP could also be detected at the HRE-containing locus of another HIF target gene, *EGLN3*. Such enrichment coincided with the detection of HIF1 α 's binding partner, HIF1 β , at the two HRE-containing loci. Both HIF1 β and YAP enrichment at these two loci were hypoxia-specific.

Importantly, the enrichment of YAP at these two loci was significantly reduced when ASPP2 was knocked out, suggesting that not only could ASPP2 modulate the physical interaction between HIF1 α and YAP, it could also affect the recruitment of YAP to the HRE-containing loci of HIF target genes, including *EGLN3* and *VEGFA*, under hypoxia.

Since YAP does not bind to DNA directly, it must be recruited by another transcription factor to its target gene loci. Given that YAP's recruitment to the two

HRE-containing loci of *EGLN3* and *VEGFA* were hypoxia-specific and coincided with HIF1 β enrichment, it was reasonable to speculate that it was recruited by HIF1 α to these loci. However, since the loci used in ChIP-qPCR experiments were 100 to 500 base pairs in length, they also contain binding sites for other transcription factors to which YAP may also bind. For example, TEAD and SMAD motifs were also identified in vicinity of the HRE-containing locus of *EGLN3* used in the ChIP experiments. Further experiments, such as re-ChIP, would be needed to validate whether YAP was recruited by HIF to the *EGLN3* and *VEGFA* loci. More experiments, such as the use of a dominant negative version of YAP, would also be worthwhile for confirming if YAP is the major route through which ASPP2 selectively modulates the expression of HIF target genes such as *EGLN3*.

Overall, this chapter identifies a potential ASPP2-mediated mechanism - through regulating YAP's nuclear localization and its nuclear interaction with HIF1 α - that may explain the selective fine-tuning of HIF target gene expression by ASPP2. Experiments will be needed in the future to further characterize how ASPP2 affects global DNA occupancy of YAP and to validate whether such effects represent a causal mechanism that selectively regulates HIF target gene expression.

7

Final discussion and perspectives

7.1 The role of ASPP2 in HIF signaling

Our work showed that, for the first time, that ASPP2 modulates the gene expression of a subset of HIF targets and controlling YAP subcellular localization may represent a potential mechanism that may explain the regulation of ASPP2 on the expression of HIF target *EGLN3* [Figure 7.1]. This potential mechanism also implies that the modulatory effect of ASPP2 on HIF targets may not be oxygen-dependent or hypoxia-specific, as supported by our qPCR results using DMOG.

Our work also clearly demonstrated that the interplay between ASPP2, FIH, and HIFs is more complex than we originally hypothesized, i.e. as a competitor for FIH hydroxylation. On the one hand, we found that ASPP2 can modulate a subset of HIF targets irrespective of FIH status. On the other hand, we also identified a group of HIF targets whose expression are sensitive to ASPP2 depending on FIH status. However, whether and how ASPP2-FIH interaction affects HIF target gene expression remain an unanswered question.

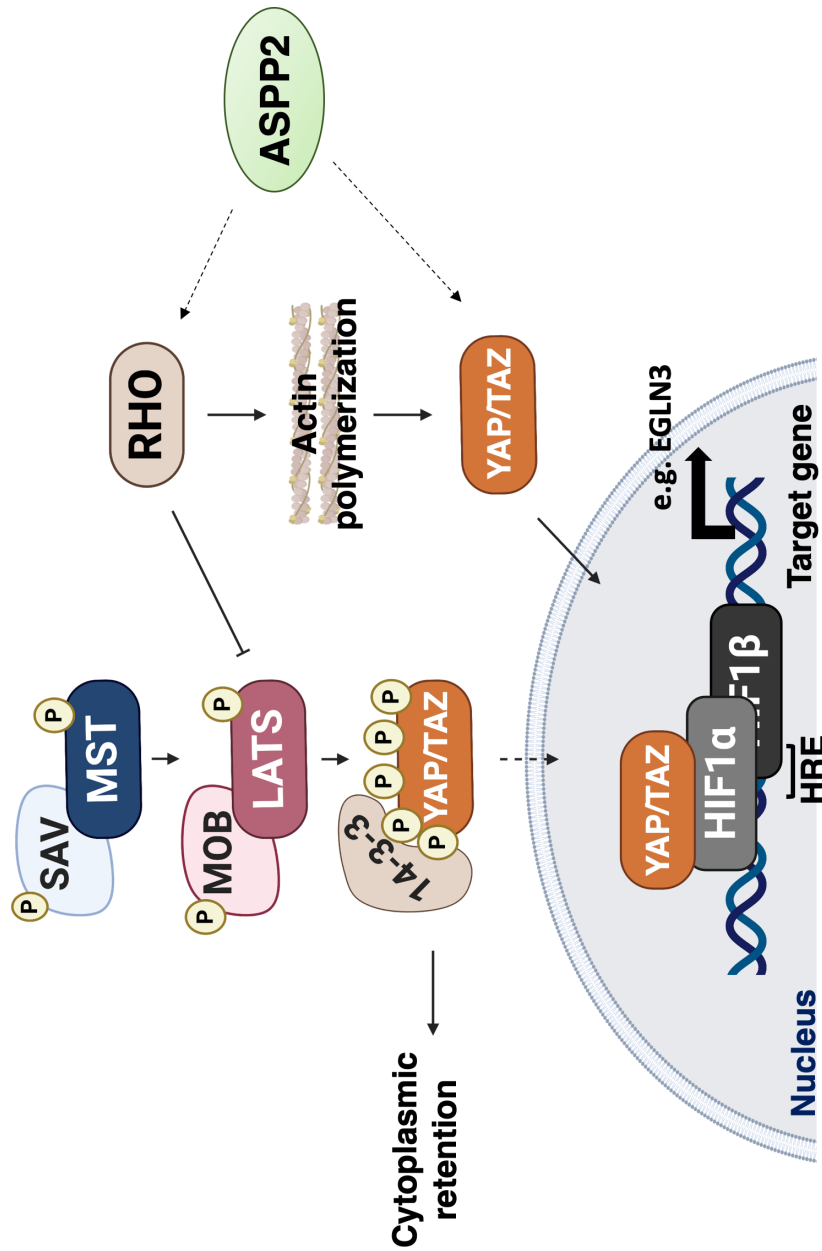


Figure 7.1: Controlling YAP subcellular localization represents one of the potential mechanisms that may explain the regulation of ASPP2 HIF target gene expression. Schematic showing the potential ways of how ASPP2 could affect HIF target gene expression through YAP. HRE: Hypoxia response element.

Mechanistic studies presented in this thesis showed that ASPP2 did not influence the stability and subcellular localization of HIF1 α . Using *EGLN3* as the candidate ASPP2-sensitive HIF target gene, we also found that ASPP2 did not affect the DNA-binding activity of both HIF1 α and its obligate heterodimeric partner HIF1 β at HRE-containing loci within and near *EGLN3* gene locus under hypoxia. Interaction between ASPP2 and HIF1 α was also not detected.

Crosstalk between HIF and Hippo signalling was recently reported as the Hippo effector YAP was demonstrated to act as a transcriptional co-activator of HIF1 α (Ma et al. 2015). In this study, we found that ASPP2 regulated YAP's nuclear availability and its nuclear interaction with HIF1 α . Consistent with these observation, using ChIP-qPCR we detected less YAP at the HRE-containing loci of ASPP2-regulated HIF target genes *EGLN3* and *VEGFA*. Taken together, our work suggested that regulating YAP's nuclear availability could be a potential mechanism through which ASPP2 regulates HIF transcriptional activity. Follow-up investigation would be needed to elucidate whether YAP is recruited directly by HIF1 α to promote the transactivation of its target genes, or whether YAP is recruited by other transcription factor(s), such as TEAD, in close proximity to HIF-binding sites and co-regulates HIF target gene transactivation. For this, it would be worthwhile to use ChIP-seq to understand the global DNA-binding profiles of both YAP and HIF1 α , and how they are influenced by ASPP2 under normoxia and hypoxia. These ChIP-seq results could then be integrated with the transcriptomic data conducted in the same cell lines and experimental conditions to elucidate YAP dependency of ASPP2-regulated HIF target genes.

7.2 The role of ASPP2 in the regulation of YAP

The aforementioned experimental observation suggests that ASPP2 regulates the nuclear localization of YAP. But how? Recent studies suggest that one potential mechanism is directly binding to YAP as a holoenzyme complex with PP1, thereby inducing dephosphorylation on YAP (Royer et al. 2014).

However, results from our pan-cancer gene coexpression analysis suggest that ASPP2 may also regulate YAP in an indirect manner. Using our beta coefficient-based coexpression method BCCA, we found a significant coexpression relationship between the gene expression of ASPP2 and a group of genes involved in Rho/ROCK signaling, including *ROCK1* and *ROCK2*, in the pan-cancer TCGA cohort. This suggests that ASPP2 may have a role to play in this signaling pathway in human cancers. This hypothesis is currently under investigation by our lab member Jie Yin, and preliminary results suggested that ASPP2 could interact physically with ROCK2 (one of the top ASPP2-coexpressed genes) and regulate ROCK activity. Interestingly, YAP has been demonstrated to act downstream of Rho/ROCK signaling independent of Hippo regulators LATS1/2 (Feng et al. 2014). Rho/ROCK are major regulators involved in cellular mechanosensing and activation of Rho/ROCK signaling can induce YAP nuclear localization through promoting the formation of F-actin (Ohashi et al. 2000; Sit et al. 2011), which sequesters scaffold protein angiomotin away from YAP (T. Oka et al. 2012; Shi et al. 2017). YAP is therefore free to translocate to the nucleus where it acts as a transcriptional coactivator to induce the transcription of genes involved in mechanosensing activity (Nardone et al. 2017).

Our finding of ASPP2 as a potential regulator of ROCK suggests that ASPP2 may control YAP nuclear shuttling through regulating ROCK activity [Figure 7.1]. Of note, from our phosphoproteomic analysis we detected serine residue 365 (S365) of YAP to be phosphorylated only in ASPP2-KO U2OS cells. S365 is located in YAP's transactivation domain and its phosphorylation has not been previously reported in the literature. It is possible that the phosphorylation status of this residue may modulate the transactivation activity of YAP. It would therefore be important to understand whether ASPP2 could act at different levels to regulate YAP activity [Figure 7.1].

What could be the functional significance of ASPP2-mediated regulation of YAP? One potential significance would be to fine-tune HIF transcriptional response, as discussed above. Apart from that, another likely functional significance would be to promote a transcriptional program that encodes genes involved in mechanosensing activity, such as the formation of focal adhesion, crosstalk with the extracellular matrix (ECM), and cell motility. Consistent with this notion, our pan-cancer coexpression analysis showed that ASPP2 gene expression is positively associated with the expression of genes involved in cell migration, while transcriptomic analyses showed that ASPP2-up-regulated genes, under both normoxia and hypoxia, are enriched for ECM organization. It would therefore be worthwhile to perform integrative analyses of RNA-seq and YAP ChIP-seq, with Rho/ROCK stimulation such as by treating the cells with lipids, like lysophosphatidic acid, or by changing ECM stiffness, so as to examine how ASPP2 may regulate transcriptional responses to extracellular biophysical cues.

7.3 The role of ASPP2 in cancer

Published studies of ASPP2 so far have mostly used cancer cell lines and knockout mouse models to uncover the potential functions of ASPP2 in cancer and beyond. Despite their usefulness, these experimental systems could not fully recapitulate the heterogeneity and complexity of human cancers and we only have limited knowledge about the roles that ASPP2 is serving in cancer patients.

To offer a complementary view of the roles of ASPP2 in human cancers, we constructed an *in silico* pipeline to infer the functions of ASPP2 from gene coexpression patterns in the TCGA pan-cancer cohort with features of confounders identification and correction to limit the noise from artifacts that may confound any functional relationships between ASPP2 and its coexpressing genes.

From this analysis, several major pathways that were found to positively coexpress with ASPP2 gene expression are oncogenic pathways, such as RAS and Rho GTPases signaling (Haga et al. 2016; Simanshu et al. 2017), Hippo signaling (through the regulation of YAP) (Zanconato et al. 2016, 2019), cell migration (Desgrosellier et al. 2010; Givant-Horwitz et al. 2004), and DNA repair (Hazan et al. 2019; Nicolai et al. 2019; Ochs et al. 2019). Some of these pathways, such as RAS and Hippo/YAP, are known to be associated with ASPP2 based on published reports, while the other ASPP2 coexpressing pathways would require further experimental investigations. On the other hand, we did not observe significant associations between ASPP2 gene expression and p53 signaling-related pathways, such as p53-mediated apoptosis and senescence programs. In addition, amplification and overexpression of *TP53BP2*, the gene that encodes ASPP2, were also observed in a subset of cancer population

in multiple TCGA cancer cohorts, whereas deletion, underexpression, and loss-of-function mutations of *TP53BP2* are rare in the same cohorts. Such observation above are seemingly contradictory to the reported role of ASPP2 as a tumor suppressor and a pro-apoptotic activator of p53.

This contradiction could potentially be explained by the concept of coselection proposed by Sir Peter Ratcliffe (Ratcliffe 2013). The functions of ASPP2 are evolved to serve physiological purposes, but not for pathophysiological conditions like cancers. Therefore during tumorigenesis, different functions of ASPP2 could be conducive, detrimental, or simply neutral to the further development of a cancer cell, depending on the intra- and extracellular contexts. Using the role of ASPP2 in RAS signaling as an example, in pre-malignant cells, up-regulation of ASPP2 expression may stimulate RAS activity, which is oncogenic, but would also come with a penalty as ASPP2 relays aberrant growth signals to p53, which activates tumor suppressive programs. The overriding role of ASPP2 in pre-malignant cells would therefore likely be tumor suppressive, as the said penalty likely outweighs the oncogenic advantage. On the other hand, in malignant cells, where wild-type p53 functions are frequently disabled by various mechanisms (Kandoth et al. 2013; Sherr et al. 2002), the p53-related penalty that comes with ASPP2 activation is relieved. This may be one of the reasons why a subset of cancer populations in multiple TCGA cancer cohorts has upregulated ASPP2 gene expression, which likely confers growth advantages after these cancer cells have evaded the coselection of penalty brought about by ASPP2 upregulation. Therefore it is possible that ASPP2 could play an oncogenic role in cancers in a context-depending manner, given its associations with the aforementioned oncogenic pathways. Consistent with this

notion, we found that ASPP2 expression positively correlates with the aggressiveness of breast cancer subtypes [Figure 7.1A] and p53 mutation status [Figure 7.1B], while high ASPP2 gene expression is associated with poorer relapse-free survival in patients with the more aggressive triple-negative subtype [Figure 7.2A] but not in those with the less aggressive luminal subtype [Figure 7.2B].

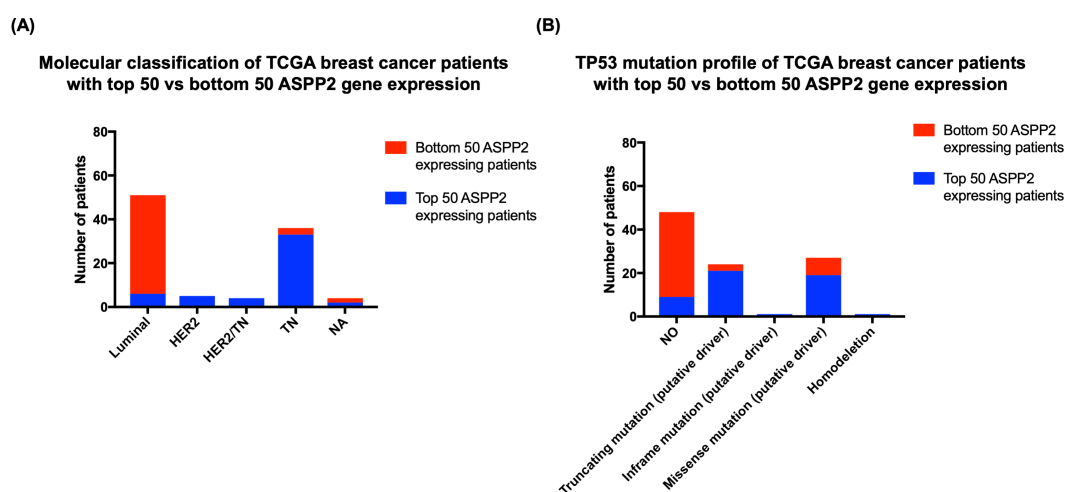


Figure 7.2: ASPP2 expression positively correlates with more aggressive breast cancer subtypes and p53 mutations. (A) Bar plot showing the molecular classification of TCGA breast cancer patients with the top 50 and bottom 50 ASPP2 gene expression. TN: triple negative; HER2/TN: estrogen and progesterone receptor negative but HER2 status missing; NA: missing subtype information; (B) Bar plot showing the mutational profile of *TP53* in TCGA breast cancer patients with the top 50 and bottom 50 ASPP2 gene expression. *TP53* mutation information was obtained using the OncoPrint function on cBioPortal (J. Gao et al. 2013).

Another possibility that could explain ASPP2's connection to the said oncogenic pathways in the pan-cancer TCGA cohort is the existence of ASPP2 splice variants in human cancers (Van Hook et al. 2017; Kampa-Schittenhelm et al. 2013; Samuels-Lev et al. 2001; Takahashi et al. 2004). *TP53BP2* contains several alternative start sites and several truncated forms of ASPP2 have been identified. The N-terminal truncated variant of ASPP2, BBP, was shown to be less capable of stimulating the

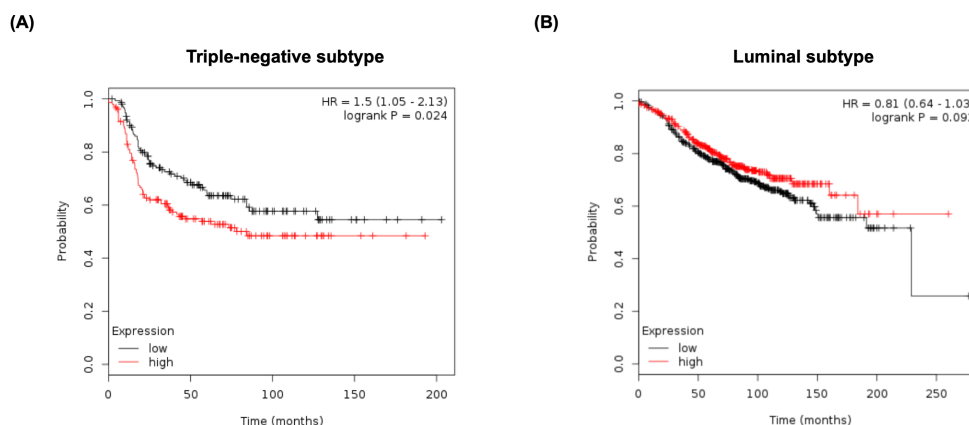


Figure 7.3: High ASPP2 gene expression is associated with poorer relapse-free survival in breast cancer patients with basal-like subtype. Kaplan-Meier (KM) survival analysis of ASPP2 expression in 255 breast cancer patients with (A) the triple-negative subtype and (B) the luminal subtype. Patients were split into two groups by KM survival curve based on their ASPP2 gene expression. The group with low ASPP2 expression (black) represents patients with the bottom 25% ASPP2 expression whereas the group with high ASPP2 expression (red) represents patients with the top 25% ASPP2 expression. Survival analysis was conducted using Kaplan-Meier Plotter (Györfy et al. 2010). HR: hazard ratio which measures the difference between the black and red curves in the above KM plots; logrankP: p-value from the log-rank test which measures the statistical significance of HR.

apoptotic function of p53 (Samuels-Lev et al. 2001). Another N-terminally truncated version of ASPP2, δ N-ASPP2, was found to be upregulated in human breast cancer tissues and displayed a dominant-negative phenotype by antagonizing stress-induced p53 functions and promoting cell proliferation (Van Hook et al. 2017). In addition, a C-terminally truncated splice variant of ASPP2 found in acute leukaemia samples, ASPP2 κ , lacks p53-binding sites and was also shown a dominant-negative effect on p53 functions. Interestingly, mammary-specific activation of the expression of a truncated version of TRP53BP2, encoded by *Trp53bp2* exons 13–18, in mice with a concomitant mammary-specific loss of E-cadherin resulted in mammary

tumors that resemble human invasive lobula carcinoma (Kas et al. 2017). It would therefore be important to analyze the alternative splicing in the TCGA pan-cancer cohort to understand the prevalence of ASPP2 splice variants and whether they are associated with the oncogenic pathways identified in the pan-cancer coexpression analysis in this study.

7.4 Conclusion

To conclude, this work demonstrates a novel role of ASPP2 as a regulator that modulates the expression of a subset of HIF target genes, potentially through regulating the crosstalk between HIF1 α and transcriptional cofactor YAP. Future study will try to elucidate the potential mechanisms, e.g. ROCK activity, that underlie the regulation of YAP by ASPP2, as well as the functional consequences of HIF-YAP crosstalk in cancers.

Our work also identifies hitherto unknown coexpressed pathways for ASPP2 at the pan-cancer level. This paves the road for future investigations that aim at uncovering new functions of ASPP2 in human cancers.

References

- Adamovich, Yaarit, Benjamin Ladeuix, Marina Golik, Maarten P. Koeners, and Gad Asher. 2017. "Rhythmic Oxygen Levels Reset Circadian Clocks through HIF1 α ." *Cell Metabolism* 25, no. 1 (January): 93–101.
- Adams, R. L P, and A. Rinaldi. 1987. "Effect of echinomycin on DNA methylation." *FEBS Letters* 215, no. 2 (May): 266–268.
- Ahn, Jinwoo, In Ja L. Byeon, Chang Hyeock Byeon, and Angela M. Gronenborn. 2009. "Insight into the structural basis of pro- and antiapoptotic p53 modulation by ASPP proteins." *Journal of Biological Chemistry* 284, no. 20 (May): 13812–13822.
- Akiri, Gal, Edmond Sabo, Hagit Dafni, Zehava Vadasz, Yelena Kartvelishvily, Noga Gan, Ofra Kessler, et al. 2003. "Lysyl oxidase-related protein-1 promotes tumor fibrosis and tumor progression in vivo." *Cancer research* 63, no. 7 (April): 1657–66.
<http://www.ncbi.nlm.nih.gov/pubmed/12670920>.
- Alam, Muhammad S. 2018. "Proximity Ligation Assay (PLA)." *Current Protocols in Immunology* 123, no. 1 (November): e58.
<http://doi.wiley.com/10.1002/cpim.58>.
- Allen, Benjamin L., and Dylan J. Taatjes. 2015. *The Mediator complex: A central integrator of transcription*, March.
- Altman, Brian J., Annie L. Hsieh, Arjun Sengupta, Saikumari Y. Krishnanaiah, Zachary E. Stine, Zandra E. Walton, Arvin M. Gouw, et al. 2015. "MYC Disrupts the Circadian Clock and Metabolism in Cancer Cells." *Cell Metabolism* 22 (6): 1009–1019.
- An, Won G., Meera Kanekal, M. Celeste Simon, Emin Maltepe, Mikhail V. Blagosklonny, and Leonard M. Neckers. 1998. "Stabilization of wild-type p53 by hypoxia-inducible factor 1 α ." *Nature* 392, no. 6674 (March): 405–408.
<http://www.ncbi.nlm.nih.gov/pubmed/9537326>
<http://www.nature.com/articles/32925>.
- Appelhoff, Rebecca J, Ya-Min Tian, Raju R Raval, Helen Turley, Adrian L Harris, Christopher W Pugh, Peter J Ratcliffe, and Jonathan M Gleadle. 2004. "Differential function of the prolyl hydroxylases PHD1, PHD2, and PHD3 in the regulation of hypoxia-inducible factor." *The Journal of biological chemistry* 279, no. 37 (September): 38458–65. <http://www.ncbi.nlm.nih.gov/pubmed/15247232>.

- Arany, Z., L. E. Huang, R. Eckner, S. Bhattacharya, C. Jiang, M. A. Goldberg, H. F. Bunn, and D. M. Livingston. 1996. "An essential role for p300/CBP in the cellular response to hypoxia." *Proceedings of the National Academy of Sciences* 93, no. 23 (November): 12969–12973.
<http://www.ncbi.nlm.nih.gov/pubmed/8917528>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC24030>
<http://www.pnas.org/cgi/doi/10.1073/pnas.93.23.12969>.
- Astuti, Dewi, Christopher J. Ricketts, Rasheduzzaman Chowdhury, Michael A. McDonough, Dean Gentle, Gail Kirby, Susanne Schlisio, et al. 2011. "Mutation analysis of HIF prolyl hydroxylases (PHD/EGLN) in individuals with features of pheochromocytoma and renal cell carcinoma susceptibility." *Endocrine-Related Cancer* 18, no. 1 (February): 73–83.
- Awad, Aline, Sokhavuth Sar, Ronan Barré, Clotilde Cariven, Mickael Marin, Jean Pierre Salles, Christophe Erneux, Didier Samuel, and Ama Gassama-Diagne. 2013. "FSHIP2 regulates epithelial cell polarity through its lipid product, which binds to Dlg1, a pathway subverted by hepatitis C virus core protein." *Molecular Biology of the Cell* 24, no. 14 (July): 2171–2185.
- Bagchi, Sonchita, Robert Fredriksson, and Åsa Wallén-Mackenzie. 2015. "In Situ Proximity Ligation Assay (PLA)," 149–159.
http://link.springer.com/10.1007/978-1-4939-2742-5_15.
- Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, and P. Horvath. 2007. "CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes." *Science* 315, no. 5819 (March): 1709–1712.
<http://www.ncbi.nlm.nih.gov/pubmed/17379808>
<http://www.sciencemag.org/cgi/doi/10.1126/science.1138140>.
- Bartkova, Jirina, Nousin Rezaei, Michalis Liontos, Panagiotis Karakaidos, Dimitris Kletsas, Natalia Issaeva, Leandros Vassilios F. Vassiliou, et al. 2006. "Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints." *Nature* 444 (7119): 633–637.
- Batie, Michael, Juliany Frost, Mark Frost, James W. Wilson, Pieta Schofield, and Sonia Rocha. 2019. "Hypoxia induces rapid changes to histone methylation and reprograms chromatin." *Science* 363 (6432): 1222–1226.
- Bellot, Grégory, Raquel Garcia-Medina, Pierre Gounon, Johanna Chiche, Danièle Roux, Jacques Pouyssegur, and Nathalie M Mazure. 2009. "Hypoxia-induced autophagy is mediated through hypoxia-inducible factor induction of BNIP3 and BNIP3L via their BH3 domains." *Molecular and cellular biology* 29, no. 10 (May): 2570–81.
<http://www.ncbi.nlm.nih.gov/pubmed/19273585>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2682037>.
- Bergamaschi, D., Y. Samuels, B. Jin, S. Duraisingham, T. Crook, and X. Lu. 2004. "ASPP1 and ASPP2: Common Activators of p53 Family Members." *Molecular and Cellular Biology* 24, no. 3 (February): 1341–1350.
- Bergamaschi, Daniele, Yardena Samuels, Nigel J. O'Neil, Giuseppe Trigiant, Tim Crook, Jung Kuang Hsieh, Daniel J. O'Connor, et al. 2003a. "iASPP oncoprotein is a key inhibitor of p53 conserved from worm to human." *Nature Genetics* 33, no. 2 (February): 162–167.

- Bergamaschi, Daniele, Yardena Samuels, Nigel J O'Neil, Giuseppe Trigiant, Tim Crook, Jung-Kuang Hsieh, Daniel J O'Connor, et al. 2003b. "iASPP oncoprotein is a key inhibitor of p53 conserved from worm to human." *Nature genetics* 33, no. 2 (February): 162–7. <http://www.ncbi.nlm.nih.gov/pubmed/12524540>.
- Bergamaschi, Daniele, Yardena Samuels, Alexandra Sullivan, Marketa Zvelebil, Hilde Breyssens, Andrea Bisso, Giannino Del Sal, et al. 2006. "iASPP preferentially binds p53 proline-rich region and modulates apoptotic function of codon 72-polymorphic p53." *Nature Genetics* 38, no. 10 (October): 1133–1141.
- Beroukhi, Rameen, Craig H. Mermel, Dale Porter, Guo Wei, Soumya Raychaudhuri, Jerry Donovan, Jordi Barretina, et al. 2010. "The landscape of somatic copy-number alteration across human cancers." *Nature* 463, no. 7283 (February): 899–905.
- Bertran, M. Teresa, Stéphane Mouilleron, Yanxiang Zhou, Rakhi Bajaj, Federico Uliana, Ganesan Senthil Kumar, Audrey van Drogen, et al. 2019. "ASPP proteins discriminate between PP1 catalytic subunits through their SH3 domain and the PP1 C-tail." *Nature Communications* 10, no. 1 (December).
- Biegging, Kathryn T., Stephano Spano Mello, and Laura D. Attardi. 2014. *Unravelling mechanisms of p53-mediated tumour suppression*.
- Blagosklonny, Mikhail V., Won G. An, Larisa Y. Romanova, Jane Trepel, Tito Fojo, and Len Neckers. 1998. "p53 inhibits hypoxia-inducible factor-stimulated transcription." *Journal of Biological Chemistry* 273, no. 20 (May): 11995–11998.
- Blau, J, H Xiao, S McCracken, P O'Hare, J Greenblatt, and D Bentley. 1996. "Three functional classes of transcriptional activation domain." *Molecular and cellular biology* 16, no. 5 (May): 2044–55.
<http://www.ncbi.nlm.nih.gov/pubmed/8628270>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC231191>.
- Bollen, Mathieu, Wolfgang Peti, Michael J. Ragusa, and Monique Beullens. 2010. *The extended PP1 toolkit: Designed to create specificity*, August.
- Bolotin, Alexander, Benoit Quinquis, Alexei Sorokin, and S Dusko Ehrlich. 2005. "Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin." *Microbiology* 151, no. 8 (August): 2551–2561.
<http://www.ncbi.nlm.nih.gov/pubmed/16079334>
<http://www.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.28048-0>.
- Bray, Freddie, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L. Siegel, Lindsey A. Torre, and Ahmedin Jemal. 2018. "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries." *CA: A Cancer Journal for Clinicians* 68, no. 6 (November): 394–424.
<http://doi.wiley.com/10.3322/caac.21492>.
- Bray, Freddie, Ahmedin Jemal, Nathan Grey, Jacques Ferlay, and David Forman. 2012. "Global cancer transitions according to the Human Development Index (2008-2030): a population-based study." *The Lancet. Oncology* 13, no. 8 (August): 790–801.
<http://www.ncbi.nlm.nih.gov/pubmed/22658655>.

- Bruick, R. K., and S L McKnight. 2001. "A Conserved Family of Prolyl-4-Hydroxylases That Modify HIF." *Science* 294, no. 5545 (November): 1337–1340.
<http://www.ncbi.nlm.nih.gov/pubmed/11598268>
<http://www.sciencemag.org/cgi/doi/10.1126/science.1066373>.
- Buday, László, and Julian Downward. 2008. "Many faces of Ras activation." *Biochimica et biophysica acta* 1786, no. 2 (December): 178–87.
<http://www.ncbi.nlm.nih.gov/pubmed/18541156>.
- Buhr, Ethan D., and Joseph S. Takahashi. 2013. "Molecular Components of the Mammalian Circadian Clock," 3–27.
- Bullard, James H, Elizabeth Purdom, Kasper D Hansen, and Sandrine Dudoit. 2010. "Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments." *BMC Bioinformatics* 11, no. 1 (December): 94.
<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-94>.
- Buti, Ludovico, Eric Spooner, Annemarthe G. Van Der Veen, Rino Rappuoli, Antonello Covacci, and Hidde L. Ploegh. 2011. "Helicobacter pylori cytotoxin-associated gene A (CagA) subverts the apoptosis-stimulating protein of p53 (ASPP2) tumor suppressor pathway of the host." *Proceedings of the National Academy of Sciences of the United States of America* 108, no. 22 (May): 9238–9243.
- Cai, Ling, Qiwei Li, Yi Du, Jonghyun Yun, Yang Xie, Ralph J. Deberardinis, and Guanghua Xiao. 2017. "Genomic regression analysis of coordinated expression." *Nature Communications* 8, no. 1 (December).
- Cai, Xinjiang, and Yanhong Zhang. 2006. "Molecular evolution of the ankyrin gene family." *Molecular biology and evolution* 23, no. 3 (March): 550–8.
<http://www.ncbi.nlm.nih.gov/pubmed/16280541>.
- Cao, Yongheng, Tomio Hamada, Tadashi Matsui, Takayasu Date, and Kuniyoshi Iwabuchi. 2004. "Hepatitis C virus core protein interacts with p53-binding protein, 53BP2/Bbp/ASPP2, and inhibits p53-mediated apoptosis." *Biochemical and Biophysical Research Communications* 315, no. 4 (March): 788–795.
- Carroll, Dana. 2011. "Genome engineering with zinc-finger nucleases." *Genetics* 188, no. 4 (August): 773–782.
- Cary, Lynne Csiszar, Michael Goebel, Bartholomew G. Corsaro, Hwei-Gene Wang, Elliot Rosen, and M.J. Fraser. 1989. "Transposon mutagenesis of baculoviruses: Analysis of Trichoplusia ni transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses." *Virology* 172, no. 1 (September): 156–169.
<http://www.ncbi.nlm.nih.gov/pubmed/2549707>
<https://linkinghub.elsevier.com/retrieve/pii/0042682289901177>.
- Cavadas, Miguel A. S., Marion Mesnieres, Bianca Crifo, Mario C. Manresa, Andrew C. Selfridge, Carsten C. Scholz, Eoin P. Cummins, Alex Cheong, and Cormac T. Taylor. 2015. "REST mediates resolution of HIF-dependent gene expression in prolonged hypoxia." *Scientific Reports* 5, no. 1 (December): 17851.
<http://www.nature.com/articles/srep17851>.

- Cerami, Ethan, Jianjiong Gao, Ugur Dogrusoz, Benjamin E. Gross, Selcuk Onur Sumer, Bülent Arman Aksoy, Anders Jacobsen, et al. 2012. "The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data." *Cancer Discovery* 2, no. 5 (May): 401–404.
- Chachami, Georgia, Efrosyni Paraskeva, José Manuel Mingot, Georgia G. Braliou, Dirk Görlich, and George Simos. 2009. "Transport of hypoxia-inducible factor HIF-1 α into the nucleus involves importins 4 and 7." *Biochemical and Biophysical Research Communications* 390, no. 2 (December): 235–240.
- Chakraborty, Abhishek A., Tuomas Laukka, Matti Myllykoski, Alison E. Ringel, Matthew A. Booker, Michael Y. Tolstorukov, Yuzhong Jeff Meng, et al. 2019. "Histone demethylase KDM6A directly senses oxygen to control chromatin and cell fate." *Science* 363 (6432): 1217–1222.
- Chan, H. M., and N. B. La Thangue. 2001. "p300/CBP proteins: HATs for transcriptional bridges and scaffolds." *Journal of Cell Science* 114 (13): 2363–2373.
- Chan, Mun Chiang, Nicholas E. Ilott, Johannes Schödel, David Sims, Anthony Tumber, Kerstin Lippl, David R. Mole, et al. 2016. "Tuning the Transcriptional Response to Hypoxia by Inhibiting Hypoxia-inducible Factor (HIF) Prolyl and Asparaginyl Hydroxylases." *Journal of Biological Chemistry* 291, no. 39 (September): 20661–20673. <http://www.jbc.org/content/291/39/20661.full#ref-43>.
- Chen, Changhu, Nabendu Pore, Alireza Behrooz, Faramarz Ismail-Beigi, and Amit Maity. 2001. "Regulation of *glut1* mRNA by Hypoxia-inducible Factor-1." *Journal of Biological Chemistry* 276, no. 12 (March): 9519–9525. <http://www.ncbi.nlm.nih.gov/pubmed/11120745>
<http://www.jbc.org/lookup/doi/10.1074/jbc.M010144200>.
- Chen, Jianfang, Zhenyu Ding, Yonghai Peng, Feng Pan, Jianjun Li, Lan Zou, Yanling Zhang, and Houjie Liang. 2014. "HIF-1 α inhibition reverses multidrug resistance in colon cancer cells via downregulation of MDR1/P-Glycoprotein." *PLoS ONE* 9, no. 6 (June).
- Chen, Shuo, Jiale Wu, Shan Zhong, Yuntong Li, Ping Zhang, Jingyi Ma, Jingshan Ren, et al. 2019. "IASPP mediates p53 selectivity through a modular mechanism fine-tuning DNA recognition." *Proceedings of the National Academy of Sciences of the United States of America* 116, no. 35 (August): 17470–17479.
- Cho, K R, and B Vogelstein. 1992. "Genetic alterations in the adenoma–carcinoma sequence." *Cancer* 70, no. 6 Suppl (September): 1727–31. <http://www.ncbi.nlm.nih.gov/pubmed/1516027>.
- Choy, Mun-Kit, Mehregan Movassagh, Martin R. Bennett, and Roger S.-Y. Foo. 2009. "PKB/Akt activation inhibits p53-mediated HIF1A degradation that is independent of MDM2." *Journal of Cellular Physiology*: n/a–n/a. <http://doi.wiley.com/10.1002/jcp.21980>.

- Cockman, Matthew E., James D. Webb, Holger B. Kramer, Benedikt M. Kessler, and Peter J. Ratcliffe. 2009. "Proteomics-based Identification of Novel Factor Inhibiting Hypoxia-inducible Factor (FIH) Substrates Indicates Widespread Asparaginyl Hydroxylation of Ankyrin Repeat Domain-containing Proteins." *Molecular & Cellular Proteomics* 8, no. 3 (March): 535–546.
<http://www.ncbi.nlm.nih.gov/pubmed/18936059><http://www.ncbi.nlm.nih.gov/pubmedcentral/nih.gov/articlerender.fcgi?artid=PMC2649815><http://www.mcponline.org/lookup/doi/10.1074/mcp.M800340-MCP200>.
- Cockman, Matthew E., James D. Webb, and Peter J. Ratcliffe. 2009. "FIH-dependent asparaginyl hydroxylation of ankyrin repeat domain-containing proteins." In *Annals of the New York Academy of Sciences*, 1177:9–18. Blackwell Publishing Inc.
- Cockman, Matthew E., David E Lancaster, Ineke P Stolze, Kirsty S Hewitson, Michael A McDonough, Mathew L Coleman, Charlotte H Coles, et al. 2006. "Posttranslational hydroxylation of ankyrin repeats in IkappaB proteins by the hypoxia-inducible factor (HIF) asparaginyl hydroxylase, factor inhibiting HIF (FIH)." *Proceedings of the National Academy of Sciences of the United States of America* 103, no. 40 (October): 14767–72.
<http://www.ncbi.nlm.nih.gov/pubmed/17003112><http://www.ncbi.nlm.nih.gov/pubmedcentral/nih.gov/articlerender.fcgi?artid=PMC1578504>.
- Coleman, Mathew L., Michael A. McDonough, Kirsty S. Hewitson, Charlotte Coles, Jasmin Mecinović, Mariola Edelmann, Kristina M. Cook, et al. 2007. "Asparaginyl Hydroxylation of the Notch Ankyrin Repeat Domain by Factor Inhibiting Hypoxia-inducible Factor." *Journal of Biological Chemistry* 282, no. 33 (August): 24027–24038. <http://www.jbc.org/content/282/33/24027.long>.
- Cong, Le, F Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D Hsu, et al. 2013. "Multiplex genome engineering using CRISPR/Cas systems." *Science (New York, N.Y.)* 339, no. 6121 (February): 819–23.
<http://www.ncbi.nlm.nih.gov/pubmed/23287718><http://www.ncbi.nlm.nih.gov/pubmedcentral/nih.gov/articlerender.fcgi?artid=PMC3795411>.
- Cong, Weili, Tomonori Hirose, Yutaka Harita, Akio Yamashita, Keiko Mizuno, Hisashi Hirano, and Shigeo Ohno. 2010. "ASPP2 Regulates Epithelial Cell Polarity through the PAR Complex." *Current Biology* 20, no. 15 (August): 1408–1414.
<http://www.ncbi.nlm.nih.gov/pubmed/20619648><https://linkinghub.elsevier.com/retrieve/pii/S0960982210007694>.
- Corn, Paul G., M. Stacey Ricci, Kimberly A. Scata, Andrew M. Arsham, M. Celeste Simon, David T. Dicker, and Wafik S. El-Deiry. 2005. "Mxi1 is induced by hypoxia in a HIF-1-dependent manner and protects cells from c-Myc-induced apoptosis." *Cancer Biology & Therapy* 4, no. 11 (November): 1285–1294.
<http://www.ncbi.nlm.nih.gov/pubmed/16319523><http://www.tandfonline.com/doi/abs/10.4161/cbt.4.11.2299>.
- Cradick, Thomas J., Eli J. Fine, Christopher J. Antico, and Gang Bao. 2013. "CRISPR/Cas9 systems targeting β -globin and CCR5 genes have substantial off-target activity." *Nucleic Acids Research* 41, no. 20 (November): 9584–9592.
<http://www.ncbi.nlm.nih.gov/pubmed/23939622><http://www.ncbi.nlm.nih.gov/pubmedcentral/nih.gov/articlerender.fcgi?artid=PMC3814385><https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt714>.

- Croft, David, Gavin O'Kelly, Guanming Wu, Robin Haw, Marc Gillespie, Lisa Matthews, Michael Caudy, et al. 2011. "Reactome: a database of reactions, pathways and biological processes." *Nucleic acids research* 39, no. Database issue (January): 691–7. <http://www.ncbi.nlm.nih.gov/pubmed/21067998><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3013646>.
- Dames, Sonja A., Maria Martinez-Yamout, Roberto N. De Guzman, H. Jane Dyson, and Peter E. Wright. 2002. "Structural basis for Hif-1 α /CBP recognition in the cellular hypoxic response." *Proceedings of the National Academy of Sciences of the United States of America* 99, no. 8 (April): 5271–5276.
- Deltcheva, Elitza, Krzysztof Chylinski, Cynthia M. Sharma, Karine Gonzales, Yanjie Chao, Zaid A. Pirzada, Maria R. Eckert, Jörg Vogel, and Emmanuelle Charpentier. 2011. "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." *Nature* 471, no. 7340 (March): 602–607. <http://www.ncbi.nlm.nih.gov/pubmed/21455174><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3070239><http://www.nature.com/articles/nature09886>.
- Denison, Michael S., and Scott R. Nagy. 2003. "ACTIVATION OF THE ARYL HYDROCARBON RECEPTOR BY STRUCTURALLY DIVERSE EXOGENOUS AND ENDOGENOUS CHEMICALS." *Annual Review of Pharmacology and Toxicology* 43, no. 1 (April): 309–334. <http://www.annualreviews.org/doi/10.1146/annurev.pharmtox.43.100901.135828>.
- Desgrosellier, Jay S., and David A. Chersesh. 2010. *Integrins in cancer: Biological implications and therapeutic opportunities*, January.
- Ding, Sheng, Xiaohui Wu, Gang Li, Min Han, Yuan Zhuang, and Tian Xu. 2005. "Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice." *Cell* 122, no. 3 (August): 473–83. <http://www.ncbi.nlm.nih.gov/pubmed/16096065>.
- Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: ultrafast universal RNA-seq aligner." *Bioinformatics* 29, no. 1 (January): 15–21. <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bts635>.
- Doench, John G, Nicolo Fusi, Meagan Sullender, Mudra Hegde, Emma W Vaimberg, Katherine F Donovan, Ian Smith, et al. 2016. "Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9." *Nature Biotechnology* 34, no. 2 (February): 184–191. <http://www.nature.com/articles/nbt.3437>.
- Donner, Aaron Joseph, Stephanie Szostek, Jennifer Michelle Hoover, and Joaquin Maximiliano Espinosa. 2007. "CDK8 is a stimulus-specific positive coregulator of p53 target genes." *Molecular cell* 27, no. 1 (July): 121–33. <http://www.ncbi.nlm.nih.gov/pubmed/17612495><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2936241>.
- Downward, Julian. 2003. *Targeting RAS signalling pathways in cancer therapy*, January.

- Drutel, Guillaume, Markus Kathmann, Anne Heron, Jean Charles Schwartz, and Jean Michel Arrang. 1996. "Cloning and selective expression in brain and kidney of ARNT2 homologous to the Ah receptor nuclear translocator (ARNT)." *Biochemical and Biophysical Research Communications* 225, no. 2 (August): 333–339.
- Egloff, M P, D F Johnson, G Moorhead, P T Cohen, P Cohen, and D Barford. 1997. "Structural basis for the recognition of regulatory subunits by the catalytic subunit of protein phosphatase 1." *The EMBO journal* 16, no. 8 (April): 1876–87. <http://www.ncbi.nlm.nih.gov/pubmed/9155014><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1169791>.
- Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein. 1998. "Cluster analysis and display of genome-wide expression patterns." *Proceedings of the National Academy of Sciences* 95, no. 25 (December): 14863–14868. <http://www.pnas.org/cgi/doi/10.1073/pnas.95.25.14863>.
- Ema, Masatsugu, Shinichiro Taya, Noboru Yokotani, Kazuhiro Sogawa, Youichi Matsuda, and Yoshiaki Fujii-Kuriyama. 1997. "A novel bHLH-PAS factor with close sequence similarity to hypoxia-inducible factor 1 α regulates the VEGF expression and is potentially involved in lung and vascular development." *Proceedings of the National Academy of Sciences of the United States of America* 94, no. 9 (April): 4273–4278.
- Epstein, A C, J M Gleadle, L A McNeill, K S Hewitson, J O'Rourke, D R Mole, M Mukherji, et al. 2001. "C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation." *Cell* 107, no. 1 (October): 43–54. <http://www.ncbi.nlm.nih.gov/pubmed/11595184>.
- Fearon, E R, and B Vogelstein. 1990. "A genetic model for colorectal tumorigenesis." *Cell* 61, no. 5 (June): 759–67. <http://www.ncbi.nlm.nih.gov/pubmed/2188735>.
- Feng, Xiaodong, Maria Sol Degese, Ramiro Iglesias-Bartolome, Jose P. Vaque, Alfredo A. Molinolo, Murilo Rodrigues, M. Raza Zaidi, et al. 2014. "Hippo-independent activation of YAP by the GNAQ uveal melanoma oncogene through a Trio-regulated Rho GTPase Signaling Circuitry." *Cancer Cell* 25, no. 6 (June): 831–845.
- Ferbeyre, G., E. de Stanchina, A. W. Lin, E. Querido, M. E. McCurrach, G. J. Hannon, and S. W. Lowe. 2002. "Oncogenic ras and p53 Cooperate To Induce Cellular Senescence." *Molecular and Cellular Biology* 22, no. 10 (May): 3497–3508.
- Firestein, Ron, Adam J Bass, So Young Kim, Ian F Dunn, Serena J Silver, Isil Guney, Ellen Freed, et al. 2008. "CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity." *Nature* 455, no. 7212 (September): 547–51. <http://www.ncbi.nlm.nih.gov/pubmed/18794900><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2587138>.
- Fraser, M J, T Ciszczon, T Elick, and C Bauser. 1996. "Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera." *Insect molecular biology* 5, no. 2 (May): 141–51. <http://www.ncbi.nlm.nih.gov/pubmed/8673264>.

- Fraser, Malcolm J., Lynne Cary, Kitima Boonvisudhi, and Hwei-Gene Heidi Wang. 1995. "Assay for Movement of Lepidopteran Transposon IFP2 in Insect Cells Using a Baculovirus Genome as a Target DNA." *Virology* 211, no. 2 (August): 397–407. <http://www.ncbi.nlm.nih.gov/pubmed/7645244><https://linkinghub.elsevier.com/retrieve/pii/S0042682285714225>.
- Freedman, S. J., Z.-Y. J. Sun, F. Poy, A. L. Kung, D. M. Livingston, G. Wagner, and M. J. Eck. 2002. "Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1." *Proceedings of the National Academy of Sciences* 99, no. 8 (April): 5367–5372. <http://www.ncbi.nlm.nih.gov/pubmed/11959990><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC122775><http://www.pnas.org/cgi/doi/10.1073/pnas.082117899>.
- Freedman, Steven J., Zhen Yu J. Sun, Andrew L. Kung, Dennis S. France, Gerhard Wagner, and Michael J. Eck. 2003. "Structural basis for negative regulation of hypoxia-inducible factor-1 α by CITED2." *Nature Structural Biology* 10, no. 7 (July): 504–512.
- Fromont-Racine, M, A E Mayes, A Brunet-Simon, J C Rain, A Colley, I Dix, L Decourty, et al. 2000. "Genome-wide protein interaction screens reveal functional networks involving Sm-like proteins." *Yeast (Chichester, England)* 17, no. 2 (June): 95–110. <http://www.ncbi.nlm.nih.gov/pubmed/10900456><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2448332>.
- Fu, Yanfang, Jennifer A Foden, Cyd Khayter, Morgan L Maeder, Deepak Reyon, J Keith Joung, and Jeffry D Sander. 2013. "High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells." *Nature Biotechnology* 31, no. 9 (September): 822–826. <http://www.ncbi.nlm.nih.gov/pubmed/23792628><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3773023><http://www.nature.com/articles/nbt.2623>.
- Fu, Yanfang, Jeffry D Sander, Deepak Reyon, Vincent M Cascio, and J Keith Joung. 2014. "Improving CRISPR-Cas nuclease specificity using truncated guide RNAs." *Nature biotechnology* 32, no. 3 (March): 279–284. <http://www.ncbi.nlm.nih.gov/pubmed/24463574><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3988262>.
- Galbraith, Matthew D., Mary A. Allen, Claire L. Bensard, Xiaoxing Wang, Marie K. Schwinn, Bo Qin, Henry W. Long, et al. 2013. "XHIF1A employs CDK8-mediator to stimulate RNAPII elongation in response to hypoxia." *Cell* 153, no. 6 (June): 1327.
- Gao, Galen F, Joel S Parker, Sheila M Reynolds, and The Genomic. 2019. "Before and After: Comparison of Legacy and Harmonized TCGA Genomic Data Commons' Data." <https://doi.org/10.1016/j.cels.2019.06.006>.
- Gao, Jianjiong, Bülent Arman Aksoy, Ugur Dogrusoz, Gideon Dresdner, Benjamin Gross, S. Onur Sumer, Yichao Sun, et al. 2013. "Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal." *Science Signaling* 6, no. 269 (April).

- Garneau, Josiane E., Marie-Ève Dupuis, Manuela Villion, Dennis A. Romero, Rodolphe Barrangou, Patrick Boyaval, Christophe Fremaux, Philippe Horvath, Alfonso H. Magadán, and Sylvain Moineau. 2010. "The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA." *Nature* 468, no. 7320 (November): 67–71. <http://www.nature.com/articles/nature09523>.
- Gasiunas, G., R. Barrangou, P. Horvath, and V. Siksnys. 2012. "Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria." *Proceedings of the National Academy of Sciences* 109, no. 39 (September): E2579–E2586. <http://www.ncbi.nlm.nih.gov/pubmed/22949671><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3465414><http://www.pnas.org/cgi/doi/10.1073/pnas.1208507109>.
- Genevet, Alice, and Nicolas Tapon. 2011. "The Hippo pathway and apico–basal cell polarity." *Biochemical Journal* 436, no. 2 (June): 213–224. <https://portlandpress.com/biochemj/article/436/2/213/88306/The-Hippo-pathway-and-apicobasal-cell-polarity>.
- Giatromanolaki, A., M. I. Koukourakis, E. Sivridis, H. Turley, K. Talks, F. Pezzella, K. C. Gatter, and A. L. Harris. 2001. "Relation of hypoxia inducible factor 1 α and 2 α in operable non-small cell lung cancer to angiogenic/molecular profile of tumours and survival." *British Journal of Cancer* 85, no. 6 (September): 881–890.
- Gillies, Robert J., Paul A. Schornack, Timothy W. Secomb, and Natarajan Raghunand. 1999. *Causes and Effects of Heterogeneous Perfusion in Tumors*, August.
- Gillot, Sébastien, and Xin Lu. 2011. "The ASPP proteins complex and cooperate with p300 to modulate the transcriptional activity of p53." *FEBS Letters* 585, no. 12 (June): 1778–1782. <https://www.sciencedirect.com/science/article/pii/S0014579311002596>.
- Givant-Horwitz, Vered, Ben Davidson, and Reuven Reich. 2004. "Laminin-induced signaling in tumor cells: The role of the Mr 67,000 laminin receptor." *Cancer Research* 64, no. 10 (May): 3572–3579.
- Goldman, Mary, Brian Craft, Mim Hastie, Kristupas Repečka, Akhil Kamath, Fran McDade, Dave Rogers, Angela N. Brooks, Jingchun Zhu, and David Haussler. 2019. "The UCSC Xena platform for public and private cancer genomics data visualization and interpretation." *bioRxiv*: 326470. <https://www.biorxiv.org/content/10.1101/326470v5>.
- Gorina, Svetlana, and Nikola P. Pavletich. 1996. "Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2." *Science* 274, no. 5289 (November): 1001–1005.
- Gossen, M., S. Freundlieb, G. Bender, G. Müller, W. Hillen, and H. Bujard. 1995. "Transcriptional activation by tetracyclines in mammalian cells." *Science* 268, no. 5218 (June): 1766–1769. <http://www.ncbi.nlm.nih.gov/pubmed/7792603><http://www.sciencemag.org/cgi/doi/10.1126/science.7792603>.
- Greenhough, Alexander, Clare Bagley, Kate J Heesom, David B Gurevich, David Gay, Mark Bond, Tracey J Collard, et al. 2018. "Cancer cell adaptation to hypoxia involves a HIF-GPRC5A-YAP axis." *EMBO Molecular Medicine* 10, no. 11 (November). <https://onlinelibrary.wiley.com/doi/abs/10.15252/emmm.201708699>.

- Griffiths, E. A., S. A. Pritchard, H. R. Valentine, N. Whitchelo, P. W. Bishop, M. P. Ebert, P. M. Price, I. M. Welch, and C. M.L. West. 2007. "Hypoxia-inducible factor-1 α expression in the gastric carcinogenesis sequence and its prognostic role in gastric and gastro-oesophageal adenocarcinomas." *British Journal of Cancer* 96, no. 1 (January): 95–103.
- Gu, Y Z, S M Moran, J B Hogenesch, L Wartman, and C A Bradfield. 1998. "Molecular characterization and chromosomal localization of a third alpha-class hypoxia inducible factor subunit, HIF3alpha." *Gene expression* 7 (3): 205–13. <http://www.ncbi.nlm.nih.gov/pubmed/9840812><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6151950>.
- Guzy, Robert D., Beatrice Hoyos, Emmanuel Robin, Hong Chen, Liping Liu, Kyle D. Mansfield, M. Celeste Simon, Ulrich Hammerling, and Paul T. Schumacker. 2005. "Mitochondrial complex III is required for hypoxia-induced ROS production and cellular oxygen sensing." *Cell Metabolism* 1, no. 6 (June): 401–408.
- Györfy, Balazs, Andras Lanczky, Aron C. Eklund, Carsten Denkert, Jan Budczies, Qiyuan Li, and Zoltan Szallasi. 2010. *An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients*, October.
- Haeussler, Maximilian, Kai Schönig, Hélène Eckert, Alexis Eschstruth, Joffrey Mianné, Jean-Baptiste Renaud, Sylvie Schneider-Maunoury, et al. 2016. "Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR." *Genome Biology* 17, no. 1 (December): 148. <http://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1012-2>.
- Haga, Raquel B., and Anne J. Ridley. 2016. *Rho GTPases: Regulation and roles in cancer cell biology*, October.
- Hahn, William C., Christopher M. Counter, Ante S. Lundberg, Roderick L. Beijersbergen, Mary W. Brooks, and Robert A. Weinberg. 1999. "Creation of human tumour cells with defined genetic elements." *Nature* 400, no. 6743 (July): 464–468. <http://www.ncbi.nlm.nih.gov/pubmed/10440377><http://www.nature.com/articles/22780>.
- Hammond, E. M., N. C. Denko, M. J. Dorie, R. T. Abraham, and A. J. Giaccia. 2002. "Hypoxia Links ATR and p53 through Replication Arrest." *Molecular and Cellular Biology* 22, no. 6 (March): 1834–1843.
- Hanahan, Douglas, and Robert A Weinberg. 2000. "The Hallmarks of Cancer." *Cell* 100, no. 1 (January): 57–70. [https://www.cell.com/cell/fulltext/S0092-8674\(00\)81683-9?cc=y](https://www.cell.com/cell/fulltext/S0092-8674(00)81683-9?cc=y).
- . 2011. "Hallmarks of Cancer: The Next Generation." *Cell* 144, no. 5 (March): 646–674. <http://www.ncbi.nlm.nih.gov/pubmed/21376230><https://linkinghub.elsevier.com/retrieve/pii/S0092867411001279>.
- Hänzelmann, Sonja, Robert Castelo, and Justin Guinney. 2013. "GSVA: Gene set variation analysis for microarray and RNA-Seq data." *BMC Bioinformatics* 14, no. 1 (January): 7. <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-14-7>.

- Hayashi, Masami, Masahiro Sakata, Takashi Takeda, Toshiya Yamamoto, Yoko Okamoto, Kenjiro Sawada, Akiko Kimura, et al. 2004. "Induction of glucose transporter 1 expression through hypoxia-inducible factor 1 α under hypoxic conditions in trophoblast-derived cells." *Journal of Endocrinology* 183, no. 1 (October): 145–154. <http://www.ncbi.nlm.nih.gov/pubmed/15525582><https://joe.bioscientifica.com/view/journals/joe/183/1/1830145.xml>.
- Hazan, Idit, Jonathan Monin, Britta A.M. Bouwman, Nicola Crosetto, and Rami I. Aqeilan. 2019. "Activation of Oncogenic Super-Enhancers Is Coupled with DNA Repair by RAD51." *Cell Reports* 29, no. 3 (October): 560–572.
- Heinz, Sven, Christopher Benner, Nathanael Spann, Eric Bertolino, Yin C. Lin, Peter Laslo, Jason X. Cheng, Cornelis Murre, Harinder Singh, and Christopher K. Glass. 2010. "Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities." *Molecular Cell* 38, no. 4 (May): 576–589. <https://linkinghub.elsevier.com/retrieve/pii/S1097276510003667>.
- Helmlinger, Gabriel, Fan Yuan, Marc Dellian, and Rakesh K. Jain. 1997. "Interstitial pH and pO₂ gradients in solid tumors in vivo: High-resolution measurements reveal a lack of correlation." *Nature Medicine* 3 (2): 177–182.
- Helps, Nicholas R., Hazel M. Barker, Stephen J. Elledge, and Patricia T.W. Cohen. 1995. "Protein phosphatase 1 interacts with p53BP2, a protein which binds to the tumour suppressor p53." *FEBS Letters* 377, no. 3 (December): 295–300.
- Helps, Nicholas R., Xinmei Luo, Hazel M Barker, and Patricia T W Cohen. 2000. *NIMA-related kinase 2 (Nek2), a cell-cycle-regulated protein kinase localized to centrosomes, is complexed to protein phosphatase 1*. Technical report.
- Hengartner, C J, V E Myer, S M Liao, C J Wilson, S S Koh, and R A Young. 1998. "Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases." *Molecular cell* 2, no. 1 (July): 43–53. <http://www.ncbi.nlm.nih.gov/pubmed/9702190>.
- Heroes, Ewald, Bart Lesage, Janina Görnemann, Monique Beullens, Luc Van Meervelt, and Mathieu Bollen. 2013. "The PP1 binding code: a molecular-lego strategy that governs specificity." *FEBS Journal* 280, no. 2 (January): 584–595. <http://doi.wiley.com/10.1111/j.1742-4658.2012.08547.x>.
- Hirose, K, M Morita, M Ema, J Mimura, H Hamada, H Fujii, Y Saijo, O Gotoh, K Sogawa, and Y Fujii-Kuriyama. 1996. "cDNA cloning and tissue-specific expression of a novel basic helix-loop-helix/PAS factor (Arnt2) with close sequence similarity to the aryl hydrocarbon receptor nuclear translocator (Arnt)." *Molecular and cellular biology* 16, no. 4 (April): 1706–13. <http://www.ncbi.nlm.nih.gov/pubmed/8657146><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC231157>.
- Hirsilä, Maija, Peppi Koivunen, Volkmar Günzler, Kari I Kivirikko, and Johanna Myllyharju. 2003. "Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor." *The Journal of biological chemistry* 278, no. 33 (August): 30772–80. <http://www.ncbi.nlm.nih.gov/pubmed/12788921>.

- Hoadley, Katherine A., Christina Yau, Toshinori Hinoue, Denise M. Wolf, Alexander J. Lazar, Esther Drill, Ronglai Shen, et al. 2018. "Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer." *Cell* 173, no. 2 (April): 291–304.
- Holmquist-Mengelbier, Linda, Erik Fredlund, Tobias Löfstedt, Rosa Noguera, Samuel Navarro, Helén Nilsson, Alexander Pietras, et al. 2006. "Recruitment of HIF-1 α and HIF-2 α to common target genes is differentially regulated in neuroblastoma: HIF-2 α promotes an aggressive phenotype." *Cancer cell* 10, no. 5 (November): 413–23. <http://www.ncbi.nlm.nih.gov/pubmed/17097563>.
- Honaker, James, Gary King, and Matthew Blackwell. 2011. "Amelia II: A program for missing data." *Journal of Statistical Software* 45 (7): 1–47.
- Horak, P., A. R. Crawford, D. D. Vadysirisack, Z. M. Nash, M. P. DeYoung, D. Sgroi, and L. W. Ellisen. 2010. "Negative feedback control of HIF-1 through REDD1-regulated ROS suppresses tumorigenesis." *Proceedings of the National Academy of Sciences* 107, no. 10 (March): 4675–4680. <http://www.ncbi.nlm.nih.gov/pubmed/20176937><http://www.ncbi.nlm.nih.gov/pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2842042><http://www.pnas.org/cgi/doi/10.1073/pnas.0907705107>.
- Hsu, Patrick D, David A Scott, Joshua A Weinstein, F Ann Ran, Silvana Konermann, Vineeta Agarwala, Yinqing Li, et al. 2013. "DNA targeting specificity of RNA-guided Cas9 nucleases." *Nature Biotechnology* 31, no. 9 (September): 827–832. <http://www.nature.com/articles/nbt.2647>.
- Hu, C.-J., L.-Y. Wang, L. A. Chodosh, B. Keith, and M. C. Simon. 2003. "Differential Roles of Hypoxia-Inducible Factor 1 (HIF-1) and HIF-2 in Hypoxic Gene Regulation." *Molecular and Cellular Biology* 23, no. 24 (December): 9361–9374.
- Hu, Cheng-Jun, Aneesa Sataur, Liyi Wang, Hongqing Chen, and M. Celeste Simon. 2007. "The N-Terminal Transactivation Domain Confers Target Gene Specificity of Hypoxia-inducible Factors HIF-1 α and HIF-2 α ." *Molecular Biology of the Cell* 18, no. 11 (November): 4528. <http://www.ncbi.nlm.nih.gov/pubmed/17804822><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2043574>.
- Huang, L. E., J. Gu, M. Schau, and H. F. Bunn. 1998. "Regulation of hypoxia-inducible factor 1 is mediated by an O₂-dependent degradation domain via the ubiquitin-proteasome pathway." *Proceedings of the National Academy of Sciences* 95, no. 14 (July): 7987–7992. <http://www.ncbi.nlm.nih.gov/pubmed/9653127><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC20916><http://www.pnas.org/cgi/doi/10.1073/pnas.95.14.7987>.
- Hughes, Timothy R., Matthew J. Marton, Allan R. Jones, Christopher J. Roberts, Roland Stoughton, Christopher D. Armour, Holly A. Bennett, et al. 2000. "Functional discovery via a compendium of expression profiles." *Cell* 102, no. 1 (July): 109–126.
- Imamura, Takaaki, Hirotoshi Kikuchi, Maria-Teresa Herraiz, Do-Youn Park, Yusuke Mizukami, Mari Mino-Kenduson, Maureen P. Lynch, et al. 2009. "HIF-1 α and HIF-2 α have divergent roles in colon cancer." *International Journal of Cancer* 124, no. 4 (February): 763–771. <http://doi.wiley.com/10.1002/ijc.24032>.

- Ito, Takashi, Tomoko Chiba, Ritsuko Ozawa, Mikio Yoshida, Masahira Hattori, and Yoshiyuki Sakaki. 2001. "A comprehensive two-hybrid analysis to explore the yeast protein interactome." *Proceedings of the National Academy of Sciences of the United States of America* 98, no. 8 (April): 4569–4574.
- Ivan, M., T. Haberberger, D. C. Gervasi, K. S. Michelson, V. Gunzler, K. Kondo, H. Yang, et al. 2002. "Biochemical purification and pharmacological inhibition of a mammalian prolyl hydroxylase acting on hypoxia-inducible factor." *Proceedings of the National Academy of Sciences* 99, no. 21 (October): 13459–13464. <http://www.pnas.org/cgi/doi/10.1073/pnas.192342099>.
- Iwabuchi, K, P L Bartel, B Li, R Marraccino, and S Fields. 1994. "Two cellular proteins that bind to wild-type but not mutant p53." *Proceedings of the National Academy of Sciences of the United States of America* 91, no. 13 (June): 6098–102. <http://www.ncbi.nlm.nih.gov/pubmed/8016121><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC44145>.
- Iwabuchi, Kuniyoshi, Bin Li, Hillary F. Massa, Barbara J. Trask, Takayasu Date, and Stanley Fields. 1998. "Stimulation of p53-mediated transcriptional activation by the p53-binding proteins, 53BP1 and 53BP2." *Journal of Biological Chemistry* 273, no. 40 (October): 26061–26068.
- Iyer, N V, L E Kotch, F Agani, S W Leung, E Laughner, R H Wenger, M Gassmann, et al. 1998. "Cellular and developmental control of O₂ homeostasis by hypoxia-inducible factor 1 alpha." *Genes & development* 12, no. 2 (January): 149–62. <http://www.ncbi.nlm.nih.gov/pubmed/9436976><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC316445>.
- Iyer, N V, S W Leung, and G L Semenza. 1998. "The human hypoxia-inducible factor 1alpha gene: HIF1A structure and evolutionary conservation." *Genomics* 52, no. 2 (September): 159–65. <http://www.ncbi.nlm.nih.gov/pubmed/9782081>.
- Jaakkola, P., D. R. Mole, Y.-M. Tian, M. I. Wilson, J. Gielbert, S. J. Gaskell, A. v. Kriegsheim, et al. 2001. "Targeting of HIF-alpha to the von Hippel-Lindau Ubiquitylation Complex by O₂-Regulated Prolyl Hydroxylation." *Science* 292, no. 5516 (April): 468–472. <http://www.ncbi.nlm.nih.gov/pubmed/11292861><http://www.sciencemag.org/cgi/doi/10.1126/science.1059796>.
- Janke, Kirsten, Ulf Brockmeier, Katja Kuhlmann, Martin Eisenacher, Jan Nolde, Helmut E Meyer, Heimo Mairbäurl, and Eric Metzen. 2013. "Factor inhibiting HIF-1 (FIH-1) modulates protein interactions of apoptosis-stimulating p53 binding protein 2 (ASPP2)." *Journal of cell science* 126, no. Pt 12 (June): 2629–40. <http://www.ncbi.nlm.nih.gov/pubmed/23606740>.
- Jia, Yu, Hui-Yan Li, Jue Wang, Ying Wang, Peng Zhang, Ning Ma, and Shi-Jing Mo. 2019. "Phosphorylation of 14-3-3ζ links YAP transcriptional activation to hypoxic glycolysis for tumorigenesis." *Oncogenesis* 8, no. 5 (May): 31. <http://www.nature.com/articles/s41389-019-0143-1>.

- Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A Doudna, and Emmanuelle Charpentier. 2012. "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." *Science (New York, N.Y.)* 337, no. 6096 (August): 816–21.
<http://www.ncbi.nlm.nih.gov/pubmed/22745249>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6286148>.
- Johnson, W Evan, Cheng Li, and Ariel Rabinovic. 2007. "Adjusting batch effects in microarray expression data using empirical Bayes methods." *Biostatistics (Oxford, England)* 8, no. 1 (January): 118–27.
<http://www.ncbi.nlm.nih.gov/pubmed/16632515>.
- Joung, J. Keith, and Jeffry D. Sander. 2013. *TALENs: A widely applicable technology for targeted genome editing*, January.
- Kaelin, William G., and Peter J. Ratcliffe. 2008. *Oxygen Sensing by Metazoans: The Central Role of the HIF Hydroxylase Pathway*, May.
- Kallio, P J, K Okamoto, S O'Brien, P Carrero, Y Makino, H Tanaka, and L Poellinger. 1998. "Signal transduction in hypoxic cells: inducible nuclear translocation and recruitment of the CBP/p300 coactivator by the hypoxia-inducible factor-1alpha." *The EMBO journal* 17, no. 22 (November): 6573–86.
<http://www.ncbi.nlm.nih.gov/pubmed/9822602>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1171004>.
- Kaluzová, Milota, Stefan Kaluz, Michael I Lerman, and Eric J Stanbridge. 2004. "DNA damage is a prerequisite for p53-mediated proteasomal degradation of HIF-1alpha in hypoxic cells and downregulation of the hypoxia marker carbonic anhydrase IX." *Molecular and cellular biology* 24, no. 13 (July): 5757–66.
<http://www.ncbi.nlm.nih.gov/pubmed/15199132>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC480909>.
- Kampa-Schittenhelm, Kerstin M, Barbara Illing, Figen Akmut, Michael Walter, Charles D Lopez, and Marcus M Schittenhelm. 2013. "Identification Of C-Terminal Truncated Splice Variants Of The Apoptosis-Stimulating Protein Of p53-2 (ASPP2) In Acute Leukemia Lacking The p53-Binding Site." *Blood* 122, no. 21 (November): 1270–1270.
- Kampa, Kerstin M, Jared D Acoba, Dexi Chen, Joel Gay, Hunjoo Lee, Kelly Beemer, Emerson Padiernos, et al. 2009. *Apoptosis-stimulating protein of p53 (ASPP2) heterozygous mice are tumor-prone and have attenuated cellular damage-response thresholds*. Technical report 11.
- Kandoth, Cyriac, Michael D. McLellan, Fabio Vandin, Kai Ye, Beifang Niu, Charles Lu, Mingchao Xie, et al. 2013. "Mutational landscape and significance across 12 major cancer types." *Nature* 502 (7471): 333–339.
- Kapoor, Avnish, Matthew S Goldberg, Lara K Cumberland, Kajan Ratnakumar, Miguel F Segura, Patrick O Emanuel, Silvia Menendez, et al. 2010. "The histone variant macroH2A suppresses melanoma progression through regulation of CDK8." *Nature* 468, no. 7327 (December): 1105–9.
<http://www.ncbi.nlm.nih.gov/pubmed/21179167>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3057940>.

- Kas, Sjors M., Julian R. De Ruiter, Koen Schipper, Stefano Annunziato, Eva Schut, Sjoerd Klarenbeek, Anne Paulien Drenth, et al. 2017. "Insertional mutagenesis identifies drivers of a novel oncogenic pathway in invasive lobular breast carcinoma." *Nature Genetics* 49, no. 8 (August): 1219–1230.
- Kasper, Lawryn H, Fayçal Boussouar, Kelli Boyd, Wu Xu, Michelle Biesen, Jerold Rehg, Troy A Baudino, John L Cleveland, and Paul K Brindle. 2005. "Two transactivation mechanisms cooperate for the bulk of HIF-1-responsive gene expression." *The EMBO Journal* 24, no. 22 (November): 3846–3858.
<http://emboj.embopress.org/cgi/doi/10.1038/sj.emboj.7600846>.
- Kasper, Lawryn H, and Paul K Brindle. 2006. "Mammalian Gene Expression Program Resiliency: The Roles of Multiple Coactivator Mechanisms in Hypoxia-Responsive Transcription." *Cell Cycle* 5 (2): 142–146. <https://www.tandfonline.com/action/journalInformation?journalCode=kccy20http://www.landesbioscience.com/journals/cc/abstract.php?id=2353>.
- Kearse, Michael G., and Jeremy E. Wilusz. 2017. *Non-AUG translation: A new start for protein synthesis in eukaryotes*, September.
- Kewley, Robyn J., Murray L. Whitelaw, and Anne Chapman-Smith. 2004. *The mammalian basic helix-loop-helix/PAS family of transcriptional regulators*.
- Kim, H, G Claps, A Möller, D Bowtell, X Lu, and Z A Ronai. 2014. "Siah2 regulates tight junction integrity and cell polarity through control of ASPP2 stability." *Oncogene* 33, no. 15 (April): 2004–2010.
<http://www.ncbi.nlm.nih.gov/pubmed/23644657><http://www.ncbi.nlm.nih.gov/pubmedcentral/nih.gov/articlerender.fcgi?artid=PMC3917971><http://www.nature.com/articles/onc2013149>.
- Kim, Jung Whan, Irina Tchernyshyov, Gregg L. Semenza, and Chi V. Dang. 2006. "HIF-1-mediated expression of pyruvate dehydrogenase kinase: A metabolic switch required for cellular adaptation to hypoxia." *Cell Metabolism* 3, no. 3 (March): 177–185.
- Kim, S., D. Kim, S. W. Cho, J. Kim, and J.-S. Kim. 2014. "Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins." *Genome Research* 24, no. 6 (June): 1012–1019.
<http://www.ncbi.nlm.nih.gov/pubmed/24696461><http://www.ncbi.nlm.nih.gov/pubmedcentral/nih.gov/articlerender.fcgi?artid=PMC4032847><http://genome.cshlp.org/cgi/doi/10.1101/gr.171322.113>.
- Kimura, Hideo, Alessandro Weisz, Tsutomu Ogura, Yoshiaki Hitomi, Yukiko Kurashima, Kouichi Hashimoto, Fulvio D'Acquisto, Masatoshi Makuuchi, and Hiroyasu Esumi. 2001. "Identification of hypoxia-inducible factor 1 ancillary sequence and its function in vascular endothelial growth factor gene induction by hypoxia and nitric oxide." *Journal of Biological Chemistry* 276, no. 3 (January): 2292–2298.
- Kinzler, K W, and B Vogelstein. 1996. "Lessons from hereditary colorectal cancer." *Cell* 87, no. 2 (October): 159–70. <http://www.ncbi.nlm.nih.gov/pubmed/8861899>.

- Knuesel, Matthew T, Krista D Meyer, Carrie Bernecky, and Dylan J Taatjes. 2009. "The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function." *Genes & development* 23, no. 4 (February): 439–51.
<http://www.ncbi.nlm.nih.gov/pubmed/19240132>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2648653>.
- Koch, Sina, and Lena Claesson-Welsh. 2012. *Signal transduction by vascular endothelial growth factor receptors*.
- Koivunen, Peppi, Maija Hirsilä, Volkmar Günzler, Kari I Kivirikko, and Johanna Myllyharju. 2004. "Catalytic properties of the asparaginyl hydroxylase (FIH) in the oxygen sensing pathway are distinct from those of its prolyl 4-hydroxylases." *The Journal of biological chemistry* 279, no. 11 (March): 9899–904.
<http://www.ncbi.nlm.nih.gov/pubmed/14701857>.
- Koivunen, Peppi, Maija Hirsilä, Anne M. Remes, Ilmo E. Hassinen, Kari I. Kivirikko, and Johanna Myllyharju. 2007. "Inhibition of hypoxia-inducible factor (HIF) hydroxylases by citric acid cycle intermediates: Possible links between cell metabolism and stabilization of HIF." *Journal of Biological Chemistry* 282, no. 7 (February): 4524–4532.
- Kondo, Keiichi, William Y Kim, Mirna Lechpammer, and William G Kaelin. 2003. "Inhibition of HIF2 α Is Sufficient to Suppress pVHL-Defective Tumor Growth." Edited by Christopher Kemp. *PLoS Biology* 1, no. 3 (December): e83.
<https://dx.plos.org/10.1371/journal.pbio.0000083>.
- Kong, Dehe, Eun Jung Park, Andrew G. Stephen, Maura Calvani, John H. Cardellina, Anne Monks, Robert J. Fisher, Robert H. Shoemaker, and Giovanni Melillo. 2005. "Echinomycin, a small-molecule inhibitor of hypoxia-inducible factor-1 DNA-binding activity." *Cancer Research* 65, no. 19 (October): 9047–9055.
- Koonin, Eugene V, Kira S Makarova, and Feng Zhang. 2017. "Diversity, classification and evolution of CRISPR-Cas systems." *Current Opinion in Microbiology* 37 (June): 67–78. <http://www.ncbi.nlm.nih.gov/pubmed/28605718>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5776717>
<https://linkinghub.elsevier.com/retrieve/pii/S1369527417300231>.
- Koukourakis, Michael I, Alexandra Giatromanolaki, Efthimios Sivridis, Constantinos Simopoulos, Helen Turley, Kate Talks, Kevin C Gatter, and Adrian L Harris. 2002. "Hypoxia-inducible factor (HIF1A and HIF2A), angiogenesis, and chemoradiotherapy outcome of squamous cell head-and-neck cancer." *International journal of radiation oncology, biology, physics* 53, no. 5 (August): 1192–202. <http://www.ncbi.nlm.nih.gov/pubmed/12128120>.
- Kulshreshtha, Ritu, Manuela Ferracin, Sylwia E Wojcik, Ramiro Garzon, Hansjuerg Alder, Francisco J Agosto-Perez, Ramana Davuluri, et al. 2007. "A MicroRNA Signature of Hypoxia †." *MOLECULAR AND CELLULAR BIOLOGY* 27 (5): 1859–1867. <http://mcb.asm.org/>.
- Lando, D., Daniel J Peet, Dean A Whelan, Jeffrey J Gorman, and Murray L Whitelaw. 2002. "Asparagine Hydroxylation of the HIF Transactivation Domain: A Hypoxic Switch." *Science* 295, no. 5556 (February): 858–861.
<http://www.ncbi.nlm.nih.gov/pubmed/11823643>
<http://www.sciencemag.org/cgi/doi/10.1126/science.1068592>.

- Lando, David, Daniel J Peet, Jeffrey J Gorman, Dean A Whelan, Murray L Whitelaw, and Richard K Bruick. 2002. "FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor." *Genes & development* 16, no. 12 (June): 1466–71.
<http://www.ncbi.nlm.nih.gov/pubmed/12080085>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC186346>.
- Langton, Paul F., Julien Colombani, Birgit L. Aerne, and Nicolas Tapon. 2007. "Drosophila ASPP Regulates C-Terminal Src Kinase Activity." *Developmental Cell* 13, no. 6 (December): 773–782.
- Langton, Paul F., Julien Colombani, E. H Y Chan, Alexander Wepf, Matthias Gstaiger, and Nicolas Tapon. 2009. "The dASPP-dRASSF8 Complex Regulates Cell-Cell Adhesion during Drosophila Retinal Morphogenesis." *Current Biology* 19, no. 23 (December): 1969–1978.
- Lapi, Eleonora, Silvia Di Agostino, Sara Donzelli, Hilah Gal, Eytan Domany, Gideon Rechavi, Pier Paolo Pandolfi, et al. 2008. "PML, YAP, and p73 Are Components of a Proapoptotic Autoregulatory Feedback Loop." *Molecular Cell* 32, no. 6 (December): 803–814.
<https://linkinghub.elsevier.com/retrieve/pii/S1097276508008411>.
- Lee, Andrew C., Brett E. Fenster, Hideki Ito, Kazuyo Takeda, Nancy S. Bae, Tazuko Hirai, Zu Xi Yu, Victor J. Ferrans, Bruce H. Howard, and Toren Finkel. 1999. "Ras proteins induce senescence by altering the intracellular levels of reactive oxygen species." *Journal of Biological Chemistry* 274, no. 12 (March): 7936–7940.
- Lee, Jason S., Yunho Kim, Jinhyuk Bhin, Hi Jai R. Shin, Hye Jin Nam, Seung Hoon Lee, Jong Bok Yoon, et al. 2011. "Hypoxia-induced methylation of a pontin chromatin remodeling factor." *Proceedings of the National Academy of Sciences of the United States of America* 108, no. 33 (August): 13510–13515.
- Leissing, Thomas Markus. 2017. "STUDIES ON FIH INTERACTIONS." PhD diss.
- Li, Bo, and Colin N Dewey. 2011. "RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome." *BMC Bioinformatics* 12, no. 1 (December): 323.
<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-323>.
- Li, Junan, Anjali Mahajan, and Ming-Daw Tsai. 2006. "Ankyrin repeat: a unique motif mediating protein-protein interactions." *Biochemistry* 45, no. 51 (December): 15168–78. <http://www.ncbi.nlm.nih.gov/pubmed/17176038>.
- Liao, Yang, Gordon K Smyth, and Wei Shi. 2019. "The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads." *Nucleic Acids Research* 47 (8). <https://github.com/alexdobin/STAR>.
- Liberzon, Arthur, Aravind Subramanian, Reid Pinchback, Helga Thorvaldsdóttir, Pablo Tamayo, and Jill P. Mesirov. 2011. "Molecular signatures database (MSigDB) 3.0." *Bioinformatics* 27, no. 12 (June): 1739–1740.
- Linehan, W. Marston, Jeffrey S. Rubin, and Donald P. Bottaro. 2009. *VHL loss of function and its impact on oncogenic signaling networks in clear cell renal cell carcinoma*, April.

- Liu, Chen-Ying, Xianbo Lv, Tingting Li, Yanping Xu, Xin Zhou, Shimin Zhao, Yue Xiong, Qun-Ying Lei, and Kun-Liang Guan. 2011. "PP1 Cooperates with ASPP2 to Dephosphorylate and Activate TAZ." *Journal of Biological Chemistry* 286, no. 7 (February): 5558–5566. <http://www.jbc.org/content/286/7/5558.full>.
- Liu, Y, J A Ranish, R Aebersold, and S Hahn. 2001. "Yeast nuclear extract contains two major forms of RNA polymerase II mediator complexes." *The Journal of biological chemistry* 276, no. 10 (March): 7169–75. <http://www.ncbi.nlm.nih.gov/pubmed/11383511>.
- Llanos, Susana, Christophe Royer, Min Lu, Daniele Bergamaschi, Wen Hwa Lee, and Xin Lu. 2011. "Inhibitory Member of the Apoptosis-stimulating Proteins of the p53 Family (iASPP) Interacts with Protein Phosphatase 1 via a Noncanonical Binding Motif *."
- Lopez, C. D., Y. Ao, L. H. Rohde, T. D. Perez, D. J. O'Connor, X. Lu, J. M. Ford, and L. Naumovski. 2000. "Proapoptotic p53-Interacting Protein 53BP2 Is Induced by UV Irradiation but Suppressed by p53." *Molecular and Cellular Biology* 20, no. 21 (November): 8018–8025.
- Luo, Weibo, Ryan Chang, Jun Zhong, Akhilesh Pandey, and Gregg L. Semenza. 2012. "Histone demethylase JMJD2C is a coactivator for hypoxia-inducible factor 1 that is required for breast cancer progression." *Proceedings of the National Academy of Sciences of the United States of America* 109, no. 49 (December).
- Luse, Donal S. 2013. *The RNA polymerase II preinitiation complex: Through what pathway is the complex assembled?*
- Ma, Biao, Yan Chen, Ling Chen, Hongcheng Cheng, Chenglong Mu, Jie Li, Ruize Gao, et al. 2015. "Hypoxia regulates Hippo signalling through the SIAH2 ubiquitin E3 ligase." *Nature cell biology* 17, no. 1 (January): 95–103. <http://www.ncbi.nlm.nih.gov/pubmed/25438054>.
- Madan, Esha, Taylor M. Parker, Christopher J. Pelham, Antonio M. Palma, Maria L. Peixoto, Masaki Nagane, Aliya Chandaria, et al. 2019. "HIF-transcribed p53 chaperones HIF-1 α ." *Nucleic acids research* 47, no. 19 (November): 10212–10234.
- Mahon, P C, K Hirota, and G L Semenza. 2001. "FIH-1: a novel protein that interacts with HIF-1 α and VHL to mediate repression of HIF-1 transcriptional activity." *Genes & development* 15, no. 20 (October): 2675–86. <http://www.ncbi.nlm.nih.gov/pubmed/11641274> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC312814>.
- Maier, Bert, Sabrina Wendt, Jens T. Vanselow, Thomas Wallaeh, Silke Reischl, Stefanie Oehmke, Andreas Schlosser, and Aehim Kramer. 2009. "A large-scale functional RNAi screen reveals a role for CK2 in the mammalian circadian clock." *Genes and Development* 23, no. 6 (March): 708–718.
- Mak, Victor C.Y., Lee Lee, Michelle K.Y. Siu, Oscar G.W. Wong, Xin Lu, Hextan Y.S. Ngan, Esther S.Y. Wong, and Annie N.Y. Cheung. 2013. "Downregulation of ASPP2 in choriocarcinoma contributes to increased migratory potential through Src signaling pathway activation." *Carcinogenesis* 34, no. 9 (September): 2170–2177.

- Makino, Yuichi, Renhai Cao, Kristian Svensson, Göran Bertilsson, Mikael Asman, Hirotohi Tanaka, Yihai Cao, Anders Berkenstam, and Lorenz Poellinger. 2001. "Inhibitory PAS domain protein is a negative regulator of hypoxia-inducible gene expression." *Nature* 414, no. 6863 (November): 550–554.
<http://www.ncbi.nlm.nih.gov/pubmed/11734856>
<http://www.nature.com/articles/35107085>.
- Malik, Sohail, and Robert G. Roeder. 2000. *Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells*, June.
- . 2010. *The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation*.
- Marxsen, Jan H, Petra Stengel, Kathrin Doege, Pekka Heikkinen, Terhi Jokilehto, Thomas Wagner, Wolfgang Jelkmann, Panu Jaakkola, and Eric Metzen. 2004. "Hypoxia-inducible factor-1 (HIF-1) promotes its degradation by induction of HIF- α -prolyl-4-hydroxylases." *The Biochemical journal* 381, no. Pt 3 (August): 761–7. <http://www.ncbi.nlm.nih.gov/pubmed/15104534>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1133886>.
- Masson, N, C Willam, P H Maxwell, C W Pugh, and P J Ratcliffe. 2001. "Independent function of two destruction domains in hypoxia-inducible factor- α chains activated by prolyl hydroxylation." *The EMBO journal* 20, no. 18 (September): 5197–206. <http://www.ncbi.nlm.nih.gov/pubmed/11566883>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC125617>.
- Masson, Norma, Rachelle S Singleton, Rok Sekirnik, David C Trudgian, Lucy J Ambrose, Melroy X Miranda, Ya-Min Tian, Benedikt M Kessler, Christopher J Schofield, and Peter J Ratcliffe. 2012. "The FIH hydroxylase is a cellular peroxide sensor that modulates HIF transcriptional activity." *EMBO reports* 13:251–257.
- Matallanas, David, David Romano, Karen Yee, Katrin Meissl, Lucia Kucerova, Daniela Piazzolla, Manuela Baccarini, J. Keith Vass, Walter Kolch, and Eric O'Neill. 2007. "RASSF1A Elicits Apoptosis through an MST2 Pathway Directing Proapoptotic Transcription by the p73 Tumor Suppressor Protein." *Molecular Cell* 27, no. 6 (September): 962–975.
- Maxwell, Patrick H., Michael S. Wiesener, Gin-Wen Chang, Steven C. Clifford, Emma C. Vaux, Matthew E. Cockman, Charles C. Wykoff, Christopher W. Pugh, Eamonn R. Maher, and Peter J. Ratcliffe. 1999. "The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis." *Nature* 399, no. 6733 (May): 271–275.
<http://www.ncbi.nlm.nih.gov/pubmed/10353251>
<http://www.nature.com/articles/20459>.
- Maynard, Mindy A., Andrew J. Evans, Tomoko Hosomi, Shuntaro Hara, Michael A.S. Jewett, and Michael Ohh. 2005. "Human HIF-3 α 4 is a dominant-negative regulator of HIF-1 and is down-regulated in renal cell carcinoma." *FASEB Journal* 19, no. 11 (September): 1396–1406.

- Maynard, Mindy A., Heng Qi, Jacky Chung, Eric H.L. Lee, Yukihiro Kondo, Shuntaro Hara, Ronald C. Conaway, Joan W. Conaway, and Michael Ohh. 2003. "Multiple splice variants of the human HIF-3 α locus are targets of the von Hippel-Lindau E3 ubiquitin ligase complex." *Journal of Biological Chemistry* 278, no. 13 (March): 11032–11040.
- McClelland, Mark L, Tim M Soukup, Scot D Liu, Jonathan H Esensten, Felipe de Sousa e Melo, Murat Yaylaoglu, Soren Warming, Merone Roose-Girma, and Ron Firestein. 2015. "Cdk8 deletion in the Apc(Min) murine tumour model represses EZH2 activity and accelerates tumourigenesis." *The Journal of pathology* 237, no. 4 (December): 508–19. <http://www.ncbi.nlm.nih.gov/pubmed/26235356>.
- McDermott, Martina S.J., Alexander A. Chumanovich, Chang Uk Lim, Jiaxin Liang, Mengqian Chen, Serena Altília, David Oliver, et al. 2017. "Inhibition of CDK8 mediator kinase suppresses estrogen dependent transcription and the growth of estrogen receptor positive breast cancer." *Oncotarget* 8 (8): 12558–12575.
- Mereiter, Stefan, Ana Magalhães, Barbara Adamczyk, Chunsheng Jin, Andreia Almeida, Lylia Drici, Maria Ibáñez-Vea, et al. 2016. "Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer." *Biochimica et Biophysica Acta - General Subjects* 1860, no. 8 (August): 1795–1808.
- Meyerson, Matthew, Stacey Gabriel, and Gad Getz. 2010. "Advances in understanding cancer genomes through second-generation sequencing." *Nature Reviews Genetics* 11, no. 10 (October): 685–696. <http://www.nature.com/articles/nrg2841>.
- Miki, Rika, Koji Kadota, Hidemasa Bono, Yosuke Mizuno, Yasuhiro Tomaru, Piero Carninci, Masayoshi Itoh, et al. 2001. "Delineating developmental and metabolic pathways in vivo by expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays." *Proceedings of the National Academy of Sciences of the United States of America* 98, no. 5 (February): 2199–2204.
- Moeller, Benjamin J., Yiting Cao, Chuan Y. Li, and Mark W. Dewhirst. 2004. "Radiation activates HIF-1 to regulate vascular radiosensitivity in tumors: Role of reoxygenation, free radicals, and stress granules." *Cancer Cell* 5, no. 5 (May): 429–441.
- Moeller, Benjamin J., Rachel A. Richardson, and Mark W. Dewhirst. 2007. *Hypoxia and radiotherapy: Opportunities for improved outcomes in cancer treatment*, June.
- Muller, Patricia A.J., and Karen H. Vousden. 2014. *Mutant p53 in cancer: New functions and therapeutic opportunities*, March.
- Murphy, Michael P. 2009. *How mitochondria produce reactive oxygen species*, January.
- Mylonis, Ilias, Georgia Chachami, Efrosyni Paraskeva, and George Simos. 2008. "Atypical CRM1-dependent nuclear export signal mediates regulation of hypoxia-inducible factor-1 α by MAPK." *Journal of Biological Chemistry* 283, no. 41 (October): 27620–27627.
- Nanni, Simona, Valentina Benvenuti, Annalisa Grasselli, Carmen Priolo, Aurora Aiello, Stefania Mattiussi, Claudia Colussi, et al. 2009. "Endothelial NOS, estrogen receptor β , and HIFs cooperate in the activation of a prognostic transcriptional pattern in aggressive human prostate cancer." *Journal of Clinical Investigation* 119, no. 5 (May): 1093–1108.

- Nardone, Giorgia, Jorge Oliver-De La Cruz, Jan Vrbsky, Cecilia Martini, Jan Pribyl, Petr Skládal, Martin Pešl, et al. 2017. "ARTICLE YAP regulates cell mechanics by controlling focal adhesion assembly." *Nature Communications* 8. www.nature.com/naturecommunications.
- Naumovski, L, and M L Cleary. 1996. "The p53-binding protein 53BP2 also interacts with Bcl2 and impedes cell cycle progression at G2/M." *Molecular and Cellular Biology* 16, no. 7 (July): 3884–3892.
- Nicolai, Sara, Robert Mahen, Giuseppe Raschellà, Alberto Marini, Marco Pieraccioli, Michal Malewicz, Ashok R. Venkitaraman, and Gerry Melino. 2019. "ZNF281 is recruited on DNA breaks to facilitate DNA repair by non-homologous end joining." *Oncogene*.
- Niehrs, Christof, and Nicolas Pollet. 1999. "Synexpression groups in eukaryotes." *Nature* 402, no. 6761 (December): 483–487. <http://www.nature.com/articles/990025>.
- Ochs, Fena, Gopal Karemore, Ezequiel Miron, Jill Brown, Hana Sedlackova, Maj Britt Rask, Marko Lampe, et al. 2019. "Stabilization of chromatin topology safeguards genome integrity." *Nature* 574, no. 7779 (October): 571–574.
- Ohashi, Kazumasa, Kyoko Nagata, Midori Maekawa, Toshimasa Ishizaki, Shuh Narumiya, and Kensaku Mizuno. 2000. "Rho-associated kinase ROCK activates LIM-kinase 1 by phosphorylation at threonine 508 within the activation loop." *Journal of Biological Chemistry* 275, no. 5 (February): 3577–3582.
- Oka, T., A. P. Schmitt, and M. Sudol. 2012. "Opposing roles of angiomin-like-1 and zona occludens-2 on pro-apoptotic function of YAP." *Oncogene* 31, no. 1 (January): 128–134.
- Oka, Tsutomu, Virginia Mazack, and Marius Sudol. 2008. "Mst2 and Lats kinases regulate apoptotic function of Yes kinase-associated protein (YAP)." *Journal of Biological Chemistry* 283, no. 41 (October): 27534–27546.
- Pablo, Casas. 2019. *Package 'funModeling' Type Package Title Exploratory Data Analysis and Data Preparation Tool-Box*. Technical report.
- Pattanayak, Vikram, Steven Lin, John P Guilingier, Enbo Ma, Jennifer A Doudna, and David R Liu. 2013. "High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity." *Nature Biotechnology* 31, no. 9 (September): 839–843. <http://www.ncbi.nlm.nih.gov/pubmed/23934178>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3782611>
<http://www.nature.com/articles/nbt.2673>.
- Peek, Clara Bien, Daniel C. Levine, Jonathan Cedernaes, Akihiko Taguchi, Yumiko Kobayashi, Stacy J. Tsai, Nicolle A. Bonar, Maureen R. McNulty, Kathryn Moynihan Ramsey, and Joseph Bass. 2017. "Circadian Clock Interaction with HIF1 α Mediates Oxygenic Metabolism and Anaerobic Glycolysis in Skeletal Muscle." *Cell Metabolism* 25, no. 1 (January): 86–92.
- Peng, Meixi, Dan Yang, Yixuan Hou, Shuiqing Liu, Maojia Zhao, Yilu Qin, Rui Chen, Yong Teng, and Manran Liu. 2019. "Intracellular citrate accumulation by oxidized ATM-mediated metabolism reprogramming via PFKF and CS enhances hypoxic breast cancer cell invasion and metastasis." *Cell Death and Disease* 10, no. 3 (March).

- Perez-Perri, Joel I., Veronica L. Dengler, K. Audrey Audetat, Ahwan Pandey, Elizabeth A. Bonner, Marjeta Urh, Jacqui Mendez, et al. 2016. "The TIP60 Complex Is a Conserved Coactivator of HIF1A." *Cell Reports* 16, no. 1 (June): 37–47.
- Pollard, Patrick J., Christoph Loenarz, David R. Mole, Michael A. McDonough, Jonathan M. Gleadle, Christopher J. Schofield, and Peter J. Ratcliffe. 2008. "Regulation of Jumonji-domain-containing histone demethylases by hypoxia-inducible factor (HIF)-1 α ." *Biochemical Journal* 416, no. 3 (December): 387–394.
- Pollard, Patrick J., Noel C. Wortham, and Ian P.M. Tomlinson. 2003. *The TCA cycle and tumorigenesis: The examples of fumarate hydratase and succinate dehydrogenase*.
- Pouyssegur, Jacques, Frédéric Dayan, and Nathalie M. Mazure. 2006. *Hypoxia signalling in cancer and approaches to enforce tumour regression*, May.
- Ran, F Ann, Patrick D Hsu, Jason Wright, Vineeta Agarwala, David A Scott, and Feng Zhang. 2013. "Genome engineering using the CRISPR-Cas9 system." *Nature protocols* 8, no. 11 (November): 2281–2308.
<http://www.ncbi.nlm.nih.gov/pubmed/24157548>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3969860>.
- Ratcliffe, Peter J. 2013. "Oxygen sensing and hypoxia signalling pathways in animals: the implications of physiology for cancer." *The Journal of physiology* 591, no. 8 (April): 2027–42. <http://www.ncbi.nlm.nih.gov/pubmed/23401619>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3634517>.
- Raval, Raju R, Kah Weng Lau, Maxine G B Tran, Heidi M Sowter, Stefano J Mandriota, Ji-Liang Li, Christopher W Pugh, Patrick H Maxwell, Adrian L Harris, and Peter J Ratcliffe. 2005. "Contrasting properties of hypoxia-inducible factor 1 (HIF-1) and HIF-2 in von Hippel-Lindau-associated renal cell carcinoma." *Molecular and cellular biology* 25, no. 13 (July): 5675–86.
<http://www.ncbi.nlm.nih.gov/pubmed/15964822>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1157001>.
- Ravi, R, B Mookerjee, Z M Bhujwalla, C H Sutter, D Artemov, Q Zeng, L E Dillehay, A Madan, G L Semenza, and A Bedi. 2000. "Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1 α ." *Genes & development* 14, no. 1 (January): 34–44.
<http://www.ncbi.nlm.nih.gov/pubmed/10640274>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC316350>.
- Ravi, Rajani, Bijoyesh Mookerjee, Zaver M. Bhujwalla, Carrie Hayes Sutter, Dmitri Artemov, Qinwen Zeng, Larry E. Dillehay, Ashima Madan, Gregg L. Semenza, and Atul Bedi. 2000. "Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1 α ." *Genes and Development* 14, no. 1 (January): 34–44.
- Renan, M J. 1993. "How many mutations are required for tumorigenesis? Implications from human cancer data." *Molecular carcinogenesis* 7 (3): 139–46.
<http://www.ncbi.nlm.nih.gov/pubmed/8489711>.
- Rhodes, Daniel R., Jianjun Yu, K. Shanker, Nandan Deshpande, Radhika Varambally, Debashis Ghosh, Terrence Barrette, Akhilesh Pandey, and Arul M. Chinnaiyan. 2004. *ONCOMINE: A Cancer Microarray Database and Integrated Data-Mining Platform*, January.

- Robinson, Mark D, Davis J McCarthy, and Gordon K Smyth. 2010. "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data." *Bioinformatics (Oxford, England)* 26, no. 1 (January): 139–40.
<http://www.ncbi.nlm.nih.gov/pubmed/19910308>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2796818>.
- Ronson, George E., Ann Liza Piberger, Martin R. Higgs, Anna L. Olsen, Grant S. Stewart, Peter J. McHugh, Eva Petermann, and Nicholas D. Lakin. 2018. "PARP1 and PARP2 stabilise replication forks at base excision repair intermediates through Fbh1-dependent Rad51 regulation." *Nature Communications* 9, no. 1 (December).
- Royer, Christophe, Sofia Koch, Xiao Qin, Jaroslav Zak, Ludovico Buti, Ewa Dudzic, Shan Zhong, Indrika Ratnayaka, Shankar Srinivas, and Xin Lu. 2014. "ASPP2 links the apical lateral polarity complex to the regulation of YAP activity in epithelial cells." *PloS one* 9 (10): e111384.
<http://www.ncbi.nlm.nih.gov/pubmed/25360797>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4216074>.
- Saadat, Iraj, Hideaki Higashi, Chikashi Obuse, Mayumi Umeda, Naoko Murata-Kamiya, Yasuhiro Saito, Huaisheng Lu, et al. 2007. "Helicobacter pylori CagA targets PAR1/MARK kinase to disrupt epithelial cell polarity." *Nature* 447, no. 7142 (May): 330–333.
- Sachdev, Shrikesh, Alexander Hoffmann, and Mark Hannink. 1998. "Nuclear Localization of I κ B α Is Mediated by the Second Ankyrin Repeat: the I κ B α Ankyrin Repeats Define a Novel Class of cis-Acting Nuclear Import Sequences." *Molecular and Cellular Biology* 18, no. 5 (May): 2524–2534.
- Salama, Rafik, Norma Masson, Peter Simpson, Lina Katrin Sciesielski, Min Sun, Ya Min Tian, Peter John Ratcliffe, and David Robert Mole. 2015. "Heterogeneous effects of direct hypoxia pathway activation in kidney cancer." *PLoS ONE* 10, no. 8 (August).
- Samuels-Lev, Yardena, Daniel J. O'Connor, Daniele Bergamaschi, Giuseppe Trigiane, Jung Kuang Hsieh, Shan Zhong, Isabelle Campargue, Louie Naumovski, Tim Crook, and Xin Lu. 2001. "ASPP proteins specifically stimulate the apoptotic function of p53." *Molecular Cell* 8 (4): 781–794.
- Sánchez-Puig, Nuria, Dmitry B Veprintsev, and Alan R Fersht. 2005. "Binding of natively unfolded HIF-1 α ODD domain to p53." *Molecular cell* 17, no. 1 (January): 11–21. <http://www.ncbi.nlm.nih.gov/pubmed/15629713>.
- Schmid, Tobias, Jie Zhou, Roman Köhl, and Bernhard Brüne. 2004. "p300 relieves p53-evoked transcriptional repression of hypoxia-inducible factor-1 (HIF-1)." *Biochemical Journal* 380, no. 1 (May): 289–295.
- Schmid, Virginia, Veronique N. Lafleur, Olivia Lombardi, Ran Li, Rafik Salama, Leandro Colli, Hani Choudhry, Stephen Chanock, Peter J. Ratcliffe, and David R. Mole. 2019. "Co-incidence of RCC-susceptibility polymorphisms with HIF cis-acting sequences supports a pathway tuning model of cancer." *Scientific Reports* 9, no. 1 (December): 1–17.

- Schmierer, Bernhard, Béla Novák, and Christopher J Schofield. 2010. *Hypoxia-dependent sequestration of an oxygen sensor by a widespread structural motif can shape the hypoxic response - a predictive kinetic model*. Technical report.
<http://www.biomedcentral.com/1752-0509/4/139>.
- Schmittgen, Thomas D., and Kenneth J. Livak. 2008. "Analyzing real-time PCR data by the comparative CT method." *Nature Protocols* 3, no. 6 (May): 1101–1108.
- Schmitz, Klaus Jürgen, Carmen Ina Müller, Henning Reis, Hakan Alakus, Günther Winde, Hideo Andreas Baba, Jeremias Wohlschlaeger, Bharat Jasani, Joachim Fandrey, and Kurt Werner Schmid. 2009. "Combined analysis of hypoxia-inducible factor 1 alpha and metallothionein indicates an aggressive subtype of colorectal carcinoma." *International Journal of Colorectal Disease* 24 (11): 1287–1296.
- Schödel, Johannes, Spyros Oikonomopoulos, Jiannis Ragoussis, Christopher W. Pugh, Peter J. Ratcliffe, and David R. Mole. 2011. "High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq." *Blood* 117, no. 23 (June).
- Schofield, C J, and Z Zhang. 1999. "Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes." *Current opinion in structural biology* 9, no. 6 (December): 722–31.
<http://www.ncbi.nlm.nih.gov/pubmed/10607676>.
- Scholz, H, H J Schurek, K U Eckardt, and C Bauer. 1990. "Role of erythropoietin in adaptation to hypoxia." *Experientia* 46, nos. 11-12 (December): 1197–201.
<http://www.ncbi.nlm.nih.gov/pubmed/2253723>.
- Selak, Mary A., Sean M. Armour, Elaine D. MacKenzie, Houda Boulahbel, David G. Watson, Kyle D. Mansfield, Yi Pan, M. Celeste Simon, Craig B. Thompson, and Eyal Gottlieb. 2005. "Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF- α prolyl hydroxylase." *Cancer Cell* 7, no. 1 (January): 77–85.
- Semenova, E., M. M. Jore, K. A. Datsenko, A. Semenova, E. R. Westra, B. Wanner, J. van der Oost, S. J. J. Brouns, and K. Severinov. 2011. "Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence." *Proceedings of the National Academy of Sciences* 108, no. 25 (June): 10098–10103. <http://www.ncbi.nlm.nih.gov/pubmed/21646539><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3121866><http://www.pnas.org/cgi/doi/10.1073/pnas.1104144108>.
- Semenza, G L, and G L Wang. 1992. "A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation." *Molecular and cellular biology* 12, no. 12 (December): 5447–54. <http://www.ncbi.nlm.nih.gov/pubmed/1448077><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC360482>.
- Semenza, Gregg L. 2002. "Physiology meets biophysics: visualizing the interaction of hypoxia-inducible factor 1 alpha with p300 and CBP." *Proceedings of the National Academy of Sciences of the United States of America* 99, no. 18 (September): 11570–2. <http://www.ncbi.nlm.nih.gov/pubmed/12186981><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC129309>.

- Semenza, Gregg L. 2012. "Hypoxia-inducible factors in physiology and medicine." *Cell* 148, no. 3 (February): 399–408.
<http://www.ncbi.nlm.nih.gov/pubmed/22304911>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3437543>.
- . 2010. *Oxygen homeostasis*, May.
- Sena, J. A., L. Wang, and C.-J. Hu. 2013. "BRG1 and BRM Chromatin-Remodeling Complexes Regulate the Hypoxia Response by Acting as Coactivators for a Subset of Hypoxia-Inducible Transcription Factor Target Genes." *Molecular and Cellular Biology* 33, no. 19 (October): 3849–3863.
- Sergushichev, Alexey A. 2016. "An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation." *bioRxiv*: 060012.
<http://dx.doi.org/10.1101/060012>
<https://www.biorxiv.org/content/10.1101/060012v1>.
- Serrano, Manuel, Athena W. Lin, Mila E. McCurrach, David Beach, and Scott W. Lowe. 1997. "Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16(INK4a)." *Cell* 88, no. 5 (March): 593–602.
- Sharma, Kirti, Sebastian Schmitt, Caroline G. Bergner, Stefka Tyanova, Nirmal Kannaiyan, Natalia Manrique-Hoyos, Karina Kongi, et al. 2015. "Cell type- and brain region-resolved mouse brain proteome." *Nature Neuroscience* 18, no. 12 (November): 1819–1831.
- Shen, Chuan, Rameen Beroukhi, Steven E. Schumacher, Jing Zhou, Michelle Chang, Sabina Signoretti, and William G. Kaelin. 2011. "Genetic and functional studies implicate HIF1a as a 14q kidney cancer suppressor gene." *Cancer Discovery* 1, no. 3 (August): 222–235.
- Sherr, Charles J., and Frank McCormick. 2002. *The RB and p53 pathways in cancer*.
- Shi, Xianle, Zixi Yin, Bin Ling, Lingling Wang, Chang Liu, Xianhui Ruan, Weiyu Zhang, and Lingyi Chen. 2017. "Rho differentially regulates the hippo pathway by modulating the interaction between amot and Nf2 in the blastocyst." *Development (Cambridge)* 144, no. 21 (November): 3957–3967.
- Shin, Dong Hoon, Shan Hua Li, Seung-Won Yang, Byung Lan Lee, Myung Kyu Lee, and Jong-Wan Park. 2009. "Inhibitor of nuclear factor-kappaB alpha derepresses hypoxia-inducible factor-1 during moderate hypoxia by sequestering factor inhibiting hypoxia-inducible factor from hypoxia-inducible factor 1alpha." *The FEBS journal* 276, no. 13 (July): 3470–80. <http://www.ncbi.nlm.nih.gov/pubmed/19456861>.
- Simanshu, Dharendra K., Dwight V. Nissley, and Frank McCormick. 2017. "RAS Proteins and Their Regulators in Human Disease." *Cell* 170, no. 1 (June): 17–33.
<https://linkinghub.elsevier.com/retrieve/pii/S0092867417306505>.
- Sit, Soon Tuck, and Ed Manser. 2011. *Rho GTPases and their role in organizing the actin cytoskeleton*, March.
- Smythies, James A, Min Sun, Norma Masson, Rafik Salama, Peter D Simpson, Elizabeth Murray, Viviana Neumann, et al. 2019. "Inherent DNA -binding specificities of the HIF -1 α and HIF -2 α transcription factors in chromatin." *EMBO reports* 20, no. 1 (January).
<https://onlinelibrary.wiley.com/doi/abs/10.15252/embr.201846401>.

- Sottocornola, Roberta, Christophe Royer, Virginie Vives, Luca Tordella, Shan Zhong, Yihua Wang, Indrika Ratnayaka, et al. 2010. "ASPP2 Binds Par-3 and Controls the Polarity and Proliferation of Neural Progenitors during CNS Development." *Developmental Cell* 19, no. 1 (July): 126–137.
<http://www.ncbi.nlm.nih.gov/pubmed/20619750>
<https://linkinghub.elsevier.com/retrieve/pii/S1534580710002935>.
- Stein, Claudia, Anaïs Flore Bardet, Guglielmo Roma, Sebastian Bergling, Ieuan Clay, Alexandra Ruchti, Claudia Agarinis, et al. 2015. "YAP1 Exerts Its Transcriptional Control via TEAD-Mediated Activation of Enhancers." *PLoS genetics* 11, no. 8 (August): e1005465. <http://www.ncbi.nlm.nih.gov/pubmed/26295846>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4546604>.
- Stiehl, Daniel P, Renato Wirthner, Jens Köditz, Patrick Spielmann, Gieri Camenisch, and Roland H Wenger. 2006. "Increased prolyl 4-hydroxylase domain proteins compensate for decreased oxygen levels. Evidence for an autoregulatory oxygen-sensing system." *The Journal of biological chemistry* 281, no. 33 (August): 23482–91. <http://www.ncbi.nlm.nih.gov/pubmed/16790428>.
- Stolze, Ineke P, Ya-Min Tian, Rebecca J Appelhoff, Helen Turley, Charles C Wykoff, Jonathan M Gleadle, and Peter J Ratcliffe. 2004. "Genetic analysis of the role of the asparaginyl hydroxylase factor inhibiting hypoxia-inducible factor (FIH) in regulating hypoxia-inducible factor (HIF) transcriptional target genes [corrected]." *The Journal of biological chemistry* 279, no. 41 (October): 42719–25.
<http://www.ncbi.nlm.nih.gov/pubmed/15302861>.
- Subramanian, Aravind, Pablo Tamayo, Vamsi K Mootha, Sayan Mukherjee, Benjamin L Ebert, Michael A Gillette, Amanda Paulovich, et al. 2005. "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles." *Proceedings of the National Academy of Sciences of the United States of America* 102, no. 43 (October): 15545–50.
<http://www.ncbi.nlm.nih.gov/pubmed/16199517>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1239896>.
- Sullivan, A, and X Lu. 2007. "ASPP: a new family of oncogenes and tumour suppressor genes." *British journal of cancer* 96, no. 2 (January): 196–200.
<http://www.ncbi.nlm.nih.gov/pubmed/17211478>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2359998>.
- Sun, Hong-Cheng, Zheng-Jun Qiu, Jun Liu, Jing Sun, Tao Jiang, Ke-Jian Huang, Min Yao, and Chen Huang. 2007. "Expression of hypoxia-inducible factor-1 alpha and associated proteins in pancreatic ductal adenocarcinoma and their impact on prognosis." *International journal of oncology* 30, no. 6 (June): 1359–67.
<http://www.ncbi.nlm.nih.gov/pubmed/17487356>.
- Suzuki, Hiroyuki, Akihiro Tomida, and Takashi Tsuruo. 2001. "Dephosphorylated hypoxia-inducible factor 1 α as a mediator of p53-dependent apoptosis during hypoxia." *Oncogene* 20, no. 41 (September): 5779–5788.
<http://www.ncbi.nlm.nih.gov/pubmed/11593383>
<http://www.nature.com/articles/1204742>.

- Szkarczyk, Damian, Andrea Franceschini, Stefan Wyder, Kristoffer Forslund, Davide Heller, Jaime Huerta-Cepas, Milan Simonovic, et al. 2015. "STRING v10: Protein-protein interaction networks, integrated over the tree of life." *Nucleic Acids Research* 43, no. D1 (January): D447–D452.
- Taatjes, Dylan J, Anders M Näär, Frank Andel, Eva Nogales, and Robert Tjian. 2002. "Structure, function, and activator-induced conformations of the CRSP coactivator." *Science (New York, N.Y.)* 295, no. 5557 (February): 1058–62. <http://www.ncbi.nlm.nih.gov/pubmed/11834832>.
- Takahashi, Naoko, Shinya Kobayashi, Xu Jiang, Koji Kitagori, Kenichi Imai, Yurina Hibi, and Takashi Okamoto. 2004. "Expression of 53BP2 and ASPP2 proteins from TP53BP2 gene by alternative splicing." *Biochemical and Biophysical Research Communications* 315, no. 2 (March): 434–438. <http://www.ncbi.nlm.nih.gov/pubmed/14766226> <https://linkinghub.elsevier.com/retrieve/pii/S0006291X04000944>.
- Talks, K. L., H. Turley, K. C. Gatter, P. H. Maxwell, C. W. Pugh, P. J. Ratcliffe, and A. L. Harris. 2000. "The expression and distribution of the hypoxia-inducible factors HIF-1 α and HIF-2 α in normal human tissues, cancers, and tumor-associated macrophages." *American Journal of Pathology* 157 (2): 411–421.
- Tarhonskaya, Hanna, Adam P Hardy, Emily A Howe, Nikita D Loik, Holger B Kramer, James S O McCullagh, Christopher J Schofield, and Emily Flashman. 2015. "Kinetic Investigations of the Role of Factor Inhibiting Hypoxia-inducible Factor (FIH) as an Oxygen Sensor." *The Journal of biological chemistry* 290, no. 32 (August): 19726–42. <http://www.ncbi.nlm.nih.gov/pubmed/26112411> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4528135>.
- Theodoropoulos, Vasilios E., Andreas C. Lazaris, Ioannis Kastriotis, Chariclia Spiliadi, George E. Theodoropoulos, Vasiliki Tsoukala, Efstratios Patsouris, and Frank Sofras. 2005. "Evaluation of hypoxia-inducible factor 1 α overexpression as a predictor of tumour recurrence and progression in superficial urothelial bladder carcinoma." *BJU International* 95, no. 3 (February): 425–431.
- Tian, H, S L McKnight, and D W Russell. 1997. "Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells." *Genes & development* 11, no. 1 (January): 72–82. <http://www.ncbi.nlm.nih.gov/pubmed/9000051>.
- Tian, Ya-Min, Kar Kheng Yeoh, Myung Kyu Lee, Tuula Eriksson, Benedikt M. Kessler, Holger B. Kramer, Mariola J. Edelmann, et al. 2011a. "Differential Sensitivity of Hypoxia Inducible Factor Hydroxylation Sites to Hypoxia and Hydroxylase Inhibitors." *Journal of Biological Chemistry* 286, no. 15 (April): 13041–13051. <http://www.ncbi.nlm.nih.gov/pubmed/21335549> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3075650> <http://www.jbc.org/lookup/doi/10.1074/jbc.M110.211110>.
- . 2011b. "Differential Sensitivity of Hypoxia Inducible Factor Hydroxylation Sites to Hypoxia and Hydroxylase Inhibitors." *Journal of Biological Chemistry* 286, no. 15 (April): 13041–13051. <http://www.ncbi.nlm.nih.gov/pubmed/21335549> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3075650> <http://www.jbc.org/lookup/doi/10.1074/jbc.M110.211110>.

- Tidow, Henning, Antonina Andreeva, Trevor J. Rutherford, and Alan R. Fersht. 2007a. "Solution Structure of ASPP2 N-terminal Domain (N-ASPP2) Reveals a Ubiquitin-like Fold." *Journal of Molecular Biology* 371, no. 4 (August): 948–958.
- . 2007b. "Solution structure of ASPP2 N-terminal domain (N-ASPP2) reveals a ubiquitin-like fold." *Journal of molecular biology* 371, no. 4 (August): 948–58. <http://www.ncbi.nlm.nih.gov/pubmed/17594908>.
- Tsai, Shengdar Q, Zongli Zheng, Nhu T Nguyen, Matthew Liebers, Ved V Topkar, Vishal Thapar, Nicolas Wyvekens, et al. 2015. "GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases." *Nature Biotechnology* 33, no. 2 (February): 187–197. <http://www.nature.com/articles/nbt.3117>.
- Turnquist, Casmir, Yihua Wang, David T Severson, Shan Zhong, Bin Sun, Jingyi Ma, Stefan N Constaninescu, et al. 2014. "STAT1-induced ASPP2 transcription identifies a link between neuroinflammation, cell polarity, and tumor suppression." *Proceedings of the National Academy of Sciences of the United States of America* 111, no. 27 (July): 9834–9. <http://www.ncbi.nlm.nih.gov/pubmed/24958857> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4103354>.
- Tusher, Virginia Goss, Robert Tibshirani, and Gilbert Chu. 2001. "Significance analysis of microarrays applied to the ionizing radiation response." *Proceedings of the National Academy of Sciences of the United States of America* 98, no. 9 (April): 5116–5121.
- Tyanova, Stefka, Tikira Temu, Pavel Sinitcyn, Arthur Carlson, Marco Y Hein, Tamar Geiger, Matthias Mann, and Jürgen Cox. 2016. "The Perseus computational platform for comprehensive analysis of (prote)omics data." *Nature Methods* 13, no. 9 (September): 731–740. <http://www.nature.com/articles/nmeth.3901>.
- Uchida, Tokujiro, Fabrice Rossignol, Michael A Matthay, Rémi Mounier, Sylvianne Couette, Eric Clottes, and Christine Clerici. 2004. "Prolonged hypoxia differentially regulates hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha expression in lung epithelial cells: implication of natural antisense HIF-1alpha." *The Journal of biological chemistry* 279, no. 15 (April): 14871–8. <http://www.ncbi.nlm.nih.gov/pubmed/14744852>.
- Uhlen, M., L. Fagerberg, B. M. Hallstrom, C. Lindskog, P. Oksvold, A. Mardinoglu, A. Sivertsson, et al. 2015. "Tissue-based map of the human proteome." *Science* 347, no. 6220 (January): 1260419–1260419. <http://www.sciencemag.org/cgi/doi/10.1126/science.1260419>.
- Van Hook, Kathryn, Zhiping Wang, Dexi Chen, Casey Nold, Zhiyi Zhu, Pavana Anur, Hun Joo Lee, et al. 2017. "ΔN-ASPP2, a novel isoform of the ASPP2 tumor suppressor, promotes cellular survival." *Biochemical and Biophysical Research Communications* 482, no. 4 (January): 1271–1277.
- Varelas, Xaralabos. 2014. *The hippo pathway effectors TAZ and YAP in development, homeostasis and disease*.
- Vassilev, Lyubomir T., Binh T. Vu, Bradford Graves, Daisy Carvajal, Frank Podlaski, Zoran Filipovic, Norman Kong, et al. 2004. "In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2." *Science* 303, no. 5659 (February): 844–848.

- Vaupel, Peter, Arnulf Mayer, and Michael Höckel. 2004. "Tumor Hypoxia and Malignant Progression." *Methods in Enzymology* 381:335–354.
- Vazquez, Alexei, Alessandro Flammini, Amos Maritan, and Alessandro Vespignani. 2003. "Global protein function prediction from protein-protein interaction networks." *Nature Biotechnology* 21, no. 6 (June): 697–700.
- Vives, Virginie, Jian Su, Shan Zhong, Indrika Ratnayaka, Elizabeth Slee, Robert Goldin, and Xin Lu. 2006. "ASPP2 is a haploinsufficient tumor suppressor that cooperates with p53 to suppress tumor growth." *Genes and Development* 20, no. 10 (May): 1262–1267.
- Vlaminck, Benoit, Sebastien Toffoli, Benjamin Ghislain, Catherine Demazy, Martine Raes, and Carine Michiels. 2007. "Dual effect of echinomycin on hypoxia-inducible factor-1 activity under normoxic and hypoxic conditions." *FEBS Journal* 274, no. 21 (November): 5533–5542.
- Vo, Ngan, and Richard H Goodman. 2001. "CREB-binding Protein and p300 in Transcriptional Regulation*." <http://www.jbc.org/>.
- Vollmers, Christopher, Satchidananda Panda, and Luciano DiTacchio. 2009. "Correction: A High-Throughput Assay for siRNA-Based Circadian Screens in Human U2OS Cells." *PLoS ONE* 4, no. 2 (February).
- Voronin, D. A., and E. V. Kiseleva. 2008. *Functional role of proteins containing ankyrin repeats*, February.
- Vyas, Sejal, Elma Zaganjor, and Marcia C Haigis. n.d. "Mitochondria and Cancer."
- Walton, Zandra E., Chirag H. Patel, Rebekah C. Brooks, Yongjun Yu, Arig Ibrahim-Hashim, Malini Riddle, Alessandra Porcu, et al. 2018. "Acid Suspends the Circadian Clock in Hypoxia through Inhibition of mTOR." *Cell* 174, no. 1 (June): 72–87.
- Wan, W., K. Peng, M. Li, L. Qin, Z. Tong, J. Yan, B. Shen, and C. Yu. 2017. "Histone demethylase JMJD1A promotes urinary bladder cancer progression by enhancing glycolysis through coactivation of hypoxia inducible factor 1 α ." *Oncogene* 36, no. 27 (July): 3868–3877.
- Wang, Guang L., Bing Hua Jiang, Elizabeth A. Rue, and Gregg L. Semenza. 1995. "Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O₂ tension." *Proceedings of the National Academy of Sciences of the United States of America* 92, no. 12 (June): 5510–5514.
- Wang, Tim, Jenny J. Wei, David M. Sabatini, and Eric S. Lander. 2014. "Genetic Screens in Human Cells Using the CRISPR-Cas9 System." *Science* 343, no. 6166 (January): 80–84. <http://www.ncbi.nlm.nih.gov/pubmed/24336569><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3972032><http://www.sciencemag.org/lookup/doi/10.1126/science.1246981>.
- Wang, Xingwen, Miao Yu, Kunming Zhao, Mengmeng He, Wenjie Ge, Yuhui Sun, Yihua Wang, Haizhu Sun, and Ying Hu. 2016. "Upregulation of MiR-205 under hypoxia promotes epithelial-mesenchymal transition by targeting ASPP2." *Cell death & disease* 7 (12): e2517. <http://www.ncbi.nlm.nih.gov/pubmed/27929537><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5261019>.

- Wang, Y., N. Godin-Heymann, X. Dan Wang, D. Bergamaschi, S. Llanos, and X. Lu. 2013. "ASPP1 and ASPP2 bind active RAS, potentiate RAS signalling and enhance p53 activity in cancer cells." *Cell Death and Differentiation* 20, no. 4 (April): 525–534.
- Wang, Yihua, Xiao Dan Wang, Eleonora Lapi, Alexandra Sullivan, Wei Jia, You Wen He, Indrika Ratnayaka, et al. 2012. "Autophagic activity dictates the cellular response to oncogenic RAS." *Proceedings of the National Academy of Sciences of the United States of America* 109, no. 33 (August): 13325–13330.
- Wang, Zhiping, Yuangang Liu, Maho Takahashi, Kathryn Van Hook, Kerstin M. Kampa-Schittenhelm, Brett C. Sheppard, Rosalie C. Sears, Philip J.S. Stork, and Charles D. Lopez. 2013. "N terminus of ASPP2 binds to Ras and enhances Ras/Raf/MEK/ERK activation to promote oncogene-induced senescence." *Proceedings of the National Academy of Sciences of the United States of America* 110, no. 1 (January): 312–317.
- Waypa, Gregory B., Kimberly A. Smith, and Paul T. Schumacker. 2016. *O2 sensing, mitochondria and ROS signaling: The fog is lifting*, February.
- Webb, James D., Andrea Murányi, Christopher W. Pugh, Peter J. Ratcliffe, and Mathew L. Coleman. 2009a. "MYPT1, the targeting subunit of smooth-muscle myosin phosphatase, is a substrate for the asparaginyl hydroxylase factor inhibiting hypoxia-inducible factor (FIH)." *Biochemical Journal* 420 (2): 327–336.
- . 2009b. "MYPT1, the targeting subunit of smooth-muscle myosin phosphatase, is a substrate for the asparaginyl hydroxylase factor inhibiting hypoxia-inducible factor (FIH)." *The Biochemical journal* 420, no. 2 (May): 327–33. <http://www.ncbi.nlm.nih.gov/pubmed/19245366>.
- Weinstein, John N., Eric A Collisson, Gordon B Mills, Kenna R Mills Shaw, Brad A Ozenberger, Kyle Ellrott, Ilya Shmulevich, Chris Sander, and Joshua M Stuart. 2013. "The Cancer Genome Atlas Pan-Cancer analysis project." *Nature Genetics* 45, no. 10 (October): 1113–1120. <http://www.nature.com/articles/ng.2764>.
- Wienert, Beeke, Stacia K Wyman, Christopher D Richardson, Charles D Yeh, Pinar Akcakaya, Michelle J Porritt, Michaela Morlock, et al. 2019. "Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq." *Science (New York, N.Y.)* 364, no. 6437 (April): 286–289. <http://www.ncbi.nlm.nih.gov/pubmed/31000663> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC6589096>.
- Wiesener, Michael S, Jan Steffen Jürgensen, Christian Rosenberger, Charlotte K Scholze, Jan H Hörstrup, Christina Warnecke, Stefano Mandriota, et al. 2003. "Widespread hypoxia-inducible expression of HIF-2alpha in distinct cell populations of different organs." *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 17, no. 2 (February): 271–3. <http://www.ncbi.nlm.nih.gov/pubmed/12490539>.

- Wilkins, Sarah E, Jaana Hyvärinen, Johana Chicher, Jeffrey J Gorman, Daniel J Peet, Rebecca L Bilton, and Peppi Koivunen. 2009. "Differences in hydroxylation and binding of Notch and HIF-1 α demonstrate substrate selectivity for factor inhibiting HIF-1 (FIH-1)." *The international journal of biochemistry & cell biology* 41, no. 7 (July): 1563–71. <http://www.ncbi.nlm.nih.gov/pubmed/19401150>.
- Wu, Lani F., Timothy R. Hughes, Armaity P. Davierwala, Mark D. Robinson, Roland Stoughton, and Steven J. Altschuler. 2002. "Large-scale prediction of *Saccharomyces cerevisiae* gene function using overlapping transcriptional clusters." *Nature Genetics* 31 (3): 255–265.
- Wu, Yaling, Dingbin Tang, Na Liu, Wei Xiong, Huanwei Huang, Yang Li, Zhixiong Ma, et al. 2017. "Reciprocal Regulation between the Circadian Clock and Hypoxia Signaling at the Genome Level in Mammals." *Cell Metabolism* 25, no. 1 (January): 73–85.
- X.D., Wang, Lapi E., Sullivan A., Ratnayaka I., Goldin R., Hay R., Lu X., et al. 2011. "SUMO-modified nuclear cyclin D1 bypasses Ras-induced senescence." *Cell Death and Differentiation* 18 (2): 304–314.
<http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L51046710%20http://dx.doi.org/10.1038/cdd.2010.101>.
- Xia, Xiaobo, Madeleine E Lemieux, Wei Li, Jason S Carroll, Myles Brown, X Shirley Liu, and Andrew L Kung. 2009. "Integrative analysis of HIF binding and transactivation reveals its role in maintaining histone methylation homeostasis." *Proceedings of the National Academy of Sciences of the United States of America* 106, no. 11 (March): 4260–5. <http://www.ncbi.nlm.nih.gov/pubmed/19255431%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2657396>.
- Xiong, Gaofeng, Rachel L. Stewart, Jie Chen, Tianyan Gao, Timothy L. Scott, Luis M. Samayoa, Kathleen O'Connor, Andrew N. Lane, and Ren Xu. 2018. "Collagen prolyl 4-hydroxylase 1 is essential for HIF-1 α stabilization and TNBC chemoresistance." *Nature Communications* 9, no. 1 (December).
- Xu, Ke, Yueping Zhan, Zeting Yuan, Yanyan Qiu, Haijing Wang, Guohua Fan, Jie Wang, et al. 2019. "Hypoxia Induces Drug Resistance in Colorectal Cancer through the HIF-1 α /miR-338-5p/IL-6 Feedback Loop." *Molecular Therapy* (October).
- Yamamoto, Yutaka, Mutsuko Ibusuki, Yasuhiro Okumura, Teru Kawasoe, Kazuharu Kai, Kenichi Iyama, and Hirotaka Iwase. 2008. "Hypoxia-inducible factor 1 α is closely linked to an aggressive phenotype in breast cancer." *Breast Cancer Research and Treatment* 110, no. 3 (August): 465–475.
- Yan, Qin, Steven Bartz, Mao Mao, Lianjie Li, and William G Kaelin. 2007. "The hypoxia-inducible factor 2 α N-terminal and C-terminal transactivation domains cooperate to promote renal tumorigenesis in vivo." *Molecular and cellular biology* 27, no. 6 (March): 2092–102.
<http://www.ncbi.nlm.nih.gov/pubmed/17220275%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1820491>.

- Yao, Tso Pang, Suk P. Oh, Miriam Fuchs, Nai Dong Zhou, Lian Ee Ch'ng, David Newsome, Roderick T. Bronson, En Li, David M. Livingston, and Richard Eckner. 1998. "Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300." *Cell* 93, no. 3 (May): 361–372.
- Yeh, Tzu-Lan, Thomas M Leissing, Martine I Abboud, Cyrille C Thinnes, Onur Atasoylu, James P Holt-Martyn, Dong Zhang, et al. 2017. "Molecular and cellular mechanisms of HIF prolyl hydroxylase inhibitors in clinical trials." *Chemical science* 8, no. 11 (November): 7651–7668.
<http://www.ncbi.nlm.nih.gov/pubmed/29435217>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5802278>.
- Yoshinari, Tomoko, Hiromasa Okada, Akihiro Yamada, Daisuke Uemura, Hirofumi Oka, Hiroyuki Suda, and Akira Okura. 1994. "Inhibition of Topoisomerase II by a Novel Antitumor Cyclic Depsipeptide, BE-22179." *Japanese Journal of Cancer Research* 85 (5): 550–555.
- Yu, F., S. B. White, Q. Zhao, and F. S. Lee. 2001. "HIF-1 binding to VHL is regulated by stimulus-sensitive proline hydroxylation." *Proceedings of the National Academy of Sciences* 98, no. 17 (August): 9630–9635.
<http://www.ncbi.nlm.nih.gov/pubmed/11504942>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC55503>
<http://www.pnas.org/cgi/doi/10.1073/pnas.181341498>.
- Yu, Fa Xing, Bin Zhao, and Kun Liang Guan. 2015. *Hippo Pathway in Organ Size Control, Tissue Homeostasis, and Cancer*, November.
- Yu, Guangchuang, Li Gen Wang, Yanyan Han, and Qing Yu He. 2012. "ClusterProfiler: An R package for comparing biological themes among gene clusters." *OMICS A Journal of Integrative Biology* 16, no. 5 (May): 284–287.
- Yun, Zhong, Heather L. Maecker, Randall S. Johnson, and Amato J. Giaccia. 2002. "Inhibition of PPAR γ 2 gene expression by the HIF-1-regulated gene DEC1/Stra13: A mechanism for regulation of adipogenesis by hypoxia." *Developmental Cell* 2 (3): 331–341.
- Zack, Travis I., Steven E. Schumacher, Scott L. Carter, Andrew D. Cherniack, Gordon Saksena, Barbara Tabak, Michael S. Lawrence, et al. 2013. "Pan-cancer patterns of somatic copy number alteration." *Nature Genetics* 45, no. 10 (October): 1134–1140.
- Zanconato, Francesca, Michelangelo Cordenonsi, and Stefano Piccolo. 2016. *YAP/TAZ at the Roots of Cancer*, June.
- . 2019. *YAP and TAZ: a signalling hub of the tumour microenvironment*, August.
- Zatloukal, Bernhard, Iris Kufferath, Andrea Thueringer, Ulf Landegren, Kurt Zatloukal, and Johannes Haybaeck. 2014. "Sensitivity and Specificity of In situ Proximity Ligation for Protein Interaction Analysis in a Model of Steatohepatitis with Mallory-Denk Bodies." Edited by David R. Borchelt. *PLoS ONE* 9, no. 5 (May): e96690. <https://dx.plos.org/10.1371/journal.pone.0096690>.

- Zhang, Huafeng, Marta Bosch-Marce, Larissa A Shimoda, Yee Sun Tan, Jin Hyen Baek, Jacob B Wesley, Frank J Gonzalez, and Gregg L Semenza. 2008. "Mitochondrial autophagy is an HIF-1-dependent adaptive metabolic response to hypoxia." *The Journal of biological chemistry* 283, no. 16 (April): 10892–903.
<http://www.ncbi.nlm.nih.gov/pubmed/18281291>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2447655>.
- Zhang, Huafeng, Ping Gao, Ryo Fukuda, Ganesh Kumar, Balaji Krishnamachary, Karen I. Zeller, Chi V. Dang, and Gregg L. Semenza. 2007. "HIF-1 Inhibits Mitochondrial Biogenesis and Cellular Respiration in VHL-Deficient Renal Cell Carcinoma by Repression of C-MYC Activity." *Cancer Cell* 11, no. 5 (May): 407–420.
<http://www.ncbi.nlm.nih.gov/pubmed/17482131>
<https://linkinghub.elsevier.com/retrieve/pii/S1535610807001158>.
- Zhao, Bin, Xiaomu Wei, Weiquan Li, Ryan S. Udan, Qian Yang, Joungmok Kim, Joe Xie, et al. 2007. "Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control." *Genes and Development* 21, no. 21 (November): 2747–2761.
- Zheng, Xiaofeng, Sarah Linke, José M. Dias, Xiaowei Zheng, Katarina Gradin, Tristan P. Wallis, Brett R. Hamilton, et al. 2008. "Interaction with factor inhibiting HIF-1 defines an additional mode of cross-coupling between the Notch and hypoxia signaling pathways." *Proceedings of the National Academy of Sciences of the United States of America* 105, no. 9 (March): 3368–3373.
- Zhong, Hua, Angelo M. De Marzo, Erik Laughner, Michael Lim, David A. Hilton, David Zagzag, Peter Buechler, William B. Isaacs, Gregg L. Semenza, and Jonathan W. Simons. 1999. "Overexpression of hypoxia-inducible factor 1 α in common human cancers and their metastases." *Cancer Research* 59, no. 22 (November): 5830–5835.
- Zhu, Zhiyi, Jason Ramos, Kerstin Kampa, Shanthi Adimoolam, Mint Sirisawad, Zhiyong Yu, Dexi Chen, Louie Naumovski, and Charles D. Lopez. 2005. "Control of ASPP2/53BP2L protein levels by proteasomal degradation modulates p53 apoptotic function." *Journal of Biological Chemistry* 280, no. 41 (October): 34473–34480.

Appendices

Appendices

Supporting information for chapter 3	ii
Supporting information for chapter 5	v
Supporting information for chapter 6	lxv
RScript for gene coexpression analysis	lxvii

Supporting information for chapter 3

Table S1: Top 20 potential off-target sites

Potential Off-target Gene	Locus Description	Off-target Score	Mismatch Count	Mismatch Position
TM6SF1	intron	1.3520939	3	*.*.*.....
RP11-547D23.1-AL163953.3	intergenic	1.2992480	4	*.*.....*.*.....
RNA5SP76-SLC30A10	intergenic	0.7972059	4	*.....*.*.*.....
CTD-2076M15.1-ACTRT1	intergenic	0.7673077	2	*.....*
AC008069.1-AC080094.1	intergenic	0.7039157	4	*..*.*.*.....
CREBRF	intron	0.6510065	4	..*.*.*.....*
RP11-397A16.3-RP11-214L13.1	intergenic	0.6510065	4	..***.....*
ERC2-MIR3938	intergenic	0.6392300	4	..**.....*.*.....
RP11-230B22.1-ATG4C	intergenic	0.5616337	4	*.*.*.*.....
SLC35C1	exon	0.5423370	4	..*.*.*.*.....
RP11-83J21.3-LAMC3	intergenic	0.5357684	4	**.....*.*.....
DTHD1-RP11-431M7.2	intergenic	0.4852276	4	*.....*.*.....*
RP5-905H7.3	exon	0.4851653	4	..*.*.*.*.....
TLR4-RP11-281A20.2	intergenic	0.4804010	4	..*.*.*.*.....
FAM154A	intron	0.4593488	4	...**.....*.*.....
CNTFR	intron	0.4034128	4	...*.*.....*

Table S1 – continued from previous page

Potential Off-target Gene	Locus Description	Off-target Score	Mismatch Count	Mismatch Position
Y_RNA-FP236240.2	intergenic	0.3922253	4	. * ... * ... * *
PKNOX1-CBS	intergenic	0.3922253	4	. * ... * ... * *
RP11-154H17.1-AL356693.1	intergenic	0.3870832	4	* * .. * *
RP11-810P8.1	intron	0.3358772	4	.. * ... * * * .

Abbreviations: * - Mismatch.

Supporting information for chapter 5

Table S2: Top 500 *TP53BP2* pan-cancer co-expressed genes

Gene	Mean Score	Co-expression	Cytoband	Rank
RNF169	0.5921381		11q13.4	1
RIF1	0.5807083		2q23.3	2
ROCK2	0.5793359		2p25.1	3
LATS1	0.5782472		6q25.1	4
AFF4	0.5756305		5q31.1	5
PPP4R2	0.5729126		3p13	6
STAG1	0.5698688		3q22.3	7
REST	0.5663049		4q12	8
JMJD1C	0.5616653		10q21.3	9
STRN	0.5604258		2p22.2	10
ASXL2	0.5585594		2p23.3	11
SCYL2	0.5535308		12q23.1	12
TAOK1	0.5522611		17q11.2	13
ROCK1	0.5512954		18q11.1	14
ASH1L	0.5510979		1q22	15
SBNO1	0.5499641		12q24.31	16
PIK3CA	0.5485517		3q26.32	17
RSF1	0.5479929		11q14.1	18
HIPK3	0.5473325		11p13	19
NCOA2	0.5447782		8q13.3	20
BPTF	0.5443881		17q24.2	21
IL6ST	0.5440402		5q11.2	22
ZNF281	0.5434938		1q32.1	23
NUP153	0.5426259		6p22.3	24
PTPN11	0.5422216		12q24.13	25
ADNP	0.5414924		20q13.13	26
TTC37	0.5413579		5q15	27
NBPF10	0.5407785		1q21.1	28
RAPGEF6	0.5400620		5q31.1	29
FNIP2	0.5392465		4q32.1	30
EPC1	0.5388824		10p11.22	31
ATF2	0.5379708		2q31.1	32
TNKS2	0.5372014		10q23.32	33

Continued on next page

Table S2 – *Continued from previous page*

DOCK7	0.5350482	1p31.3	34
USP12	0.5344868	13q12.13	35
NR2C2	0.5337032	3p25.1	36
SHPRH	0.5330658	6q24.3	37
KIDINS220	0.5329284	2p25.1	38
ADAM17	0.5304614	2p25.1	39
PPTC7	0.5292244	12q24.11	40
GPATCH8	0.5285114	17q21.31	41
CELF1	0.5279963	11p11.2	42
MED13	0.5276039	17q23.2	43
ANKRD17	0.5268508	4q13.3	44
PIKFYVE	0.5268237	2q34	45
KIF2A	0.5261479	5q12.1	46
DPP8	0.5260814	15q22.31	47
EIF3A	0.5255133	10q26.11	48
TOR1AIP2	0.5237682	1q25.2	49
SLK	0.5236964	10q24.33	50
RBM27	0.5235468	5q32	51
HIPK1	0.5228119	1p13.2	52
SP3	0.5225223	2q31.1	53
DHX9	0.5219084	1q25.3	54
TJP1	0.5211016	15q13.1	55
LMTK2	0.5210116	7q21.3	56
CTTNBP2NL	0.5206032	1p13.2	57
ADAM10	0.5205644	15q21.3	58
SMC3	0.5200643	10q25.2	59
KIAA1109	0.5196084	4q27	60
TGFBRAP1	0.5191508	2q12.2	61
BIRC6	0.5187668	2p22.3	62
RASA1	0.5180771	5q14.3	63
DHX33	0.5178873	17p13.2	64
ZDHHC20	0.5178768	13q12.11	65
C10orf118	0.5173517	10q25.3	66
C11orf30	0.5168695	11q13.5	67
EP300	0.5163407	22q13.2	68
DPY19L3	0.5160842	19q13.11	69

Continued on next page

Table S2 – *Continued from previous page*

SON	0.5157551	21q22.11	70
KCTD20	0.5149463	6p21.31	71
DDX21	0.5138593	10q22.1	72
PDZD8	0.5133198	10q26.11	73
DDX3X	0.5132691	Xp11.4	74
MED13L	0.5122569	12q24.21	75
HECTD1	0.5122366	14q12	76
USP34	0.5114262	2p15	77
PBRM1	0.5106079	3p21.1	78
GMCL1	0.5099372	2p13.3	79
KIAA2018	0.5089049	3q13.2	80
PIK3C2A	0.5078038	11p15.1	81
MIB1	0.5076834	18q11.2	82
RBBP5	0.5076350	1q32.1	83
ZFP91	0.5061379	11q12.1	84
SMARCA5	0.5058647	4q31.21	85
MED1	0.5050768	17q12	86
SMG1	0.5045590	16p12.3	87
PAFAH1B2	0.5044229	11q23.3	88
TET3	0.5042387	2p13.1	89
TPR	0.5035451	1q31.1	90
NIPBL	0.5028392	5p13.2	91
STT3B	0.5027309	3p23	92
KIAA0947	0.5018417	5p15.32	93
LRCH3	0.5015524	3q29	94
CBL	0.5014996	11q23.3	95
SBF2	0.5014309	11p15.4	96
KDM3B	0.5014053	5q31.2	97
TEAD1	0.5012230	11p15.3	98
ARHGEF12	0.5010221	11q23.3	99
KLHDC10	0.5000958	7q32.2	100
HIVEP1	0.5000016	6p24.1	101
EPC2	0.4999442	2q23.1	102
BMPR2	0.4999121	2q33.1	103
SETD2	0.4997548	3p21.31	104
GTF2I	0.4996153	7q11.23	105

Continued on next page

Table S2 – *Continued from previous page*

ERBB2IP	0.4989473	5q12.3	106
GTF3C4	0.4987735	9q34.13	107
REV3L	0.4986784	6q21	108
APC	0.4984853	5q22.2	109
MGAT5	0.4984086	2q21.2	110
PHF3	0.4981708	6q12	111
TNPO3	0.4973046	7q32.1	112
RSPRY1	0.4972594	16q13	113
TRIM44	0.4970913	11p13	114
MDN1	0.4970459	6q15	115
TRIO	0.4968264	5p15.2	116
ARID1A	0.4959874	1p36.11	117
SEC24B	0.4957138	4q25	118
ATXN1L	0.4956665	16q22.2	119
CREB1	0.4953412	2q33.3	120
UBXN7	0.4949896	3q29	121
VPS13B	0.4944550	8q22.2	122
MAP3K2	0.4939294	2q14.3	123
RC3H2	0.4937183	9q33.2	124
BACH1	0.4932682	21q21.3	125
FBXW11	0.4915440	5q35.1	126
GOLIM4	0.4902349	3q26.2	127
SF3B1	0.4894388	2q33.1	128
RNF111	0.4893858	15q22.1	129
TAF2	0.4892536	8q24.12	130
ARID1B	0.4889297	6q25.3	131
MYO9A	0.4888010	15q23	132
ZCCHC11	0.4883865	1p32.3	133
DDX46	0.4879949	5q31.1	134
ZFR	0.4872318	5p13.3	135
PRKAR2A	0.4870410	3p21.31	136
ARL5B	0.4868633	10p12.31	137
CEP170	0.4865164	1q43	138
BDP1	0.4863926	5q13.2	139
BRWD3	0.4860432	Xq21.1	140
RRAGC	0.4854127	1p34.3	141

Continued on next page

Table S2 – *Continued from previous page*

RALGAPA2	0.4838087	20p11.23	142
HEATR5B	0.4825220	2p22.2	143
YLPM1	0.4818035	14q24.3	144
SKIL	0.4817857	3q26.2	145
SIRT1	0.4815192	10q21.3	146
SETX	0.4807559	9q34.13	147
PLEKHM3	0.4806224	2q33.3	148
GON4L	0.4804385	1q22	149
CASP8AP2	0.4804317	6q15	150
IREB2	0.4803165	15q25.1	151
RBM12	0.4797167	20q11.22	152
WDR26	0.4794681	1q42.12	153
IDE	0.4790717	10q23.33	154
TCF12	0.4786403	15q21.3	155
SERINC5	0.4785372	5q14.1	156
LIMS1	0.4781098	2q12.3	157
TAB2	0.4774804	6q25.1	158
RAB3GAP1	0.4769593	2q21.3	159
RLF	0.4769205	1p34.2	160
ELK3	0.4764648	12q23.1	161
NPAT	0.4764483	11q22.3	162
TOP1	0.4760689	20q12	163
CLCN3	0.4754765	4q33	164
PTPLB	0.4752735	3q21.1	165
NUP98	0.4752564	11p15.4	166
MKL2	0.4749574	16p13.12	167
HELZ	0.4745980	17q24.2	168
ZC3H13	0.4742654	13q14.13	169
ZNF644	0.4741834	1p22.2	170
ZKSCAN1	0.4741332	7q22.1	171
TAF1	0.4741299	Xq13.1	172
TRRAP	0.4735517	7q22.1	173
FAM168A	0.4732789	11q13.4	174
CLOCK	0.4728778	4q12	175
TRIP12	0.4728732	2q36.3	176
WAC	0.4728593	10p12.1	177

Continued on next page

Table S2 – *Continued from previous page*

UBR5	0.4726048	8q22.3	178
NCOR1	0.4724496	17p11.2	179
PUM1	0.4716643	1p35.2	180
GOLGA4	0.4715018	3p22.2	181
SLC30A6	0.4709149	2p22.3	182
CAMSAP1	0.4704447	9q34.3	183
CCNI	0.4701720	4q21.1	184
RASSF3	0.4701226	12q14.2	185
WIPF2	0.4697157	17q21.2	186
ASAP1	0.4690040	8q24.22	187
ZNF609	0.4684839	15q22.31	188
SMURF2	0.4679489	17q24.1	189
MFHAS1	0.4678799	8p23.1	190
PDPK1	0.4677712	16p13.3	191
CLPX	0.4674317	15q22.31	192
KDM5A	0.4664272	12p13.33	193
MFAP3	0.4663651	5q33.2	194
MAML2	0.4660614	11q21	195
DENND4A	0.4656137	15q22.31	196
SEC23IP	0.4655755	10q26.11	197
WAPAL	0.4654965	10q23.2	198
DOCK1	0.4654122	10q26.2	199
ARID4B	0.4652658	1q42.3	200
OTUD4	0.4652434	4q31.21	201
RNF38	0.4650658	9p13.2	202
TOP2B	0.4647399	3p24.2	203
MAN2A1	0.4647366	5q21.3	204
PPP2R5E	0.4647240	14q23.2	205
USP32	0.4647043	17q23.2	206
ARIH1	0.4645188	15q24.1	207
GCC2	0.4644903	2q12.3	208
PRKDC	0.4644728	8q11.21	209
UHMK1	0.4644266	1q23.3	210
NF1	0.4643928	17q11.2	211
RNMT	0.4643607	18p11.21	212
DAAM1	0.4640994	14q23.1	213

Continued on next page

Table S2 – *Continued from previous page*

FAM115A	0.4639508	7q35	214
SP4	0.4635459	7p15.3	215
NCKAP1	0.4631390	2q32.1	216
SMCHD1	0.4631284	18p11.32	217
SECISBP2L	0.4629619	15q21.1	218
ZNF827	0.4627676	4q31.22	219
PJA2	0.4627574	5q21.3	220
SOS1	0.4626876	2p22.1	221
ALMS1	0.4625076	2p13.1	222
ANAPC1	0.4623260	2q13	223
UGCG	0.4622292	9q31.3	224
SMC5	0.4621129	9q21.12	225
MACF1	0.4618585	1p34.3	226
NCOA6	0.4612106	20q11.22	227
PRPF40A	0.4608375	2q23.3	228
TSC22D2	0.4607586	3q25.1	229
AKAP10	0.4604841	17p11.2	230
ANKRD50	0.4604452	4q28.1	231
FAM63B	0.4604397	15q21.3-q22.1	232
DIP2B	0.4600334	12q13.12	233
USP42	0.4599063	7p22.1	234
AKAP9	0.4593652	7q21.2	235
CHD6	0.4591730	20q12	236
FUBP1	0.4590918	1p31.1	237
CDK12	0.4584144	17q12	238
ZRANB1	0.4579024	10q26.13	239
SPEN	0.4576177	1p36.21	240
ACAP2	0.4575848	3q29	241
TTC17	0.4571271	11p12	242
CHD2	0.4568468	15q26.1	243
TRIP11	0.4568376	14q32.12	244
ZBTB10	0.4565638	8q21.13	245
YES1	0.4565479	18p11.32	246
APAF1	0.4562185	12q23.1	247
GSK3B	0.4560579	3q13.33	248
CEP350	0.4558588	1q25.2	249

Continued on next page

Table S2 – *Continued from previous page*

AKAP2	0.4557134	9q31.3	250
NEK4	0.4556329	3p21.1	251
NOTCH2	0.4555456	1p12	252
ARHGAP5	0.4554774	14q12	253
HIP1	0.4547418	7q11.23	254
HIVEP2	0.4546886	6q24.2	255
PHF20	0.4544887	20q11.22	256
LEPROT	0.4544270	1p31.3	257
DHX15	0.4542401	4p15.2	258
WNK1	0.4541768	12p13.33	259
SYNJ2	0.4534758	6q25.3	260
NCOA1	0.4531798	2p23.3	261
ZNF146	0.4529759	19q13.12	262
ARFGEF2	0.4528676	20q13.13	263
WDFY3	0.4528624	4q21.23	264
ACBD3	0.4528547	1q42.12	265
KLHL28	0.4528476	14q21.2	266
CAND1	0.4527141	12q14.3	267
ITCH	0.4523401	20q11.22	268
YAP1	0.4523354	11q22.1	269
SOCS5	0.4521322	2p21	270
ATF6	0.4519240	1q23.3	271
MAN1A2	0.4518114	1p12	272
PIK3R4	0.4514795	3q22.1	273
RAPGEF2	0.4512646	4q32.1	274
CTDSPL2	0.4511946	15q15.3	275
UPF2	0.4511560	10p14	276
KLHL8	0.4511063	4q22.1	277
OXSRI	0.4509341	3p22.2	278
FUBP3	0.4503895	9q34.11	279
FOXJ3	0.4499381	1p34.2	280
NUP160	0.4499335	11p11.2	281
PPP1R12A	0.4498188	12q21.31	282
SEC24A	0.4490687	5q31.1	283
USP47	0.4489641	11p15.3	284
MEX3C	0.4485418	18q21.2	285

Continued on next page

Table S2 – *Continued from previous page*

MARCH6	0.4484907	5p15.2	286
SPAG9	0.4482865	17q21.33	287
ATP11B	0.4480738	3q26.33	288
HNRNPR	0.4480510	1p36.12	289
RALGAPB	0.4479886	20q11.23	290
ESF1	0.4479827	20p12.1	291
VPS13C	0.4475367	15q22.2	292
RREB1	0.4475252	6p24.3	293
KIAA2026	0.4472004	9p24.1	294
PATL1	0.4468787	11q12.1	295
GNA13	0.4468746	17q24.1	296
DOCK9	0.4467118	13q32.3	297
SPTY2D1	0.4466949	11p15.1	298
PPFIA1	0.4465212	11q13.3	299
POLR2B	0.4462542	4q12	300
DICER1	0.4459550	14q32.13	301
ATAD2B	0.4459187	2p23.3	302
MYCBP2	0.4458332	13q22.3	303
MAP3K7	0.4443676	6q15	304
CDC73	0.4443239	1q31.2	305
VHL	0.4442464	3p25.3	306
EIF4G2	0.4437687	11p15.4	307
TMF1	0.4436370	3p14.1	308
OSBPL8	0.4436239	12q21.2	309
TNPO1	0.4436026	5q13.2	310
ANKRD28	0.4432089	3p25.1	311
KIRREL	0.4426094	1q23.1	312
ARID4A	0.4423671	14q23.1	313
DHX36	0.4421749	3q25.2	314
SPIN1	0.4419101	9q22.1	315
RANBP2	0.4417104	2q13	316
MORC3	0.4413302	21q22.12	317
CLASP1	0.4408314	2q14.3	318
KIF13A	0.4406343	6p22.3	319
APPBP2	0.4405457	17q23.2	320
BCLAF1	0.4404504	6q23.3	321

Continued on next page

Table S2 – *Continued from previous page*

AKAP11	0.4403568	13q14.11	322
ZBTB11	0.4403312	3q12.3	323
MATR3	0.4399439	5q31.2	324
ARHGAP21	0.4395554	10p12.1	325
CDK13	0.4394935	7p14.1	326
HIF1A	0.4389245	14q23.2	327
XRN2	0.4388822	20p11.22	328
TMOD3	0.4387356	15q21.2	329
FBXO11	0.4384147	2p16.3	330
KIF1B	0.4382909	1p36.22	331
SENP1	0.4377651	12q13.11	332
TOPORS	0.4372601	9p21.1	333
MKLN1	0.4371698	7q32.3	334
ERC1	0.4370317	12p13.33	335
CHD9	0.4370017	16q12.2	336
PRPF4B	0.4368032	6p25.2	337
MAPK1	0.4366388	22q11.22	338
AKAP13	0.4364149	15q25.3	339
ASCC3	0.4362266	6q16.3	340
FAM73A	0.4362171	1p31.1	341
UTP14C	0.4360203	13q14.3	342
MTF1	0.4357806	1p34.3	343
ALS2	0.4355715	2q33.1	344
SOS2	0.4352312	14q21.3	345
ZNF800	0.4348724	7q31.33	346
SPATA13	0.4348593	13q12.12	347
ZBTB43	0.4345365	9q33.3	348
HEATR1	0.4344707	1q43	349
MAP4K5	0.4340487	14q22.1	350
EAF1	0.4339956	3p25.1	351
CREBBP	0.4338254	16p13.3	352
SPOPL	0.4337127	2q22.1	353
IQGAP1	0.4335512	15q26.1	354
CKAP5	0.4333995	11p11.2	355
SPTBN1	0.4332758	2p16.2	356
ELF1	0.4332358	13q14.11	357

Continued on next page

Table S2 – *Continued from previous page*

SUZ12	0.4331575	17q11.2	358
CHD8	0.4331124	14q11.2	359
QSER1	0.4329463	11p13	360
ATXN7	0.4326012	3p14.1	361
WDR47	0.4324911	1p13.3	362
SMARCAD1	0.4323960	4q22.3	363
UBR1	0.4322227	15q15.2	364
BBX	0.4317893	3q13.12	365
ACTR2	0.4314381	2p14	366
USP6NL	0.4312224	10p14	367
G3BP2	0.4311209	4q21.1	368
TSTD2	0.4308250	9q22.33	369
ARL13B	0.4306385	3q11.1	370
ANKIB1	0.4305270	7q21.2	371
ZNF148	0.4298702	3q21.2	372
MAML1	0.4297139	5q35.3	373
C5orf51	0.4293370	5p13.1	374
IGF2R	0.4290524	6q25.3	375
NBPF9	0.4288576	1q21.2	376
POLR3A	0.4286350	10q22.3	377
TRIM33	0.4284984	1p13.2	378
CTNND1	0.4284178	11q12.1	379
LRBA	0.4284049	4q31.3	380
MGA	0.4282945	15q15.1	381
TRIM56	0.4282588	7q22.1	382
MARCH8	0.4282038	10q11.21- q11.22	383
GAPVD1	0.4281366	9q33.3	384
OSBPL11	0.4280364	3q21.2	385
LRRC58	0.4278730	3q13.33	386
CASC5	0.4276826	15q15.1	387
FAM168B	0.4273548	2q21.1	388
TOPBP1	0.4272982	3q22.1	389
AVL9	0.4272484	7p14.3	390
EXOC5	0.4272108	14q22.3	391
DNAJB14	0.4271676	4q23	392

Continued on next page

Table S2 – *Continued from previous page*

ZMYM2	0.4268776	13q12.11	393
UBE2H	0.4261793	7q32.2	394
NUP205	0.4259135	7q33	395
NEDD1	0.4258244	12q23.1	396
NUFIP2	0.4257799	17q11.2	397
VPRBP	0.4257112	3p21.2	398
ATP2A2	0.4256457	12q24.11	399
PAPOLG	0.4255103	2p16.1	400
YME1L1	0.4254922	10p12.1	401
FBXO38	0.4254527	5q32	402
SETD5	0.4254028	3p25.3	403
KIAA1430	0.4249433	4q35.1	404
KDM5B	0.4249268	1q32.1	405
PCNX	0.4247894	14q24.2	406
SAMD4B	0.4245388	19q13.2	407
SNX30	0.4234440	9q32	408
TBCEL	0.4231919	11q23.3	409
TET2	0.4230792	4q24	410
PKN2	0.4229317	1p22.2	411
ATF7IP	0.4228683	12p13.1	412
MPDZ	0.4228477	9p23	413
DLG1	0.4226264	3q29	414
NBPF14	0.4225762	1q21.2	415
UBR3	0.4224335	2q31.1	416
SLC25A24	0.4224294	1p13.3	417
VEZF1	0.4219516	17q22	418
TMEM184C	0.4216885	4q31.23	419
WDR11	0.4216305	10q26.12	420
RASAL2	0.4214674	1q25.2	421
LNPEP	0.4213053	5q15	422
NPEPPS	0.4212145	17q21.32	423
CLIP1	0.4211528	12q24.31	424
ZC3H7A	0.4208549	16p13.13	425
AQR	0.4208529	15q14	426
JAK1	0.4207981	1p31.3	427
ZNF623	0.4206443	8q24.3	428

Continued on next page

Table S2 – *Continued from previous page*

PPP1R15B	0.4205218	1q32.1	429
MOSPD2	0.4204830	Xp22.2	430
ZFX	0.4203030	Xp22.11	431
USP24	0.4202507	1p32.3	432
NFKB1	0.4202445	4q24	433
ITSN2	0.4200210	2p23.3	434
EPB41L5	0.4196935	2q14.2	435
DCP1A	0.4194783	3p21.1	436
ATP13A3	0.4194251	3q29	437
SCAPER	0.4193898	15q24.3	438
DYRK2	0.4193107	12q15	439
SMAD5	0.4190685	5q31.1	440
HEATR5A	0.4186380	14q12	441
UHRF1BP1L	0.4185025	12q23.1	442
ZNF292	0.4182739	6q14.3	443
KLF3	0.4176561	4p14	444
PTAR1	0.4171282	9q21.12	445
NFAT5	0.4170989	16q22.1	446
SIN3A	0.4164397	15q24.2	447
RAB3GAP2	0.4163679	1q41	448
RB1CC1	0.4163172	8q11.23	449
TAB3	0.4160966	Xp21.2	450
RPRD1B	0.4160157	20q11.23	451
EIF2AK3	0.4159788	2p11.2	452
GIGYF2	0.4157338	2q37.1	453
INTS2	0.4150933	17q23.2	454
PIAS1	0.4145771	15q23	455
PDS5B	0.4141426	13q13.1	456
CHD1	0.4141272	5q21.1	457
FBXW2	0.4141168	9q33.2	458
FEM1B	0.4138843	15q23	459
UBR2	0.4138765	6p21.1	460
ATF7	0.4138725	12q13.13	461
KLHL20	0.4136415	1q25.1	462
KIAA1432	0.4136405	9p24.1	463
CLINT1	0.4135749	5q33.3	464

Continued on next page

Table S2 – *Continued from previous page*

FRS2	0.4134652	12q15	465
BAZ1B	0.4133592	7q11.23	466
HIAT1	0.4132540	1p21.2	467
MRE11A	0.4131695	11q21	468
LEMD3	0.4130422	12q14.3	469
TGOLN2	0.4129033	2p11.2	470
MLLT10	0.4127274	10p12.31	471
RBAK	0.4126461	7p22.1	472
AKIRIN1	0.4124356	1p34.3	473
TNRC6B	0.4121954	22q13.1	474
CWC22	0.4119366	2q31.3	475
KDM3A	0.4118207	2p11.2	476
BTAF1	0.4118099	10q23.32	477
TP53BP1	0.4116297	15q15.3	478
VPS54	0.4116016	2p14	479
ZNF507	0.4115637	19q13.11	480
THRAP3	0.4112485	1p34.3	481
MBD5	0.4112331	2q23.1	482
JMY	0.4111545	5q14.1	483
MED14	0.4111036	Xp11.4	484
NEK1	0.4109425	4q33	485
ZNF710	0.4107991	15q26.1	486
PACSIN2	0.4103214	22q13.2	487
RPS6KA3	0.4102412	Xp22.12	488
RYBP	0.4101136	3p13	489
ATRX	0.4100063	Xq21.1	490
KIF5B	0.4098426	10p11.22	491
CDKN2AIP	0.4097860	4q35.1	492
MEF2A	0.4097672	15q26.3	493
AHCTF1	0.4095784	1q44	494
ANO6	0.4094313	12q12	495
GXYLT1	0.4094206	12q12	496
TCF20	0.4092783	22q13.2	497
SSH1	0.4089101	12q24.11	498
FNBP1L	0.4088215	1p22.1	499
ZMYM4	0.4086688	1p34.3	500

Continued on next page

Table S2 – *Continued from previous page*

Table S3: Bottom 500 *TP53BP2* pan-cancer co-expressed genes

Gene	Mean Co-expression Score	Cytoband	Rank
TRAPPC5	-0.5641108	19p13.2	9601
NDUFB7	-0.5556812	19p13.12	9600
GADD45GIP1	-0.5545619	19p13.13	9599
ATP5D	-0.5442990	19p13.3	9598
NDUFA11	-0.5420726	19p13.3	9597
MRPL54	-0.5334826	19p13.3	9596
NDUFA13	-0.5282361	19p13.11	9595
TCEB2	-0.5275904	16p13.3	9594
FAU	-0.5230529	11q13.1	9593
C16orf13	-0.5223201	16p13.3	9592
C19orf70	-0.5203293	19p13.3	9591
NDUFA2	-0.5199268	5q31.3	9590
NDUFA3	-0.5194957	19q13.42	9589
C19orf24	-0.5168080	19p13.3	9588
ZNHIT1	-0.5152304	7q22.1	9587
BLOC1S1	-0.5134502	12q13.2	9586
MRPL41	-0.5124603	9q34.3	9585
NDUFS8	-0.5123982	11q13.2	9584
SERF2	-0.5102215	15q15.3	9583
NDUFA7	-0.5098475	19p13.2	9582
C19orf53	-0.5080393	19p13.13	9581
MRPL27	-0.5079228	17q21.33	9580
SSNA1	-0.5052770	9q34.3	9579
UQCR10	-0.5046744	22q12.2	9578
MRPL55	-0.5046701	1q42.13	9577
AURKAIP1	-0.5032636	1p36.33	9576
ROMO1	-0.5028424	20q11.22	9575
NDUFB10	-0.5022617	16p13.3	9574
EDF1	-0.5013114	9q34.3	9573
MRPL23	-0.5004285	11p15.5	9572
NDUFB1	-0.5002707	14q32.12	9571
NAA10	-0.4979542	Xq28	9570
DPM3	-0.4976588	1q22	9569

Continued on next page

Table S3 – *Continued from previous page*

TMEM219	-0.4959423	16p11.2	9568
NDUFS7	-0.4940351	19p13.3	9567
MYEOV2	-0.4925622	2q37.3	9566
KRTCAP2	-0.4912874	1q22	9565
ARL6IP4	-0.4905720	12q24.31	9564
COPE	-0.4898722	19p13.11	9563
NDUFB11	-0.4890141	Xp11.3	9562
THAP3	-0.4889335	1p36.31	9561
POLR2I	-0.4886406	19q13.12	9560
LSMD1	-0.4885840	17p13.1	9559
TEX264	-0.4877133	3p21.2	9558
C11orf83	-0.4875968	11q12.3	9557
FAM195A	-0.4872934	16p13.3	9556
COX6B1	-0.4850836	19q13.12	9555
UQCR11	-0.4841950	19p13.3	9554
UQCRQ	-0.4838159	5q31.1	9553
MRPL52	-0.4832794	14q11.2	9552
THAP7	-0.4828180	22q11.21	9551
TRAPPC2L	-0.4821887	16q24.3	9550
MRPS34	-0.4817534	16p13.3	9549
MRPL53	-0.4813386	2p13.1	9548
PGLS	-0.4810213	19p13.11	9547
COX5B	-0.4797233	2q11.2	9546
COMMD4	-0.4795462	15q24.2	9545
ZNF524	-0.4791914	19q13.42	9544
PHPT1	-0.4785045	9q34.3	9543
POLR2L	-0.4776084	11p15.5	9542
CISD3	-0.4766362	17q12	9541
SSR4	-0.4765244	Xq28	9540
UXT	-0.4762390	Xp11.23	9539
C9orf142	-0.4752755	9q34.3	9538
PQBP1	-0.4739485	Xp11.23	9537
INO80E	-0.4734071	16p11.2	9536
RPS15	-0.4731317	19p13.3	9535
TMEM208	-0.4715869	16q22.1	9534
UBL5	-0.4707101	19p13.2	9533

Continued on next page

Table S3 – *Continued from previous page*

ZNF593	-0.4701268	1p36.11	9532
ZMAT5	-0.4699680	22q12.2	9531
ZDHHC24	-0.4687858	11q13.2	9530
NUBP2	-0.4682254	16p13.3	9529
MPG	-0.4678713	16p13.3	9528
APRT	-0.4671703	16q24.3	9527
RPL35	-0.4670901	9q33.3	9526
NUDT22	-0.4669999	11q13.1	9525
MIF	-0.4669650	22q11.23	9524
DGCR6L	-0.4660794	22q11.21	9523
C19orf60	-0.4656032	19p13.11	9522
ALKBH7	-0.4655108	19p13.3	9521
NOSIP	-0.4651439	19q13.33	9520
RNF181	-0.4649175	2p11.2	9519
TIMM13	-0.4640496	19p13.3	9518
ATP5I	-0.4636705	4p16.3	9517
ISOC2	-0.4636281	19q13.42	9516
POLR2F	-0.4631898	22q13.1	9515
COQ4	-0.4631237	9q34.11	9514
SCAND1	-0.4627150	20q11.23	9513
NDUFB8	-0.4624796	10q24.31	9512
RNASEK	-0.4623730	17p13.1	9511
C19orf43	-0.4622973	19p13.13	9510
ATP5G1	-0.4620551	17q21.32	9509
FKBP2	-0.4617445	11q13.1	9508
RPS19BP1	-0.4609539	22q13.1	9507
RPL36	-0.4603301	19p13.3	9506
BCL7C	-0.4593718	16p11.2	9505
TRAPPC1	-0.4593196	17p13.1	9504
RPL18	-0.4589649	19q13.33	9503
FBXL15	-0.4576389	10q24.32	9502
HIGD2A	-0.4575681	5q35.2	9501
FAM96B	-0.4573062	16q22.1	9500
FDX1L	-0.4570879	19p13.2	9499
RPP21	-0.4568937	6p22.1	9498
TIMM17B	-0.4564933	Xp11.23	9497

Continued on next page

Table S3 – *Continued from previous page*

PSMB3	-0.4564569	17q12	9496
RNASEH2C	-0.4563633	11q13.1	9495
LSM7	-0.4559117	19p13.3	9494
COX8A	-0.4557513	11q13.1	9493
C14orf2	-0.4542583	14q32.33	9492
COX4I1	-0.4527902	16q24.1	9491
RABAC1	-0.4523293	19q13.2	9490
ELOF1	-0.4519743	19p13.2	9489
CHMP2A	-0.4516279	19q13.43	9488
SSSCA1	-0.4514717	11q13.1	9487
NDUFS3	-0.4514490	11p11.2	9486
MRPL40	-0.4514282	22q11.21	9485
MRPS11	-0.4509840	15q25.3	9484
MRPS15	-0.4509178	1p34.3	9483
COX6A1	-0.4508026	12q24.31	9482
RPL13	-0.4499553	16q24.3	9481
SNRPD2	-0.4499470	19q13.32	9480
LSM4	-0.4490155	19p13.11	9479
RFXANK	-0.4474498	19p13.11	9478
PEX16	-0.4473430	11p11.2	9477
TAF10	-0.4471897	11p15.4	9476
CCDC107	-0.4466685	9p13.3	9475
RPL28	-0.4464991	19q13.42	9474
AP2S1	-0.4456811	19q13.32	9473
PGP	-0.4455567	16p13.3	9472
ETFB	-0.4455141	19q13.41	9471
MRPL38	-0.4450694	17q25.1	9470
TRMT112	-0.4449604	11q13.1	9469
ANAPC11	-0.4447462	17q25.3	9468
PSMG3	-0.4446981	7p22.3	9467
PRELID1	-0.4441475	5q35.3	9466
BOLA2	-0.4437778	16p11.2	9465
MRPS12	-0.4435739	19q13.2	9464
MRPL24	-0.4430825	1q23.1	9463
MRPL12	-0.4429470	17q25.3	9462
RPLP2	-0.4426016	11p15.5	9461

Continued on next page

Table S3 – *Continued from previous page*

TRPT1	-0.4423744	11q13.1	9460
NENF	-0.4416779	1q32.3	9459
ARF5	-0.4411779	7q32.1	9458
TXN2	-0.4411055	22q12.3	9457
PFDN5	-0.4407517	12q13.13	9456
NME3	-0.4407042	16p13.3	9455
SAC3D1	-0.4405316	11q13.1	9454
JOSD2	-0.4392652	19q13.33	9453
SIRT6	-0.4372611	19p13.3	9452
SF3B5	-0.4363606	6q24.2	9451
VPS28	-0.4363157	8q24.3	9450
ATP5G2	-0.4361808	12q13.13	9449
PSMB10	-0.4357993	16q22.1	9448
HSD17B10	-0.4354094	Xp11.22	9447
MRPS24	-0.4352239	7p13	9446
SLC25A39	-0.4351586	17q21.31	9445
COMTD1	-0.4350475	10q22.2	9444
POLR2J	-0.4347688	7q22.1	9443
DRAP1	-0.4343935	11q13.1	9442
UROD	-0.4340142	1p34.1	9441
GPX4	-0.4339959	19p13.3	9440
DYNLRB1	-0.4337526	20q11.22	9439
C12orf10	-0.4325812	12q13.13	9438
BANF1	-0.4319769	11q13.1	9437
FAM58A	-0.4316278	Xq28	9436
ATP6V0B	-0.4312847	1p34.1	9435
NSMCE1	-0.4306475	16p12.1	9434
RPS9	-0.4300853	19q13.42	9433
CCDC101	-0.4295082	16p11.2	9432
DCXR	-0.4294966	17q25.3	9431
FIS1	-0.4294000	7q22.1	9430
CCDC124	-0.4289737	19p13.11	9429
BRMS1	-0.4289299	11q13.2	9428
C7orf55	-0.4283881	7q34	9427
POP7	-0.4280776	7q22.1	9426
TMEM134	-0.4277986	11q13.2	9425

Continued on next page

Table S3 – *Continued from previous page*

PRDX5	-0.4277560	11q13.1	9424
MRP63	-0.4276414	13q12.11	9423
MTX1	-0.4276257	1q22	9422
ABHD14A	-0.4272004	3p21.2	9421
BAD	-0.4271790	11q13.1	9420
ATPIF1	-0.4270048	1p35.3	9419
YIF1A	-0.4269659	11q13.2	9418
CCDC12	-0.4264989	3p21.31	9417
MRPS26	-0.4263270	20p13	9416
GPX1	-0.4260913	3p21.31	9415
POP5	-0.4258143	12q24.31	9414
SURF2	-0.4248410	9q34.2	9413
DPM2	-0.4245511	9q34.11	9412
ATP5J2	-0.4244595	7q22.1	9411
ENDOG	-0.4243472	9q34.11	9410
SIVA1	-0.4237592	14q32.33	9409
USE1	-0.4235686	19p13.11	9408
CLPP	-0.4232917	19p13.3	9407
NPRL2	-0.4230708	3p21.31	9406
NDUFA1	-0.4224496	Xq24	9405
C19orf25	-0.4216275	19p13.3	9404
NDUFS5	-0.4215660	1p34.3	9403
MRPL28	-0.4212755	16p13.3	9402
CCDC94	-0.4210663	19p13.3	9401
HCFC1R1	-0.4205491	16p13.3	9400
CHCHD5	-0.4204128	2q14.1	9399
MRPL20	-0.4200068	1p36.33	9398
RPL29	-0.4199826	3p21.2	9397
NDUFA8	-0.4197494	9q33.2	9396
EIF3K	-0.4197170	19q13.2	9395
CHCHD10	-0.4196307	22q11.23	9394
SEC61B	-0.4194693	9q22.33	9393
OAZ1	-0.4193966	19p13.3	9392
STRA13	-0.4191679	17q25.3	9391
MSRB2	-0.4181846	10p12.2	9390
NHP2	-0.4177708	5q35.3	9389

Continued on next page

Table S3 – *Continued from previous page*

RPL39	-0.4175169	Xq24	9388
ATP5H	-0.4173104	17q25.1	9387
STUB1	-0.4172070	16p13.3	9386
FASTK	-0.4171406	7q36.1	9385
SDF2L1	-0.4167349	22q11.21	9384
PSENEN	-0.4158080	19q13.12	9383
UBA52	-0.4154379	19p13.11	9382
LSM10	-0.4149271	1p34.3	9381
RUVBL2	-0.4146340	19q13.33	9380
TMEM223	-0.4142955	11q12.3	9379
IMP3	-0.4137866	15q24.2	9378
GUK1	-0.4137827	1q42.13	9377
C17orf89	-0.4130611	17q25.3	9376
AIP	-0.4127063	11q13.2	9375
ATOX1	-0.4126674	5q33.1	9374
TREX1	-0.4126448	3p21.31	9373
NDUFB2	-0.4123957	7q34	9372
OST4	-0.4120470	2p23.3	9371
NT5C	-0.4115632	17q25.1	9370
SRA1	-0.4112643	5q31.3	9369
MAP2K2	-0.4109130	19p13.3	9368
TIMM8B	-0.4108383	11q23.1	9367
SNRNP25	-0.4103937	16p13.3	9366
RPLP1	-0.4103365	15q23	9365
NTHL1	-0.4099301	16p13.3	9364
SURF1	-0.4096324	9q34.2	9363
LAGE3	-0.4094270	Xq28	9362
DHRS4L2	-0.4090447	14q11.2	9361
MRPL43	-0.4089932	10q24.31	9360
MRPL17	-0.4082907	11p15.4	9359
C9orf16	-0.4079667	9q34.11	9358
CHCHD1	-0.4073579	10q22.2	9357
RPL38	-0.4072974	17q25.1	9356
FAM3A	-0.4070733	Xq28	9355
FAM86A	-0.4069858	16p13.3	9354
MRPL34	-0.4068495	19p13.11	9353

Continued on next page

Table S3 – *Continued from previous page*

SCO2	-0.4068439	22q13.33	9352
HINT2	-0.4066612	9p13.3	9351
CDK2AP2	-0.4058073	11q13.2	9350
TMEM141	-0.4057742	9q34.3	9349
NOL12	-0.4054700	22q13.1	9348
BUD31	-0.4052979	7q22.1	9347
OCEL1	-0.4052175	19p13.11	9346
HAUS7	-0.4051402	Xq28	9345
BCAP31	-0.4049974	Xq28	9344
QTRT1	-0.4048869	19p13.2	9343
CINP	-0.4046469	14q32.31	9342
EXOSC4	-0.4045887	8q24.3	9341
NME2	-0.4045668	17q21.33	9340
WDR74	-0.4044324	11q12.3	9339
POLD4	-0.4041882	11q13.2	9338
NDUFS6	-0.4040684	5p15.33	9337
MRPL2	-0.4039518	6p21.1	9336
GEMIN7	-0.4028964	19q13.32	9335
SNRPA	-0.4024144	19q13.2	9334
NARFL	-0.4020480	16p13.3	9333
NAPA	-0.4020204	19q13.33	9332
COX7B	-0.4016453	Xq21.1	9331
C11orf48	-0.4009791	11q12.3	9330
C11orf31	-0.4006603	11q12.1	9329
TRAPPC6A	-0.3995635	19q13.32	9328
PHKG2	-0.3995547	16p11.2	9327
RPS19	-0.3994902	19q13.2	9326
AUP1	-0.3990953	2p13.1	9325
CCS	-0.3987468	11q13.2	9324
NEDD8	-0.3974281	14q12	9323
PPP1CA	-0.3973033	11q13.2	9322
C1orf86	-0.3972648	1p36.33	9321
CUTA	-0.3972053	6p21.32	9320
ZFPL1	-0.3970628	11q13.1	9319
DNAJC4	-0.3969986	11q13.1	9318
DPP7	-0.3967389	9q34.3	9317

Continued on next page

Table S3 – *Continued from previous page*

RPL27A	-0.3963595	11p15.4	9316
C12orf57	-0.3962860	12p13.31	9315
RER1	-0.3959966	1p36.32	9314
GFER	-0.3954864	16p13.3	9313
NMRAL1	-0.3954411	16p13.3	9312
CD320	-0.3949268	19p13.2	9311
ACOT8	-0.3949059	20q13.12	9310
B4GALT7	-0.3948043	5q35.3	9309
TIMM10	-0.3945732	11q12.1	9308
RBX1	-0.3937031	22q13.2	9307
RPS21	-0.3933171	20q13.33	9306
TMEM147	-0.3925820	19q13.12	9305
RBM42	-0.3915576	19q13.12	9304
MRPL21	-0.3911665	11q13.3	9303
SIL1	-0.3911585	5q31.2	9302
ATP5E	-0.3904411	20q13.32	9301
C6orf1	-0.3901266	6p21.31	9300
DDT	-0.3900013	22q11.23	9299
DUS1L	-0.3898122	17q25.3	9298
FKBP8	-0.3895462	19p13.11	9297
ATP6V1F	-0.3894778	7q32.1	9296
OGFOD2	-0.3888309	12q24.31	9295
YDJC	-0.3888112	22q11.21	9294
PPIB	-0.3886904	15q22.31	9293
WDR13	-0.3880790	Xp11.23	9292
ZFAND2B	-0.3874918	2q35	9291
TSSC4	-0.3873687	11p15.5	9290
EBP	-0.3872752	Xp11.23	9289
CNPY2	-0.3868657	12q13.3	9288
IDH3G	-0.3868386	Xq28	9287
UQCRC1	-0.3865366	3p21.31	9286
CCDC85B	-0.3862910	11q13.1	9285
HSPBP1	-0.3862315	19q13.42	9284
USMG5	-0.3861251	10q24.33	9283
B3GAT3	-0.3861127	11q12.3	9282
NDUFAF3	-0.3860675	3p21.31	9281

Continued on next page

Table S3 – *Continued from previous page*

MPST	-0.3860147	22q12.3	9280
TXNL4A	-0.3859561	18q23	9279
WIBG	-0.3852385	12q13.2	9278
NDUFC1	-0.3838132	4q31.1	9277
DGUOK	-0.3836524	2p13.1	9276
DDX49	-0.3824226	19p13.11	9275
PSME2	-0.3821816	14q12	9274
RPS16	-0.3821572	19q13.2	9273
PSMB4	-0.3820760	1q21.3	9272
RPL36AL	-0.3818839	14q21.3	9271
PIN1	-0.3811626	19p13.2	9270
C1orf122	-0.3810605	1p34.3	9269
RAB4B	-0.3803320	19q13.2	9268
SLC39A3	-0.3801893	19p13.3	9267
PFDN6	-0.3800525	6p21.32	9266
CAPNS1	-0.3800061	19q13.12	9265
PSMB6	-0.3799525	17p13.2	9264
TIMM50	-0.3793927	19q13.2	9263
MMAB	-0.3793036	12q24.11	9262
RNF25	-0.3791721	2q35	9261
RANGRF	-0.3787405	17p13.1	9260
SNRNP35	-0.3786240	12q24.31	9259
PTGES2	-0.3783233	9q34.11	9258
OVCA2	-0.3782694	17p13.3	9257
TMEM222	-0.3780903	1p36.11	9256
C19orf10	-0.3775037	19p13.3	9255
MXD3	-0.3773229	5q35.3	9254
TMUB1	-0.3757885	7q36.1	9253
UBE2M	-0.3755773	19q13.43	9252
GMPPA	-0.3755725	2q35	9251
DHPS	-0.3750474	19p13.13	9250
TFPT	-0.3748261	19q13.42	9249
RPL32	-0.3748169	3p25.2	9248
CIB1	-0.3747505	15q26.1	9247
PSMB1	-0.3746642	6q27	9246
SHARPIN	-0.3746477	8q24.3	9245

Continued on next page

Table S3 – *Continued from previous page*

RPL34	-0.3746420	4q25	9244
ASPSCR1	-0.3745454	17q25.3	9243
AES	-0.3740742	19p13.3	9242
PHB	-0.3732951	17q21.33	9241
RAD23A	-0.3731848	19p13.13	9240
COX6C	-0.3727864	8q22.2	9239
ADRM1	-0.3726794	20q13.33	9238
MPV17L2	-0.3725629	19p13.11	9237
ZNF511	-0.3714669	10q26.3	9236
RPS14	-0.3710378	5q33.1	9235
MLST8	-0.3709196	16p13.3	9234
NOC4L	-0.3708756	12q24.33	9233
RPS11	-0.3706258	19q13.33	9232
C7orf50	-0.3706039	7p22.3	9231
PPIA	-0.3705874	7p13	9230
MCAT	-0.3696448	22q13.2	9229
ATP5L	-0.3693684	11q23.3	9228
COX7A2	-0.3690729	6q14.1	9227
TMEM11	-0.3689157	17p11.2	9226
FLYWCH2	-0.3688459	16p13.3	9225
CDK5	-0.3684701	7q36.1	9224
EIF3G	-0.3683133	19p13.2	9223
RPS2	-0.3683031	16p13.3	9222
GSTZ1	-0.3682252	14q24.3	9221
ORMDL2	-0.3681309	12q13.2	9220
SHFM1	-0.3680346	7q21.3	9219
NDUFV1	-0.3678734	11q13.2	9218
FIBP	-0.3677544	11q13.1	9217
ZDHHC12	-0.3676974	9q34.11	9216
RPL19	-0.3674865	17q12	9215
ECHS1	-0.3674730	10q26.3	9214
APOA1BP	-0.3673729	1q22	9213
DDX56	-0.3672295	7p13	9212
COPS6	-0.3671173	7q22.1	9211
PMVK	-0.3670051	1q21.3	9210
HRAS	-0.3670046	11p15.5	9209

Continued on next page

Table S3 – *Continued from previous page*

PPP4C	-0.3667293	16p11.2	9208
NDUFA6	-0.3658558	22q13.2	9207
C1orf123	-0.3657272	1p32.3	9206
SDHAF2	-0.3647982	11q12.2	9205
ZNF580	-0.3645921	19q13.42	9204
FAM50A	-0.3645639	Xq28	9203
CLTB	-0.3644755	5q35.2	9202
RPUSD1	-0.3643994	16p13.3	9201
URM1	-0.3643912	9q34.11	9200
NOP10	-0.3638580	15q14	9199
NUTF2	-0.3638468	16q22.1	9198
RPL18A	-0.3636283	19p13.11	9197
VKORC1	-0.3635178	16p11.2	9196
MRPS21	-0.3627950	1q21.2	9195
SLC25A28	-0.3627081	10q24.2	9194
RPL8	-0.3619390	8q24.3	9193
DNAJC30	-0.3617379	7q11.23	9192
TIAF1	-0.3615052	17q11.2	9191
THOC6	-0.3614634	16p13.3	9190
EIF5A	-0.3613222	17p13.1	9189
POLR2H	-0.3613037	3q27.1	9188
TPRA1	-0.3612156	3q21.3	9187
PARK7	-0.3608474	1p36.23	9186
G6PC3	-0.3605289	17q21.31	9185
NR2C2AP	-0.3600439	19p13.11	9184
ICT1	-0.3592420	17q25.1	9183
EIF4EBP3	-0.3591204	5q31.3	9182
E4F1	-0.3590018	16p13.3	9181
IMP4	-0.3588037	2q21.1	9180
PIN4	-0.3583728	Xq13.1	9179
TRMU	-0.3583041	22q13.31	9178
TOMM6	-0.3582688	6p21.1	9177
ANKRD54	-0.3577356	22q13.1	9176
PTOV1	-0.3571070	19q13.33	9175
COX17	-0.3568766	3q13.33	9174
DMAP1	-0.3565087	1p34.1	9173

Continued on next page

Table S3 – *Continued from previous page*

COX5A	-0.3563470	15q24.2	9172
FAHD2A	-0.3563443	2q11.1	9171
HDDC3	-0.3563010	15q26.1	9170
PPP1R7	-0.3561834	2q37.3	9169
LMAN2	-0.3559879	5q35.3	9168
DGCR6	-0.3556813	22q11.21	9167
LEKHJ1	-0.3555538	19p13.3	9166
H2AFJ	-0.3554735	12p12.3	9165
MRPL14	-0.3554184	6p21.1	9164
DHRS4	-0.3547074	14q11.2	9163
GNPTG	-0.3537155	16p13.3	9162
GPS2	-0.3537148	17p13.1	9161
RPL14	-0.3532889	3p22.1	9160
NDUFV2	-0.3532783	18p11.22	9159
JMJD8	-0.3532333	16p13.3	9158
RNF126	-0.3531662	19p13.3	9157
ERI3	-0.3530842	1p34.1	9156
MRPL11	-0.3523429	11q13.2	9155
RPL24	-0.3521698	3q12.3	9154
EEF1D	-0.3521018	8q24.3	9153
EXOSC5	-0.3514759	19q13.2	9152
SNRPC	-0.3513565	6p21.31	9151
MIIP	-0.3512770	1p36.22	9150
TP53I13	-0.3511437	17q11.2	9149
COX16	-0.3511005	14q24.2	9148
GABARAP	-0.3509948	17p13.1	9147
RPUSD3	-0.3509695	3p25.3	9146
RPL37A	-0.3509549	2q35	9145
PSMD13	-0.3506560	11p15.5	9144
CST3	-0.3506096	20p11.21	9143
EIF6	-0.3504160	20q11.22	9142
PRMT7	-0.3502140	16q22.1	9141
C21orf33	-0.3500666	21q22.3	9140
SSU72	-0.3500556	1p36.33	9139
TOMM7	-0.3499229	7p15.3	9138
CSNK2B	-0.3498084	6p21.33	9137

Continued on next page

Table S3 – *Continued from previous page*

GPS1	-0.3495443	17q25.3	9136
NHP2L1	-0.3495051	22q13.2	9135
RFNG	-0.3494963	17q25.3	9134
BOLA1	-0.3492415	1q21.2	9133
MRPL4	-0.3490396	19p13.2	9132
ARL2	-0.3485354	11q13.1	9131
MOB2	-0.3483037	11p15.5	9130
MRPL22	-0.3479208	5q33.2	9129
UBL7	-0.3478385	15q24.1	9128
CLN6	-0.3478160	15q23	9127
RPS29	-0.3476047	14q21.3	9126
NUDT16L1	-0.3474625	16p13.3	9125
TXNDC17	-0.3472957	17p13.1	9124
SCNM1	-0.3472117	1q21.3	9123
PSMC3	-0.3469897	11p11.2	9122
TMEM205	-0.3468145	19p13.2	9121
BSG	-0.3460898	19p13.3	9120
ADCK5	-0.3460777	8q24.3	9119
SEPW1	-0.3458001	19q13.33	9118
ZBTB48	-0.3454368	1p36.31	9117
TRADD	-0.3454008	16q22.1	9116
TECR	-0.3452565	19p13.12	9115
CYHR1	-0.3451989	8q24.3	9114
HEXDC	-0.3449760	17q25.3	9113
SRM	-0.3449445	1p36.22	9112
ALKBH4	-0.3444990	7q22.1	9111
PFN1	-0.3444819	17p13.2	9110
COX7C	-0.3442678	5q14.3	9109
TUFM	-0.3440914	16p11.2	9108
BOLA3	-0.3437944	2p13.1	9107
PSMB7	-0.3434103	9q33.3	9106
TBL3	-0.3433726	16p13.3	9105
PDZD11	-0.3430249	Xq13.1	9104
RPS10	-0.3427688	6p21.31	9103
ATP5O	-0.3427385	21q22.11	9102

Continued on next page

Table S3 – *Continued from previous page*

Table S4: REACTOME pathways coexpressed with *TP53BP2* at the pan-cancer level

REACTOME Pathway	FDR	NES	Group
REACTOME SYNTHESIS OF PIPS AT THE PLASMA MEMBRANE	0.0039551	2.181167	PI
REACTOME PI METABOLISM	0.0039551	2.161078	PI
REACTOME SIGNALING BY FGFR1 IN DISEASE	0.0039551	2.115694	RTK
REACTOME SUMOYLATION	0.0039551	2.108851	SUMO
REACTOME HDMS DEMETHYLATE HISTONES	0.0039551	2.086502	CHROMATIN
REACTOME FGFR1 MUTANT RECEPTOR ACTIVATION	0.0039551	2.055558	RTK
REACTOME SUMOYLATION OF DNA DAMAGE RESPONSE AND REPAIR PROTEINS	0.0039551	2.008626	MITOSIS
REACTOME INSULIN RECEPTOR SIGNALLING CASCADE	0.0039551	1.967273	RTK
REACTOME IRS MEDIATED SIGNALLING	0.0039551	1.965679	RTK
REACTOME DOWNSTREAM SIGNAL TRANSDUCTION	0.0039551	1.940625	RTK
REACTOME SIGNALING BY TYPE 1 INSULIN LIKE GROWTH FACTOR 1 RECEPTOR IGF1R	0.0040765	1.896016	RTK
REACTOME TRANSCRIPTIONAL REGULATION OF WHITE ADIPOCYTE DIFFERENTIATION	0.0039551	1.884853	TRANSCRIPTION
REACTOME RUNX1 INTERACTS WITH CO FACTORS WHOSE PRECISE EFFECT ON RUNX1 TARGETS IS NOT KNOWN	0.0039551	1.874975	TRANSCRIPTION

Table S4 – continued from previous page

REACTOME Pathway	FDR	NES	Group
REACTOME REGULATION OF LIPID METABOLISM BY PEROXISOME PROLIFERATOR ACTIVATED RECEPTOR ALPHA PPARALPHA	0.0039551	1.868084	LIPID
REACTOME CIRCADIAN CLOCK	0.0039551	1.858978	CIRCADIAN
REACTOME CHROMATIN ORGANIZATION	0.0039551	1.848799	CHROMATIN
REACTOME RESOLUTION OF SISTER CHROMATID COHESION	0.0039551	1.843366	MITOSIS
REACTOME PKMTS METHYLATE HISTONE LYSINES	0.0053071	1.838985	CHROMATIN
REACTOME SIGNALING BY HIPPO	0.0096679	1.832423	HIPPO
REACTOME BMAL1:CLOCK NPAS2 ACTIVATES CIRCADIAN GENE EXPRESSION	0.0081274	1.815190	TRANSCRIPTION
REACTOME REGULATION OF CHOLESTEROL BIOSYNTHESIS BY SREBP SREBF	0.0069570	1.803377	LIPID
REACTOME SIGNALING BY VEGF	0.0039551	1.801179	RTK
REACTOME NUCLEAR PORE COMPLEX NPC DISASSEMBLY	0.0058936	1.800474	NUCLEAR IMPORT
REACTOME NUCLEAR IMPORT OF REV PROTEIN	0.0059127	1.796270	NUCLEAR IMPORT
REACTOME SUMOYLATION OF CHROMATIN ORGANIZATION PROTEINS	0.0069570	1.790163	CHROMATIN
REACTOME TRANSCRIPTIONAL ACTIVATION OF MITOCHONDRIAL BIOGENESIS	0.0058168	1.788484	TRANSCRIPTION
REACTOME LAMININ INTERACTIONS	0.0094843	1.777972	EXTRACELLULAR

Table S4 – continued from previous page

REACTOME Pathway	FDR	NES	Group
REACTOME EXPORT OF VIRAL RIBONUCLEOPROTEINS FROM NUCLEUS	0.0084157	1.777226	NUCLEAR IMPORT
REACTOME HOST INTERACTIONS WITH INFLUENZA FACTORS	0.0083984	1.771280	NUCLEAR IMPORT
REACTOME SUMOYLATION OF TRANSCRIPTION COFACTORS	0.0092446	1.765833	TRANSCRIPTION
REACTOME SIGNALING BY SCF KIT	0.0096679	1.754468	RTK
REACTOME FACTORS INVOLVED IN MEGAKARYOCYTE DEVELOPMENT AND PLATELET PRODUCTION	0.0039551	1.751310	PLATELET
REACTOME SIGNALING BY RHO GTPASES	0.0039551	1.750643	RHO
REACTOME FCGAMMA RECEPTOR FCGR DEPENDENT PHAGOCYTOSIS	0.0050303	1.744419	RTK
REACTOME SUMOYLATION OF RNA BINDING PROTEINS	0.0096013	1.742245	NUCLEAR IMPORT
REACTOME SIGNALING BY PDGF	0.0098947	1.718372	RTK
REACTOME SIGNALING BY MET	0.0087050	1.673680	RTK
REACTOME MITOTIC PROMETAPHASE	0.0039551	1.667712	MITOSIS
REACTOME RHO GTPASE CYCLE	0.0048263	1.633584	RHO
REACTOME MITOTIC SPINDLE CHECKPOINT	0.0085068	1.624928	MITOSIS
REACTOME ORGANELLE BIOGENESIS AND MAINTENANCE	0.0039551	1.615241	ORGANELLE
REACTOME RHO GTPASE EFFECTORS	0.0039551	1.613101	RHO

Table S4 – continued from previous page

REACTOME Pathway	FDR	NES	Group
REACTOME INTRA GOLGI AND RETROGRADE GOLGI TO ER TRAF-FIC	0.0085068	1.520883	GOLGI
REACTOME FATTY ACID METABOLISM	0.0069570	-1.494461	LIPID
REACTOME HIV INFECTION	0.0053241	-1.507445	RNA, PRO
REACTOME HOST INTERACTIONS OF HIV FACTORS	0.0041992	-1.569975	RNA, PRO
REACTOME TRANSCRIPTIONAL REGULATION BY RUNX2	0.0040765	-1.612389	RNA, PRO
REACTOME SIGNALING BY THE B CELL RECEPTOR BCR	0.0040765	-1.626275	RNA, PRO
REACTOME HEDGEHOG OFF STATE	0.0040123	-1.646936	RNA, PRO
REACTOME INTERLEUKIN 1 SIGNALING	0.0060707	-1.653314	RNA, PRO
REACTOME CLEC7A DECTIN 1 SIGNALING	0.0040123	-1.679690	RNA, PRO
REACTOME TRANSCRIPTION COUPLED NUCLEOTIDE EXCISION REPAIR TC NER	0.0074800	-1.702776	RNA, PRO
REACTOME METABOLISM OF NUCLEOTIDES	0.0092446	-1.730195	NUCLEOTIDES
REACTOME DOWNSTREAM TCR SIGNALING	0.0039551	-1.776261	RNA, PRO
REACTOME RRNA MODIFICATION IN THE NUCLEUS AND CYTOSOL	0.0051472	-1.779175	RNA, PRO
REACTOME MAPK6 MAPK4 SIGNALING	0.0039551	-1.802064	RNA, PRO
REACTOME HEDGEHOG ON STATE	0.0039551	-1.832738	RNA, PRO
REACTOME SIGNALING BY NOTCH4	0.0039551	-1.847067	RNA, PRO

Table S4 – continued from previous page

REACTOME Pathway	FDR	NES	Group
REACTOME DNA REPLICATION	0.0040797	-1.847446	RNA, PRO
REACTOME FCERI MEDIATED NF KB ACTIVATION	0.0039551	-1.849522	RNA, PRO
REACTOME TP53 REGULATES METABOLIC GENES	0.0039551	-1.864623	ETC
REACTOME BIOSYNTHESIS OF THE N GLYCAN PRECURSOR DOLICHOL LIPID LINKED OLIGOSACCHARIDE LLO AND TRANSFER TO A NASCENT PROTEIN	0.0052914	-1.875283	RNA, PRO
REACTOME DEGRADATION OF BETA CATENIN BY THE DESTRUCTION COMPLEX	0.0039551	-1.936190	RNA, PRO
REACTOME DOWNSTREAM SIGNALING EVENTS OF B CELL RECEPTOR BCR	0.0039551	-1.950186	RNA, PRO
REACTOME CYCLIN A:CDK2 ASSOCIATED EVENTS AT S PHASE ENTRY	0.0039551	-1.951863	RNA, PRO
REACTOME GLUTATHIONE CONJUGATION	0.0074010	-1.970012	GLUTATHIONE
REACTOME PCP CE PATHWAY	0.0039551	-1.985695	RNA, PRO
REACTOME CELLULAR RESPONSE TO HYPOXIA	0.0039551	-1.993999	RNA, PRO
REACTOME UCH PROTEINASES	0.0039551	-2.032226	RNA, PRO
REACTOME TRNA MODIFICATION IN THE NUCLEUS AND CYTOSOL	0.0085031	-2.037361	RNA, PRO
REACTOME REGULATION OF RAS BY GAPS	0.0039551	-2.060455	RNA, PRO

Table S4 – continued from previous page

REACTOME Pathway	FDR	NES	Group
REACTOME INSERTION OF TAIL ANCHORED PROTEINS INTO THE ENDOPLASMIC RETICULUM MEMBRANE	0.0084157	-2.067241	ER
REACTOME SWITCHING OF ORIGINS TO A POST REPLICATIVE STATE	0.0039551	-2.069228	RNA, PRO
REACTOME ABC FAMILY PROTEINS MEDIATED TRANSPORT	0.0039551	-2.071899	RNA, PRO
REACTOME REGULATION OF MITOTIC CELL CYCLE	0.0039551	-2.090383	RNA, PRO
REACTOME MRNA SPLICING MINOR PATHWAY	0.0039551	-2.093824	RNA, PRO
REACTOME REGULATION OF MRNA STABILITY BY PROTEINS THAT BIND AU RICH ELEMENTS	0.0039971	-2.094841	RNA, PRO
REACTOME DETOXIFICATION OF REACTIVE OXYGEN SPECIES	0.0039551	-2.136171	GLUTATHIONE
REACTOME REGULATION OF PTEN STABILITY AND ACTIVITY	0.0039551	-2.137195	RNA, PRO
REACTOME ACTIVATION OF NF KAPPAB IN B CELLS	0.0039551	-2.146550	RNA, PRO
REACTOME RUNX1 REGULATES TRANSCRIPTION OF GENES INVOLVED IN DIFFERENTIATION OF HSCS	0.0039551	-2.150408	RNA, PRO
REACTOME PREFOLDIN MEDIATED TRANSFER OF SUBSTRATE TO CCT TRIC	0.0039551	-2.150997	PROTEIN FOLDING
REACTOME GLUCONEOGENESIS	0.0039551	-2.170585	GLUCONEOGENESIS
REACTOME DNA REPLICATION PRE INITIATION	0.0039551	-2.173181	RNA, PRO
REACTOME REGULATION OF RUNX2 EXPRESSION AND ACTIVITY	0.0039551	-2.207988	RNA, PRO

Table S4 – continued from previous page

REACTOME Pathway	FDR	NES	Group
REACTOME COOPERATION OF PREFOLDIN AND TRIC CCT IN ACTIN AND TUBULIN FOLDING	0.0039551	-2.215171	PROTEIN FOLDING
REACTOME ASYMMETRIC LOCALIZATION OF PCP PROTEINS	0.0039551	-2.238305	RNA, PRO
REACTOME G1 S DNA DAMAGE CHECKPOINTS	0.0039551	-2.264999	RNA, PRO
REACTOME DEGRADATION OF AXIN	0.0039551	-2.285513	RNA, PRO
REACTOME MITOCHONDRIAL FATTY ACID BETA OXIDATION	0.0039551	-2.294324	LIPID
REACTOME ORC1 REMOVAL FROM CHROMATIN	0.0039551	-2.295348	RNA, PRO
REACTOME CDK MEDIATED PHOSPHORYLATION AND REMOVAL OF CDC6	0.0039551	-2.296186	RNA, PRO
REACTOME APC C:CDH1 MEDIATED DEGRADATION OF CDC20 AND OTHER APC C:CDH1 TARGETED PROTEINS IN LATE MITOSIS EARLY G1	0.0039551	-2.320455	RNA, PRO
REACTOME ABC TRANSPORTER DISORDERS	0.0039551	-2.329077	RNA, PRO
REACTOME TNFR2 NON CANONICAL NF KB PATHWAY	0.0039551	-2.329230	RNA, PRO
REACTOME ASSEMBLY OF THE PRE REPLICATIVE COMPLEX	0.0039551	-2.341547	RNA, PRO
REACTOME ACTIVATION OF APC C AND APC C:CDC20 MEDIATED DEGRADATION OF MITOTIC PROTEINS	0.0039551	-2.347144	RNA, PRO
REACTOME REGULATION OF RUNX3 EXPRESSION AND ACTIVITY	0.0039551	-2.355961	RNA, PRO
REACTOME HEDGEHOG LIGAND BIOGENESIS	0.0039551	-2.371423	RNA, PRO

Table S4 – continued from previous page

REACTOME Pathway	FDR	NES	Group
REACTOME DECTIN 1 MEDIATED NONCANONICAL NF KB SIGNALING	0.0039551	-2.378197	RNA, PRO
REACTOME ANTIGEN PROCESSING CROSS PRESENTATION	0.0039551	-2.378437	RNA, PRO
REACTOME DEGRADATION OF GLI1 BY THE PROTEASOME	0.0039551	-2.380802	RNA, PRO
REACTOME STABILIZATION OF P53	0.0039551	-2.408723	RNA, PRO
REACTOME INFECTIOUS DISEASE	0.0085068	-2.410352	RNA, PRO
REACTOME DEFECTIVE CFTR CAUSES CYSTIC FIBROSIS	0.0039551	-2.455517	RNA, PRO
REACTOME REGULATION OF APOPTOSIS	0.0039551	-2.467025	RNA, PRO
REACTOME SCF SKP2 MEDIATED DEGRADATION OF P27 P21	0.0039551	-2.473740	RNA, PRO
REACTOME DEGRADATION OF DVL	0.0039551	-2.495093	RNA, PRO
REACTOME METABOLISM OF POLYAMINES	0.0039551	-2.514469	RNA, PRO
REACTOME CROSS PRESENTATION OF SOLUBLE EXOGENOUS ANTIGENS ENDOSOMES	0.0039551	-2.551769	RNA, PRO
REACTOME THE ROLE OF GTSE1 IN G2 M PROGRESSION AFTER G2 CHECKPOINT	0.0039551	-2.609565	RNA, PRO
REACTOME PROTEIN LOCALIZATION	0.0043798	-2.613087	MITOCHONDRIAL PROTEIN IMPORT
REACTOME AUF1 HNRNP D0 BINDS AND DESTABILIZES MRNA	0.0039551	-2.617612	RNA, PRO
REACTOME VIF MEDIATED DEGRADATION OF APOBEC3G	0.0039551	-2.635161	RNA, PRO

Table S4 – continued from previous page

REACTOME Pathway	FDR	NES	Group
REACTOME FBXL7 DOWN REGULATES AURKA DURING MITOTIC ENTRY AND IN EARLY MITOSIS	0.0039551	-2.676244	RNA, PRO
REACTOME NEGATIVE REGULATION OF NOTCH4 SIGNALING	0.0039551	-2.695555	RNA, PRO
REACTOME ACTIVATION OF THE MRNA UPON BINDING OF THE CAP BINDING COMPLEX AND EIFS AND SUBSEQUENT BINDING TO 43S	0.0039551	-2.822320	RNA, PRO
REACTOME SIGNALING BY ROBO RECEPTORS	0.0053143	-2.894811	RNA, PRO
REACTOME INFLUENZA INFECTION	0.0047134	-2.906947	RNA, PRO
REACTOME MITOCHONDRIAL PROTEIN IMPORT	0.0039551	-3.020590	MITOCHONDRIAL PROTEIN IMPORT
REACTOME METABOLISM OF AMINO ACIDS AND DERIVATIVES	0.0068715	-3.063525	RNA, PRO
REACTOME RRNA PROCESSING IN THE NUCLEUS AND CYTOSOL	0.0051472	-3.157595	RNA, PRO
REACTOME RRNA PROCESSING	0.0051472	-3.169546	RNA, PRO
REACTOME NONSENSE MEDIATED DECAY NMD	0.0040765	-3.276510	RNA, PRO
REACTOME COMPLEX I BIOGENESIS	0.0039551	-3.296420	ETC
REACTOME THE CITRIC ACID TCA CYCLE AND RESPIRATORY ELECTRON TRANSPORT	0.0043316	-3.315416	ETC
REACTOME REGULATION OF EXPRESSION OF SLITS AND ROBOS	0.0048263	-3.433504	RNA, PRO
REACTOME SELENOAMINO ACID METABOLISM	0.0040765	-3.471286	RNA, PRO

Table S4 – continued from previous page

REACTOME Pathway	FDR	NES	Group
REACTOME MITOCHONDRIAL TRANSLATION	0.0040123	-3.523240	MITOCHONDRIAL TRANSLATION
REACTOME EUKARYOTIC TRANSLATION INITIATION	0.0040945	-3.564249	RNA, PRO
REACTOME TRANSLATION	0.0073182	-3.608736	RNA, PRO
REACTOME NONSENSE MEDIATED DECAY NMD INDEPENDENT OF THE EXON JUNCTION COMPLEX EJC	0.0040123	-3.631644	RNA, PRO
REACTOME RESPIRATORY ELECTRON TRANSPORT ATP SYNTHESIS BY CHEMIOSMOTIC COUPLING AND HEAT PRODUCTION BY UNCOUPLING PROTEINS	0.0039551	-3.637282	ETC
REACTOME RESPIRATORY ELECTRON TRANSPORT	0.0039551	-3.646728	ETC
REACTOME SRP DEPENDENT COTRANSLATIONAL PROTEIN TARGETING TO MEMBRANE	0.0040765	-3.666987	RNA, PRO
Abbreviations: NES: Normalized enrichment score; FDR: False discovery rate; PI: Phosphoinositol metabolism; RTK: Receptor-tyrosine-kinase/RAS/PI3K signaling; Lipid: Lipid metabolism; Circadian: Circadian signaling; Hippo: Hippo signaling; Extracellular: Extracellular matrix organization; Rho: Rho signaling; Golgi: Golgi transport; RNA,PRO: RNA-polymerase-and-proteasome-related; ETC: Electron transport chain			

Table S5: Differentially expressed genes up-regulated in ASPP2-KO U2OS cells compared to WT U2OS cells

Gene	LogFC	AdjP
FKBP9P1	7.519491	0.0012536
NPIPA8	7.220239	0.0412669
PGAM1	6.422920	0.0163308
SLITRK6	5.580896	0.0000576
BEND4	5.346424	0.0004178
SLC6A12	5.268754	0.0007457
C12orf50	4.833205	0.0000000
PTPN7	4.750665	0.0000690
KRT17	4.682808	0.0022325
ZP1	4.312774	0.0076378
ESPN	4.297829	0.0216513
COLEC11	4.166002	0.0210252
TBC1D10C	4.105243	0.0000018
MYL10	3.924842	0.0000032
TMEM271	3.887678	0.0002288
ABCC9	3.760231	0.0015134
C1QTNF1	3.748670	0.0000109
SLC49A3	3.626664	0.0013420
MMP10	3.616010	0.0413299
IGFBP4	3.606143	0.0000000
ZNF132	3.551635	0.0066590
GPAT2	3.534981	0.0000000
LOC100288846	3.364237	0.0190758
MIR663AHG	3.284122	0.0000803
IFT172	3.249232	0.0000210
LINC01234	3.222541	0.0000000
TUBA3E	3.210387	0.0000000
GAGE1	3.131857	0.0003590
LIPH	3.067080	0.0022912
TUBA3D	3.015348	0.0000000
COL20A1	2.993807	0.0177147
OSCAR	2.977117	0.0000014
LAPTM5	2.974930	0.0021414

Continued on next page

Table S5 – *Continued from previous page*

SLCO2A1	2.974706	0.0000899
AURKC	2.972428	0.0000000
CRHR1	2.915662	0.0000174
KLK6	2.908467	0.0000000
KCNK10	2.886195	0.0000000
TGM5	2.834804	0.0313422
LRRC32	2.832680	0.0017432
TRPV3	2.813229	0.0000392
ATP1A2	2.811205	0.0009442
CABYR	2.808783	0.0000040
ESM1	2.788529	0.0000065
CDKL3	2.771580	0.0066919
GRM4	2.768575	0.0056057
PRSS33	2.735361	0.0089453
AOAH	2.706209	0.0139860
LPAR4	2.667026	0.0038065
CFAP53	2.651302	0.0259988
LINC01835	2.641889	0.0049789
NKAIN4	2.636526	0.0430259
HOXB2	2.620633	0.0000106
KNDC1	2.598159	0.0014659
SPATA31C2	2.591806	0.0000000
KLK10	2.584463	0.0021501
TG	2.563195	0.0258857
SERPINF1	2.541905	0.0000000
MYH15	2.532931	0.0007696
AMN1	2.523100	0.0040103
HES6	2.520025	0.0000000
GPSM3	2.483751	0.0000020
CD22	2.466659	0.0017351
FBXL14	2.459771	0.0286629
SSC5D	2.457033	0.0296048
DBH-AS1	2.453502	0.0000839
ENOSF1	2.444839	0.0000000
ITGB2-AS1	2.407844	0.0002717
ALDH8A1	2.406283	0.0104351

Continued on next page

Table S5 – *Continued from previous page*

CSMD2	2.402159	0.0069769
CTXND1	2.394222	0.0094290
KSR1	2.391775	0.0061047
MGST1	2.375441	0.0000000
H2AFJ	2.360137	0.0000000
ASTN1	2.354250	0.0000644
TMEM132E	2.351673	0.0000224
DIP2A-IT1	2.340972	0.0097321
FZD10-AS1	2.337615	0.0481568
GCC2-AS1	2.317513	0.0005675
TSPAN11	2.295661	0.0000129
LINC02657	2.291956	0.0009967
ATP8B1	2.290473	0.0037056
LINC02547	2.274898	0.0007419
ARL14	2.269412	0.0000037
TTYH1	2.268504	0.0000000
STK32A	2.262764	0.0000039
SULT1A1	2.259965	0.0000000
GRIP2	2.241869	0.0027809
LGALS9	2.236324	0.0033394
IFITM10	2.227867	0.0002945
INPP5D	2.226315	0.0000000
CACNA2D4	2.225967	0.0000001
GCGR	2.223863	0.0008694
HSD17B3	2.221202	0.0000947
GALNT9	2.202382	0.0137586
PSCA	2.173543	0.0283032
LINC00707	2.169222	0.0000024
ARHGEF33	2.154588	0.0000025
LSP1	2.148798	0.0005300
LGR6	2.145552	0.0412057
LOC101928982	2.143942	0.0410228
DLX4	2.136579	0.0000034
ITGAX	2.120129	0.0034971
MUC20	2.119528	0.0000239
ND6	2.113880	0.0000000

Continued on next page

Table S5 – *Continued from previous page*

SBSN	2.112505	0.0035363
LINC00520	2.099979	0.0074811
LINC01833	2.088635	0.0000003
LGI3	2.087926	0.0011459
ACRBP	2.086491	0.0035045
CHRNA9	2.055276	0.0090357
TUBG2	2.054654	0.0000010
LRRC29	2.052742	0.0038436
DRAXIN	2.050454	0.0000000
NGEF	2.040777	0.0000080
MARCO	2.032642	0.0000032
LMO3	2.021278	0.0094430
LOC100129534	2.020490	0.0117548
RNASE1	2.012811	0.0000001
MUC15	2.010606	0.0280143
MGP	2.008087	0.0273298
Abbreviations: LogFC - log2 fold change (ASPP2-KO - WT); Adj p - adjusted p-value.		

Table S6: Differentially expressed genes down-regulated in ASPP2-KO U2OS cells compared to WT U2OS cells

Gene	LogFC	Adj p
MAGEA10	-10.731242	0.0000000
GTSF1	-10.672512	0.0000000
PHYHD1	-10.094430	0.0000000
CDKN2B	-9.480942	0.0000000
FAM155B	-9.193216	0.0000000
NOS3	-8.589157	0.0000000
TP73-AS1	-8.119187	0.0000002
CHST15	-8.068568	0.0000000
NLRP2	-7.892976	0.0000000
ZNF85	-7.764811	0.0000047
GLUL	-7.702542	0.0000000
CHDH	-7.079899	0.0000007

Continued on next page

Table S6 – *Continued from previous page*

SLC25A20	-7.061763	0.0000012
HLA-DPB1	-7.014643	0.0000000
FBLIM1	-6.807850	0.0000000
ALX4	-6.791383	0.0000240
NACAD	-6.764719	0.0000007
BOLA2B	-6.709422	0.0001095
NPW	-6.594288	0.0001713
CAVIN3	-6.473846	0.0000012
NMB	-6.271382	0.0000000
TAFA4	-6.235059	0.0000371
MPPE1	-5.989277	0.0008382
C8orf31	-5.969357	0.0000000
AGAP2-AS1	-5.902178	0.0006660
PPP1R1B	-5.873689	0.0000001
GIPC3	-5.768020	0.0000759
A4GALT	-5.756563	0.0007758
SHISAL1	-5.712728	0.0000000
LHX2	-5.703542	0.0000000
TSKS	-5.606302	0.0000147
FAM174B	-5.471442	0.0000000
ADD2	-5.403707	0.0000000
HSPB2	-5.206245	0.0000000
ZNF93	-5.013733	0.0051342
SELENBP1	-4.970788	0.0000000
FAM72A	-4.968691	0.0005973
DUSP27	-4.919314	0.0007492
ATOH8	-4.886121	0.0000000
ZNF675	-4.838840	0.0044240
CCDC36	-4.792071	0.0000183
KISS1R	-4.732981	0.0013963
MLXIPL	-4.682046	0.0000009
KCTD4	-4.521148	0.0000028
NRCAM	-4.457228	0.0000449
LOC105370854	-4.455166	0.0000001
RARRES2	-4.441710	0.0000001
HSPB7	-4.292291	0.0010675

Continued on next page

Table S6 – *Continued from previous page*

DNAH2	-4.260729	0.0317250
CRYBA2	-4.099409	0.0005811
LYPD5	-4.070454	0.0007744
RPP25	-3.962442	0.0000193
COL3A1	-3.944377	0.0000000
EPYC	-3.874584	0.0002019
ACSS1	-3.873176	0.0007389
MT1G	-3.812435	0.0000008
CHI3L1	-3.789338	0.0311723
ADRB3	-3.779239	0.0427506
MGC16275	-3.743751	0.0001086
APOL4	-3.726791	0.0000713
ARSD	-3.718774	0.0003755
FENDRR	-3.695768	0.0001344
ATG9B	-3.682937	0.0000371
KBTBD8	-3.618318	0.0126132
ADAMTS5	-3.586585	0.0003603
TLR8-AS1	-3.581769	0.0469957
ASCC2	-3.505208	0.0000000
MAPK8IP2	-3.499462	0.0003153
MPZL2	-3.490161	0.0003348
WFIKKN1	-3.420038	0.0112023
SLC38A3	-3.392183	0.0019864
OR51B5	-3.339498	0.0000033
KLC3	-3.331285	0.0498540
HHIPL1	-3.306720	0.0025272
GPRIN2	-3.281397	0.0027176
SAA1	-3.234129	0.0000002
IL32	-3.215962	0.0000000
GJA4	-3.205653	0.0039113
DPP6	-3.184966	0.0054557
ZNF503-AS1	-3.176102	0.0006503
GAS2	-3.109473	0.0159299
LTB	-3.088727	0.0000000
MYOD1	-3.049273	0.0065111
OR51B4	-3.021450	0.0010769

Continued on next page

Table S6 – *Continued from previous page*

WDR72	-2.991600	0.0007460
LINC01639	-2.920809	0.0004376
SNORA80B	-2.900680	0.0466485
TMEM163	-2.888743	0.0032865
GPR12	-2.888170	0.0008365
MS4A15	-2.888091	0.0351298
TRPM6	-2.872127	0.0000005
COL8A1	-2.863605	0.0000000
GPC5	-2.829424	0.0067501
PTGER2	-2.795702	0.0000563
SORD	-2.753407	0.0065111
AATK	-2.738942	0.0255541
SLC22A14	-2.682623	0.0291836
PKIA-AS1	-2.645748	0.0027176
MEGF6	-2.642905	0.0061524
DPY19L2P4	-2.642256	0.0032018
PDE11A	-2.637257	0.0064551
TNFRSF1B	-2.626938	0.0229477
LUM	-2.598141	0.0000000
NMNAT3	-2.566435	0.0024358
TMEM220	-2.563727	0.0000012
PLPP7	-2.538648	0.0000810
LAMP3	-2.538168	0.0007207
BASP1	-2.528960	0.0007327
DEF6	-2.515684	0.0000344
SLC6A10P	-2.514669	0.0002903
LLGL2	-2.506279	0.0171469
RGS20	-2.489143	0.0083206
DOK6	-2.485871	0.0000012
RIMKLA	-2.473740	0.0000000
BDKRB1	-2.472214	0.0008158
RNY4	-2.447304	0.0162880
ABCC10	-2.446322	0.0007457
SNORA79	-2.437763	0.0002278
RMRP	-2.418719	0.0010072
DLX3	-2.416742	0.0005398

Continued on next page

Table S6 – *Continued from previous page*

MMP13	-2.413315	0.0000006
DAPK1	-2.411338	0.0006422
GLB1L3	-2.407978	0.0007456
TRAF1	-2.398753	0.0095479
OR7E14P	-2.363781	0.0050627
RNY3	-2.353454	0.0306302
SHISA9	-2.311688	0.0007958
TJP3	-2.302659	0.0004454
MPZL3	-2.293780	0.0002896
SNORA80A	-2.275738	0.0467078
COL16A1	-2.271526	0.0000000
PPP1R3F	-2.259725	0.0144799
SCARNA13	-2.250137	0.0003846
STEAP1	-2.245264	0.0000002
ADGRE1	-2.210602	0.0000015
ARMC4	-2.204286	0.0000177
DNM3	-2.196870	0.0000029
RBPMS	-2.186138	0.0000003
LOC105375614	-2.174990	0.0416033
FAT3	-2.173649	0.0001809
BIRC3	-2.151642	0.0000002
HOXD10	-2.150292	0.0000000
UNC13D	-2.147255	0.0000000
POGLUT2	-2.146710	0.0025742
CHRNA1	-2.145566	0.0000000
KISS1	-2.145253	0.0000000
MAP3K20-AS1	-2.118359	0.0069263
DYNC1I1	-2.114432	0.0000022
DNAJC12	-2.112913	0.0003190
CNN1	-2.108983	0.0000000
GPR20	-2.098097	0.0000009
HOXD11	-2.097129	0.0000000
LINC00665	-2.074523	0.0000000
UPK1A-AS1	-2.039620	0.0020473
SIX6	-2.038732	0.0067246
ROS1	-2.029412	0.0370479

Continued on next page

Table S6 – Continued from previous page

ZNF100	-2.011345	0.0082071
PDZRN3	-2.002719	0.0046201
Abbreviations: LogFC - log2 fold change (ASPP2-KO - WT); Adj p - adjusted p-value.		

Table S7: List of ASPP2-downregulated translation-related genes

Gene	LogFC	Adj p	Gene	LogFC	Adj p
RPS15A	0.7937559	3.612346e-08	RPS3A	0.3305693	5.139912e-04
RPL9	0.6973881	3.907030e-06	RPL23	0.3267340	8.788416e-04
EIF2S3	0.6699825	4.175311e-05	EIF3F	0.3251224	8.218586e-04
RPL23A	0.6259558	4.716908e-04	RPL6	0.2864418	1.339321e-02
RPL14	0.6143115	8.384768e-05	RPLP1	0.2859211	3.159677e-03
RPL24	0.5957201	3.036949e-02	RPS2	0.2732092	2.996426e-02
RPL37	0.5916702	1.814929e-05	RPL3	0.2714651	2.226579e-03
RPL11	0.5627319	5.304879e-07	RPL31	0.2384208	2.965702e-02
RPS9	0.5606139	6.495483e-03	RPS13	0.2379008	1.621578e-03
SRP14	0.5481878	1.064517e-04	RPS16	0.2306782	4.677466e-03
EIF3B	0.5264379	1.513251e-03	RPL19	0.2292612	3.760216e-03
RPL28	0.4925589	1.973683e-03	RPL27	0.2073685	1.090333e-02
EIF4EBP1	0.4583743	9.770672e-03	RPL35A	0.2058964	5.261190e-03
EIF2B4	0.4535660	4.194579e-03	RPS20	0.2035288	7.328543e-03
RPS5	0.4178597	1.580941e-03	RPL15	0.2017433	2.580815e-02
RPS18	0.3824310	3.776053e-02	SRP72	0.2014206	3.313679e-02
RPS12	0.3695356	6.208362e-06	RPL32	0.1830494	4.535565e-02
EIF4B	0.3659506	4.786315e-02	RPLP0	0.1618654	2.965702e-02
ETF1	0.3641843	6.992873e-04			

Abbreviations: LogFC - log2 fold change (ASPP2-KO - WT); Adj p - adjusted p-value.

Table S8: Enriched KEGG pathways for candidate ASPP2 interactors identified by IP-MS

Description	FDR	Leading edges
Hippo signaling pathway	0.00e+00	FRMD6, MPP5, PARD3, PARD6B, PARD6G, PPP1CA, PPP1CB, PPP1CC, PPP2R2A, PRKCI, WWTR1, YAP1
Tight junction	9.30e-06	MPP5, PARD3, PARD6B, PARD6G, PPP2R2A, PRKCI
Endocytosis	5.02e-05	CAV1, IQSEC2, PARD3, PARD6B, PARD6G, PRKCI
mRNA surveillance pathway	1.70e-04	PPP1CA, PPP1CB, PPP1CC, PPP2R2A
Hippo signaling pathway - multiple species	1.70e-04	FRMD6, WWTR1, YAP1
Human papillomavirus infection	1.70e-04	MPP5, PARD3, PARD6B, PARD6G, PPP2R2A, PRKCI
Platelet activation	3.70e-04	PPP1CA, PPP1CB, PPP1CC, PRKCI
Dopaminergic synapse	3.80e-04	PPP1CA, PPP1CB, PPP1CC, PPP2R2A
Insulin signaling pathway	4.00e-04	PPP1CA, PPP1CB, PPP1CC, PRKCI
Adrenergic signaling in cardiomyocytes	4.10e-04	PPP1CA, PPP1CB, PPP1CC, PPP2R2A
Long-term potentiation	7.50e-04	PPP1CA, PPP1CB, PPP1CC
Amphetamine addiction	7.50e-04	PPP1CA, PPP1CB, PPP1CC
Rap1 signaling pathway	1.10e-03	PARD3, PARD6B, PARD6G, PRKCI
Focal adhesion	1.10e-03	CAV1, PPP1CA, PPP1CB, PPP1CC

Table S8 – continued from previous page

Description	FDR	Leading edges
Proteoglycans in cancer	1.10e-03	CAV1, PPP1CA, PPP1CB, PPP1CC
Inflammatory mediator regulation of TRP channels	1.40e-03	PPP1CA, PPP1CB, PPP1CC
Insulin resistance	2.10e-03	PPP1CA, PPP1CB, PPP1CC
Oocyte meiosis	2.50e-03	PPP1CA, PPP1CB, PPP1CC
Vascular smooth muscle contraction	2.50e-03	PPP1CA, PPP1CB, PPP1CC
Alcoholism	3.90e-03	PPP1CA, PPP1CB, PPP1CC
Oxytocin signaling pathway	4.30e-03	PPP1CA, PPP1CB, PPP1CC
Cellular senescence	4.60e-03	PPP1CA, PPP1CB, PPP1CC
cGMP-PKG signaling pathway	4.70e-03	PPP1CA, PPP1CB, PPP1CC
Axon guidance	5.60e-03	PARD3, PARD6B, PARD6G
Herpes simplex infection	6.10e-03	PPP1CA, PPP1CB, PPP1CC
cAMP signaling pathway	7.20e-03	PPP1CA, PPP1CB, PPP1CC
Regulation of actin cytoskeleton	8.00e-03	PPP1CA, PPP1CB, PPP1CC

Abbreviations: FDR - False discovery rate.

Table S9: List of candidate ASPP2-regulated phosphosites identified by phosphoproteomic analysis

Gene	Phosphosite	Gene	Phosphosite
ABCF1	S24	MSN	S429
ACBD3	S344	MTA2	S52
ACTR10	S201	MTHFD1	T314, S318
ADNP	S289	MYH9	S1111
AHNAK	S819, S5393	NAA15	T584
AHNAK2	S5714	NAP1L1	T39, S141
AKAP12	S386, T1760	NAP1L4	S5, T117
AKAP2	T698, S704	NASP	T390, S751, T756
AKAP2	T698	NCAPD2	S1333
AMFR	S522	NCAPH	T203, T222
AMPD2	T188	NES	S820, S945
ANXA2; ANXA2P2	S26	NFIC	S334
ARFGAP3	S454	NOP56	S465
BAIAP2	S428	NUCKS1	S19
C1QBP	S87, T165	OGFR	S487, S505, T529

Continued on next page

Table S9 – Continued from previous page

CALU	S277, S300	PAFAH1B2	S2
CALU	S277	PAK1	S220, T229, T230
CANX	S86	PALM2- AKAP2; PALM2	S376
CD44	S254	PCBP1	S111, S171
CDK11B; CDC2L1; CDK11A	S238		
CEP170	S1112	PDLIM5	S309, S313
CNOT10	S606	PGD	S465
COL4A3BP	S138	PLIN3	S291
CRKL	S107	PPP1R12A	S853, S856
CSRP1	S81	PRDX1	T183
CTCF	S224	PRDX6	S83
CTNNA1; CTNNA2	T658	PRKD2	T714
CTNNB1	S29, T556	PRKDC	T18, T2638
CTTN	T399	PSAT1	S20
CTTNBP2NL	S484	PSMC4	T25
DAP	S49, T56	PSMD2	T24, S363
DDX42	S751	PSMD4	S269

Continued on next page

Table S9 – Continued from previous page

DVL3	S48	PTPN12	S372, S670
DYNC1L2	S205	PYCR1	S270
EEF1A1; EEF1A1P5	T226, S416	RANBP2	S2741
EEF2	S38	RECQL	S58
EIF4A1	S394	RIF1	S1127, S1552, S1554
EIF4B	S523	ROCK2	S1132, S1137
EIF4EBP1	S86	RPL18	S163
EML4	S16	RPL31	T104
EPB41L2	S683	RPLP2	S16
EPB41L3	T468	RTN4	S11
EPRS	S1000	SAFB2; SAFB	S227, S234
ERCC6L	S1098	SAMD4A	T670
FAM21C; FAM21A	S877	SAMD4B	S274, S279, S281
FAM91A1	T353, S355	SEC16A	S2039
FAM91A1	S355	SEPTIN7	S333
FAS	T193	SERBP1	S237, S323
FASN	S265, S1261	SERPINH1	S139

Continued on next page

Table S9 – Continued from previous page

FASN	S1261	SET	S15
FIP1L1	S496	SF3B1	S400
FLNA	S1041, S2026	SHMT2	S29
FLNC	S1939, T2575	SKA3	S124
FNBP1L	T496	SMN2; SMN1	T25
FXR1	T398, T475	SNX8	T452
FXR1	T398	SPTAN1	S817, S1075
GAPDH	T182	SPTBN1	T779, S817, S1041
GORASP2	S381	SQSTM1	S328
GOT1	S116	STRAP	S335
GTF3C1	S568, S748	STRN	S190, S294
HDAC1	S348	SUB1	S104
HEATR6	S336, S337	SUMO1	T10
HLA-A	S311, S343, S359	SYNCRIP	T162
HLA-B	S352	TBC1D4	S726
HMGCS1	T476	TJP2	S1044, S1045

Continued on next page

Table S9 – Continued from previous page

HN1	S41, S42, S87, S88	TLK2	S182
HNRNPA1	S299, S300	TMPO	S66
HNRNPA2B1	S212	TNKS1BP1	S498, S899
HNRNPK	S214, S430	TOR1AIP1	S156
HNRNPL	T98	TP53BP1	S301, S310, S484
HNRNPM	S204, S365	TPD52	S32
HNRNPU	T267	TPR	S1838, S2037
HSP90AA1	S453, S468, S470, T645	TRIM28	S437, S453
HSP90AB1	S470	TUBA1B; TUBA1C; TUBA1A; TUBA3C; TUBA3E	T56
HSPA9	S148, S164	TUBB	T15, T272
HSPB1	S49, S176	TUBB4A; TUBB; TUBB4B; TUBB2A; TUBB2B	S60, S216, T256

Continued on next page

Table S9 – Continued from previous page

HSPD1	S70, T117, T163, S253, S256	TUBB4A; TUBB4B	S40
HUWE1	S2887, T3830	TWF1	S143
IL18	S72	UBA1	S824, S855
ILF3	S190, S476, S486	UBAP2L	S482
KHSRP	S184	UBR4	S2871
KIAA1217	S1030	UCHL3	S125
KIF2C	S632	USP11	S867
LDB1	T270, S341	USP14	S228
LIG1	S66	USP7	S49
LIMA1	S230	VCL	S596, T603
LMNA	S406, S546	VIM	S83, T336, S412
MAGED2	S265	WASF2	S292
MAP1B	S1156, S1219	XRCC1	S430
MAP1B	S1219	YAF2	S94, S99, S112
MAP2K2	T17	YAP1	S365
MAP4	T644	YBX3	S79

Continued on next page

Table S9 – *Continued from previous page*

MAP4K4	S666, S668	YTHDF3	S383
MATR3; DK- FZp686L22104	S722	ZC3H15	S356
MB21D2	S436		
MCL1	S162, S376, T378		
Abbreviations: S - Serine; T - Threonine.			

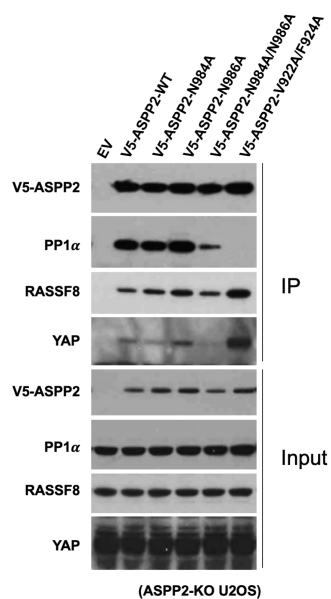


Figure S1: Validation of IP-MS hits by WB. Western blot analysis of the co-immunoprecipitates from ASPP2-KO U2OS cells transiently transfected with empty vector (EV, lane 1), V5-ASPP2-WT (lane 2), and four other V5-tagged ASPP2 mutants (lanes 3-6), using the indicated antibodies (performed and blotted by Dr. Thomas Leissing, University of Oxford). Lanes 3-6 were blotted for experimental purposes beyond the scope of this study.

Supporting information for chapter 6

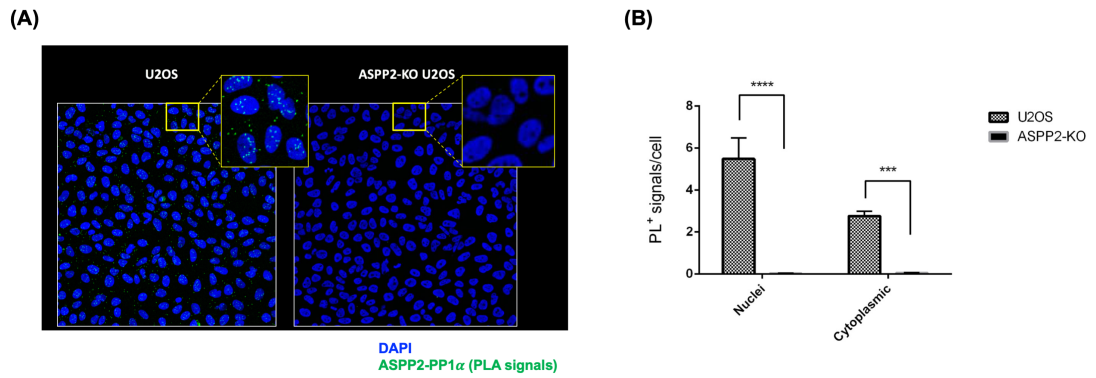


Figure S2: Detection of ASPP2-PP1 α interaction using PLA. (A) Representative immunofluorescence staining of PLA signals (green), with nuclei counterstained with DAPI (blue) in WT and ASPP2-KO U2OS cells. (B) Top: Bar plot showing the average nuclear and cytoplasmic PLA signals per cell. Each data point represents nuclear PLA signals in an individual cell from the analysis of three separate images taken from WT or ASPP2-KO U2OS in biological duplicates.

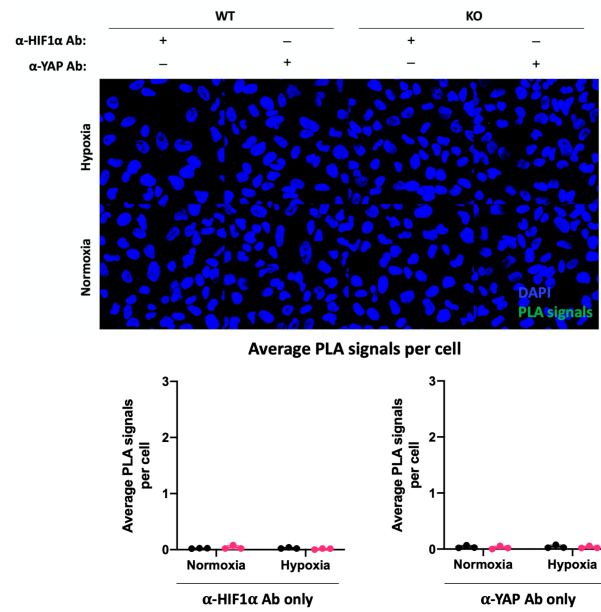


Figure S3: Controls for PLA for the detection of YAP-HIF1 α interaction. Top: Representative immunofluorescence staining of PLA signals (green) using anti-YAP antibody only or anti-HIF1 α antibody only, with nuclei counterstained with DAPI (blue) in WT and ASPP2-KO U2OS cells incubated for 6 hours under 21% O₂ (normoxia) or 0.5% O₂ (hypoxia). Bottom: Bar plots showing the average nuclear PLA signals per cell. Each data point represents nuclear PLA signals in an individual cell from the analysis of five separate images taken from biological duplicates of WT or ASPP2-KO U2OS cells incubated for 6 hours under 21% O₂ (normoxia) or 0.5% O₂ (hypoxia). Representative images are in the top panel.

RScript for gene coexpression analysis

```
#RScript for pancancer gene coexpression analysis#
```

```
#### loading libraries ####
```

```
library(data.table)
library(tibble)
library(dplyr)
library(stringr)
library(ggplot2)
library(ggrepel)
library(Hmisc)
library(mice)
```

```
#### data input ####
```

```
#### File: Pan-cancer batch-effect-normalized mRNA gene expression values with BOTH tumor and normal from Xena browser####
```

```
#### values are log2(x+1), in which x is batch effects normalized RSEM
```

```
#### detailed information is listed in:
```

```
#### https://docs.google.com/document/d/1q-7Tkzd7pci4Rz-\_lswASRMRzYrbgx1FTTfAWOyHbmk/edit#heading=h.b4dyuxs3v1x5
```

```
syn4976369 = fread("EB++AdjustPANCAN_IlluminaHiSeq_RNASeqV2.geneExp.tsv", sep =  
"t", header=T, stringsAsFactors = F)
```

```
syn4976369[1:10, 1:10]
```

```
syn4976369[, 1] = word(syn4976369[, 1], 1, sep = "\\") #remove "\" from the first column that  
contains gene identifiers
```

```
head(syn4976369[, 1])
```

```
#check the class of each column to see if the input contains features of the correct class (i.e. gene  
expression values as numeric, geneid as characters)
```

```
table(sapply(syn4976369, class))
```

```
which(sapply(syn4976369, class)=="factor") #applies only if there is factor class identified in the  
above step
```

```
n_occur <- data.frame(table(syn4976369$sample)) #check if there are duplicated gene names  
after the name correction above
```

```
n_occur <- data.frame(table(syn4976369$gene_id)) #check if there are duplicated gene names  
after the name correction above
```

```
n_occur[n_occur$Freq > 1,]
```

```
#### look at which rows contain the duplicated gene, then draw histogram to see gene expression  
distribution across samples to see which row to remove
```

```
syn4976369 = syn4976369[-16302,] #remove the row that contains duplicated gene
```

```
discard = which(syn4976369$sample %in% "?") #check which geneid is a question mark instead of  
proper geneid
```

```
# discard those genesnames with ?
```

```
if (length(discard) > 0) {
```

```
  syn4976369 = syn4976369[-discard,]
```

```
}
```

```
dat1 = syn4976369 #backup syn4976369 just in case
```

```
dat1 = as.data.frame(t(dat1[, -1])) #transpose df so that the genes now as columns, patient samples  
as rows
```

```
colnames(dat1) = syn4976369$sample #annotate column names with geneid
```

```
# colnames(dat1) = syn4976369$gene_id #annotate column names with geneid
```

```
dat1[1:5, 1:5] #check if colnames and rownames make sense
```

```
table(sapply(dat1, class)) #check again the feature class just in case some features incorrectly  
converted into factors
```

```
dim(dat1) #check the dimension of our dataset
```

```
#import the meta data for annotation of the samples
```

```

meta =
read.csv("Survival_SupplementalTable_S1_20171025_xena_sp",sep='\t',header=T,stringsAsFactor
s = F)
rownames(dat1) = gsub("\\.", "-", rownames(dat1)) #replace the "." in sample names with "-" to
match those in the meta file
dat1[1:5,1:5]
dat1$cancer = meta$cancer.type.abbreviation[match(rownames(dat1),meta$sample)] #create a
column to store cancer type info
head(dat1$cancer) #the cancer info should start with "ACC"
table(dat1$cancer) #look at the distribution of cancer types across samples
write.csv(table(dat1$cancer),"20190830_numbers_per_cancer_pre_filtering.csv")
# check if the dataset contains missing data (>1 methods included, but keep only the fastest
method, preferably graphical)
# method 1
library(funModeling)
dat_status = df_status(dat1[, -20503])
missing_dat = dat1[!complete.cases(dat1), , drop = FALSE]
# visualize missing data using Amelia missmap function
library(Amelia)
missmap(missing_dat)
# selecting genes with NAs
genes_withNA = subset(dat_status$variable, dat_status$p_na > 0)
# find columns containing genes with NA
ind = which(colnames(missing_dat) %in% genes_withNA)
missing_dat = missing_dat[,ind]
table(missing_dat$cancer)
table(dat_status$p_na)
# select the rows with NA cancer types
ind = which(dat1$cancer %in% NA)
dat1 = dat1[-ind,]

# selecting genes with >5% NAs
dat_status = df_status(dat1[, -20503])
genes_withNA = subset(dat_status$variable, dat_status$p_na > 0)
ind = which(colnames(missing_dat) %in% genes_withNA)
ind = c(ind, 20503)
missing_dat_5 = missing_dat[,ind]
table(missing_dat_5$cancer)

# selecting genes with >10% NAs
genes_withNA = subset(dat_status$variable, dat_status$p_na > 10)
ind = which(colnames(missing_dat) %in% genes_withNA)
ind = c(ind, 20503)
missing_dat_10 = missing_dat[,ind]
table(missing_dat_10$cancer)

# find where those genes are in the dataset and remove them
ind = which(colnames(dat1) %in% genes_withNA)
dat1 = dat1[, -ind]
dim(dat1) #3045 genes are removed
dat_status = df_status(dat1)
dat_imputed = dat1
# there seems to be a small subset of patients with a lot of NAs (>4000), we will remove these
patients from further analysis
# missing_dat = dat1[!complete.cases(dat1), , drop = FALSE]
# table(missing_dat$cancer)

```

```

# dat_status = df_status(missing_dat)
# dat_imputed = dat1[complete.cases(dat1), , drop = FALSE] # keep those samples with complete
cases, i.e. no NAs
# we removed 299 patients (173 LAML, 33 PRAD and 38 STAD)
dim(dat_imputed) #we now have 10770 samples and 17458 genes
table(sapply(dat_imputed,class)) #check class

#move the "cancer" column to the front
which(colnames(dat_imputed) == "cancer")
dat_imputed = dat_imputed[,c(16335,1:16334)]
# dat_imputed = dat_imputed[,colSums(dat_imputed[, -1]) > 0] #keep genes (columns) with column
sum > 0
# now check again to confirm no missing data
# dat_status = df_status(dat_imputed, print_results = F)
# no genes are removed
#### upper quartile normalization ####
dat_imputed_backup = dat_imputed #backup data first
write.csv(dat_imputed,"20190830_post_NA_EB+
+AdjustPANCAN_IlluminaHiSeq_RNASeqV2.geneExp.tsv")

#### how many normal tissues per cancer
library(stringr)
dat_imputed$number = str_sub(rownames(dat_imputed),start=-2)
table(dat_imputed$number)
TP = c("01","02","05")
TM = c("06","07")
NT = "11"
dat_norm$group = ifelse(dat_norm$number %in% TP,"TP",
                        ifelse(dat_norm$number %in% TM, "TM", "NT"))
with(dat_imputed,table(cancer,number))

#### extract and store cancer type meta in another df
cancer_type_meta = as.data.frame(dat_imputed[,1])
rownames(cancer_type_meta) = rownames(dat_imputed)

#### now look at how the data looks like (whether normalization is done/done well) by box-and-
whisker plot
library(dplyr)
library(reshape2)
library(ggplot2)

x = melt(dat_imputed[,3:ncol(dat_imputed)],id.vars = "cancer")
mads = aggregate(. ~cancer+variable, data=x, mad, na.rm=TRUE)
cancers = unique(mads$cancer)
df_mads = subset(mads,mads$cancer==cancers[1])
pancan_mads_summary = as.matrix(summary(df_mads$value))
for (i in 2:length(cancers)){
  df_mads = subset(mads,mads$cancer==cancers[i])
  summ = as.matrix(summary(df_mads$value))
  pancan_mads_summary = cbind(pancan_mads_summary,summ)
}
pancan_mads_summary = as.data.frame(pancan_mads_summary)
colnames(pancan_mads_summary) = cancers

ind = which(df_mads$value > summary(df_mads$value)[3])

```

```

cancers = unique(dat_imputed$cancer)

i=32
dat = subset(dat_imputed, dat_imputed$cancer == cancers[i])
dat = dat[,4:ncol(dat)]
mn = apply(dat, 1, mean)
sd = apply(dat, 1, sd)
med = apply(dat, 1, median)
q1 = apply(dat, 1, FUN = function(i) quantile(i, probs = 0.25, na.rm = T))
q3 = apply(dat, 1, FUN = function(i) quantile(i, probs = 0.75, na.rm = T))
min = apply(dat, 1, FUN = min)
max = apply(dat, 1, FUN = max)
toplot = cbind.data.frame(mn, sd, med, q1, q3, min, max)

colnames(toplot) = c("mean", "sd", "median", "q1", "q3", "min", "max")
toplot$sample = rownames(toplot)
p <- ggplot(toplot, aes(x = sample, y = mean))
p <- p + geom_errorbar(aes(ymin = min, ymax = max), linetype = 1, width = 0.5) #main range
p <- p + geom_crossbar(aes(y = median, ymin = q1, ymax = q3), linetype = 1, fill = 'white') #box
p <- p + geom_errorbar(aes(ymin = mean - sd, ymax = mean + sd), linetype = 3, width = 0.5) #sd range
p <- p + geom_point() # mean
p + theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
  labs(title = cancers[i], x = "Samples", y = "Gene expression (Normalized RSEM values)")

#### calculate the median absolute deviation for each gene in each cancer for removing genes
with low variance before subsequent analyses ####
x = melt(dat_norm[,3:ncol(dat_norm)], id.vars = "cancer")
mads_pancan = aggregate(.~variable, data = x[, -1], mad, na.rm = TRUE)
summary(mads_pancan)
ind = which(mads_pancan$value < 0.65)
keep = unique(as.character(mads_pancan$variable[ind]))
ind = which(colnames(dat_norm) %in% keep)
dat_mads_filtered = dat_norm[, ind]

#####
####task: TP53BP2 expression normal vs tumor
####annotate normal and tumor samples
library(stringr)
dat_norm$number = str_sub(rownames(dat_norm), start = -2)
table(dat_norm$number)
TP = c("01", "02", "05")
TM = c("06", "07")
NT = "11"
dat_norm$group = ifelse(dat_norm$number %in% TP, "TP",
  ifelse(dat_norm$number %in% TM, "TM", "NT"))
table(dat_norm$group)
dat_norm = dat_norm[, c(1, 13253, 2:13251)]
dat_norm[1:5, 1:5]
table(sapply(dat_norm, class))
p = ggplot(subset(dat_norm, subset = dat_norm$cancer == "BRCA"),
  aes(TP53BP2, colour = group)) + geom_density()
p = ggplot(dat_norm,
  aes(TP53BP2, colour = group)) + geom_density() + facet_wrap(~ cancer, ncol = 4, nrow = 8)

p

```

```
####task: TP53BP2 expression over/underexpression using reference population (mean
expression in population diploid for TP53BP2)
CNV_dat = fread("Gistic2_CopyNumber_Gistic2_all_thresholded.by_genes",sep = "\t",header = T,
stringsAsFactors = F)
table(sapply(CNV_dat,class))
CNV_dat_t = as.data.frame(t(CNV_dat[,-1]))
table(sapply(CNV_dat_t,class))
colnames(CNV_dat_t) = CNV_dat$Sample
CNV_dat_t[1:5,1:5]
CNV_dat_t$cancer = dat_norm$cancer[match(rownames(CNV_dat_t),rownames(dat_norm))]
ind = which(colnames(CNV_dat_t) %in% c("TP53BP2","cancer"))
CNV_target = CNV_dat_t[,ind]
CNV_target = CNV_target[complete.cases(CNV_target), , drop = FALSE]
CNV_target$mrna = dat_norm$TP53BP2[match(rownames(CNV_target),rownames(dat_norm))]
colnames(CNV_target) = c("cnv","cancer","mrna")
pancan_expressionStatus = data.frame(cancer = NA, zscore = NA, status = NA)
total_expression_table = data.frame(cancer = NA, status = NA, freq = NA, p = NA)
for( i in 1:length(cancers)){
  df = subset(CNV_target, CNV_target$cancer == cancers[i])
  diploid = subset(df, df$cnv == 0)
  df$zscore = (df$mrna - mean(diploid$mrna)) / sd(diploid$mrna)
  df$status = ifelse(df$zscore >= 2, "overexpressed",
                    ifelse(df$zscore <= -2, "underexpressed", "normal"))
  pancan_expressionStatus = rbind.data.frame(pancan_expressionStatus, df[,c(2,4,5)])
  expression_table = as.data.frame(t(table(df$status)))
  expression_table$Var1 = cancers[i]
  prob_expression_table = as.data.frame(t(prop.table(table(df$status))))
  prob_expression_table$Var1 = cancers[i]
  expression_table$p = prob_expression_table$Freq
  colnames(expression_table) = c("cancer","status","freq","p")
  total_expression_table = rbind.data.frame(total_expression_table,expression_table)
}
total_expression_table = total_expression_table[-1,]
pancan_expressionStatus = pancan_expressionStatus[-1,]
pancan_expressionStatus$mrna =
dat_norm$TP53BP2[match(rownames(pancan_expressionStatus),rownames(dat_norm))]
# function for number of observations
give.n <- function(x){
  return(c(y = mean(x)*-0.6, label = length(x)))
  # experiment with the multiplier to find the perfect position
}

p = ggplot(pancan_expressionStatus,
  aes(y = mrna, x = status, fill = status)) + geom_violin(trim = FALSE) + facet_wrap(~ cancer,
ncol = 7, nrow = 5)
p + scale_fill_brewer(palette = "Dark2")+ stat_summary(fun.data = "mean_sdl", fun.args = list(mult
= 1), geom = "pointrange", color = "black") +
  labs(y = "ASPP2 normalized RSEM values") + stat_summary(fun.data = give.n, geom = "text",
fun.y = median,
  position = position_dodge(width = 0.75), size = 3.5) +
  theme_bw() + theme(axis.text.x = element_blank(),axis.title.x = element_blank(),panel.grid =
element_blank())

total_expression_table$cancer = factor(total_expression_table$cancer, levels =
c("CHOL","LIHC","BRCA","LUAD","UCEC",
"OV","SKCM","CESC","COAD","LUSC",
```

```

"ACC","PAAD","READ","STAD","ESCA","HNSC",
"UVM","PCPG","BLCA","TGCT","GBM","THYM",
"UCS","SARC","THCA","KIRC","DLBC","KIRP",
"LGG","PRAD","KICH","MESO"))

library(tidyverse)
ggplot(total_expression_table %>% arrange(status, desc(p)),
  aes(x=status,y=p, fill = status)) +
  geom_bar(stat="identity", position="dodge") +
  facet_wrap(. ~ cancer, ncol = 7, nrow = 5) +
  theme_classic() +
  theme(panel.spacing=unit(0,"pt"),
    panel.border=element_rect(colour="grey50", fill=NA))+ scale_fill_brewer(palette = "Dark2")+
  theme_bw() + theme(axis.text.x = element_blank(),axis.title.x = element_blank(),panel.grid =
  element_blank())

#add patient age and sex to the dataset
dat_norm$age =
meta$age_at_initial_pathologic_diagnosis[match(rownames(dat_norm),meta$sample)]
dat_norm$sex = meta$gender[match(rownames(dat_norm),meta$sample)]
summary(dat_norm$age)
table(dat_norm$sex)
dat_norm = dat_norm[,c(13178,13179,1:13177)]
dat_norm[1:5,1:5]
which(colnames(dat_norm) == "TP53BP2")
dat_norm = dat_norm[,c(1:4,11955,5:11954,11956:ncol(dat_norm))]
table(sapply(dat_norm,class))

#### Regression model ####
#### we use g with a gaussian fit, i.e. linear regression, with sex, age and TP53BP2 as predictors.
Regression performed individually with every single gene as outcome variable.
#### we collect beta coefficients of the predictors as well as p-values from the model;

#### check scna confounding effects from same chromosome
pancan_cnv = fread("X:\\James Chan\\Gistic2_CopyNumber_Gistic2_all_data_by_genes",sep =
"\t",header=T,stringsAsFactors = F)
head(colnames(pancan_cnv))
head(rownames(pancan_cnv))
pancan_cnv[1:5,1:5]
rownames(pancan_cnv) = pancan_cnv$Sample
ind = which(rownames(pancan_cnv) %in% colnames(dat_mads_filtered))
pancan_cnv = pancan_cnv[ind,]
genes = pancan_cnv$Sample
ind = which(colnames(pancan_cnv) %in% dat_mads_filtered$samples)
pancan_cnv = pancan_cnv[,..ind]
t_pancan_cnv = t.data.frame(pancan_cnv)
t_pancan_cnv[1:5,1:5]
colnames(t_pancan_cnv) = genes
t_pancan_cnv[1:5,1:5]
t_pancan_cnv = as.data.frame(t_pancan_cnv)
ind = which(dat_mads_filtered$samples %in% rownames(t_pancan_cnv))
variables_2 = dat_mads_filtered[ind,]
variables_2$number = str_sub(variables_2$samples,start=-2)
table(variables_2$number)
TP = c("01","02","03")

```



```

TM = c("06","07")
NT = "11"
variables_2$group = ifelse(variables_2$number %in% TP,"TP",
                           ifelse(variables_2$number %in% TM, "TM", "NT"))
variables_2 = distinct(variables_2,samples,.keep_all = T)
rownames(variables_2) = variables_2$samples
table(variables_2$group)
variables_2 = subset(variables_2,variables_2$group == "TP")

ind = which(rownames(t_pancan_cnv) %in% rownames(variables_2))
t_pancan_cnv = t_pancan_cnv[ind,]
ind = which(colnames(variables_2) %in% colnames(t_pancan_cnv))
variables_2 = variables_2[,ind]

variables_2$cancer = meta$cancer.type.abbreviation[match(rownames(variables_2),
meta$sample)]
variables_2[1:5,1:5]
variables_2 = variables_2[,c(9607,1:9606)]
head(variables_2$cancer)

t_pancan_cnv$cancer = meta$cancer.type.abbreviation[match(rownames(t_pancan_cnv),
meta$sample)]
head(t_pancan_cnv$cancer)

t_pancan_cnv = t_pancan_cnv[match(rownames(variables_2), rownames(t_pancan_cnv)), ]
head(rownames(t_pancan_cnv))
head(rownames(variables_2))
t_pancan_cnv = t_pancan_cnv[,match(colnames(variables_2), colnames(t_pancan_cnv)) ]
head(colnames(t_pancan_cnv))
head(colnames(variables_2))
table(sapply(t_pancan_cnv,class))
head(t_pancan_cnv$cancer)
variables_2[1:10,1:10]
t_pancan_cnv[1:10,1:10]
tail(colnames(t_pancan_cnv))
tail(colnames(variables_2))
head(colnames(t_pancan_cnv))
head(colnames(variables_2))
#### performing linear regression ####
table(sapply(variables_2,class))
variables_2 = variables_2[complete.cases(variables_2),]
ind = which(rownames(t_pancan_cnv) %in% rownames(variables_2))
t_pancan_cnv = t_pancan_cnv[ind,]
ind = which(colnames(t_pancan_cnv) %in% colnames(variables_2))
t_pancan_cnv = t_pancan_cnv[,ind]

t_pancan_cnv = t_pancan_cnv[match(rownames(variables_2), rownames(t_pancan_cnv)), ]
tail(rownames(t_pancan_cnv))
tail(rownames(variables_2))
t_pancan_cnv = t_pancan_cnv[,match(colnames(variables_2), colnames(t_pancan_cnv)) ]
tail(colnames(t_pancan_cnv))
tail(colnames(variables_2))
head(colnames(t_pancan_cnv))
head(colnames(variables_2))
cancers = unique(variables_2$cancer)

```

```

y = list()
for (i in 1:length(cancers)){
  var = subset(variables_2,subset = variables_2$cancer==cancers[i])
  var = log2(var[,-1] + 1)
  var[,1:ncol(var)] = scale(var[,1:ncol(var)])
  x = which(is.na(var), arr.ind=TRUE)
  x = unique(x[,2])
  y = cbind(y,x)
  print(x)
}
y = unique(unlist(y))

y = list()
for (i in 1:length(cancers)){
  cnv = subset(variables_2,subset = variables_2$cancer==cancers[i])
  cnv = cnv[,-1]
  cnv[,1:ncol(cnv)] = scale(cnv[,1:ncol(cnv)])
  x = which(is.na(cnv), arr.ind=TRUE)
  x = unique(x[,2])
  y = cbind(y,x)
  print(x)
}
y = unique(unlist(y))
pan_cancer_p_values_3 = matrix(ncol=0,nrow=ncol(variables_2[,2:ncol(variables_2)]))
pan_cancer_c_values_3 = matrix(ncol=0,nrow=ncol(variables_2[,2:ncol(variables_2)]))
cancers = unique(variables_2$cancer)
cancers = cancers[complete.cases(cancers)]
for (i in 1:length(cancers)){
  var = subset(variables_2,subset = variables_2$cancer==cancers[i])
  var = log2(var[,-1] + 1) #log2 transform
  var[,1:ncol(var)] = scale(var[,1:ncol(var)]) #standardization
  cnv = subset(t_pancan_cnv,subset = t_pancan_cnv$cancer==cancers[i])
  cnv = cnv[,-1]
  cnv[,1:ncol(cnv)] = scale(cnv[,1:ncol(cnv)])
  print(cancers[i])

  p_values_3 = matrix(nrow=0,ncol=1)
  c_values_3 = matrix(nrow=0,ncol=1)

  for (gene in 1:ncol(var)){
    mlm3 = lm(var[,gene] ~ cnv$TP53BP2 + cnv[,gene] + var$TP53BP2, data = var)
    p3 = summary(mlm3)$coefficients["var$TP53BP2",4]
    p_values_3 = rbind(p_values_3,p3)
    c3 = summary(mlm3)$coefficients["var$TP53BP2",1]
    c_values_3 = rbind(c_values_3,c3)
  }
  rownames(p_values_3) = colnames(var)[1:ncol(var)]
  rownames(c_values_3) = colnames(var)[1:ncol(var)]
  pan_cancer_p_values_3 = cbind(pan_cancer_p_values_3,p_values_3)
  pan_cancer_c_values_3 = cbind(pan_cancer_c_values_3,c_values_3)
  colnames(pan_cancer_p_values_3)[i] = cancers[i]
  colnames(pan_cancer_c_values_3)[i] = cancers[i]
  print(cancers[i])
}
pancan_c_val_aspp2 = melt(pan_cancer_c_values_3)
pancan_p_val_aspp2 = melt(pan_cancer_p_values_3)

```

```

pancan_all_melt_aspp2 = cbind.data.frame(pancan_c_val_aspp2,pancan_p_val_aspp2$value)
colnames(pancan_all_melt_aspp2) = c("gene","cancer","coefficient","pval")
pancan_all_melt_aspp2$adjp = p.adjust(pancan_all_melt_aspp2$pval,method = "fdr")
pancan_all_melt_aspp2$score = ifelse(pancan_all_melt_aspp2$adjp < 0.05,
pancan_all_melt_aspp2$coefficient,0)
sapply(pancan_all_melt_aspp2,class)
pancan_all_melt_aspp2$cancer = as.character(pancan_all_melt_aspp2$cancer)
cancers = unique(as.character(pancan_all_melt_aspp2$cancer))
uniquegene = unique(as.character(pancan_all_melt_aspp2$gene))
wg = matrix(nrow=0,ncol=2)
for (i in 1:length(uniquegene)){
  gene_ind =
subset(pancan_all_melt_aspp2,subset=pancan_all_melt_aspp2$gene==uniquegene[i])

  sum_score = sum(gene_ind$score)
  avg_score = sum_score/length(cancers)
  wg = rbind.data.frame(wg,avg_score)
}

rownames(wg) = uniquegene
wg = as.data.frame(wg)
wg$gene = rownames(wg)
wg <- wg[order(wg$avg_score, decreasing=TRUE),]

targets <- wg$avg_score
names(targets) = wg$gene
#### GSEA performed by GSEA Desktop version 4.0.1 ####

```