UNIVERSITAT DE BARCELONA

Role of DPPA3 in hypoxia and tumour dormancy in cancer

Estefania Cuesta Borràs

Programa de Doctorat en Biomedicina

Facultat de Biologia,

UNIVERSITAT DE BARCELONA

ROLE OF DPPA3 IN HYPOXIA AND TUMOUR DORMANCY IN CANCER

Memòria presentada per

Estefania Cuesta Borràs

Per optar al grau de Doctora per la Universitat de Barcelona

Treball realitzat sota la direcció del Dr. Héctor García Palmer i la Dra. Isabel Puig Borreil i tutelat per la Dra. Montserrat Jaumot Pijoan en el laboratori de Cèl·lules Mare i Càncer dins el programa de Recerca Translacional de l'Institut d'Oncologia de la Vall d'Hebron (VHIO).

Dr. Héctor García Palmer	Dra. Isabel Puig Borreil	Dra. Montserrat Jaumot Pijoan
Director	Directora	Tutora

Estefania Cuesta Borràs

Doctoranda

Barcelona, setembre 2019

UNIVERSITAT DE BARCELONA

VHIO VALL D'HEBRON Institute of Oncology

No paris mai de lluitar, esforçar-te i sobretot, creure en tu mateix.
Tu i només tu ets capaç d'aconseguir tot allò que et proposis.

ACKNOWLEDGEMENTS

Sembla mentida que amb tantes pàgines escrites no trobi les paraules per descriure tot el que sento en pensar en tota la gent i emocions úniques que m'han acompanyat durant aquests darrers cinc anys. Ha estat un camí que no ha estat fàcil, però cadascuna de les experiències que he viscut, tant a nivell professional com personal, m'han portat a ser una persona més forta, autònoma i capaç d'afrontar nous obstacles. I tot això no tindria sentit sense cap de les persones que m'han acompanyat durant aquests anys.

Quan vaig començar al lab, el VHIO encara estava creixent. Tot i això, ja hi havia la Irene, l'Oriol i la Lorena. Irene, me conociste cuando aún no tenía ni idea de cómo procesar un tumor. Gracias a tu paciencia y apoyo durante todo este tiempo he aprendido una barbaridad de cosas, eres imprescindible. Oriol, encara que a vegades et mataria, en el fons no sé què faria sense tu, ja ho saps ;). Lorena, has sido mucho más que una compañera, me has apoyado en los peores momentos, preocupándote siempre por mi, dentro y fuera del lab, te quiero (un poco). Ahora ya no me puedes llamar pequeño saltamontes, aunque yo te dejo si quieres... Jordi, tot i que sempre et fiquis amb mi, sé que ho has amb carinyo, sempre et capaç de treure'ns un somriure. Laia, de les últimes en entrar al grup però no per això menys important. Tot el contrari, la teva forma de ser (i els teus gustos musicals ;)) són genials, el teu recolzament i empatia han estat realment importants aquests últims mesos. Àlex i Clàudia, els iogurins del lab, els millors (i pitjors potser...) moments de la tesi encara estan per venir, i n'estic convençuda que ho fareu genial. Moltes gràcies per la vostra ajuda sempre que ho he necessitat.

Héctor, moltíssimes gràcies per confiar en mi i donar-me l'oportunitat única de créixer al teu grup, com a professional però encara més com a persona. I Isa, no sé que hagués fet sense tu, ja ho saps. Has estat sempre allà, has escoltat les meves mil i una històries, i sempre, sempre, m'has recolzat. Has estat la millor mentora, suport i referent que mai hauria pogut tenir. Ets la millor i un exemple a seguir.

Però resulta que fora del lab hi havia gent espectacular. Gemma, ràpidament des que vas arribar, encara no sé exactament com, però hi va haver un feeling i connexió que ens van portar al que som ara, amigues per damunt de companyes, disposades a ajudar-nos en el que faci falta. El teu suport ha estat imprescindible durant tot aquest temps, ets una de les millors coses que em puc emportar del VHIO. Mireia, quina sort vaig tenir d'estar al mateix despatx que tu i poder-te conèixer, ets un sol. Sandra y Garazi, sois geniales, con vuestras charlas y cafés los domingos pasan mucho más rápido. Queralt, ets la bomba, ja ho saps, m'encanta com de divertida pots arribar a ser, poder-te explicar totes les novetats de la

setmana, que em donis consell i em recolzis sempre que calgui. Laia, tu també ets una iogurina, però amb molta experiència ja, espero que celebrem com toca aquest final d'etapa (i continuació de la teva!). Hi ha moltíssima gent al VHIO i VHIR, tant predocs com tècnics i postdocs que valen tantíssim la pena, un exemple de suport i cooperació admirable.

Guillem, Pelli i Marina, tot i el pas del temps i les grans distàncies, sempre demostreu que sou allà pel que calgui. Marina, sé que passaran els anys però tu sempre hi seràs i et preocuparàs per mi, ets una gran amiga.

Bet, des d'aquell primer dia a la sala de cultius del PRBB fins ara... amigues i companyes de tesi, explicant-nos tots els drames però també passant moments genials (i els que ens queden...!). No paro de pensar en tot el que farem quan acabem. I Aina, has estat una gran sorpresa per mi, una persona genial que vull conservar. Elo, pienso mucho en ti, y en las ganas que tengo de escaparme para verte otra vez. A pesar de la distancia, ahí estás, dispuesta a escucharme y ayudarme en lo que necesite. Aleix, tot i la distància que hi hagi hagut entremig, gràcies per haver estat una constant.

I les meves amigues de la uni, les meves nenes, què hagués fet sense vosaltres? Com més passa el temps, més me n'adono com d'importants sou per mi. Sou la meva família i el meu suport en tot moment. Laura, Sara, Natàlia i Nina, ni totes les paraules del món podrien descriure com us necessito i estimo. Molta gent ve i va, però vosaltres sempre hi sou i sereu. Berta, conxi, amiga incondicional, estic tan contenta d'haver-te conegut i pensar tots els moments que hem viscut, i tots els que encara ens queden per viure...

Pau, ets una de les persones que més em coneix, que més m'ha estimat i estimo. M'has fet créixer tant com a persona, i m'has recolzat tantíssim durant tots aquests anys que és impensable escriure aquest text sense que tu hi siguis present. Estic tan orgullosa de tu com tu de mi.

I per acabar, la meva família. Papa i mama, tot i els problemes que hagin pogut sorgir, totes les coses bones i dolentes que hem passat, no importen perquè vosaltres m'estimeu i sou allà. Mama, gràcies per haver-me ajudat durant tot aquest temps, de cuidar-me i estimar-me. Ets tan forta i capaç de tot el que et proposis, t'estimo. Iaia Magda, ets la iaia més enrotllada i divertida que ningú podria tenir, em fa tan feliç tenir-te. Iaia Blasa, siempre te tendré en el corazón, te quiero. Tata, no sé si ets més amiga, germana o una segona mare. És increïble la de voltes que dóna la vida, però passi el que passi, el teu amor està per sobre de totes les coses. Ets una de les persones més importants de la meva vida, t'estimo molt.

INDEX

FIGURE INDEX

INTRODUCTION

BACKGROUND

RESULTS

DISCUSSION

TABLE INDEX

INTRODUCTION

MATERIALS AND METHODS

RESULTS

ANNEX

ABSTRACT

Colorectal cancer (CRC) is a leading cause of death worldwide and tumour recurrence is a frequent complication that arises from minimal residual disease and shows up after a period of clinical dormancy. Slow-cycling cancer cells (SCCC), also called dormant tumour cells, have been shown to be responsible for tumour relapse due to their enhanced chemoresistance and tumour-initiating capacity. Although the recent efforts invested in the characterization of SCCC, our knowledge about the mechanisms underlying tumour dormancy is still limited. Thanks to the identification, isolation and molecular characterization of SCCC in our laboratory, we identified a set of pluripotency factors overexpressed in these cells, among them, DPPA3. In the present work, we characterized for the first time the role of DPPA3 in the biology of tumour dormancy. DPPA3 is an epigenetic factor essential for early development and predominantly expressed in embryonic stem cells (ESCs) and primordial germ cells (PGCs). Its function is linked to the protection of imprinted loci and transposable elements in the genome from active demethylation processes. Besides maintaining a repressive state in specific loci, DPPA3 is also related to the passive demethylation phenomenon observed in cells at these developmental stages. By the use of genetically modified CRC cell lines we revealed a central role of DPPA3 promoting cell dormancy. We unmasked its capacity controlling the response to hypoxia as a key mechanism to govern cancer cell phenotype. DPPA3 overexpression stimulated the hypoxia program by increasing hypoxia inducible factor 1 subunit alpha (HIF1α) protein levels and enhancing its transcriptional activity. Besides, overexpression of DPPA3 alone was sufficient to induce a G2/M-phase cell cycle arrest and reduce tumour growth. Interestingly, DPPA3 enhanced chemoresistance to CRC standard of care adjuvant chemotherapies. Finally, a cohort of CRC patients with high expression of *DPPA3* or enriched in a DPPA3 signature showed a shorter disease-free survival. Altogether, these results pioneer the importance of DPPA3 in cancer and contribute to the understanding of tumour malignancy associated to hypoxia, chemoresistance and dormancy, the unravelling of which is of foremost importance to progress in the battle against the disease.

RESUM

El càncer colorectal (CRC) presenta una elevada mortalitat arreu del món i la recurrència és una complicació freqüent que sorgeix com a conseqüència de la malaltia mínima residual que es presenta després d'un període de latència clínica. Diversos estudis han descrit la presència de cèl·lules canceroses amb una baixa taxa proliferativa (*slow-cycling cancer cells*, SCCC) responsables de la recaiguda del pacient a causa de la seva elevada quimioresistència i capacitat regenerativa de tumors. Tot i l'esforç invertit en la caracterització de les SCCC, anomenades també cèl·lules dorments, el nostre coneixement sobre els mecanismes moleculars subjacents a la seva latència és encara limitat. Gràcies a la identificació, aïllament i caracterització molecular de les SCCC feta al nostre laboratori, vam poder identificar un conjunt de factors de pluripotència sobreexpressats en aquestes cèl·lules, entre ells, DPPA3. DPPA3 és un factor epigenètic essencial per al desenvolupament embrionari i expressat en cèl·lules mare embrionàries (ESCs) i cèl·lules germinals primordials (PGCs). La seva funció està relacionada amb la protecció en loci improntats i elements transposables del genoma de processos de desmetilació activa. A part de mantenir un estat repressiu en regions concretes del genoma, DPPA3 també està relacionat amb processos de desmetilació passiva durant el desenvolupament embrionari. Amb l'ús de línies de CRC modificades genèticament per sobreexpressar exògenament o reduir els nivells endògens de DPPA3, vam poder esbrinar el rol que exerceix aquest factor en els programes d'hipòxia i latència tumoral. La sobreexpressió de DPPA3 en línies de CRC va induir el programa d'hipòxia augmentant els nivells de proteïna i activitat transcripcional de HIF1α. A més, també va ser suficient per generar una parada del cicle cel·lular en la fase G2/M *in vitro* i un retard del creixement tumoral i major quimioresistència a tractaments d'adjuvància estàndards en CRC *in vivo*. Finalment, vam observar que una elevada expressió de *DPPA3* en pacients de CRC està associada amb una major quimioresistència. El conjunt d'aquests resultats contribueixen a la comprensió de la rellevància de DPPA3 en càncer i la malignitat tumoral associada a hipòxia, quimioresistència i latència tumoral.

ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
4E-BP1	Eukaryotic translation initiation factor 4E-binding protein 1
5-FU	5-fluorouracil
ABC	ATP-binding cassette
AMPK	5′-AMP-activated protein kinase
APC	Adenomatous polyposis coli
APOBEC1	Apolipoprotein B mRNA editing enzyme catalytic subunit 1
ARNT	Asaryl hydrocarbon nuclear translocator
AUC	Area under the curve
bHLH	Basic-helix-loop-helix
BMP	Bone morphogenetic protein
BrdU	5-bromo-2'-deoxyuridine
CA9	Carbonic anhydrase 9
CBC	Columnar base cell
CDK	Cyclin-dependent kinase
CDKN1A	Cyclin dependent kinase inhibitor 1A
CRC	Colorectal cancer
CSC	Cancer stem cell
CUL2	Cullin 2
DNMT1	DNA methyltransferase 1
DNMT3B	DNA methyltransferase 3 beta
DOX	Doxycycline
DPPA3	Developmental Pluripotency Associated 3
DTC	Disseminated tumour cell
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EHMT2	Euchromatic histone lysine methyltransferase 2
eIF2α	Eukaryotic initiation factor 2 alpha
EIF4A1	Eukaryotic translation initiation factor 4A1
EIF4E	Elongation initiation factor 4E
EIF4F	Eukaryotic initiation factor 4F
EIF4G1	Eukaryotic translation initiation factor 4 gamma 1
EIF4G2	Eukaryotic translation initiation factor 4 gamma 2

ELOB	Elongin B
ELOC	Elongin C
EMT	Epithelial-to-mesenchymal transition
EPO	Erythropoietin
ER–	Estrogen receptor negative
ER+	Estrogen receptor positive
ERK	Extracellular signal-regulated kinases
ES	Enrichment score
ESC	Embryonic stem cell
FC	Fold change
FIH1	Factor-inhibiting HIF1α
G9a	Euchromatic histone lysine methyltransferase 2
GCNT3	Glucosaminyl (N-acetyl) transferase 3, mucin type
GDF3	Growth differentiation factor 3
H2BeGFP	Histone H2B-enhanced green fluorescent protein
HEPH	Hephaestin
HES	Hairy/enhancer of split
HIF	Hypoxia inducible factor
HIF1A	Hypoxia inducible factor 1 subunit alpha
HMT	Histone methyl transferase
HNSCC	Head and neck squamous-cell carcinoma
HRE	Hypoxia response element
HSP90	Heat shock protein 90
IPO5	Importin 5
IRES	Internal ribosomal entry sites
ISC	Intestinal stem cell
KDM	Histone lysine demethylase
KO	Knockout
LDHA	Lactate dehydrogenase A
LEF	Lymphoid enhancer factors
LGR5	Leucine-rich G protein-coupled receptor 5
LRCs	Label-retaining cells
Luc	Luciferase
MAPK	Mitogen-activated kinase
MDM2	MDM2 proto-oncogene
MDR1	Multidrug resistance protein 1

MRD	Minimal residual disease
mTOR	Mammalian target of rapamycin
MUC2	Mucin 2
MYC	MYC proto-oncogene, BHLH transcription factor
NANOG	Nanog homeobox
NCID	NOTCH intracellular domain
OCT4	POU Class 5 Homeobox 1
ODD	Oxygen-dependent degradation
OLFM4	Olfactomedin 4
PAS	Per and sim
PcG	Polycomb group
PCNA	Proliferating cell nuclear antigen
PDK1	Pyruvate dehydrogenase kinase 1
PDX	Patient-derived xenograft
PERK	Protein kinase R (PKR)-like endoplasmic reticulum kinase
PFS	Progression-free survival
PGC	Primordial germ cell
PHD	Prolyl hydroxylase domain
PRC1	Polycomb repressive complex 1
RARB	Retinoic acid receptor beta
RB1	RB transcriptional corepressor 1
RBX1	Ring-box 1
RCCC	Rapid-cycling cancer cells
Rluc	*Renilla reniformis* luciferase
ROS	Reactive oxygen species
S6	Ribosomal protein S6
S6K1	Ribosomal Protein S6 Kinase B1
SCCC	Slow-cycling cancer cells
SLC2A1	Solute carrier family 2 member 1
SLC2A3	Solute carrier family 2 member 3
SOX2	SRY-box 2
TA	Transit-amplifying
TAD	Transactivation domain
TCF	T-cell factor
TET	Ten-eleven Translocation
TGCT	Testicular germ cell tumour

TGFβ	Tumour growth factor β
TP53	Tumour Protein P53
TRE	Tetracycline responsive element
TSS	Transcription start site
UBE2D1	Ubiquitin conjugating enzyme E2 D1
UHRF1	Ubiquitin like with PHD and ring finger domains 1
uPAR	Urokinase plasminogen activator receptor
UPR	Unfolded protein response
UTR	Untranslated region
VEGFA	Vascular endothelial growth factor A
VHL	Von Hippel-Lindau tumor suppressor
WT	Wild-type

INTRODUCTION

1. The colonic mucosa

The surface of the colonic mucosa is essentially flat and contains deep finger-like invaginations that increase in depth towards the rectum and reach the muscularis propia. These pits are termed colonic crypts and millions of them populate the colon. The intestinal crypts contain multipotential stem cells capable of regenerating all intestinal cell types [1]. Three major terminally differentiated types of cells are present in the colonic crypts: the absorptive enterocytes, the mucous-secreting goblet cells and the hormone-secreting enteroendocrine cells [2,3]. Stem cells present at the colonic crypt base produce the proliferative cells located in the transit-amplifying (TA) compartment. Cells in the TA zone eventually commit to one of the three mature cell lineages, which move towards the top of the crypt and eventually undergo cell death [1] (Figure I1).

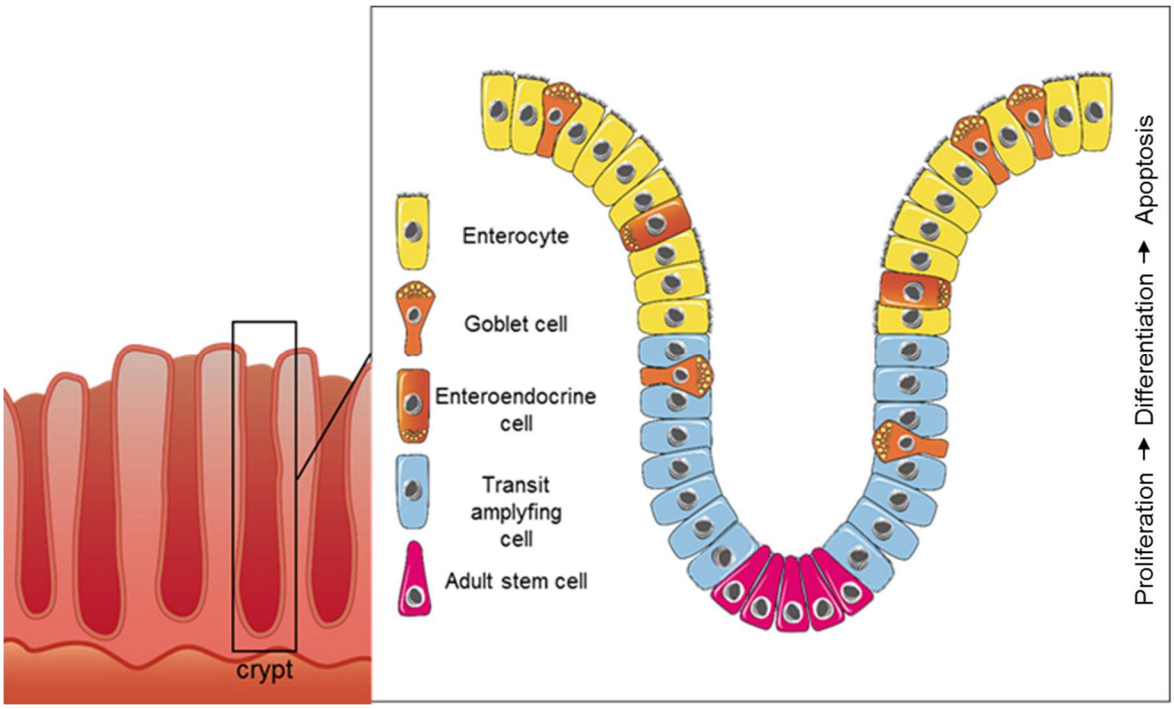


Figure I1. Schematic representation of the colonic crypt. Adult stem cells at crypt bottoms generate the cells located in the transit-amplifying zone, where most of cell production occurs. As cells migrate towards the top of the crypt they differentiate into mature cell types: enterocytes, goblet and enteroendocrine cells.

Several markers can be used to track the differentiation state in colonic cells. The number of goblet cells is much higher in the large than the small intestine. Mucin 2 (MUC2), a highly O-glycosylated protein is broadly used as a marker of goblet-cell differentiation [4]. Glucosaminyl (N-acetyl) transferase 3, mucin type (GCNT3) is involved in the biosynthesis of mucins and can also be used as a marker of cell differentiation [5]. HEPH is used as another marker of differentiated cells and it is necessary for the transport of dietary iron from the absorptive cells

to the circulatory system and it is expressed in the small and large intestine [6,7]. Thus, the system is highly dependent on a continued supply of the different cellular types. The intestinal stem cells (ISCs) present at the bottom of the crypt are in charge of the continued fuelling of these demands [1,8,9].

1.1. Intestinal stem cell compartment

The architecture in the small intestinal crypt is the following: the bottom of the crypt is made up by columnar base cells (CBCs) that are mingled between Paneth cells, which are not present in the colon. The +4 cells are localized right above the Paneth cells. CBCs, Paneth and +4 cells together form the so-called stem cell zone. ISCs generate the proliferating progenitors located in the TA zone, which commit toward the different intestinal lineages as they migrate upwards onto the tip of the villus [1,8,9]. Two different ISC models arouse from the study of the intestinal crypt:

- **The +4 model**: it was first mentioned in 1965 [10] and later on confirmed by Potten and colleagues by cell tracing experiments which proposed that the crypts originate from a cell positioned between the +2 to +7 position (on average, +4) [11]. *BMI1* is the most well-known marker of the +4 cells and encodes a component of the Polycomb group (PcG) proteins, mostly known for their repressive functions [9]. Although being recognised as a +4 cell marker, *BMI1* expression has also been observed in CBCs [12].
- **The Stem Cell Zone model**: it was first described in 1970 and identified CBCs as a specific type of cycling cells immersed between the Paneth cells, which constitute the niche for the CBCs. CBCs proliferate and generate a progeny that exits the niche and commit to the different lineages [13]. The leucine-rich G protein-coupled receptor 5 (*LGR5*) is a Wnt/β-catenin target gene and a selective marker of CBCs [12], an observation that underlines the importance of this signalling pathway in the stem cell compartment. LGR5-positive cells can generate self-renewing intestinal organoids *in vitro* [14]. Olfactomedin 4 (OLFM4) is another target of Wnt/β-catenin signalling and a marker of CBCs [15]. CD133, also known as Prominin-1, has been described to mark CBCs. However, it is likely to be expressed throughout the stem cell and TA compartment [16].

Despite the crypt structure in the small intestine is of great importance for the understanding of the functioning in the colonic crypts, we must be aware of some important discrepancies between the small and the large intestinal crypts. Mainly the small intestine harbours Paneth and +4/BMI1-positive cells while the colon does not. However, LGR5-positive cells do exist in the colon crypt base and define the colonic stem cells. Elevated NOTCH signalling is also associated with colon stem cells [17,18]. Stemness is a state that is lost when cells leave the

stem cell zone, but can also be regained when differentiated cells re-enter the niche. Thus, the tight balance between specific signalling pathways present in the colonic cells is necessary for stem cell maintenance, regeneration and differentiation [19–21].

1.2. Essential crypt signalling pathways

The epithelium and mesenchyme cooperate to generate the dynamic complex observed in intestinal crypts. Together, they provide the key signals that regulate intestinal fate determination: Wnt/β-catenin, NOTCH, Hedgehog, epidermal growth factor (EGF) and bone morphogenetic proteins (BMP) signalling [22] (Figure I2). For simplicity, only the signalling pathways essential for the interpretation of the results section will be described.

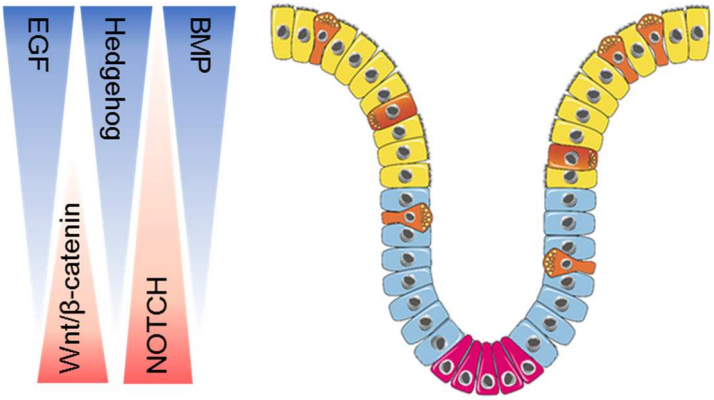


Figure I2. Signalling pathways regulating intestinal cell fate determination. The tight balance between the different signalling pathways including Wnt/β-catenin, NOTCH, EGF, BMP and Hedgehog regulate stem cell activity, being Wnt/β-catenin and NOTCH more active in the crypt base while EGF, BMP and Hedgehog signalling gradually increase towards the top of the crypt.

1.2.1. Wnt/β-catenin signalling pathway

Wnt/β-catenin signalling is pivotal in cell proliferation, polarity and fate determination during embryonic development as well as the maintenance of stem cells in adult tissues. For this reason, alterations in this pathway can result in several diseases including cancer [23].

β-catenin is the main effector of the canonical Wnt signalling. When not stimulated, β-catenin locates to the cytoplasmic membrane interacting with cadherins, which increase its stability. β-catenin can be released from cadherins through post-translational modifications like phosphorylation. Once in the cytoplasm, β-catenin is rapidly degraded by the Wnt degradation complex. This process takes place thanks to GSK-3 activity, which together with axin and adenomatous polyposis coli (APC), phosphorylate β-catenin and send it to degradation in the proteasome [24]. Wnt ligands activate the canonical β-catenin signalling by interacting with the seven-span transmembrane protein Frizzled and the single-span transmembrane protein

LRP. Wnt stimulation inhibits β-catenin degradation pathway, making possible its entrance to the nucleus. Once there, β-catenin acts as a transcriptional co-activator binding T-cell factor/lymphoid enhancer factors (TCF/LEF) family [25–27].

Nuclear β-catenin gradually decreases along the intestinal crypt, being higher at the base and inhibited at the top, where cells differentiate [28]. This gradient underlines the importance of the signalling in the maintenance of the stem cell population and the proliferation of TA cells in both colonic and intestinal crypts.

1.2.2. NOTCH signalling pathway

NOTCH signalling plays a fundamental role in the determination of cell fate and regulates the balance between cell proliferation, differentiation and apoptosis [29,30].

NOTCH is a transmembrane protein that is encoded by four different genes (*NOTCH1-4*) in mammals and it is proteolically cleaved by the interaction with their ligands (Delta/Delta-like/Jagged/Serrate) [31]. Upon ligand interaction, a γ-secretase protease cleaves the receptotor, releasing the NOTCH intracellular domain (NCID) to the cytoplasm, which finally translocates into the nucleus. Once there, NCID forms a complex with one of the CSL transcription factors and co-activator MAML-1 to induce target gene transcription [30,32]. Hairy/enhancer of split (*HES*) are the best-characterized NOTCH target genes, especially *HES1* in the intestine [33]. Depletion of *Hes1* in mice causes a significant increase in the secretory cell lineage [33]. HES1 supresses the expression of *KLF4*, a transcriptional repressor that is expressed in the differentiated cells of the intestine [34,35]. Elevated NOTCH1 activity is present in the proliferative TA zone of the colonic crypt [36] and suppression of NOTCH signalling directs intestinal enterocyte progenitors to differentiate towards the secretory lineage [31].

2. Colorectal cancer: overview and molecular pathogenesis

Colorectal cancer (CRC) emerges as a consequence of the accumulation of acquired genetic and epigenetic changes that lead to a transformation of the healthy epithelium to a malignant adenocarcinoma. Fearon and Vogelstein proposed the classic tumour progression model that involves a stepwise formation of CRC. They described that it initiates as benign adenomatous polyps in the mucosa that, with time, accumulate genetic mutations that eventually transform them into carcinomas which show high invasiveness and metastatic ability. The term used to define this process is called adenoma-carcinoma sequence and involves the inactivation of tumour suppressor genes like *APC* or *tumour protein P53* (*TP53*), while oncogenes like *KRAS* are activated [37,38]. Nevertheless, new genomic and epigenomic approaches have led to great insights into the nature of CRC. Apart from the genes already identified as drivers of the

disease, other genes such as *PI3KCA* and *BRAF* show an elevated mutation rate [39,40]. Even though cancer cells often exhibit many characteristics of stem cells, the precise cellular origin of colorectal cancer is still unknown. Two opposite models have been proposed: (1) the "top-down" model, in which the tumour initiates at the top of the crypt and then spreads laterally and consequently downwards towards the normal crypt [41]; and (2) the "bottom-up" model, which proposed that the stem cell at the bottom of the crypt is the tumour initiating cell that eventually populates the entire crypt [42].

2.1. CRC cancer recurrence and metastasis

A frequent complication in CRC is tumour recurrence due to resilient cancer cells that eventually resume growth. Around 95% of stage I CRC patients will not relapse upon treatment, whereas in 20%–25% of stage II and 40%–50% of stage III patients, disease will show up again [43]. Most conventional and target-directed drugs (e.g. oxaliplatin, 5-fluorouracil, and irinotecan) eliminate active proliferating cancer cells. Due to the intratumoural heterogeneity, in many occasions, a population of chemoresistant slow-cycling or dormant residual cells is left behind after several rounds of chemotherapy. This small population able to hide and protect itself against live-threating environmental conditions is nowadays one of the main battles that cancer research is still fighting [44]. Tumour recurrence can appear in the site of primary tumour formation or as a metastasis in a distant organ. The liver is one of the most significant targets for organ-specific metastases and a major cause of mortality in patients with CRC [45]. Lungs represent another frequent colonized organ by CRC metastases: blood is drained from the colon through the portal system to the liver. From the liver the next organ is the lung, via the heart [46]. Altogether, colorectal tumour recurrence either within the primary tumour site or in a distant organ represents a major cause of death in patients with advanced CRC. Hence, the biological mechanisms underlying the nature of chemoresistant cells represents a field of continuous investigations.

3. Dormancy and Slow-Cycling Cancer Cells

3.1. Overview and definition of dormancy

In the classical view of cancer, unlimited and uncontrollable cell proliferation were attributed to the disease. However, a growing body of data has demonstrated that cancer cells are able not only to proliferate but also to alternate reversible periods of *slow-cycliness* or dormancy with periods of rapid growth. Dormancy periods are often indolent and can be overlooked by clinicians until the minimal residual disease (MRD) initiates proliferation when is detected as a recurrence. Metastases originated from disseminated tumour cells (DTCs) that persist after treatment as a MRD usually undergo a period of dormancy [47]. Classically, clinical dormancy

is described as the time that it takes for cancer to relapse after clinical remission, either as a local recurrence or as a metastatic disease [48,49]. Clinically detectable distant metastases can occur within a time period ranging from weeks to decades. The lapse of time between primary tumour detection and metastatic relapse is often defined as latency and the duration of metastatic latency varies between cancer types (Figure I3). Lung cancer is an example of short latency resulting in 5-year survival rates around 17% [50]. Sequential metastasis to liver and lungs is often observed in CRC progression, and more than 85% of recurrences are detected within the first 4 years of follow-up in advanced tumours [51]. Therefore, this particular type of cancer shows medium latency and aggressiveness, resulting in a 5-year survival rate of 65% [50]. Prostate cancer is a well-known example of a tumour type with very long latency in which nearly 100% of diagnosed patients survive the first 5 years, and 82% are still alive after 15 years [50]. In the case of breast cancer and in contrast to the other types of carcinomas discussed above, it can be classified as either medium or long latent disease depending on the volume, stage, and molecular subtype of the primary tumour [52].

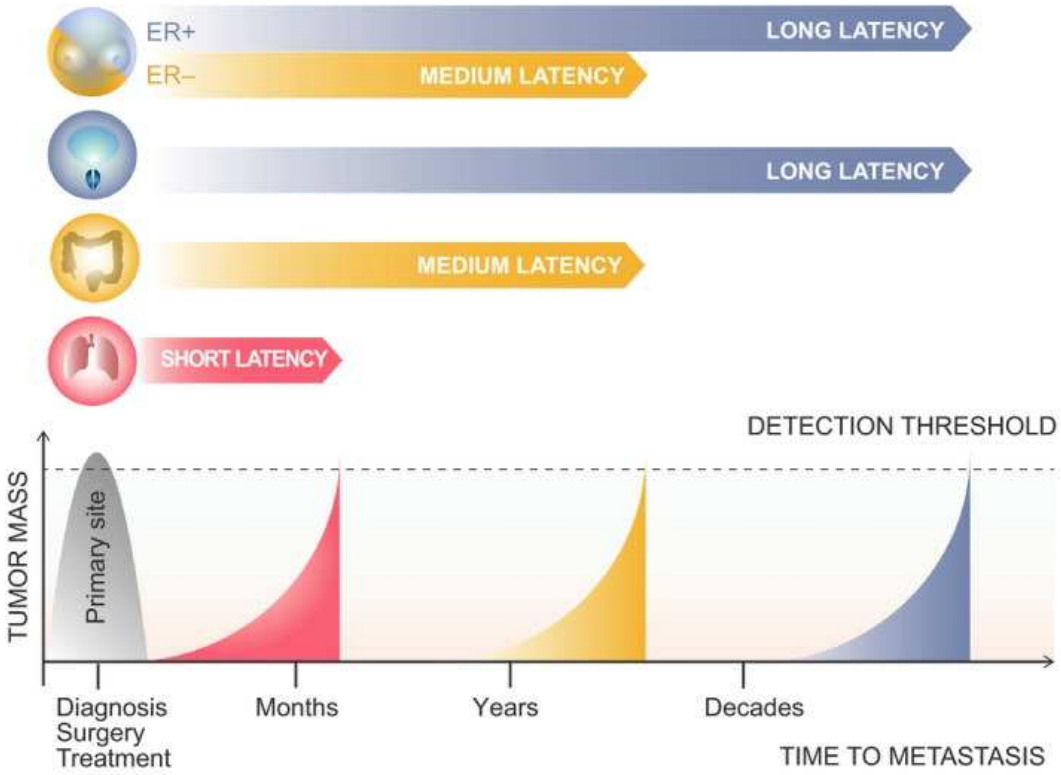


Figure I3. Temporal course of cancer metastasis. Metastatic relapse might occur within months, years or decades after primary tumour diagnosis, removal and treatment. Depending on the cancer type, length of the latency will vary: short for lung cancer (red), middle for CRC and estrogen receptor negative (ER–) breast cancer (yellow), and long for prostate cancer and estrogen receptor positive (ER+) breast cancer (blue). Dashed line indicates threshold of detection symptomatic metastases [52].

Dormancy is a reversible state in which cells stop proliferating due to a myriad of molecular mechanisms that can occur at the individual cellular level or as a consequence of the biological behaviour of whole tumoural mass (Figure I4):

- **Cancer cell dormancy or quiescence**: a single cancer cell arrests cell cycle in a reversible way due to cell-intrinsic mechanisms.
- **Tumour dormancy**: the whole tumour mass stops growing because of a balanced growth/apoptosis rate as a response to cell-extrinsic mechanisms.

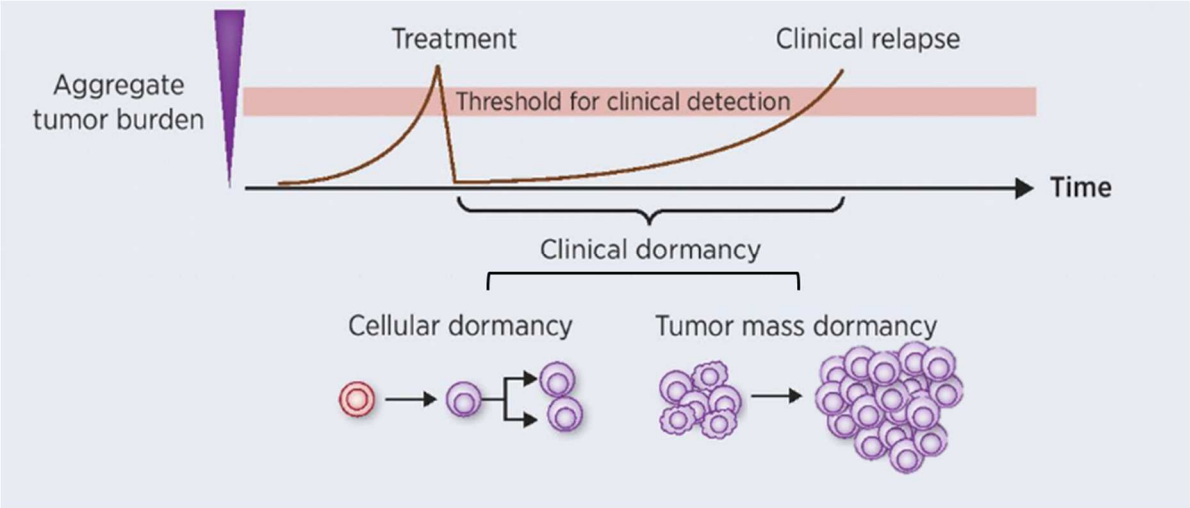


Figure I4. Mechanisms of cancer dormancy. Top: Clinical dormancy represents the lapse of time necessary for the residual cancer cells to regrow and become detectable after clinical remission. Bottom: Cancer dormancy can be physiologically distinguished as cellular dormancy/quiescence or tumour dormancy. Quiescent cells can resume their proliferation and re-enter cell cycle. In a dormant tumour, there is a balance of the proliferative versus apoptotic rate. When proliferation outweighs apoptosis, the cellular mass expands and re-generates a detectable tumour. Adapted from [49].

3.2. Contextualization of dormancy

Reversibility is a key feature of dormancy that makes possible the transition to this state in different contexts [53]:

- **Primary cancer dormancy**: cancer cells in the primary tumour can alternate their cycling state in order to confer them an increased cell fitness among the whole cancer population.
- **Metastatic dormancy**: DTCs can spread from the primary tumour to distant sites and remain there in a dormant state until favourable microenvironmental conditions arise. Accumulating evidence suggests that tumour cells can already disseminate from the primary region at the onset of tumour development even before the full acquisition of malignant properties [47].

- **Therapy-induced dormancy**: it consists in the enrichment of dormant cells upon cancer drug treatment. Cells can eventually awake giving rise to an even more aggressive disease.
- **Immune-mediated dormancy**: tumours get rid of immunogenic cells and remaining dormant cells are able to evade the immunologic response.

3.3. Molecular mechanisms driving dormancy

The **niche** that surrounds tumour cells plays a critical role in the induction of a slow-cycling state. The different niches that induce the entry to dormancy are:

- **Bone marrow stem cell niche**: it is the most common organ for DTCs homing in colon, breast, prostate and head and neck cancers and it favours maintenance of dormancy [54].
- **Perivascular niche**: while stable microvasculature creates a dormant niche, sprouting neovasculature induces micrometastatic outgrowth. It has been reported that some dormant cancer cells can inhabit the perivascular niches [55].
- **Cancer stem cell (CSC) niche**: the ability of some CSCs to remain quiescent can be due to the niche where they are located. This niche can induce the dormancy of other non-CSCs [53].
- **Metastatic niche**: the colonised organ presents a stressful environment, making DTCs enter dormancy to overcome these conditions [56,57].

Angiogenic dormancy defines the inability of cancer cells to induce a successful angiogenic process leading them to dormancy. In this situation, there is a balance between proliferation and cell death caused by a hypoxic microenvironment. The angiogenic switch leads to the acquisition of the necessary angiogenic potential that results in the escape from the slow-cycling state towards the initiation of a proliferative, growing tumoural mass [58].

The tumour **microenvironment** is another dormancy inductor. Some key factors governing dormancy entry are epidermal growth factor receptor (EGFR), urokinase plasminogen activator receptor (uPAR), extracellular signal-regulated kinases (ERK) and p38 mitogen-activated kinase (MAPK). The p38 MAPKs group belongs to the stress-activated MAPK family and consists of four different members in mammals (p38α, p38β, p38γ and p38δ). Whereas p38α and p38β are widely expressed and better characterized, p38γ and p38δ appear to have a more tissue specific expression pattern, being restricted to skeletal muscle in the case of p38γ and to testes, pancreas, kidney and small intestine for p38δ [59,60]. The p38 MAPK pathway has been shown to be activated in dormant tumour cells [61,62] and its signalling in hematopoietic stem cells restricts proliferation while it promotes entry to dormancy [63]. Disruption of the uPAR

complex activates p38 MAPK signalling and the inhibition of the Raf-MEK-ERK pathway. High p38/ERK signalling ratio results in cell cycle arrest at G0-G1 controlled by the upregulation of *TP53*, *NR2F1* and *BHLHB3* and the downregulation of *FOXM1* and *JUN*, which promote G1 exit [61,62,64,65]. Oxygen deficit is another microenvironmental cue that stimulates the entry to dormancy in order to allow the survival of cells by reducing their rate of oxygen and energy consumption [48].

Given the reversible nature of dormant cells, **epigenetic** modifications have been proposed to exert profound effects in the entry, maintenance and reactivation of quiescence [66]. Environmental pressure increases with time and drug treatments, making tumoural cells evolve and gain epigenetic modifications that drive them to a dormant state. DNA methylation and histone post-translational modifications constitute some of these epigenetic mechanisms regulating dormancy. Treatment with 5-azacytidine demethylating agent results in a decreased expression of *DNA methyltransferase 1* (*DNMT1*) and *FOXM1* (involved in G0/G1 exit) in haematological and epithelial tumour cells. On the other hand, 5-azacytidine treatment induces the expression of *retinoic acid receptor beta* (*RARB)* and *cyclin dependent kinase inhibitor 1A* (*CDKN1A)*, p38-known targets in dormant head and neck squamous-cell carcinoma (HNSCC) cells [59,67]. The promoter of the *RARB* nuclear receptor is commonly hyper-methylated in several types of cancers, while this gene is re-expressed in dormancy models, suggesting that epigenetic regulation can orchestrate the regulation of these programs. NR1H3 is another nuclear receptor induced in dormant cells although its specific role remains unknown [59]. In light of these results, it has been suggested that specific RARβ agonists, among others, or inhibitors of FOXM1 or DNMT1 could be used to reprogram tumour cells into dormancy [66]. Retinoic acid, tumour growth factor β (TGFβ) and BMPs constitute some other examples that affect chromatin structure, dictate cell fate and are important inductors of dormancy [68–71].

3.4. Dormancy and chemoresistance

The virulence of dormant cancer cells relies on their reversible status and high resistance to conventional chemotherapeutic agents, which are usually designed to target hyper-proliferative cells. As a result, chemotherapies often lead to an enrichment of dormant cancer cells. Apart from reduced cell proliferation, other mechanisms like autophagy and stress-related pathways (e.g. p38 signalling or unfolded protein response, UPR) are thought to influence on their survival [72].

Given that conventional anti-proliferative chemotherapies fail to eliminate dormant tumour cells, new molecular approaches for their eradication are emerging [72].

- Dormancy inductors, which keep cells dormant.
- Molecules that eradicate cells already dormant.
- Awakening dormant tumour cells with conventional anti-proliferative therapies. This approach is controversial, as treatment of highly proliferative lesions does not always completely halt their progression. Moreover, residual DTCs are hetereogeneous. Therefore, awakening them would possibly make us face highly aggressive cells that resist to most current therapies.

3.5. Hypoxia as an inductor of dormancy

3.5.1. Effects of tumour hypoxia on angiogenic dormancy

Tumours require a functional vasculature just like normal tissues. Therefore, in areas devoid of enough vasculature to feed the uncontrolled proliferating population of cancer cells, cell death may take place, leading to an equilibrium between proliferation and apoptosis that keeps the tumour mass dormant [72]. The concept of angiogenic dormancy refers to the balance of proliferation versus cell death due to the lack of angiogenesis. Microscopic tumours (<2 mm diameter) have been shown to be avascular due to the inability of the cells to stimulate the formation of blood vessels. Lack of angiogenesis results in a worse blood supply that favours a hypoxic microenvironment, which may be causally related to low proliferation rates and thus tumour dormancy [48].This situation can eventually revert and undergo an angiogenic switch once these cells acquire the ability to become vascularized [47].

3.5.2. Hypoxia and cell cycle arrest

Tumour hypoxic areas reduce the rate of oxygen consumption by decreasing proliferation. Hence, tumour cells in hypoxic regions enter a slow-cycling state but yet are viable for prolonged periods of time [73]. Immunohistochemical analyses have demonstrated that hypoxic tumour cells are in a slow-cycling state as most of the population is negative for classic proliferation markers (e.g. proliferating cell nuclear antigen, PCNA, 5-bromo-2'-deoxyuridine, BrdU). Hypoxia-induced quiescence is mainly attributed to a G0/G1 cell cycle arrest or to a disproportionally long G1 phase [48,74]. It has been described that hypoxia decreases cyclin-dependent kinase (CDK) activity, which positively regulate S phase entry through RB transcriptional corepressor 1 (RB1) phosphorylation. The accumulation of hypophosphorylated RB1 (active form) blocks S phase entry and cell growth [75]. RB1 activity is also regulated by the activity of the PP1 phosphatase [76] and hypoxia increases PP1-mediated RB1 dephosphorylation, thus representing an additional block for cell cycle entry [75].

Induction of hypoxia inducible factor 1 subunit alpha (HIF1α), a transcription factor key in the hypoxia program, is sufficient to induce cell cycle arrest [77,78]. Many studies have been

subjected to understand the molecular mechanisms underlying HIF1α on cell proliferation. Some of them have shown that HIF1α counteracts MYC proto-oncogene, BHLH transcription factor (MYC) effects on proliferation by several mechanisms: HIF1α can directly interact with MYC and displace it from its DNA binding sites leading to the derepression of *CDKN1A* and *CDKN1B*, involved in inhibition of cell cycle [79]. In addition, HIF1α can induce the expression of *MXI1*, a MYC antagonist [80].

Histone lysine demethylases (KDMs) are sensitive to oxygen levels and some of them are directly induced by HIF1α. HIF1α upregulates *KDM4B* in CRC cells, which results in lower H3K9me3 in the promoters of HIF1α-induced genes. KDM4B is involved in cell proliferation, apoptosis and cell cycle arrest [81].

The endoplasmic reticulum is a specialized organelle for the maturation of proteins that are destined for membrane expression or secretion [82]. Endoplasmic reticulum stress is any perturbation that compromises the protein folding functionality of the endoplasmic reticulum [83]. Low oxygen levels induce endoplasmic reticulum stress in solid tumours, as hypoxia leads to an accumulation of unfolded proteins and generates reactive oxygen species (ROS) [84]. The UPR is highly conserved and occurs as a consequence of endoplasmic reticulum stress. UPR allows cells to cope with cellular stress triggered by chemical or environmental factors. This pathway controls multiple downstream processes including protein production, maturation and degradation, cell metabolism and cell death [82]. Protein kinase R (PKR)-like endoplasmic reticulum kinase (PERK) signalling is a key component of the UPR and phosphorylates eukaryotic initiation factor 2 alpha (eIF2α) [85], which inhibits global translation while it promotes the translation of specific mRNAs encoding stress-responsive factors, making cells more tolerant to hypoxic conditions [86,87]. PERK activation, however, is independent of HIF1α [84]. Hypoxia-induced PERK activation has been shown to be necessary for the induction of autophagy [88], a mechanism involved in the survival of dormant cells [89]. Abrogation of PERK signalling decreases autophagy thereby sensitizing cells to hypoxia-induced cell death [90]. Moreover, PERK induces a G0-G1 cell cycle arrest and promotes cell survival *in vitro* and inhibits tumour growth in CRC cells [91].

4. Hypoxia

4.1. Definition of hypoxia and basic concepts

Hypoxia is a microenvironmental parameter that is encountered in many pathological conditions like ischemic diseases, atherosclerosis and cancer but also in physiological processes like embryonic development [92–94]. Normoxia, which is supposed to describe “normal” oxygen levels, mostly refers to the oxygen concentrations used in *in vitro* studies,

which is about 20-21%. However, the oxygen concentrations present in peripheral tissues is much lower [95]. These observations have driven to the use of another term that better describes normal physiologic oxygen levels: physoxia. For experimental studies, physoxic conditions are proposed to be 5% oxygen. The physiological hypoxia is the specific oxygen concentration at which each tissue triggers its hypoxic response, which can be by vasodilatation and/or upregulation of hypoxia response genes. The levels of oxygen required to elicit the hypoxia response might vary depending on the tissue affected since normal tissues have different median oxygen levels. In the pathological hypoxia, there is a persistence of poor oxygenation in the affected tissue, which disrupts normal homeostasis. Pathological hypoxia (usually applies below 1%) takes place in solid tumours, in which O_2 levels can drop from 4.2 to 0.3 % (almost all falling below 2%). These data is summarised in Table I1. Yet in many occasions, tumours are still able to persist and even expand. Hypoxia-resistant tumour cells may become quiescent and, eventually, can be selected over other cells that have not adapted properly to this situation [96,97]. Therefore, it is of great importance understanding the mechanisms that make tumours able to survive and evolve under these conditions.

% Oxygen	Description
20-21	Oxygen air at normal atmospheric pressure
9.5	Arterial blood oxygen concentration
5	Physoxia: physiological oxygen level in peripheral tissues with an average of approximately 6%. For experimental studies, 5% is the proposed compromise since this is often used
2	Physiological hypoxia: the hypoxia response is elicited in a specific tissue. The exact oxygen level to trigger the hypoxic response may vary between tissues/cell types
1	Pathological hypoxia: persistence of poor oxygenation

Table I1. Approximate oxygen levels in tissues in physiologic and pathologic conditions. An approximate value has been used for each condition. Adapted from [96].

Pseudohypoxia is another term that was first used to describe hypoxia-like conditions in the pathophysiological context of diabetes [98]. It refers to hypoxia-driven phenotypes (e.g. increased invasion, metabolic reprogramming and stem cell-like characteristics) observed when oxygen levels are normal. In this scenario, aberrant expression of HIF proteins and their target genes in an oxygenated environment create a pseudohypoxic phenotype. Of note, pseudohypoxic phenotypes are often observed in cancer [99].

4.2. The Hypoxia-Inducible Factor

The Hypoxia-Inducible Factor (HIF) is the master regulator of hypoxia at the cellular level and its activity is crucial to elicit the hypoxia response. It is an obligate heteromeric transcription

factor composed of an oxygen labile alpha unit (HIF1α, HIF2α or HIF3α) plus a beta stable subunit (HIF1β or asaryl hydrocarbon nuclear translocator, ARNT). While the alpha subunit is only stable under hypoxic conditions, protein levels of the beta subunit are more stable and independent of O_2. Little is known of HIF3α (*HIF3A* gene) due to its late discovery and restricted expression [100]. HIF1α (*hypoxia inducible factor 1 subunit alpha*, *HIF1A* gene) was first described in 1995 by Semenza and colleagues [101] and two years later, independent groups identified HIF2α (*EPAS1* gene) [102,103]. Whereas *HIF1A* is broadly expressed in all cells, *EPAS1* expression is more restricted, being more abundant in blood vessels [104]. In low oxygen conditions, HIF1α subunits are stabilized, dimerize with HIF1β and translocate to the nucleus [105,106]. Once there, HIF heterodimers recognize and bind to hypoxia response element (HRE) consensus sequences (G/ACGTG), triggering the expression of those genes involved in the hypoxia program [104] (Figure I5). Analysis of HIF-binding motifs has demonstrated the existence of sequences other than the consensus HRE that are susceptible to bind HIF complexes. While a proportion of these sites clusters with annotated promoters, around 30% of HIF1α and 50% of HIF2α binding sites are more than 10 kb away from their nearest annotated gene [107,108]. In cancer, induction of HIF activity results in the activation of genes involved in metabolic reprogramming, cell proliferation, migration, angiogenesis, apoptosis and drug resistance [109].

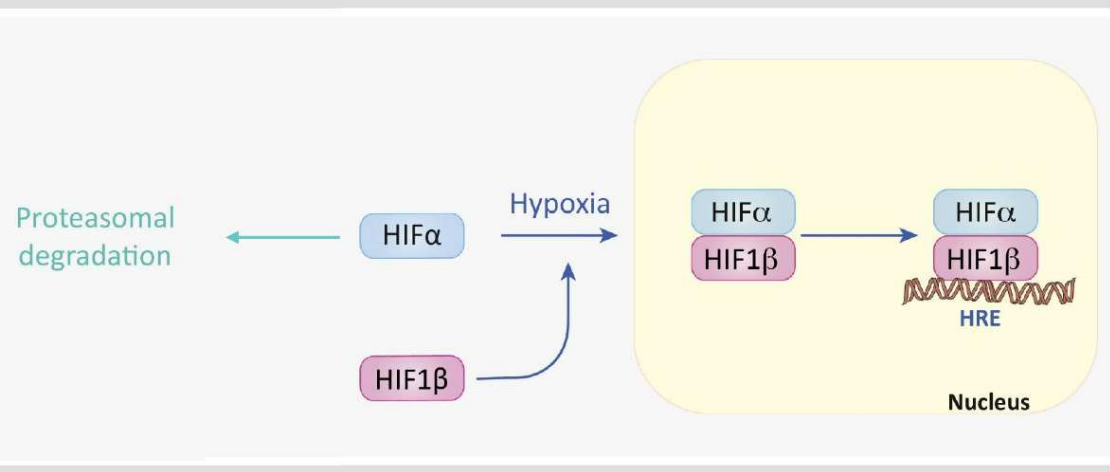


Figure I5. Schematic representation of HIFα stabilization and transactivation in hypoxia. While in normoxic conditions HIF1α is constitutively sent to proteasomal degradation, hypoxia inhibits HIF1α degradation. Stabilized HIF1α dimerizes with HIF1β, translocates to the nucleus and transactivates target genes bearing HREs. Adapted from [110].

4.2.1. Structure of HIF family members

The alpha and beta subunits that constitute the HIF heterodimer belong to the basic-helix-loop-helix (bHLH)- per and sim (PAS) protein family as their structures are related to two

nuclear proteins found in *Drosophila*, PAS, and have a bHLH motif [101]. bHLH-PAS proteins have three functional domains: bHLH, PAS and TAD (transactivation domain), which can regulate the transcription of several family members including their own [111]. The amino-terminal bHLH is necessary for DNA binding and the PAS motifs are required for heterodimerization. In the alpha subunit, HIF1α and HIF2α show a 48% of amino acid sequence similarity, whereas the HIF3α paralog is not that similar [106]. Many of the *HIF3A* splicing variants carry the N-TAD but lack the C-TAD [112,113]. The high degree of similarity between HIF1α and HIF2α explains their common capability to bind the HIF1β unit and indistinguishable DNA sequences [106]. HIF1α and HIF2α have two TADs required for the activation of HIF target genes: the N-terminal TAD (N-TAD) and the C-terminal TAD (C-TAD). The C-TAD acts to regulate the transactivation of target genes common to both HIF1α and HIF2α through coactivator recruitment. The N-TADs of HIF1α and HIF2α confer target selectivity between these two family members [106]. The C-TADs of HIF1α and HIF2α interact with the p300/CBP transcriptional coactivators [114]. p300 and CBP are multidomain proteins with histone acetyltransferase activity that bind the transactivation domains of a variety of transcription factors [115]. In contrast to HIF1β, all HIFα subunits contain oxygen-dependent degradation (ODD) domains overlapping the N-TAD that regulate HIFα stability depending on the oxygen tension [105,106,116]. The ODD contains a key asparagine and proline residues that are targeted for hydroxylation in normoxic conditions [106] (Figure I6).

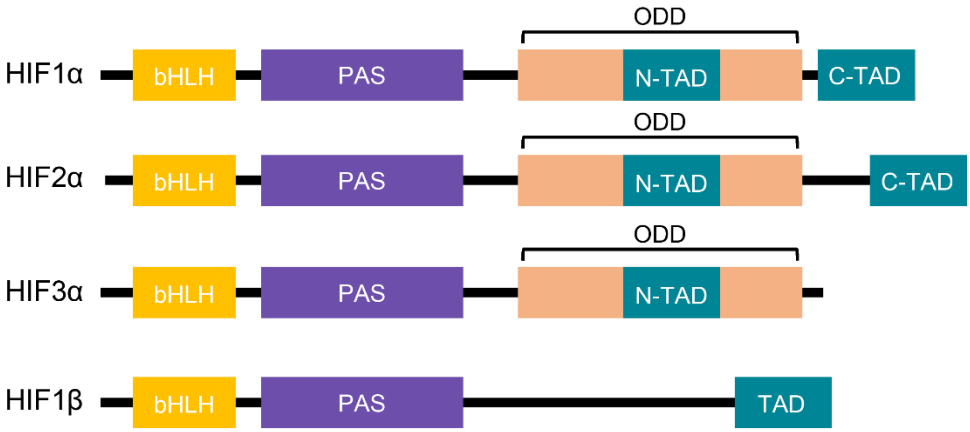


Figure I6. Functional domains of bHLH-PAS family proteins that form the HIF heterodimer. bHLH is involved in DNA binding, PAS in protein-protein interactions, ODD in oxygen-dependent degradation and TADs in transcriptional activation. Whereas both HIF1α and HIF2α have C-TAD and N-TAD, HIF3α has only an N-TAD. HIF1β does not have the ODD. Adapted from [117].

4.2.2. HIF1α-interacting transcription factors

Besides HIF1β, a variety of transcription factors can interact with HIF1α, widening the functional range of HIF1α. TP53 has been shown to interact with HIF1α and induce HIF1α

degradation [118] while β-catenin can bind to HIF1α and enhance its transcription. This dynamic role of β-catenin is thought to be a mechanism for adapting to hypoxia by constraining tumour cell growth [119]. HIF1α and TGFβ signalling can synergise by direct association of HIF1α and SMAD3 to activate *vascular endothelial growth factor A* (*VEGFA*) transcription [120]. MYC is another candidate partner of HIF1α: HIF1α-MYC weak complexes can be formed *in vivo* resulting in a decreased MYC activity and de-repression of *CDKN1A* along with cell cycle arrest [79]. HIF1α can also compete with MYC for SP1 binding and repress *MSH2* and *MSH6* expression leading to an increased genomic instability [121].

4.3. Carbonic anhydrase 9 as a biomarker of hypoxia

Carbonic anhydrase 9 (CA9) is an enzyme involved in pH regulation that catalyses the reversible conversion of CO_2 to bicarbonate and proton [122]. It is expressed during embryonic development, which takes place in a relatively hypoxic environment [123]. In addition, pH regulation by CA9 is a mechanism that can be adopted by tumour cells, which, although maintaining a glycolytic metabolism, are able to regulate the intracellular pH within the physiological limits by extruding lactate and protons out of the cell. Protons accumulation outside cells further contributes to acidosis, a hallmark of solid tumours [124].

The HRE consensus sequence acts as a *cis*-acting element upstream of *CA9* transcription start site that is necessary for the activation of its transcription [122,125]. In contrast to most other hypoxia-inducible genes, HRE in *CA9* gene is exclusively bound by HIF1α [122].

CA9 has been widely used as a biomarker of tumour hypoxia [124,125] but also as a prognostic factor predictive of survival in certain cancers such as astrocytoma, sarcoma, cervix, breast, non-small cell lung and HNSCC [126–131]. In most cases, CA9 and HIF1α protein patterns coincide (HIF1α +/CA9+). However, there are cases in which they do not [132,133]. Besides the HRE, there are other cis-regulatory elements in the *CA9* promoter. The methylation status of specific CpG sites in the *CA9* promoter have been shown to be important for its expression [134–136]. A study showed that HCT116 CRC cells, albeit having a hypermethylated *CA9* promoter, when transfected with an exogenous *CA9* promoter construct, this was strongly induced by hypoxia, suggesting that although the transcriptional machinery required for *CA9* expression is intact, the methylation state on its promoter affects its transcription [122,137].

4.4. Regulation of HIF1α

The activity of HIF1α can be regulated at transcriptional, translational and post-translational levels by multiple molecular pathways. The following sections describe them in more detail.

4.4.1. Oxygen-dependent regulation of HIF1α

Although *HIF1A* is constitutively transcribed and synthesized, it undergoes rapid degradation in well-oxygenated environments having a half-life of approximately 5 minutes [138]. This degradation is mainly due to post-translational modifications, which are the best-understood mechanisms regulating HIF1α stability [105].

4.4.1a. Von Hippel-Lindau tumour suppressor-mediated degradation pathway

The prolyl hydroxylase domain (PHD)-containing enzymes and Von Hippel-Lindau tumour suppressor (VHL) are the two major proteins involved in this pathway. Three isoforms for PHD exist: PHD1 (*EGLN2* gene), PHD2 (*EGLN1* gene) and PHD3 (*EGLN3* gene). PHD2 is the main HIF1α regulator in normoxia among the three isoforms [139]. PHDs use oxygen as substrate and ascorbate, Fe^{2+} and α-ketoglutarate (αKG, also known as 2-oxoglutarate, or 2-OG) as co-substrates. Active PHDs recognize and catalyse the hydroxylation of the two proline residues located in the ODD of HIF1α. These hydroxylated prolines in the ODD trigger VHL-mediated ubiquitination and proteasomal degradation of HIF1α [110,140] (Figure I7).

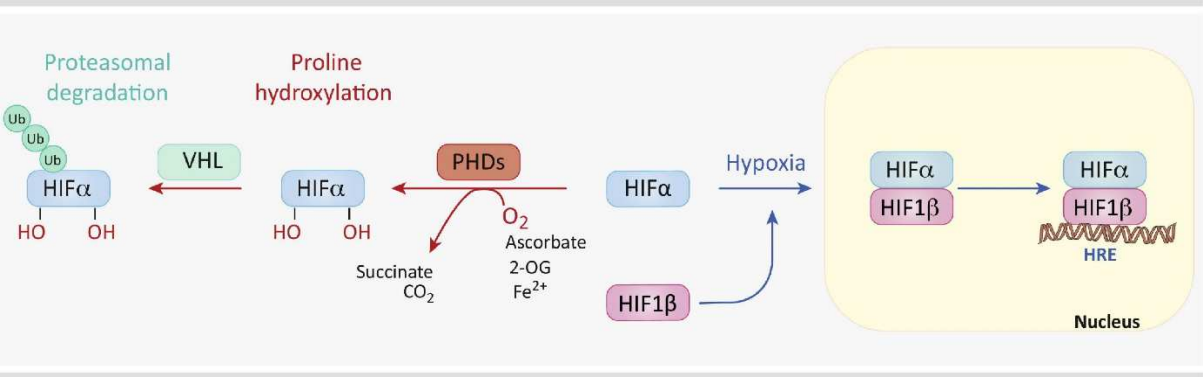


Figure I7. Schematic representation of HIFα stabilization and transactivation in hypoxia. In normoxia, PHDs catalyse the oxygen-dependent hydroxylation (-OH) of HIF1α using 2-OG, ascorbate, and Fe^{2+} as co-substrates. This post-translational modification allows the binding of VHL to HIF1α. VHL subsequently ubiquitinates HIF1α to send it to proteasomal degradation. Hypoxia inhibits PHD activity, which leads to HIF1α stabilization, dimerization with HIF1β, nuclear translocation, and transactivation of target genes with consensus HREs. Adapted from [110].

In HIF1α ubiquitination process, VHL forms a multiprotein complex with elongin C (ELOC), elongin B (ELOB), the RING E3 ubiquitin ligase cullin 2 (CUL2) and ring-box 1 (RBX1) [141–143]. CUL2 acts as a scaffold that brings together the substrate (HIF1α) and the E2-conjugating enzyme. The ubiquitin conjugating enzyme E2 D1 (UBE2D1) is an E2-conjugating enzyme that has been shown to mediate HIF1α degradation [144]. CUL2 then catalyses the transfer of ubiquitins from the E2-conjugating enzyme to HIF1α [145–147] (Figure I8).

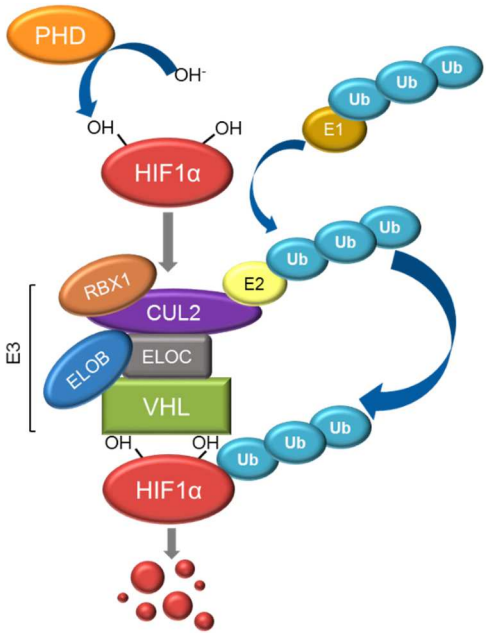


Figure I8. Hydroxylation and ubiquitination of HIF1α in normoxia. In normoxia, PHDs are active and catalyse the hydroxylation of the two proline residues in the ODD of HIF1α. This permits the ubiquitination of HIF1α by the coordinated activity of the ubiquitin-activating E1, the ubiquitin-conjugating E2 and the ubiquitin-ligase E3 complex composed by ELOB, ELOC, RBX1, CUL2 and VHL. Then, HIF1α is degraded by the proteasome-dependent pathway. Adapted from [146].

4.4.1b. Factor-inhibiting HIF1α

Factor-inhibiting HIF1α (FIH1) in an oxygen-sensing asparaginyl hydroxylase that hydroxylates HIF1α causing a disruption between this protein and the coactivators p300/CBP, resulting in an impaired HIF1α transcriptional activity [148].

4.4.1c. Mitochondria as a HIF1α regulator

It has been observed by genetic and pharmacological approaches that the mitochondrial electron transport chain is required for hypoxic HIF1α stabilization. A theory that might explain these observations is that under hypoxia, mitochondria, with their high rate of oxygen consumption, leave the cell devoid of oxygen. This blocks PHDs' activity, resulting in HIF1α stabilization [149,150]. Another model postulates that, in conditions of moderate hypoxia (around 1.5% O_2), mitochondria can stimulate the production of cellular ROS that attenuate PHDs' activity and increase HIF1α protein stability [151].

4.4.2. Oxygen-independent regulation of HIF1α

4.4.2a. Intermediate metabolites

The increase of some metabolites can also affect the stability of HIF1α. High levels of succinate caused by the inhibition of the succinate deshydrogenase, a component of the Krebs

cycle, inhibit PHDs thereby stabilizing HIF1α [152]. Pyruvate and lactate can also contribute to the pseudohypoxic phenotype [153–155] whereas αKG, together with the PHD cofactors ascorbate and Fe^{2+}, destabilize HIF1α at dose-dependent concentrations [156].

4.4.2b. Cap-dependent translation of HIF1α and growth factor signalling pathways

Mammalian target of rapamycin (mTOR) and MAPK/ERK pathways induce a cap-dependent translation of HIF1α. On one hand, the PI3K–AKT–mTOR signalling results in the phosphorylation and activation of the ribosomal protein S6 kinase b1 (S6K1, Figure I9), which in turn, phosphorylates the ribosomal protein S6 (S6) and induces protein translation. On the other hand, mTOR phosphorylates and inhibits elongation initiation factor 4E (EIF4E)-binding proteins like eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1), an inhibitor of translational initiation [105,157,158].

Certain growth factors can activate the MAPK/ERK cascade, which phosphorylates 4E-BP1, S6K1 and MAPK interacting protein kinases (MNK), resulting in an increased rate of HIF1α translation. Besides being involved in HIF1α translation, ERK is also involved in its transcriptional activation by the phosphorylation of the p300/CBP coactivator. This facilitates the formation of the HIF1α-p300/CBP complex and increases HIF1α transcriptional activity [105,157,158].

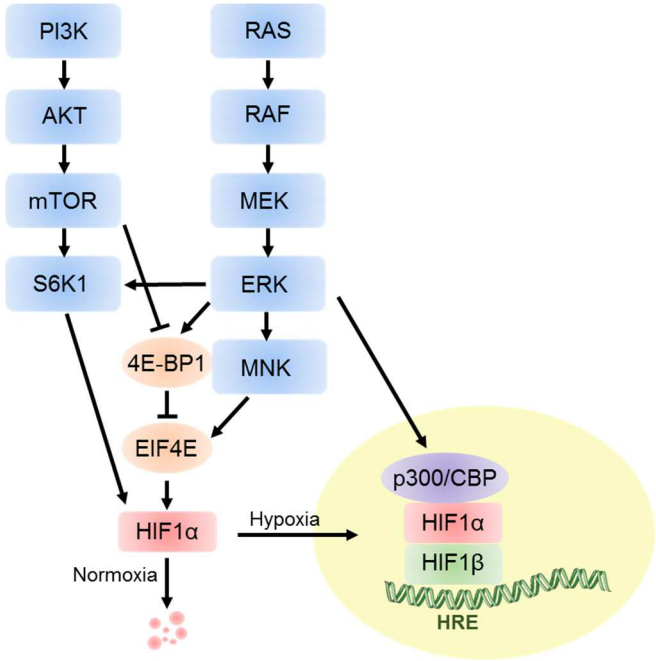


Figure I9. Regulation of HIF1α translation by mTOR and MAPK-ERK pathways. Both PI3K-AKT-mTOR and MAPK-ERK pathways converge to enhance HIF1α synthesis. Besides, ERK phosphorylates and activates the p300/CBP transcriptional coactivator. Adapted from [105]

4.4.2c. MDM2 proto-oncogene

It has been reported that HIF1α can bind TP53 to allow MDM2 proto-oncogene (MDM2)-mediated ubiquitination of HIF1α followed by proteasomal degradation [118].

4.4.2d. Heat shock protein 90

Heat shock protein 90 (HSP90) is a chaperone that prevents the aggregation of unfolded proteins [159]. HSP90 binds HIF1α and induces conformational changes in its structure so it can fit and dimerize with HIF1β, thereby increasing its transcriptional activity [160].

4.4.3. Regulation of *HIF1A* expression

HIF1A is constitutively expressed in many cell types including cancer cells. The promoter of *HIF1A* includes some putative sites in which specific activators or repressors of gene transcription can bind. Several HREs are among these putative sites in the *HIF1A* promoter, suggesting an positive feedback loop [161]. Oxidative stress can also regulate HIF1α at the transcriptional level by the activation of transcription factors that bind *HIF1A* promoter and enhance its expression [162,163].

4.5. Metabolic reprogramming in hypoxia

Under low oxygen situations, healthy cells switch from oxidative to glycolytic metabolism. Tumour cells, instead, use glycolysis more readily even when sufficient oxygen is available. HIF-1α is a master regulator of glycolysis and plays an important role as an activator of aerobic glycolysis and lactate production [164]. During reprogramming of cellular metabolism, HIF1α upregulates *pyruvate dehydrogenase kinase 1* (*PDK1*), which phosphorylates and inactivates the catalytic subunit of the pyruvate dehydrogenase, thereby inhibiting the conversion of pyruvate to acetyl-CoA necessary for the activity of the Krebs cycle [110]. Transcription of lactate dehydrogenase A (*LDHA*), which catalyses the conversion of pyruvate to lactate, is also regulated by HIF1α [165]. The combined action of PDK1 and LDHA blocks the transformation of pyruvate to acetyl-CoA and increases flux from pyruvate to lactate, imposing glycolysis over oxidation [166]. The rewiring of oxidative to glycolytic metabolism is accompanied by an increased extracellular glucose uptake by GLUT1 and GLUT3 glucose transporters, which codifying genes are HIF1α direct targets [164].

4.6. mTOR signalling and hypoxia

The mTOR kinase complex is fundamental for cell growth by stimulating cap-dependent mRNA translation as well as nucleotide and lipid biosynthesis while inhibiting protein turnover by blocking autophagy, lysosomal biogenesis and proteasomal degradation. Besides its

importance in normal tissue homeostasis, it also plays a critical role in tumour growth. Rapamycin-sensitive mTORC1 and insensitive mTORC2 are the two main complexes that form the mTOR system [110]. While mTORC1 is important for protein translation, active mTORC2 signalling inhibits apoptosis, promotes survival and targets proteins important for cytoskeletal organization [167].

mTORC1 is tightly regulated by the coordination of energy stress, DNA damage and hypoxia and regulates cell growth and proliferation by promoting many anabolic processes [168]. mTORC1 has five main components: mTOR, the catalytic subunit, RPTOR, mLST8, PRAS40 and DEPTOR. PRAS40 and DEPTOR are negative regulators of mTORC1 pathway through their physical interaction. When mTORC1 activity is low, PRAS40 and DEPTOR inhibit it, whereas upon mTORC1 activation, it phosphorylates and inactivates PRAS40 and DEPTOR. Among its known targets, mTORC1 phosphorylates and inactivates 4E-BP1, allowing the initiation of cap-dependent translation. mTORC1 also phosphorylates and activates S6K1, which in turn phosphorylates S6, thereby inducing ribosomal biogenesis [169] (Figure I10).

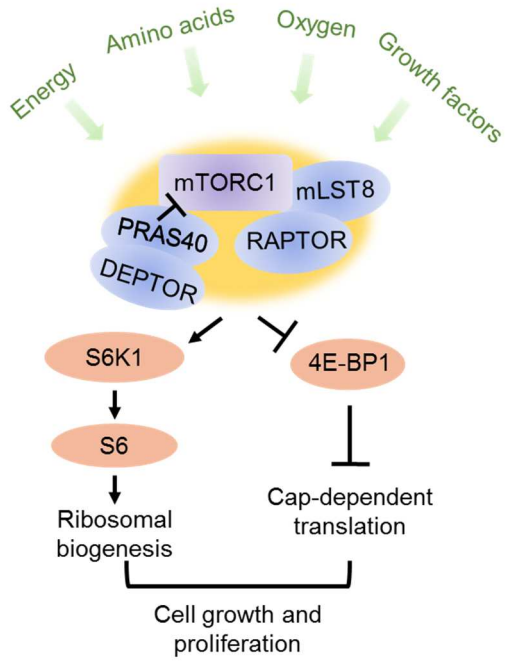


Figure I10. Physiological signals affecting the activation of mTORC1. Energy, amino acids, oxygen and growth factors activate mTORC1 signalling, which regulate a number of cellular processes including cell growth and proliferation.

Hypoxia represses mTOR signalling through a HIF1α-mediated blockade upstream mTOR effectors. This repression is necessary to assure cell survival under low-energy conditions [110]. Various hypoxia-related mTOR inhibitory loops that can depend or not on HIF1α have been described (Figure I11):

- HIF1α directly transactivates *REDD1*. Active REDD1 phosphorylates and inactivates mTORC1. REDD1 activation additionally stablishes a negative feedback loop upon HIF1α by limiting mitochondrial ROS formation [170].
- RHEB, a GTP-binding protein that binds and stimulates mTORC1, can be bound and inhibited by HIF1α. *BNIP3* is a HIF1α transcriptional target and it blocks autophagy by inhibiting mTOR signaling through binding RHEB [171].
- Hypoxic activation of the DNA damage-responsive kinase ATM enhances HIF1α-REDD1 signalling thus reinforcing mTORC1 inhibition [168].
- PML has been shown to be induced by HIF1α in breast cancer [172]. In low oxygen conditions, PML binds and inhibits mTORC1 [173,174].
- 5′-AMP-activated protein kinase (AMPK) is a major sensor of cellular energy levels. High adenosine monophosphate/adenosine triphosphate ratio activates AMPK, which in turn promotes catabolic pathways to generate energy while attenuating energy-consuming processes [175]. Hypoxia stimulates AMPK activity, which in turn inhibits mTOR through activation of the TSC1–TSC2 tumour suppressor complex [176,177].

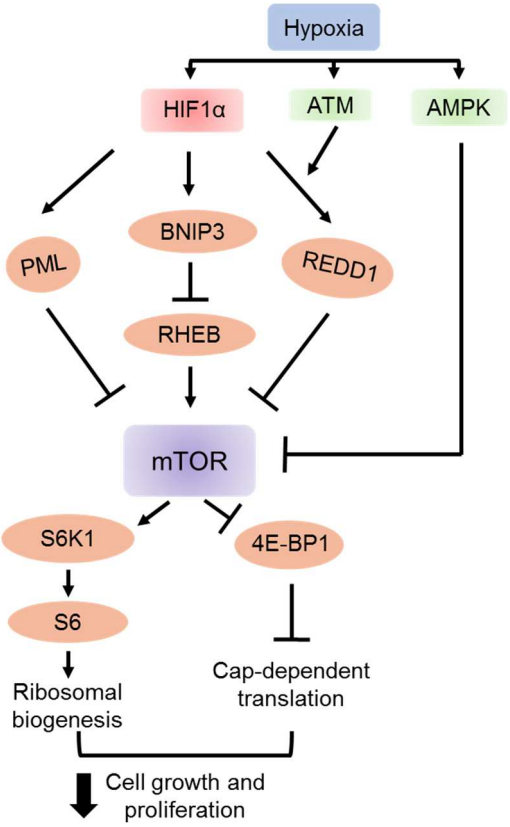


Figure I11. Summary of hypoxia-induced mechanisms known to inhibit mTOR activity. Active mTOR, as a part of mTORC1, stimulates proliferation and cell growth in normal oxygen tensions through the phosphorylation of S6, necessary for ribosomal biogenesis, and 4E-BP1, an inhibitor of the complex necessary to initiate cap-dependent translation. A drop in oxygen levels results in HIF1α stabilization,

which, in turn, induces the transcription of *PML*, *BNIP3* and *REDD1*. Hypoxic activation of ATM enhances HIF1α-REDD1 signalling. The AMPK energy sensor is another mTOR inhibitor that gets active under hypoxia. The coordinated action of AMPK, ATM, PML, BNIP3 and REDD1 attenuate mTOR signalling, thereby decreasing cell growth and proliferation. Adapted from [82].

4.7. Cap-dependent and –independent initiation of protein translation

4.7.1. Cap-dependent initiation of translation

Translation of most mammalian mRNAs is mediated by their 5' cap structure [178]. In the canonical cap-dependent process of translation, eukaryotic initiation factor 4F (EIF4F) binds 5'-capped mRNA. EIF4F is composed of EIF4E, the scaffold eukaryotic translation initiation factor 4 gamma 1 (EIF4G1) and the mRNA helicase eukaryotic translation initiation factor 4A1 (EIF4A1). Hypophosphorylated 4E-BP1 binds EIF4E, thereby preventing the interaction between the EIF4E and the EIF4G and inhibiting initiation of translation [179]. mTORC1 produces an inactivating phosphorylation on 4E-BP1 that allows the ensemble of the EIF4F complex, which can then initiate mRNA translation [180,181] (Figure I12).

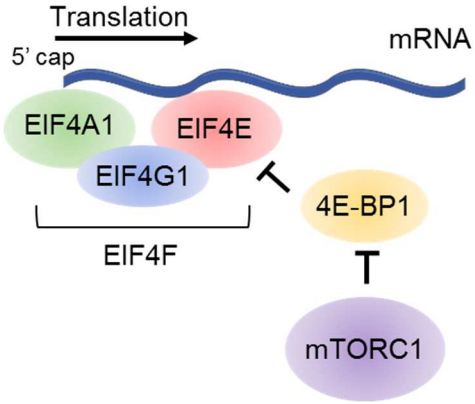


Figure I12. Control of cap-dependent initiation of translation by mTORC1. Hypophosphorylated 4E-BP1 binds EIF4E displacing it from EIF4G1 and inhibiting cap-dependent translation. mTORC1 phosphorylates and inactivates 4E-BP1 making it be unable to bind EIF4E. This results in the assembly of the whole EIF4F complex on the cap 5' permitting mRNA translation. Adapted from [182].

4.7.2. Cap-independent initiation of translation

A strategy to ensure selective translation of the appropriate transcripts in conditions where mTOR signalling is not active relies on the cap-independent initiation of translation mediated by internal ribosomal entry sites (IRES). In this mechanism, the 40S ribosomal subunit is recruited to mRNA near the initiation codon thanks to the IRES, generally located at the 5′-untranslated region (UTR). Under certain stresses such as hypoxia, mTOR activity is reduced,

which permits 4E-BP1 activation (hypophosphorylated state). Active 4E-BP1 sequesters EIF4E, thereby suppressing cap-dependent mRNA translation and promoting translation of IRES-containing mRNAs [183]. EIF4G1 is thought to be involved in the translation of many IRES-containing mRNAs [184]. Importantly, some cellular mRNAs can utilize a dual mechanism of initiation (cap or IRES) including those that orchestrate angiogenesis (*VEGFA*) or hypoxia responses (*HIF1A*) [185,186].

It has been demonstrated that 4E-BP1 and EIF4G1 are strongly overexpressed in the majority of advanced breast cancers. In these tumours both factors facilitate cap-independent mRNA translation in response to hypoxia, promoting translation of IRES-containing mRNAs and ensuring tumour survival and angiogenesis [187] (Figure I13). Eukaryotic translation initiation factor 4 gamma 1 (EIF4G2) is another translation initiation factor that can only mediate cap-independent translation as it lacks the terminus necessary for EIF4E interaction. EIF4G2 has been reported to mediate IRES-dependent translation in human embryonic stem cells (ESCs) [188] and translation of specific mRNAs in quiescent cells and immature oocytes through interactions with a specific complex that binds the 3′-UTR [189].

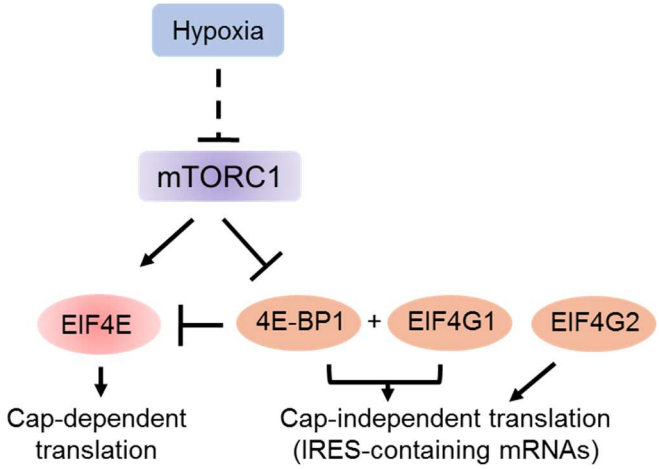


Figure I13. Mechanisms mediating cap-independent initiation of translation. Hypoxia downmodulates mTORC1 activity, which leads to increased hypophosphorylated 4E-BP1 and repression of the cap-dependent translation mediated by EIF4E. On the other hand, 4E-BP1 together with EIF4G1 has been shown to mediate a hypoxia-activated switch to facilitate cap-independent mRNA translation. Additionally, EIF4G2 is another factor that collaborates in cap-independent translation mechanisms.

4.8. Epigenetics in hypoxia

4.8.1. Histone methylation

Hypoxia affects chromatin structure and modulates the methylation of many histones [109,190], indicating that it affects the enzymatic activity of KDMs and histone methyltransferases (HMTs).

4.8.1a. Hypoxia-mediated effects on KDMs

KDMs can either stimulate or repress gene transcription depending on the affected histone residue (Table I2). KDMs require oxygen and αKG generated in the Krebs cycle in order to catalyse their enzymatic reactions. Conversely, the expression of many KDMs increases in low oxygen conditions [109]. Different studies have uncovered some hypoxia-mediated effects on KDMs:

- *KDM3A* has a putative HRE sequence at its promoter [191]. HIF1α induces the expression of *KDM3A* and this acts as a hypoxia signal amplifier as it enhances the expression of HIF1α target genes. In addition, KDM3A directly interacts with HIF1α and decreases H3K9me2 levels on promoters of genes necessary to enhance hypoxia-driven glycolysis [191,192].
- *KDM4B* promoter also contains a HRE sequence [191] and is induced by HIF1α in CRC cells and decreases H3K9me3 levels in the promoters of HIF1α-induced genes. Indeed, *KDM4B* expression positively correlates with *CA9* expression. KDM4B is involved in cell proliferation, apoptosis and cell cycle arrest [81,193].
- *KDM5B* and *KDM4C* harbour HREs in their promoter regions and are upregulated in hypoxia. KDM4C interacts with HIF1α to enhance its binding to HREs in genes involved in cell metabolism such as *LDHA*, *PDK1* and *solute carrier family 2 member 1* (*SLC2A1*, also called GLUT1) and eventually induce their transcription [194] (Table I2).

KDM	Histone marks
KDM3A	H3K9me1, H3K9me2
KDM2B	H3K4me3, H3K36me1, H3K26me2
KDM4B	H3K9me2, H3K9me3, H3K36me2, H3K36me3
KDM5B	H3K4me2, H3K4me3
KDM6B	H3K27me, H3K27me3
KDM4C	H3K9me2, H3K9me3, H3K36me2, H3K36me3

Table I2. HIF-dependent KDMs and histone targets. Adapted from [109].

Several histone methylation marks increase in hypoxia. Hepa1-6 cells exposed to 0.2% O_2 exhibit increased H3K4me2, H3K4me3, H3K79me3, H3K27me3 and H3K9me2 [195]. Another study on mouse macrophages showed a global increase in H3K9me2, H3K9me3 and H3K36me3 [196]. Despite the global increase observed in H3K9me3 in low oxygen conditions, hypoxia has been shown to induce KDM4B-mediated H3K9me3 demethylation at the promoter of specific hypoxia-inducible genes in CRC cells [193]. Therefore, although more work is needed to fully understand the consequences of hypoxia on histone methylation, it is widely accepted that low O_2 levels affect the methylation status of several histone residues [190] (Figure I14).

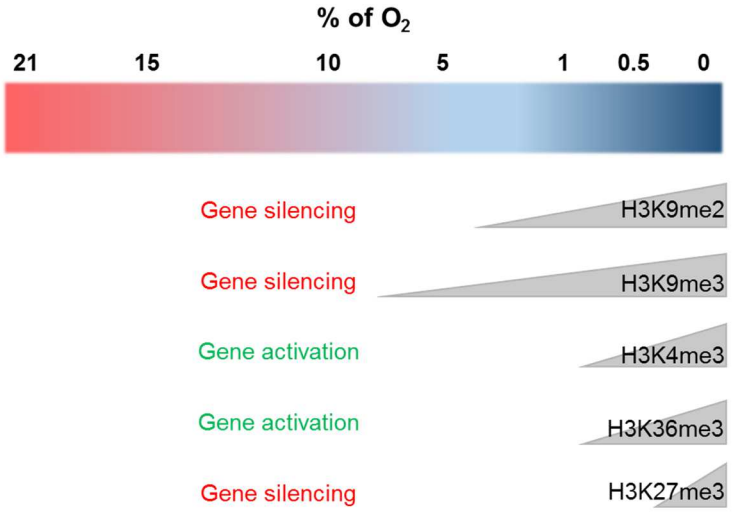


Figure I14. Increase of histone methylation marks in hypoxia. The methylation of either gene activating or silencing histone marks increases at certain O_2 concentrations. Adapted from [190].

4.8.1b. Hypoxia-mediated effects on histone methyltransferases

Hypoxia-mediated increase in histone methylation is also mediated by an increased activity of histone methyltransferases. Hypoxia increases euchromatic histone lysine methyltransferase 2 (EHMT2, also referred as G9a) protein levels, leading to a higher dimethylation of H3K9 [197,198], while impairing demethylation processes of this mark [197].

4.8.2. DNA methylation

Despite there is some controversy regarding the specific effects that hypoxia causes on DNA methylation, a long list of studies point to an alteration of DNA methylation under low oxygen conditions. Demethylation of *HIF1A* promoter, which contains HRE sequences, leads to an auto-transactivation of *HIF1A* gene, and an increased expression of HIF1α target genes in CRC cells. In addition, inhibition of DNA methyltransferases (DNMTs) by 5-aza-deoxycitidine only in hypoxic conditions elevates the expression of hypoxia-induced genes [199].

Hypermethylation also blocks binding of HIF1 complex to *CA9*, *STC2*, *EPO* and *BNIP3* promoters or enhancers and represses their expression in a variety of cancer cell lines [200–203].

Both global DNA hypo- and hypermethylation have been shown to occur in hypoxic conditions depending on the models studied. Colorectal and melanoma cancer cells show an inverse correlation between the magnitude of tumour hypoxia and methylation levels (by measuring 5mC) in cell lines and xenograft models [204]. Wu and co-workers evidenced a positive regulation and activity of the Ten-eleven Translocation (TET) enzymes TET1 and TET3 and induction of DNA hydroxymethylation in hypoxia, which leads to an increased TNFα-p38-MAPK pathway to enhance breast cancer stemness [109,205].

Conversely, another study showed that hypoxia causes global DNA hypermethylation through 5hmC loss in several murine and human cell lines. This hypermethylation is due to a loss of TETs activity [206]. Benign prostate epithelial cells and cardiac fibroblasts exposed to prolonged hypoxia also exhibit global DNA hypermethylation and increased levels of DNMTs such as DNMT1 or DNMT3B [109,207,208].

4.9. Hypoxia and cancer

Pathological hypoxia is a common microenvironmental factor in tumours that facilitates cancer cell survival and propagation. Hypoxia promotes overexpression of HIF1α and HIF2α subunits and activation of their downstream targets that are responsible for blood vessel formation, metastasis, and resistance to treatment [209] (Figure I15).

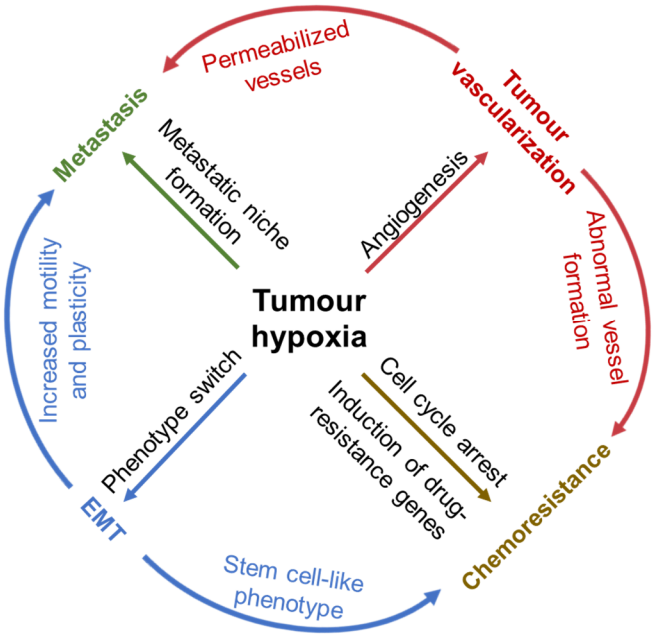


Figure I15. Role of hypoxia in tumour behaviour. Hypoxia stimulates tumour angiogenesis by sprouting of the pre-existing vessels. New blood vessels facilitate metastasis formation by cancer cells leaving the primary tumour site. Hypoxic cancer cells also undergo epithelial-to-mesenchymal transition (EMT) acquiring a plastic and mobile phenotype. Chemoresistance of patients is caused by EMT-related stemness of cancer cells, hypoxia-induced cell cycle arrest in G1 phase and upregulation of genes that confer drug-resistance. Hindered drug diffusion due to anomalous vascularity is another mechanism of chemoresistance. Adapted from [209].

4.9.1. Blood vessel formation

Hypoxia occurs in both physiological and pathological conditions. In cancer, the uncontrolled cell growth in solid tumours eventually results in a neoplasm devoid of oxygen in some areas. Cancer cells that are under a chronic exposure to low oxygen conditions enhance the adaptation and survival of a subset of hypoxia-resistant cells, leading to an increased blood vessel formation, aggressiveness, metastatic potential and chemoresistance [209,210].

In order to ameliorate the low oxygen conditions, tumours stimulate blood vessel formation. This process is crucial in embryogenesis, where blood vessels are formed *de novo* by vasculogenesis followed by angiogenesis, a process in which the pre-existing vasculature is

used to generate new blood vessels. In contrast, abnormal angiogenesis occurs in tumours due to hyperproliferating cancer cells that surpass their blood supply. Consequently, hypoxic cells induce the expression of pro-angiogenic factors like VEGFA [209,211]. Some anti-angiogenic therapies are used in the clinic, like bevacizumab, which targets VEGFA and is used to treat metastatic CRC [212,213]. However, the long-term exposure to this agent can be a doubled-edged sword, as it reduces tumour growth while it increases the invasiveness and metastatic ability of cancer cells [214].

4.9.2. Metastasis

Metastasis is another hallmark of hypoxic cancerous cells. Enhanced angiogenesis facilitates the extravasation, circulation and relocation of tumour cells to distant tissues. EMT is active during embryogenesis but also in tumour hypoxia, where it enhances the invasive and migratory behaviour of cells from many types of solid tumours [209]. It has been observed that HIF1α is mainly present at the front edge of the invading tumour in gastric cancer biopsies [215]. In line with this, CA9 is crucial in regulating the intra- and extracellular pH, thereby affecting the survival and invasion of cancer cells [216].

4.9.3. Resistance to treatment

Hypoxia positively correlates with poor prognosis in cancer patients. Elevated HIF1α and HIF2α turns head-and-neck tumours more resistant to chemotherapy [217]. Similar results are observed in oropharyngeal patients, in which high HIF1α levels relate to a lower chance to achieve complete remission after irradiation [218]. Inactivation of HIF1α in embryonic fibroblasts increases their susceptibility to carboplatin and etopisode chemotherapeutic agents [219]. Hypoxia confers treatment resistance through a variety of processes such as:

- Induction of cell cycle arrest [220,221].
- Inhibition of apoptosis and senescence [220,222].
- Control of autophagy, TP53 and mitochondrial activity [220,222]. HIF1α has been shown to antagonize TP53-mediated apoptosis [223,224]. For instance, HIF1α suppresses TP53 activation in response to 5-fluorouracil (5-FU) treatment in gastric cancer cells [224]. HIF1α can reduce mitochondrial activity through diverse mechanisms such as alteration of cytochrome C oxidase expression [225], suppression of the Krebs cycle [226,227] or suppression of ROS production [224].
- Lowering drug delivery and cellular uptake through high acidity and drug extrusion of the cell. The multidrug resistance protein 1 (MDR1, *ABCB1* gene) is a membrane-resident P-glycoprotein that belongs to the family of ATP-binding cassette (ABC) transporters and a HIF1α-target [228,229]. MDR1 acts as an efflux pump that decreases

the intracellular concentration of chemotherapeutic drugs. MDR1 contribution to hypoxia-induced drug resistance has been observed in glioma, gastric cancer, breast and colon cancer cells [222].

- Decreasing cytotoxicity of several chemotherapeutics as oxygen is required for the cytotoxic effect of some chemotherapeutical agents [230].

5. DPPA3

5.1. DPPA3 dynamics in embryonic development

DPPA3 (Developmental Pluripotency Associated 3), also called stellar (stella in mice) or PGC7, is a maternal factor essential for early development and predominantly expressed in ESCs, located in the inner cell mass of the blastocyst, and in primordial germ cells (PGCs), the precursors of male and female gametes [231].

DPPA3 is present in two separate developmental processes: in the pre-implantation zygote and during emrbyo's PGCs development. During the pre-implantation stage, the zygote undergoes an epigenetic reprogramming that involves an active DNA demethylation on their genomes. DPPA3 is pivotal in this process as it protects retrotransposable elements and asymmetrically methylated imprinted genes from TET-mediated active demethylation, which involves DNA conversion of 5mC to 5hmC [232–234]. Genomic imprinting is a crucial process during embryonic development and leads to the parental-dependent expression of a small subset of genes via differential methylation and silencing of one parental allele [235]. TETs mediate a part of the demethylation process observed in the maternal and paternal genomes in the zygote [236]. TET1 is specifically expressed in ESCs and PGCs. TET2 is also expressed in ESCs, and the catalytic activity of both TET1 and TET2 is required for normal differentiation during ESC lineage specification [237]. TET3 is the only amongst the three enzymes that is present immediately after fertilization in the zygote and mediates the global erasure of 5mC [232,238–240]. Contrarily to the high active demethylation process that takes place in the paternal genome of the zygote, the maternal genome is protected from the enzymatic activity of TET3 at certain imprinted regions such as *Peg1*, *Peg3*, *Peg10*. Nevertheless, it has been reported that, although not in the same extent than that of the maternal genome, some paternally-imprinted loci (e.g. *H19*, *Rasgrf1*) are also protected by DPPA3 from active TET-mediated demethylation [231]. DPPA3 recognizes and binds H3K9me2 repressive marks located nearby imprinted genes and protects these loci from active demethylation processes (Figure I16). A crucial evidence necessary to understand this mechanistic is the lack of histones in the paternal genome. While the maternal genome contains histones, the paternal DNA is tightly packaged with protamine. Moreover, the maternal genome contains considerable levels of

H3K9me2, reason for which Nakamura and colleagues examined whether there was any relation between the presence of DPPA3 and H3K9me2 marks. G9a dimethylates H3K9 and its depletion in murine ES cells has shown that H3K9me2 is crucial for the tight association of DPPA3 with nucleosomes [231]. Besides active demethylation, passive DNA demethylation involving DNMT1 nuclear exclusion has also been reported during pre-implantation development [241,242].

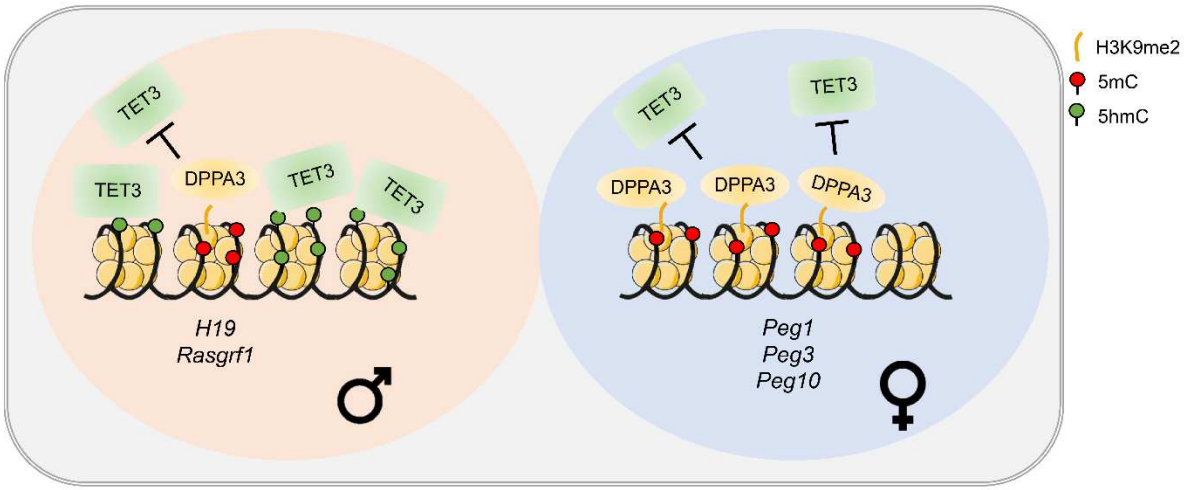


Figure I16. Mechanism of action of DPPA3 in zygote genomes soon after fertilization. DPPA3 preferentially binds the maternal (right) genome given the high content of H3K9me2 at specific loci, which must be protected from TET3-mediated demethylation (e.g. *Peg1*, *Peg3*, *Peg10*). Meanwhile, the paternal (left) genome is widely demethylated by TET3, with some exceptions (*H19*, *Rasgrf1*). Adapted from [243].

DPPA3 is also expressed in PGCs during the process of germ cell specification between embryonic day 7.25 (E7.25) and E15.5 in both male and female mouse gonads. During the specification process, PGCs undergo a wave of epigenetic remodelling in which most of their epigenetic marks are erased. While male mice stop expressing *Dppa3* at E15.5, *Dppa3* expression is resumed in the immature oocytes of new-born female mice and it continues until oocytes are mature [231]. *Dppa3*-deficient females are infertile but *Dppa3*-deficient males do not show any abnormality in reproductive activity [231,244]. Epigenetic reprogramming in PGCs at E7.25 involves a genome-wide loss of 5mC that is more global than that occurring in zygotes, since genome imprints are erased and the demethylation of transposable elements is more extensive [233,245]. Whereas median methylation levels at CpGs are 85% in sperm, 75% in ESCs and 73.2% in E13.5 embryos, those in E13.5 male and female PGCs are only 16.3% and 7.8% respectively [246]. BMP triggers the start of the specification process (around E6.0) by signalling PGCs from the extra-embryonic ectoderm. PGCs then shut down the somatic transcriptional program, re-express pluripotency factors like SOX2 and get prepared for the upcoming epigenetic reprogramming. At this point, PGCs start to migrate to the hindgut until the genital

ridges where they initiate sexually dimorphic development [247,248]. Migrating PGCs arrest at the G2 phase of cell cycle and become transcriptionally quiescent [233]. Upon PGC specification, *Dnmt3a*, *Dnmt3b* and *Uhrf1*, involved in DNA methylation, are transcriptionally repressed [249]. Meanwhile, PGCs continue expressing *Dnmt1*. GLP and G9a, which mediate H3K9 dimethylation, are also repressed throughout this process [250]. Therefore, from the start of the specification until E12.5, the PGCs have almost negligible DNA methyltransferase and H3K9 methyltransferase activity. In addition to passive DNA demethylation, active demethylation processes mediated by base excision repair-related proteins together with TET1 and TET2 activity have also been linked to 5mC erasure in PGCs [233,246,251]. Despite the specific role attributed to DPPA3 during the pre-implantation stage in embryonic development, its function in PGCs has not been fully elucidated. Dppa3 is necessary for the proper demethylation of the retrotransposon genes during PGCs reprogramming [234]. Transposable elements occupy 46% of the human genome and contribute to various biological processes including chromosome function, genome integrity and epigenetic regulation of specific genes [252,253]. Therefore, it has been postulated that higher methylation levels in these sequences may affect physiology of cells [234] (Figure I17).

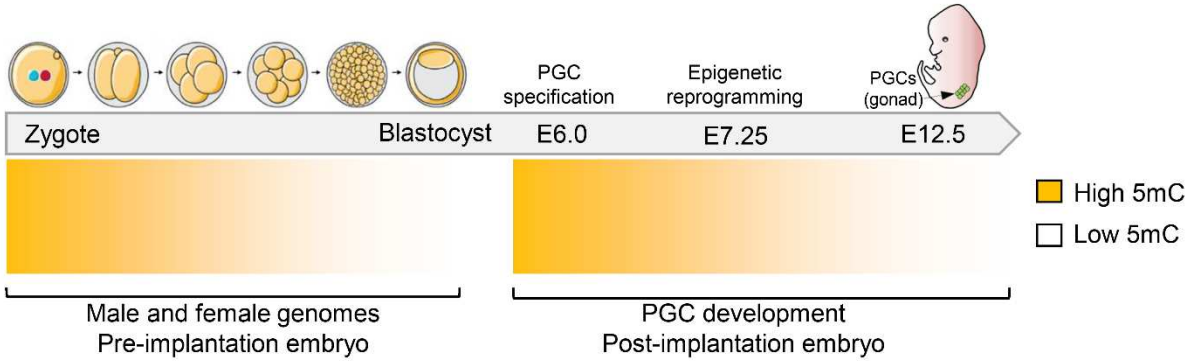


Figure I17. DNA methylation dynamics in mouse pre-implantation and germ cell development. Top left: schematic of pre-implantation stages from the zygote to the blastocyst stage. Bottom left: genomic 5mC levels gradually decrease during the development of the pre-implantation embryo. While 5mC levels in the paternal genome decrease faster, 5mC in the maternal genome show slower dynamics. Top right: PGCs development in the post-implantation embryo including their specification and epigenetic reprogramming. Bottom right: 5mC dynamics in the PGCs compartment. Specified PGCs re-acquire the expression of pluripotency factors and prepare for the epigenetic reprogramming that takes place after E7.75 in which PGCs undergo a profound genome-wide loss of 5mC. Adapted from [233].

5.2. DPPA3 structure, partners and effects on DNA methylation

DPPA3 maps on chromosome 12p13, a region that harbours several pluripotency-related genes such as *apolipoprotein B mRNA editing enzyme catalytic subunit 1* (*APOBEC1)*, *growth differentiation factor 3* (*GDF3)* and *nanog homeobox* (*NANOG)* [254] (Figure I18). Human DPPA3 consists of a 159-amino acid protein whose sequence is moderately conserved and contains a nuclear localization signal and a nuclear export signal [231].

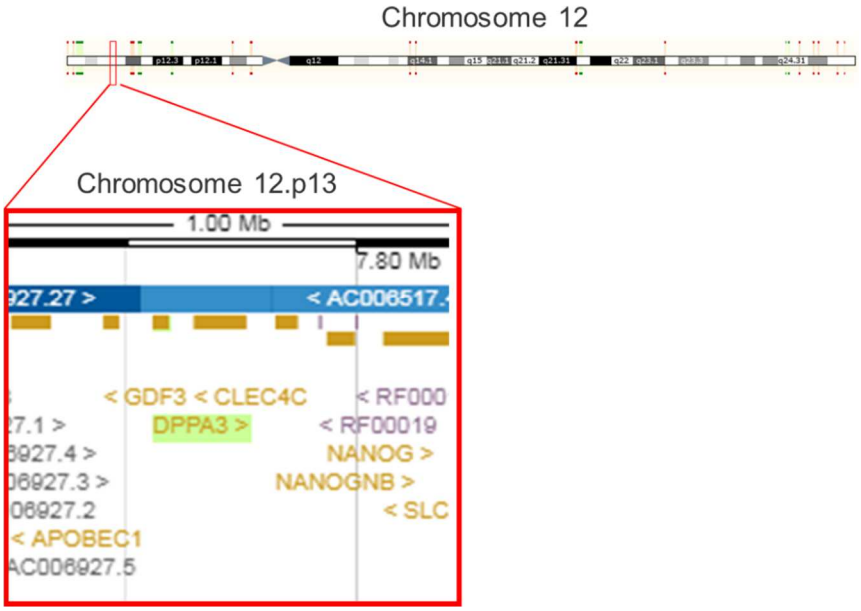


Figure I18. *DPPA3* location in the human genome. *DPPA3* flanks *GDF3*, *APOBEC1* and *NANOG* pluripotency genes [255].

The N-terminal region of DPPA3 is required for its binding to H3K9me2 while the C-terminal half mediates its ability to exclude TET3 from the maternal pronucleus [256].

It has been shown that DPPA3 is an intrinsically unfolded protein [257]. Unfolded proteins show increased structural flexibility thereby being more easily to adapt to a variety of binding partners to perform their function. This flexibility can also help to play a role in proteins associated with protein–protein binding, protein–DNA binding, protein–RNA binding, cell-cycle control, cell signalling and longevity of proteins [258,259].

The most well-known partners of DPPA3 are importin 5 (IPO5) and ubiquitin like with PHD and ring finger domains 1 (UHRF1). IPO5 is a nuclear transport shuttle that mediates DPPA3 nuclear localization [231]. On the other hand, DPPA3 competes with DNMT1 for UHRF1 binding. DNMT1 catalyses the transfer of methyl groups to specific CpG structures in DNA and is responsible for the maintenance of DNA methylation during its replication ensuring the fidelity of inherited epigenetic patterns. UHRF1 has been recently shown to directly interact with

DNMT1 [260]. UHRF1 recognises unmodified arginine 2 in histone H3 (H3R2), H3K9me2 and H3K9me3. This allows targeting of UHRF1 and DNMT1 to genomic loci mainly located at replicating heterochromatic regions [261]. DPPA3 binding to UHRF1 causes DNMT1 disruption from chromatin triggering global passive DNA demethylation. In addition, DPPA3 binding to UHRF1 disrupts this protein from chromatin because DPPA3 binds a region of UHRF1 necessary for the recognition of the histone H3 tail and competes for the interaction between UHRF1 and the histone H3 tail [262] (Figure I19).

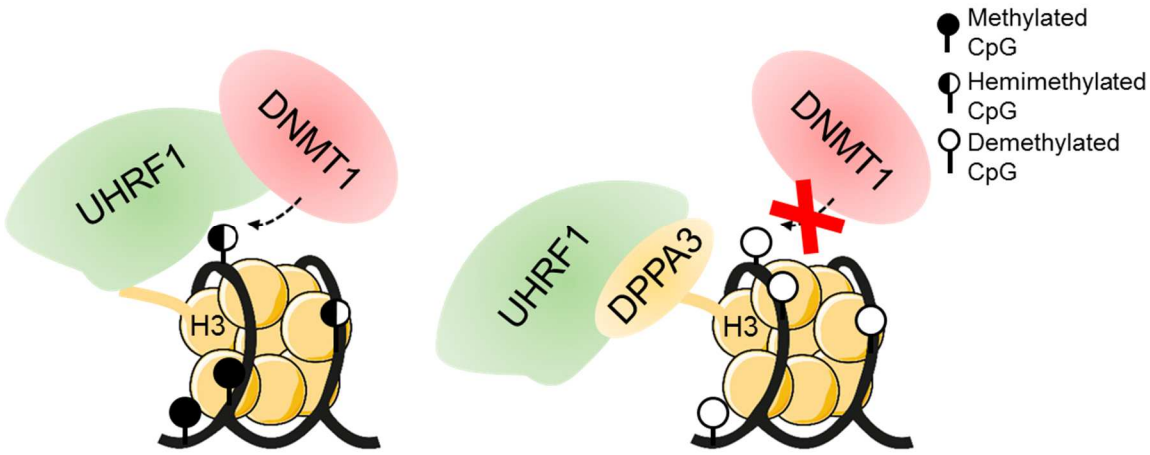


Figure I19. Effects of DPPA3 on DNA methylation maintenance. Left: UHRF1 recruits DNMT1 by recognising hemimethylated CpG sequences and specific modifications in the histone H3 (H3R2, H3K9me2 or H3K9me3). Recruited DNMT1 methylates the newly synthetized DNA strand thereby assuring the proper 5mC inheritance pattern. Right: DPPA3 recognizes H3K9me2 marks and binds UHRF1, which cannot longer bind modified residues on the H3 and is unable to recruit DNMT1. Absence of recruited DNMT1 leads to a global genome hypomethylation along cell divisions.

A proteomic analysis identified additional DPPA3-interactor proteins with a variety of molecular functions like RNA processing (e.g. YBX1, HNRNPR and U2AF2), chromatin organization (RBBP4, DMAP1, and HP1BP3), protein translation (KPNB1, KPNA1, YWHAG, and NUPL1) and cell cycle (CDK2, UHRF1, HSPA2, MCM6, AKAP8L and RBBP4) [257]. Both HSPA2 and AKAP8L are required for the G2 to M phase transition [263,264].

5.3. DPPA3 and cancer

DPPA3 is expressed at very low levels in most human adult tissues. However, its expression is enriched in ovary and testis, which contain pluripotent cells [254]. In cancer, DPPA3 has been mainly linked to testicular germ cell tumours (TGCTs), where *DPPA3*, together with *NANOG* and *GDF3*, is overexpressed compared to normal testis [254,265]. TGCTs are a heterogeneous group of neoplasms that originate from germ cells and can be divided in three types: teratomas and yolk sac carcinomas (type I), seminomas and non-seminomas (type II) and spermatocytic seminomas (type III) [266]. The seminoma class of TGCTs has a morphological resemblance to PGCs and it is characterized by chromosome duplications at p12. In addition, high expression of *DPPA3* has also been localized in some breast carcinoma samples [267].

Another work suggested a link between global DNA hypomethylation induced by DPPA3 overexpression to cellular transformation of a mouse fibroblast cell line. They also showed that higher levels of DPPA3 lead to an increased metastatic ability in a melanoma cell line and hypothesized that it could be due to the activation of metastasis-related genes [268].

BACKGROUND

One of the main interests of the Stem Cell and Cancer Group is to understand tumour dormancy as a key phenomenon affecting cancer relapse and drug resistance. The team recently published a large dataset describing the biology of slow cycling cancer cells (SCCC) [269]. The study identified in different cancer models a small proportion of non-proliferative dormant tumour cells which slow-cycling nature was evidenced by their ability to retain a specific label (the histone H2B fused with the enhanced green fluorescent protein, GFP, H2BeGFP) incorporated in their genome in pulse-chase experiments upon doxycycline (DOX) treatment. This approach was already used in other studies [270–273], although most of them were developed in transgenic mouse models and only a few in human tissues, tumours in particular. Based on this methodology, SCCC could be identified in colorectal, melanoma and glioblastoma patient tumour samples and cell lines, grown either *in vivo* as tumour xenografts or *in vitro* as three-dimensional (3D) structures (Figure B1).

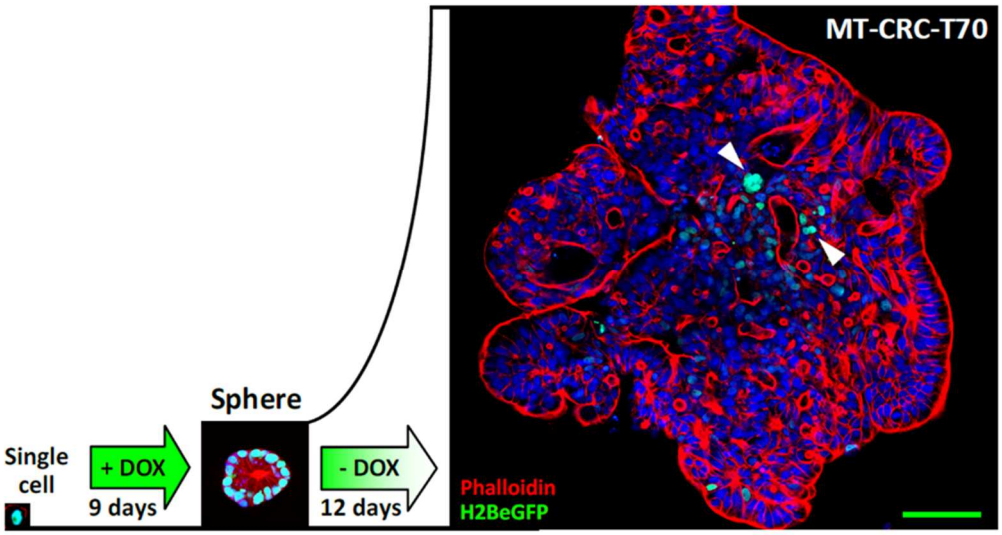


Figure B1. Representative immunofluorescence picture of H2BeGFP-infected SW1222 megacolonies. SW1222 CRC cells grew embedded in Matrigel and were generated from single-cell suspensions. DOX pulse-chase experiments were performed to evaluate label-retaining SCCC (green). Phalloidin (red) was used to visualize cellular organization. Arrowheads point SCCCs. Hoechst was used as nuclear counterstain. Scale bars: 100 μm [269].

The identification and isolation of SCCC permitted the study of their distinctive gene expression profile and evidenced a cell-autonomous transcriptional program, which was common across tumour types as biologically diverse as colorectal carcinoma, melanoma and glioblastoma. General processes such as proliferation and metabolism were negatively enriched in SCCC, whereas those related to drug detoxification, stemness, hypoxia or crosstalk with the immune system were positively enriched. SCCC accumulated at the G2/M phases of the cell cycle, which could be explained by the decreased expression of a set of genes involved in the replication machinery, including cyclins and cyclin-dependent kinases,

genes involved in DNA replication, centromere/kinetochore assembly machinery or chromosome segregation. In addition, SCCC were negative for the senescence/DNA damage marker phospho-histone H2AX, excluding the possibility that G2/M arrest was caused by DNA damage. In the case of CRC SCCC, they were also negative for the differentiation markers MUC2, cytokeratin-20, chromogranin-A and lysozyme, evidencing their undifferentiated state. Of note, SCCC exhibited tumour-initiating capacity and could generate rapid-cycling cancer cells (RCCC) and *vice versa*, implicating that *slow cycliness* was a transient and reversible state.

A common Pan-Cancer-SCCC signature was generated from the gene expression analysis performed on CRC, melanoma and glioblastoma SCCC. The signature was common across all models irrespectively of their intrinsic differences according to tumour type, individual patient traits, mutation repertoire or experimental model analysed (*in vivo* or *in vitro*). The Pan-Cancer-SCCC signature score was applied to the gene expression available from cancer cell lines downloaded from Cancer Therapeutics Research Portal v2 website (http://portals.broadinstitute.org/ctrp/). Of note, those cell lines enriched in the Pan-Cancer-SCCC signature were more resistant to standard-of-care chemotherapeutics such as paclitaxel, docetaxel, topotecan or SN-38. Finally, another signature was generated from CRC samples (CRC-SCCC signature) and applied on a cohort of CRC patients treated with 5-FU–based adjuvant chemotherapy. Those classified as positive for the CRC-SCCC signature had significantly shorter disease-free survival after adjustment for other known prognostic risk factors. CRC-SW1222 xenografts and *in vitro* experiments performed in colorectal, melanoma and glioblastoma cancer cells demonstrated that SCCC were more resistant to chemotherapy-induced apoptosis and increased their proportion upon treatment. Interestingly, SCCC isolated from different tumour types exhibited a higher expression of drug resistance genes compared to RCCC counterparts, including drug efflux pumps.

Altogether, the data provided in this work complements a growing body of evidence indicating the importance of dormancy in chemoresistance and cancer recurrence (Figure B2). Therefore, the unravelling of the biological core network governing dormancy is of foremost importance for identifying new drug targets for SCCC eradication, fighting chemoresistance and patients relapse. In response, Puig and colleagues performed a systematic analysis of the gene expression profiles distinctive of SCCC.

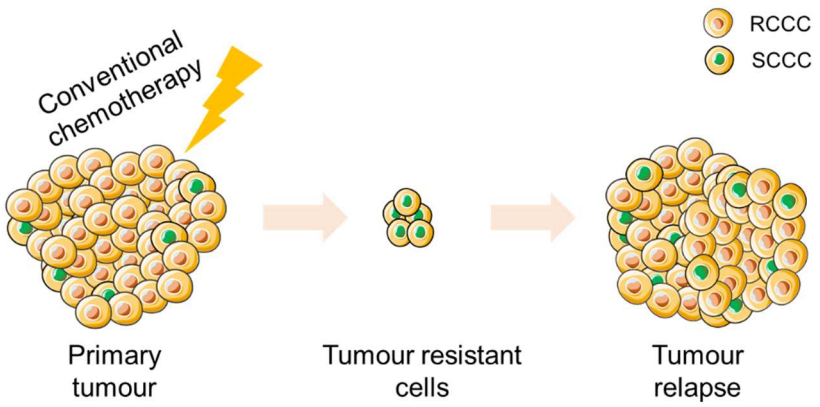


Figure B2. SCCC are responsible for tumour recurrence due to their enhanced drug resistance. Cancer cells within the same tumour show distinct tumour initiating potential and resistance to conventional chemotherapy. After first-line treatment, chemoresistant SCCC remain as a minimal residual disease and eventually regrow leading to a detectable clinical tumour relapse composed of RCCC and SCCC.

A wide range of pluripotency and epigenetic factors were upregulated in SCCC, including Tet Methylcytosine Dioxygenase 2 (TET2), which catalyses the conversion of the 5mC DNA repressive mark to 5hmC. Hydroxylated 5mC can then undergo successive oxidation steps until its complete demethylation, thereby enabling gene expression [274]. Gene expression studies uncovered that TET2, which has been described to be a tumour suppressor in haematological disorders [275], was essential for the survival of those cancer cells transiting to and/or consolidating a slow-cycling state in different types of cancer. Elevated 5hmC levels generated by TET2 predicted tumour relapse and worse patient survival across different types of cancer including CRC.

Apart from the identification of TET2 as a crucial factor mediating SCCC survival, gene expression analysis uncovered a set of other genes upregulated in SCCC ([269], unpublished data) that had been described in pluripotency and DNA demethylation processes in ESCs and PGCs during embryogenesis [269,276–278] (Figure B3). Interestingly, DPPA3 was one of those genes.

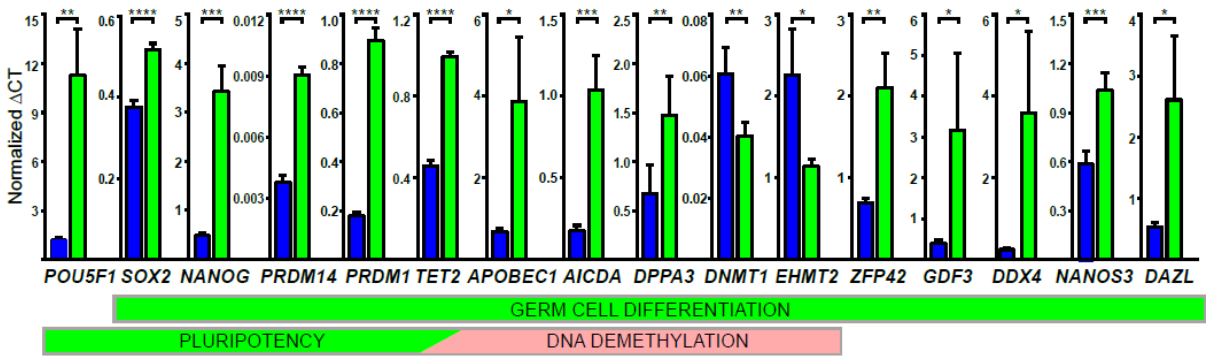


Figure B3. Expression of genes in SCCC and RCCC. RT-qPCR analysis showing the expression in SCCC (green bars) and RCCC (blue bars) isolated from a patient-derived xenograft (PDX) xenograft tumour, of pluripotency and DNA methylation-related genes involved in germ cell differentiation. Error bars show geometric mean +/- standard deviation of triplicates [269], unpublished data.

Just very few studies linked DPPA3 to cancer at that time. Higher expression of *DPPA3* and/or gene amplification was described in undifferentiated tumours such as TGCTs [254,265]. A possible relationship between the overexpression of DPPA3 and the metastatic ability of a melanoma cell line had been suggested [268]. In summary, DPPA3 was not formally linked to cancer yet. Interestingly, other epigenetic factors had been described as central to control phenotypes distinctive of SCCC such as drug-resistance, tumour-initiating capacity or multipotency.

For all these reasons, we decided to focus our investigations on DPPA3 as an interesting candidate ruling the peculiar biology of dormant tumour cells. It was at this particular moment that I joined the laboratory and the development of this study was proposed to me as a thesis project. Of course, I accepted the challenge!

OBJECTIVES

Tumour relapse is a frequent complication in cancer and a major cause of death that can occur after prolonged periods of latency. During this disease-free stage, a reservoir of cancer cells remain in a non-proliferative dormant state that confers them intrinsic chemoresistance. The virulence of these cells relies on their plastic ability to reverse their non-proliferative state to generate a growing mass that is eventually diagnosed as a recurrence. Although many efforts are currently being invested to elucidate the molecular mechanisms underlying tumour dormancy, just few investigations have identified drug targets for their effective eradication. DPPA3 is an epigenetic factor that we found to be upregulated in CRC-SCCC cells, suggesting that it might be important for their biology.

The goal of this project was to study the implication of DPPA3 in tumour dormancy. The specific objectives were as follows:

- Investigate the role of DPPA3 in CRC.
- Study the implication of DPPA3 in the molecular mechanisms driving tumour dormancy.
- Assess the effects of DPPA3 on tumour cell chemoresistance and its impact on disease progression of CRC patients.

MATERIALS AND METHODS

1. Cell lines

The following cell lines were used in this study:

- SW1222 and HT29: human CRC cell lines. Used for *in vitro* and *in vivo* experiments.
- DLD1, HCA7, HCT116, LIM245, RKO, SW480, SW620 and TC310X6 are human CRC cell lines that were used together with SW1222 and HT29 cell lines for comparison of endogenous *DPPA3* mRNA levels.
- HEK293T: human embryonic kidney cell line. Used for virus production and transient DPPA3 overexpression experiments.

For generation of the doxycycline (DOX)-inducible DPPA3 overexpressing (DPPA3-OE) cells, SW1222 and HT29 cells were transduced with lentiviruses expressing the human DPPA3 protein fused with the sequences encoding CMYC and FLAG tags (pSIN-TRE-DPPA3-CMYC-FLAG-hPGK-rtTA2-P2A-mCherry plasmid described in *Plasmids* section). SW1222 and HT29 cells infected with the empty lentiviral vector (pSIN-TRE-hPGK-rtTA2-P2A-mCherry) were used as control cells. Once recovered from transduction, cells were pelleted and sorted using fluorescence-activated cell sorting (FACS) with a FACS Aria II cell sorter (BD BioSciences) obtaining a pool of cells expressing the highest levels of mCherry. Individual DPPA3-OE clones were generated from expanded single cells. Clones generated from DPPA3-OE SW1222 or HT29 cells were used in all overexpression experiments unless otherwise specified. To induce DPPA3 expression in *in vitro* experiments, DPPA3-OE and control SW1222 and HT29 cells were treated with 5 µg/ml of DOX during the last 5 days of the experiment.

DPPA3 knockdown SW1222 cell derivatives were generated by expressing different non-targeting shRNA (shC) and DPPA3 TRC-shRNA (shDPPA3). After transduction, cells were selected with Puromycin (Life Technologies) (1 µg/ml). One shC cell line and three shDPPA3 cell lines (shDPPA3-1 to 3) were generated using different shRNA constructs (specified in *Plasmids* section).

DOX-inducible expression of H2BeGFP has already been described [269]. Briefly, SW1222 wild-type (WT), shC and shDPPA3 cells were transduced with lentiviruses expressing H2BeGFP protein (pSIN-TRE-H2BeGFP-hPGK-rtTA2 vector, described in *Plasmids* section).

DPPA3 knockout (DPPA3 KO) and control (Scramble) cell lines were generated using a CRISPR-Cas9 system. SW1222 cells were transfected with pSpCas9-sgRNADPPA3guide2-2A-GFP or pSpCas9-sgRNAscramble-2A-GFP constructs (described in *Plasmids* section) using linear polyethylenimine (PEI 25000, Polysciences, Inc.). 48 hours post-transfection, cells were pelleted and sorted using FACS with a FACS Aria II cell sorter (BD BioSciences). Single

cells were expanded to obtain individual clones. Genomic DNA was isolated from edited clones and non-edited SW1222 control cells. A 450bp fragment of exon 1 of *DPPA3* was amplified by PCR using genomic DPPA3 CRISPR Screening-primers (Table MM1) and PCR products were analysed by Sanger sequencing. To validate biallelic mutation in each clone, genomic DNA was isolated from edited clones and non-edited SW1222 cells. Exon 1 of *DPPA3* was PCR amplified using the DPPA3 CRISPR Screening-primers (Table MM1) used for PCR amplification of exon 1 commented before. The PCR products were A-tailed and cloned into pCR4-TOPO vector (Invitrogen) following the manufacturer's instructions. Individually cloned PCR products were then analysed by Sanger sequencing using the DPPA3 CRISPR Screening Forward primer (Table MM1).

2. Cell culture and treatments

2.1. Three-dimensional (3D) cell culture: SW1222 megacolonies

DPPA3-OE, shDPPA3, DPPA3 KO and respective control SW1222 cells were maintained in Dulbecco's modified eagle medium (DMEM; Biowest) supplemented with 10% Fetal bovine serum (FBS) (complete DMEM) and 1% P/S (Life Technologies). Megacolonies embedded in Matrigel from SW1222 cells were generated from single-cells resuspended in complete DMEM and mixed 1:1 with Corning Matrigel Basement Membrane Matrix (BD Bioscience). Drops of 25 µl of the mixture were seeded in a 48-MW plate and incubated for 30 minutes at 37ºC without medium to allow the matrix solidification. Then, complete 750 µl of DMEM was added and changed twice weekly. Cells were cultured until they reached a well-differentiated heterogeneous structure (2-3 weeks of culture), where they were used for posterior analyses (RNA, immunofluorescence and apoptosis analysis). For expression of the human Histone2B fused to the eGFP (H2BeGFP), megacolonies from SW1222 cells infected with H2BeGFP lentiviral vector (described in *Plasmids* section) were treated with 5 µg/ml of DOX (pulse) for 9 days and released of treatment for 12 days DOX (chase). For DPPA3 overexpression, megacolonies from SW1222 cells infected with the pSIN-TRE-DPPA3-CMYC-FLAG-hPGK-rtTA2-P2A-mCherry construct and control cells were treated with 5 µg/ml of DOX during the last 5 days of the cell culture.

2.2. Two-dimensional (2D) cell culture and hypoxic treatment

Cell lines were cultured in two dimensions and maintained in Dulbecco's modified eagle medium (DMEM, Biowest) supplemented with 10% Fetal bovine serum (FBS) (complete DMEM) and 1% penicillin/streptomycin (P/S) (Life Technologies). Normoxic cultured cells were grown at atmospheric O_2 concentrations (21%) with 5% CO_2. Hypoxic treatment of cells followed normoxic culture and consisted in the incubation of cells during 24 hours in an

atmosphere of 0.5% O_2, 5% CO_2, and 94.5% N_2 in a custom-made seal chamber which was filled with the gaseous mixture and then properly sealed and tempered at 37°C. After hypoxic treatment, cells were reoxygenated for approximately 1 minute and used for posterior analysis (protein, DNA, and RNA analysis, BrdU incorporation assays and assessment of *CA9* promoter activity).

3. Plasmids

3.1. DPPA3 overexpression

3.1.1. DOX-inducible overexpression

For generation of the lentiviral vector pSIN-TRE-DPPA3-CMYC-FLAG-hPGK-rtTA2-P2A-mCherry, the coding sequence of human DPPA3 fused to CMYC and FLAG sequences from the pCMV6-DPPA3 plasmid (Origene plasmid ID: RC214676) was amplified by PCR with specific primers (Table MM2) containing NheI and PacI restriction sites. The resulting PCR product was then subcloned into the NheI-PacI cut pSIN-TRE-hPGK-rtTA2-P2A-mCherry plasmid to obtain the final pSIN-TRE-DPPA3-CMYC-FLAG-hPGK-rtTA2-P2A-mCherry construct.

3.1.2. Transient overexpression

For transient overexpression experiments in HEK293T cells, pCMV6-Entry and pCMV6-DPPA3 (NM_199286) plasmids from Origene (ID: PS100001 and RC214676, respectively) were used.

3.1.3. *DPPA3* knockdown

shRNA knockdown control (shC) or against *DPPA3* transcript (shDPPA3-1 to 3) (NM_199286.2) were performed using the Mission shRNA Lentiviral vector system (pLKO1-puro TRC1.5 and TRC2; Sigma Aldrich). The specific shRNA constructs are included in Table MM3.

3.1.4. *DPPA3* knockout

A *DPPA3*-specific sgRNA oligo was cloned into the pSpCas9(BB)-2A-GFP (px458) expression vector (Addgene, plasmid ID: 48138), which bicistronically expresses sgRNA and Cas9 nuclease, to generate the pSpCas9-sgRNADPPA3guide2-2A-GFP construct. The *DPPA3*-specific sgRNA sequence, which targets the first exon of *DPPA3*, was determined by the CRISPR Design Tool (http://crispr.mit.edu) and validated by T7E1 endonuclease assay.

sgRNA scramble sequence was obtained from Origene. Oligos sequences used are annotated in Table MM1.

3.2. H2BeGFP expression

The lentiviral vector pSIN-TRE-H2BeGFP-rtTA2, which encodes H2BeGFP (Addgene Plasmid ID: 11680), previously generated in our laboratory [269] was used for H2BeGFP expression in SW1222 cells.

3.3. Luciferase reporter constructs

3.3.1. Wild-type and HRE-mutated CA9promoter-pGL3 constructs

To check *CA9* promoter activity, a 232bp fragment (-173/+39bp *CA9* promoter) of the pEZX-PF02.1 plasmid (Genecopoeia) containing the human *CA9* promoter region (-1476/+42bp) was used as template and amplified by PCR with specific primers (Table MM4) containing XhoI and HindIII restriction sites. The resulting PCR product was then subcloned upstream of a *Firefly* luciferase (Luc) open reading frame into the XhoI-HindIII cut promoter-less Luc reporter pGL3 basic vector (Promega, plasmid ID: E1751) to obtain the final CA9promoter-pGL3 construct. Site-directed mutagenesis at core HRE sequence of human *CA9* promoter has been used in other works [279–281] and was performed in the CA9promoter-pGL3 reporter construct using the QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent Technologies) following the manufacturer's instructions. Mutated oligo used for site-directed mutagenesis was designed using QuikChange Primer Design (https://www.agilent.com/store/primerDesignProgram) (Table MM4). For simplicity, the non-mutated CA9promoter-pGL3 construct is referred as wild-type (WT)-CA9promoter-pGL3 and the HRE-mutated (mut) construct as HREmut-CA9promoter-pGL3. The primers used for *CA9* promoter (-173/+39bp) cloning and site-directed mutagenesis are described in Table MM4. The identity of all constructs was verified by Sanger sequencing.

3.3.2. TOP/FOP*Flash* vectors

TOP/FOP*Flash* system has been widely used to evaluate β-catenin-dependent signalling events that drive the expression of T-cell factor (TCF) [282,283]. The TOP*Flash* vector (Addgene, plasmid ID: 24307) used in this work is a TCF reporter plasmid containing seven TCF-binding sites upstream of a *Firefly* Luc reporter gene. The FOP*Flash* plasmid (Addgene, plasmid ID: 12457) was used as a negative control for TOP*Flash* activity and contains eight mutated TCF-binding sites upstream of a *Firefly* Luc open reading frame.

3.3.3. Renilla Luciferase control reporter vector

The pRLTK plasmid (Promega, plasmid ID: E2241) containing the *Renilla reniformis* luciferase (RLuc) gene under the control of the *thymidine kinase* promoter was used for transfection reference in all luciferase reporter assays.

4. Luciferase reporter assays

4.1. *CA9* promoter activity

To monitor the transcriptional activity of HIF1α upon DPPA3 overexpression, *CA9* promoter activity was analysed by co-transfecting HEK293T cells with either pCMV6-Entry or pCMV6-DPPA3 plasmids together with the WT- or HREmut-CA9promoter-pGL3 constructs and the pRLTK plasmid for transfection reference. 48h after transfection, cells underwent hypoxia treatment (0.5% O_2) for 24h. Then, Luc and RLuc activities were separately measured with the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions. Luc activity was normalized to RLuc. Experiments were performed in triplicates, and the figures represent the average of three experiments ± SD.

4.2. TOP/FOP*Flash* assay

For analysis of the effects of DPPA3 on β-catenin/TCF transcriptional activity, HEK293T cells were co-transfected with either pCMV6-Entry or pCMV6-DPPA3 plasmids together with the TOP*Flash* or FOP*Flash* constructs and the pRLTK plasmid for transfection reference. 48h after transfection, Luc and RLuc activities were separately measured with the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions. Luc activity was normalized to RLuc. Experiments were performed in triplicates, and the figures represent the average of three experiments ± SD.

5. Lentiviral production and infection

To infect cell lines lentiviruses were produced in HEK293T cells using standard procedures and psPAX2 and pMD2.G (Addgene, plasmids ID: 12260 and 12259, respectively) packaging vectors. 48h after transfection, the supernatant was collected and filtered. This supernatant was then used to infect cells directly.

6. Protein analysis

6.1. Total extracts

For both *in vitro* and *in vivo* experiments, total cell extracts from shDPPA3 or DPPA3-OE and respective control SW1222 or HT29 cells were obtained in SDS lysis buffer (1% SDS, 10 mM EDTA pH 8 and 1 mM dithiothreitol, DTT) containing protease inhibitors (Roche). In the case of tumour xenografts from *in vivo* experiments, samples were homogenized in SDS lysis buffer using a pestle. Following cell lysis and protein extraction, samples were sonicated. Protein was quantified using the Pierce BCA Protein Assay Kit (Thermo Scientific).

6.2. Subcellular fractionation

shDPPA3, DPPA3-OE and respective control SW1222 or HT29 cells grown *in vitro* were lysed in buffer A (10 mM HEPES pH 7.8, 10 mM KCl, 1.5 mM $MgCl_2$, and 0.5 mM DTT) for 10 minutes at 4ºC. Then, 1/3 volume of 10% Triton X-100 was added. Samples were centrifuged at 15000 g at 4ºC for 1 minute to separate cytoplasmic and nuclear fractions. For cytoplasmic protein extraction, supernatant was incubated with 1/9 volume of buffer B (0.3M HEPES pH 7.8, 1.4M KCl, and 30mM $MgCl_2$) for 30 minutes at 4ºC. Samples were centrifuged at 15000 g at 4ºC for 15 minutes and supernatant (cytoplasmic fraction) was kept. For nuclear protein extraction, pelleted samples were washed twice in buffer A and resuspended in 1/5 volume of buffer C (20 mM HEPES pH 7.8, 25% glycerol, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA and 0.5 mM DTT) for 30 minutes at 4ºC. Then, samples were centrifuged at 15000 g for 15 minutes at 4ºC. Supernatant (nucleoplasm) was collected and pellet (chromatin) was washed twice in buffer C and resuspended in SDS lysis buffer (1% SDS, 10 mM EDTA pH 8). Chromatin was sonicated to homogenize it. Protein was quantified using the Pierce BCA Protein Assay Kit (Thermo Scientific).

6.3. Western Blotting

Western blots were performed according to standard procedures. Briefly, samples were mixed with 5X loading buffer and boiled at 95ºC for 5 minutes. Depending on the protein of interest, 5 to 100 ug of lysates were loaded per lane in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) at different percentages of polyacrylamide concentration, ranging from 8 to 15%. Gels were run in Tris-Glycine-SDS (TGS) buffer and proteins were transferred to a nitrocellulose membrane (BioRad) for 90 to 120 minutes (depending on the molecular weight of the protein of interest) in transfer buffer. Ponceau S was used prior to blocking to ensure proper transfer of proteins to membrane. The solution was removed and washed several times with distilled water to remove residual staining solution. Membranes

were blocked with Tris-Buffered Saline (TBS)-Tween and 5% non-fat milk or bovine serum albumin (BSA) depending on the antibody used for 1 hour at RT. Primary antibodies were added to fresh blocking solution and incubated overnight at 4ºC at specified dilutions (Table MM5). After three 10-minute washes with TBS-Tween buffer, membranes were incubated for 1 hour at room temperature (RT) with Horseradish Peroxidase (HRP)-conjugated secondary antibodies (Table MM6) diluted with fresh blocking solution. After further washes with TBS-Tween buffer, membranes were incubated with SuperSignal West Pico Chemiluminescent Substrate (ThermoFisher Scientific) and exposed to autoradiography films (Fujifilm).

Western Blotting buffers

5X loading buffer: 225 mM Tris pH 6.8, 50% glycerol, 5% SDS, 0.07% bromophenol blue, 260 mM DTT

TGS buffer: 25 mM Tris-OH pH 8.3, 192 mM glycine, 5% SDS

Transfer buffer: 50 mM Tris-OH, 396 mM glycine, 0.1 % SDS, 20% ethanol

TBS-Tween: 25 mM Tris-Hcl pH 7.5, 137 mM NaCl, 0.1% Tween-20

Ponceau S: 0.1% Ponceau S, 5% acetic acid

7. Cell cycle analysis

$6x10^6$ DPPA3-OE and control SW1222 and HT29 cells growing in 2D under normoxic conditions (21% O_2) were treated with 5 µg/ml of DOX during the last 5 days to induce DPPA3 expression. Then, cells were trypsinized, centrifuged and resuspended in 0.5 ml of cold PBS. Then, 1.5 ml of ice-cold absolute ethanol was added while gently vortexing the cell pellet and fixed for at least 30 minutes on ice. Cells were washed three times with ice-cold PBS and resuspended in 1 ml of ice-cold PBS. 100 µl of ribonuclease (100 µg/ml, DNase free, Sigma-Aldrich) were added to the samples at RT for 5 minutes. Then, 400 µl of propidium iodide (50 µg/ml in PBS) was added and incubated at 4ºC for 1 hour. Finally, cells were pelleted and resuspended in 500 µl of ice-cold PBS. Cell cycle profile was analysed by flow cytometry using a Navios Flow Cytometer (Beckman Coulter) and FCS express software (De Novo Software). Experiments were performed in triplicates, and the figures represent the average of three experiments ± SD.

8. BrdU analysis

8.1. *In vitro* BrdU labelling and analysis

DPPA3-OE and control SW1222 cells growing in 2D under normoxic conditions (21% O_2) were treated with 5 µg/ml of DOX during the last 5 days to induce DPPA3 expression. During the last 24 hours, half of the samples of each condition underwent hypoxia treatment (0.5% O_2), while the other half were maintained in normoxic conditions. Cells were labelled with 10 µM 5-bromo-2'-deoxyuridine (BrdU) (Sigma) for 6 hours under either normoxic or hypoxic conditions before being trypsinized. Then, $1x10^6$ cells were fixed in 1 ml of ice-cold 70% ethanol in PBS for at least 1 hour. Fixed cells were incubated with 500 µl of 2 N HCl-0.5% Triton X-100 to denature DNA for 30 minutes at RT and neutralized with 500 µl of 0.1 M sodium borate for 2 minutes. Cells were washed with 150 µl of 1% BSA in PBS, pelleted and resuspended in 50 µl of 1% BSA-0.5% Tween-20 in PBS and 10 µl of FITC-anti-BrdU (BD Biosciences) (Table MM5) for 1 hour at RT protected from light. Cells were washed once with 150 µl of 1% BSA in PBS, pelleted and resuspended in 500 µl of PBS. Fluorescence was measured by flow cytometry using a Navios Flow Cytometer (Beckman Coulter) and FCS express software (De Novo Software). Experiments were performed in triplicates, and the figures represent the average of three experiments ± SD. Non-specific binding of antibody was measured by staining cells cultured in the absence of BrdU.

8.2. *In vivo* BrdU labelling

DPPA3-OE and control HT29 cells were injected subcutaneously into both flanks of groups of 3 NOD-SCID mice per experimental condition. When tumours were detectable (within two weeks after the injection), mice were treated with DOX (2 mg/ml; Sigma-Aldrich) *ad libitum* in drinking water containing 5% sucrose (Sigma-Aldrich) to induce DPPA3 expression until experimental end-point. When tumours were detectable, mice were twice injected i.p. with 50 mpk of BrdU (Sigma). When mice were euthanized, a part of the xenograft tumours was fixed for BrdU immunohistochemical staining (described in *Immunohistochemistry and immunofluorescence* section).

9. RNA analysis

9.1. Phenol-chloroform RNA extraction

For *in vitro* RNA extraction of DPPA3-OE and control SW1222 cells growing in 2D, cells were washed twice with PBS and lysed with 1 ml of Trizol reagent (Invitrogen). For *in vitro* RNA extraction of SW1222 cells growing embedded in Matrigel, megacolonies were first harvested

using Matrigel Recovery Solution (BD Bioscience) as described in *Megacolonies extraction from Matrigel* section, followed by direct collection into 1 ml of Trizol reagent (Invitrogen). For *in vivo* RNA extraction, snap-frozen tumour xenograft samples were homogenized using a pestle in 1 ml of Trizol reagent (Invitrogen). Cell lysates from *in vitro* or *in vivo* experiments were mixed with 200 ul of RNase-free chloroform and incubated at RT for 2 minutes. The solution was centrifuged at 13000 rpm for 15 minutes at 4ºC and the supernatant was transferred to a new sterile 1.5 ml microtube and mixed with 500 ul of RNase-free isopropanol and incubated at RT for 10 minutes. RNA was precipitated by centrifugation at 13000 rpm for 30 minutes at 4ºC. Pellets were washed once with 800 ul of 75% RNase-free ethanol and centrifuged at 13000 rpm for 5 minutes at 4ºC. If required, RNA was reprecipitated by resuspending the pellet with 300 ul of H_2O, 30 ul of sodium acetate 3 M (Ambion) and 660 ul of ethanol, all of them RNase-free. Then, washed again with 800 ul of 75% RNase-free ethanol and centrifuged at 13000 rpm for 5 minutes at 4ºC. After evaporation of ethanol traces, the pellet was gently resuspended in DEPC-water and stored at -80ºC prior to Nanodrop quantification.

9.2. Quantitative RT-PCR

For quantitative RT-PCR (qRT-PCR) analysis, RNA integrity was checked by running an aliquot of the RNA sample on a denaturing agarose gel stained with ethidium bromide. Extracted RNA was first DNase-digested (New England Biolabs) followed by RNA retrotranscription with SuperScript III Reverse Transcriptase (Invitrogen) following manufacturer's instructions. Analyses were carried out in triplicates with 40 ng of cDNA using FastStart Universal Master Mix (Roche) in a 7900HT qPCR System cycler (Applied Biosystems). Specific pairs of primers (Table MM7) were designed and used to detect the indicated transcripts. Relative gene expression was determined by the comparative C_T method [284]. We applied geNorm algorithms [284,285]. We applied geNorm algorithms [285] to select TATA-binding protein (*TBP*) and peptidylprolyl isomerase A (Cyclophilin A, *PPIA*) as most stable reference transcripts. The geometric means of the expression values for both housekeeping genes were used to normalize the expression and to calculate the normalized SD of all transcripts analysed. Experiments were performed in triplicates, and the figures represent the average of three experiments ± SD.

9.3. Microarrays

Microarray gene expression analysis was performed on DPPA3-OE, shDPPA3 and *DPPA3* KO SW1222 CRC cells and respective controls. For all conditions, cells were grown as Matrigel-embedded megacolonies. In the case of DPPA3-OE and respective control

conditions, cells were treated with 5 µg/ml of DOX during the last 5 days to induce DPPA3 expression. In order to study the short-term effects on gene expression of SW1222 cells by DPPA3 induction, DPPA3-OE and control cells cultured in 2D treated with 5 µg/ml of DOX during 4 hours were included in the microarray analysis. Total RNA of biological triplicates was isolated by phenol-chloroform extraction as described above. Transcriptomes were determined on a genome wide Human Gene 1.0 ST Array (Affimetrix). RNA integrity was confirmed in an Agilent 2100 Bioanalyzer using RNA Nano-chips. Total RNA was linearly amplified using Ovation® Pico WTA System V2 (Nugen). The resulting ssDNA was used to hybridize Affimetrix microarrays. Hybridization data was acquired using the Affimetrix GeneChip/GeneTitan platforms. We used Partek Genomics Suite 6.6 software (Partek Inc.) to normalize raw CEL files in different combinations as indicated using Robust Multichip Average (RMA) algorithm. Normalized expression values were used to determine the fold change (FC) expression between the respective conditions (DPPA3 overexpression versus control, shDPPA3 versus shC, and DPPA3 KO versus scramble) and its statistical significance in parametric two-tailed paired sample t test (p-value). To generate lists of differentially expressed genes between the indicated conditions, normalized lists of differentially expressed genes between the indicated conditions were cut-off at ± 1.2 FC between conditions and at a significance level of p-value <0.05. For analysis of *DPPA3* transcript expression (Figure 16A) we took advantage of CRC patient-derived xenograft (PDX) samples previously generated in the laboratory from PDX samples derived from colon carcinoma tissues (detailed in *Generation of patient-derived xenograft models* section).

9.4. Gene expression profiling of FFPE patients' tumour samples (nCounter)

DPPA3 transcript expression across chemo-treated stage II/III CRC tumours (Figure R25A) was studied on gene expression profiling data previously generated in the laboratory with FFPE tumour samples from the VHIO (Vall d'Hebron Institute of Oncology) cohort. Briefly, hematoxylin and eosin (H/E) staining was performed in each FFPE tumour tissue of a collection of 53 CRC samples included in the VHIO cohort. Areas enriched with tumour tissue were identified and a minimum of two FFPE tumour tissue cores (1 mm diameter) were collected. RNA was purified using the Roche HighPure FFPE Micro Kit, and approximately 100 ng of total RNA was used to measure expression of *DPPA3* using the NanoString Technologies nCounter Platform. Raw data was log base 2 transformed and normalized using five house-keeping transcripts in the nSolver 2.0 software.

10. DNA analysis

10.1. DNA extraction

DPPA3-OE and control SW1222 cells growing in 2D treated with 5 µg/ml of DOX during the last 5 days to induce DPPA3 expression and undergoing or not hypoxic treatment were lysed with 200-500 µl of lysis buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 0.1% Triton X-100, 0.5% Tween-20, 200 µg/ml proteinase K and 50 µg/ml RNase) depending on cell number and incubated overnight at 55ºC. Proteinase K was inactivated by incubating samples at 95ºC for 20 minutes. Phenol:Chloroform:Isoamyl Alcohol 25:24:1 (Sigma) was added in equal volumes to each sample. The samples were mixed gently and centrifuged for 5 minutes at 13000 rpm. The upper phase was transferred to a new sterile 1.5 ml microtube. Equal volume of chloroform was added and each sample was centrifuged for 5 minutes at 13000 rpm. The upper phase was once again transferred to a new sterile 1.5 ml microtube. DNA samples were precipitated using 1/9 sample volume of 3 M sodium acetate (Ambion) and 2.5 volumes of cold (–20ºC) absolute ethanol. Subsequently, samples were centrifuged at 13000 rpm for 30 minutes at 4°C. Ethanol was removed and each DNA pellet was washed with cold (–20 °C) 70% ethanol (250 µl). Samples were centrifuged at 13000 rpm for 10 minutes, ethanol was removed and the DNA pellet was dried. The DNA sample was then resuspended with appropriate amount of double distilled water and incubated 10 min at 56ºC. The DNA was stored at –20 °C for dot blot or methylome analysis.

10.2. Dot blot

DPPA3-OE and control SW1222 cells growing in 2D under normoxic conditions (21% O_2) were treated with 5 µg/ml of DOX during the last 5 days to induce DPPA3 expression in DPPA3 SW1222 cells. During the last 24 hours, half of the control and DPPA3 samples underwent hypoxia treatment (0.5% O_2), while the other half were maintained in normoxic conditions. DNA was then extracted and 1300 ng and 2600 ng of genomic DNA (gDNA) were used for 5mC and 5hmC analyses respectively and diluted in 65 µl of H_2O. 65 µl of Denature solution 2X was added to samples, which were then incubated at 95ºC for 10 minutes. Immediately after that, samples were put on ice and mixed with 130 µl of cold (4ºC) Neutralized Buffer 2X and let chill on ice for 10 minutes. Then, DNA serial dilutions were prepared in a 96-MW plate on ice by pipetting 260 µl of denatured and neutralized DNA samples into the first column of the plate. The remaining wells were filled with 130 µl of H_2O and 2-fold serial dilutions were performed by mixing 130 µl of each sample with 130 µl of H_2O in the next well. DNA samples were spotted on a positively charged nylon membrane (GE healthcare) in an assembled Bio-Dot apparatus (Bio-Rad). NaOH 0.4 M was used to fix DNA samples to the membrane. The

blotted membrane was washed with SSC buffer 2X for 5 minutes, dried on a Whatman cellulose filter paper for 10 minutes, UV-crosslinked (energy 120.000 µJ/cm^2) for 1 minute and blocked with PBS-5% non-fat milk-0.1% Tween-20 blocking solution for 1 hour at RT. Membranes were then incubated with the primary antibody at specified dilutions (Table MM5). For 5hmC detection, membranes were incubated with primary antibody for 1 hour at RT while in the case of 5mC membranes were incubated with primary antibody overnight at 4ºC. Next, membranes were washed three times, 5 minutes each, in PBS and incubated in appropriate HRP-conjugated secondary antibody solutions for 1 hour at RT at specified dilutions (Table MM6). Membranes were once again washed three times, 5 minutes each, in PBS, developed with SuperSignal West Pico Chemiluminescent Substrate (ThermoFisher Scientific) and exposed to autoradiography films (Fujifilm).

Dot blot buffers

Denature solution 2X: 0.8 M NaOH, 20 mM EDTA

Neutralized Buffer 2X: 2M ammonium acetate pH 7 (4ºC)

SSC buffer 2X: 0.3 M NaCl, 30 mM sodium citrate

10.3. Methylome

Genomic DNA of biological duplicates was extracted from normoxic control and DPPA3-OE SW1222 cells growing in 2D treated with 5 µg/ml of DOX for five days was quantified with the Qubit dsDNA BR Assay kit and 600 ng of DNA was bisulfite-treated using the EZ DNA Methylation-Gold kit (Zymo Research). Then, the Infinium HD Methylation Assay (Illumina) (amplification, fragmentation, precipitation, hybridization, wash, extension, staining, and imaging) was performed according to the manufacturer's explicit specifications. Infinium Methylation EPIC BeadChip arrays were processed using R software (v.3.5.1) and minfi package (v.1.28.4) (minfi). Raw IDAT files were normalized using single-sample normal-exponential out-of-band method (ssNOOB). Probes with poor performance, probes containing SNPs, and cross-reactive probes (Pidsley) were filtered out prior to downstream analyses. Differentially methylated probes were assessed using minfi. For each probe, association between methylation and phenotype was tested using F-test with variance shrinkage. P-values were then adjusted for multiple testing using Bonferroni correction and probes were considered significant when the adjusted p-value < 0.05.

11. Immunohistochemistry and immunofluorescence

11.1. Megacolonies immunofluorescence staining and analysis

Matrigel-embedded megacolonies were grown on coverslips. Then, cultures were fixed with 4% PFA during 1 hour at RT and permeabilised with 1% Triton X-100/PBS for three hours. Samples were incubated with blocking solution (0.1% Triton X-100 and 3% BSA in PBS) overnight at 4ºC. Primary antibodies (Table MM5) with their corresponding dilution in blocking solution were added and incubated 24 hours at RT. Then, megacolonies were incubated with the corresponding secondary Alexa-Fluor antibody (Table MM6) overnight at RT. Hoechst 33342 (5 µg/ml) was used as counterstaining to detect cell nuclei in all samples. A Nikon C2+ Confocal Microscope was used to visualize fluorescence and images acquired using NIS-Elements Advanced Research software. MUC2 staining was calculated from dividing MUC2 integrated density by the sectioned megacolony area.

11.2. Immunohistochemical staining and analysis

For immunohistochemical analyses, Formalin-Fixed Paraffin-Embedded (FFPE) tissue sections were routine deparaffinated, rehydrated and treated with 1mM EDTA pH 9 for CA9 staining or with 10 mM sodium citrate buffer pH 6 for p-H3 (Ser10), p-H2AX (Ser139), MUC2, CK20 or BrdU stainings. For all antibodies, after the blocking of endogenous peroxidase activity, slides were permeabilized with 1% Tween-20 in PBS for 15 min. Then, tissue specimens were blocked for 1 hour with 3% BSA in PBS. For BrdU staining, after antigen retrieval, DNA was denatured to single strands by immersing paraffin-sections in 2N HCl for 15 minutes at RT and then neutralizing with 0.1 M borate buffer pH 8.5 for 15 minutes at RT. Sections were permeabilized and blocked using 0.5% Triton X-100 and 5% normal donkey serum diluted in PBS for 1 hour at RT. After blocking, sections were then incubated with corresponding primary antibodies (Table MM5) diluted in blocking solution at 4ºC overnight. After washing, sections were incubated with corresponding HRP-conjugated secondary antibodies (Table MM6) at corresponding dilutions for 1 hour at RT. After washing, Chromogen DAB/substrate reagent (DAKO) was added onto the slides and incubated up to 10 minutes. Finally, the slides were counterstained with haematoxylin, dehydrated, and mounted. NanoZoomer 2.0-HT Digital slide scanner C9600 (Hamamatsu Photonics K.K.) was used to visualize and assess the immunostainings. Image quantification was performed using the QuPath open source software. The percentage of tumour cells positive for CA9, p-H3 (Ser10), or p-H2AX (Ser139) staining of each tumoural section was used for CA9, p-H3 (Ser10), or p-H2AX (Ser139) quantifications respectively. For MUC2 quantification, the MUC2-positive area of each tumoural section was divided by the tumoural area. For BrdU quantification,

percentages of BrdU-positive tumour cells of each tumoural section were used to quantify the fold-change (FC) between each specimen in the irinotecan group versus the average of those in the vehicle group. For CK20 quantification in orthotopic tumour xenograft models, CK20-positive tumour cells foci were counted in each lung section. Foci containing ≤4 tumour cells were considered small, those having 5-30 tumour cells were considered medium whereas foci with >30 tumour cells were considered large. To calculate the number of tumoural cells in primary tumours from orthotopic experiments, the number of tumoural cells of each tumoural section was divided by the tumoural area.

For immunofluorescence detection of FFPE sections, antigen retrieval was performed in 10 mM sodium citrate buffer (pH 6). Slides were then permeabilized with 1% Tween-20 in PBS for 15 min. Then, tissue specimens were blocked for 1 hour with PBS containing 3% of BSA and incubated with H2AK119ub1 antibody at specified dilution (Table MM5) in blocking solution overnight at 4 ºC. Finally, slides were incubated with the corresponding Alexa Fluor secondary antibody (Table MM6) at a dilution of 1:200 for 1 hour at RT. Hoechst 33342 (5 µg/ml) was used as counterstaining to detect cell nuclei in all samples. A Nikon C2+ Confocal Microscope was used to visualize fluorescence and images acquired using NIS-Elements Advanced Research software. Image quantification was performed using the ImageJ open source software. At least five standard confocal images were taken for each section. H2AK119ub1 levels were calculated from dividing H2AK119ub1 integrated density by the tumoural area.

11.3. Alcian blue staining and analysis

Alcian blue staining was performed on FFPE tumour xenograft sections to identify goblet cells by the Translational Molecular Pathology service of VHIR (Vall d'Hebron Institut de Recerca). Briefly, following deparaffinization, FFPE sections were hydrated and incubated in absolute ethanol for 3 minutes, acetic acid for 5 minutes, 1% Alcian blue for 30 minutes and nuclear fast red for 10 minutes. 5-minutes PBS washes were performed 3 times after each step. Samples were then dehydrated and coverslipped with mounting media. NanoZoomer 2.0-HT Digital slide scanner C9600 (Hamamatsu Photonics K.K.) was used to visualize and assess the Alcian blue staining. Image quantification was performed using the ImageJ open source software. Alcian blue staining was analysed by quantifying the percentage of Alcian blue-positive area of each tumoural section.

12. Megacolonies extraction from Matrigel

Cultured Matrigel-embedded SW1222 megacolonies growing in 48-MW plates were washed twice with ice-cold PBS and harvested with a total of 400 µl/well of ice-cold Matrigel Recovery

Solution (BD Bioscience) as per manufacturer's protocol. Harvested megacolonies were incubated on ice for 1 hour, pelleted and washed twice with ice-cold PBS. Extracted megacolonies were subsequently used for SCCC and RCCC isolation or RNA analysis.

12.1. Isolation of SCCC and RCCC from Matrigel

To obtain SCCC and RCCC from H2BeGFP-infected shC and shDPPA3 SW1222 Matrigel-embedded megacolonies, DOX pulse-chased cultures (described in *Apoptosis* section) were harvested using Matrigel Recovery Solution (BD Bioscience) as described in *Megacolonies extraction from Matrigel* section. After extraction from Matrigel, megacolonies were filtered through a 100 µm cell strainer (Corning) to purify megacolonies bigger than 100 µm of diameter. Megacolonies were collected from the cell strainer surface and dissociated using trypsin-EDTA to obtain a single cell suspension. Finally, single-cells were resuspended in sorting medium: 4 mM Glutamine (Life Technologies), 20% FBS, 1% penicillin/streptomycin, 10 µM Y-27632 (Calbiochem) diluted in CO2-independent medium (Life Technologies). DAPI (2 µg/ml, Roche) was added to exclude dead cells and cellular aggregates in all cell suspensions. Finally, live SCCC and RCCC were sorted using a MoFlo Legacy cell sorter (Beckman Coulter). The population of cells retaining an H2BeGFP signal equivalent to that observed in cells continuously exposed to DOX, were considered and sorted as SCCC. The RCCC fraction comprised around 10-20% of all cells with an H2BeGFP signal one order of magnitude lower than the SCCC fraction.

13. Apoptosis assays

Megacolonies from shC and shDPPA3 cells infected with H2BeGFP were grown in Matrigel and treated with 5 µg/ml of DOX (pulse) for 9 days. After 12 days without DOX (chase), cells were treated with 20 µM of oxaliplatin for an additional 5 days before apoptosis was measured. Five days after treatment, single-cells were recovered from Matrigel cultures using Matrigel Recovery Solution (BD Bioscience) as described in *Isolation of SCCC and RCCC from Matrigel* section and SCCC and RCCC single-cells were evaluated for apoptosis using the Annexin V-APC kit (Bender MedSystems). Dead cells were detected as DAPI positive (2 µg/ml, Roche). Finally, cells were analysed by flow cytometry using a Navios Flow Cytometer (Beckman Coulter). Experiments were performed in triplicates, and figures represent the average of three experiments ± SD.

14. *In vivo* experiments

14.1. Subcutaneous tumour xenografts

NOD-SCID (NOD.CB17-Prkdcscid/NcrCrl) and Nude (Crl:NU-Foxn1nu) were purchased from Charles River Laboratories. $1x10^6$ DPPA3-OE and control HT29 or SW1222 cells or shDPPA3 and shC SW1222 cells resuspended in 50 µl of PBS were mixed with 50 µl of Matrigel and injected subcutaneously into both flanks of 3-6 mice per experimental group. In the case of DPPA3-OE and control HT29 or SW1222 cells, when tumours were detectable (within two weeks after the injection), mice were treated with DOX (2 mg/ml; Sigma-Aldrich) *ad libitum* in drinking water containing 5% sucrose (Sigma-Aldrich) to induce DPPA3 expression until specified. When matching end-point criteria, mice were euthanized and parts of the xenograft tumours were snap-frozen for RNA/protein extraction or fixed for histological analysis. Tumour growth was monitored by caliper measurement three times per week and volume was estimated using the following formula: $V = (\text{length} \times \text{width}^2)/2$, where length represents the largest tumour diameter and width represents the perpendicular tumour diameter.

14.2. Orthotopic tumour xenografts

$1x10^6$ DPPA3-OE and control HT29 cells or $2x10^6$ DPPA3-OE and control SW1222 cells resuspended in 50 µl of PBS were prepared to inject into the cecal wall of 16-19 NOD-SCID mice per experimental group. We anesthetized mice with continuous vaporized isoflurane and disinfected the mouse's abdominal area with an iodine swab. For injection, the cecum was brought outside the abdomen onto a moist sterile gauze and tumour cells were orthotopically injected into the cecal wall of mice. After injection, we returned the cecum to its normal anatomic position within the abdomen. Closure of the abdomen was performed by approximating fascial edges before closing the skin with nylon sutures. To induce DPPA3 expression, 5 days after injection, mice were treated with DOX (2 mg/ml; Sigma-Aldrich) *ad libitum* in drinking water containing 5% sucrose (Sigma-Aldrich) until experimental end-point. When matching end-point criteria, mice were euthanized and the cecum containing the primary tumour was extracted, weighed and fixed for histological analysis. Livers and lungs were also extracted and fixed for histological analysis.

14.3. *In vivo* chemoresistance and tumour regrowth

$1x10^6$ DPPA3-OE and control HT29 cells were prepared as described in *Subcutaneous tumour xenografts* section and injected subcutaneously into both flanks of 9 NOD-SCID mice per experimental group. When tumours were detectable (within two weeks after the injection), mice were treated with DOX (2 mg/ml; Sigma-Aldrich) *ad libitum* in drinking water containing

5% sucrose (Sigma-Aldrich) to induce DPPA3 expression until experimental end-point. Experimental groups were treated with either vehicle, oxaliplatin (5 mpk), 5-FU (40 mpk) or irinotecan (50 mpk). Oxaliplatin and irinotecan were administered once per week while 5-FU was administered twice per week, all of them by intraperitoneal injection. After mice were euthanized, parts of the xenograft tumours were snap-frozen for RNA/protein extraction or fixed for histological analysis. Tumours were measured three times per week and volume was estimated using the formula specified in *Subcutaneous tumour xenografts* section.

For the tumour regrowth experiment (Figures R24A and B), DOX and chemotherapy were withdrawn from mice in vehicle and irinotecan experimental groups (6 mice per group). Tumour growth monitoring was done three times per week until 10 days after, where mice were euthanized.

14.4. Analysis of Area Under the Curve

Area under the curve (AUC) of each tumour was calculated using the GraphPad Prism software in order to monitor the tumour growth rate in experiments performed with subcutaneous xenograft models. We used AUC as an alternative method to measure tumour volumes and compare tumour growth curves in subcutaneous tumour xenografts experiments. This single numerical value obtained by AUC calculation has been shown to be easy to obtain for individual curves, reflects the entire tumour growth curve through a single number, can be easily modified to obtain data for defined sections of the growth curve, and allows easier comparisons between groups [286].

14.5. Analysis of Progression-Free Survival

Progression-Free Survival (PFS) of mice was calculated in the tumour regrowth experiment performed with the DPPA3-OE and control HT29 cells (Figure R24C). According to the Response Evaluation Criteria in Solid Tumors (RECIST) [287], a percentage change in tumour size of about 20% from baseline (time of DOX and irinotecan treatments withdrawal) determined the progression of the disease. A second order polynomial (quadratic) equation was applied to calculate the elapsed time between baseline tumour size and a 20% increase in tumour volume using the GraphPad software. Significance was calculated using the log-rank (Mantel-Cox) test.

15. Generation of patient-derived xenograft models

We took advantage of PDX models previously generated and described in the laboratory [288]. Briefly, colon carcinoma tissues obtained upon surgery were washed 3 times in cold PBS solution and incubated overnight in DMEM/F12 (Gibco) containing a cocktail of antibiotics and

antifungals (penicillin (250 U/ml), streptomycin (250 µg/ml), fungizone (10 µg/ml), kanamycin (10 µg/ml), gentamycin (50 µg/ml), and nystatin (5 µg/ml; Sigma-Aldrich). Enzymatic digestion was performed using collagenase (1.5 mg/ml; Sigma-Aldrich) and DNase I (20 µg/ml; Sigma-Aldrich) in a medium supplemented with a cocktail of antibiotics and antifungals (penicillin (250 units/ml), streptomycin (250 µg/ml), fungizone (10 µg/ml), kanamycin (10 µg/ml), gentamycin (50 µg/ml), and nystatin (5 µg/ml; Sigma-Aldrich) during 1 hour at 37°C with intermittent pipetting every 15 minutes to disperse cells. The dissociated sample was then filtered (100 µm pore size) and washed with fresh medium. Red blood cells were lysed by brief exposure to ammonium chloride and the sample was washed again. Finally, cells were used for subcutaneous injections into NOD-SCID mice to generate subcutaneous tumours further used for histological and RNA analysis (Figure 16A-C).

16. DPPA3 gene expression signature

Based on the microarray gene expression analysis performed in DPPA3-OE, shDPPA3 and *DPPA3* KO SW1222 CRC cells growing as Matrigel-embedded megacolonies or cultured in 2D (DPPA3-OE cells), we derived the DPPA3 gene expression signature (DPPA3 Sig). We considered as DPPA3-induced genes (50 in total) those positively regulated (≥1.2 FC between conditions, p-value < 0.05) in DPPA3-OE cells that were common with those negatively regulated (≤-1.2 FC between conditions, p-value < 0.05) in shDPPA3 or *DPPA3* KO cells in at least 3 of the 4 different conditions (DPPA3-OE, shDPPA3 and DPPA3 KO megacolonies and 2D-cultured DPPA3-OE cells). We considered as DPPA3-repressed genes (32 in total) those negatively regulated (≤-1.2 FC between conditions, p-value < 0.05) in DPPA3-OE cells that were common with those positively regulated (≥1.2 FC between conditions, p-value < 0.05) in shDPPA3 or *DPPA3* KO cells in at least 3 out of the 4 different conditions. All genes included in DPPA3 Sig are listed in Table A1. To apply the DPPA3 Sig on the clinical cohort used (GSE39582), we calculated the median expression of the 50 DPPA3-induced genes minus the median expression of the 32 DPPA3-repressed genes in all patient samples. This calculation generates a unique enrichment score (ES) for each sample. We selected the upper quartile to define positivity for the signature (DPPA3 Sig High).

16.1. Correlation between DPPA3 and dormancy signatures

To calculate the correlation between the ES of the DPPA3 Sig and a dormancy signature already published and validated by Kim et al. in breast cancer patients [289], we used samples from treated stage II and III CRC patients of the clinical cohort GSE39582 (n = 200) (Figure R26B). The dormancy signature ES was calculated following the same procedure used in the DPPA3 Sig (described in *DPPA3 gene expression signature* section). Correlation between the

enrichment scores of DPPA3 and dormancy signatures was calculated applying Pearson's test using GraphPad Prism software.

17. Clinical Cohorts

cBioPortal for Cancer Genomics (public): We examined *DPPA3* (NM_199286) expression across cancer types using the free web database known as cBioPortal (http://cbioportal.org) that allows users to explore, analyse and visualize the multidimensional cancer genomics data. The web portal provided information for the tumour samples from The Cancer Genome Atlas (TCGA) Pan-Cancer Atlas resource.

Oncomine (public): The transcript expression levels of *DPPA3* (NM_199286) in seminoma and prostate cancer samples were analysed using Oncomine Compendium of Expression Array data (https://www.oncomine.org/resource/login.html). Oncomine automatically computes differential expression profiles for cancer types and subtypes so that they can be easily queried for a gene of interest [290]. Significance was evaluated by a parametric two-tailed sample t test with Welch correction using GraphPad Prism software and the p-value for statistic significance was set up as 0.05. The dataset, reporter ID and platform used were as follows:

- Normal tissue and seminoma samples: Korkola Seminoma Statistics, 231385_at, Human Genome U133B Array.
- Normal tissue and prostate cancer samples: Varambilly Seminoma Statistics, 231385_at, Human Genome U133 Plus 2.0 Array.

CANCERTOOL (public): The correlation between transcript expression levels of *DPPA3* (NM_199286) and *CA9* (NM_001216) in prostate cancer primary tumour and metastasis samples (Figure R16D) were analysed using CANCERTOOL data (http://web.bioinformatics.cicbiogune.es/CANCERTOOL/). CANCERTOOL is a web-based interface that provides rapid and comprehensive visualization of gene expression data for the gene(s) of interest in well-annotated cancer datasets. It also carries out gene-to-gene correlations in multiple datasets at the same time or using preset patient groups [291]. The study name, GEO accession number and platform used were as follows: Taylor, GSE21034, Affimetrix Human Exon 1.0 ST Array.

GSE39582 (public): Affymetrix Human Genome U133 Plus 2.0 Array data was downloaded from GEO website (https://www.ncbi.nlm.nih.gov/geo). We used Partek Genomics Suite 6.6 software (Partek Inc.) to normalize raw CEL files using RMA algorithm. Normalized expression values were used to compare the expression between the respective patient samples. The DPPA3 Sig ES was calculated and samples at the upper quartile were classified as positive (High) while the rest were classified as negative (Medium/Low) (Figures R26A and B). We

applied the survival analysis to patients that received adjuvant chemotherapy (n = 200) or not (n = 251) and that had outcome data available, corresponding to high-risk stage II or stage III cases. Progression-free survival (PFS) was defined as months from diagnosis to relapse due to any cause. We performed Cox Proportional Hazards univariable modelling using GraphPad Prism software.

VHIO (private): *DPPA3* expression was evaluated by NanoString Technologies nCounter Platform (described in *Gene expression profiling of FFPE patients' tumour samples (nCounter)* section) on a cohort of 53 tumours from high-risk stage II or stage III CRC patients. Samples were plotted based on *DPPA3* expression and we observed two clearly differentiated groups (Figure R25A). Hence, samples in the upper group were classified as high whereas those samples in the lower group were classified as low/negative. Overall survival (OS) was shown as years from diagnosis to death due to any cause while PFS was defined as years from diagnosis to relapse due to any cause. We performed Cox Proportional Hazards univariable modelling using GraphPad Prism software.

18. Functional Gene Set Enrichment Analysis

Gene set enrichment analyses (GSEA) and leading-edge analyses were performed with the GSEA platform of the Broad Institute (http://www.broadinstitute.org/gsea) that included calculation of the corresponding p-values and enrichment scores. We used custom gene sets and publicly available gene sets (Molecular Signatures Database v6.2;http://www.broadinstitute.org/gsea). Custom gene sets: SCCC PAN-CANCER UPREGULATED and SCCC PAN-CANCER DOWNREGULATED gene sets included genes differentially expressed (FC≥1.3 and FC≤-1.3 respectively, p-value <0.05) in the gene expression analysis of the different cancer models used (CRC, melanoma and glioblastoma) for the generation of the PanCancer-SCCC signature [269]. The list of genes is shown in Table A2. PRIMORDIAL GERM CELLS Gene set was composed of genes typical of germ cells (*APOBEC1, PRDM1, DDX4, DAZL, DPPA3, ZFP42, GDF3, AICDA, ACR, POU5F1, PRDM14, NANOS3, EHMT2, DNMT1, ACRBP, NANOG, TET2, SOX2*).

Gene expression analysis of DPPA3-OE SW1222 megacolonies (detailed in *Microarrays* section) and SW1222 SCCC previously generated in the laboratory [269] were used for GSEA analysis (Figures R3, R4 and R5A). For visualization of GSEA results on dot plots (Figures R3A and R4), we used ggplot2 package (R software).

19. Statistics

The statistical significance of different groups of values obtained in several experiments across the work is described in each corresponding figure legend. We evaluated the correlation between CA9 protein levels content and *DPPA3* transcript expression (Figure 16C) in CRC patient samples, and between *DPPA3* and *CA9* transcript expression (Figure 16D) in prostate cancer samples using a Pearson´s test to evaluate significance. Correlation between the enrichment scores of DPPA3 and dormancy signatures was calculated applying Pearson's test. For unpaired two-tailed t tests, Welch's correction was applied in figures R1, R15A-B, R16B, R18E, R20A-C, R21A, R24D, R26D, R28A (DPPA3 overexpression, bottom), R29B and R30B-C. For ANOVA analyses, Dunnett's correction was applied in figures R2B, R23B, R28A (DPPA3 knockout and knockdown, top and middle), R29A; Tukey's correction was applied in figures R8D, R9B, R10B, R11A, R12A-B, R13B, R18B, R19A; and Sidak's correction was applied in figures R9D, R10A, R17A-B, R23C, R24B, R28B, R29C, R30A. For simplicity, only the biologically relevant significant differences were shown in figure R9D. Codes with asterisks were used to indicate different levels of statistical significance: $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$; $****P \leq 0.0001$.

20. Study approval

Experiments with mice were conducted following the European Union's animal care directive (86/609/CEE) and were approved by the Ethical Committee of Animal Experimentation of the VHIR - the Vall d'Hebron Research Institute (approval ID, 06/12 CEEA, 87/12 CEEA, 17/15 CEEA and 18/15 CEEA). Human tumour samples for PDX were obtained after approval from Ethics Committee of the Vall d'Hebron University Hospital, Barcelona, Spain (approval ID, PR(IR)79/2009). Written informed consent was signed by all patients.

Oligo name	Direction	Sequence (5' to 3')	
sgRNA DPPA3	Forward	**CACCG**TTTAATCCAACCTACATCCC	sgRNA guides + BbsI adapters
sgRNA DPPA3	Reverse	**AAAC**GGGATGTAGGTTGGATTAAA**C**	
sgRNA Scramble	Forward	**CACCG**GCACTACCAGAGCTAACTCA	
sgRNA Scramble	Reverse	**AAAC**TGAGTTAGCTCTGGTAGTGC**C**	
DPPA3 exon 1 screening	Forward	AGATATGACATCCGACAGATTAT	CRISPR Screening
DPPA3 exon 1 screening	Reverse	CCCCTATAGTCAGATAGTCAAGAA	

Table MM1. List of oligo sequences used for the generation of control (scramble) and *DPPA3* knockout cell lines and validation (DPPA3 CRISPR Screening) of *DPPA3* mutation in exon 1. BbsI adapter sequences of guide sgRNAs is underlined and highlighted in bold.

Oligo name	Direction	Sequence (5' to 3')
NheI-DPPA3	Forward	TTTGCTAGCTACCGAGGAGATCTGCC
PacI-DPPA3	Reverse	GGTTAATTAAACAGCTATGACCGCGG

Table MM2. Primers used for DPPA3-CMYC-FLAG cloning from the pCMV6-DPPA3 vector (Origene, RC214676) into the pSIN-TRE-hPGK-rtTA2-P2A-mCherry vector to generate the DPPA3-OE cell lines.

Oligo name	Sequence (5' to 3')
shC (non-silencing) (SHC202)	CCGGCAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTT GTTTTT
shDPPA3-1 (TRCN0000167070)	CCGGCGTTGATAAAGAACCTTAGTACTCGAGTACTAAGGTTCTTTATCAAC GTTTTTTG
shDPPA3-2 (TRCN0000172771)	CCGGCCATGTGTCTTAGAAGCCCAACTCGAGTTGGGCTTCTAAGACACAT GGTTTTTTG
shDPPA3-3 (TRCN0000254440)	CCGGCTATGCTAGTATAGACTATACCTCGAGGTATAGTCTATACTAGCATA GTTTTTG

Table MM3. List of oligo sequences used for the generation of shC and shDPPA3 cell lines.

Oligo name	Direction	Sequence (5' to 3')	
XhoI-*CA9* promoter	Forward	CCGCTCGAGACCTGCCCCTCACTCCAC	*CA9* promoter cloning
HindIII-*CA9* promoter	Reverse	GGGAAGCTTGGCTGACTGTGGGGTGTC	
HRE-mutated *CA9* promoter oligo	Reverse	GCTCTCGTTTCCAATGC**TG**GTACAGCCCGTAC ACAC	*CA9* promoter mutagenesis

Table MM4. Primers used for the cloning of the wild-type *CA9* promoter region (-1476/+42bp) into the pGL3 basic luciferase reporter (*CA9* promoter cloning), and for the generation of the same vector mutagenized at the HRE core sequence (*CA9* promoter mutagenesis). AC bases from the consensus HRE sequence were mutagenized to obtain the final mutated sequence. Mutagenized bases are underlined and highlighted in bold in the HRE-mutated *CA9* promoter oligo sequence.

Target	Host specie	Provider	Reference	Dilution
5hmC	Rabbit	Active Motif	39769	DB 1/10000
5mC	Mouse	Calbiochem	16233D3	DB 1/1000
BrdU	Rat	Abcam	ab6326	IHC 1/50
	Mouse	BD Biosciences	347583	FC 1/5
CA9	Rabbit	Novus Biologicals	NB100-417	WB 1/4000 IHC, IF 1/1000
CD31	Rabbit	Abcam	ab28364	WB 1/500
CK20	Mouse	Dako	M7019	IHC 1/100
DNMT1	Rabbit	EpiGentek	A-1700	WB 1/4000
ERK1/2 MAPK	Rabbit	Cell Signaling	9102	WB 1/5000
Flag	Rabbit	Sigma	F7425	WB 1/2000
	Mouse	Sigma	F1804	WB 1/2000
GAPDH	Mouse	Lab (hybridome)	Clone FF26A/F9	WB 1/10000
GFP	Mouse	Santa Cruz	sc-9996	IF 1/100
H2AK119ub1	Rabbit	Cell Signaling	8240	IHC 1/1500
H3	Rabbit	Abcam	ab1791	WB 1/20000
H3K9me2	Rabbit	Millipore	07-441	WB 1/2000
HIF1α	Mouse	BD Transduction	610958	WB 1/1000
Lamin B1	Rabbit	Abcam	ab16048	WB 1/10000
MUC2	Mouse	BD Pharmingen	555926	IHC, IF 1/100
NCID1	Rabbit	Cell Signaling	4147	WB 1/1000
p38 MAPK	Rabbit	Cell Signaling	9212	WB 1/1000
p-ERK1/2 MAPK (Thr202/Tyr204)	Rabbit	Cell Signaling	9101	WB 1/5000
p-H2AX (Ser139)	Rabbit	Cell Signaling	2577	WB 1/5000 IHC 1/100
p-H3 (Ser10)	Rabbit	Millipore	06-570	WB 1/4000 IHC 1/100
p-p38 MAPK (Thr180/Tyr182)	Rabbit	Cell Signaling	9211	WB 1/1000
p-S6 Ribosomal Protein (Ser240/244)	Rabbit	Cell Signaling	2215	WB 1/2000
S6 Ribosomal Protein	Mouse	Cell Signaling	2317	WB 1/5000
UHRF1	Mouse	BD Transduction	612264	WB 1/5000
α-Tubulin	Mouse	Sigma	T9026	WB 1/10000

Table MM5. Primary antibodies used, their commercial information and dilutions used for each application are listed. DB, Dot blot; IHC, Immunohistochemistry; FC, Flow cytometry; WB, Western blot; IF, Immunofluorescence.

Antibody	Provider	Reference	Application
Goat anti-mouse IgG HRP conjugated	ThermoFisher	G-21040	WB, IHC
Donkey anti-rabbit IgG HRP conjugated	Jackson ImmunoResearch	711-035-152	
Goat Anti-Rat IgG HRP conjugated	Abcam	ab205720	IHC
Alexa Fluor 488 goat anti-mouse IgG	ThermoFisher	A11001	IF, IHC
Alexa Fluor 488 goat anti-rabbit IgG	ThermoFisher	A21206	
Alexa Fluor 555 goat anti-mouse IgG	ThermoFisher	A31570	
Alexa Fluor 555 goat anti-rabbit IgG	ThermoFisher	A31572	

Table MM6. Secondary antibodies used, their commercial information and application are listed. IHC, Immunohistochemistry; WB, Western blot; IF, Immunofluorescence.

Oligo name	Direction	Sequence (5' to 3')
CA9	Forward	GGCTACAGCTGAACTTCCGA
	Reverse	AATTCAGCTGGACTGGCTCA
CUL2	Forward	GTG ATC GGC CCGTCCTTCTC
	Reverse	GTAGTGTTGAAATCTGTCTCCGC
DPPA3	Forward	CTACATCCCAGGGTCTCC
	Reverse	TGATAGTCAAGTTACTAAGGTTCT
DPPA3-FLAG	Forward	CAAGACACCAAGCCACTTC
	Reverse	CGTCGTCATCCTTGTAATCC
EGLN1	Forward	AGACTGGGATGCCAAGGTAAG
	Reverse	ATGTCAGCAAACTGGGCTTTG
EIF4EBP1	Forward	CCCGGGAGGTACCAGGATCA
	Reverse	ATCCCTTGGGGGTGTTTTGG
EIF4G1	Forward	TTCTTAATCGAGGGGGTGCTG
	Reverse	TCATTTGTGTCGCTTGTGGC
EIF4G2	Forward	TCATGCCTCCCACACAATCG
	Reverse	GCCTCAGGCTAATCTGGCTTTT
ELOB	Forward	AGCACGGTGTTCGAACTGAA
	Reverse	TGGTCATCCTTGTACAGCCG
ELOC	Forward	CGGGACTGACGAGAAACTAC
	Reverse	TCCTTCACAGCCACCATAGG
GCNT3	Forward	TGTTCACATTTGCTGCCACG
	Reverse	GTGGAGGAGGACACAATCCTTT
GDF3	Forward	TCCCGAGACTTATGCTAC
	Reverse	GGAGGAAGCTTGGGAAAT
HEPH	Forward	GGAGAAAGGCCCTGTAACCA
	Reverse	TGGCCCACATTACTC TTTGCT
HIF1A	Forward	TTCCTTCTCTTCTCCGCGTG
	Reverse	ACTTATCTTTTTCTTGTCGTTCGC
LDHA	Forward	ACGTGCATTCCCGATTCCTT
	Reverse	GGGGGTCTGTTCTTCCTTTAGA
LGR5	Forward	GACTATGCCTTTGGAAACC
	Reverse	GGAGCCCATCAAAGCATT
MUC2	Forward	CACCTGGCTGTGCTTAAC
	Reverse	TGTAGGCATCGCTCTTCT
OLFM4	Forward	GTATGGCGTCTAGGGATT
	Reverse	AGGTTTCTTCCAGGCATG
PDK1	Forward	TCCAGAGTCTTCAGGAGCTT
	Reverse	ACCATGTTCTTCTAGGCCTTTCAT
PPIA	Forward	CTCCTTTGAGCTGTTTGCAG
	Reverse	CACCACATGCTTGCCATCC
PRKAA1	Forward	GAACAAGTTGTGGCTCACCC
	Reverse	TCTGGGCCTGCATACAATCTTC

RBX1	Forward	GGCAGCGATGGATGTGGATA
	Reverse	AGGGCTACTGCATTCCACTTT
SLC2A1	Forward	TGGCATCAACGCTGTCTTCT
	Reverse	AGCCAATGGTGGCATACACA
TBP	Forward	CGGCTGTTTAACTTCGCTTC
	Reverse	CACACGCCAAGAAACAGTGA
TET1	Forward	CCTGGACTTCTGTGCTCAT
	Reverse	CCCAAAGAGCGGTTATCT
TET2	Forward	GTTAGAAAGGAGACCCGA
	Reverse	CAGGAGCAAAGGCAAGTA
TET3	Forward	CCGAGAAGAAGAAGATTCAG
	Reverse	GACAATCCACCCTTCAGAGA
UBE2D1	Forward	AGTGATCTACAGCGCGATCC
	Reverse	TGCGCTATCAGGAGTCATCTC
VEGFA	Forward	CATTGGAGCCTTGCCTTG
	Reverse	CGCTGATAGACATCCATG
VHL	Forward	ATCCACAGCTACCGAGGTCA
	Reverse	GGTTAACCAGAAGCCCATCGT

Table MM7. Primers used for qRT-PCR analysis.

RESULTS

1. Generation of DPPA3 cancer cell models and gene expression analysis

The molecular characterization of SCCC performed in our laboratory [269] uncovered a distinctive core set of epigenetic and pluripotency factors of unknown role in the biology of tumour dormancy. DPPA3 was one of these factors presenting high expression in SCCC. It is a key factor in embryonic development that controls the epigenetic landscape in pre-implantation embryos and PGCs [231]. Interestingly, DPPA3 has not been formally related to cancer yet. Its high expression in SCCC and PGCs that present traits of undifferentiation and pluripotency made us consider DPPA3 as an attractive candidate to be characterized in cancer and dormancy.

As a first overview, we evaluated *DPPA3* expression across different types of tumours in the open source database cBioPortal and observed that TGCTs were the cancer type that highest expressed the epigenetic factor (Figure R1A). Although this gene is expressed at very low levels in most human adult tissues, its expression is enriched in ovary and testis, which contain pluripotent cells [254]. In order to assess that the elevated transcript levels of *DPPA3* in TGCTs were not due to its high basal expression in normal testis, we compared *DPPA3* transcript levels between normal testicular tissue and tumour samples in the Oncomine database. We observed that seminomas, a TGCTs subtype, showed higher *DPPA3* mRNA levels than normal testis (Figure R1B). Indeed, the genomic region 12p13 that harbours *DPPA3, NANOG* and *GDF3* pluripotency genes has been shown to be amplified in TGCTs [267]. Of note, we also observed that despite prostate tumours were not among the ones expressing highest *DPPA3* mRNA levels (Figure R1A) the gene was induced in prostate tumour samples compared to normal tissue (Figure R1B). Although CRC samples did not exhibit high levels of *DPPA3* mRNA compared to other cancer types (Figure R1A), our initial results in SCCC derived from CRC cells encouraged us to study its role in this type of cancer.

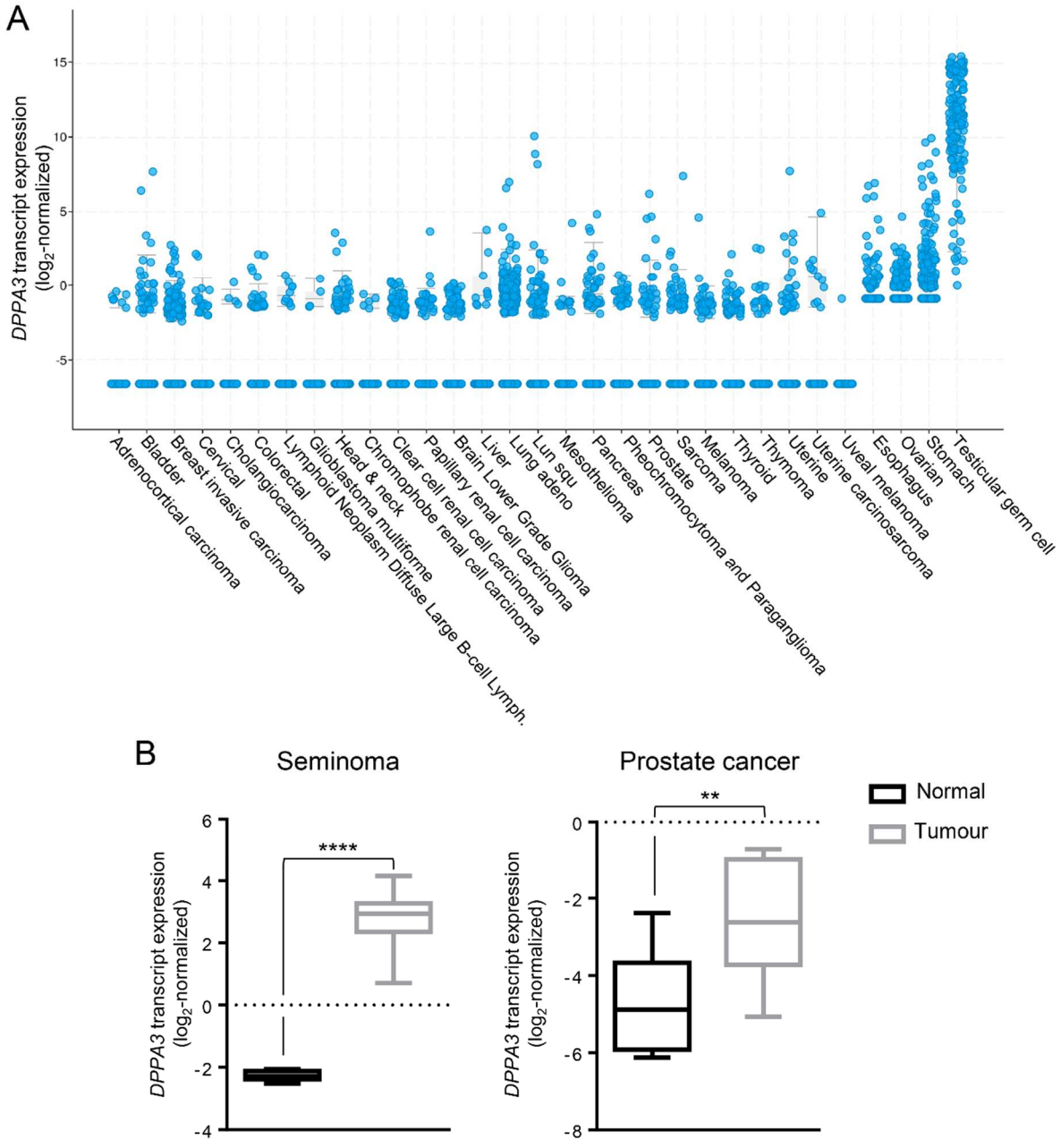


Figure R1. DPPA3 is upregulated in seminoma and prostate tumours. (A, B) *DPPA3* transcript levels represented as log_2-normalized values (A) across 31 The Cancer Genome Atlas (TCGA) cancer types, (B, left) in normal testis (normal, n = 6) and seminoma (tumour, n = 12) samples, and (B, right) in prostate gland (normal, n = 6) and prostate cancer samples (tumour, n = 12). (B) $^{**}P \leq 0.01$; $^{****}P \leq 0.0001$ of unpaired two-tailed t test.

We first analysed *DPPA3* mRNA levels by qRT-PCR in several CRC cell lines grown *in vitro* and generally observed very low levels with borderline Ct values in some cases. SW1222 and DLD1 CRC cells expressed the highest levels of *DPPA3* (Figure R2A). SW1222 retain an elevated capacity to differentiate to distinct intestinal cell lineages forming complex glandular structures when grown as three-dimensional (3D) structures [292]. We therefore selected this

cell line as an initial model presenting sufficient *DPPA3* expression and certain multipotency capacity. We knocked down its expression in SW1222 cells using three different shRNA molecules targeting *DPPA3* transcript. Due to the lack of specific antibodies against human DPPA3 protein we could not check the loss of *DPPA3* expression by Western blot. Alternatively, we checked its expression in our knockdown clones by qRT-PCR (Figure R2B). We also generated two different *DPPA3* knockout (KO) clones by means of CRISPR-Cas9 system using a guide RNA targeting a sequence located at the first exon of *DPPA3*. *DPPA3* KO clones were validated by Sanger sequencing given the impossibility to check KO efficiency by any technique that required the use of an antibody against the endogenous protein. Both *DPPA3* KO clones showed indels in the locus where the CRISPR guide oligo was directed, resulting in frameshift mutations in the first exon of *DPPA3* gene (Figure R2C). These mutations eventually leaded to a premature stop codon within the open reading frame (either in the first or second exon), resulting in the disruption of DPPA3 protein expression. Finally, aiming to mimic its elevated expression in SCCC, we generated SW1222 and HT29 CRC cell lines overexpressing DPPA3 (DPPA3-OE cells). We transfected the cells with a plasmid containing the complete *DPPA3* cDNA fused with a Flag tag and under the control of a constitutive promoter. Although being able to select SW1222 and HT29 CRC cells with geneticin after transfecting them with the DPPA3 plasmid, we were unable to visualize the Flag tag by Western blot. A plausible explanation for that would be that these cells were somehow depleting exogenous DPPA3 expression but still maintaining their resistance to geneticin. Therefore, DPPA3 may be deleterious and overexpressing cells were counter-selected due to unknown mechanisms. In order to overcome this limitation, we took advantage of an all-in-one lentivirus than was previously used in our laboratory for the doxycycline (DOX)-inducible expression of the histone 2B fused with the enhanced GFP [269]. In this case, we cloned the human *DPPA3* cDNA fused to a Flag tag right after the promoter containing a tetracycline responsive element (TRE). As this vector constitutively expressed mCherry fluorescent protein, we selected the infected cells by fluorescence-activated cell sorting (FACS). We decided to use as control cells infected with the empty lentiviral mCherry vector (without DPPA3-Flag) treated with DOX, mainly because two reasons: first, DOX impacts on cell phenotype (e.g. by slowing down their proliferative rate [293,294], so we could not distinguish between direct DPPA3-dependent or indirect side-effects. Moreover, we realized that our all-in-one lentivirus system had a low degree of leakage leading to a basal expression of DPPA3 that was high enough to be noticeable by qRT-PCR but not by Western blot. Despite of this small degree of leakage, we considered that it might have an impact on cells taking into consideration DPPA3 function: an epigenetic factor that acts genome-wide by binding histone marks. Therefore, DOX was added in all experiments performed in control and DPPA3-OE cells unless otherwise specified. For simplicity, DOX treatment is not commented. In order to

homogenise DPPA3 expression in our cultures, we generated clones from HT29 and SW1222 cell lines infected with the DPPA3-OE lentivirus. We evaluated its exogenous expression by Western blot and observed that HT29 cells expressed higher DPPA3 levels than SW1222 cells (Figure R2D).

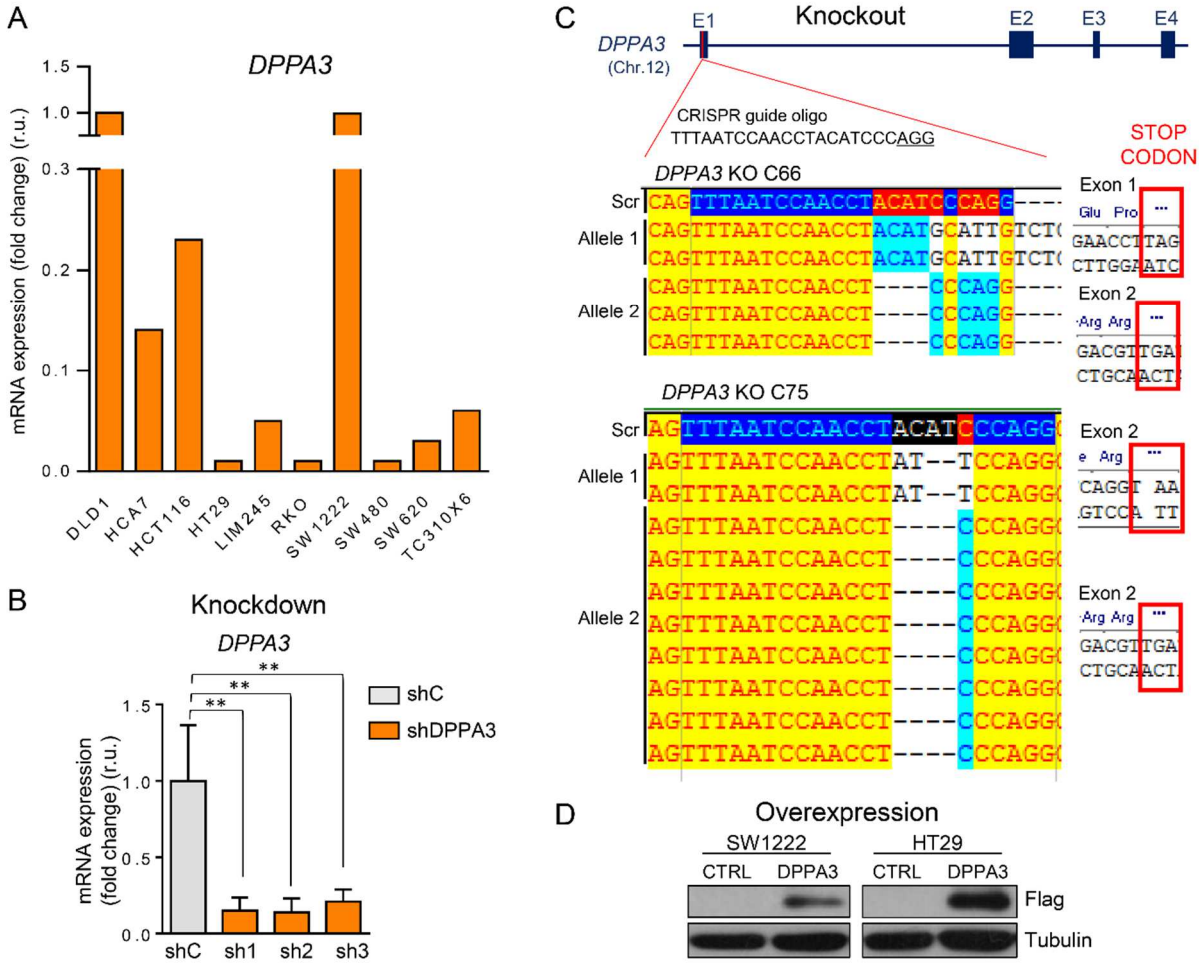


Figure R2. Generation of DPPA3 cancer cell models. (A, B) qRT-PCR showing *DPPA3* mRNA levels in (A) the different CRC cell lines indicated and (B) shC (grey) and shDPPA3 (sh1-3, orange) SW1222 CRC cell lines. Experiments were performed in triplicates, and data represent the average of three experiments ± SD. r.u., relative units. $^{**}P \leq 0.01$ of 1-way ANOVA. (A, B) Fold changes are relative to DLD1 in panel A and to shC in panel B. Expression levels were normalized to an endogenous control. (C) Diagram depicting CRISPR-Cas9 system-based strategy used for *DPPA3* knockout in SW1222 CRC cells. A *DPPA3*-specific CRISPR guide oligo (sequence shown) was designed to target the first exon of the gene. CRISPR/Cas9-derived mutations lead to a stop codon (***, right) either in the first or second exon of *DPPA3* in both alleles of the two *DPPA3* knockout clones (*DPPA3* KO C66 and C75) generated. (D) Western blot showing DPPA3 exogenous levels induced by DOX treatment in SW1222 and HT29 control (CTRL) and DPPA3-OE (DPPA3) cells using a Flag antibody. Tubulin was used as loading control.

Standard two-dimensional (2D) cell culture conditions fail to recapitulate the original tumour architecture or microenvironmental gradients and are not designed to retain the cellular heterogeneity of parental tumours. Three-dimensional (3D) cell cultures better recapitulate physiologic conditions due to the ability of cells to generate a heterogeneous structure able to self-organize and differentiate towards different cell lineages [14,295–298]. SW1222 cells exemplify this observation since they differentiate into the several cell lineages present in colon epithelia when grown in 3D as Matrigel-embedded tumour megacolonies [299]. In an attempt to understand the role of DPPA3 in the biology of cancer, we performed a microarray gene expression analysis using the different SW1222 CRC models generated (overexpression, knockdown and knockout) grown in 3D as tumour megacolonies. DPPA3 overexpression modulated gene expression programs related to various biological processes such as cell cycle, hypoxia and drug resistance (Figure R3A). Given that *DPPA3* was induced in SW1222 SCCC, we generated two gene sets with the upregulated and downregulated genes (FC≥1.3 and FC≤-1.3 respectively, p-value <0.05) from the Pan-Cancer SCCC signature, which was generated with the expression profile of SCCC isolated from CRC, melanoma and glioblastoma models [269]. Interestingly, Pan-Cancer SCCC upregulated genes were positively enriched in DPPA3-OE cells whereas those down-modulated in Pan-Cancer SCCC were negatively enriched (Figure R3B). These results indicated that DPPA3 could control key biological processes that define the identity of SCCC by regulating the expression of several sets of target genes.

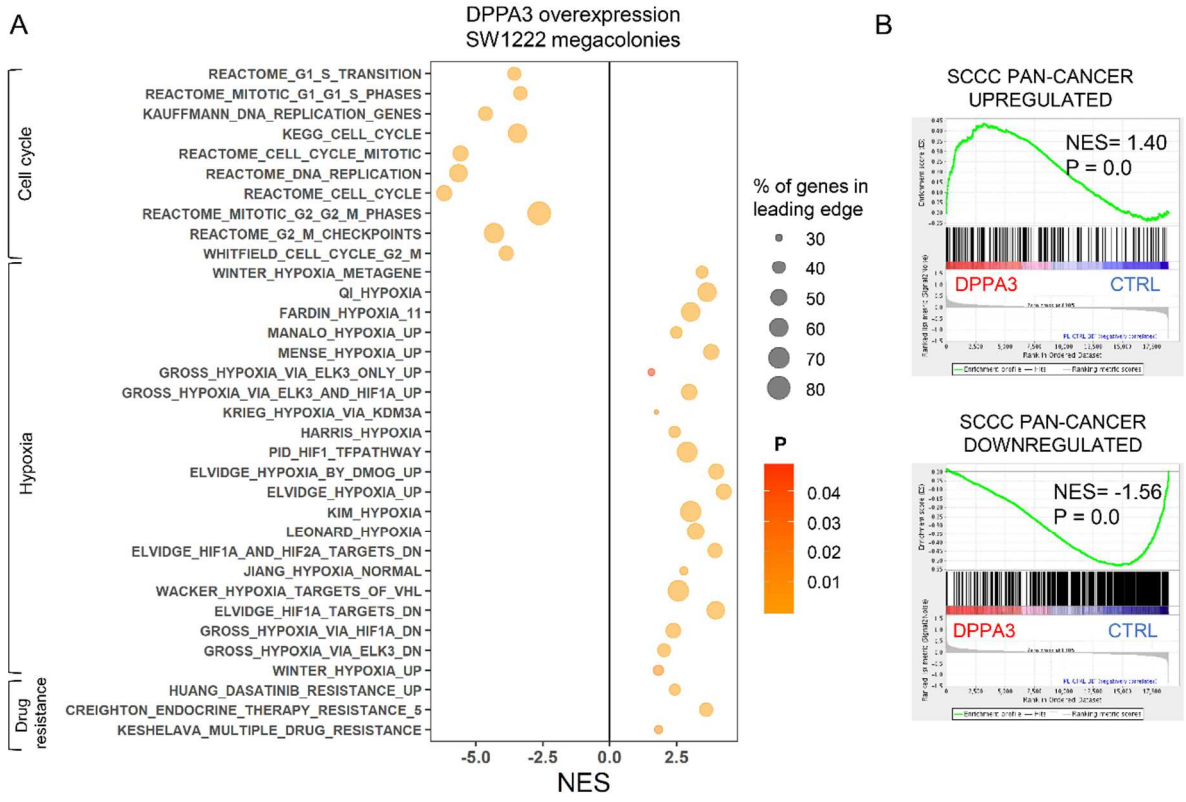


Figure R3. DPPA3 modulates different biological processes and promotes a transcriptional profile similar to that observed in SCCC. (A) Dot plot representing the normalized enrichment scores of cell cycle, hypoxia and drug resistance-related pathways enriched in DPPA3-OE SW1222 megacolonies compared to control cells determined from gene set enrichment analysis (GSEA) results based on curated gene sets. The size of the dot represents the percentage of genes in leading edge, and the colour represents the adjusted p-value, ranging from 0.01 to 0.04. (B) GSEA plots showing enrichment of the specified gene sets in the expression profiles of DPPA3-OE (DPPA3) versus control (CTRL) SW1222 megacolonies. (A, B) NES, normalized enrichment score; P, p-value.

Indeed, a closer look to the data evidenced that SW1222 SCCC and SW1222 DPPA3-OE cells showed a similar gene expression profile with an equivalent pattern in gene sets related to cell cycle, hypoxia and drug resistance (Figure R4).

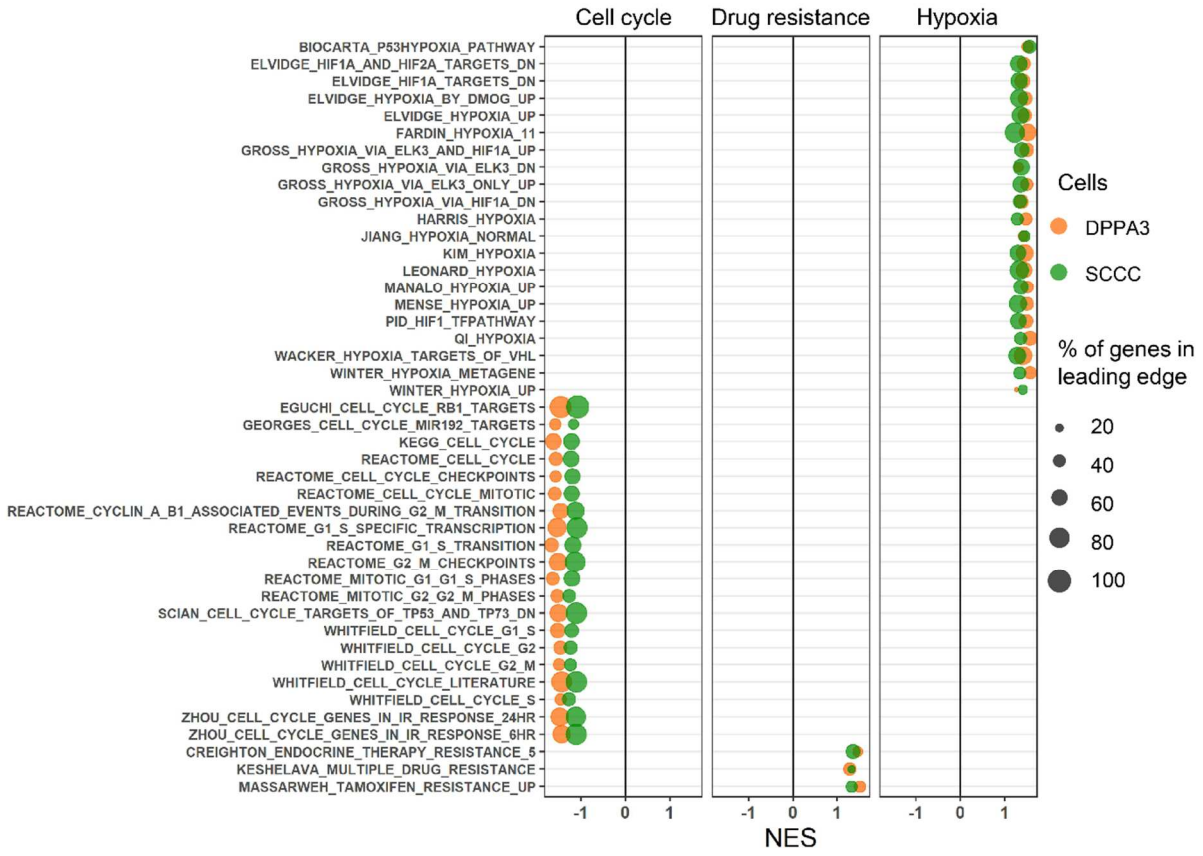


Figure R4. DPPA3-OE cells and SCCC share similar genetic programs related to cell cycle, chemoresistance and hypoxia. Dot plot representing the normalized enrichment score of hypoxia, cell cycle and drug resistance-related pathways enriched in DPPA3-OE SW1222 megacolonies compared to control cells (orange dots), and SW1222 SCCC compared to RCCC (green dots) determined from GSEA results based on curated gene sets. The size of the dot represents the percentage of genes in leading edge. NES, normalized enrichment score.

In addition, leading edge analysis revealed that DPPA3 overexpression was sufficient to affect the expression of many genes regulated in SW1222 SCCC (Figure R5A), confirming the importance of this epigenetic factor in SCCC.

Given the higher expression of genes related to the hypoxia response in SCCC of SW1222 megacolonies, we assessed the cellular response to changes in oxygen levels in these structures. Immunofluorescence assays performed with the hypoxia marker CA9 revealed that SCCC (GFP-positive cells) were in CA9-positive areas (Figure R5B). This result indicated that SCCC probably locate within regions devoid of high oxygen concentrations. In summary, our DPPA3-OE model mimicked the transcriptional program of SCCC, being therefore a suitable model to study the role of this epigenetic factor in SCCC.

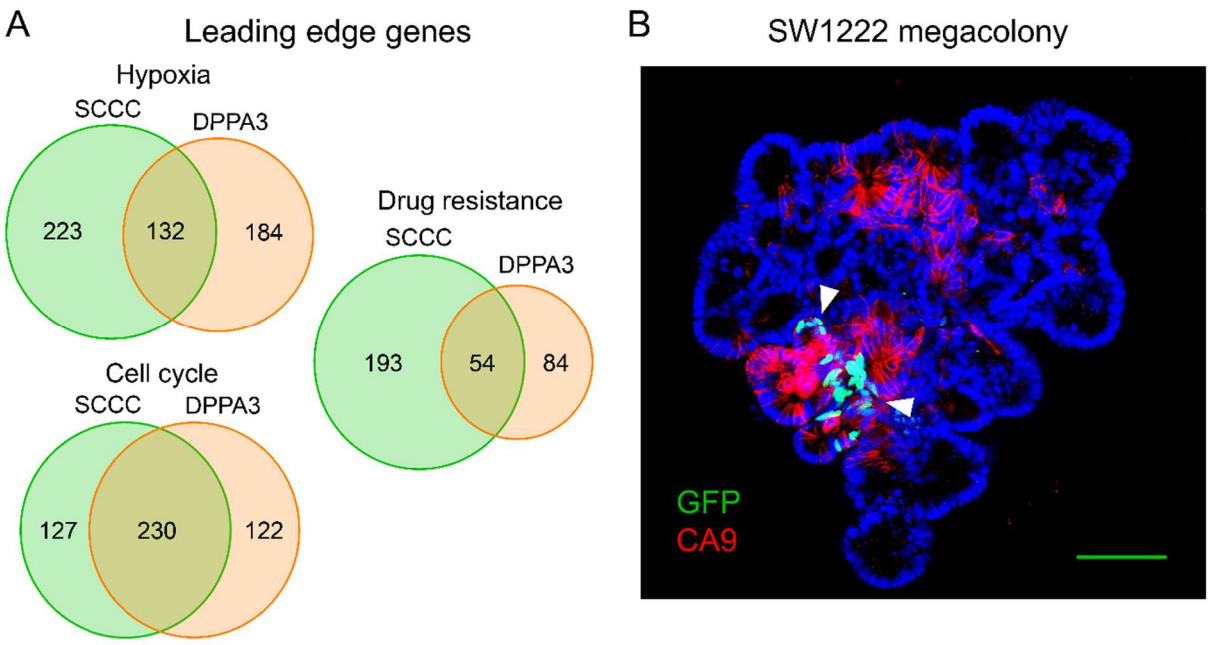


Figure R5. Common genes modulated in SCCC and DPPA3-OE cells related with hypoxia, drug resistance and cell cycle associated with the hypoxic microenvironment of SCCC. (A) Venn diagrams showing numbers of common leading edge genes between SCCC and DPPA-OE megacolonies obtained in GSEA analysis used in Figure R4. (B) Representative immunofluorescence image showing SCCC (arrowheads) labelled with an anti-GFP antibody and CA9 distribution in H2BeGFP-infected SW1222 megacolonies growing embedded in Matrigel. Scale bars: 100 µm.

2. DPPA3 regulates the hypoxia program

Gene expression analysis unmasked a possible effect of DPPA3 mediating the cellular response to hypoxia. Of note, previous data generated in our laboratory demonstrated that SCCC express genes related to this condition [269]. Given the importance of hypoxia in tumour biology, chemoresistance and patient survival, we decided to investigate the role of DPPA3 in this biological process.

For a comprehensive understanding of DPPA3 role in the cellular response to hypoxia, we used 2D *in vitro* assays. This approach allowed us to homogenise the effects mediated by changes in oxygen levels on the whole bulk of cultured cells, especially after noticing that only limited areas of tumour megacolonies presented a hypoxic microenvironment (Figure R5B). Results obtained in 2D models can be subsequently validated in more physiological conditions such as 3D *in vitro* assays or *in vivo* cancer models.

We first measured the expression of *DPPA3* mRNA in SW1222 CRC cells cultured in hypoxic (0.5% O_2) and normoxic (21% O_2) conditions. Increased transcripts of endogenous *DPPA3* were detected by qRT-PCR in hypoxia (Figure R6A). Of note, this increase also occurred in

DPPA3-OE cells (Figure R6B), in which exogenous *DPPA3* expression is solely under the control of an artificial inducible-promoter that has lost the endogenous transcriptional regulatory networks. To distinguish between endogenous and exogenous *DPPA3* mRNAs we used a primer pair that specifically bound the end of the *DPPA3* cDNA and the Flag sequence. We observed that hypoxia increased the levels of exogenous *DPPA3* (Figure R6C) , suggesting an mRNA stabilization mechanism rather than a promoter upregulation of gene expression. In fact, mRNA stabilization is a common mechanism for augmenting the expression of many genes such as *VEGFA* and *SLC2A1* in response to hypoxia [300–302]. Consequently, we observed that hypoxia promoted an increase of exogenous DPPA3 protein levels and, in particular, its chromatin-bound pool in SW1222 and HT29 cells (Figure R6D). This was a specially interesting result, since the biological effects of DPPA3 rely on its binding to H3K9me2 [256], an epigenetic mark previously described as increased by hypoxia [197,198]. In addition, hypoxia promotes a remodelling of chromatin structure and methylation of many histones [109,190]. We confirmed that hypoxia increased H3K9me2 mark in SW1222 cells (Figure R6E). Importantly, exogenous DPPA3 also promoted an accumulation of H3K9me2 in the same cells both in normoxic and hypoxic conditions.

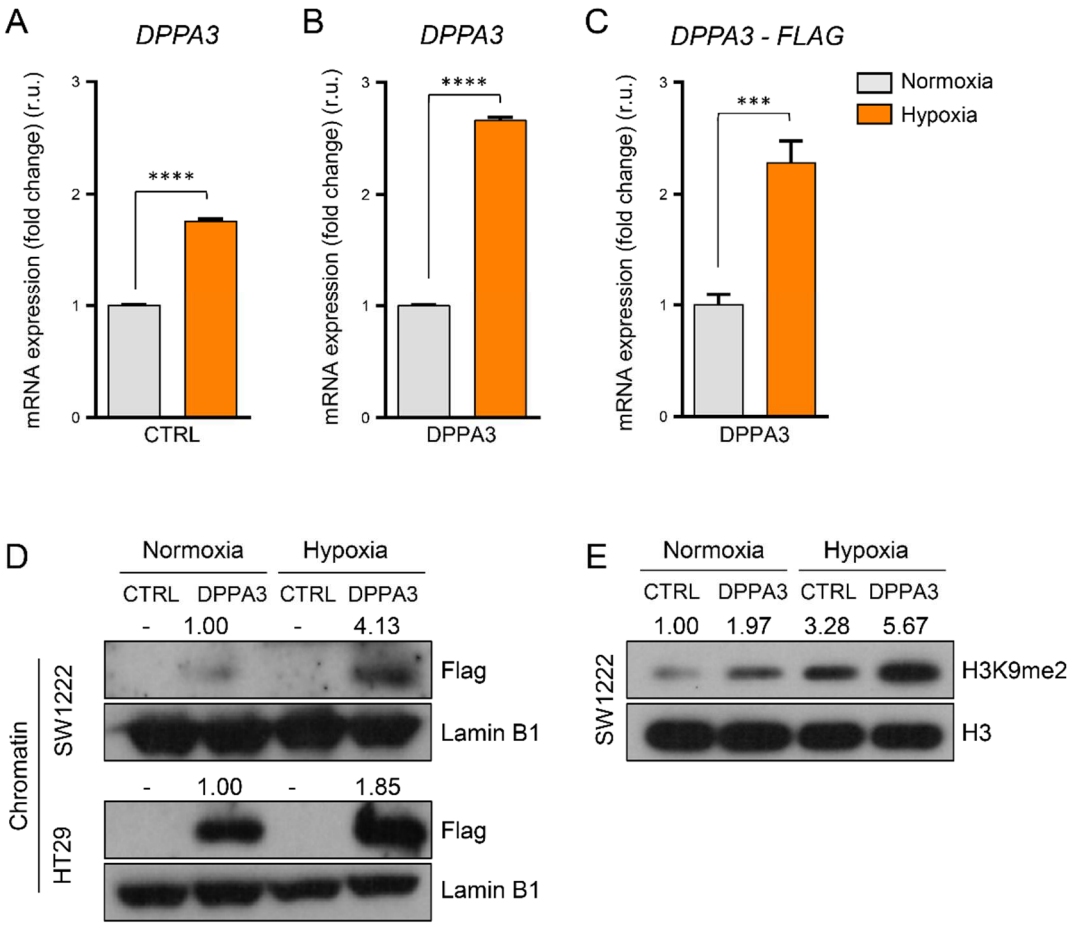


Figure R6. DPPA3 mRNA and protein levels are modulated in hypoxia together with H3K9me2. (A, B) qRT-PCR showing *DPPA3* mRNA levels of (A) control (CTRL) and (B) DPPA3-OE (DPPA3) SW1222 cells in normoxic or hypoxic conditions. (C) qRT-PCR showing exogenous *DPPA3* mRNA levels of DPPA3-OE SW1222 cells in normoxic or hypoxic conditions. (A-C) Grey bars, normoxic condition; orange bars, hypoxic condition. Expression levels were normalized to an endogenous control and expressed relative to normoxic condition. Experiments were performed in triplicates, and data represent the average of three experiments ± SD. r.u., relative units; ***$P \leq 0.001$; ****$P \leq 0.0001$ of unpaired two-tailed t test. (D) Western blot showing exogenous protein levels of DOX-induced DPPA3 in the chromatin fraction of control and DPPA3-OE SW1222 and HT29 cells in normoxic and hypoxic conditions using a Flag antibody. Lamin B1 was used as loading control. (E) Western blot showing H3K9me2 levels of normoxic and hypoxic control (CTRL) and DPPA3-OE (DPPA3) SW1222 cells. Histone H3 was used as loading control. (D, E) Numbers show quantification by densitometry of DPPA3-Flag and H3K9me2 protein levels, which were normalized to lamin B1 or histone H3 and expressed relative to control normoxic cells.

Hypoxia blocks mTOR signalling that is central for protein synthesis but a very expensive process from the energetic point of view. Hypoxic cells activate alternative processes to sustain protein synthesis and cell survival [110]. We confirmed that hypoxia decreased the levels of phosphorylation of S6 ribosomal protein (Figure R7A), a canonical target of mTORC1 activity [169]. Similarly, DPPA3 overexpression also decreased p-S6 levels in normoxia. It has been described that AMPK is a major sensor of cellular energy levels that gets activated upon energy imbalance and down-modulates the mTOR pathway [176,177,303] (Figure R7B). DPPA3 overexpression upregulated *PRKAA1*, the gene encoding for the alpha 1 subunit of AMPK (Figure R7C). Active mTOR allows proliferation and cell growth under normoxic conditions in part through the inactivation of 4E-BP1, an inhibitor of the complex necessary to initiate cap-dependent translation [169] (Figure R7B). A study using breast cancer cells demonstrated that hypoxia orchestrates a switch from cap-dependent to cap-independent translation through the cooperation between 4E-BP1 and EIF4G1. In this scenario, 4E-BP1 allows a hypoxia-mediated inhibition of cap-dependent translation and, together with EIF4G1, increases the selective translation of IRES-containing mRNAs that include key proangiogenic, hypoxia and survival mRNAs [187]. EIF4G2 is another initiation factor of cap-independent translation that has been reported to mediate IRES-dependent translation in human ESCs [188] and translation of specific mRNAs in quiescent cells and immature oocytes [189]. Evaluation of 4E-BP1, EIF4G1 and EIF4G2 expression (*EIF4EBP1*, *EIF4G1* and *EIF4G2* genes, respectively) by qRT-PCR in SW1222 tumour xenografts evidenced an upregulation of *EIF4G1* and *EIF4G2* by DPPA3 (Figure R7C). Altogether, these data suggested that DPPA3 overexpression in CRC cells results in a switch from cap-dependent to cap-independent translation as a consequence of blocking mTOR activity, a phenomenon commonly observed in hypoxia.

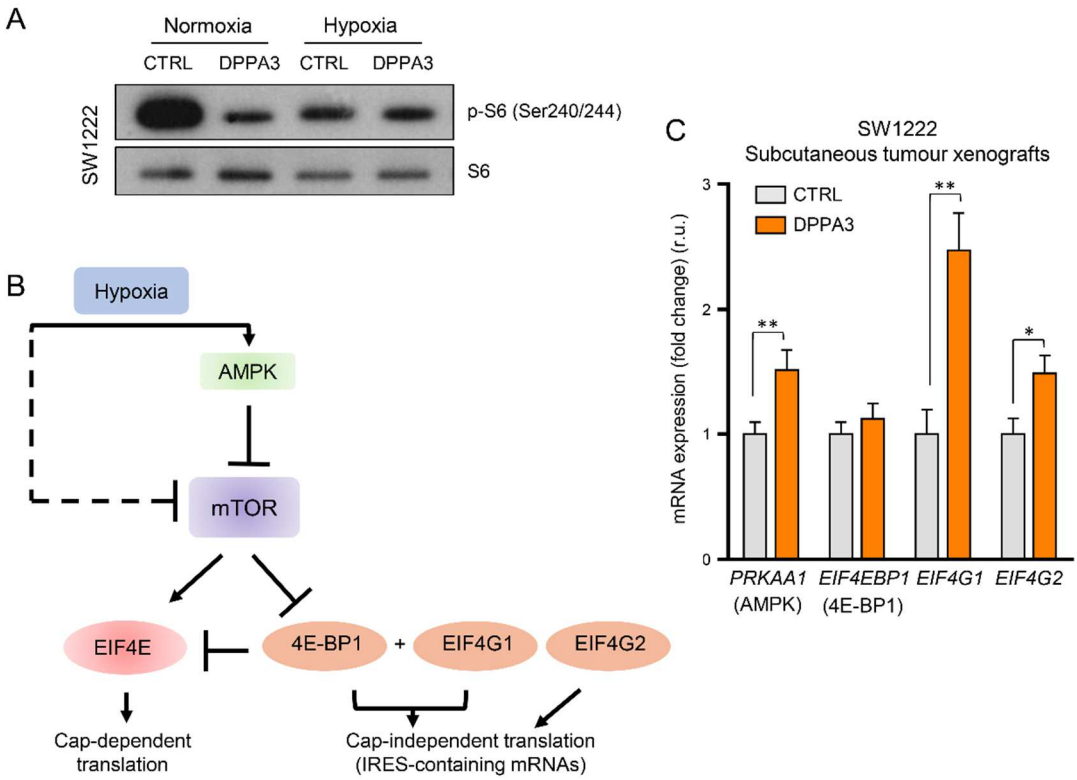


Figure R7. mTOR signalling is blocked in DPPA3-OE cells. (A) Western blot showing p-S6 (Ser240/244) levels in DPPA3-OE (DPPA3) and control (CTRL) SW1222 cells in normoxic and hypoxic conditions. Total S6 was used as loading control. (B) Diagram depicting the effects of hypoxia on AMPK and the mTOR pathway activity and downstream effects. AMPK energy sensor, which inhibits mTOR, is activated under hypoxic conditions. In addition, hypoxia downmodulates mTORC activity through other mechanisms, leading to increased hypophosphorylated (active) 4E-BP1 and repression of the cap-dependent translation mediated by EIF4E. 4E-BP1 together with EIF4G1 have been shown to mediate a hypoxia-activated switch to facilitate cap-independent mRNA translation. EIF4G2 is another factor that collaborates in cap-independent translation mechanisms. (C) Expression of the indicated genes was evaluated by qRT-PCR in control (CTRL) and DPPA3-OE (DPPA3) SW1222 subcutaneous xenografts. Expression levels were normalized to an endogenous control and expressed relative to control xenografts. Three tumours per condition were used, and data represent the average ± SD. r.u., relative units; $*P \leq 0.05$; $**P \leq 0.01$ of unpaired two-tailed t test.

3. DPPA3 promotes HIF1α protein stability and activity

Since HIF1α is a central effector of hypoxia signalling cascade we evaluated its protein levels upon DPPA3 overexpression. DPPA3 increased HIF1α protein levels in normoxic SW1222 and HT29 cells (Figure R8A). While this increase augmented in hypoxic SW1222 cells, the contrary was observed in HT29 cells. As a transcriptional factor, HIF1α recognizes and binds consensus HRE sequences to activate transcription of target genes [304]. For this reason, we confirmed that the increase of total HIF1α took place in the chromatin (Figure R8B).

Importantly, although HIF1α protein levels were not affected by DPPA3 depletion in shDPPA3 cells, CA9, a direct HIF1α target gene, was less expressed in knockdown than in control cells (Figure R8C). These results indicated that DPPA3 was necessary for the complete activation of the hypoxia program.

We also found that the increase of HIF1α protein levels by DPPA3 overexpression was not due to an induction of *HIF1A* gene expression since its mRNA levels were equivalent in DPPA3-OE and control cells (Figure R8D). However, whereas hypoxia reduced the expression of *HIF1A* mRNA in control cells, this effect did not occur in DPPA3-OE cells.

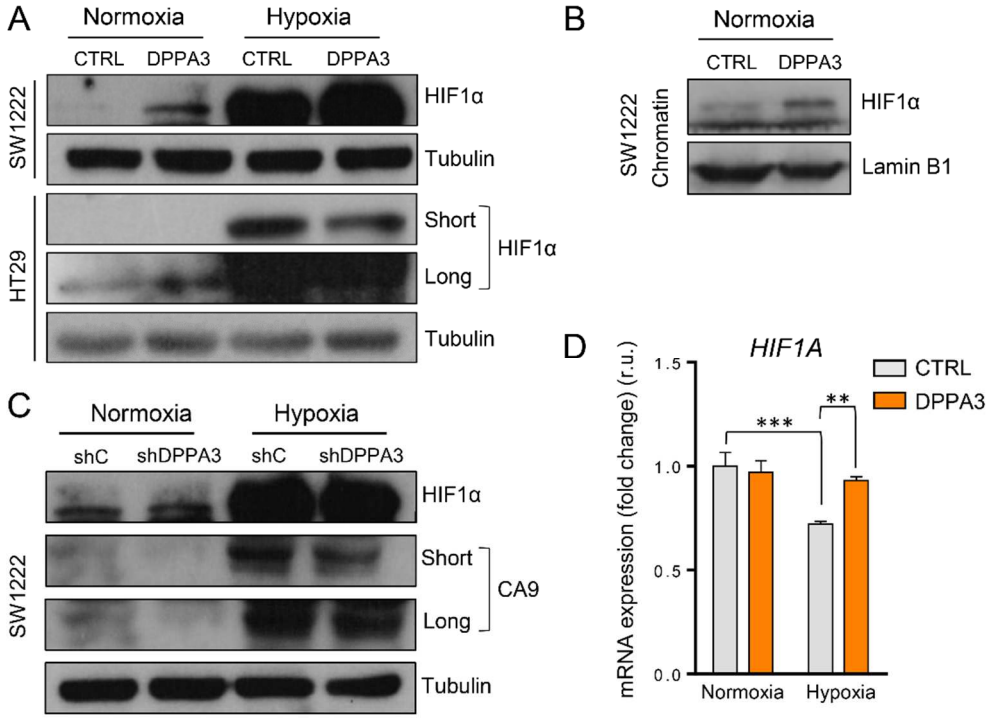


Figure R8. DPPA3 overexpression modulates HIF1α protein levels. (A) Western blot showing HIF1α total levels in control (CTRL) and DPPA3-OE (DPPA3) SW1222 and HT29 cells in normoxic and hypoxic conditions. (B) Western blot showing HIF1α levels in the chromatin fraction of control and DPPA3-OE normoxic SW1222 cells. Lamin B1 was used as loading control. (C) Western blot showing HIF1α and CA9 total levels in shC and shDPPA3 SW1222 cells in normoxic and hypoxic conditions. shDPPA3 clone 2 (sh2) cell line was used in this analysis. (D) qRT-PCR showing *HIF1A* mRNA levels in control (CTRL) and DPPA3-OE (DPPA3) normoxic or hypoxic SW1222 cells. Expression levels were normalized to an endogenous control and expressed relative to control normoxic cells. Experiments were performed in triplicates, and data represent the average of three experiments ± SD. r.u., relative units; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$ of 2-way ANOVA test. (A, C) Tubulin was used as loading control. For better visualization of results, short and long exposures of the autoradiography film are shown.

In summary, *DPPA3* mRNA stabilization in hypoxia highlights its importance in this biological process. Nuclear accumulation of HIF1α by DPPA3 may account for the activation of the hypoxia transcriptional response. In addition, upregulation of H3K9me2 levels together with a reduced mTOR activity might contribute to a hypoxic phenotype in cells overexpressing DPPA3 epigenetic factor. At this point we had confirmed that DPPA3 could modulate the hypoxia cascade upstream of HIF1α. Our next step was to dissect the molecular mechanisms driven by DPPA3 in this process.

One of the major pathways involved in the modulation of HIF1α is at the post-translational level: under conditions where oxygen is available, HIF1α gets hydroxylated and subsequently marked for ubiquitination to be sent to the proteasome [110]. It has been demonstrated that prompt ubiquitination of HIF1α occurs in the nuclear compartment upon reoxygenation (1 minute in normoxia) of hypoxic cells. In this situation, HIF1α is rapidly degraded with the concomitant appearance of higher molecular weight bands reminiscent of the ubiquitinated forms of HIF1α (Ub-HIF1α) [305]. Since ubiquitination represents a pivotal mechanism regulating HIF1α availability and given that DPPA3 modulates HIF1α protein levels, we sought to determine whether DPPA3 affected HIF1α stability through the ubiquitin-proteasome pathway. With that aim, we checked Ub-HIF1α in hypoxic SW1222 cells reoxygenated for one minute prior protein extraction. Only in this situation and not in normoxia we could detect HIF1α post-translational modifications. Western blot analysis revealed that chromatin-bound Ub-HIF1α (slower-migrating HIF1α) in reoxygenated cells dropped upon DPPA3 overexpression in both CRC cell lines, SW1222 and HT29 (Figure R9A). This result indicated that DPPA3 affected the HIF1α-degradation circuit, leading us to study the status of key regulators of HIF1α ubiquitination. Prolyl hydroxylases (PHD1-3) mark HIF1α for ubiquitination under conditions of oxygen availability, being PHD2 (*EGLN1* gene) the most abundant in high oxygen conditions [140]. *EGLN1* promoter contains a HRE and its expression has been reported to increase in hypoxia as a negative feedback loop mechanism [306]. While overexpression of DPPA3 induced *EGLN1* expression in normoxic cells (Figure R9B), hypoxia accentuated this effect, alluding to a negative feedback loop initiated as a desensitization mechanism to hypoxia that was strengthened by DPPA3.

In light of these results, we decided to broaden our search by checking the expression of other components required for HIF1α degradation. In the process of HIF1α ubiquitination, VHL forms a multiprotein complex with ELOC, ELOB, CUL2 and RBX1 [141–143] (Figure R9C). CUL2 acts as a scaffold that brings together HIF1α and an E2-conjugating enzyme and catalyses the transfer of ubiquitins from the E2-conjugating enzyme to HIF1α [145–147] while UBE2D1 is an E2-conjugating enzyme that has been shown to mediate HIF1α degradation [144]. *VHL* and *ELOC* expression increased in hypoxic DPPA3-OE cells, possibly again as a negative feedback loop

mechanism to compensate the exacerbated hypoxic response that might take place in DPPA3-OE cells (Figure R9D). On the other hand, the expression of *CUL2* and *UBE2D1* decreased in DPPA3-OE cells regardless of oxygen levels. Considering the importance of both CUL2 and UBE2D1 within the HIF1α degradation complex, it is reasonable to think that their down-modulation might exert profound effects on HIF1α stability.

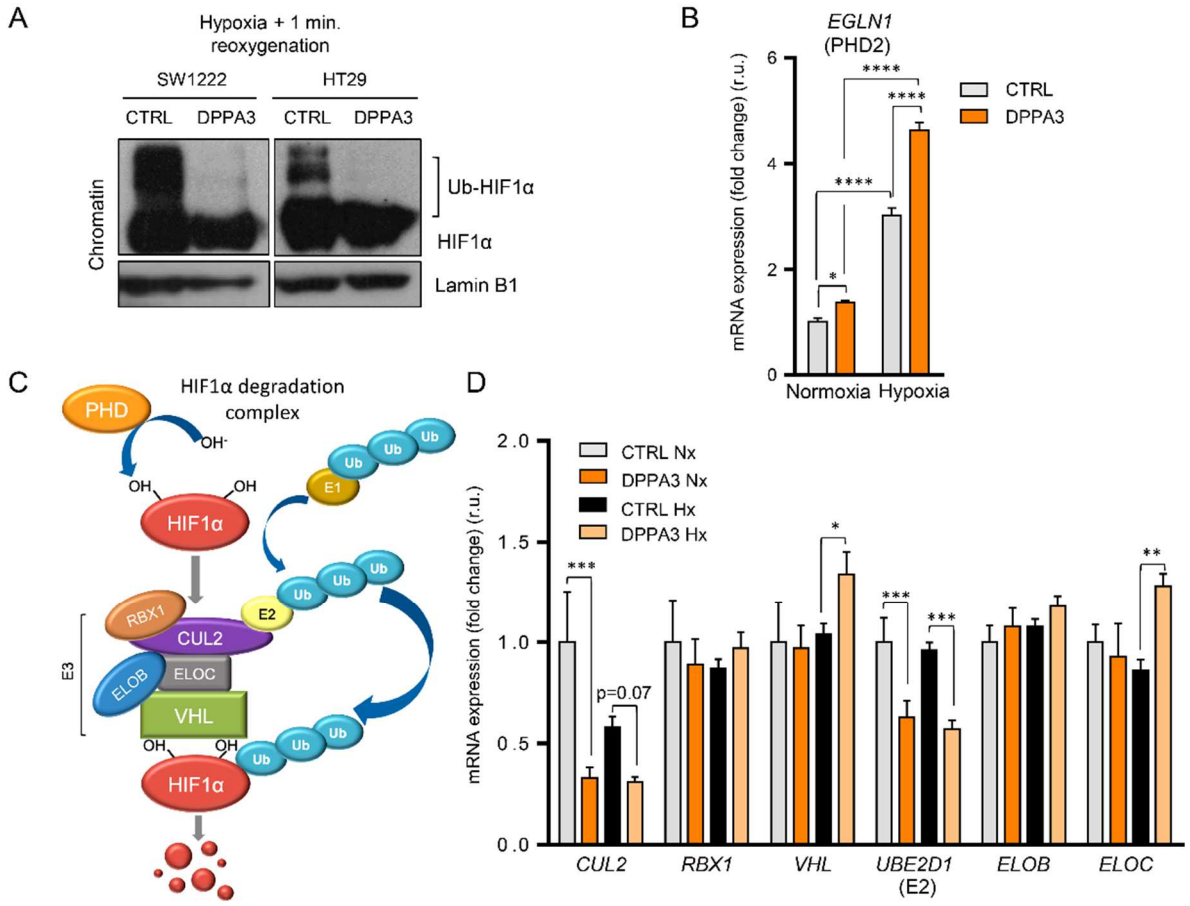


Figure R9. DPPA3 down-modulates HIF1α ubiquitination. (A) Western blot showing protein levels of HIF1α and its ubiquitinated forms (Ub-HIF1α) in the chromatin fraction of control (CTRL) and DPPA3-OE (DPPA3) SW1222 and HT29 cells in normoxic and hypoxic conditions. Lamin B1 was used as loading control. (B) qRT-PCR showing *EGLN1* mRNA levels in control (CTRL, grey) and DPPA3-OE (DPPA3, orange) normoxic and hypoxic SW1222 cells. (C) Diagram representing components (PHD, E1, E2 and E3) necessary for the hydroxylation and ubiquitination of HIF1α in normoxia. (D) Expression of the indicated genes was evaluated by qRT-PCR in control (CTRL) and DPPA3-OE (DPPA3) normoxic (Nx, grey and dark orange bars) and hypoxic (Hx, black and light orange bars) SW1222 cells. (B, D) Expression levels were normalized to an endogenous control and expressed relative to control normoxic cells. Experiments were performed in triplicates, and data represent the average of three experiments ± SD. r.u., relative units; p, p-value; **P* ≤ 0.05; ***P* ≤ 0.01; ****P* ≤ 0.001; *****P* ≤ 0.0001 of 2-way ANOVA test.

In conclusion, DPPA3 might stabilize HIF1α at the protein level by reducing the expression of two main components of the ubiquitin-degradation complex, CUL2 and UBE2D1, eventually decreasing its activity.

As a consequence of HIF1α upregulation by DPPA3, we expected an increased expression of HIF1α/hypoxia target genes. Indeed, we detected by qRT-PCR a higher expression of *LDHA*, *PDK1, SLC2A1, CA9* and *VEGFA* mRNAs in DPPA3-OE cells (Figure R10A). *LDHA* and *PDK1* were upregulated in normoxic and hypoxic conditions, whereas the *SLC2A1, CA9* and *VEGFA* were solely induced in hypoxia, suggesting a possible cooperation between DPPA3 and HIF1α. *CA9* is a direct HIF1α target gene that presents a HRE [122]. We assessed *CA9* promoter activity in hypoxic HEK293T cells transiently overexpressing DPPA3 using a luciferase reporter construct. We tested two different constructs, one with a wild-type (WT) HRE sequence (WT-CA9promoter-pGL3) and another with a point mutation in it (HREmut-CA9promoter-pGL3). Mutations at the HRE sequence of *CA9* promoter have already been reported to abolish promoter's activity [136,281]. We confirmed this observation as the mutant construct did not induce luciferase expression in control hypoxic cells (Figure R10B). Interestingly, we found that DPPA3 overexpression significantly increased *CA9* promoter activity. This induction was completely ablated when the promoter sequence was mutated in its HRE, indicating that DPPA3 can regulate the expression of hypoxic target genes specifically through HIF1α transcriptional activity. Altogether, these results suggested that DPPA3 participates in the cellular response to hypoxia upstream of HIF1α, increasing HIF1α protein levels and enhancing its transcriptional activity.

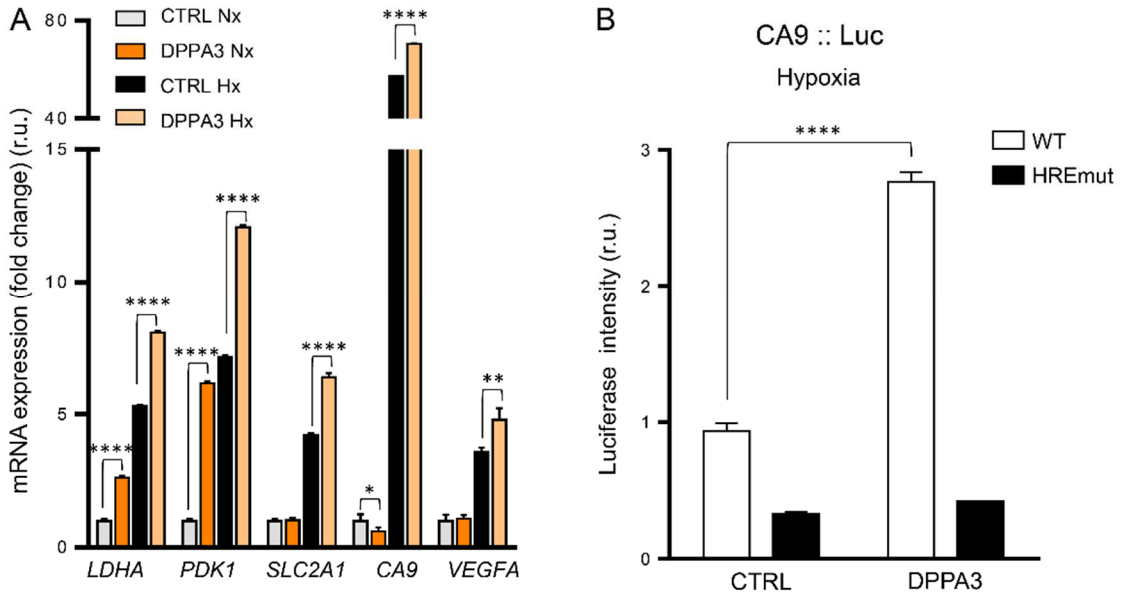


Figure R10. DPPA3 increases HIF1α transcriptional activity. (A) Expression of the indicated genes was evaluated by qRT-PCR in control and DPPA3-OE normoxic (grey and dark orange bars, respectively) and control and DPPA3-OE hypoxic (black and light orange bars, respectively) SW1222 cells. Expression levels were normalized to an endogenous control and expressed relative to control normoxic cells. (B) Luciferase reporter assay to assess *CA9* promoter activity on HEK293T cells co-transfected with either pCMV6-Entry (CTRL) or pCMV6-DPPA3 (DPPA3) plasmids together with the wild-type or HREmut-CA9promoter-pGL3 constructs (WT and HREmut, respectively). 48 hours after transfection, cells underwent hypoxia treatment (0.5% O_2) for 24h and luciferase activity was assessed. Values were expressed relative to cells co-transfected with the pCMV6-Entry and WT-CA9promoter-pGL3 constructs. (A, B) Experiments were performed in triplicates, and data represent the average of three experiments ± SD. r.u., relative units; **P* ≤ 0.05; ***P* ≤ 0.01; *****P* ≤ 0.0001 of 2-way ANOVA test.

4. DPPA3 causes a genome-wide demethylation

It has been shown that hypoxia can induce the loss of global DNA methylation in some types of cancers such as CRC and melanoma [204]. Moreover, binding of HIF1α to its target sequences can be determined by their methylation status in some gene promoters, including *HIF1A* and *CA9* [134–136,161]. Of note, DPPA3 overexpression in somatic cells causes genome-wide demethylation [231,262,307,308]. Bearing this information in mind, we hypothesized that DPPA3 could mediate, at least in part, the genome de-methylation induced by the hypoxia program. To demonstrate this hypothesis, we analysed the effect of DPPA3 activity on genome methylation status. We first confirmed that hypoxia induced a reduction in global 5mC levels [204] in SW1222 cells by dot blot analysis (Figure R11A). In addition, and in line with previous publications [231,262,307,308], we observed that DPPA3 overexpression decreased global 5mC

levels in normoxic conditions and caused a modest decrease in hypoxia. In order to evaluate the changes in the DNA methylation pattern mediated by DPPA3 across genome, we performed a methylome analysis in normoxic SW1222 cells (Figure R11B). Methylation of individual loci was determined by average β values that ranges from 0.00 (unmethylated) to 1.00 (completely methylated). Biological replicates of DPPA3-OE and control cells showed an elevated degree of similarity, confirming the technical robustness of the technology and the high specificity of DPPA3-induced methylation patterns. Subsequent data analysis evidenced a bimodal distribution of the CG dinucleotide methylation with a low-methylation peak that was found at β values ranging from 0.00 to 0.20 and a high-methylation peak that covered the interval from 0.80 to 1.00. As in control cells, the bimodal methylation distribution was also observed in DPPA3-OE cells. Nevertheless, the high-methylation peak corresponding to DPPA3-OE cells (β ≥ 0.80 in control cells) was shifted to lower methylation values, indicating a DPPA3-induced demethylation and confirming the dot blot results. Interestingly, most of the differences in methylation occurred in inter-genic and gene body regions (Figure R11C). Specifically, a 46% of the hypomethylation caused by DPPA3-OE occurred at gene bodies and a 25% at inter-genic regions. The rest (24%) occurred in regions that extended from 1500 bp upstream until the transcription start site (TSS), plus a 5% of overlapping probes (probes that overlapped different transcribed regions). Despite DPPA3 overexpression caused a major hypomethylation, there were some hypermethylated sequences compared to control cells. In this case, 54% of hypermethylation occurred at inter-genic regions, while a 19% occurred in gene bodies and a 27% in regions extending from 1500 bp upstream until the TSS. These data demonstrate that DPPA3 induces a global genome demethylation in CRC cells, suggesting a molecular demethylation mechanism similar to that already described in oocytes and somatic cells [231,262,307–309].

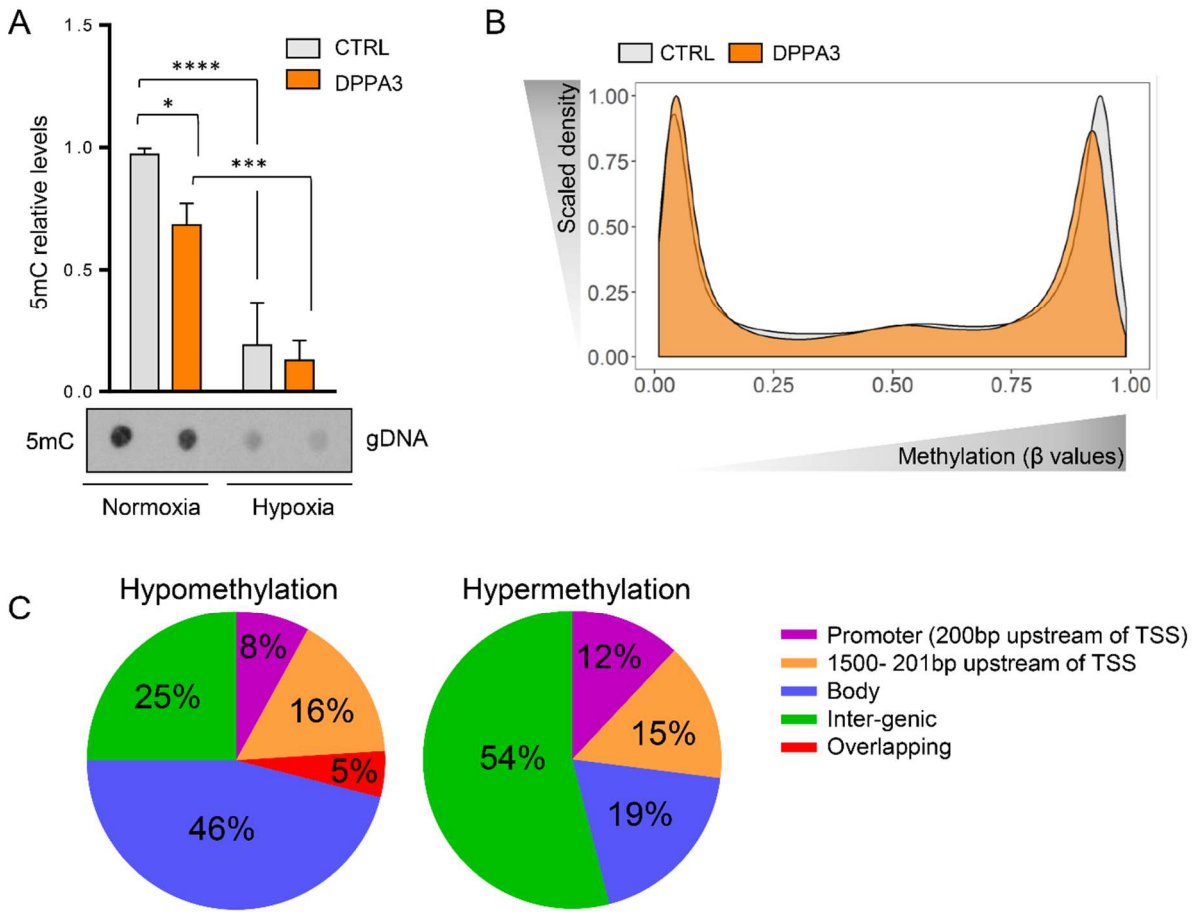


Figure R11. DPPA3 causes a genome-wide hypomethylation. (A) Dot blot analysis showing 5mC levels of genomic DNA (gDNA) extracted from control (CTRL, grey) and DPPA3-OE (DPPA3, orange) normoxic and hypoxic SW1222 cells. Experiments were performed in triplicates, and data represent the average of three experiments ± SD. **P* ≤ 0.05; ****P* ≤ 0.001; *****P* ≤ 0.0001 of 2-way ANOVA test. (B) Density plot showing scaled density of DNA methylation levels (as β values) for control (grey) and DPPA3-OE (orange) SW1222 cells. (C) Pie charts showing percentages of differentially methylated probes across different genomic regions, as indicated, in DPPA3-OE versus control SW1222 cells. Overlapping probes correspond to those that overlapped different transcribed regions.

Demethylating activity of TET enzymes on DNA has been reported to be involved in the hypoxia response, being enhanced or repressed depending on the cellular model studied [206,310]. As *TET1* expression and enzymatic activity has been shown to be higher in neuroblastoma cells under hypoxic conditions [310], we wanted to assess whether the lower 5mC levels observed upon DPPA3 expression and in hypoxia were due to active demethylation mechanisms. We observed by qRT-PCR analysis that hypoxia induced *TET1* and *TET2* expression, while DPPA3 overexpression down-modulated the hypoxia-mediated induction of *TET1* and diminished *TET2* expression in normoxia (Figure R12A). Interestingly, we observed that 5hmC and 5mC exhibited a similar pattern: DPPA3 slightly decreased 5hmC levels while hypoxia magnified these effects (Figure R12B). Increased *TET1* and *TET2*

expression in hypoxic conditions was opposite to what could be expected of 5hmC levels. These results suggested that an active and TET-mediated demethylation was not the main mechanism promoting the reduction observed in 5mC levels upon DPPA3 overexpression or oxygen starvation.

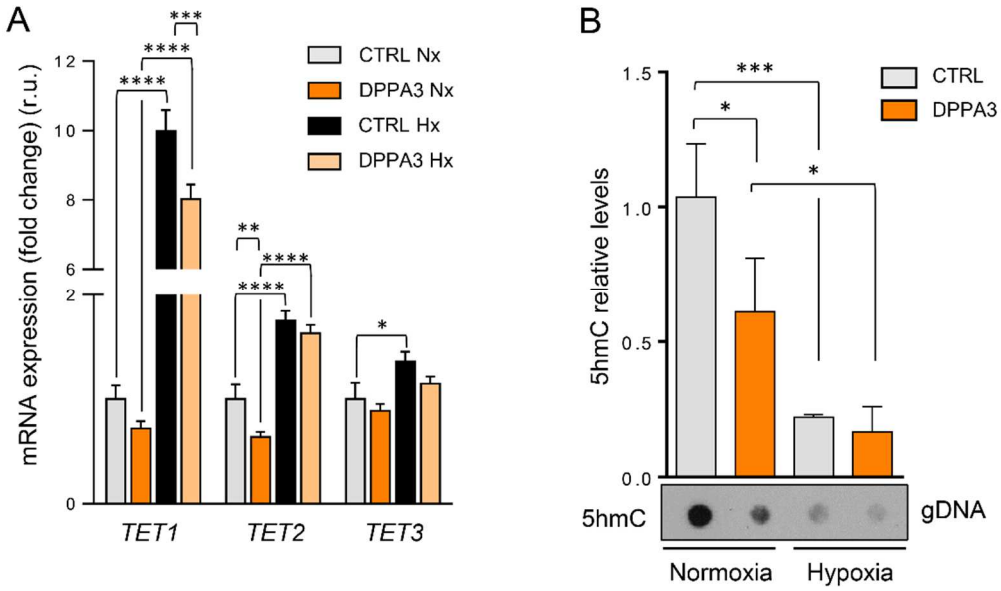


Figure R12. Genome hypomethylation upon DPPA3 overexpression is not caused by active demethylation processes. (A) qRT-PCR showing *TET1*, *TET2* and *TET3* mRNA levels in control (CTRL) and DPPA3-OE (DPPA3) normoxic (grey and dark orange bars, respectively) and control and DPPA3-OE hypoxic (black and light orange bars, respectively) SW1222 cells. Expression levels were normalized to an endogenous control and expressed relative to control normoxic cells. r.u., relative units. (B) Dot blot analysis showing 5hmC levels of gDNA extracted from control and DPPA3-OE normoxic and hypoxic SW1222 cells. (A, B) Experiments were performed in triplicates, and data represent the average of three experiments ± SD. **P* ≤ 0.05; ***P* ≤ 0.01; ****P* ≤ 0.001; *****P* ≤ 0.0001 of 2-way ANOVA test.

DNMT1 is a maintenance DNA methyltransferase that requires UHRF1 for its recruitment to hemimethylated DNA during DNA replication [261]. It has been shown in different cell models including mouse ESCs, oocytes and human somatic cells that high levels of DPPA3 prevent DNMT1 from effectively maintaining DNA methylation by disrupting UHRF1-mediated recruitment of DNMT1 to chromatin [231,262,307–309]. It has also been demonstrated that DPPA3 disrupts UHRF1 association with chromatin by directly interacting with the histone binding domain of UHRF1 and competing for the interaction between UHRF1 and the histone H3 tail [262] (Figure R13A). Given the global reduction in 5mC levels upon DPPA3 overexpression, we sought to demonstrate if this could be a passive demethylation phenomenon mediated by UHRF1 and DNMT1 interplay in cancer cells. We first checked total UHRF1 and DNMT1 protein levels in SW1222 cancer cells. Although not statistically significant, there was an

upwards trend in UHRF1 levels upon DPPA3 overexpression, which decreased in hypoxic conditions (Figure R13B). Of note, low UHRF1 levels in hypoxia have already been documented in hepatocellular carcinoma cells [311]. Hypoxia caused a drop in DNMT1 total protein levels, which has also been reported in certain ovarian cancer cell lines [312], while DPPA3 overexpression did not cause remarkable effects in both conditions.

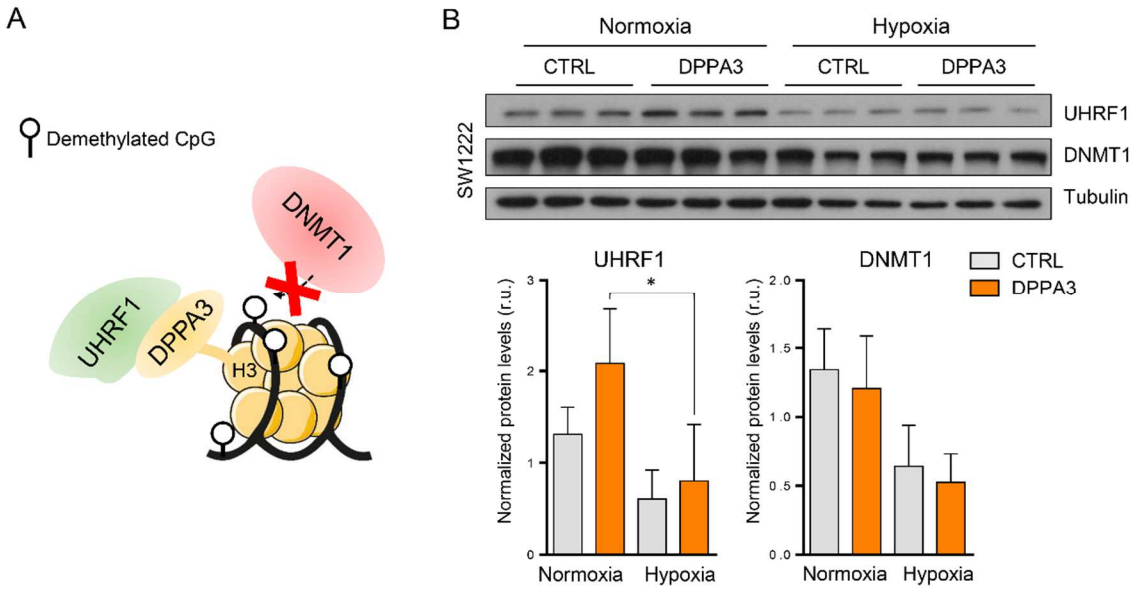


Figure R13. Effects of DPPA3 and hypoxia on total UHRF1 and DNMT1 protein levels. (A) Diagram depicting DPPA3-mediated DNMT1 exclusion from chromatin. DPPA3 recognizes H3K9me2 marks and binds UHRF1, which cannot longer recruit DNMT1. This situation leads to a global genome hypomethylation along cell divisions. (B) (Top) Western blots and quantification by densitometry (bottom) of total protein levels of UHRF1 and DNMT1 from control (CTRL) and DPPA3-OE (DPPA3) SW1222 cells cultured in normoxic and hypoxic conditions. Tubulin was used as loading control. UHRF1 and DNMT1 protein levels were normalized to tubulin and expressed relative to control normoxic cells. Experiments were performed in triplicates, and data represent the average of three experiments ± SD. r.u., relative units. *$P \leq 0.05$ of 2-way ANOVA test.

DNMT1 and UHRF1 bind chromatin to maintain DNA methylation across cell divisions. As we had only assessed total DNMT1 and UHRF1 total levels regardless of their cellular localisation, we considered important to check their distribution across cellular compartments. We observed that hypoxia lowered UHRF1 protein levels in chromatin and cytoplasmic fractions of SW1222 and HT29 cells (Figures R14A and C). DPPA3 overexpression did not affect the levels of chromatin-bound UHRF1. However, it remarkably increased the cytosolic fraction of UHRF1, thereby explaining the higher levels of total UHRF1 observed in DPPA3-OE cells (Figure R13B). In the case of HT29 cells, chromatin-bound DNMT1 was displaced to cytoplasm upon DPPA3 overexpression in both normoxic and hypoxic conditions. While the same situation occurred in hypoxic SW1222 cells, in normoxic conditions, the slight decrease

of DNMT1 in chromatin was due to its displacement to the nucleoplasm rather than cytoplasm (Figures R14A and B). *DPPA3* knockdown did not lead to major changes in cytoplasmic UHRF1 and DNMT1 (Figure R14C). However, we observed that, in hypoxia, this condition increased the levels of chromatin-bound DNMT1.

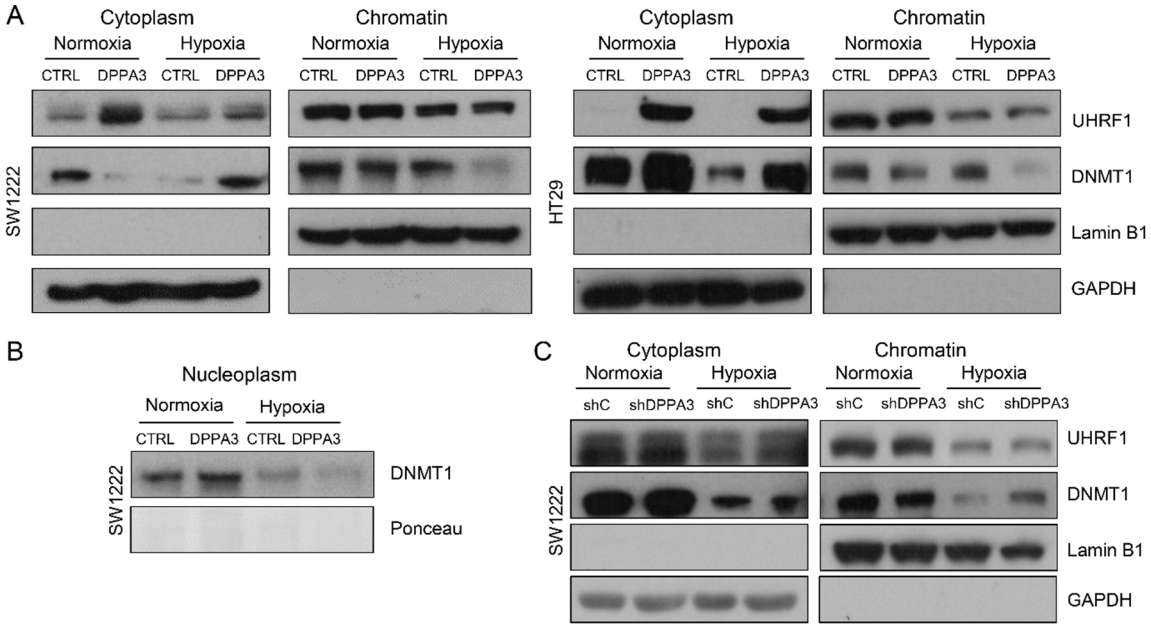


Figure R14. DPPA3 excludes DNMT1 from chromatin. Western blots showing protein levels of (A) UHRF1 and DNMT1 in the cytoplasm and chromatin fractions of control (CTRL) and DPPA3-OE (DPPA3) SW1222 and HT29 cells, (B) DNMT1 in the nucleoplasm of control and DPPA3-OE SW1222 cells, and (C) UHFR1 and DNMT1 in the cytoplasm and chromatin fractions of shC and shDPPA3 SW1222 cells cultured in normoxic and hypoxic conditions. shDPPA3 clone 2 (sh2) cell line was used in this analysis. (A, C) Lamin B1 and GAPDH were used as loading controls. (B) Ponceau S was used to stain the blot membrane as an alternative loading control.

In summary, these results indicate that the genome reduction of 5mC promoted by DPPA3 activity in CRC cells might be a passive demethylation process mediated by a DPPA3 displacement of DNMT1 from chromatin to cytoplasm.

5. DPPA3 accentuates the hypoxic phenotype in tumour xenografts

Physoxic levels range from 3 to 7.4%. Pathological hypoxia is a main characteristic of most tumours, where oxygen concentrations can drop to 0.3% [96,313]. Since *in vitro* models represent a first but not realistic approximation to what takes place in tumours, we moved *in vivo* to validate the effects exerted by DPPA3 in the hypoxic response. We performed subcutaneous injections of DPPA3-OE, shDPPA3 and their corresponding control SW1222 cells. While DPPA3 overexpression caused an increase in CA9 protein levels evaluated by immunohistochemistry (Figure R15A), its depletion did not show significant effects (Figure

R15B). We checked in tumour xenografts the expression of those HIF1α targets previously analysed *in vitro* (*LDHA*, *PDK1*, *SLC2A*, *CA9* and *VEGFA*) and observed that all of them were upregulated by DPPA3 (Figure R15C). As VEGFA contributes to tumour neovascularization [209], we checked by Western blot analysis the levels of the endothelial marker CD31, which increased in DPPA3-OE tumour xenografts (Figure R15D). These results suggested a promotion of tumour vascularisation, possibly through a DPPA3-mediated potentiation of the hypoxia program. Altogether, these data demonstrate that DPPA3 sensitizes tumour xenografts to low oxygen environments presumably by enhancing a hypoxia response.

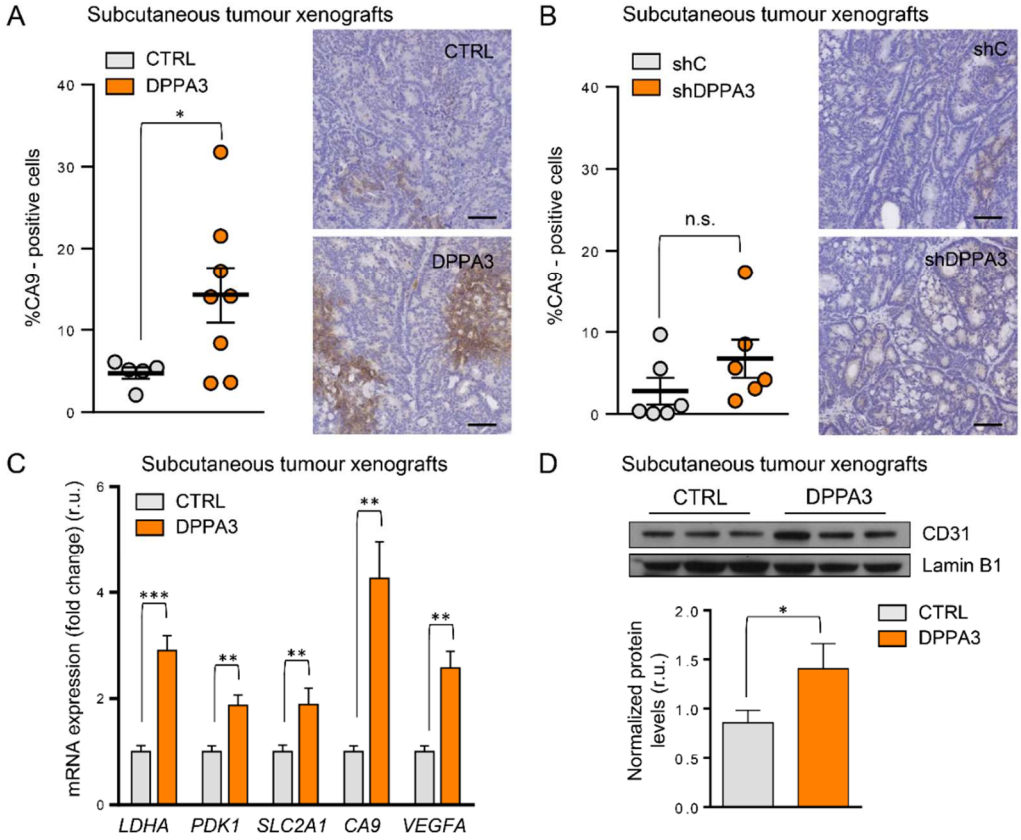


Figure R15. DPPA3 accentuates the hypoxic phenotype in tumour xenografts. (A, B) CA9 levels were evaluated by immunohistochemistry (IHC) in (A) DPPA3-OE (DPPA3) or (B) shDPPA3 SW1222 cells grown as subcutaneous xenografts (n = 5-8 tumours per condition) and their respective controls. Representative IHC images are shown. shDPPA3 clone 2 cell line was used in this analysis. Data are represented as mean ± SEM. n.s. not significant; **P* ≤ 0.05, of unpaired two-tailed t test. Scale bars: 100 µm. (C) Expression of the indicated genes was evaluated by qRT-PCR in control and DPPA3-OE SW1222 subcutaneous xenografts. Expression levels were normalized to an endogenous control and expressed relative to control xenografts. (D) (Top) Western blot showing CD31 total protein levels and (bottom) its quantification by densitometry from control and DPPA3-OE SW1222 subcutaneous xenografts. Lamin B1 was used as loading control. CD31 protein levels were normalized to lamin B1 and expressed relative to control tumours. (C, D) Three tumours per condition were used, and data represent the average ± SD. r.u., relative units; **P* ≤ 0.05, ***P* ≤ 0.01; ****P* ≤ 0.001 of unpaired two-tailed t test.

In light of these observations, we decided to explore the correlation between DPPA3 expression and the hypoxic response by analysing CA9 expression in human tumour samples. Firstly, we took advantage of gene expression data previously generated in our laboratory from a cohort of patient-derived xenografts (PDX) coming from 32 untreated stage II and III CRC primary tumours. After plotting their respective *DPPA3* normalized expression, we considered DPPA3-high those samples found in the upper quartile, while those in the lower quartile were considered DPPA3-low (Figure R16A). Immunohistochemistry analysis of CA9 revealed higher staining of this protein in DPPA3-high tumours than DPPA3-low specimens (Figure R16B) and a positive correlation between *DPPA3* expression and CA9 protein levels in these same samples was observed (Figure R16C). Additionally, we used the free open database CANCERTOOL to visualize any correlation between *DPPA3* and *CA9* expression in other types of cancers. We previously observed that prostate tumours presented higher *DPPA3* expression compared with normal tissue (Figure R1B). Interestingly, evidence of hypoxic phenotype in prostate cancer has been documented through detection of specific molecular markers [314]. Based on these notions, we decided to study the correlation between *DPPA3* and *CA9* expression in a cohort of 150 prostate cancer samples. Interestingly, there was a positive correlation between the expression of both genes in primary tumours and metastases (Figure R16D). Altogether, these results uncovered a possible yet unknown scenario in which DPPA3 might regulate the tumour response to hypoxia in cancer patients.

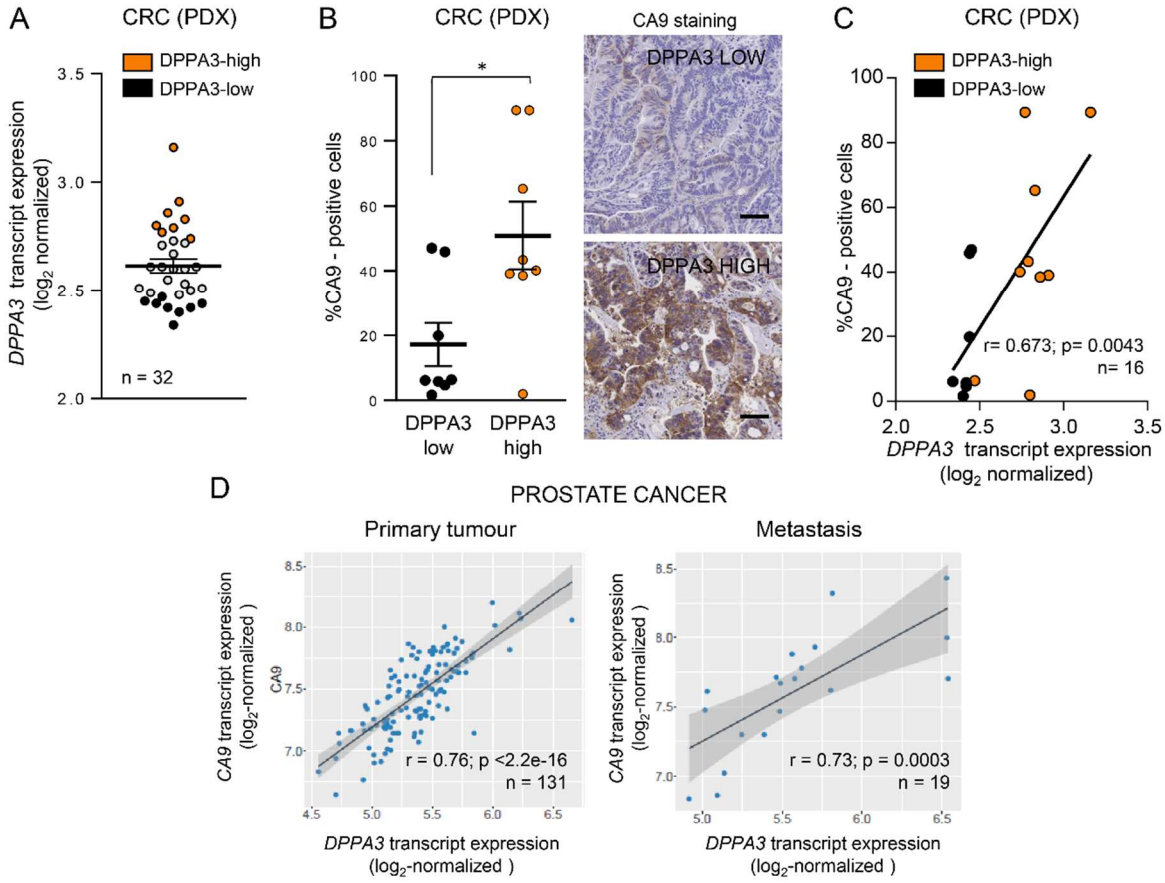


Figure R16. DPPA3 and CA9 expression correlate in colorectal and prostate tumours. (A) Stratification of 32 colorectal cancer (CRC) patient-derived xenograft (PDX) samples according to the quartiles of *DPPA3* transcript expression analysed by microarrays, being DPPA3-high those in the upper quartile (orange dots) and DPPA3-low those in the lower quartile (black dots). (B) (Left) Quantification and (right) representative images of CA9 levels evaluated by immunohistochemistry in DPPA3-high and low CRC PDX samples (n = 8 PDX per condition) used in panel A. **P* ≤ 0.05 of unpaired two-tailed t test. Scale bars: 100 μm. (C) Dispersion plot correlating the percentage of CA9-positve cells of samples used in panel B versus *DPPA3* mRNA normalized expression values. Black line represents linear regression. (D) Correlation analysis between *CA9* and *DPPA3* transcript expression in primary tumour (left) and metastasis (right) samples from prostate cancer patients (GSE21034). Plotted values correspond to the log_2-normalized gene expression values for the two genes for each patient. Black line represents linear regression and dark grey area indicates the limits of the confidence intervals. (A, B) Data are represented as mean ± SEM. (A, C, D) n, number of samples. (C, D) r, Pearson´s correlation coefficient; p, p-value.

6. DPPA3 overexpression is negatively associated with cell cycle progression and tumour growth

Many studies have documented that hypoxia causes a cell cycle arrest and prevention of cancer cell proliferation [48,209]. As showed before, several cell cycle-related gene sets were negatively enriched in DPPA3-OE megacolonies (Figure R3A). These results made as wonder if DPPA3 affected cancer cell proliferation and tumour growth. Indeed, we observed that DPPA3 impacted on cell cycle progression causing an accumulation of normoxic SW1222 and HT29 cells at G2/M phases (Figure R17), mimicking the dynamics observed in SCCC [269].

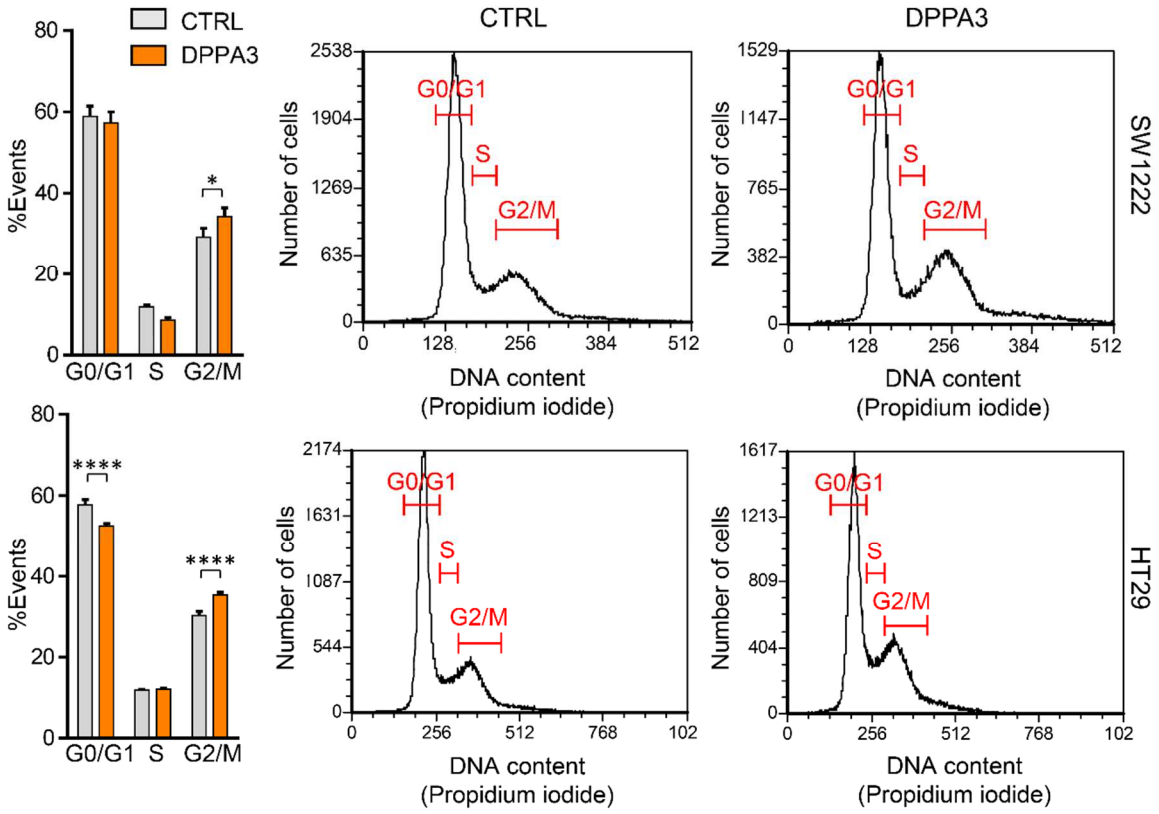


Figure R17. DPPA3-OE cells accumulate at the G2/M phases of the cell cycle. (Right) Cell cycle analysis and (left) quantification of normoxic control (CTRL) and DPPA3-OE (DPPA3) SW1222 and HT29 cells. Frequency histograms show the distribution of cells in the three major phases of the cell cycle (G0/G1, S, G2/M). *$P \leq 0.05$; ****$P \leq 0.0001$ of 2-way ANOVA test.

BrdU incorporation into DNA and phosphorylation of histone H3 at Ser10 are broadly used to analyse the status of S and mitosis phases of cell cycle respectively [315,316] (Figure R18A). We performed a BrdU assay in which we evaluated its incorporation into DNA by flow cytometry in normoxic and hypoxic control and DPPA3-OE SW1222 cells. Hypoxia diminished BrdU incorporation in both types of cells (Figure R18B) while a minor percentage of DPPA3-OE cells incorporated BrdU compared to their controls in both normoxic and hypoxic conditions.

We also evaluated the presence of phosphorylated histone H3 at Ser10, a marker of mitotic phase. Upon exit from mitosis, a global dephosphorylation of histone H3 takes place [317]. We observed by Western blot that hypoxia promoted a reduction of phosphorylated histone H3 at Ser10 in both DPPA3-OE SW1222 and HT29 cells (Figure R18C). However, DPPA3 overexpression caused opposite effects on phosphorylated histone H3 in both cell lines: decreasing in SW1222 while increasing in HT29 cells. Both scenarios indicated that DPPA3 overexpression affected cell cycle dynamics at G2/M phase, either by halting cell cycle progression before entering mitosis, where histone H3 phosphorylation occurs, or right after entry into mitosis, leading to an accumulation of the mark.

A stop at G2/M checkpoint has been described in response to DNA damage for allowing a complete DNA repair before progressing to S-phase or mitosis [318]. To exclude the possibility that the accumulation of DPPA-OE cells at G2/M phases of cell cycle was a consequence of DNA damage, we assessed the phosphorylation levels of histone H2AX at Ser139 [319] (Figure R18D). We observed that DPPA3 overexpression in SW1222 and HT29 cells did not affect H2AX phosphorylation *in vitro*. Similarly, hypoxia treatment did not exert any changes in H2AX phosphorylation. These results were further validated *in vivo* in HT29 cells growing as subcutaneous tumour xenografts (Figure R18E). We therefore concluded that DPPA3 promoted an accumulation at G2/M-phase due to causes other than DNA damage.

The role of DPPA3 as a negative regulator of cell cycle was confirmed upon analysing the gene expression profile distinctive of DPPA3-OE SW1222 megacolonies. Several MYC target genes (Figure R18F) and cell cycle genes (Table R1) were down-modulated in DPPA3-OE cells. Of note, the expression of these same genes was also reduced in SCCC [269] (Figure R18F). These data aroused the possibility that DPPA3 could be involved in an epigenetic reprogramming of cancer cells crucial for entry to dormancy. We therefore sought to delve deeper into this issue.

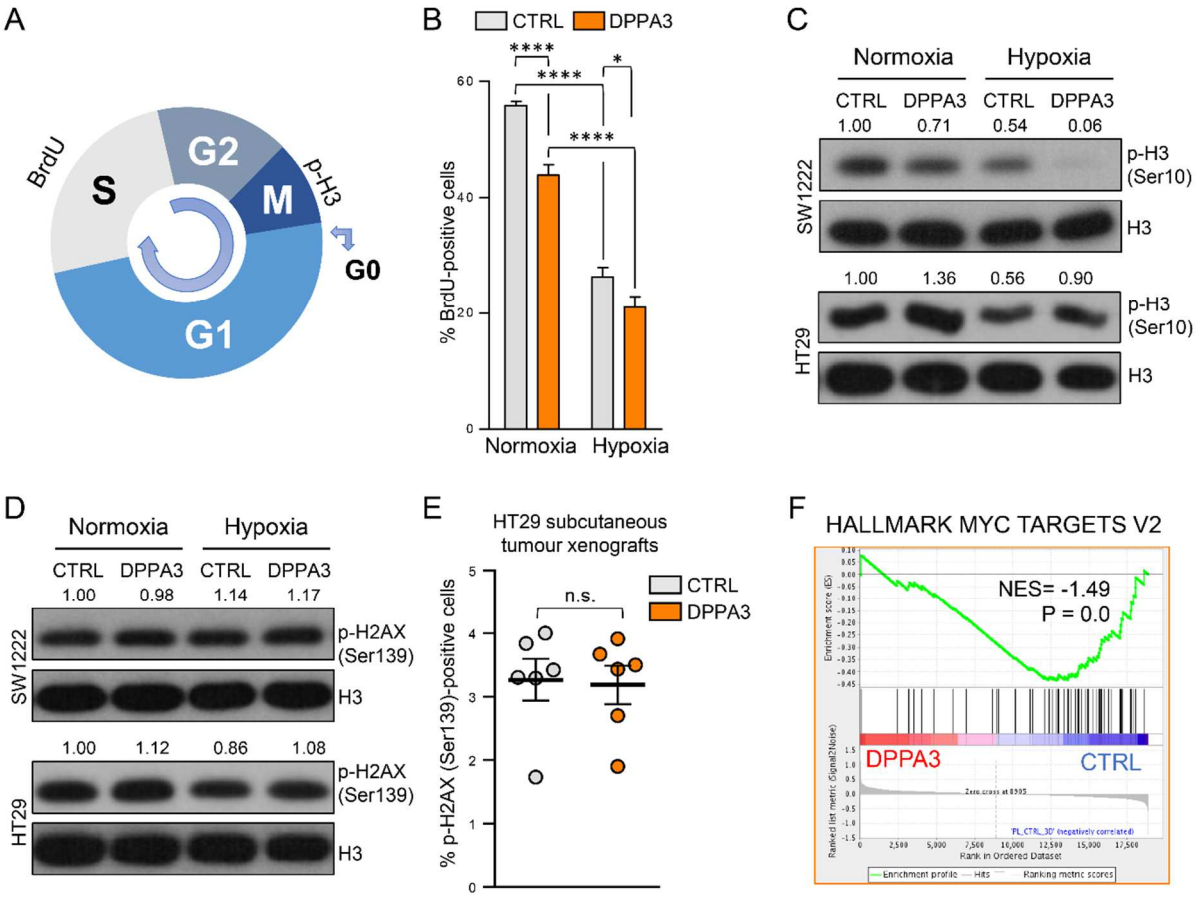


Figure R18. DPPA3 overexpression affects cell cycle progression. (A) Diagram representing the different phases of eukaryotic cell cycle. Phosphorylation of histone H3 (p-H3) at Ser10 takes place in the nuclei of cells during the M-phase (mitosis). Upon exit from mitosis, there is a global dephosphorylation of histone H3. BrdU can be incorporated into the newly synthesized DNA of replicating cells during the S-phase (synthesis), in which DNA is replicated. (B) Quantification of BrdU analysis performed with normoxic and hypoxic control (CTRL) and DPPA3-OE (DPPA3) SW1222 cells. **P* ≤ 0.05; *****P* ≤ 0.0001 of 2-way ANOVA test. (C, D) Western blots showing (C) p-histone H3 (Ser10) and (D) p-H2AX (Ser139) levels of normoxic and hypoxic control and DPPA3-OE SW1222 cells. Histone H3 was used as loading control. Numbers show quantification by densitometry of p-histone H3 and p-H2AX levels, which were normalized to histone H3 and expressed relative to control normoxic condition. (E) Quantification of p-H2AX (Ser139)-positive cells evaluated by immunohistochemistry in DPPA3-OE (DPPA3) and control (CTRL) HT29 subcutaneous tumour xenografts (n = 6 tumours per condition). Data are represented as mean ± SEM. Unpaired two-tailed t test. n.s. not significant. (F) GSEA plot showing enrichment of the indicated publicly available gene set in the expression profile of DPPA3-OE versus control SW1222 megacolonies. NES, normalized enrichment score; P, p-value.

SW1222 megacolonies

Gene symbol	SCCC FC	SCCC P	DPPA3 FC	DPPA3 P	Gene assignment
CDK1	-1.54	0.00	-1.41	0.01	Cyclin Dependent Kinase 1
CDK2	-1.43	0.00	-1.20	0.01	Cyclin Dependent Kinase 2
CCNB1	-1.97	0.00	-1.29	0.00	Cyclin B1
CCNB2	-1.70	0.00	-1.24	0.04	Cyclin B2
CCNA2	-1.68	0.00	-1.22	0.03	Cyclin A2
E2F1	-1.24	0.00	-1.20	0.03	E2F Transcription Factor 1
CHEK1	-1.54	0.00	-1.46	0.00	Checkpoint Kinase 1
PLK1	-1.66	0.00	-1.26	0.04	Polo Like Kinase 1
AURKA	-1.50	0.00	-1.49	0.00	Aurora Kinase A

Puig et al. *JCI*. 2018

Table R1. DPPA3 overexpression down-modulates the expression of cell cycle genes. Fold change table for gene expression of cell cycle regulators in SW1222 SCCC versus RCCC (green box) and DPPA3-OE versus control SW1222 megacolonies (orange box). FC, fold change. P, p-value.

High p38/ERK signalling activity is considered a hallmark of tumour dormancy [61,64,65] and hypoxia is known to regulate p38 and ERK signalling pathways in cancer cells [320–325]. Therefore, we assessed if DPPA3 overexpression altered those pathways by checking the phosphorylation status of p38 and ERK MAPKs. Both hypoxia and DPPA3 overexpression increased the phosphorylation of p38 in SW1222 and HT29 cells (Figure R19A). Whereas no remarkable changes in p-ERK levels were observed in control cells undergoing hypoxic treatment, those overexpressing DPPA3, especially SW1222 cells, exhibited higher p-ERK levels than hypoxic controls. Both hypoxia and DPPA3 overexpression under normoxic conditions increased p-p38/p-ERK ratio in SW1222 and HT29 cell lines, suggesting that both scenarios skewed cellular phenotype towards dormancy. DPPA3 overexpression, however, could not increase the elevated p38/ERK signalling observed in hypoxic cells. Of note, DPPA3 overexpression in SW1222 megacolonies upregulated a series of p38 target genes (Figure R19B).

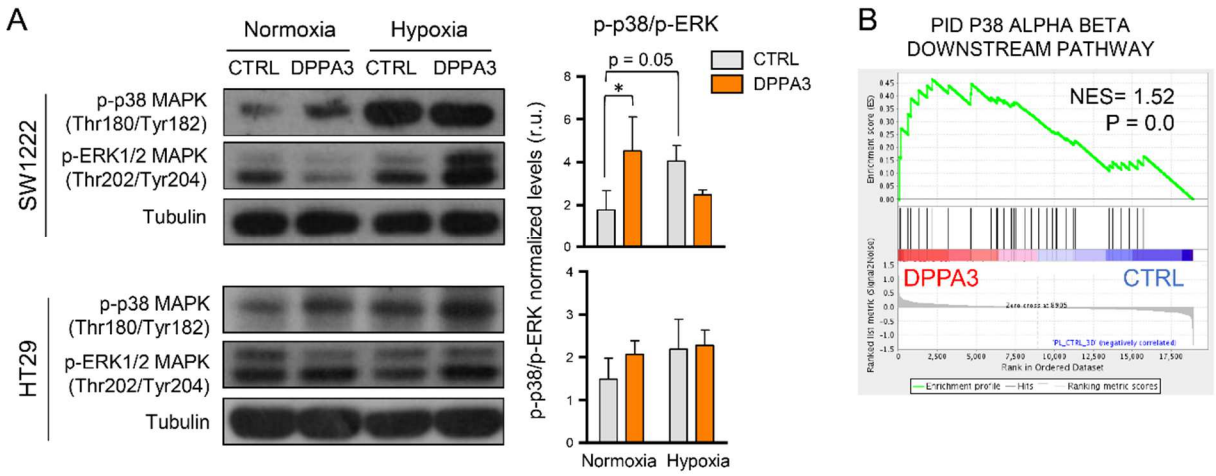


Figure R19. DPPA3 overexpression increases p38/ERK activity ratio. (A) (Left) Western blots showing phosphorylated p38 and ERK1/2 MAPKs at indicated residues in normoxic and hypoxic control (CTRL) and DPPA3-OE (DPPA3) SW1222 and HT29 cells. Tubulin was used as loading control. (Right) Quantification of p-p38/p-ERK1/2 ratio. P-p38 and p-ERK1/2 were normalized to tubulin. Then, the ratio of normalized p-p38/p-ERK1/2 levels was calculated and expressed relative to control normoxic condition. Experiments were performed in triplicates, and data represent the average of three experiments ± SD. r.u., relative units. **P* ≤ 0.05 of 2-way ANOVA test. (B) GSEA plot showing enrichment of the indicated gene set in the expression profile of DPPA3-OE versus control SW1222 megacolonies. NES, normalized enrichment score. (A, B) P, p-value.

Further to this, a panel of genes associated with the dormancy phenotype was regulated upon DPPA3 overexpression (Table R2), including *RARB*, *NR1H3*, *CDKN1A, CTSD, DDR1, STAT3, NDRG1, BUB1, CKS2, DNMT1, FOXM1 and APEX1* (Figure R19D). Importantly, some of these genes (*RARB*, *NR1H3*, *CDKN1A, CTSD, DDR1*, *STAT3 and NDRG1*) are linked to active p38 signalling [59,289,326], suggesting a connection between DPPA3 and dormancy by modulating genetic programs responsible for this phenotype.

DPPA3-OE SW1222 megacolonies

	Gene	FC	P	Gene assignment	Reference
Upregulated	*RARB*	1.68	0.00	Retinoic Acid Receptor Beta	Adam, A. P. et al. Can Res (2009); Tsai, H.C. et al. Cancer Cell (2012); Kim, R. S. et al. PLoS ONE (2012)
	NR1H3	1.26	0.00	Nuclear Receptor Subfamily 1 Group H Member 3	Adam, A. P. et al. Can Res (2009)
	CDKN1A	1.24	0.10	Cyclin Dependent Kinase Inhibitor 1A	Adam, A. P. et al. Can Res (2009); Tsai, H.C. et al. Cancer Cell (2012); Kim, R. S. et al. PLoS ONE (2012)
	CTSD	1.67	0.00	Cathepsin D	Kim, R. S. et al. PLoS ONE (2012)
	DDR1	1.22	0.01	Discoidin Domain Receptor Tyrosine Kinase 1	Kim, R. S. et al. PLoS ONE (2012)
	STAT3	1.23	0.10	Signal Transducer And Activator Of Transcription 3	Kim, R. S. et al. PLoS ONE (2012)
	NDRG1	1.72	0.00	N-Myc Downstream Regulated 1	Kobayashi et al. J Exp Med (2011)
Downregulated	*BUB1*	-1.33	0.00	BUB1 Mitotic Checkpoint Serine/Threonine Kinase	Kim, R. S. et al. PLoS ONE (2012)
	CKS2	-1.25	0.03	CDC28 Protein Kinase Regulatory Subunit 2	Kim, R. S. et al. PLoS ONE (2012)
	DNMT1	-1.49	0.00	DNA Methyltransferase 1	Kim, R. S. et al. PLoS ONE (2012)
	FOXM1	-1.28	0.00	Forkhead Box M1	Kim, R. S. et al. PLoS ONE (2012)
	APEX1	-1.20	0.05	Apurinic/Apyrimidinic Endodeoxyribonuclease 1	Kim, R. S. et al. PLoS ONE (2012)

Table R2. DPPA3 modulates the expression of cancer cell dormancy genes. Fold change table for gene expression in DPPA3-OE versus control SW1222 megacolonies of factors upregulated (red box) or downregulated (blue box) in dormancy cancer models. FC, fold change; P, p-value.

We used heterotopic and orthotopic *in vivo* models for assessing if the effect of DPPA3 blocking cell cycle progression *in vitro* would ultimately affect tumour growth. In all cases, DOX was administered in mice drinking water *ad libitum* along tumour growth. For the heterotopic model, we injected subcutaneously DPPA3-OE and control SW1222 and HT29 cancer cells. We decided to calculate the area under the curve (AUC) for each growing tumour as a method to quantify the effects of the different experimental conditions. This single numerical value was described as easier and more comprehensive for quantifying tumour growth dynamics [286]. DPPA3 overexpression in subcutaneous SW1222 xenografts caused a decreased yet not significant tumour growth (Figure R20A), possibly due to the high variability in control xenografts and reduced number of samples. In the case of HT29 cells, DPPA3 overexpression significantly slowed down tumour growth (Figure R20B). Finally, no differences in tumour growth were appreciable in shDPPA3 tumour xenografts (Figure R20C).

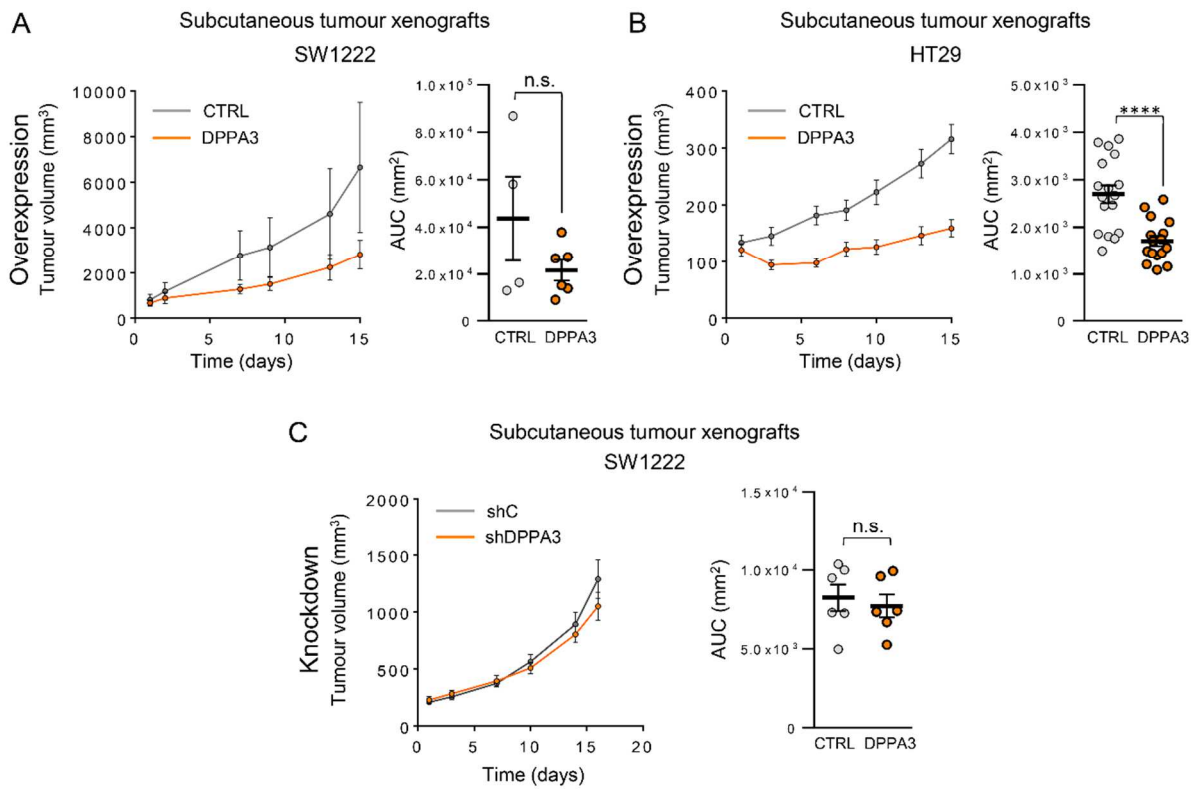


Figure R20. DPPA3 overexpression decreases tumour growth. (A-C) (Left) Tumour growth curves and (right) AUC scatter plots of each tumour growth curve of the indicated subcutaneous xenografts. (Left) Each point in tumour growth curves represents the mean ± SEM. (Right) Data are represented as mean ± SEM. n.s. not significant; ***$P \leq 0.001$ of unpaired two-tailed t test. (A, B) Control (CTRL) and DPPA3-OE (DPPA3) (A) SW1222 (n = 4-6 tumours per condition) or (B) HT29 (n = 16-18 tumours per condition) xenografts. (C) shC and shDPPA3 SW1222 xenografts (n = 6 tumours per condition). shDPPA3 clone 2 cell line was used in this analysis.

We used an orthotopic injection of cells in the cecum wall of mice as a second approach for evaluating the effects of DPPA3 overexpression. This was a more physiologic model in which cells were injected into an equivalent location to that of primary tumours in CRC patients. Furthermore, this orthotopic model also allowed us to determine the metastatic potential of the injected cancer cells, since they can spread from the primary site to distant organs taking advantage of the high irrigation of the cecum [288]. DPPA3-OE primary tumours originated from both SW1222 and HT29 cells were smaller than controls (Figure R21A). Whereas SW1222 did not present any metastatic potential, HT29 easily metastasized to the lungs. We first observed that DPPA3 reduced the total number of CK20-positive lung metastasis foci (Figure R21B). In addition, DPPA3 also reduced the size of the metastatic lesions. These results suggest that the delay in cell cycle progression promoted by DPPA3 overexpression (Figures R17 and R18) caused a reduction of tumour and metastasis growth.

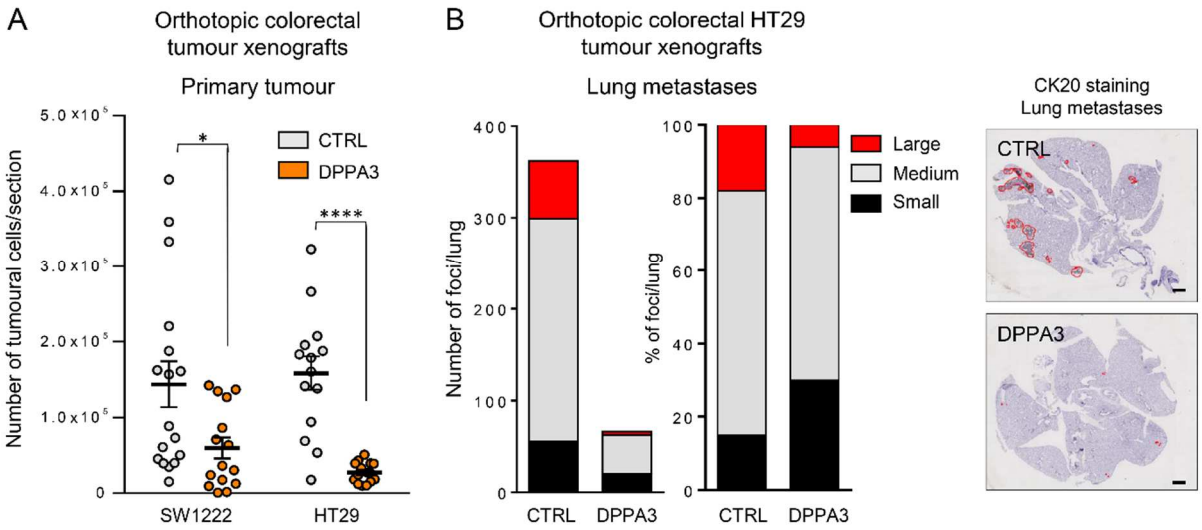


Figure R21. DPPA3 overexpression reduces the number of metastatic foci and increases the proportion of small lesions. (A) Quantification of total tumoural cells in haematoxylin-stained primary tumour sections from control (CTRL) and DPPA3-OE (DPPA3) SW1222 and HT29 cells orthotopically injected in mice (n = 14-17 tumours per condition). Data are represented as mean ± SEM. **P* ≤ 0.05; ****P* ≤ 0.001 of unpaired two-tailed t test. (B) Bar graphs showing (left) total number and (middle) percentage of large, medium and small metastatic foci per lung of mice orthotopically injected with control and DPPA3-OE HT29 cells. (Right) Representative images of lungs from mice injected with control and DPPA3-OE HT29 cells showing metastatic foci (circled in red) stained with the CK20 colon carcinoma marker. Scale bars: 1 mm.

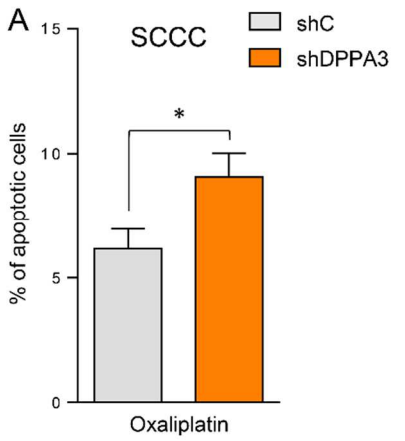
Altogether, the accumulated data obtained *in vitro* and *in vivo* indicates that DPPA3 plays a relevant role modulating the entry of tumour cells into a slow-cycling status and eventually dormancy.

7. DPPA3 increases drug resistance

Treatments targeting proliferating cancer cells often miss dormant tumour cells as they are not actively dividing [44,53,269]. Moreover, it is known that hypoxia is a crucial mediator of chemo- and radioresistance [327,328]. HIF overexpression in clinical samples is associated with therapeutic resistance and decreased survival following ionizing radiation or chemotherapy [329]. Given the enrichment of gene sets related to drug resistance together with the involvement of DPPA3 in hypoxia and dormancy programs (Figure R3), we hypothesized that this factor could affect resistance of tumours to conventional chemotherapies.

We used 3D *in vitro* models as a first approach to study if DPPA3 was involved in drug resistance. Since this factor was overexpressed in SCCC, we assessed if it affected their chemoresistant properties. Indeed, depletion of DPPA3 sensitized SCCC present in SW1222 megacolonies to chemotherapy (Figure R22A). In addition, we observed that DPPA3

increased the expression of ATP-binding cassette transporters, which are implicated in cancer cell multidrug resistance [330] (Figure R22B).



B

DPPA3-OE SW1222 megacolonies

Gene symbol	FC	P	Gene assignment
ABCB1	1.87	0.00	ATP Binding Cassette Subfamily B Member 1
ABCC2	1.31	0.01	ATP Binding Cassette Subfamily C Member 2
ABCC3	1.20	0.04	ATP Binding Cassette Subfamily C Member 3
ABCC4	1.30	0.03	ATP Binding Cassette Subfamily C Member 4

Figure R22. DPPA3 depletion sensitizes SCCC to oxaliplatin treatment while its overexpression upregulates drug resistance genes. (A) Apoptosis assay of shC and shDPPA3 SCCC extracted from oxaliplatin-treated SW1222 megacolonies. shDPPA3 clone 2 cell line was used in this analysis. Experiments were performed in triplicates, and data are represented as mean ± SD. **P* ≤ 0.05 of unpaired two-tailed t test. (B) Fold change table for gene expression of ATP-binding cassette transporters in DPPA3-OE versus control SW1222 megacolonies. FC, fold change; P, p-value.

These initial results *in vitro* encouraged us to test the role of DPPA3 in chemoresistance in mouse models. Mice were subcutaneously injected with control and DPPA3-OE HT29 cells and treated with DOX in drinking water all along the experiment. When tumours were established mice were regularly treated with the standard-of-care chemotherapeutic agents oxaliplatin, 5-FU or irinotecan. While oxaliplatin and 5-FU-treated control tumours showed a significant smaller AUC compared to the vehicle condition, this effect was abolished when DPPA3 was overexpressed (Figure R23A and B). However, both types of tumours were sensitive to irinotecan, even though the difference was more remarkable in control than in DPPA3-OE tumours.

We had previously observed increased phosphorylation levels of the histone H3 in DPPA3-OE HT29 cells *in vitro* (Figure R18C), which could be indicative of a halt in the mitotic phase of cell cycle. We next validated if this mark was associated to a decreased growth in HT29 subcutaneous tumour xenografts. Although not statistically significant in all treatments, we observed that chemotherapy, apart from decreasing tumour growth, it also caused a general induction of histone H3 phosphorylation (Figure R23C). We also observed that DPPA3-OE vehicle tumours showed p-histone H3 levels slightly higher than control vehicles.

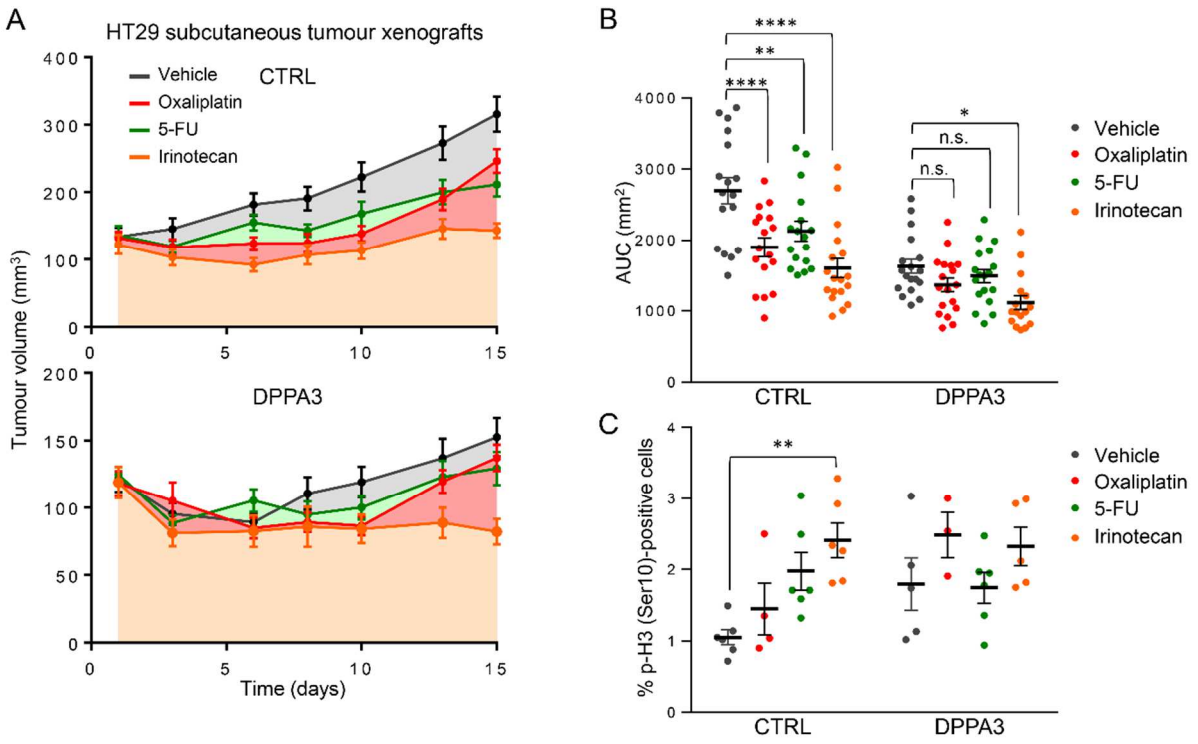


Figure R23. DPPA3-overexpressing tumour xenografts are more resistant to chemo-treatment. (A) Growth curves of control (CTRL) and DPPA3-OE (DPPA3) HT29 subcutaneous xenografts from vehicle, oxaliplatin, 5-FU and irinotecan-treated mice. Each point represents the mean ± SEM of 17-18 tumours. (B) Scatter plot showing the AUC of each tumour growth curve from panel A of each indicated group. (C) Quantification of p-histone H3 (Ser10)-positive cells evaluated by immunohistochemistry in CTRL and DPPA3 HT29 subcutaneous xenografts from vehicle, oxaliplatin, 5-FU and irinotecan-treated mice (n = 6 tumours per condition). (B, C) Data are represented as mean ± SEM. n.s. not significant; **P* ≤ 0.05, ***P* ≤ 0.01; *****P* ≤ 0.0001 of 2-way ANOVA test.

Chemoresistance can also be evaluated by analysing the capacity of the resistant surviving tumour cells to regrow after an effective phase of treatment. As irinotecan was the only drug that significantly affected tumour growth in both conditions, we specifically assessed tumour regrowth kinetics in irinotecan-treated tumours after treatment removal (Figure R24A). To facilitate interpretation of the results, DOX was withdrawn to stop DPPA3 overexpression and equalize the proliferation rate of both control and DPPA3-OE residual cells. Then, we evaluated the regrowth capacity by comparing the AUC of each tumour (treatment-withdrawn and vehicle) for each condition, control and DPPA3-OE. While differences in tumour growth were still noticeable between vehicle and treatment-withdrawn control tumours, these were lost in tumours overexpressing DPPA3 (Figure R24B).

We measured tumour regrowth using adapted RECIST criteria [287] (increase of 20% of tumour volume) and observed that progression-free survival (PFS) was significantly shorter in mice with DPPA3-OE tumours after withdrawal of irinotecan treatment (Figure R24C). In

conclusion, control tumours showed delayed regrowth after irinotecan treatment compared to DPPA3 condition, suggesting that DPPA3 promotes an accumulation of resistant residual tumoural cells with the capacity to induce cancer relapse.

Considering the high chemoresistance exhibited by SCCC [269], we wondered if these cells were responsible of the earlier relapse observed in DPPA3-OE tumour xenografts after irinotecan treatment. We performed a BrdU pulse-chase to label SCCC as previously described [331] . We quantified SCCC enrichment in tumour xenografts after irinotecan treatment by calculating the fold increase of BrdU-positive cells (label retaining cells, LRCs) in irinotecan-treated versus vehicle tumours. Importantly, DPPA3-OE specimens were enriched in BrdU LRCs upon irinotecan administration, which could have influenced their regrowth after treatment release (Figure R24D).

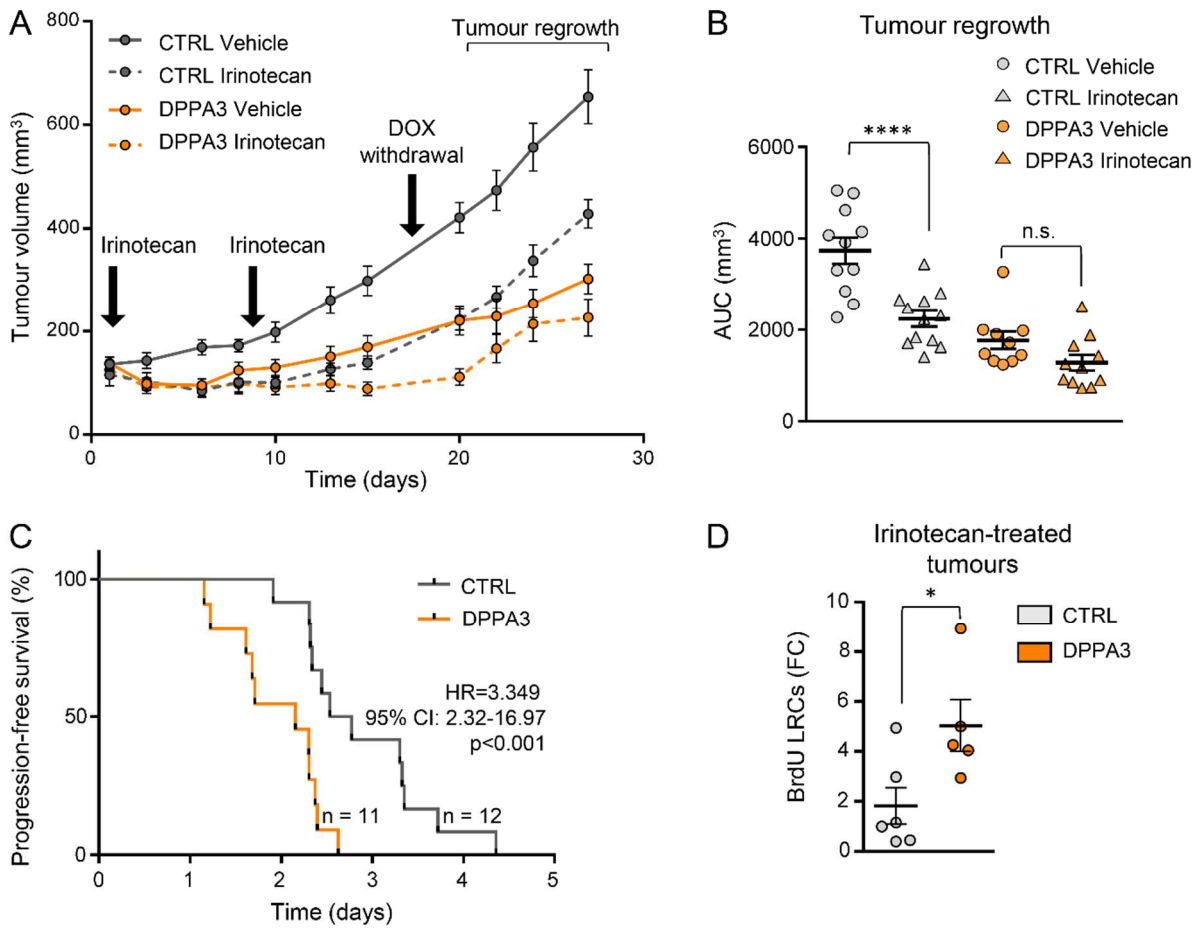


Figure R24. DPPA3 overexpression enhances drug resistance in tumour xenografts. (A) Growth curves of control (CTRL, grey) and DPPA3-OE (DPPA3, orange) HT29 subcutaneous xenografts from vehicle (solid line) and irinotecan (dashed line)-treated mice. The animals were continuously treated with DOX and received a total of two doses of irinotecan until both treatments were withdrawn. Time elapsed between treatment withdrawal until experimental end-point determined tumour regrowth. Each point represents the mean ± SEM of 10-11 tumours. (B) Scatter plot showing the AUC of each tumour growth curve along the regrowth phase from panel A (7 days). Data are represented as mean ± SEM. n.s. not significant; ****$P \leq 0.0001$ of 2-way ANOVA test. (C) The survival curve represents progression-free survival (PFS) percentages showing the impact of irinotecan on the regrowth of control or DPPA3-OE HT29 xenografts. A 20% increase in tumour volume after treatment release was considered as regrowth or progression. Significance was calculated using the log-rank (Mantel-Cox) test. HR, hazard ratio; p, log-rank p-value. (D) Quantification of BrdU Label-Retaining Cells (LRCs) evaluated by immunohistochemistry. Fold changes (FC) are relative to one of the control (CTRL) irinotecan-treated xenografts (n = 5-6 tumours per condition). Percentages of BrdU LRCs of CTRL and DPPA3-OE (DPPA3) irinotecan-treated xenografts are corrected by the average percentage of the corresponding BrdU vehicle xenograft group. Data are represented as mean ± SEM. *P ≤ 0.05, of unpaired two-tailed t test.

In summary, all these data indicate that DPPA3 overexpression in CRC cells confers an increased resistance to conventional chemotherapy by two possible strategies: (i) a passive mechanism associated with the low proliferative rate of DPPA3-OE cells and an accumulation of SCCC in tumours, and/or (ii) an active mechanism related to the overexpression of a variety of drug detoxifying membrane pumps.

To address if these effects could be relevant in the clinic, we took advantage of gene expression data previously generated in our laboratory using an nCounter platform (NanoString Technologies). This technology has shown excellent robustness and reproducibility as well as a high sensitivity of target detection even at very low input RNA amounts [332]. In these analyses, the expression of several genes was studied, including *DPPA3*, in primary tumour samples of a cohort of 53 baseline stage II and III 5-FU-based adjuvant chemo-treated CRC patients. We stratified samples according to *DPPA3* expression and observed a scatter plot with two differentiated groups, one expressing higher levels of *DPPA3* (DPPA3-high) and another expressing null or low levels of the gene (DPPA3-low/neg) (Figure R25A). Interestingly, patients expressing elevated levels of *DPPA3* (DPPA3-high) showed a significant shorter PFS (Figure R25B).

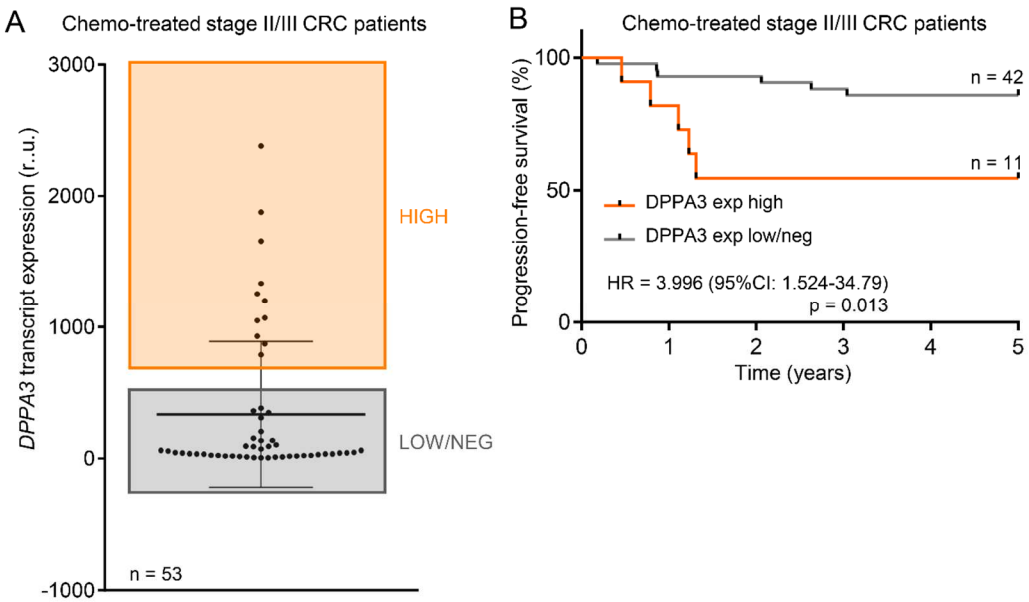


Figure R25. High expression of *DPPA3* predicts poor progression-free survival in CRC chemo-treated patients. (A) Stratification of 53 primary tumours from chemo-treated stage II/III colorectal cancer (CRC) patients (VHIO cohort) according to *DPPA3* transcript expression analysed with the nCounter platform. Two differentiated groups were observed: samples showing higher transcript expression of *DPPA3* were classified as high (orange box) whereas those expressing lower levels or were negative for *DPPA3* expression were classified as low/negative (grey box). Data are represented as mean ± SD. r.u., relative units. (B) Survival curves represent progression-free survival percentages of the same cohort stratified in panel A according to *DPPA3* transcript expression. HR, hazard ratio. Cox proportional hazards model.

We then generated a gene expression signature distinctive of DPPA3 (DPPA3 Sig) by an integrative analysis of the microarray profiles obtained from DPPA3-OE, shDPPA3 and *DPPA3* KO SW1222 CRC cells. We observed that those CRC patients treated with 5-FU-based adjuvant chemotherapy showed shorter PFS when their stage II or III primary tumours were enriched for the DPPA3 Sig (Figure R26A and R26B) whereas those non-treated in the same cohort (GSE39582 dataset) did not. Unfortunately, these results were not validated by immunohistochemistry since specific anti-DPPA3 antibodies were not available to detect the human protein.

Knowing that DPPA3 may promote the expression of dormancy genes, and that it was sufficient to decrease tumour growth and enhance resistance to chemotherapy in xenograft models, we sought to assess if CRC chemo-treated patients classified as positive for a DPPA3 Sig were also enriched in a genetic dormancy program. For this reason, we studied if there was a correlation between the DPPA3 Sig and a dormancy signature already published and validated in breast cancer patients [289] (Figure R26C). We checked the enrichment score (ES) of tumours analysed in Figure R26B (chemo-treated stage II and III CRC patients) in both DPPA3 and dormancy signatures and found a weak positive correlation between them. Given that the correlation was limited, we checked the enrichment in the dormancy signature comparing DPPA3 Sig-high versus DPPA3 Sig-low CRC tumours (Figure R26D), excluding DPPA3 Sig-medium samples. Interestingly, DPPA3 Sig-high tumours were positively enriched in the dormancy signature. All these data suggests that DPPA3 could enhance drug resistance by regulating an expression profile that commits tumours to a dormant phenotype.

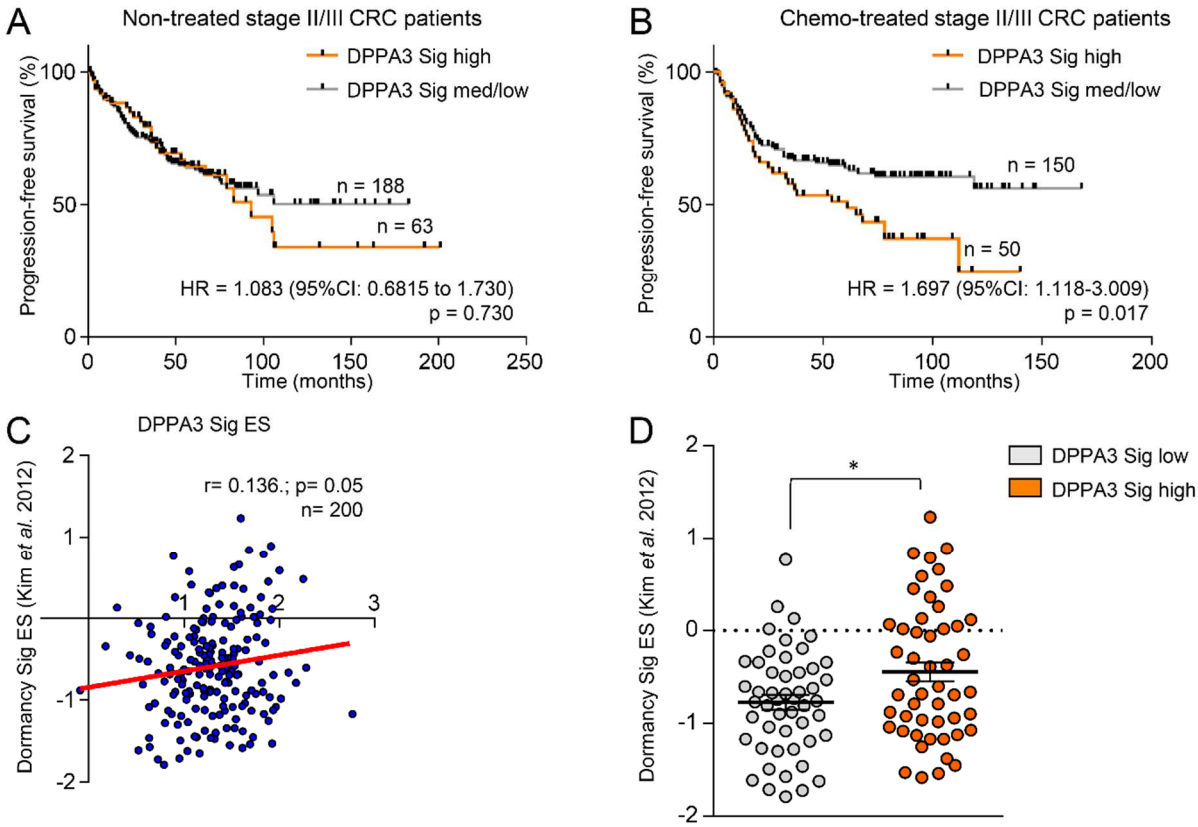
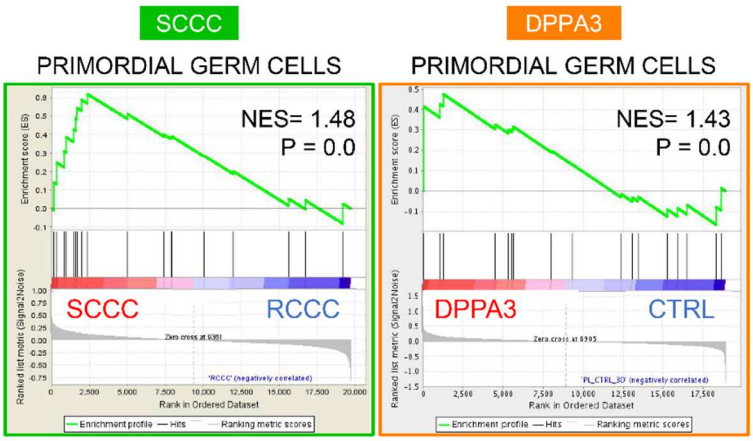


Figure R26. DPPA3 signature predicts poor progression-free survival in CRC chemo-treated patients and is associated with a dormancy signature. (A, B) Progression-free survival percentages of (A) non-treated (n=251) and (B) chemo-treated (n=200) stage II/III colorectal cancer (CRC) patients (GSE39582) according to DPPA3 signature (Sig) enrichment score (ES). Samples with ES in the upper quartile were classified as high (orange lines) while the rest were classified as medium (med) and low (grey lines). HR, hazard ratio. Cox proportional hazards model. (C) Dot plot correlating ES of DPPA3 versus dormancy signatures in tumour samples from chemo-treated stage II/III CRC patients (GSE39582, n=200) analysed in panel B according to their gene expression profiles. Red line represents linear regression. r, Pearson´s correlation coefficient; p, p-value. (D) The dormancy Sig ES of tumour samples from chemo-treated stage II/III CRC patients (GSE39582) classified as DPPA3 Sig high (upper quartile, n = 50) or low (lower quartile, n = 50) were plotted. Data are represented as mean ± SEM. **P* ≤ 0.05 of unpaired two-tailed t test.

8. DPPA3 modulates cell fate commitment

DPPA3 expression is boosted in undifferentiated cells at specific stages of embryonic development, specifically in ESCs and PGCs [233]. We found that both SW1222 SCCC and DPPA3-OE SW1222 megacolonies upregulated the expression of genes involved in germ cell development, a stage in which these cells exhibit a low degree of differentiation (Figure R27).



Puig et al. *JCI*. 2018

Figure R27. Germ cell gene sets are enriched in SCCC and DPPA3-OE cells. GSEA plots showing enrichment of the indicated custom gene set in the expression profile of SW1222 SCCC versus RCCC (SCCC, green square) and DPPA3-OE versus control SW1222 megacolonies (DPPA3, orange square). NES, normalized enrichment score; P, p-value.

In addition, SW1222 SCCC were negative for colonic differentiation markers like MUC2 [269]. Interestingly, hypoxia is a microenvironmental factor that inhibits differentiation of cancer cells [333,334], including the SW1222 CRC cell line [335]. Given the expression pattern of DPPA3 during embryonic development and its effects on hypoxia, we considered the possibility that this factor could influence cell fate commitment. We evaluated DPPA3-mediated effects on differentiation by analysing the levels of the mucin marker MUC2 in SW1222 megacolonies of the different SW1222-DPPA3 models: overexpressing (DPPA3) and depleted (*DPPA3* KO and shDPPA3). Whereas *DPPA3*-depleted megacolonies showed an increased number of MUC2-positive cells, its overexpression drastically reduced them (Figure R28A). The expression of several differentiation and stemness genes was also affected by the modulation of *DPPA3* (Figure R28B). While high levels of DPPA3 resulted in an increased expression of stemness markers and decreased expression of cell differentiation markers, *DPPA3* depletion caused the opposite effects.

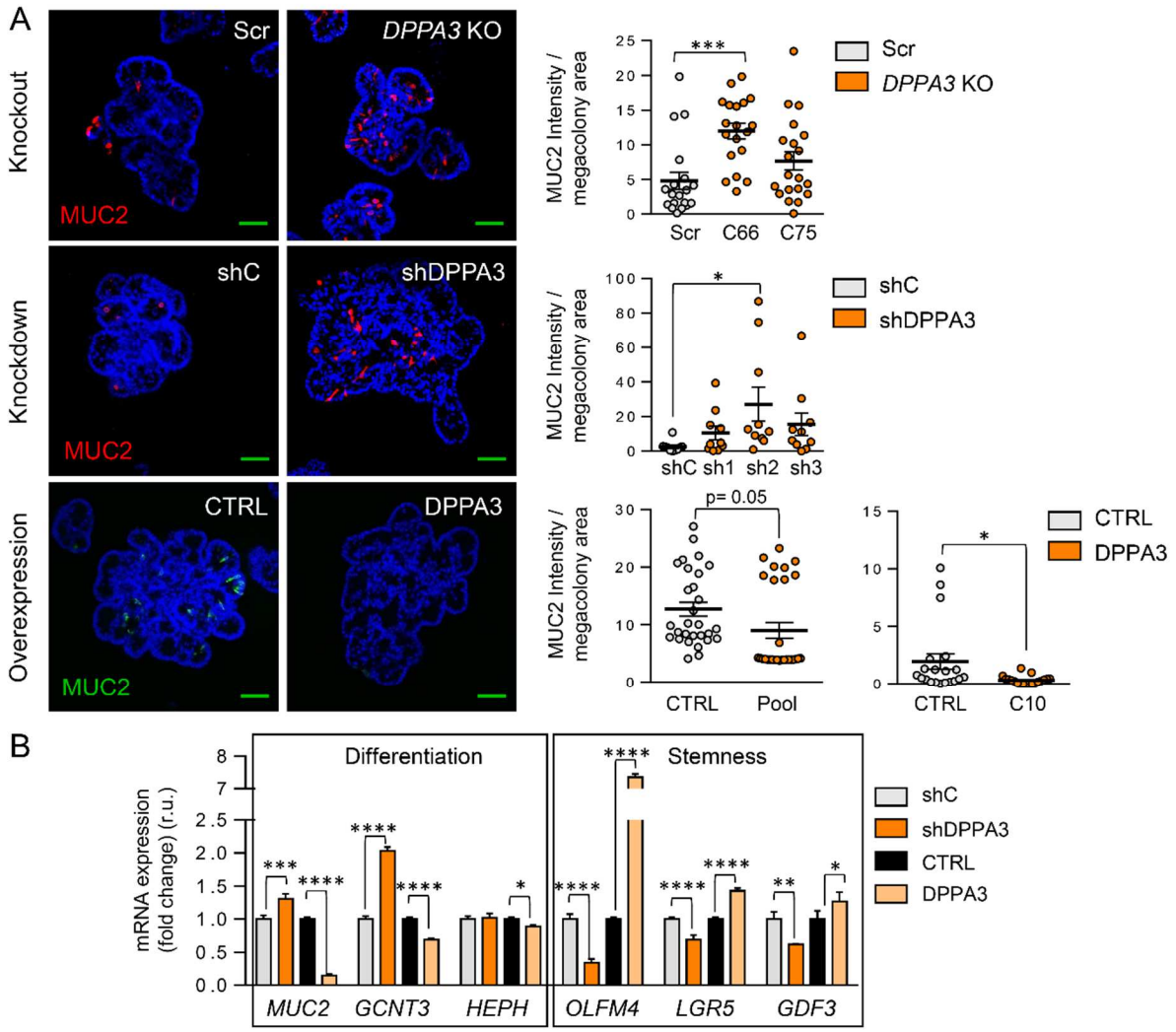


Figure R28. DPPA3 controls cell fate commitment *in vitro*. (A) (Left) Representative images and (right) quantification of MUC2 evaluation by immunofluorescence analysis in *DPPA3* knockout (*DPPA3* KO, top), shDPPA3 (middle) or overexpressing (DPPA3, bottom) SW1222 megacolonies with their respective controls: scramble (Scr, top), shC (middle) or empty vector (CTRL, bottom). For knockout models, two *DPPA3* KO clones (C66 and C75) were used. For knockdown models, shDPPA3 clones 1 to 3 (sh1-3) were used. For overexpression models, DPPA3 pool and clone 10 (C10) cell lines were used. Each dot in the MUC2 quantification plots represents a megacolony. 10-30 megacolonies per condition were analysed. Data are represented as mean ± SEM. **P* ≤ 0.05; ****P* ≤ 0.001 of 1-way ANOVA for knockout and knockdown models, and unpaired two-tailed t test for overexpression models. Scale bars: 100 µm. (B) Expression of the differentiation and stemness indicated genes was evaluated by qRT-PCR in shDPPA3 and DPPA3-OE (DPPA3) SW1222 megacolonies and respective controls. Expression levels were normalized to an endogenous control and expressed relative to respective control cells. Experiments were performed in triplicates, and data represent the average of three experiments ± SD. r.u., relative units. **P* ≤ 0.05; ****P* ≤ 0.001; *****P* ≤ 0.0001 of 1-way ANOVA. (A, B) p, p-value.

Similar results were observed *in vivo*, where *DPPA3* knockdown in SW1222 tumour xenografts caused an increased mucinous differentiation evaluated by Alcian blue staining (Figure R29A), while its overexpression decreased cell differentiation towards MUC2-positive cells (Figure R29B). In addition, shDPPA3 tumour xenografts resembled SW1222 megacolonies in terms of expression of stemness and differentiation markers (Figure R29C).

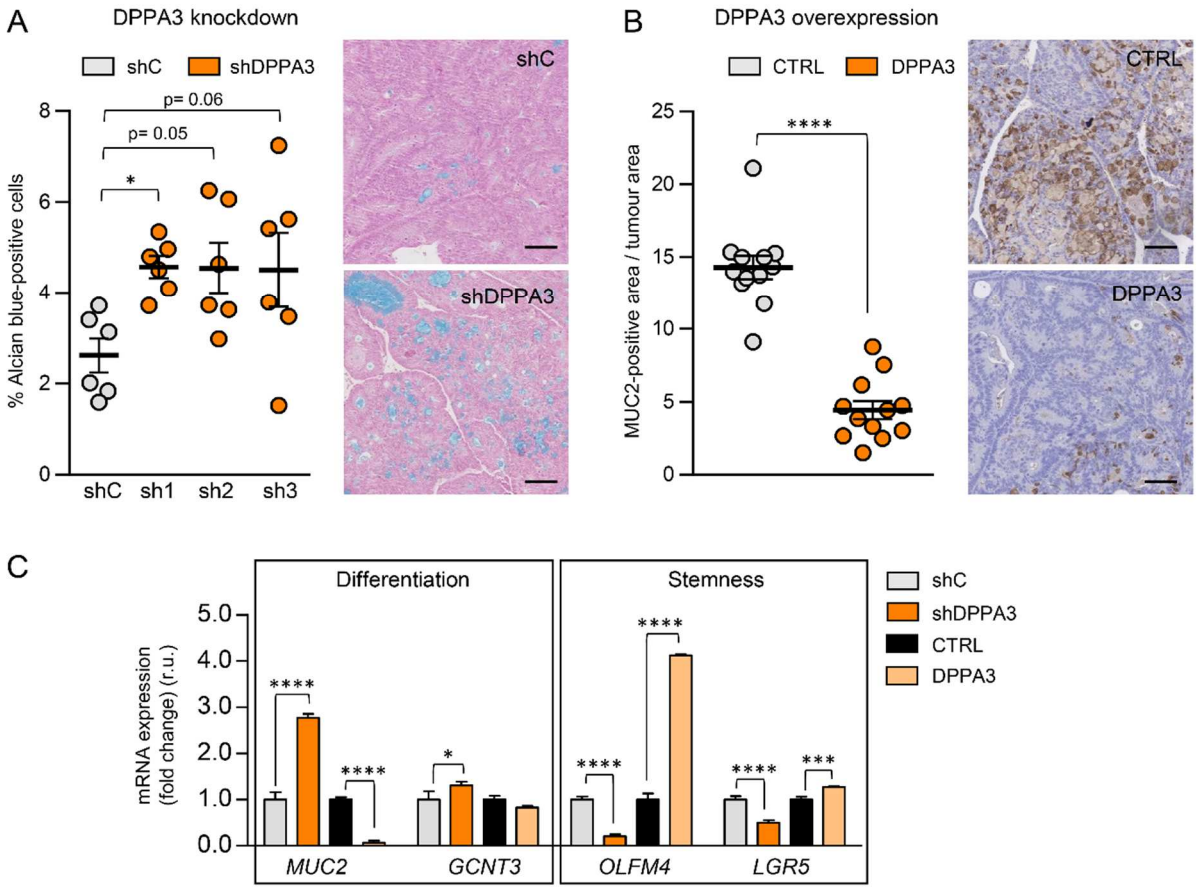


Figure R29. DPPA3 controls cell fate commitment *in vivo*. (A) (Left) Quantification and (right) representative images of Alcian blue staining in shDPPA3 clones (sh1-3) of SW1222 cells grown as subcutaneous xenografts (n = 6 tumours per condition). **P* ≤ 0.05 of 1-way ANOVA. (B) (Left) Quantification and (right) representative images of MUC2 evaluation by immunohistochemistry analysis in control (CTRL) and DPPA3-OE (DPPA3) SW1222 subcutaneous xenografts (n = 12 tumours per condition). Pool DPPA3-OE cells were used in this analysis. *****P* ≤ 0.0001 of unpaired two-tailed t test. (A, B) Data are represented as mean ± SEM. Scale bars: 100 µm. (C) Expression of differentiation and stemness indicated genes was evaluated by qRT-PCR in shDPPA3 and DPPA3-OE (DPPA3) SW1222 subcutaneous xenografts and respective controls. Expression levels were normalized to an endogenous control and expressed relative to respective control xenografts. Three tumours per condition were used, and data represent the average ± SD. r.u., relative units. **P* ≤ 0.05; ****P* ≤ 0.001; *****P* ≤ 0.0001 of 1-way ANOVA.

Since the expression of *LGR5* and *OLFM4* stemness markers was positively regulated by DPPA3, we aimed to study if this factor affected Wnt/β-catenin or NOTCH signalling pathways, as they modulate the expression of both genes in normal colonic crypt [12,15,336]. To investigate if the transcriptional activity of β-catenin was modulated by DPPA3, we used the TOP*Flash*/FOP*Flash* luciferase reporter system. Luciferase activity was assessed in HEK293T cells co-transfecting the TCF wild-type (TOP) or mutated (FOP) binding site reporter with the DPPA3 expression vector. DPPA3 overexpression increased the luciferase activity in cells transfected with the TOP*Flash* reporter (Figure R30A), indicating a positive modulation of the Wnt/β-catenin pathway activity. We also evaluated by Western blot the levels of NOTCH1 intracellular domain (NICD1), which is released when the pathway is activated [337]. NICD1 increased in DPPA3-OE SW1222 cells (Figure R30B), suggesting a higher NOTCH signalling activity. Altogether, these data gave us some hints regarding a possible regulation of *LGR5* and *OLFM4* expression by DPPA3 through the modulation of Wnt/β-catenin and NOTCH signalling pathways respectively.

PcG proteins also play a crucial role regulating cell fate, in this case by generally locking genes into a silenced state [338]. In fact, the activity of certain PcG proteins is required to maintain cell lineage identity at the colonic stem level by sustaining Wnt/β-catenin activity [339]. Two main polycomb repressive complexes (PRCs) are present in mammals: PRC1 and PRC2 [340,341]. To evaluate if DPPA3 modulated PcG activity, we assessed levels of monoubiquitination at lysine 119 of H2AK (H2AK119ub1) mediated by the repressive RING1 component of the PRC1 complex [340]. We analysed by immunofluorescence H2AK119ub1 levels in SW1222 cells grown subcutaneously as tumour xenografts. Depletion of *DPPA3* resulted in lower amounts of H2AK119ub1, suggesting a reduced RING1 activity (Figure R30C). In line with this, gene expression analysis showed that knockdown of *DPPA3* in SW1222 megacolonies increased the expression of genes repressed by PcG proteins (Figure R30D), suggesting that DPPA3 could eventually modulate the activity of PcG proteins through unknown mechanisms.

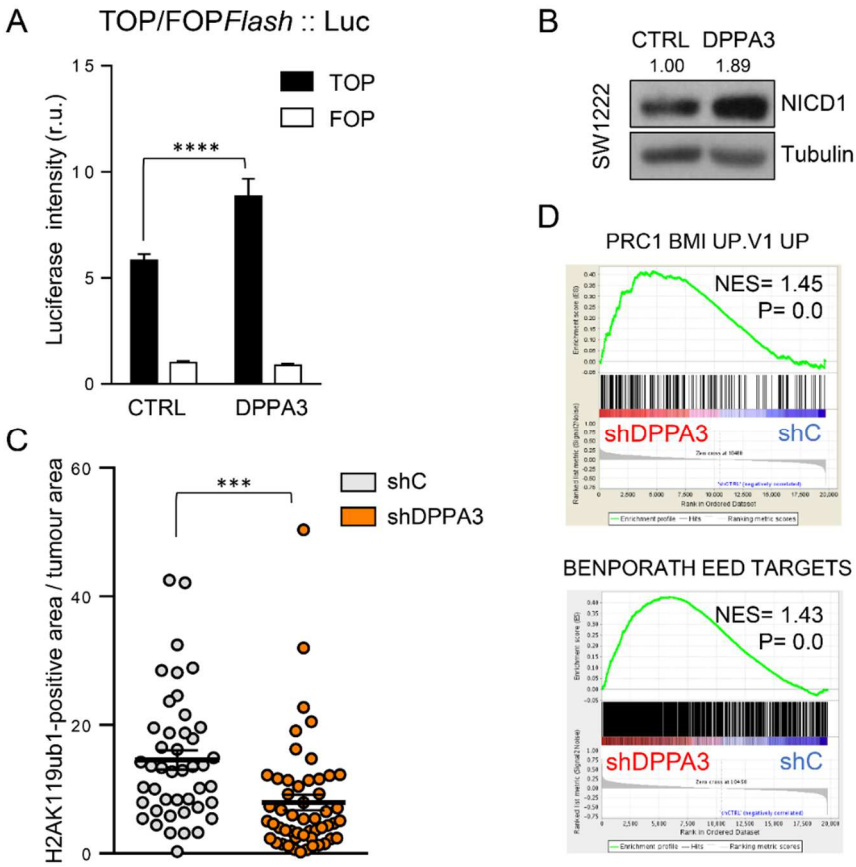


Figure R30. DPPA3 overexpression enhances NOTCH and Wnt signalling while *DPPA3* knockdown affects the activity of PcG proteins. (A) Normoxic HEK293T cells were co-transfected with either pCMV6-Entry (CTRL) or pCMV6-DPPA3 (DPPA3) plasmids together with the wild-type (TOP*Flash*) or mutated (FOP*Flash*) β-catenin/TCF reporter plasmids and luciferase activity was measured 48 hours later. Experiments were performed in triplicates, and data represent the average of three experiments ± SD. Values were expressed relative to cells co-transfected with the pCMV6-Entry and FOP*Flash* constructs. r.u., relative units. ****$P \leq 0.0001$ of 2-way ANOVA. (B) Western blot showing total NCID1 levels of normoxic control (CTRL) and DPPA3-OE (DPPA3) SW1222 cells. Tubulin was used as loading control. NCID protein levels were normalized to tubulin and expressed relative to control condition. (C) H2AK119ub1 evaluation by immunofluorescence analysis in shDPPA3 and shC SW1222 subcutaneous xenografts (n = 6 tumours per condition). shDPPA3 clone 2 cell line was used in this analysis. Five images were taken per section and each dot represents an image. Data are represented as mean ± SEM. ***$P \leq 0.001$ of unpaired two-tailed t test. (D) GSEA plots showing enrichment of the indicated gene sets in the expression profile of shDPPA3 versus shC SW1222 megacolonies. NES, normalized enrichment score; P, p-value.

In summary, these results indicate that DPPA3 promotes a stemness status through the regulation of different programs implicated in cell fate determination and differentiation including Wnt, NOTCH and PcG pathways.

DISCUSSION

1. Molecular determinants of slow-cycling and stem cell phenotypes

Cancer is a prevalent disease in western countries and tumour recurrence is a frequent complication leading to advanced metastatic stages and patients´ death. Effective standard-of-care chemotherapy can select the most resilient tumour cells that finally form the undetectable minimal residual disease (MRD). MRD can eventually regrow after a prolonged period of clinical dormancy provoking patients´ relapse [47]. Slow-cycling or dormant tumour cells are considered to be responsible of such recurrence given their enhanced chemoresistance and ability to remain latent over long periods of time. Unfortunately, no effective therapy has been developed for eradicating dormant tumour cells yet. In response, our team has studied the biological determinants of cancer cell dormancy aiming to identify new drug targets for eliminating these chemoresistant cells. Our results confirmed that *slow-cycliness* can be regarded as a reversible process in which cancer cells, under the pressure of certain environmental cues such as lack of nutrients, low oxygen levels or chemotherapy, adopt an undifferentiated phenotype, stop proliferating and exhibit an increased drug resistance. Most importantly, we identified a core set of factors overexpressed in SCCC that could be key regulators of their distinctive phenotype [269].

Our group recently demonstrated that TET2 epigenetic enzyme is crucial for SCCC survival [269] whereas NANOG controls their self-renewal capacity (unpublished data). These and other pluripotency factors such as SOX2, OCT4 and PRDM1 were also highly expressed in SCCC. Interestingly, some of these factors that are central for the self-renewal and pluripotency of embryonic stem cells are also instrumental in the plasticity of cancer cells when acquiring a stem cell-like phenotype. The expression and activity of these factors has been described as relevant for prostate, breast, lung, bladder, colorectal and renal cancer cells among others. Furthermore, gene expression studies from patients´ tumour samples showed that transcription signatures associated with stem cell pathways identify aggressive cancers and inform on poor outcomes to therapy [342,343].

In addition to these pluripotency factors, SCCC also showed high expression of genes related with germ cell specification such as *APOBEC1*, *EHMT2*, *GDF3* or *DPPA3*. We could even detect elevated expression of markers distinctive of spermatogenesis such as *ACR* or *ACRBP* (unpublished data). This data suggested that SCCC were executing a gene expression program distinctive of primordial germ cell specification typical of early embryonic development. Indeed, for those cancers of germ cell origin like TGCTs, molecular and cellular

events that occur during tumour formation, including the expression of ESC-specific pluripotency factors, recapitulate human fetal development [344]. Actually, the genomic region 12p13 that harbours *DPPA3* together with other pluripotency genes such as *NANOG* and *GDF3* is amplified in the seminoma class of TGCTs [254,267]. Moreover, we corroborated that *DPPA3* was highly expressed in prostate tumours and seminomas when compared with control normal tissue.

All these results encouraged us to select *DPPA3* gene for studying its potential role determining the phenotype of SCCC and its relevance in cancer development. We first observed that *DPPA3* expression was very low in different CRC cell lines as expected for cells of somatic origin. In fact, high DPPA3 levels and epigenetic activity are only detected during embryonic development in a context of multipotency [231]. Even though its residual expression, we observed that *DPPA3* gene silencing by knockdown or CRISPR technology promoted phenotypical changes in CRC cells. We concluded that despite its low expression, DPPA3 is able to control the methylation status of the cancer genome, the post-translational modification of histones in the chromatin and the gene expression program of cancer cells.

2. DPPA3 regulates the hypoxia program

Our group described the whole gene expression profile distinctive of SCCC and their positive enrichment in genes responsible for hypoxia response [269]. We now show the preferential location of SCCC in regions devoid of high oxygen concentrations when CRC cells grow as 3D megacolonies. These observations are in line with the data presented by Fluegen and co-workers, suggesting that primary tumour hypoxic microenvironments induce a dormancy program that gives rise to a subpopulation of chemoresistant dormant tumour cells [345].

Importantly, we showed that DPPA3 induces the expression of many hypoxia-related genes that are also highly expressed in SW1222 CRC-SCCC. Furthermore, we observed a stabilization of *DPPA3* mRNA and higher protein levels in hypoxic cells. This phenomenon has been recurrently described for many important components of the hypoxia response (e.g. VEGFA, EPO and GLUT1) [300–302,346]. Altogether, we revealed DPPA3 as a hypoxia-induced factor with a potential role in the execution of the hypoxia program.

mTORC1, a central component of the mTOR signalling and fundamental for cell growth and proliferation [168], is shut down in hypoxic cells as its repression is necessary for the survival of cells under low-energy circumstances [110]. DPPA3 increased the expression of *PRKAA1*, involved in cellular energy sensing and capable of repressing mTOR pathway [176,177]. In fact, we observed that DPPA3 overexpression reduced the amounts of phosphorylated S6

ribosomal protein as a direct target of mTOR signalling similarly to those detected in hypoxic cells. In order to continue synthetizing proteins necessary for their survival, hypoxic cells must use alternative pathways to cap-dependent translation. The expression of *EIF4G1*, *EIF4G2* and *EIF4EBP1* has been reported to be involved in a hypoxic-mediated switch from cap-dependent to cap-independent mRNA translation [187,188]. Both *EIF4G1* and *EIF4G2* were upregulated in DPPA3-OE SW1222 cells grown as tumour xenografts while *EIF4EBP1* expression was unaltered. Hypophosphorylated 4E-BP1 (*EIF4EBP1* gene) inhibits cap-dependent translation by binding EIF4E [347]. mTOR signalling phosphorylates and inhibits 4E-BP1 *in vitro*. Moreover, 4E-BP1 can trigger hypoxia-mediated inhibition of cap-dependent mRNA translation [187]. Therefore, it would be interesting to evaluate the phosphorylation status of 4E-BP1 in the presence of DPPA3. Another option could involve the elegant approach used by Braunstein and colleagues to determine the switch from cap-dependent to cap-independent initiation of translation [187]. In this experiment, they used plasmids expressing bi-cistronic mRNA reporters containing a Renilla luciferase controlled by a cap-dependent sequence and a downstream Firefly luciferase under an IRES-dependent sequence. Quantification of the two different forms of luciferase in our DPPA3-OE cells could help us understand the mechanisms by which these cells are able to sustain protein synthesis. In summary, the upregulation of *PRKAA1*, *EIF4G1* and *EIF4G2* suggests that DPPA3-OE tumour xenografts use alternative mechanisms to cope with the limitation of the canonical protein translation as it occurs in hypoxia. However, further studies are needed to validate this hypothesis.

Since HIF1α transcription factor orchestrates the cellular response to hypoxia we decided to study its potential crosstalk with DPPA3. HIF1α is continuously synthetized but very rapidly degraded in the presence of oxygen. Its stability is controlled by post-translational modifications, including protein hydroxylation, subsequent ubiquitination and proteasome degradation [105]. The importance of the ubiquitin-proteasome pathway has been demonstrated in pancreatic cancer cells in which *VHL* hypermethylation results in HIF1α stabilization and nuclear translocation, leading to the upregulation of its target genes such as *CA9* and *SLC2A1* [348]. We detected a down-modulation of key components of the ubiquitin-proteasome pathway (*CUL2* and *UBE2D1*) in DPPA3-OE cells regardless of oxygen concentrations. In fact, DPPA3 overexpression reduced HIF1α ubiquitination in SW1222 and HT29 cells after reoxygenation. We actually found that HIF1α protein levels increased in DPPA3-OE SW1222 CRC cells in both normoxic and hypoxic conditions. These results indicate that DPPA3 might reduce the activity of protein degradation complexes and consequently increase the stability of key cell signalling factors, including HIF1α as central for responding to hypoxia.

Despite a lower HIF1α ubiquitination was also observed in DPPA3-OE HT29 cells, HIF1α protein levels increased in normoxic but not in hypoxic conditions. This apparent discrepancy

in hypoxia may be due to negative feedback loops controlling HIF1α levels and activity. Indeed, the HIF system is under the control of a large number of negative and positive feedback systems to ensure a tight and flexible control of its signalling activity in response to hypoxia [349]. The reduced levels of HIF1α observed in hypoxic DPPA3-OE HT29 cells compared to hypoxic control could be due to a negative feedback mechanism for compensating an exacerbated response to hypoxia. This hypothesis would explain that, in contrast to SW1222 cells, hypoxic DPPA3-OE HT29 cells were unable to increase the protein levels of the HIF1α target CA9.

It would be interesting to evaluate in the future the status of some key regulators of these feedback mechanisms such as the levels of microRNA miR-17-92 transcript. This short non-coding RNA molecule with hairpin structure targets *HIF1A* mRNA resulting in its translational repression [350]. Other negative feedback regulation mechanisms involve the induction of *EGLN1* and *VHL* by HIF1α [349,351]. We observed that the transcription of *EGLN1*, *VHL* and another component of the oxygen-dependent degradation complex of HIF1α, *ELOC*, incremented in DPPA3-OE hypoxic SW1222 cells. It would be plausible to argue that the increased expression of these genes was a result of HIF1α stabilisation promoted by DPPA3. In this scenario, *ELOC* induction could be involved in the negative feedback circuit of the hypoxia response driven by HIF1α. It has also been reported that aHIF, a HIF1α antisense RNA molecule that contains several putative HREs in its promoter, is induced under hypoxia and destabilizes *HIF1A* mRNA in human lung epithelial cells [352]. This could represent another negative feedback mechanism responsible for the lower *HIF1A* mRNA levels observed in hypoxic control SW1222 cells.

Furthermore, positive feedback loops may also be responsible for DPPA3 capacity of upregulating HIF1α protein levels and activity in cancer cells. Methylation of a CpG dinucleotide located at the HRE of *HIF1A* promoter prevents HIF1α binding to this sequence and autotransactivation. Demethylation of this region in CRC cell lines and primary tumour samples enables HIF1α binding to its own promoter resulting in the activation of *HIF1A* expression [199]. It is possible that the genomic demethylating activity of DPPA3 affects this and other HREs facilitating HIF1α binding. In fact, DPPA3 increased *CA9* promoter activity and gene expression together with other HIF1α target genes such as *LDHA*, *PDK1*, *SLC2A1* and *VEGFA*, clearly indicating its capacity to enhance HIF1α transcriptional activity.

Promoter methylation has been proven to affect HIF1α transcriptional activity [199]. For instance, HRE hypermethylation blocks HIF complex binding at *erythropoietin* (*EPO*) enhancer and represses its expression [202]. Actually, although CA9 can be used as a marker of hypoxia, it has been reported that its expression does not always correlate with low oxygen levels since

a high CpG methylation status of its promoter can be a dominant repressive mark [122]. Indeed, studies in gastric and renal cancer cells showed that lower methylation levels at *CA9* promoter are associated with increased expression of the gene [134,200]. We observed that in well-established SW1222 tumour xenografts with similar sizes, hence comparable oxygen tensions, CA9 protein levels varied depending on DPPA3 expression. In this case, DPPA3 demethylating activity could increase the availability of permissive HRE ready to bind HIF1α and augment the expression of *CA9* and other hypoxia-responsive genes. Future analyses of the methylation status of *CA9* promoter region in these specimens may help to confirm this hypothesis. In summary, the genome-wide epigenetic effects caused by DPPA3 may reorganize cancer cells' chromatin for triggering an adaptive response to low oxygen levels.

Although *LDHA* and *PDK1* were induced in normoxic DPPA3-OE cells, *SLC2A1*, *CA9* and *VEGFA* were not upregulated in these conditions. It has been reported that several transcription factors can modulate HIF1α-mediated induction of some hypoxia target genes by binding to *cis*-acting motifs present in their promoters. In this regard, Alfranca and colleagues demonstrated that hypoxia induced the phosphorylation of JUN factor by MAPK8, which enables the functional cooperation of JUN and HIF1α for the induction of *VEGFA* [353]. Another report demonstrated that JUN is also required for the hypoglycaemia-mediated induction of *VEGFA* expression [354]. Therefore, we should take into consideration that additional transcriptional factors are required to work in cooperation with HIF1α for the effective induction of specific hypoxia response genes. A reduced activity of some of these additional transcriptional factors in normoxia may explain why DPPA3 overexpression alone in cultured cancer cells could only induce an incomplete hypoxia program without promoting the expression of *SLC2A1*, *CA9* and *VEGFA* genes. In contrast, SW1222 tumour xenografts are superior models recapitulating hypoxic conditions observed in human tumours. In this scenario, the expression of *SLC2A1*, *CA9* and *VEGFA* genes was enhanced upon DPPA3 overexpression, evidencing a more physiological and complete molecular program that drives the hypoxia response *in vivo*.

Methylome analysis corroborated the lower global methylation levels observed in DPPA3-OE cells by dot blot analysis. It has been described that colorectal and melanoma cancer cell lines cultured under hypoxic conditions exhibit a loss of global methylation [204]. Another group demonstrated that in hepatocellular carcinoma, hypoxia induces genomic DNA demethylation in CpG islands by reducing the levels of S-adenosylmethionine, a major biological methyl donor and a critical marker of genomic methylation status [355]. However, the effects of hypoxia on DNA methylation may be tissue-dependent, as other studies have reported a DNA hypermethylation in this scenario. For instance, an increased gene-specific methylation due to the upregulation of *DNMT3B* has been observed in a hypoxia-adapted benign prostate

epithelial cell line [207]. Another publication showed that prolonged hypoxia in cardiac fibroblasts resulted in a global hypermethylation caused by an increased expression of *DNMT1* and *DNMT3B* [208]. TET activity has also been linked to changes in DNA methylation upon oxygen starvation. Thienpont and co-workers reported that hypoxia promotes hypermethylation in some cancer cell lines and that this phenomenon is due to reduced TET activity [206]. Specifically, they observed in eight different cancer cell lines (none of them colorectal) considerable lower 5hmC levels in hypoxic conditions while no significant changes occurred in 5mC. In contrast, another work on neuroblastoma cells demonstrated that hypoxia increases *TET1* expression and transcriptional activity leading to higher 5hmC global levels including some canonical HIF-responsive genes, and that this facilitates DNA demethylation and HIF binding [310]. Therefore, there is still some controversy regarding the alterations in DNA methylation mediated by hypoxia. These may possibly vary depending on the cellular model used and experimental conditions, including oxygen levels, exposure time to hypoxia and type of hypoxic treatments.

Several publications have evidenced that the hypoxic tumour microenvironment might re-model chromatin structure to induce epigenetic changes that alter gene expression in cancer cells [356]. Taking into consideration that hypoxia induced a genome-wide hypomethylation in our model, we wanted to assess if this phenomenon was mediated by a higher activity of TET proteins. Interestingly, DPPA3 overexpression, apart from decreasing global 5mC, caused a drop in 5hmC levels. Surprisingly, although the expression of TETs, mainly *TET1*, skyrocketed in hypoxia, 5hmC levels dropped, resembling 5mC pattern. Altogether, the comparable dynamics of 5mC and 5hmC made us hypothesize that, rather than a modulation of TETs activity, the profound drop in 5hmC levels across the genome was a consequence of low 5mC levels, the substrate of TETs' enzymatic activity.

UHRF1 recruits DNMT1 to chromatin to maintain genome methylation across cell divisions. Interestingly, both factors are repressed in hypoxia in different cancer models [311,312,357]. This is in line with the hypoxia-induced hypomethylation observed in some cancers [204,355]. In concert with these data, we observed that hypoxia caused a decrease in total UHRF1 and DNMT1 protein levels in SW1222 cells. Binding of UHRF1 to histone H3 followed by UHRF1-mediated histone H3 ubiquitination is required for the recruitment of DNMT1 onto replication foci [261,358,359]. It has also been described that DPPA3 binds UHRF1's PHD region, necessary for the recognition of the histone H3 tail, thereby competing with the interaction between UHRF1 and the histone H3 tail [262]. Importantly, UHRF1 can change its partner, switching from DNMT1 to DPPA3 and causing a nuclear delocalization of DNMT1 that leads to a passive genome demethylation [262,309]. Indeed, we validated in CRC cells that overexpression of DPPA3 caused an exclusion of DNMT1 from chromatin while DPPA3 depletion increased

chromatin-bound DNMT1. Although we did not observe any remarkable changes in chromatin-associated UHRF1, it is possible that the elevated levels of exogenous DPPA3 competed with DNMT1 for UHRF1 binding. In this scenario, DNMT1 would lose its scaffold (UHRF1) to bind DNA while DPPA3 would relocate UHRF1 along the genome keeping it associated with chromatin. Further studies including co-immunoprecipitation analysis between DPPA3 and UHRF1 are required to validate this possibility.

In summary, we can speculate that the global low methylation levels observed in normoxic DPPA3-OE cells were due to a nuclear exclusion of DNMT1. Given that the expression of some HIF target genes can be affected by promoter methylation [199], it is possible that DPPA3-mediated hypomethylation facilitates their expression. In this manner, DPPA3-demethylating activity could generate a pseudohypoxic phenotype in normoxia. Under low oxygen levels circumstances, DPPA3 overexpression might enhance the hypoxic program by amplifying DNA demethylation excluding from chromatin the reduced levels of DNMT1. Nonetheless, further experiments are necessary to confirm that DNMT1 loss-of-activity is the central mediator of DNA demethylation caused by DPPA3 overexpression in CRC cells.

DPPA3 associates to chromatin by binding H3K9me2 [256]. In fact, we observed that this epigenetic mark increased in hypoxic SW1222 cells. In this line, it has been described that hypoxia increases the expression of H3K9 methyltransferases like G9a and inhibits H3K9me2 demethylation [195–198]. Other publications have highlighted the importance of a fine-tuned control of histone methylation at specific loci, as the histone demethylase KDM3A can act as a hypoxia signal amplifier by decreasing H3K9me2 levels in the promoters of hypoxia-induced genes [191,192]. Similarly, the hypoxic induction of KDM4B decreases H3K9me3 in the promoters of HIF1α-induced genes and positively correlates with *CA9* expression [81,191,193]. Although H3K9me2/me3 marks have been implicated in heterochromatin formation and gene silencing, some controversial results indicate that gene repression may not always occur. A large-scale analysis performed in different human and mouse cell lines suggested that H3K9me3 is enriched at active gene promoters [360]. In addition, Barski and co-workers demonstrated a modest correlation between H3K9me3 and H3K9me2 levels and gene silencing, since they detected highly localized H3K9me3 peaks in some active genes [361].

Given the increased H3K9me2 levels as a consequence of hypoxia and DPPA3 overexpression together with a genome-wide hypomethylation, we propose two possible epigenetic mechanisms by which DPPA3 might regulate the hypoxia program (Figure D1): (i) a *cis*-regulation of gene expression in which DPPA3 binds H3K9me2 marks at HREs inducing DNMT1 chromatin exclusion and a consequent demethylation of these regions. The second mechanism involves (ii) a *trans*-regulation of gene expression where DPPA3 associates with

other genomic regions and sequesters UHRF1. As a consequence, UHRF1 would not be able to bind H3K9me2 marks nearby HRE to recruit DNMT1. In either case, these effects would ultimately promote a global exclusion of DNMT1 from chromatin, affecting the methylation status of HREs.

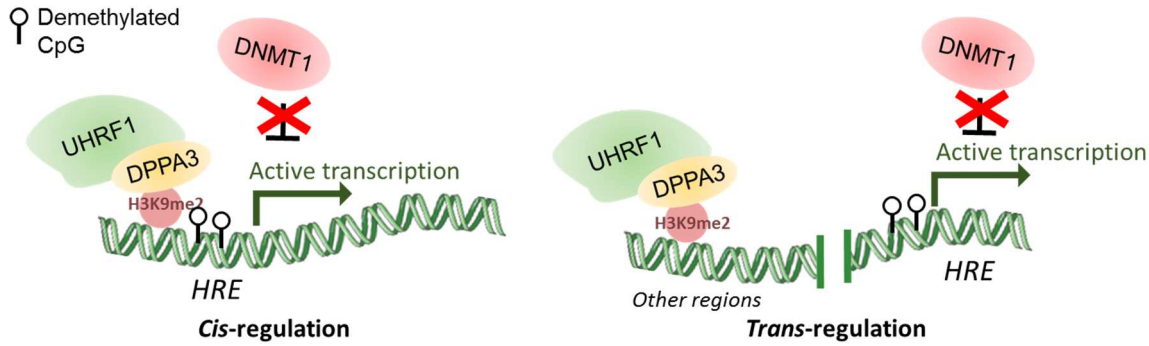


Figure D1. *Cis* and *trans* models mediating DNMT1 exclusion from HREs. DPPA3 binds H3K9me2 and hijacks UHRF1, which can no longer recruit DNMT1 to maintain DNA methylation. This can take place in HREs (left) or in other genomic regions (right). In either case, it might ultimately affect the methylation status of HREs and lead to expression of genes involved in the hypoxia program.

To discern if any of these possibilities takes place in our model, we should first study H3K9me2 and DPPA3 distribution across the genome by chromatin immunoprecipitation-sequencing (ChIP-seq) experiments. In this way, we could identify those hypoxia target genes potentially regulated by DPPA3 bound to H3K9me2 in their regulatory regions. Integrative analysis of ChIP-seq experiments and the gene expression profiles associated to DPPA3 overexpression would be helpful in clarifying DPPA3 activity over the genome.

It is widely accepted that hypoxia provokes a reorganization of chromatin structure by epigenetic mechanisms affecting histones and DNA. Our analyses evidenced that DPPA3 overexpression exerted global effects on genome methylation and showed that an elevated percentage of hypomethylation events occurred in gene bodies. Lienert et al. demonstrated that during neuronal differentiation, a gain of H3K9me2 localized in gene bodies did not necessarily coincide with lower transcription of the corresponding gene [362]. There is the possibility that hypoxia and DPPA3 overexpression result in the accumulation of H3K9me2 preferentially at gene bodies of hypoxia-responsive genes. This would promote the hypomethylation of the neighbouring regions due to DNMT1 exclusion.

We observed that most hypermethylation events occurred at inter-genic regions. This might be in line with the observation that DPPA3 is necessary for protection of DNA methylation at specific regions during embryonic development, some of which are located in inter-genic regions [363]. On the other hand, as we mentioned above, demethylation of promoter regions is

crucial to allow gene transcription of some hypoxia-inducible genes. Promoter hypomethylation observed in DPPA3-OE cells may affect the expression of several genes, affecting different cellular processes like the hypoxic response. However, a deeper characterization of the affected regions is needed to understand the biological significance of this phenomenon. In addition, many HIF-binding sites (around 30%) are located at long chromosomal distances, even more than 10 kb away from hypoxia inducible transcripts [107,108]. This scenario challenges the identification of those HIF1α target genes regulated by DPPA3 activity, since their epigenomic regulation might be more complicated than just checking the promoter methylation status.

Our experiments with tumour xenografts helped to confirm that in a more physiological model DPPA3 is actually relevant for enhancing the molecular response to hypoxia. In this sense, DPPA3 reduced the expression of *CUL2* and *UBE2D1*, impairing the activity of the VHL-mediated degradation circuit of HIF1α. As a result, HIF1α might become less ubiquitinated and degraded. In this manner, DPPA3 would contribute to manage the hypoxia response by stabilizing HIF1α protein levels in growing tumours. In addition, the genome-wide demethylation caused by DPPA3 could increase HRE availability for HIF1α binding, promoting the expression of hypoxia target genes. VEGFA is a HIF1α target and central pro-angiogenic factor [209,210] that was induced in DPPA3-OE tumour xenografts. Interestingly, we observed an increased number of CD31 endothelial cells in DPPA3-OE tumour xenografts, suggesting an enhanced angiogenesis that could be promoted by an active DPPA3/HIF1α/VEGFA axis. Angiogenesis is a hallmark of hypoxic tumours, in which a subset of hypoxia-resistant cancer cells can lead to an increased blood vessel formation, aggressiveness and chemoresistance. However, in many cases, the angiogenic process in pathological conditions like tumour progression may lead to advanced but dysfunctional vascularization, creating a situation where hyper-proliferating cancer cells surpass their blood supply and become hypoxic [209,210]. Therefore, a deeper analysis on the quality of vasculature's structure would help us to elucidate whether DPPA3 overexpression contributes to the formation of new functional blood vessels or, rather, a disorganized vasculature unable to support tumour growth.

Our analyses in real patients´ tumours indicated a positive correlation between *DPPA3* and *CA9* expression equivalent to that observed in our cell and mouse models. Our results also showed that the expression of *DPPA3* mRNA in human tumours is generally below detection levels when using conventional technologies such as microarrays or qRT-PCR. Fortunately, the use of the nCounter Platform significantly increased the sensitivity to detect *DPPA3* transcripts. These showed a positive correlation with CA9 protein levels in tumour samples from CRC patients. In addition, we could further validate this observation in prostate cancer

primary tumours and metastases. Interestingly, it was in prostate cancer, which is associated with a hypoxic phenotype [364], where *DPPA3* expression was high enough to be detectable using conventional microarrays. These observations suggest that our experimental results demonstrating the central role of DPPA3 in hypoxia could also be relevant for the clinical evolution of patients suffering different types of solid tumours.

In conclusion, this study pioneers the importance of DPPA3 in tumour hypoxia. In brief, DPPA3 accelerates the adaptive cell response against low oxygen microenvironments via an increase of HIF1α protein levels and transcriptional activity together with a global genome hypomethylation. These results also support other publications that remark the importance of epigenetics in the modulation of tumour cells' behaviour under hypoxic stress. Nevertheless, much work remains to be done for elucidating the exact molecular mechanisms by which DPPA3 sensitizes CRC cells to low oxygen conditions and if this phenomenon occurs in other cancer types.

3. DPPA3 as an inductor of dormancy

Slow-cycliness and the capacity to remain viable for long periods of time are the main traits of dormant tumour cells. DPPA3-OE cells exhibited a transcriptional program similar in some aspects to that observed in SCCC, including a generalized shutdown of genes involved in cell cycle progression [269]. Hypoxia has been shown to reduce cell proliferation as an early response to oxygen deficit in the tumour. Hence, cancer cells in hypoxic areas are usually dormant but remain viable for prolonged periods of time [73]. Some groups have demonstrated by immunocytochemical techniques that hypoxic tumour cells are negative for proliferation markers [48]. Interestingly, DPPA3 overexpression decreased BrdU incorporation into DNA in SW1222 cells, although this effect was less potent than hypoxia.

Some studies have shown that hypoxia causes a disproportionately long G1 phase or arrest in G0/G1 phases of cell cycle [48]. CDKN1B is known to inhibit G1/S transition in hypoxia, although there is some controversy about whether it does so in a HIF1α-dependent or independent manner [365–367]. However, *CDKN1B* expression was not upregulated in DPPA3-OE megacolonies, indicating that the decreased proliferation observed in these cells may be independent of this mechanism. Nevertheless, we observed that DPPA3 overexpression reduced the phosphorylation of histone H3 at Ser10, a marker of entry into mitosis[317], in SW1222 cells. In contrast, we observed opposite results in HT29 cells. These apparently controversial data might be explained by the mutational status of *TP53* in these cells lines. While SW1222 cells are *TP53*-WT, HT29 cell line carries a mutant form of the *TP53* gene that affects the response to the cytotoxic action of certain treatments [368,369]. TP53 controls G2

arrest prior to entry into mitosis [370]. It is plausible that, via unknown mechanisms, overexpression of DPPA3 in SW1222 cells triggers a TP53-mediated halt in cell cycle prior to mitosis, leading to decreased levels of phosphorylated histone H3. Conversely, in the case of *TP53*-mutant HT29 cells, G2/M checkpoint might not be active causing cell cycle progression until mitosis, in which phosphorylation of histone H3 occurs. Either scenario culminates in the accumulation of cells at the G2/M phases of cell cycle. These observations together with the generalised down-modulation of cell cycle genes similar to SCCC indicate that the mechanism by which DPPA3 hampers cell cycle progression may be comparable to that occurring in SCCC [269].

A proteomic analysis performed in HEK293T cells looking for DPPA3 interactors identified several proteins involved in cell cycle, among which HSPA2 and AKAP8L are required for G2 to M phase transition [263,264]. An interaction between DPPA3 and these factors, which could be assessed by co-immunoprecipation analysis, represents another possibility to explain the G2/M arrest observed. Indeed, HSPA2 has been shown to be upregulated in several human cancers, while its silencing leads to growth arrest and decreased tumourigenic potential [371].

HIFs regulate the activity of MYC, an oncogene central in many cancers that exerts major effects in cell proliferation. It has been demonstrated in different cell models that enforced expression of *MYC* in quiescent cells induces cell cycle entry whereas its inhibition causes cycling cells to stop proliferating [372]. Various groups have shown that HIF1α counteracts MYC effects on proliferation by several mechanisms, one of which consists on binding to MAX and disrupting MYC-MAX complexes [373]. In addition, HIF1α displaces MYC from DNA binding sites and de-represses *CDKN1A* [79]. We observed that a set of MYC target genes was negatively enriched in DPPA3-OE cells. Indeed, several MYC target genes related with active proliferation such as *CCNA2*, *CCNB1*, *CDK1*, *CDK2* and *E2F1* [372] were down-modulated in DPPA3-OE megacolonies. Although our results suggest that DPPA3 could reduce cell proliferation by decreasing MYC activity, more work should be performed to prove this hypothesis.

Depletion of UHRF1 has been proven to arrest HCT116 CRC and melanoma cells at G2/M [374,375]. Indeed, UHRF1 promotes cellular proliferation and is currently regarded as an oncogene [376,377]. Tien and colleagues associated the G2/M arrest observed in UHRF1-depleted HCT116 cells with the activation of DNA damage repair pathway [374]. However, unaltered p-H2AX levels in DPPA3-OE cells indicated that the G2/M cell cycle arrest observed was not associated to DNA damage. Moreover, DPPA3 promoted the disruption of UHRF1 and DNMT1 without a significant decrease of total UHRF1 protein levels. These results

indicate that the G2/M arrest occurring in DPPA3-OE cells might not be related with DNA damage due a depletion of UHFR1.

DNMT1, UHRF1's partner, is usually silent in dormant cells together with other transcription factor including FOXM1 [59]. In fact, we observed that DPPA3 overexpression decreased the expression of *DNMT1* and *FOXM1* genes. Interestingly, down-modulation of both factors has been detected after treating haematological and epithelial tumour cells with the hypomethylating agent 5-azacytidine [67] while *RARB* and *CDKN1A* were induced via p38 signalling in dormant HNSCC cells [59,289]. These and other dormancy-related factors including *NR1H3* [59], *NDRG1* [326], *CTSD*, *DDR1*, *STAT3*, *BUB1*, *CKS2* and *APEX1* [289] were induced upon DPPA3 overexpression. Indeed, some of them are induced by p38 signalling [59,289,326], which was more active upon DPPA3 overexpression in CRC cells. In addition, we observed that both hypoxia and DPPA3 overexpression incremented p38/ERK signalling activity ratio. Actually, DTCs originating from hypoxic microenvironments in primary tumours carry a dormancy signature that includes high p-p38/p-ERK, while cancer cells originating from normoxic regions do not [378]. It has also been demonstrated in HNSCC, breast, prostate, melanoma, and fibrosarcoma cancer models that high p38/ERK activity ratio predicts whether cells proliferate or enter a state of dormancy [61,65].

Altogether, our observations reveal DPPA3 as a new player involved in cancer cell dormancy. However, it is not clear yet to what extent the activation of the hypoxic response and the genome-wide demethylation promoted by DPPA3 may account for such phenotype.

As mentioned before, most hypomethylating effects mediated by DPPA3 occurred in gene bodies, which, unlike promoter methylation, negatively correlate with gene expression as they are associated with a disruption of chromatin structure [379]. Gene body hypomethylation together with hypermethylation of gene promoters in DPPA3-OE cells could be responsible for gene silencing, including regions necessary for cell cycle progression. A deeper evaluation of the differentially methylated genomic regions would help us to elucidate whether epigenetic remodelling mediated by DPPA3 is responsible for gene silencing.

One of the most striking phenotypes mediated by DPPA3 was the abrogation of tumour growth in subcutaneous xenografts and the reduction of primary tumour size in orthotopic xenograft models. The decreased BrdU incorporation into DNA together with a cell cycle arrest at G2/M phases caused by DPPA3 overexpression observed *in vitro* suggest the existence of a cell intrinsic mechanism that may account for the effects observed *in vivo*. In addition, DPPA3 overexpression reduced the number of total metastatic lesions detected in lungs in our orthotopic mouse models of CRC. Interestingly, DPPA3 increased the proportion of small in detriment of the large metastatic foci. Different mechanisms may be involved in such

phenotype. On one hand, since DPPA3-OE primary cecum tumours were much smaller, many less cells were given the opportunity to escape to the bloodstream and colonize distant organs. The slower proliferation imposed by DPPA3 could also explain the smaller size of the detected metastatic lung foci. On the other hand, an increasing body of data has evidenced that pre-invasive lesions that are not fully stablished can undergo EMT and disseminate at early stages of tumour formation. In addition, it has been described that tumour hypoxic cells dramatically reduce migration speed compared to normoxic cells [378]. One might speculate that the enhanced hypoxic program promoted by DPPA3 overexpression would not only reduce primary cecum tumour xenograft growth, but also induce a slow migratory phenotype causing that less cells were able to reach distant metastatic organs during a limited period of time.

DTCs can remain as a small group of latent cells that eventually switch to a proliferative phenotype and generate a newly detectable metastatic lesion in patients [72,380]. Indeed, the persistence in a dormant state permits DTC to evade treatments contributing to late recurrence of disease [271]. The future use of our orthotopic mouse models might help us to understand the potential role of DPPA3 on cancer cell dormancy as a phenotype responsible for an intrinsic chemoresistance during metastatic relapse.

4. DPPA3 and chemoresistance

While overexpression of DPPA3 resulted in the induction of drug resistance genes, some of which were upregulated in SCCC, its depletion sensitized SCCC to chemotherapy. DPPA3 induced the expression of genes encoding for ABC transporters (*ABCB1*, *ABCC2-4*) involved in multidrug resistance [381]. Importantly, HIF1α upregulates *ABCB1* [228,229], which decreases the intracellular concentration of chemotherapeutic agents and has been linked to hypoxia-induced drug resistance in glioma, gastric, breast and colon cancer cells [222]. These results were in line with the enhanced chemoresistance of DPPA3-OE HT29 tumour xenografts. Future analyses will clarify whether HIF1α is involved in DPPA3-mediated induction of some of these detoxifying genes.

While DPPA3 overexpression in tumour xenografts conferred resistance to oxaliplatin and 5-FU, it did not protect against irinotecan treatment. It has been described that irinotecan blocks HIF1α stabilization and induces massive death of hypoxic CRC cells *in vitro* whereas it reduces tumour growth *in vivo* [382,383]. This data might explain the specific sensitivity of DPPA3-OE tumours to irinotecan. However, we observed some degree of resistance to the cytotoxic agent in DPPA3-OE tumour xenografts compared to controls. Nevertheless, DPPA3-OE tumour xenografts regrew more rapidly after an effective irinotecan treatment and DOX withdrawal (exogenous DPPA3 silencing), possibly due to an increased proportion of

chemoresistant SCCC. Indeed, we found that DPPA3 overexpression increased the number of BrdU-retaining cells in tumour xenografts after irinotecan treatment. In this line, we previously demonstrated that oxaliplatin, another CRC standard-of-care chemotherapy, enriches the proportion of resistant SCCC in tumour xenografts, and that these cells may be responsible for tumour relapse [269].

In line with the possible role of DPPA3 in clinical chemoresistance, stage II/III chemo-treated CRC patients with primary tumours highly expressing *DPPA3* showed a shorter PFS. Since *DPPA3* expression in tumours is generally very low and its detection is challenging, the use of a DPPA3 gene signature was easier to translate into the clinics. Indeed, it predicted a poorer PFS in 5-FU-treated stage II and III CRC patients while it did not show any predictive value in non-treated CRC patients of the same cohort.

Mutations of oncogenes such as *KRAS* or *BRAF* diminish time to recurrence in CRC patients within a short periods of time [384,385]. In contrast, DPPA3 signature predicted shorter PFS only after long latency periods (5-12 years) in the cohort analysed. These results indicate that DPPA3 might commit some cancer cells in growing tumours towards a slow-cycling chemoresistant phenotype by imposing a distinctive gene expression program. These DPPA3-driven SCCC could represent a source of MRD and a seed of a future relapse. Future analyses of larger cohorts should be performed to consolidate the value of a DPPA3 signature as a clinical biomarker and specially to predict long-term relapse in CRC and other types of solid tumours.

Long-term dormancy has been described in some types of tumours including ER+ breast and prostate cancers [52]. Kim and colleagues generated a signature to help identifying those disseminated breast cancer cells prone to undergo dormancy [289]. The signature was composed of data generated from two studies: one using human cancer cells that remained latent as a consequence of angiogenic dormancy [386], and another analyzing dormant HNSCC cells with activated p38 signalling and low uPAR levels [59]. When we analyzed CRC chemo-treated patient's tumours, we observed a weak positive correlation between our DPPA3 signature and the dormancy signature generated by Kim et al., suggesting a potential role of the epigenetic factor controlling long-term dormancy. However, although cancer cells enriched in the DPPA3 signature may remain in a dormant state that shields them from chemotherapy, the poorer PFS after long latency periods observed in patients classified as DPPA3 Sig high suggests that these tumours can eventually revert its dormant state. Indeed, collective evidence indicates that the reversibility of the dormancy program in DTCs may be controlled by epigenetic mechanisms [66]. In this sense, DPPA3 might be a relevant factor ruling the bi-directional transitions between dormancy and proliferation states.

Fluegen et al. showed that primary tumour cells exposed to hypoxic microenvironments upregulate hypoxia and dormancy genes in a reversible manner. They demonstrated that hypoxia-imprinted DTCs residing in lungs were more prone to enter a dormant state and evade chemotherapy [345]. Authors also suggested that the dormancy-like response is more long-lived than the hypoxia program or that additional hypoxia-responsive pathways might be activated concomitantly with dormancy. They proposed that primary hypoxic microenvironments give rise to a subpopulation of dormant DTCs that evade therapy and that may represent the source of disease relapse and poor prognosis associated. Low oxygen levels upregulate DPPA3, which in turn, sensitizes cells to the hypoxic response, resulting in a positive feedback mechanism that may eventually drive cells into dormancy. In addition, DPPA3 regulates p38/ERK signalling ratio and the expression of dormancy markers while maintaining a pseudohypoxic response under normoxic conditions. Therefore, DPPA3 could be a plausible candidate that helps driving and maintaining dormancy of post-hypoxic DTCs. That would occur by reprogramming cells' epigenome during hypoxia exposure in primary tumours, contributing to therapy evasion and tumour recurrence. Additional experiments to better characterize the impact of DPPA3 on DTCs chemoresistance will help us to validate this hypothesis.

5. DPPA3 modulates cell differentiation

Aside from its effects in response to hypoxia, tumour growth and drug resistance, modulation of DPPA3 expression affected cell lineage commitment. This observation caught our attention, as this factor is expressed during early embryonic development in the pre-implantation embryo and in PGCs [231]. Indeed, DPPA3-OE SW1222 megacolonies showed a gene expression profile enriched in gene sets related to PGCs.

DPPA3 overexpression skewed cell fate towards an undifferentiated phenotype enhancing the expression of the *stemness* genes *OLFM4* and *LGR5*, the latter being a Wnt target and a selective marker of CBCs at the stem cell zone in the colonic crypt [12]. Wnt signalling is vital for self-renewal of intestinal stem cells at the base of the crypt [28]. Indeed, it has been shown that LGR5-positive cells can generate self-renewing intestinal organoids *in vitro* [14]. We observed an increased Wnt activity upon DPPA3 overexpression, reinforcing the idea that DPPA3 maintains a *stemness* status in cancer cells. *OLFM4* is another Wnt target and a marker of CBCs [15]. The expression of *OLFM4* in CBCs is directly dependent on NOTCH signalling, the inhibition of which causes the differentiation into secretory cells in mice, including goblet cells [336]. Actually, this pathway is crucial for cell fate determination and regulates stem cell behaviour with widespread roles in a variety of tissues, including the intestinal crypt [387,388]. Of note, DPPA3-OE SW1222 cells showed an enhanced NOTCH

signalling evidenced by higher levels of NICD. These results pointed towards a positive regulation of NOTCH and Wnt signalling pathways by DPPA3, vital for the maintenance of *stemness* features and that could be linked to the undifferentiated phenotype observed in these cells.

GDF3 was also induced by DPPA3 in SW1222 megacolonies. This gene encodes a TGFβ superfamily ligand that is expressed in mouse and human ESCs and regulates pluripotency features [389]. In addition, GDF3 has been shown to inhibit proliferation of MCF7 cells while its knockdown enables breast tumour progression [390]. Altogether, the positive regulation of several stemness factors by DPPA3 in SW1222 megacolonies may underpin the narrow connection between DPPA3 and a pluripotency circuit that eventually governs cell fate. *DPPA3* knockdown in megacolonies and xenografts reinforced this hypothesis as it resulted in increased number of mucin-positive cells *in vitro* and *in vivo* and higher expression of *MUC2* and *GCNT3*, involved in mucin biosynthesis.

Some publications have uncovered the importance of HIF during embryonic development. For instance, genetic ablation experiments in mice have demonstrated that the different HIF subunits are essential for embryonic development and survival [391]. A variety of experimental systems has evidenced that hypoxia regulates the proliferation and differentiation of ESC, leading to the idea that HIF can regulate the multipotency attributed to stem cells. In cancer, it has been proposed that HIF activity can promote the adoption of stem cell characteristics by differentiated hypoxic tumour cells [392]. For example, hypoxia confers a more immature phenotype of human neuroblastoma and breast tumour cells [393]. Curiously, robust expression of HIF1α has been detected in mouse pre-implantation embryos and in PGCs, in a fashion similar to DPPA3. Furthermore, HIF1α is expressed in adult testis and mature oocytes [395]. These observations highlight the relevance of HIF1α in a context dependent of the coordinated network of pluripotency factors.

A joint action of HIF1α and factors determinant for cell fate commitment can occur in other cellular contexts. It has been reported that hypoxia blocks the differentiation of myogenic cells and primary neural stem cells in a NOTCH-dependent manner [396]. The same authors showed that hypoxic treatment in rat embryonic carcinoma cells induced the transcription of NOTCH target genes through the stabilisation of NICD. Stabilized NICD can be recruited to a DNA-binding complex together with HIF1α to regulate gene expression.

Wnt signalling and hypoxia have also been studied in stem cells, in which a physical interaction between β-catenin and HIF1α can take place under oxygen starvation to enhance cell survival [119]. The connection between hypoxia, NOTCH, Wnt signalling and dedifferentiation could be of interest in our model, in which we observed a positive regulation of the different processes

upon DPPA3 overexpression. Determining the degree of crosstalk among them could be relevant to understand the molecular mechanisms mediating the undifferentiated phenotype observed in DPPA3-OE cells.

Finally, we noticed that changes in DPPA3 levels might modulate the activity of PcG proteins, as depletion of *DPPA3* expression induced some of their target genes (generally repressed by PcG proteins) while it decreased H2AK119ub levels, the epigenetic mark catalysed by PRC1 [340]. Of note, Wnt activity, which is elevated in DPPA3-OE cells, has been shown to be sustained by PRC1 in ISCs [339]. Additional experiments should be performed to elucidate whether DPPA3-mediated effects on the epigenome affect the expression or activity of any of the components of PRCs.

6. Representation of the working model

In conclusion, this study uncovers DPPA3 as a novel modulator of the transcriptional program that takes place in cancer cells following exposure to hypoxia. In addition, our data describes a possible role of this factor in regulating entry into dormancy. The gene expression program elicited by DPPA3 might be implicated in chemotherapy resistance and a subsequent cancer relapse. Tumour type and biology might determine whether these effects take place in only a subpopulation of tumoural cells (if the hypoxic environment occurs in specific regions of the tumour mass) or in the whole neoplasm (if the tumour type itself is associated with a hypoxic phenotype). The following diagrams (Figures D2 and D3) are intended to depict the mechanisms by which DPPA3 modulates tumour cells' phenotype.

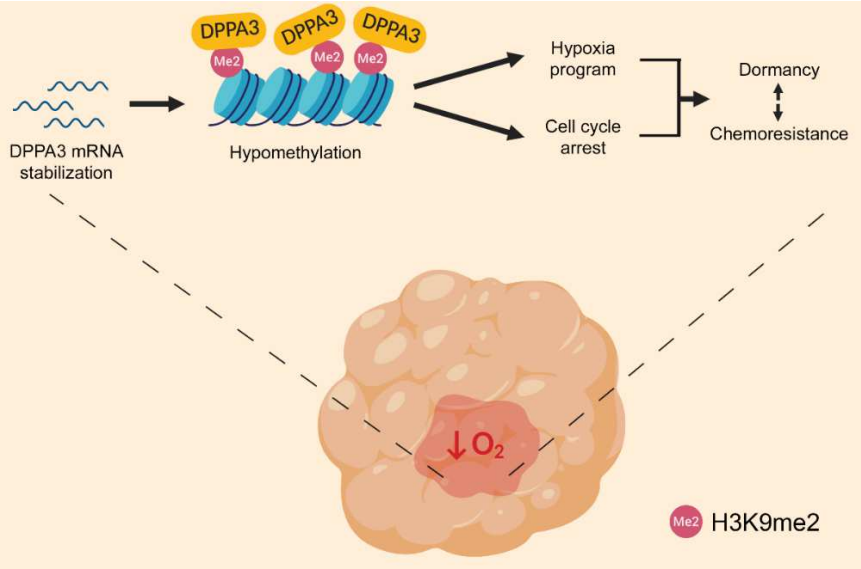


Figure D2. Effects of DPPA3 on dormancy and chemoresistance linked to tumour hypoxia. Hypoxia, either as a feature of the tumour type or as a microenvironmental factor affecting specific regions of the neoplasm, might stabilise *DPPA3* mRNA leading to increased protein levels in the nucleus, which might bind H3K9me2 and promote a genome-wide demethylation. This process may ultimately generate a reversible epigenetic reprogramming that results in an improved adaptive response of tumour cells to hypoxia, entrance into dormancy and chemotherapy evasion.

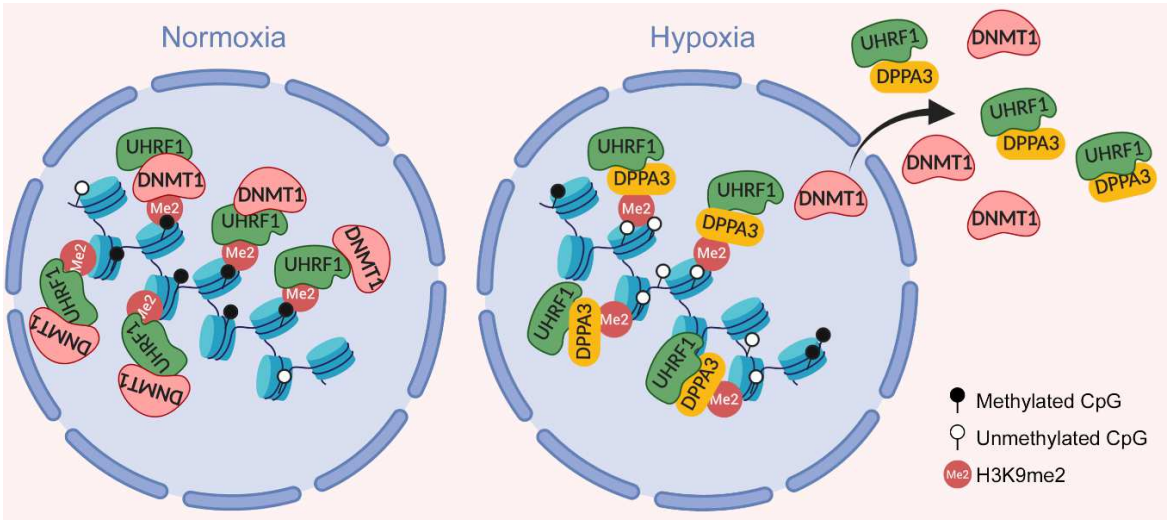


Figure D3. Molecular mechanism involved in DPPA3-mediated genome demethylation under hypoxic conditions. At the molecular level, binding of DPPA3 to H3K9me2 following exposure to low oxygen levels might result in a partner switching: DPPA3 binds UHRF1 resulting in the disruption of DNMT1-UHRF1 complexes. As a result, DNMT1 is excluded from chromatin and a genome-wide demethylation takes place, affecting global chromatin organization and the expression of specific genes. This might cause an enhancement of transcriptional programs involved in the hypoxia response and the silencing of genes necessary for cell cycle progression.

CONCLUSIONS

1. DPPA3 overexpression stabilizes HIF1α protein and induces a genome-wide demethylation enhancing the hypoxia program.
2. The expression of DPPA3 and the hypoxia marker CA9 positively correlate in tumours from colorectal and prostate cancer patients.
3. Overexpression of DPPA3 causes a cell cycle arrest at G2/M phase and modulates the expression of tumour dormancy genes inducing a dormant phenotype in CRC orthotopic primary tumours and corresponding metastases.
4. Overexpression of DPPA3 in CRC tumour xenografts enhances chemoresistance and drives tumour relapse.
5. High expression of *DPPA3* and positive enrichment of its signature in primary tumours from chemo-treated CRC patients predicts shorter disease-free survival.
6. Positive enrichment in the DPPA3 signature of chemo-treated CRC patients' primary tumours is associated with a dormancy gene expression profile.
7. Overexpression of DPPA3 regulates cell lineage commitment of CRC cells maintaining an undifferentiated phenotype characterized by the activation of stem cell pathways such as Wnt and NOTCH.

ANNEX

DPPA3-induced genes								
Gene symbol	shDPPA3		DPPA3 KO		DPPA3-OE 5 days		DPPA3-OE 4h	
	FC	*P*	*FC*	*P*	*FC*	*P*	*FC*	*P*
ADSS	1.09	0.231	-1.25	0.009	1.29	0.002	1.25	0.007
AOAH	-1.30	0.005	1.10	0.295	2.09	0.000	1.22	0.026
BNIP2	1.18	0.039	-1.38	0.001	1.25	0.010	1.27	0.006
CA12	-1.42	0.000	-1.23	0.017	1.32	0.002	-1.02	0.831
CA9	-1.52	0.005	-1.38	0.027	2.06	0.000	1.42	0.013
CLDN1	-1.65	0.000	-1.03	0.692	1.46	0.000	1.54	0.000
CNOT7	-1.15	0.231	-1.31	0.033	1.72	0.000	1.81	0.000
CTSS	-1.56	0.000	-1.11	0.182	2.13	0.000	1.45	0.000
DACH1	1.06	0.358	-1.26	0.002	1.37	0.000	1.27	0.001
DAPK2	-1.26	0.013	-1.36	0.003	1.22	0.027	-1.14	0.142
DPEP1	1.03	0.756	-1.59	0.000	1.41	0.001	1.29	0.010
EFHD2	-1.27	0.029	-1.25	0.050	1.47	0.001	-1.06	0.556
ELF4	1.11	0.254	-1.23	0.047	1.42	0.002	1.35	0.005
EREG	-1.31	0.030	1.08	0.552	1.95	0.000	1.87	0.000
FABP6	-1.18	0.118	-1.40	0.006	1.46	0.002	1.33	0.012
GHR	-1.04	0.709	-1.26	0.049	2.55	0.000	1.34	0.010
GXYLT2	-1.05	0.574	-1.38	0.001	1.93	0.000	1.27	0.008
HAVCR1	1.23	0.186	-1.81	0.002	2.16	0.000	1.67	0.004
HEXB	-1.07	0.463	-1.28	0.015	1.31	0.006	1.21	0.041
HTATIP2	-1.27	0.017	-1.18	0.095	1.75	0.000	1.64	0.000
IARS2	1.01	0.869	-1.27	0.011	1.25	0.012	1.25	0.012
IL37	-1.40	0.030	-1.89	0.001	3.18	0.000	2.04	0.000
ILK	-1.21	0.008	-1.14	0.074	1.20	0.011	1.54	0.000
JUN	-1.17	0.104	-1.52	0.001	1.58	0.000	1.30	0.013
KAT2B	1.00	0.980	-1.29	0.003	1.22	0.010	1.42	0.000
KIT	-1.15	0.358	-1.97	0.001	8.45	0.000	2.08	0.000
L3MBTL3	-1.31	0.008	1.40	0.003	1.34	0.005	1.60	0.000
LGALS1	-1.01	0.963	-1.81	0.000	2.17	0.000	1.35	0.018
LGR5	-1.20	0.024	-1.47	0.000	1.51	0.000	2.02	0.000
LRRC19	-1.18	0.100	-1.36	0.007	1.69	0.000	1.28	0.016
LSR	1.04	0.636	-1.22	0.034	1.33	0.003	1.32	0.004
LY6G6D	-1.64	0.000	1.60	0.000	2.00	0.000	1.63	0.000
MACC1	-1.26	0.002	1.03	0.671	2.01	0.000	1.52	0.000
MAML2	-1.00	0.974	-1.26	0.034	1.30	0.013	1.27	0.022
MYL3	-1.68	0.000	1.28	0.004	1.53	0.000	1.47	0.000
NT5E	1.02	0.843	-1.61	0.000	1.70	0.000	1.45	0.001
NTRK2	-1.30	0.000	-1.29	0.000	1.40	0.000	1.09	0.124
PDLIM1	-1.36	0.016	-1.38	0.017	1.29	0.042	-1.03	0.820
PLS3	-1.09	0.339	-1.33	0.008	1.48	0.000	1.76	0.000
QPCT	-1.39	0.007	1.18	0.159	1.74	0.000	1.35	0.013
SLC39A2	-1.29	0.031	-1.29	0.040	1.74	0.000	1.31	0.021
SORBS1	-1.18	0.084	-1.26	0.027	1.26	0.022	1.25	0.023
SPRY3	-1.24	0.027	1.12	0.254	1.32	0.006	1.32	0.006
TFF1	-1.31	0.007	1.28	0.016	2.17	0.000	2.06	0.000
TGFBI	-1.43	0.000	-1.46	0.000	2.26	0.000	1.62	0.000
TM4SF1	1.09	0.354	-1.34	0.008	1.48	0.001	1.27	0.017
TMEM154	-1.22	0.030	-1.34	0.005	1.25	0.016	1.13	0.151
TMEM211	-1.27	0.042	-1.07	0.552	1.37	0.011	1.53	0.001
VPS37A	1.05	0.692	-1.44	0.011	1.72	0.000	2.02	0.000
ZC3H6	1.02	0.758	-1.38	0.000	1.33	0.001	1.28	0.002

DPPA3-repressed genes								
Gene symbol	shDPPA3		DPPA3 KO		DPPA3-OE 5 days		DPPA3-OE 4h	
	FC	*P*	*FC*	*P*	*FC*	*P*	*FC*	*P*
ABCC5	1.04	0.643	1.38	0.001	-1.27	0.006	-1.24	0.011
AHRR	1.24	0.013	-1.15	0.113	-1.44	0.000	-1.43	0.000
AKR1C2	1.44	0.030	-1.10	0.562	-2.77	0.000	-1.85	0.001
CACNA1E	1.22	0.047	1.45	0.001	-1.54	0.000	-1.28	0.014
CST1	2.00	0.002	-2.47	0.000	-4.56	0.000	-1.91	0.004
CYP4F11	1.00	0.998	1.48	0.000	-2.03	0.000	-1.85	0.000
CYP4F12	1.09	0.383	1.32	0.020	-1.55	0.000	-1.41	0.003
CYP4X1	1.26	0.029	-1.48	0.002	-4.08	0.000	-2.26	0.000
DKK4	1.56	0.007	1.44	0.031	-1.38	0.040	1.05	0.715
FER1L6	-1.03	0.683	1.20	0.016	-1.21	0.010	-1.23	0.006
GAL	1.31	0.016	-1.51	0.001	-2.55	0.000	-1.29	0.022
GEMIN5	1.20	0.031	1.25	0.015	-1.29	0.005	-1.02	0.777
GRHL3	-1.25	0.004	1.27	0.004	-1.46	0.000	-1.42	0.000
HSD3B1	-1.40	0.000	1.35	0.000	-1.42	0.000	-1.51	0.000
JADE3	-1.01	0.924	1.52	0.003	-1.38	0.012	-1.28	0.044
KIF13A	1.37	0.003	1.05	0.634	-1.34	0.005	-1.37	0.003
L1TD1	1.39	0.002	-1.26	0.021	-1.72	0.000	-1.59	0.000
LCE1C	1.22	0.011	1.21	0.017	1.03	0.663	-1.20	0.014
MAOA	-1.09	0.266	1.33	0.002	-1.27	0.006	-1.71	0.000
MPV17L2	-1.01	0.930	1.30	0.017	-1.72	0.000	-1.59	0.000
MYO1A	1.31	0.007	1.34	0.007	-1.37	0.003	-1.30	0.008
PCP4	1.36	0.037	-1.00	0.973	-1.70	0.001	-2.09	0.000
PLA2G2A	1.73	0.000	-2.10	0.000	-9.64	0.000	-2.95	0.000
PRPF3	1.18	0.042	1.41	0.001	-1.32	0.002	-1.28	0.005
PYGO1	1.26	0.000	1.05	0.308	-1.34	0.000	-2.21	0.000
RNF157	1.58	0.003	-1.33	0.056	-1.96	0.000	-2.27	0.000
RYR2	-1.11	0.096	1.39	0.000	-1.44	0.000	-1.25	0.001
SLC9A4	1.46	0.000	1.42	0.001	-1.80	0.000	1.02	0.827
TPM4	-1.05	0.494	1.22	0.018	-1.38	0.000	-1.46	0.000
ZNF101	1.10	0.408	1.35	0.023	-1.42	0.007	-1.62	0.001
ZNF239	1.26	0.019	1.25	0.028	-1.19	0.061	-1.34	0.004
ZNF738	1.09	0.077	1.32	0.000	-1.54	0.000	-1.34	0.000

Table A1. Lists of the 50 DPPA3-induced (first list) and 23 DPPA3-repressed (second list) genes used for the generation of the DPPA3 Signature. Gene symbol, fold-change (*FC*) of each condition (*DPPA3* knockdown, shDPPA3; *DPPA3* knockout, DPPA3 KO; DPPA3 5 days-overexpression, DPPA3-OE 5 days; and DPPA3 4 hours-overexpression, DPPA3-OE, 4h) versus respective controls and p-value (*P*) of each condition are detailed.

SCCC PAN-CANCER - Upregulated						
ABCA5	CFLAR	GPR183	LOC541473	PCDHB18	RNA5SP38	SNORD38B
ACRC	CIR1	HBP1	LRLE1	PCID2	RNA5SP405	SNORD53
ADAM21	CLEC4A	HCG8	LRP2BP	PCMTD1	RNA5SP42	SNORD56B
ADAM32	CLK1	HERC4	LRRC37A	PDC	RNA5SP427	SNORD63
ADCY10P1	CLTC-IT1	HIP1	LRRC8C	PDK4	RNA5SP432	SNORD7
ALS2CR12	CPEB2	HIVEP2	LRRFIP1	PDXDC2P	RNA5SP449	SPDYE1
AMT	CPEB4	HMGB3P22	LRRTM2	PDZK1	RNA5SP450	SPDYE5
AMY1A	CREBRF	HNRNPU-AS1	LSMEM1	PGM2L1	RNA5SP46	SPDYE7P
AMY2B	CRIPAK	HSPA6	LYST	PGM5P2	RNA5SP474	SPDYE8P
ANGPTL1	CRYAB	IFRD1	MIR103A2	PHC3	RNA5SP68	SPINK1
ANKRD12	CSNK1G1	IGHG1	MIR21	PIGT	RNA5SP69	STX19
ANKRD20A12P	CTSK	IGHV3-20	MIR221	PKD1P1	RNA5SP76	TAF1D
ANKRD20A3	CYP3A5	IGIP	MIR222	PLA2G10	RNA5SP85	TAS2R31
ANKRD20A4	DDIT3	IGKV1OR10-1	MIR22HG	PLAC8L1	RNA5SP97	TAS2R4
ANKRD20A8P	DDX60L	IL24	MIR320C1	PLEKHH2	RNU1-16P	TM4SF20
ANKRD20A9P	DEFB109P1B	ITGB8	MIRLET7G	PLGLB2	RNU5B-1	TMEM14E
ANXA1	DGCR11	JMY	MST1	PPARGC1A	RNU6-50	TMEM51-AS1
ARHGEF10	DICER1-AS1	KCNJ13	MUC20	PRKXP1	RNU6-64P	TMF1
ARID4A	DNAJB4	KIAA2026	MYEF2	PTPRR	RNU6-79P	TPTEP1
ATG14	DNHD1	KLHL22-IT1	N4BP2L2	RBM14	RNU6-82P	TRIM52
ATXN1	DSERG1	KLHL24	NAIP	REXO1L1	RNU7-35P	TSSK3
AVIL	ERVH-4	KLHL28	NBEAL1	RNA5-8SP3	RNU7-5P	UCA1
B3GALT2	EVI2B	KMT2A	NFAT5	RNA5SP104	RNY1	UGT2B11
BIRC3	EYS	KMT2C	NKTR	RNA5SP114	RNY1P4	USP17L9P
C10orf118	FAM106A	KMT2E	NOXRED1	RNA5SP123	RNY3P6	VNN1
C11orf34	FAM138A	KRTAP19-1	NPIPA5	RNA5SP129	RPL21P44	WDR52
C1orf132	FAM138C	KRTAP21-2	NPIPB3	RNA5SP161	RPS20P27	WDR96
C1orf162	FAM214A	LEAP2	NPIPB5	RNA5SP166	RRN3	WEE1
C3orf35	FAM223A	LIMD1-AS1	NPIPB6	RNA5SP187	RRN3P1	WNK1
C3orf65	FAM25C	LINC00216	NPIPL2	RNA5SP195	RUNX2	WSB1
C8orf44	FAM63B	LINC00268	OCLM	RNA5SP196	SAMD9	XKR9
C9orf131	FBXW10	LINC00328	OCR1	RNA5SP20	SLC5A3	YPEL2
CA1	FLJ45950	LINC00537	OFD1	RNA5SP209	SLPI	ZC3H6
CAPZA2	GABBR1	LINC00597	OMG	RNA5SP217	SMG1	ZNF117
CBLB	GBP2	LINC00888	OPTN	RNA5SP229	SMG1P1	ZNF224
CCDC11	GCC2	LINC00965	OR11H12	RNA5SP266	SND1-IT1	ZNF292
CCDC144B	GEM	LINC00969	OR2J3	RNA5SP268	SNORA15	ZNF559-ZNF177
CCDC144CP	GOLGA4	LOC100130428	OR2L8	RNA5SP282	SNORA16B	ZNF587
CCDC146	GOLGA6L4	LOC100132099	OR4F21	RNA5SP289	SNORA46	ZNF83
CCDC148	GOLGA6L5	LOC100133299	OR4F29	RNA5SP295	SNORA70G	ZNF841
CCDC162P	GOLGA6L9	LOC100288807	OR4G3P	RNA5SP297	SNORD102	ZNF91
CCRL1	GOLGA8O	LOC100652783	OR7E125P	RNA5SP33	SNORD109B	ZRSR1
CEACAM5	GOLGB1	LOC101060353	OR7E87P	RNA5SP343	SNORD117	
CEACAM6	GPCPD1	LOC101060455	PAR5	RNA5SP345	SNORD13P2	
CEP95	GPR18	LOC101060558	PBOV1	RNA5SP366	SNORD21	

SCCC PAN-CANCER - Downregulated						
AAAS	CENPI	ESYT1	IMP4	MYC	PRC1	SNRNP40
ABCE1	CENPK	EXO1	IMPDH2	NAA10	PRDX3	SNRPA
ACADM	CENPN	EXOSC2	IPO11	NAA15	PRELID1	SNRPB
ACAT2	CENPO	EXOSC4	IPO4	NAE1	PREP	SNRPC
ACLY	CENPV	EXOSC8	IPO5	NARS2	PRIM1	SNRPD1
ACO1	CENPW	EXOSC9	IPO7	NASP	PRKAR2B	SNRPF
ACOT7	CEP55	FABP5	ITGB3BP	NAT10	PRKCSH	SORD
ACTB	CEP78	FADS1	ITPA	NBN	PRKDC	SPA17
ACTG1	CHAF1A	FAM169A	JAGN1	NCAPD2	PRMT1	SPAG5
ACTL6A	CHAF1B	FAM72B	KARS	NCAPD3	PRMT3	SPC25
ACTR1A	CHCHD3	FAM72D	KIAA0101	NCAPG	PRMT5	SPDL1
ADK	CHD1L	FANCD2	KIAA1524	NCAPG2	PRPF19	SPINK4
ADSL	CHEK1	FASN	KIF11	NCAPH	PRPS1	SRM
AGPS	CHEK2	FAU	KIF15	NDC1	PRPS2	SRPK1
AHCY	CHID1	FBL	KIF18A	NDC80	PRR11	SRPRB
AIF1L	CHRNA5	FBXO5	KIF20A	NDUFA11	PSMA4	SSNA1
AKIP1	CIRH1A	FDPS	KIF20B	NDUFA13	PSMA5	SSRP1
AKR1B1	CISD1	FERMT1	KIF22	NDUFA2	PSMB3	ST20
ALDH18A1	CISD2	FH	KIF23	NDUFA9	PSMB6	ST6GAL1
ALDH1B1	CIT	FIBP	KIF2C	NDUFAF2	PSMC3	STIP1
ALDH2	CKAP2	FKBP4	KIF4A	NDUFAF4	PSMD10	STMN1
ALDH7A1	CKAP2L	FKBP5	KIF4B	NDUFB6	PSMD14	STOML2
ALG6	CKAP5	FLOT2	KIFC1	NDUFB8	PSMD2	SUOX
ALG8	CKMT1A	FOXM1	KNSTRN	NDUFV1	PSMD3	SUPT16H
ALYREF	CKS1B	FUCA2	KNTC1	NEDD8	PSME3	SUV39H1
ANAPC15	CKS2	G3BP1	KPNA2	NEK2	PSMG1	TACC3
ANAPC7	CLIC1	G6PC3	KPNA3	NELFCD	PTBP1	TAGLN2
ANKRD32	CLN6	G6PD	KRT8	NHP2	PTDSS1	TALDO1
ANLN	CLPTM1L	GAL	LAP3	NIP7	PTMA	TBC1D4
ANP32A	CLSPN	GANAB	LARS2	NIPSNAP1	PTMS	TECR
ANP32E	CLUH	GART	LAS1L	NIT2	PTPLAD1	TERC
AP1B1	CMBL	GCSH	LBR	NME1	PTRHD1	TESC
AP1S1	CMSS1	GEMIN4	LGR5	NMU	PTTG1	TEX261
APEX1	CMTM7	GEMIN6	LMNB1	NOC2L	PWP1	TFAM
APH1A	COMMD4	GEN1	LMNB2	NOC3L	RAB15	TFB1M
APOBEC3B	COMMD8	GGCT	LOC100996481	NOLC1	RAB1B	TFDP1
APOC1	COPS3	GGH	LOC440311	NOP14	RAB8A	THG1L
APOO	COPZ1	GINS1	LRR1	NOP16	RACGAP1	THOC7
APRT	COQ3	GINS2	LRRC59	NOP2	RAD51	TICRR
ARHGAP11A	COX11P1	GINS4	LSM2	NPM3	RAD51AP1	TIMELESS
ARHGDIA	COX8A	GLO1	LSM3	NR2C2AP	RANBP1	TIMM10
ARL2	CPE	GMDS	LSM5	NSDHL	RANBP3	TIMM13
ARRB2	CPSF1	GMNN	LSM6	NT5DC1	RBBP4	TIMM50
ASF1B	CPVL	GMPS	LSM7	NT5DC2	RBBP7	TIMM8A
ASPM	CSE1L	GOT2	LSMD1	NUCKS1	RBL1	TIPIN
ASRGL1	CSTF2	GPI	LYAR	NUDC	RBMX	TK1
ASUN	CTPS1	GRPEL1	LYPLA1	NUDT15	RCC1	TKT
ATAD2	CTSC	GSR	M6PR	NUDT5	REXO2	TLCD1
ATIC	CWC27	GSS	MAD2L1	NUF2	RFC2	TMA16
ATP5G1	CYC1	GTF3C6	MAK16	NUP107	RFC3	TMEM106C
AURKA	DCK	GTSE1	MANEA	NUP155	RFC4	TMEM141
AURKB	DCPS	H2AFY2	MAP7	NUP188	RFC5	TMEM14C
BANF1P3	DCUN1D5	H2AFZ	Mar-01	NUP205	RFX5	TMEM97
BARD1	DDOST	H2BFS	MBOAT1	NUP210	RFXANK	TMX1

BCAP31	DDT	HADH	MCM10	NUP37	RNASEH2A	TOMM22
BDH2P1	DDX10	HARS	MCM2	NUP43	RNASEH2B	TOMM34
BIRC5	DDX21	HAT1	MCM3	NUP62CL	RNF157	TOP1MT
BLM	DDX23	HAUS1	MCM4	NUP93	RNF5	TOP2A
BLMH	DDX39A	HAUS8	MCM5	NUSAP1	RNU11	TP53
BLOC1S5-TXNDC5	DEPDC1	HDAC1	MCM6	OAZ1	RNU5D-1	TPGS2
BLVRA	DEPDC1B	HIST1H1B	MCM7	ORC1	RNVU1-8	TPX2
BOLA3	DESI1	HIST1H1D	MCM8	ORC6	RPL22L1	TRAP1
BPNT1	DFFA	HIST1H1E	MDH1	OTUD6B	RPL29	TRIP13
BRCA1	DHCR24	HIST1H2AB	MELK	OXCT1	RPL29P11	TRMT112
BRI3BP	DHCR7	HIST1H2AC	METTL1	PA2G4	RPL35P5	TSFM
BRIP1	DHFR	HIST1H2AE	MGST1	PAICS	RPL8	TSTA3
BSG	DHX33	HIST1H2AI	MGST2	PAK1IP1	RPRD1A	TTF2
BUB1	DIAPH3	HIST1H2AJ	MID1IP1	PAM16	RPS9	TTK
BUB1B	DKC1	HIST1H2AM	MIF	PARP1	RPSAP56	TUBB
BZW2	DLAT	HIST1H2BF	MINA	PARPBP	RQCD1	TUBB4B
C14orf1	DLGAP5	HIST1H2BH	MIS18A	PBK	RRM1	TUBBP1
C1QBP	DNAJC11	HIST1H2BI	MKI67	PCCB	RRM2	TUBG1
C1QTNF9B-AS1	DNAJC8	HIST1H2BL	MLF1IP	PCNA	RRP15	TUFM
C3orf37	DNMT1	HIST1H2BM	MLF2	PDCD11	RRP36	TXN2
C4orf27	DNPH1	HIST1H2BO	MMAB	PDCD2L	RRS1	TYMS
C6orf211	DPAGT1	HIST1H3A	MMS22L	PDCL3	RSU1	UBE2S
C9orf123	DPCD	HIST1H3B	MND1	PDE3B	RUVBL1	UBE2T
C9orf64	DPH5	HIST1H3F	MNF1	PDHB	RUVBL2	UBR7
CA8	DPH6	HIST1H3I	MPDU1	PDSS1	RWDD2B	UCHL3
CAD	DPP3	HIST1H3J	MPHOSPH6	PFDN6	SAE1	UHRF1
CASC5	DPY30	HIST1H4A	MRFAP1	PGD	SAMHD1	UMPS
CASP6	DRG2	HIST1H4B	MRPL13	PHB	SAMM50	UNG
CBX5	DSCC1	HIST1H4C	MRPL15	PHB2	SASS6	UQCR11
CCDC167	DTL	HIST1H4D	MRPL16	PHF5A	SCARNA12	UQCRC1
CCNA2	DUT	HIST1H4E	MRPL19	PHF6	SCARNA17	UQCRFS1P1
CCNB1	DYNLT1	HIST1H4J	MRPL20	PHGDH	SELRC1	URM1
CCNB2	E2F7	HIST1H4K	MRPL22	PIGK	SEPHS1	UROS
CCNE2	E2F8	HIST1H4L	MRPL24	PIGU	SET	USP13
CCT2	EBNA1BP2	HIST2H2AA3	MRPL28	PINX1	SF3A3	USP39
CCT3	EBP	HIST2H2AB	MRPL35	PKP2	SF3B3	UTP14A
CCT5	EBPL	HIST2H2AC	MRPL37	PLA2G2A	SGOL1	UTP15
CCT6A	ECH1	HIST2H3D	MRPL42	PLA2G4A	SHCBP1	UTP20
CCT7	ECT2	HIST2H3PS2	MRPL45	PLK1	SHMT2	UXS1
CD151	EEF1E1	HJURP	MRPL46	PLK4	SKA1	VBP1
CD81	EEF2	HMGB1	MRPL47	PM20D2	SKA2	VKORC1L1
CDC123	EFTUD2	HMGB2	MRPL50	PMPCA	SKA3	VPS25
CDC20	EI24	HMGB3	MRPL54	PNPT1	SKP2	VPS4A
CDC25A	EIF3C	HMGN1P30	MRPS11	POLA2	SLC18B1	VRK1
CDC37	EIF3G	HMGN1P38	MRPS15	POLD2	SLC1A5	WBSCR22
CDC45	EIF3I	HMGN2P3	MRPS16	POLD3	SLC25A11	WDHD1
CDC6	EIF3K	HMGN3	MRPS18A	POLE	SLC25A15	WDR12
CDC7	EIF3L	HMMR	MRPS18B	POLR1E	SLC25A39	WDR3
CDCA2	EIF4A1	HNRNPA0	MRPS2	POLR2G	SLC27A2	WDR5
CDCA3	EIF4B	HNRNPL	MRPS23	POLR3K	SLC35A4	WDR76
CDCA5	EIF5AL1	HNRNPM	MRPS27	POP1	SLC35F2	WDR77
CDCA7	EIF5AP4	HOXB9	MRPS28	POP5	SLC39A8	WIBG
CDCA7L	ELAC2	HPRT1	MRPS33	PPAT	SLC5A6	XPO5
CDCA8	ELOVL6	HRSP12	MRPS35	PPIB	SLIRP	XRCC2
CDK1	EMD	HSD17B10	MRPS7	PPIH	SLITRK6	XRCC5

CDK2	EMG1	HSD17B4	MRTO4	PPIL1	SMARCB1	XRCC6
CDK4	ENO1	HSPA1B	MSH2	PPM1G	SMC1A	YARS
CDK5	ENOPH1	IDH1	MST4	PPP1CA	SMC2	YEATS4
CDKN3	ERCC6L	IDH2	MTA2	PPP1R14B	SNAPIN	ZNF724P
CENPE	ERH	IFITM3	MTFMT	PPP2R4	SNHG1	ZWILCH
CENPF	ESCO2	IGFBP2	MTHFD1	PPP4C	SNORD74	ZWINT
CENPH	ESD	ILF2	MYBL2	PPP5C	SNRNP25	

Table A2. Lists of genes (gene symbol) used in PAN-CANCER SCCC UPREGULATED (first list) and PAN-CANCER SCCC DOWNREGULATED (second list) gene sets.

BIBLIOGRAPHY

1. Humphries, A. & Wright, N. A. Colonic crypt organization and tumorigenesis. *Nat. Rev. Cancer* **8**, 415–424 (2008).

2. Chang, W. W. L. & Leblond, C. P. Renewal of the epithelium in the descending colon of the mouse. I. Presence of three cell populations: Vacuolated-columnar, mucous and argentaffin. *Am. J. Anat.* **131**, 73–99 (1971).

3. Tsubouchi, S. & Leblond, C. P. Migration and turnover of entero-endocrine and caveolated cells in the epithelium of the descending colon, as shown by radioautography after continuous infusion of 3H-thymidine into mice. *Am. J. Anat.* **156**, 431–51 (1979).

4. Velcich, A. *et al.* Patterns of expression of lineage-specific markers during the in vitro-induced differentiation of HT29 colon carcinoma cells. *Cell Growth Differ.* **6**, 749–57 (1995).

5. Yeh, J. C., Ong, E. & Fukuda, M. Molecular cloning and expression of a novel β-1,6-N-acetylglucosaminyltransferase that forms core 2, core 4, and I branches. *J. Biol. Chem.* **274**, 3215–21 (1999).

6. Frazer, D. M. *et al.* Cloning and gastrointestinal expression of rat hephaestin: relationship to other iron transport proteins. *Am. J. Physiol. Liver Physiol.* **281**, G931-9 (2001).

7. Petrak, J. & Vyoral, D. Hephaestin - A ferroxidase of cellular iron export. *Int. J. Biochem. Cell Biol.* **37**, 1173–8 (2005).

8. Zeuner, A., Todaro, M., Stassi, G. & De Maria, R. Colorectal cancer stem cells: From the crypt to the clinic. *Cell Stem Cell* **15**, 692–705 (2014).

9. Barker, N., Bartfeld, S. & Clevers, H. Tissue-Resident Adult Stem Cell Populations of Rapidly Self-Renewing Organs. *Cell Stem Cell* **7**, 656–670 (2010).

10. Cairnie, A. B., Lamerton, L. F. & Steel, G. G. Cell proliferation studies in the intestinal epithelium of the rat. I. Determination of the kinetic parameters. *Exp. Cell Res.* **39**, 528–38 (1965).

11. Potten, C. S. Extreme sensitivity of some intestinal crypt cells to X and γ irradiation. *Nature* **269**, 518–21 (1977).

12. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003–7 (2007).

13. Cheng, H. & Leblond, C. P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine V. Unitarian theory of the origin of the four epithelial cell types. *Am. J. Anat.* **141**, 537–61 (1974).

14. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262–265 (2009).

15. Van der Flier, L. G., Haegebarth, A., Stange, D. E., van de Wetering, M. & Clevers, H. OLFM4 Is a Robust Marker for Stem Cells in Human Intestine and Marks a Subset of Colorectal Cancer

Cells. *Gastroenterology* **137**, 15–17 (2009).

16. Zhu, L. *et al.* Prominin 1 marks intestinal stem cells that are susceptible to neoplastic transformation. *Nature* **457**, 603–607 (2009).

17. Jung, P. *et al.* Isolation and in vitro expansion of human colonic stem cells. *Nat. Med.* **17**, 1225–7 (2011).

18. Hirata, A. *et al.* Dose-dependent roles for canonical Wnt signalling in de novo crypt formation and cell cycle properties of the colonic epithelium. *J. Cell Sci.* **140**, 66–75 (2013).

19. Fevr, T., Robine, S., Louvard, D. & Huelsken, J. Wnt/beta-catenin is essential for intestinal homeostasis and maintenance of intestinal stem cells. *Mol. Cell. Biol.* **27**, 7551–9 (2007).

20. Crosnier, C., Stamataki, D. & Lewis, J. Organizing cell renewal in the intestine: Stem cells, signals and combinatorial control. *Nat. Rev. Genet.* **7**, 349–59 (2006).

21. Van Es, J. H. *et al.* Notch/gamma-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature* **435**, 959–63 (2005).

22. Sailaja, B. S., He, X. C. & Li, L. The regulatory niche of intestinal stem cells. *J. Physiol.* **594**, 4827–36 (2016).

23. Clevers, H. & Nusse, R. Wnt/β-catenin signaling and disease. *Cell* **149**, 1192–205 (2012).

24. Van Es, J. H., Barker, N. & Clevers, H. You Wnt some, you lose some: Oncogenes in the Wnt signaling pathway. *Curr. Opin. Genet. Dev.* **13**, 28–33 (2003).

25. Behrens, J. *et al.* Functional interaction of β-catenin with the transcription factor LEF- 1. *Nature* **15**, 638–42 (1996).

26. Huber, O. *et al.* Nuclear localization of beta-catenin by interaction with transcription factor LEF-1. *Mech. Dev.* **59**, 3–10 (1996).

27. Molenaar, M. *et al.* XTcf-3 transcription factor mediates β-catenin-induced axis formation in xenopus embryos. *Cell* **86**, 391–9 (1996).

28. Van de Wetering, M. *et al.* The β-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241–50 (2002).

29. Artavanis-Tsakonas, S. Notch Signaling: Cell Fate Control and Signal Integration in Development. *Science.* **284**, 770–776 (1999).

30. Miyamoto, S. & Rosenberg, D. W. Role of Notch signaling in colon homeostasis and carcinogenesis. *Cancer Sci.* **102**, 1938–42 (2011).

31. Sikandar, S. S. *et al.* NOTCH signaling is required for formation and self-renewal of tumor-initiating cells and for repression of secretory cell differentiation in colon cancer. *Cancer Res.* **70**, 1469–1478 (2010).

32. Katoh, M. & Katoh, M. Notch signaling in gastrointestinal tract (review). *Int. J. Oncol.* **30**, 247–51 (2007).

33. Jensen, J. *et al.* Control of endodermal endocrine development by Hes-1. *Nat. Genet.* **24**, 36–44 (2000).

34. Ghaleb, A. M., Aggarwal, G., Bialkowska, A. B., Nandan, M. O. & Yang, V. W. Notch Inhibits Expression of the Kruppel-Like Factor 4 Tumor Suppressor in the Intestinal Epithelium. *Mol. Cancer Res.* **6**, 1920–7 (2008).

35. Dang, D. T., Pevsner, J. & Yang, V. W. The biology of the mammalian Krüppel-like family of transcription factors. *Int. J. Biochem. Cell Biol.* **32**, 1103–21 (2000).

36. Sander, G. R. & Powell, B. C. Expression of Notch Receptors and Ligands in the Adult Gut. *J. Histochem. Cytochem.* **52**, 509–16 (2004).

37. Vogelstein, B. *et al.* Genetic Alterations during Colorectal-Tumor Development. *N. Engl. J. Med.* **319**, 525–532 (1988).

38. Fearon, E. R. Molecular genetics of colorectal cancer. *Annu. Rev. Pathol.* **6**, 479–507 (2011).

39. Mei, Z., Duan, C. Y., Li, C. B., Cui, L. & Ogino, S. Prognostic role of tumor PIK3CA mutation in colorectal cancer: A systematic review and meta-analysis. *Ann. Oncol.* **27**, 1836–48 (2016).

40. Barras, D. BRAF Mutation in Colorectal Cancer: An Update. *Biomark. Cancer* **7**, 9–12 (2015).

41. Shih, I.-M. *et al.* Top-down morphogenesis of colorectal tumors. *Proc. Natl. Acad. Sci.* **98**, 2640–5 (2001).

42. Preston, S. L. *et al.* Bottom-up histogenesis of colorectal adenomas: Origin in the monocryptal adenoma and initial expansion by crypt fission. *Cancer Res.* **63**, 3819–25 (2003).

43. Merlos-Suárez, A. *et al.* The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* **8**, 511–524 (2011).

44. Burrell, R. A., McGranahan, N., Bartek, J. & Swanton, C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**, 338–45 (2013).

45. Jin, K. *et al.* Mechanisms regulating colorectal cancer cell metastasis into liver (review). *Oncol. Lett.* **3**, 11–15 (2012).

46. Riihimaki, M., Hemminki, A., Sundquist, J. & Hemminki, K. Patterns of metastasis in colon and rectal cancer. *Sci. Rep.* **6**, (2016).

47. Aguirre-Ghiso, J. A. Models, mechanisms and clinical evidence for cancer dormancy. *Nat. Rev. Cancer* **7**, 834–846 (2007).

48. Sertil, A. R. *Tumor Dormancy, Quiescence, and Senescence, Vol. 3. Tumor Dormancy, Quiescence, and Senescence: Aging, Cancer, and Noncancer Pathologies* **3**, (Springer

Netherlands, 2014).

49. Yeh, A. C. & Ramaswamy, S. Mechanisms of Cancer Cell Dormancy--Another Hallmark of Cancer? *Cancer Res.* **75**, 5014–22 (2015).

50. N, H. *et al.* SEER Cancer Statistics Review, 1975-2016. (2019). Available at: https://seer.cancer.gov/csr/1975_2014/.

51. Nguyen, D. X., Bos, P. D. & Massagué, J. Metastasis: From dissemination to organ-specific colonization. *Nat. Rev. Cancer* **9**, 274–84 (2009).

52. Gomis, R. R. & Gawrzak, S. Tumor cell dormancy. *Mol. Oncol.* **11**, 62–78 (2017).

53. Talukdar, S. *et al.* Dormancy and cancer stem cells: An enigma for cancer therapeutic targeting. in *Advances in Cancer Research* **141**, 43–84 (Elsevier Inc., 2019).

54. Gao, X., Zhang, M., Tang, Y. & Liang, X. Cancer cell dormancy: mechanisms and implications of cancer recurrence and metastasis. *Onco. Targets. Ther.* **10**, 5219–5228 (2017).

55. Ghajar, C. M. *et al.* The perivascular niche regulates breast tumour dormancy. *Nat. Cell Biol.* **15**, 807–817 (2013).

56. Linde, N., Fluegen, G. & Aguirre-Ghiso, J. A. The Relationship Between Dormant Cancer Cells and Their Microenvironment. *Adv. Cancer Res.* **132**, 45–71 (2016).

57. Patel, P. & Chen, E. I. Cancer stem cells, tumor dormancy, and metastasis. *Front. Endocrinol. (Lausanne).* **3**, 125 (2012).

58. Shaked, Y., McAllister, S., Fainaru, O. & Almog, N. Tumor dormancy and the angiogenic switch: possible implications of bone marrow- derived cells. *Curr. Pharm. Des.* **20**, 4920–33 (2014).

59. Adam, A. P. *et al.* Computational identification of a p38 SAPK-regulated transcription factor network required for tumor cell quiescence. *Cancer Res.* **69**, 5664–5672 (2009).

60. Thornton, T. M. & Rincon, M. Non-classical p38 map kinase functions: Cell cycle checkpoints and survival. *Int. J. Biol. Sci.* **5**, 44–52 (2009).

61. Aguirre-Ghiso, J. A., Estrada, Y., Liu, D. & Ossowski, L. ERK(MAPK) activity as a determinant of tumor growth and dormancy; regulation by p38(SAPK). *Cancer Res.* **63**, 1684–95 (2003).

62. Sosa, M. S., Avivar-Valderas, A., Bragado, P., Wen, H. C. & Aguirre-Ghiso, J. A. ERK1/2 and p38α/β signaling in tumor cell quiescence: opportunities to control dormant residual disease. *Clin. Cancer Res.* **17**, 5850–7 (2011).

63. Tesio, M. *et al.* Hematopoietic stem cell quiescence and function are controlled by the CYLD–TRAF2–p38MAPK pathway. *J. Exp. Med.* **212**, 525–538 (2015).

64. Aguirre-Ghiso, J. A., Ossowski, L. & Rosenbaum, S. K. Green fluorescent protein tagging of extracellular signal-regulated kinase and p38 pathways reveals novel dynamics of pathway

activation during primary and metastatic growth. *Cancer Res.* **64**, 7336–45 (2004).

65. Aguirre-Ghiso, J. A., Liu, D., Mignatti, A., Kovalski, K. & Ossowski, L. Urokinase Receptor and Fibronectin Regulate the ERKMAPK to p38MAPK Activity Ratios That Determine Carcinoma Cell Proliferation or Dormancy In Vivo. *Mol. Biol. Cell* **12**, 863–79 (2001).

66. Sosa, M. S., Bernstein, E. & Aguirre-Ghiso, J. A. Epigenetic Regulation of Cancer Dormancy as a Plasticity Mechanism for Metastasis Initiation. in *Tumor Dormancy and Recurrence* (eds. Wang, Y. & Crea, F.) 1–16 (Springer International Publishing, 2017).

67. Tsai, H. C. *et al.* Transient Low Doses of DNA-Demethylating Agents Exert Durable Antitumor Effects on Hematological and Epithelial Tumor Cells. *Cancer Cell* **21**, 430–446 (2012).

68. Wilkinson, D. S. *et al.* A Direct Intersection between p53 and Transforming Growth Factor Pathways Targets Chromatin Modification and Transcription Repression of the α-Fetoprotein Gene. *Mol. Cell. Biol.* **25**, 1200–1212 (2005).

69. Glenisson, W., Castronovo, V. & Waltregny, D. Histone deacetylase 4 is required for TGFβ1-induced myofibroblastic differentiation. *Biochim. Biophys. Acta* **1773**, 1572–82 (2007).

70. Jepsen, K. *et al.* SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron. *Nature* **450**, 415–9 (2007).

71. Cras, A. *et al.* Epigenetic patterns of the retinoic acid receptor beta2 promoter in retinoic acid-resistant thyroid cancer cells. *Oncogene* **26**, 4018–24 (2007).

72. Sosa, M. S., Bragado, P. & Aguirre-Ghiso, J. A. Mechanisms of disseminated cancer cell dormancy: An awakening field. *Nat. Rev. Cancer* **14**, 611–622 (2014).

73. Gatenby, R. A. *et al.* Oxygen tension in human tumors: in vivo mapping using CT-guided probes. *Radiology* **156**, 211–4 (2014).

74. Carmeliet, P. *et al.* Role of HIF-1α in hypoxia-mediated apoptosis, cell proliferation and tumour angiogenesis. *Nature* **394**, 485–490 (1998).

75. Krtolica, A., Krucher, N. A. & Ludlow, J. W. Hypoxia-induced pRB hypophosphorylation results from downregulation of CDK and upregulation of PP1 activities. *Oncogene* **17**, 2295–2304 (1998).

76. Vietri, M., Bianchi, M., Ludlow, J. W., Mittnacht, S. & Villa-Moruzzi, E. Direct interaction between the catalytic subunit of Protein Phosphatase 1 and pRb. *Cancer Cell Int.* **6**, 3 (2006).

77. Hackenbeck, T. *et al.* HIF-1 or HIF-2 induction is sufficient to achieve cell cycle arrest in NIH3T3 mouse fibroblasts independent from hypoxia. *Cell Cycle* **8**, 1386–95 (2009).

78. Hubbi, M. E. *et al.* A nontranscriptional role for HIF-1α as a direct inhibitor of DNA replication. *Sci. Signal.* **6**, ra10 (2013).

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

80. Corn, P. G. *et al.* Mxi1 is induced by hypoxia in a HIF-1-dependent manner and protects cells from c-Myc-induced apoptosis. *Cancer Biol. Ther.* **4**, 1285–94 (2005).

81. Antero, S., Kai, K. & Anu, K. Hypoxia-Inducible Histone Lysine Demethylases: Impact on the Aging Process and Age-Related Diseases. *Aging Dis.* **7**, 180 (2016).

82. Wouters, B. G. & Koritzinsky, M. Hypoxia signalling through mTOR and the unfolded protein response in cancer. *Nat. Rev. Cancer* **8**, 851–864 (2008).

83. Walter, P. & Ron, D. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. *Science.* **334**, 1081–1086 (2011).

84. Schönenberger, M. J. Hypoxia signaling pathways: modulators of oxygen-related organelles. *Front. Cell Dev. Biol.* **3**, (2015).

85. Koritzinsky, M. *et al.* Gene expression during acute and prolonged hypoxia is regulated by distinct mechanisms of translational control. *EMBO J.* **25**, 1114–25 (2006).

86. Koumenis, C. *et al.* Regulation of Protein Synthesis by Hypoxia via Activation of the Endoplasmic Reticulum Kinase PERK and Phosphorylation of the Translation Initiation Factor eIF2. *Mol. Cell. Biol.* **22**, 7405–16 (2002).

87. Ivanova, I. G., Park, C. V., Yemm, A. I. & Kenneth, N. S. PERK/eIF2α signaling inhibits HIF-induced gene expression during the unfolded protein response via YB1-dependent regulation of HIF1α translation. *Nucleic Acids Res.* **46**, 3878–3890 (2018).

88. Rouschop, K. M. A. *et al.* The unfolded protein response protects human tumor cells during hypoxia through regulation of the autophagy genes MAP1LC3B and ATG5. *J. Clin. Invest.* **120**, 127–41 (2010).

89. Lu, Z. *et al.* The tumor suppressor gene ARHI regulates autophagy and tumor dormancy in human ovarian cancer cells. *J. Clin. Invest.* **118**, 3917–29 (2008).

90. Rouschop, K. M. *et al.* PERK/eIF2 signaling protects therapy resistant hypoxic cells through induction of glutathione synthesis and protection against ROS. *Proc. Natl. Acad. Sci.* **110**, 4622–4627 (2013).

91. Ranganathan, A. C., Ojha, S., Kourtidis, A., Conklin, D. S. & Aguirre-Ghiso, J. A. Dual Function of Pancreatic Endoplasmic Reticulum Kinase in Tumor Cell Growth Arrest and Survival. *Cancer Res.* **68**, 3260–3268 (2008).

92. Gilany, K. & Mohtaram, V. Hypoxia: a Review. *J. Paramed. Sci.* **1**, 43–60 (2010).

93. Sluimer, J. C. *et al.* Hypoxia, Hypoxia-Inducible Transcription Factor, and Macrophages in Human Atherosclerotic Plaques Are Correlated With Intraplaque Angiogenesis. *J. Am. Coll.*

Cardiol. **51**, 1258–1265 (2008).

94. Li, J. *et al.* Altered metabolic responses to intermittent hypoxia in mice with partial deficiency of hypoxia-inducible factor-1α. *Physiol. Genomics* **25**, 450–457 (2006).

95. Carreau, A., Hafny-Rahbi, B. El, Matejuk, A., Grillon, C. & Kieda, C. Why is the partial oxygen pressure of human tissues a crucial parameter? Small molecules and hypoxia. *J. Cell. Mol. Med.* **15**, 1239–1253 (2011).

96. McKeown, S. R. Defining normoxia, physoxia and hypoxia in tumours-implications for treatment response. *Br. J. Radiol.* **87**, 20130676 (2014).

97. Bragado, P., Sosa, M. S., Keely, P., Condeelis, J. & Aguirre-Ghiso, J. A. Microenvironments Dictating Tumor Cell Dormancy. *Recent Results Cancer Res.* **195**, 25–39 (2012).

98. Williamson, J. R. *et al.* Hyperglycemic Pseudohypoxia and Diabetic Complications. *Diabetes* **42**, 801–813 (1993).

99. Kluckova, K. & Tennant, D. A. Metabolic implications of hypoxia and pseudohypoxia in pheochromocytoma and paraganglioma. *Cell Tissue Res.* **372**, 367–378 (2018).

100. Yang, X., Yang, S., Wang, C. & Kuang, S. The hypoxia-inducible factors HIF1α and HIF2α are dispensable for embryonic muscle development but essential for postnatal muscle regeneration. *J. Biol. Chem.* **292**, 5981–5991 (2017).

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

102. Tian, H., McKnight, S. L. & Russell, D. W. Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. *Genes Dev.* **11**, 72–82 (1997).

103. Ema, M. *et al.* A novel bHLH-PAS factor with close sequence similarity to hypoxia-inducible factor 1 regulates the VEGF expression and is potentially involved in lung and vascular development. *Proc. Natl. Acad. Sci.* **94**, 4273–4278 (1997).

104. Majmundar, A. J., Wong, W. J. & Simon, M. C. Hypoxia-Inducible Factors and the Response to Hypoxic Stress. *Mol. Cell* **40**, 294–309 (2010).

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

106. Dengler, V. L., Galbraith, M. D. & Espinosa, J. M. Transcriptional regulation by hypoxia inducible factors. *Crit. Rev. Biochem. Mol. Biol.* **49**, 1–15 (2014).

107. Platt, J. L. *et al.* Capture-C reveals preformed chromatin interactions between <scp>HIF</scp> -binding sites and distant promoters. *EMBO Rep.* **17**, 1410–1421 (2016).

108. Schodel, J. *et al.* High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. *Blood* **117**, e207–e217 (2011).

109. Choudhry, H. & Harris, A. L. Advances in Hypoxia-Inducible Factor Biology. *Cell Metab.* **27**, 281–298 (2018).

110. Schito, L. & Rey, S. Cell-Autonomous Metabolic Reprogramming in Hypoxia. *Trends Cell Biol.* **28**, 128–142 (2018).

111. Graham, A. M. & Presnell, J. S. Hypoxia Inducible Factor (HIF) transcription factor family expansion, diversification, divergence and selection in eukaryotes. *PLoS One* **12**, e0179545 (2017).

112. Maynard, M. A. *et al.* Multiple Splice Variants of the Human HIF-3α Locus Are Targets of the von Hippel-Lindau E3 Ubiquitin Ligase Complex. *J. Biol. Chem.* **278**, 11032–11040 (2003).

113. Pasanen, A. *et al.* Hypoxia-inducible factor (HIF)-3α is subject to extensive alternative splicing in human tissues and cancer cells and is regulated by HIF-1 but not HIF-2. *Int. J. Biochem. Cell Biol.* **42**, 1189–1200 (2010).

114. Li, H., Ko, H. P. & Whitlock, J. P. Induction of Phosphoglycerate Kinase 1 Gene Expression by Hypoxia. *J. Biol. Chem.* **271**, 21262–21267 (2002).

115. Freedman, S. J. *et al.* Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1. *Proc. Natl. Acad. Sci.* **99**, 5367–5372 (2002).

116. Bruick, R. K. A Conserved Family of Prolyl-4-Hydroxylases That Modify HIF. *Science.* **294**, 1337–1340 (2001).

117. Serocki, M. *et al.* miRNAs regulate the HIF switch during hypoxia: a novel therapeutic target. *Angiogenesis* **21**, 183–202 (2018).

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

119. Kaidi, A., Williams, A. C. & Paraskeva, C. Interaction between β-catenin and HIF-1 promotes cellular adaptation to hypoxia. *Nat. Cell Biol.* **9**, 210–217 (2007).

120. Sánchez-Elsner, T. *et al.* Synergistic Cooperation between Hypoxia and Transforming Growth Factor-β Pathways on Human Vascular Endothelial Growth Factor Gene Expression. *J. Biol. Chem.* **276**, 38527–38535 (2001).

121. Koshiji, M. *et al.* HIF-1α Induces Genetic Instability by Transcriptionally Downregulating MutSα Expression. *Mol. Cell* **17**, 793–803 (2005).

122. Kaluz, S., Kaluzová, M., Liao, S.-Y., Lerman, M. & Stanbridge, E. J. Transcriptional control of the tumor- and hypoxia-marker carbonic anhydrase 9: A one transcription factor (HIF-1) show? *Biochim. Biophys. Acta - Rev. Cancer* **1795**, 162–172 (2009).

123. Liao, S.-Y., Lerman, M. I. & Stanbridge, E. J. Expression of transmembrane carbonic anhydrases, CAIX and CAXII, in human development. *BMC Dev. Biol.* **9**, 22 (2009).

124. Ivanov, S. *et al.* Expression of Hypoxia-Inducible Cell-Surface Transmembrane Carbonic Anhydrases in Human Cancer. *Am. J. Pathol.* **158**, 905–919 (2001).

125. Wykoff, C. C. *et al.* Hypoxia-inducible expression of tumor-associated carbonic anhydrases. *Cancer Res.* **60**, 7075–83 (2000).

126. Chia, S. K. *et al.* Prognostic Significance of a Novel Hypoxia-Regulated Marker, Carbonic Anhydrase IX, in Invasive Breast Carcinoma. *J. Clin. Oncol.* **19**, 3660–3668 (2001).

127. Giatromanolaki, A. *et al.* Expression of hypoxia-inducible carbonic anhydrase-9 relates to angiogenic pathways and independently to poor outcome in non-small cell lung cancer. *Cancer Res.* **61**, 7992–8 (2001).

128. Haapasalo, J. A. Expression of Carbonic Anhydrase IX in Astrocytic Tumors Predicts Poor Prognosis. *Clin. Cancer Res.* **12**, 473–477 (2006).

129. Måseide, K. *et al.* Carbonic Anhydrase IX as a Marker for Poor Prognosis in Soft Tissue Sarcoma. *Clin. Cancer Res.* **10**, 4464–4471 (2004).

130. Loncaster, J. A. *et al.* Carbonic anhydrase (CA IX) expression, a potential new intrinsic marker of hypoxia: Correlations with tumor oxygen measurements and prognosis in locally advanced carcinoma of the cervix. *Cancer Res.* **61**, 6394–9 (2001).

131. Jonathan, R. A. *et al.* The prognostic value of endogenous hypoxia-related markers for head and neck squamous cell carcinomas treated with ARCON. *Radiother. Oncol.* **79**, 288–297 (2006).

132. Olive, P. L. *et al.* Carbonic anhydrase 9 as an endogenous marker for hypoxic cells in cervical cancer. *Cancer Res.* **61**, 8924–9 (2001).

133. Mayer, A. Carbonic Anhydrase IX Expression and Tumor Oxygenation Status Do Not Correlate at the Microregional Level in Locally Advanced Cancers of the Uterine Cervix. *Clin. Cancer Res.* **11**, 7220–7225 (2005).

134. Cho, M. *et al.* Hypomethylation of the MN/CA9 promoter and upregulated MN/CA9 expression in human renal cell carcinoma. *Br. J. Cancer* **85**, 563–567 (2001).

135. Ashida, S., Nishimori, I., Tanimura, M., Onishi, S. & Shuin, T. Effects of von Hippel-Lindau gene mutation and methylation status on expression of transmembrane carbonic anhydrases in renal cell carcinoma. *J. Cancer Res. Clin. Oncol.* **128**, 561–568 (2002).

136. Grabmaier, K., de Weijert, M., Uemura, H., Schalken, J. & Oosterwijk, E. Renal cell carcinoma-associated G250 methylation and expression: in vivo and in vitro studies. *Urology* **60**, 357–362 (2002).

137. Kaluzova, M., Kaluz, S., Lerman, M. I. & Stanbridge, E. J. DNA Damage Is a Prerequisite for p53-Mediated Proteasomal Degradation of HIF-1 in Hypoxic Cells and Downregulation of the Hypoxia Marker Carbonic Anhydrase IX. *Mol. Cell. Biol.* **24**, 5757–5766 (2004).

138. Salceda, S. & Caro, J. Hypoxia-inducible Factor 1α (HIF-1α) Protein Is Rapidly Degraded by the Ubiquitin-Proteasome System under Normoxic Conditions. *J. Biol. Chem.* **272**, 22642–22647 (1997).

139. Metzen, E. *et al.* Regulation of the prolyl hydroxylase domain protein 2 (phd2/egln-1) gene: identification of a functional hypoxia-responsive element. *Biochem. J.* **387**, 711–717 (2005).

140. Berra, E. HIF prolyl-hydroxylase 2 is the key oxygen sensor setting low steady-state levels of HIF-1 in normoxia. *EMBO J.* **22**, 4082–4090 (2003).

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

142. Lisztwan, J., Imbert, G., Wirbelauer, C., Gstaiger, M. & Krek, W. The von Hippel-Lindau tumor suppressor protein is a component of an E3 ubiquitin-protein ligase activity. *Genes Dev.* **13**, 1822–1833 (1999).

143. Stebbins, C. E. Structure of the VHL-ElonginC-ElonginB Complex: Implications for VHL Tumor Suppressor Function. *Science.* **284**, 455–461 (1999).

144. Gehrke, S. G. *et al.* UbcH5A, a member of human E2 ubiquitin-conjugating enzymes, is closely related to SFT, a stimulator of iron transport, and is up-regulated in hereditary hemochromatosis. *Blood* **101**, 3288–3293 (2003).

145. Cardote, T. A. F., Gadd, M. S. & Ciulli, A. Crystal Structure of the Cul2-Rbx1-EloBC-VHL Ubiquitin Ligase Complex. *Structure* **25**, 901-911.e3 (2017).

146. Krek, W. VHL takes HIF's breath away. *Nat. Cell Biol.* **2**, E121–E123 (2000).

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

148. Webb, J. D., Coleman, M. L. & Pugh, C. W. Hypoxia, hypoxia-inducible factors (HIF), HIF hydroxylases and oxygen sensing. *Cell. Mol. Life Sci.* **66**, 3539–3554 (2009).

149. Hagen, T., Taylor, C. T., Lam, F. & Moncada, S. Redistribution of intracellular oxygen in hypoxia by nitric oxide: effect on HIF1alpha. *Science.* **302**, 1975–1978 (2003).

150. Klimova, T. & Chandel, N. S. Mitochondrial complex III regulates hypoxic activation of HIF. *Cell Death Differ.* **15**, 660–666 (2008).

151. Chandel, N. S. Mitochondrial Regulation of Oxygen Sensing. in *Advances in Experimental Medicine and Biology* 339–354 (Springer, 2010).

152. Selak, M. A. *et al.* Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-α prolyl hydroxylase. *Cancer Cell* **7**, 77–85 (2005).

153. Sonveaux, P. *et al.* Targeting the Lactate Transporter MCT1 in Endothelial Cells Inhibits Lactate-Induced HIF-1 Activation and Tumor Angiogenesis. *PLoS One* **7**, e33418 (2012).

154. Lu, H., Forbes, R. A. & Verma, A. Hypoxia-inducible Factor 1 Activation by Aerobic Glycolysis Implicates the Warburg Effect in Carcinogenesis. *J. Biol. Chem.* **277**, 23111–23115 (2002).

155. Kim. Pyruvate promotes tumor angiogenesis through HIF-1-dependent PAI-1 expression. *Int. J. Oncol.* **38**, 571–576 (2011).

156. MacKenzie, E. D. *et al.* Cell-Permeating -Ketoglutarate Derivatives Alleviate Pseudohypoxia in Succinate Dehydrogenase-Deficient Cells. *Mol. Cell. Biol.* **27**, 3282–3289 (2007).

157. Galban, S. & Gorospe, M. Factors interacting with HIF-1alpha mRNA: novel therapeutic targets. *Curr. Pharm. Des.* **15**, 3853–3860 (2009).

158. Sampaio, C. *et al.* MAPK Signaling Up-regulates the Activity of Hypoxia-inducible Factors by Its Effects on p300. **155**, 1683–1695 (2015).

159. Minet, E. *et al.* Hypoxia-induced activation of HIF-1: role of HIF-1α-Hsp90 interaction. *FEBS Lett.* **460**, 251–256 (1999).

160. Gradin, K. *et al.* Functional interference between hypoxia and dioxin signal transduction pathways: competition for recruitment of the Arnt transcription factor. *Mol. Cell. Biol.* **16**, 5221–31 (1996).

161. Minet, E. *et al.* HIF1A Gene Transcription Is Dependent on a Core Promoter Sequence Encompassing Activating and Inhibiting Sequences Located Upstream from the Transcription Initiation Site and cis Elements Located within the 5′UTR. *Biochem. Biophys. Res. Commun.* **261**, 534–540 (1999).

162. BelAiba, R. S. *et al.* Hypoxia Up-Regulates Hypoxia-Inducible Factor-1α Transcription by Involving Phosphatidylinositol 3-Kinase and Nuclear Factor κB in Pulmonary Artery Smooth Muscle Cells. *Mol. Biol. Cell* **18**, 4691–4697 (2007).

163. Sperandio, S. *et al.* The transcription factor Egr1 regulates the HIF-1α gene during hypoxia. *Mol. Carcinog.* **48**, 38–44 (2009).

164. Yu, L., Chen, X., Sun, X., Wang, L. & Chen, S. The Glycolytic Switch in Tumors: How Many Players Are Involved? *J. Cancer* **8**, 3430–3440 (2017).

165. Semenza, G. L. *et al.* Hypoxia Response Elements in the Aldolase A, Enolase 1, and Lactate Dehydrogenase A Gene Promoters Contain Essential Binding Sites for Hypoxia-inducible Factor 1. *J. Biol. Chem.* **271**, 32529–32537 (1996).

166. Yang, C. *et al.* Analysis of Hypoxia-Induced Metabolic Reprogramming. in *Methods in*

Enzymology 425–455 (Elsevier Inc., 2014).

167. Valcourt, J. R. *et al.* Staying alive. *Cell Cycle* **11**, 1680–1696 (2012).

168. Cam, H. & Houghton, P. J. Regulation of mammalian target of rapamycin complex 1 (mTORC1) by hypoxia: causes and consequences. *Target. Oncol.* **6**, 95–102 (2011).

169. Laplante, M. & Sabatini, D. M. mTOR signaling at a glance. *J. Cell Sci.* **122**, 3589–3594 (2009).

170. Lipina, C. & Hundal, H. S. Is REDD1 a Metabolic Éminence Grise ? *Trends Endocrinol. Metab.* **27**, 868–880 (2016).

171. Burton, T. R. & Gibson, S. B. The role of Bcl-2 family member BNIP3 in cell death and disease: NIPping at the heels of cell death. *Cell Death Differ.* **16**, 515–523 (2009).

172. Ponente, M. *et al.* PML promotes metastasis of triple-negative breast cancer through transcriptional regulation of HIF1A target genes. *JCI Insight* **2**, (2017).

173. Li, Y. *et al.* Bnip3 Mediates the Hypoxia-induced Inhibition on Mammalian Target of Rapamycin by Interacting with Rheb. *J. Biol. Chem.* **282**, 35803–35813 (2007).

174. Bernardi, R. *et al.* PML inhibits HIF-1α translation and neoangiogenesis through repression of mTOR. *Nature* **442**, 779–785 (2006).

175. Hardie, D. G. Minireview: The AMP-Activated Protein Kinase Cascade: The Key Sensor of Cellular Energy Status. *Endocrinology* **144**, 5179–5183 (2003).

176. Brugarolas, J. Regulation of mTOR function in response to hypoxia by REDD1 and the TSC1/TSC2 tumor suppressor complex. *Genes Dev.* **18**, 2893–2904 (2004).

177. Liu, L. *et al.* Hypoxia-Induced Energy Stress Regulates mRNA Translation and Cell Growth. *Mol. Cell* **21**, 521–531 (2006).

178. Gallie, D. R. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev.* **5**, 2108–2116 (1991).

179. Müller, D. *et al.* 4E-BP restrains eIF4E phosphorylation. *Translation* **1**, e25819 (2013).

180. Rong, L. *et al.* Control of eIF4E cellular localization by eIF4E-binding proteins, 4E-BPs. *RNA* **14**, 1318–1327 (2008).

181. Gingras, A.-C., Raught, B. & Sonenberg, N. eIF4 Initiation Factors: Effectors of mRNA Recruitment to Ribosomes and Regulators of Translation. *Annu. Rev. Biochem.* **68**, 913–963 (1999).

182. Buchkovich, N. J., Yu, Y., Zampieri, C. A. & Alwine, J. C. The TORrid affairs of viruses: effects of mammalian DNA viruses on the PI3K–Akt–mTOR signalling pathway. *Nat. Rev. Microbiol.* **6**, 266–275 (2008).

183. Svitkin, Y. V. *et al.* Eukaryotic Translation Initiation Factor 4E Availability Controls the Switch between Cap-Dependent and Internal Ribosomal Entry Site-Mediated Translation. *Mol. Cell. Biol.* **25**, 10556–10565 (2005).

184. Hellen, C. U. T. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.* **15**, 1593–1612 (2001).

185. Lang, K. J. D., Kappel, A. & Goodall, G. J. Hypoxia-inducible Factor-1α mRNA Contains an Internal Ribosome Entry Site That Allows Efficient Translation during Normoxia and Hypoxia. *Mol. Biol. Cell* **13**, 1792–1801 (2002).

186. Stein, I. *et al.* Translation of vascular endothelial growth factor mRNA by internal ribosome entry: implications for translation under hypoxia. *Mol. Cell. Biol.* **18**, 3112–9 (1998).

187. Braunstein, S. *et al.* A Hypoxia-Controlled Cap-Dependent to Cap-Independent Translation Switch in Breast Cancer. *Mol. Cell* **28**, 501–512 (2007).

188. Yoffe, Y. *et al.* Cap-independent translation by DAP5 controls cell fate decisions in human embryonic stem cells. *Genes Dev.* **30**, 1991–2004 (2016).

189. Bukhari, S. I. A. *et al.* A Specialized Mechanism of Translation Mediated by FXR1a-Associated MicroRNP in Cellular Quiescence. *Mol. Cell* **61**, 760–773 (2016).

190. Shmakova, A., Batie, M., Druker, J. & Rocha, S. Chromatin and oxygen sensing in the context of JmjC histone demethylases. *Biochem. J.* **462**, 385–395 (2014).

191. Pollard, P. J. *et al.* Regulation of Jumonji-domain-containing histone demethylases by hypoxia-inducible factor (HIF)-1α. *Biochem. J.* **416**, 387–394 (2008).

192. Beyer, S., Kristensen, M. M., Jensen, K. S., Johansen, J. V. & Staller, P. The Histone Demethylases JMJD1A and JMJD2B Are Transcriptional Targets of Hypoxia-inducible Factor HIF. *J. Biol. Chem.* **283**, 36542–36552 (2008).

193. Fu, L. *et al.* HIF-1α-induced histone demethylase JMJD2B contributes to the malignant phenotype of colorectal cancer cells via an epigenetic mechanism. *Carcinogenesis* **33**, 1664–1673 (2012).

194. Xia, X. *et al.* Integrative analysis of HIF binding and transactivation reveals its role in maintaining histone methylation homeostasis. *Proc. Natl. Acad. Sci.* **106**, 4260–4265 (2009).

195. Johnson, A. B., Denko, N. & Barton, M. C. Hypoxia induces a novel signature of chromatin modifications and global repression of transcription. *Mutat. Res.* **640**, 174–179 (2008).

196. Tausendschön, M., Dehne, N. & Brüne, B. Hypoxia causes epigenetic gene regulation in macrophages by attenuating Jumonji histone demethylase activity. *Cytokine* **53**, 256–262 (2011).

197. Chen, H., Yan, Y., Davidson, T. L., Shinkai, Y. & Costa, M. Hypoxic stress induces dimethylated

histone H3 lysine 9 through histone methyltransferase G9a in mammalian cells. *Cancer Res.* **66**, 9009–9016 (2006).

198. Casciello, F. *et al.* G9a drives hypoxia-mediated gene repression for breast cancer cell survival and tumorigenesis. *Proc. Natl. Acad. Sci.* **114**, 7077–7082 (2017).

199. Koslowski, M., Luxemburger, U., Türeci, Ö. & Sahin, U. Tumor-associated CpG demethylation augments hypoxia-induced effects by positive autoregulation of HIF-1α. *Oncogene* **30**, 876–882 (2011).

200. Nakamura, J. *et al.* Expression of Hypoxic Marker CA IX Is Regulated by Site-Specific DNA Methylation and Is Associated with the Histology of Gastric Cancer. *Am. J. Pathol.* **178**, 515–524 (2011).

201. Law, A. Y. S. *et al.* Epigenetic and HIF-1 regulation of stanniocalcin-2 expression in human cancer cells. *Exp. Cell Res.* **314**, 1823–1830 (2008).

202. Wenger, R. H., Kvietikova, I., Rolfs, A., Camenisch, G. & Gassmann, M. Oxygen-regulated erythropoietin gene expression is dependent on a CpG methylation-free hypoxia-inducible factor-1 DNA-binding site. *Eur. J. Biochem.* **253**, 771–777 (1998).

203. Murai, M. *et al.* Aberrant methylation and silencing of the BNIP3 gene in colorectal and gastric cancer. *Clin. Cancer Res.* **11**, 1021–7 (2005).

204. Shahrzad, S., Bertrand, K., Minhas, K. & Coomber, B. L. Induction of DNA hypomethylation by tumor hypoxia. *Epigenetics* **2**, 119–125 (2007).

205. Wu, M.-Z. *et al.* Hypoxia Drives Breast Tumor Malignancy through a TET–TNFα–p38–MAPK Signaling Axis. *Cancer Res.* **75**, 3912–3924 (2015).

206. Thienpont, B. *et al.* Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature* **537**, 63–68 (2016).

207. Watson, J. A. *et al.* Generation of an epigenetic signature by chronic hypoxia in prostate cells. *Hum. Mol. Genet.* **18**, 3594–3604 (2009).

208. Watson, C. J. *et al.* Hypoxia-induced epigenetic modifications are associated with cardiac tissue fibrosis and the development of a myofibroblast-like phenotype. *Hum. Mol. Genet.* **23**, 2176–2188 (2014).

209. Muz, B., de la Puente, P., Azab, F. & Azab, A. K. The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. *Hypoxia* **3**, 83–92 (2015).

210. Vadde, R. *et al.* Role of hypoxia-inducible factors (HIF) in the maintenance of stemness and malignancy of colorectal cancer. *Crit. Rev. Oncol. Hematol.* **113**, 22–27 (2017).

211. Lawler, P. R. & Lawler, J. Molecular Basis for the Regulation of Angiogenesis by Thrombospondin-1 and -2. *Cold Spring Harb. Perspect. Med.* **2**, a006627–a006627 (2012).

212. Folkman, J. Tumor Angiogenesis: Therapeutic Implications. *N. Engl. J. Med.* **285**, 1182–1186 (1971).

213. Hurwitz, H. *et al.* Bevacizumab plus Irinotecan, Fluorouracil, and Leucovorin for Metastatic Colorectal Cancer. *N. Engl. J. Med.* **350**, 2335–2342 (2004).

214. Mountzios, G., Pentheroudakis, G. & Carmeliet, P. Bevacizumab and micrometastases: Revisiting the preclinical and clinical rollercoaster. *Pharmacol. Ther.* **141**, 117–124 (2014).

215. Rohwer, N. *et al.* HIF-1α determines the metastatic potential of gastric cancer cells. *Br. J. Cancer* **100**, 772–781 (2009).

216. Lou, Y. *et al.* Targeting Tumor Hypoxia: Suppression of Breast Tumor Growth and Metastasis by Novel Carbonic Anhydrase IX Inhibitors. *Cancer Res.* **71**, 3364–3376 (2011).

217. Koukourakis, M. I. *et al.* Hypoxia-inducible factor (HIF1A and HIF2A), angiogenesis, and chemoradiotherapy outcome of squamous cell head-and-neck cancer. *Int. J. Radiat. Oncol.* **53**, 1192–1202 (2002).

218. Aebersold, D. M. *et al.* Expression of hypoxia-inducible factor-1α: a novel predictive and prognostic parameter in the radiotherapy of oropharyngeal cancer. *Cancer Res.* **61**, 2911–6 (2001).

219. Unruh, A. *et al.* The hypoxia-inducible factor-1α is a negative factor for tumor therapy. *Oncogene* **22**, 3213–3220 (2003).

220. Das, B. *et al.* Hypoxia Enhances Tumor Stemness by Increasing the Invasive and Tumorigenic Side Population Fraction. *Stem Cells* **26**, 1818–1830 (2008).

221. Pawlik, T. M. & Keyomarsi, K. Role of cell cycle in mediating sensitivity to radiotherapy. *Int. J. Radiat. Oncol.* **59**, 928–942 (2004).

222. Rohwer, N. & Cramer, T. Hypoxia-mediated drug resistance: Novel insights on the functional interaction of HIFs and cell death pathways. *Drug Resist. Updat.* **14**, 191–201 (2011).

223. Sendoel, A., Kohler, I., Fellmann, C., Lowe, S. W. & Hengartner, M. O. HIF-1 antagonizes p53-mediated apoptosis through a secreted neuronal tyrosinase. *Nature* **465**, 577–583 (2010).

224. Rohwer, N. *et al.* Hypoxia-Inducible Factor 1α Determines Gastric Cancer Chemosensitivity via Modulation of p53 and NF-κB. *PLoS One* **5**, e12038 (2010).

225. Fukuda, R. *et al.* HIF-1 Regulates Cytochrome Oxidase Subunits to Optimize Efficiency of Respiration in Hypoxic Cells. *Cell* **129**, 111–122 (2007).

226. Papandreou, I., Cairns, R. A., Fontana, L., Lim, A. L. & Denko, N. C. HIF-1 mediates adaptation to hypoxia by actively downregulating mitochondrial oxygen consumption. *Cell Metab.* **3**, 187–197 (2006).

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

228. Comerford, K. M. *et al.* Hypoxia-inducible factor-1-dependent regulation of the multidrug resistance (MDR1) gene. *Cancer Res.* **62**, 3387–94 (2002).

229. Abraham, J., Salama, N. N. & Azab, A. K. The role of P-glycoprotein in drug resistance in multiple myeloma. *Leuk. Lymphoma* **56**, 26–33 (2015).

230. Cosse, J.-P. & Michiels, C. Tumour hypoxia affects the responsiveness of cancer cells to chemotherapy and promotes cancer progression. *Anticancer. Agents Med. Chem.* **8**, 790–7 (2008).

231. Nakamura, T. *et al.* PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat. Cell Biol.* **9**, 64–71 (2007).

232. Iqbal, K., Jin, S.-G., Pfeifer, G. P. & Szabó, P. E. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc. Natl. Acad. Sci.* **108**, 3642–3647 (2011).

233. Saitou, M., Kagiwada, S. & Kurimoto, K. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development* **139**, 15–31 (2012).

234. Nakashima, H. *et al.* Effects of Dppa3 on DNA Methylation Dynamics During Primordial Germ Cell Development in Mice. *Biol. Reprod.* **88**, 1–9 (2013).

235. Dhe-Paganon, S., Syeda, F. & Park, L. DNA methyl transferase 1: regulatory mechanisms and implications in health and disease. *Int. J. Biochem. Mol. Biol.* **2**, 58–66 (2011).

236. Wossidlo, M. *et al.* 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat. Commun.* **2**, 241 (2011).

237. Koh, K. P. *et al.* Tet1 and Tet2 Regulate 5-Hydroxymethylcytosine Production and Cell Lineage Specification in Mouse Embryonic Stem Cells. *Cell Stem Cell* **8**, 200–213 (2011).

238. Hajkova, P. Epigenetic reprogramming — taking a lesson from the embryo. *Curr. Opin. Cell Biol.* **22**, 342–350 (2010).

239. Gu, T.-P. *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–610 (2011).

240. Melamed, P., Yosefzon, Y., David, C., Tsukerman, A. & Pnueli, L. Tet Enzymes, Variants, and Differential Effects on Function. *Front. Cell Dev. Biol.* **6**, (2018).

241. Branco, M. R., Oda, M. & Reik, W. Safeguarding parental identity: Dnmt1 maintains imprints during epigenetic reprogramming in early embryogenesis. *Genes Dev.* **22**, 1567–1571 (2008).

242. Hirasawa, R. *et al.* Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev.* **22**, 1607–1616 (2008).

243. Kang, J., Kalantry, S. & Rao, A. PGC7, H3K9me2 and Tet3: regulators of DNA methylation in zygotes. *Cell Res.* **23**, 6–9 (2013).

244. Payer, B. *et al.* stella Is a Maternal Effect Gene Required for Normal Early Development in Mice. *Curr. Biol.* **13**, 2110–2117 (2003).

245. Hill, P. W. S. *et al.* Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte. *Nature* **555**, 392–396 (2018).

246. Popp, C. *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101–1105 (2010).

247. Hajkova, P. *et al.* Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117**, 15–23 (2002).

248. Lee, J. *et al.* Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. *Development* **129**, 1807–17 (2002).

249. Kurimoto, K. *et al.* Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. *Genes Dev.* **22**, 1617–1635 (2008).

250. Tachibana, M. Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev.* **19**, 815–826 (2005).

251. Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M. & Saitou, M. Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J.* **32**, 340–353 (2012).

252. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

253. Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **8**, 272–285 (2007).

254. Clark, A. T. *et al.* Human STELLAR, NANOG, and GDF3 genes are expressed in pluripotent cells and map to chromosome 12p13, a hotspot for teratocarcinoma. *Stem Cells* **22**, 169–179 (2004).

255. Ensembl. Available at: https://www.ensembl.org.

256. Nakamura, T. *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415–9 (2012).

257. Liu, H. *et al.* Comprehensive Proteomic Analysis of PGC7-Interacting Proteins. *J. Proteome Res.* **16**, 3113–3123 (2017).

258. Dunker, A. K. *et al.* The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics* **9**, S1 (2008).

259. Uversky, V. N., Gillespie, J. R. & Fink, A. L. Why are 'natively unfolded' proteins unstructured under physiologic conditions? *Proteins Struct. Funct. Genet.* **41**, 415–427 (2000).

260. Li, T. *et al.* Structural and mechanistic insights into UHRF1-mediated DNMT1 activation in the maintenance DNA methylation. *Nucleic Acids Res.* **46**, 3218–3231 (2018).

261. Liu, X. *et al.* UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nat. Commun.* **4**, 1563 (2013).

262. Du, W. *et al.* Stella protein facilitates DNA demethylation by disrupting the chromatin association of the RING finger–type E3 ubiquitin ligase UHRF1. *J. Biol. Chem.* **294**, 8907–8917 (2019).

263. Alekseev, O. M., Richardson, R. T. & O'Rand, M. G. Linker Histones Stimulate HSPA2 ATPase Activity Through NASP Binding and Inhibit CDC2/Cyclin B1 Complex Formation During Meiosis in the Mouse1. *Biol. Reprod.* **81**, 739–748 (2009).

264. Li, Y. A novel histone deacetylase pathway regulates mitosis by modulating Aurora B kinase activity. *Genes Dev.* **20**, 2566–2579 (2006).

265. Bowles, J., Teasdale, R. P., James, K. & Koopman, P. Dppa3 is a marker of pluripotency and has a human homologue that is expressed in germ cell tumours. *Cytogenet. Genome Res.* **101**, 261–265 (2003).

266. Giuliano, C., Freemantle, S. & Spinella, M. Testicular Germ Cell Tumors: A Paradigm for the Successful Treatment of Solid Tumor Stem Cells. *Curr. Cancer Ther. Rev.* **2**, 255–270 (2006).

267. Ezeh, U. I., Turek, P. J., Reijo Pera, R. A. & Clark, A. T. Human embryonic stem cell genes OCT4, NANOG, STELLAR, and GDF3 are expressed in both seminoma and breast carcinoma. *Cancer* **104**, 2255–2265 (2005).

268. Funaki, S. *et al.* Global DNA hypomethylation coupled to cellular transformation and metastatic ability. *FEBS Lett.* **589**, 4053–4060 (2015).

269. Puig, I. *et al.* TET2 controls chemoresistant slow-cycling cancer cell survival and tumor recurrence. *J. Clin. Invest.* **128**, 3887–3905 (2018).

270. Tumbar, T., Guasch, G., Greco, V. & Blanpain, C. Defining the epithelial stem cell niche in skin. *Science.* **303**, 359–363 (2004).

271. Wilson, A. *et al.* Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair. *Cell* **135**, 1118–1129 (2008).

272. Buczacki, S. J. A. a *et al.* Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* **495**, 65–9 (2013).

273. dos Santos, C. O. *et al.* Molecular hierarchy of mammary differentiation yields refined markers of mammary stem cells. *Proc. Natl. Acad. Sci.* **110**, 7123–7130 (2013).

274. Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* **30**, 733–750 (2016).

275. Mercher, T. *et al.* TET2, a tumor suppressor in hematological disorders. *Biochim. Biophys. Acta - Rev. Cancer* **1825**, 173–177 (2012).

276. Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230–1234 (2007).

277. Messerschmidt, D. M., Knowles, B. B. & Solter, D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev.* **28**, 812–828 (2014).

278. Costa, Y. *et al.* NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* **495**, 370–374 (2013).

279. Villar, D. *et al.* Cooperativity of Stress-Responsive Transcription Factors in Core Hypoxia-Inducible Factor Binding Regions. *PLoS One* **7**, e45708 (2012).

280. Grabmaier, K., de Weijert, M. C., Verhaegh, G. W., Schalken, J. A. & Oosterwijk, E. Strict regulation of CAIXG250/MN by HIF-1α in clear cell renal cell carcinoma. *Oncogene* **23**, 5624–5631 (2004).

281. Yildirim, H., Karaman, M. & Köçkar, F. The role of hypoxia response element in TGFβ-induced carbonic anhydrase IX expression in Hep3B human hepatoma cells. *Arch. Biol. Sci.* **69**, 593–601 (2017).

282. Korinek, V. *et al.* Two Members of the Tcf Family Implicated in Wnt/β-Catenin Signaling during Embryogenesis in the Mouse. *Mol. Cell. Biol.* **18**, 1248–1256 (1998).

283. Ishitani, T. *et al.* The TAK1–NLK–MAPK-related pathway antagonizes signalling between β-catenin and transcription factor TCF. *Nature* **399**, 798–802 (1999).

284. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **3**, 1101–1108 (2008).

285. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, RESEARCH0034 (2002).

286. Duan, F. *et al.* Area under the curve as a tool to measure kinetics of tumor growth in experimental animals. *J. Immunol. Methods* **382**, 224–228 (2012).

287. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).

288. Puig, I. *et al.* A Personalized Preclinical Model to Evaluate the Metastatic Potential of Patient-Derived Colon Cancer Initiating Cells. *Clin. Cancer Res.* **19**, 6787–6801 (2013).

289. Kim, R. S. *et al.* Dormancy Signatures and Metastasis in Estrogen Receptor Positive and Negative Breast Cancer. *PLoS One* **7**, e35569 (2012).

290. Rhodes, D. R. *et al.* Oncomine 3.0: Genes, Pathways, and Networks in a Collection of 18,000 Cancer Gene Expression Profiles. *Neoplasia* **9**, 166–180 (2007).

291. Cortazar, A. R. *et al.* CANCERTOOL: A Visualization and Representation Interface to Exploit Cancer Datasets. *Cancer Res.* **78**, 6320–6328 (2018).

292. Richman, P. I. & Bodmer, W. F. Control of differentiation in human colorectal carcinoma cell lines: Epithelial—mesenchymal interactions. *J. Pathol.* **156**, 197–211 (1988).

293. Ahler, E. *et al.* Doxycycline Alters Metabolism and Proliferation of Human Cell Lines. *PLoS One* **8**, e64561 (2013).

294. Onoda, T. *et al.* Doxycycline inhibits cell proliferation and invasive potential: Combination therapy with cyclooxygenase-2 inhibitor in human colorectal cancer cells. *J. Lab. Clin. Med.* **143**, 207–216 (2004).

295. Boj, S. F. *et al.* Organoid Models of Human and Mouse Ductal Pancreatic Cancer. *Cell* **160**, 324–338 (2015).

296. Hubert, C. G. *et al.* A Three-Dimensional Organoid Culture System Derived from Human Glioblastomas Recapitulates the Hypoxic Gradients and Cancer Stem Cell Heterogeneity of Tumors Found In Vivo. *Cancer Res.* **76**, 2465–2477 (2016).

297. McCracken, K. W. *et al.* Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature* **516**, 400–404 (2014).

298. Shamir, E. R. & Ewald, A. J. Three-dimensional organotypic culture: experimental models of mammalian biology and disease. *Nat. Rev. Mol. Cell Biol.* **15**, 647–664 (2014).

299. Yeung, T. M., Gandhi, S. C., Wilding, J. L., Muschel, R. & Bodmer, W. F. Cancer stem cells from colorectal cancer-derived cell lines. *Proc Natl Acad Sci U S A* **107**, 3722–3727 (2010).

300. Gantt, K. R. *et al.* The regulation of glucose transporter (GLUT1) expression by the RNA binding protein HuR. *J. Cell. Biochem.* **99**, 565–574 (2006).

301. Gorospe, M., Tominaga, K., Wu, X., Fähling, M. & Ivan, M. Post-Transcriptional Control of the Hypoxic Response by RNA-Binding Proteins and MicroRNAs. *Front. Mol. Neurosci.* **4**, (2011).

302. Jafarifar, F., Yao, P., Eswarappa, S. M. & Fox, P. L. Repression of VEGFA by CA-rich element-binding microRNAs is modulated by hnRNP L. *EMBO J.* **30**, 1324–1334 (2011).

303. De Meester, C. *et al.* Role of AMP-activated protein kinase in regulating hypoxic survival and

proliferation of mesenchymal stem cells. *Cardiovasc. Res.* **101**, 20–29 (2014).

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

305. Groulx, I. & Lee, S. Oxygen-Dependent Ubiquitination and Degradation of Hypoxia-Inducible Factor Requires Nuclear-Cytoplasmic Trafficking of the von Hippel-Lindau Tumor Suppressor Protein. *Mol. Cell. Biol.* **22**, 5319–5336 (2002).

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

307. Funaki, S. *et al.* Inhibition of maintenance DNA methylation by Stella. *Biochem. Biophys. Res. Commun.* **453**, 455–460 (2014).

308. Mulholland, C. B. *et al.* TET1 drives global DNA demethylation via DPPA3-mediated inhibition of maintenance methylation. *bioRxiv* 321604 (2018).

309. Li, Y. *et al.* Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. *Nature* **564**, 136–140 (2018).

310. Mariani, C. J. *et al.* TET1-mediated hydroxymethylation facilitates hypoxic gene induction in neuroblastoma. *Cell Rep.* **7**, 1343–1352 (2014).

311. Kim, J. H. *et al.* Downregulation of UHRF1 increases tumor malignancy by activating the CXCR4/AKT-JNK/IL-6/Snail signaling axis in hepatocellular carcinoma cells. *Sci. Rep.* **7**, 1–16 (2017).

312. Horiuchi, A. *et al.* Hypoxia upregulates ovarian cancer invasiveness via the binding of HIF-1α to a hypoxia-induced, methylation-free hypoxia response element of S100A4 gene. *Int. J. Cancer* **131**, 1755–1767 (2012).

313. Vaupel, P., Höckel, M. & Mayer, A. Detection and Characterization of Tumor Hypoxia Using pO 2 Histography. *Antioxid. Redox Signal.* **9**, 1221–1236 (2007).

314. Marignol, L., Coffey, M., Lawler, M. & Hollywood, D. Hypoxia in prostate cancer: A powerful shield against tumour destruction? *Cancer Treat. Rev.* **34**, 313–327 (2008).

315. Prigent, C. Phosphorylation of serine 10 in histone H3, what for? *J. Cell Sci.* **116**, 3677–3685 (2003).

316. Miller, M. W. & Nowakowski, R. S. Use of bromodeoxyuridine-immunohistochemistry to examine the proliferation, migration and time of origin of cells in the central nervous system. *Brain Res.* **457**, 44–52 (1988).

317. Hans, F. & Dimitrov, S. Histone H3 phosphorylation and cell division. *Oncogene* **20**, 3021–3027 (2001).

318. Calonge, T. M. & O'Connell, M. J. Turning off the G2 DNA damage checkpoint. *DNA Repair (Amst).* **7**, 136–140 (2008).

319. Sharma, A., Singh, K. & Almasan, A. Histone H2AX phosphorylation: A marker for DNA damage. *Methods Mol. Biol.* **920**, 613–26 (2012).

320. Khandrika, L. *et al.* Hypoxia-associated p38 mitogen-activated protein kinase-mediated androgen receptor activation and increased HIF-1α levels contribute to emergence of an aggressive phenotype in prostate cancer. *Oncogene* **28**, 1248–1260 (2009).

321. Lin, S.-P. *et al.* Survival of Cancer Stem Cells under Hypoxia and Serum Depletion via Decrease in PP2A Activity and Activation of p38-MAPKAPK2-Hsp27. *PLoS One* **7**, e49605 (2012).

322. Xu, L. Hypoxia-Induced Activation of p38 Mitogen-Activated Protein Kinase and Phosphatidylinositol 3'-Kinase Signaling Pathways Contributes to Expression of Interleukin 8 in Human Ovarian Carcinoma Cells. *Clin. Cancer Res.* **10**, 701–707 (2004).

323. Kim, J.-Y., Kim, Y.-J., Lee, S. & Park, J.-H. The critical role of ERK in death resistance and invasiveness of hypoxia-selected glioblastoma cells. *BMC Cancer* **9**, 27 (2009).

324. Lee, K. H., Choi, E. Y., Hyun, M. S. & Kim, J.-R. Involvement of MAPK pathway in hypoxia-induced up-regulation of urokinase plasminogen activator receptor in a human prostatic cancer cell line, PC3MLN4. *Exp. Mol. Med.* **36**, 57–64 (2004).

325. Liu, L. *et al.* ERK/MAPK activation involves hypoxia-induced MGr1-Ag/37LRP expression and contributes to apoptosis resistance in gastric cancer. *Int. J. Cancer* **127**, 820–9 (2010).

326. Kobayashi, A. *et al.* Bone morphogenetic protein 7 in dormancy and metastasis of prostate cancer stem-like cells in bone. *J. Exp. Med.* **209**, 639–639 (2012).

327. Gray, L. H., Conger, A. D., Ebert, M., Hornsey, S. & Scott, O. C. A. The Concentration of Oxygen Dissolved in Tissues at the Time of Irradiation as a Factor in Radiotherapy. *Br. J. Radiol.* **26**, 638–648 (1953).

328. Roizin-Towle, L. & Hall, E. J. Studies with bleomycin and misonidazole on aerated and hypoxic cells. *Br. J. Cancer* **37**, 254–260 (1978).

329. Wigerup, C., Påhlman, S. & Bexell, D. Therapeutic targeting of hypoxia and hypoxia-inducible factors in cancer. *Pharmacol. Ther.* **164**, 152–169 (2016).

330. Begicevic, R.-R. & Falasca, M. ABC Transporters in Cancer Stem Cells: Beyond Chemoresistance. *Int. J. Mol. Sci.* **18**, 2362 (2017).

331. Cotsarelis, G., Sun, T.-T. & Lavker, R. M. Label-retaining cells reside in the bulge area of pilosebaceous unit: Implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell* **61**, 1329–1337 (1990).

332. Veldman-Jones, M. H. *et al.* Evaluating Robustness and Sensitivity of the NanoString

Technologies nCounter Platform to Enable Multiplexed Gene Expression Analysis of Clinical Samples. *Cancer Res.* **75**, 2587–2593 (2015).

333. Santoyo-Ramos, P., Likhatcheva, M., García-Zepeda, E. A., Cristina Castañeda-Patlan, M. & Robles-Flores, M. Hypoxia-inducible factors modulate the Stemness and malignancy of colon cancer cells by playing opposite roles in canonical Wnt signaling. *PLoS One* **9**, (2014).

334. Wang, P., Wan, W., Xiong, S.-L., Feng, H. & Wu, N. Cancer stem-like cells can be induced through dedifferentiation under hypoxic conditions in glioma, hepatoma and lung cancer. *Cell Death Discov.* **3**, 16105 (2017).

335. Yeung, T. M., Gandhi, S. C. & Bodmer, W. F. Hypoxia and lineage specification of cell line-derived colorectal cancer stem cells. *Proc Natl Acad Sci U S A* **108**, 4382–4387 (2011).

336. VanDussen, K. L. *et al.* Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells. *Development* **139**, 488–497 (2012).

337. Hu, Q.-D. *et al.* F3/Contactin Acts as a Functional Ligand for Notch during Oligodendrocyte Maturation. *Cell* **115**, 163–175 (2003).

338. Ku, M. *et al.* Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains. *PLoS Genet.* **4**, e1000242 (2008).

339. Chiacchiera, F. *et al.* Polycomb Complex PRC1 Preserves Intestinal Stem Cell Identity by Sustaining Wnt/β-Catenin Transcriptional Activity. *Cell Stem Cell* **18**, 91–103 (2016).

340. Conway, E. M. & Bracken, A. P. Unraveling the Roles of Canonical and Noncanonical PRC1 Complexes. in *Polycomb Group Proteins* 57–80 (Elsevier Inc., 2017).

341. Richly, H., Aloia, L. & Di Croce, L. Roles of the Polycomb group proteins in stem cells and cancer. *Cell Death Dis.* **2**, e204 (2011).

342. Hepburn, A. C. *et al.* The induction of core pluripotency master regulators in cancers defines poor clinical outcomes and treatment resistance. *Oncogene* **38**, 4412–4424 (2019).

343. Iglesias, J. M., Gumuzio, J. & Martin, A. G. Linking Pluripotency Reprogramming and Cancer. *Stem Cells Transl. Med.* **6**, 335–339 (2017).

344. Clark, A. T. The Stem Cell Identity of Testicular Cancer. *Stem Cell Rev.* **3**, 49–59 (2007).

345. Fluegen, G. *et al.* Phenotypic heterogeneity of disseminated tumour cells is preset by primary tumour hypoxic microenvironments. *Nat. Cell Biol.* **19**, 120–132 (2017).

346. McGary, E. C., Rondon, I. J. & Beckman, B. S. Post-transcriptional Regulation of Erythropoietin mRNA Stability by Erythropoietin mRNA-binding Protein. *J. Biol. Chem.* **272**, 8628–8634 (1997).

347. Gingras, A.-C. *et al.* Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes Dev.* **13**, 1422–1437 (1999).

348. Schmitt, A. M. *et al.* VHL inactivation is an important pathway for the development of malignant sporadic pancreatic endocrine tumors. *Endocr. Relat. Cancer* **16**, 1219–1227 (2009).

349. Henze, A.-T. & Acker, T. Feedback regulators of hypoxia-inducible factors and their role in cancer biology. *Cell Cycle* **9**, 2821–2835 (2010).

350. Taguchi, A. *et al.* Identification of Hypoxia-Inducible Factor-1 as a Novel Target for miR-17-92 MicroRNA Cluster. *Cancer Res.* **68**, 5540–5545 (2008).

351. del Peso, L. *et al.* The von Hippel Lindau/Hypoxia-inducible Factor (HIF) Pathway Regulates the Transcription of the HIF-Proline Hydroxylase Genes in Response to Low Oxygen. *J. Biol. Chem.* **278**, 48690–48695 (2003).

352. Uchida, T. *et al.* Prolonged Hypoxia Differentially Regulates Hypoxia-inducible Factor (HIF)-1α and HIF-2α Expression in Lung Epithelial Cells. *J. Biol. Chem.* **279**, 14871–14878 (2004).

353. Alfranca, A. *et al.* c-Jun and Hypoxia-Inducible Factor 1 Functionally Cooperate in Hypoxia-Induced Gene Transcription. *Mol. Cell. Biol.* **22**, 12–22 (2002).

354. Textor, B., Sator-Schmitt, M., Richter, K. H., Angel, P. & Schorpp-Kistner, M. c-Jun and JunB Are Essential for Hypoglycemia-Mediated VEGF Induction. *Ann. N. Y. Acad. Sci.* **1091**, 310–8 (2006).

355. Liu, Q. *et al.* Hypoxia Induces Genomic DNA Demethylation through the Activation of HIF-1 and Transcriptional Upregulation of MAT2A in Hepatoma Cells. *Mol. Cancer Ther.* **10**, 1113–1123 (2011).

356. Dutta, B., Yan, R., Lim, S. K., Tam, J. P. & Sze, S. K. Quantitative Profiling of Chromatome Dynamics Reveals a Novel Role for HP1BP3 in Hypoxia-induced Oncogenesis. *Mol. Cell. Proteomics* **13**, 3236–3249 (2014).

357. Skowronski, K., Dubey, S., Rodenhiser, D. I. & Coomber, B. Ischemia dysregulates DNA methyltransferases and p16INK4a methylation in human colorectal cancer cells. *Epigenetics* **5**, 547–556 (2010).

358. Nishiyama, A. *et al.* Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* **502**, 249–253 (2013).

359. Qin, W. *et al.* DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. *Cell Res.* **25**, 911–929 (2015).

360. Squazzo, S. L. Suz12 binds to silenced regions of the genomein a cell-type-specific manner. *Genome Res.* **16**, 890–900 (2006).

361. Barski, A. *et al.* High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell* **129**, 823–837 (2007).

362. Lienert, F. *et al.* Genomic Prevalence of Heterochromatic H3K9me2 and Transcription Do Not

Discriminate Pluripotent from Terminally Differentiated Cells. *PLoS Genet.* **7**, e1002090 (2011).

363. Voon, H. P. J. & Gibbons, R. J. Maintaining memory of silencing at imprinted differentially methylated regions. *Cell. Mol. Life Sci.* **73**, 1871–1879 (2016).

364. Fraga, A., Ribeiro, R., Príncipe, P., Lopes, C. & Medeiros, R. Hypoxia and Prostate Cancer Aggressiveness: A Tale With Many Endings. *Clin. Genitourin. Cancer* **13**, 295–301 (2015).

365. Gardner, L. B. *et al.* Hypoxia Inhibits G 1 /S Transition through Regulation of p27 Expression. *J. Biol. Chem.* **276**, 7919–7926 (2001).

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

367. Lim, J.-H. *et al.* Bafilomycin Induces the p21-Mediated Growth Inhibition of Cancer Cells under Hypoxic Conditions by Expressing Hypoxia-Inducible Factor-1α. *Mol. Pharmacol.* **70**, 1856–1865 (2006).

368. Ravizza, R., Gariboldi, M. B., Passarelli, L. & Monti, E. Role of the p53/p21 system in the response of human colon carcinoma cells to Doxorubicin. *BMC Cancer* **4**, 92 (2004).

369. Barberi-Heyob, M. *et al.* Wild-type p53 gene transfer into mutated p53 HT29 cells improves sensitivity to photodynamic therapy via induction of apoptosis. *Int. J. Oncol.* **24**, 951–8 (2004).

370. Taylor, W. R. & Stark, G. R. Regulation of the G2/M transition by p53. *Oncogene* **20**, 1803–1815 (2001).

371. Zhang, H., Chen, W., Duan, C.-J. & Zhang, C.-F. Overexpression of HSPA2 is correlated with poor prognosis in esophageal squamous cell carcinoma. *World J. Surg. Oncol.* **11**, 141 (2013).

372. Bretones, G., Delgado, M. D. & León, J. Myc and cell cycle control. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1849**, 506–516 (2015).

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

374. Tien, A. L. L. *et al.* UHRF1 depletion causes a G2/M arrest, activation of DNA damage response and apoptosis. *Biochem. J.* **435**, 175–185 (2011).

375. Abusnina, A., Keravis, T., Yougbaré, I., Bronner, C. & Lugnier, C. Anti-proliferative effect of curcumin on melanoma cells is mediated by PDE1A inhibition that regulates the epigenetic integrator UHRF1. *Mol. Nutr. Food Res.* **55**, 1677–1689 (2011).

376. Ashraf, W. *et al.* The epigenetic integrator UHRF1: on the road to become a universal biomarker for cancer. *Oncotarget* **8**, (2017).

377. Hopfner, R. *et al.* ICBP90, a novel human CCAAT binding protein, involved in the regulation of topoisomerase IIα expression. *Cancer Res.* **60**, 121–8 (2000).

378. Nobre, A. R., Entenberg, D., Wang, Y., Condeelis, J. & Aguirre-Ghiso, J. A. The Different Routes to Metastasis via Hypoxia-Regulated Programs. *Trends Cell Biol.* **28**, 941–956 (2018).

379. Yang, X. *et al.* Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell* **26**, 577–590 (2014).

380. Hüsemann, Y. *et al.* Systemic Spread Is an Early Step in Breast Cancer. *Cancer Cell* **13**, 58–68 (2008).

381. Choi, C. H. ABC transporters as multidrug resistance mechanisms and the development of chemosensitizers for their reversal. *Cancer Cell Int.* **5**, 30 (2005).

382. Murono, K. *et al.* SN-38 overcomes chemoresistance of colorectal cancer cells induced by hypoxia, through HIF1alpha. *Anticancer Res.* **32**, 865–872 (2012).

383. Pencreach, E. *et al.* Marked Activity of Irinotecan and Rapamycin Combination toward Colon Cancer Cells In vivo and In vitro Is Mediated through Cooperative Modulation of the Mammalian Target of Rapamycin/Hypoxia-Inducible Factor-1 Axis. *Clin. Cancer Res.* **15**, 1297–1307 (2009).

384. Margonis, G. A. *et al.* Association of BRAF Mutations With Survival and Recurrence in Surgically Treated Patients With Metastatic Colorectal Liver Cancer. *JAMA Surg.* **153**, e180996 (2018).

385. Kemeny, N. E. *et al.* KRAS mutation influences recurrence patterns in patients undergoing hepatic resection of colorectal metastases. *Cancer* **120**, 3965–3971 (2014).

386. Almog, N. *et al.* Transcriptional Switch of Dormant Tumors to Fast-Growing Angiogenic Phenotype. *Cancer Res.* **69**, 836–844 (2009).

387. Román-Trufero, M. *et al.* Maintenance of Undifferentiated State and Self-Renewal of Embryonic Neural Stem Cells by Polycomb Protein Ring1B. *Stem Cells* **27**, 1559–1570 (2009).

388. Wilson, A. & Radtke, F. Multiple functions of Notch signaling in self-renewing organs and cancer. *FEBS Lett.* **580**, 2860–2868 (2006).

389. Levine, A. J. GDF3, a BMP inhibitor, regulates cell fate in stem cells and early embryos. *Development* **133**, 209–216 (2005).

390. Li, Q., Ling, Y. & Yu, L. GDF3 inhibits the growth of breast cancer cells and promotes the apoptosis induced by Taxol. *J. Cancer Res. Clin. Oncol.* **138**, 1073–1079 (2012).

391. Ramírez-Bergeron, D. L. & Simon, M. C. Hypoxia-Inducible Factor and the Development of Stem Cells of the Cardiovascular System. *Stem Cells* **19**, 279–286 (2001).

392. Keith, B. & Simon, M. C. Hypoxia-Inducible Factors, Stem Cells, and Cancer. *Cell* **129**, 465–472 (2007).

393. Axelson, H., Fredlund, E., Ovenberger, M., Landberg, G. & Påhlman, S. Hypoxia-induced dedifferentiation of tumor cells – A mechanism behind heterogeneity and aggressiveness of solid

tumors. *Semin. Cell Dev. Biol.* **16**, 554–563 (2005).

394. Prigione, A. *et al.* HIF1α Modulates Cell Fate Reprogramming Through Early Glycolytic Shift and Upregulation of PDK1-3 and PKM2. *Stem Cells* **32**, 364–376 (2014).

395. Takahashi, N. *et al.* Hypoxia Inducible Factor 1 Alpha Is Expressed in Germ Cells throughout the Murine Life Cycle. *PLoS One* **11**, e0154309 (2016).

396. Gustafsson, M. V. *et al.* Hypoxia Requires Notch Signaling to Maintain the Undifferentiated Cell State. *Dev. Cell* **9**, 617–628 (2005).