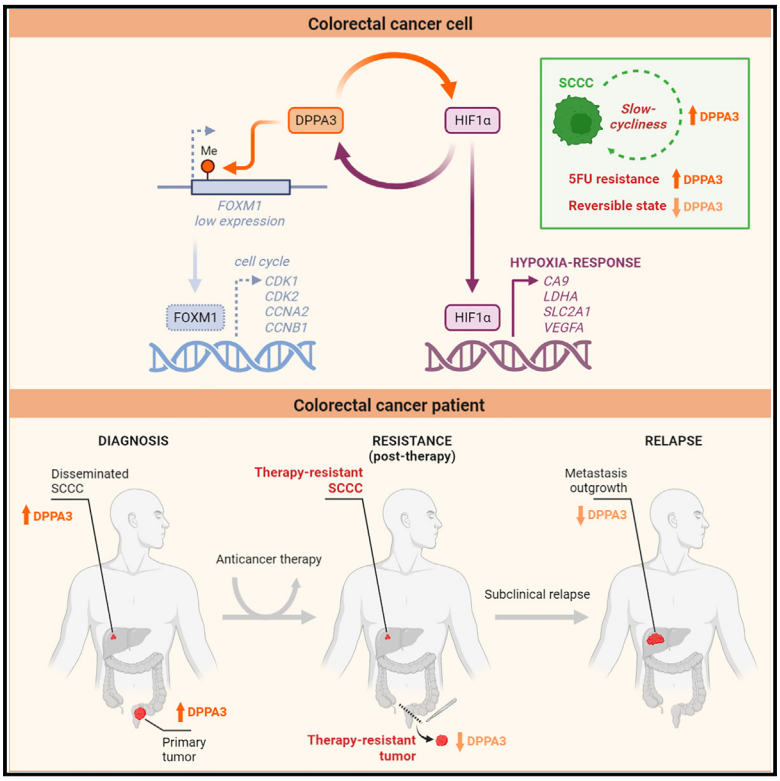
Cell Reports

Article

DPPA3-HIF1α axis controls colorectal cancer chemoresistance by imposing a slow cell-cycle phenotype

Graphical abstract



Highlights

- Overexpression of DPPA3 predicts disease progression in colorectal cancer patients
- DPPA3 is required for slow-cycling cancer cell chemoresistance
- DPPA3 mediates "slow-cycliness" by downregulating *FOXM1* expression
- *HIF1A* knockdown partly rescues cell-cycle progression of DPPA3-overexpressing cells

Authors

Estefania Cuesta-Borràs, Cándida Salvans, Oriol Arqués, ..., Jaume Capdevila, Héctor G. Palmer, Isabel Puig

Correspondence

hgpalmer@vhio.net (H.G.P.), ipuig@vhio.net (I.P.)

In brief

Cuesta-Borràs et al. show that chemotherapy induces a hypoxia response by increasing DPPA3 levels in cancer cells, which in turn promotes a chemoresistant proliferative-delayed state. Our findings identify DPPA3 as a molecular node linking hypoxia response with the acquisition of slow-cycling capacity that fosters cancer cell persistence.

Cuesta-Borràs et al., 2023, Cell Reports *42*, 112927
August 29, 2023
https://doi.org/10.1016/j.celrep.2023.112927

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

DPPA3-HIF1α axis controls colorectal cancer chemoresistance by imposing a slow cell-cycle phenotype

Estefania Cuesta-Borràs,[1,9] Cándida Salvans,[1,8,9] Oriol Arqués,[1] Irene Chicote,[1,2] Lorena Ramírez,[3] Laia Cabellos,[1] Jordi Martínez-Quintanilla,[1] Alex Mur-Espinosa,[1,8] Alejandro García-Álvarez,[3] Jorge Hernando,[3] Juan Ramón Tejedor,[4] Oriol Mirallas,[3] Elena Élez,[2,3] Mario F. Fraga,[4] Josep Tabernero,[2,3,6] Paolo Nuciforo,[2,5] Jaume Capdevila,[3,7] Héctor G. Palmer,[1,2,10,*] and Isabel Puig[1,2,10,11,*]

[1]Stem Cells and Cancer Laboratory, Vall d'Hebron Institute of Oncology (VHIO), 08035 Barcelona, Spain
[2]CIBERONC, 08029 Madrid, Spain
[3]Gastrointestinal and Endocrine Tumors Group, Medical Oncology Department, Vall d'Hebron University Hospital (HUVH), Vall d'Hebron Institute of Oncology (VHIO), Universitat Autònoma de Barcelona (UAB), 08035 Barcelona, Spain
[4]Nanomaterials and Nanotechnology Research Center (CINN), Spanish National Research Council (CSIC), Health Research Institute of the Principality of Asturias (ISPA), Spanish Biomedical Research Network in Rare Diseases (CIBERER), Institute of Oncology of Asturias (IUOPA), University of Oviedo, 33011 Oviedo, Asturias, Spain
[5]Molecular Oncology Group, Vall d'Hebron Institute of Oncology (VHIO), 08035 Barcelona, Spain
[6]UVic-UCC, IOB-Quiron, 08023 Barcelona, Spain
[7]IOB-Teknon, 08023 Barcelona, Spain
[8]University of Barcelona, Barcelona, Spain
[9]These authors contributed equally
[10]These authors contributed equally
[11]Lead contact
*Correspondence: hgpalmer@vhio.net (H.G.P.), ipuig@vhio.net (I.P.)
https://doi.org/10.1016/j.celrep.2023.112927

SUMMARY

Tumor relapse is linked to rapid chemoresistance and represents a bottleneck for cancer therapy success. Engagement of a reduced proliferation state is a non-mutational mechanism exploited by cancer cells to bypass therapy-induced cell death. Through combining functional pulse-chase experiments in engineered cells and transcriptomic analyses, we identify DPPA3 as a master regulator of slow-cycling and chemoresistant phenotype in colorectal cancer (CRC). We find a vicious DPPA3-HIF1α feedback loop that downregulates *FOXM1* expression via DNA methylation, thereby delaying cell-cycle progression. Moreover, downregulation of HIF1α partially restores a chemosensitive proliferative phenotype in DPPA3-overexpressing cancer cells. In cohorts of CRC patient samples, DPPA3 overexpression acts as a predictive biomarker of chemotherapeutic resistance that subsequently requires reduction in its expression to allow metastatic outgrowth. Our work demonstrates that slow-cycling cancer cells exploit a DPPA3/HIF1α axis to support tumor persistence under therapeutic stress and provides insights on the molecular regulation of disease progression.

INTRODUCTION

Most colorectal cancer (CRC) patients acquire a relatively rapid chemoresistance manifested as disease relapse, remaining a key obstacle to successful cancer therapy. Chemotherapy frequently leaves residual cancer cells that provide a reservoir from which recurrence arises. The term "persister" in bacteria describes a transient state that tolerates antibiotic treatment for a longer period than the majority of the rapidly killed population.[1] This concept was adopted in cancer studies, where cancer cells evading treatment are called drug-tolerant persister cells.[2] Persister cancer cells resist anticancer therapy by adopting a slow-cycling state that reverts when insults cease. Importantly, their capacity to resist therapy is caused by the slow-cycling phenotype, which is a non-mutational mechanism of therapy resistance, instead of genetic mutations. However, uncertainty remains in determining whether persister cancer cells are pre-existing or induced by chemotherapy, and which signaling pathways regulate their phenotype. Therefore, "persister cancer cells" refers to the therapeutic setting, whereas "slow-cycling cancer cells" (SCCCs) is used regardless of the therapeutic context.[3]

Unfortunately, although the mechanisms governing cancer persistence have improved our understanding,[3] they have not yet been successfully targeted in the clinical setting. The therapy-resistant slow-cycling phenotype of persister cancer cells could be ruled by specific developmental gene expression programs that protect cells from environmental stressors.[4,5]

A single-cell transcriptomic study performed in chemoresistant cancer cells identified a hypoxia-response program upregulation, suggesting that cancer persistence may reflect an adaptive stress response.[6] Accordingly, we and others observed a relationship between the hypoxia response and SCCCs.[7,8] Despite that the link of hypoxia to therapy resistance has been widely documented in the literature,[9] the role of hypoxia in regulating slow-cycling phenotype and chemoresistance remains poorly understood.

Here, by exploring a distinctive germline gene set in SCCCs, we identified DPPA3 (Developmental Pluripotency Associated 3) as a key regulator of "slow-cycliness" and chemoresistance in CRC cells. We uncovered a positive DPPA3-HIF1α feedback loop involved in disease progression and highlighted HIF1α as a potential vulnerability to target chemoresistant DPPA3-overexpressing cancer cells. We further confirmed this evidence by generating different cell line models and interrogating CRC patient samples. Altogether, our findings uncover the biological insights on the functional relevance of hypoxia signaling in the modulation of slow-cycling chemoresistant cancer cell phenotype.

RESULTS

DPPA3 is overexpressed in SCCCs and is downregulated in liver metastatic outgrowth

Diapause is a powerful tool to explore fundamental questions about the relationship between stress resistance and "slow-cycliness."[2,4,5] In embryonic diapause, primordial germ cells (GCs) pause their proliferation to undergo epigenetic reprogramming.[10] Well-identified factors involved in this process[11] became promising candidates for governing a slow-cycling phenotype leading to chemoresistance.

In doxycycline (DOX) pulse-chase experiments, only SCCCs retain the H2BeGFP label[8,12] (Figure 1A). We compared the enrichment of a GC-related gene set in GFP-labeled SCCCs versus rapid-cycling cancer cells (RCCCs) isolated in previously performed three-dimensional (3D) pulse-chase experiments[8] (Figures 1B, S1A, and S1B). Gene set enrichment analysis (GSEA) revealed higher expression of nine leading-edge core genes related to GCs in SCCCs versus RCCCs (Figure 1B). Interestingly, our group previously identified *TET2* as a key to SCCC intrinsic survival.[8] Therefore, identifying it among the leading-edge genes validated our approach to discover regulators of SCCC phenotype.

From the different genes identified, *APOBEC1* and *DPPA3* are described as regulatory genes of DNA methylation reprogramming. Given that APOBEC1 is a cytidine deaminase classified as DNA mutator,[13] it was ruled out as a candidate for regulating the slow-cycling phenotype as a non-mutational mechanism of therapy resistance. The naive pluripotency factor DPPA3 is involved not only in the DNA methylation reprogramming in primordial GCs,[14] but also in oocytes,[15] early embryos,[16,17] somatic cells,[18] and hepatocellular carcinoma,[19] suggesting its relevant role across different biological contexts. *DPPA3* expression was confirmed in SCCCs derived from the two CRC models used for the GSEA analysis, a patient-derived xenograft CRC model (T70-PDX) and the CRC-SW1222-H2BeGFP cell line pool (Figure 1C). A selected CRC SW1222-H2BeGFP clone was used to analyze *DPPA3* expression in three different cell proliferative states, SCCCs, RCCCs, and super-rapid-cycling cancer cells (sRCCCs) (Figure S1C). Pulse-chase experiments demonstrated that the higher the proliferation rate, the lower the *DPPA3* expression was, restricting its expression mostly to SCCCs (Figure 1D). As we previously described,[8] SCCCs presented higher self-renewal capacity than RCCCs because of their undifferentiated condition and reversible state (Figures 1E and 1F). Interestingly, downregulation of *DPPA3* by short hairpin in SCCCs (sh*DPPA3*) revealed a higher organoid re-initiation capacity than in control SCCCs without affecting RCCC self-renewal (Figures 1F and S1D). Moreover, we observed that sh*DPPA3* SCCCs produced a high megacolonies ($\geq$400 μm) proportion (Figure 1G). These results suggest that, despite the high self-renewal capacity observed in SCCCs, DPPA3 restricts organoid enlargement, functionally validating its role in regulating the slow-cycling phenotype.

Besides chemoresistance, SCCCs contribute to late disease relapses because of their slow-cycling nature that delays metastasis outgrowth.[20] We analyzed DPPA3 expression in 65 non-treated paired liver metastasis and primary tumor biopsies from a CRC cohort (Figure 1H; Table S1). We observed a decrease in the DPPA3-positive cell percentage in liver metastases compared with paired primary tumors in patients categorized as DPPA3-High (>20% of DPPA3-positive cells in primary tumor biopsies), but no significant changes in DPPA3-Low patients (Figures 1I and S1E; Table S1). In addition, in a subset of 27 metachronous metastases, we confirmed a higher percentage of DPPA3-positive cells in later (>2 years) than earlier ($\leq$2 years) relapsed metastases (Figure 1J). From these results, we concluded that DPPA3 could contribute to restricting metastasis outgrowth from disseminated SCCCs in CRC patients.

DPPA3 overexpression in primary tumors correlates with worse clinical outcomes in treated CRC patients

Considering the chemoresistant nature of SCCCs,[8,21,22] we next sought to analyze the role of DPPA3 in regulating this phenomenon. Standard of care for CRC patients remains 5-fluorouracil (5FU)-based chemotherapy and radiotherapy, both aimed at killing highly proliferative cells. Consistently, pulse-chase experiments confirmed that, while CRC-SW1222 RCCCs underwent 5FU-induced apoptosis, SCCCs survived and increased their proportion upon 5FU treatment (Figures 2A and 2B). Next, we observed that *DPPA3* downregulation sensitized SCCCs, but not RCCCs, to 5FU treatment (Figure 2C), confirming its prominent role in SCCC chemoresistance. Furthermore, the higher colony re-initiation capacity previously observed in sh*DPPA3* compared with short hairpin RNA (shRNA) knockdown control (shCTRL) SCCCs (Figure 1F) was impaired when colonies were treated with 5FU before isolation and re-seeding (Figure 2D). Again, these results demonstrate the key role of DPPA3 in preserving SCCC viability upon chemotherapy. Interestingly, post-5FU-treated shCTRL SCCCs generated a higher proportion of minicolonies (<50 μm) in detriment to megacolonies formation (Figure S1F). In contrast, post-5FU-treated sh*DPPA3* SCCCs produced the opposite proportions (Figure S1F), confirming the requirement to downregulate DPPA3 expression to promote proliferation and colony enlargement.

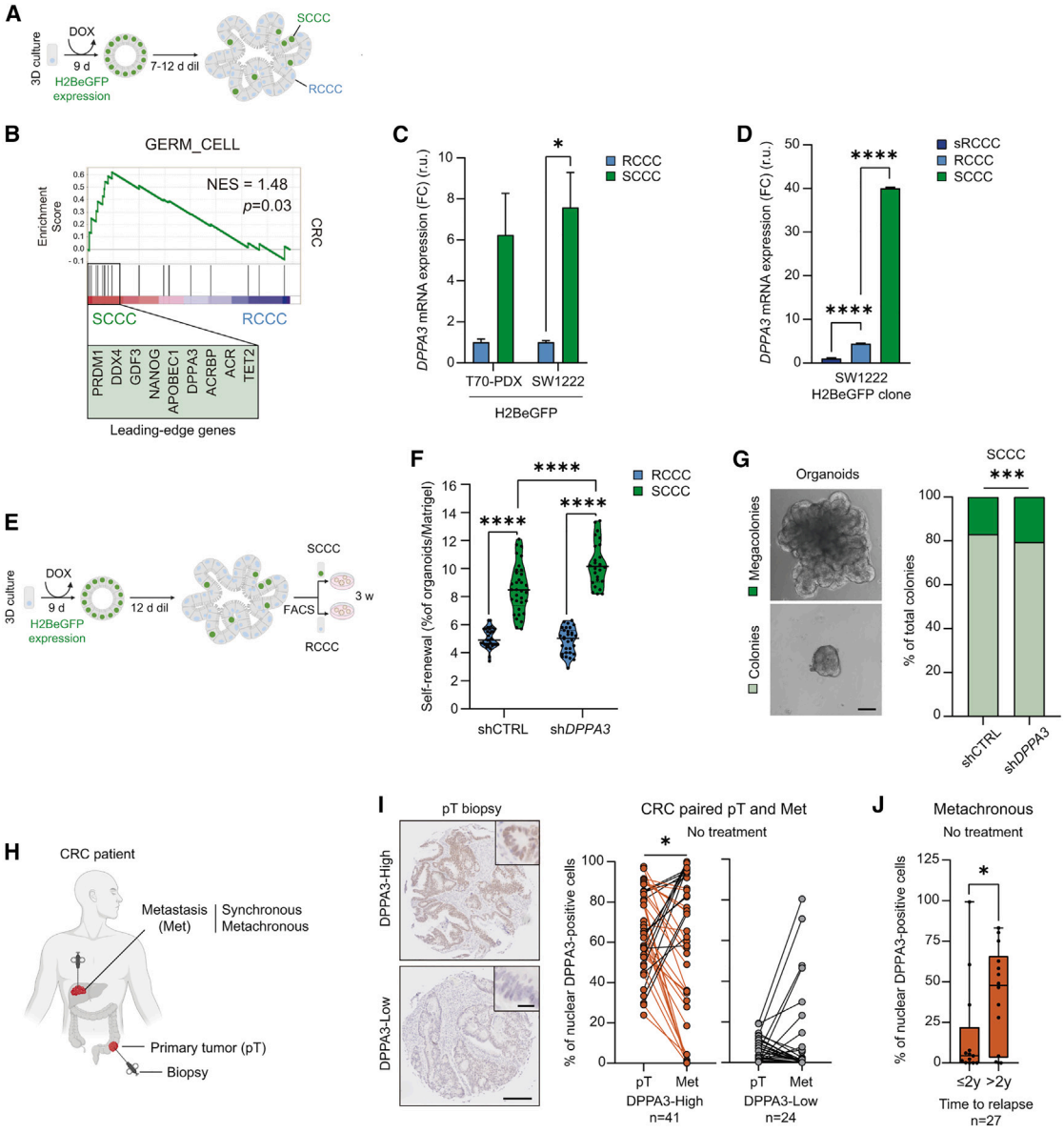


Figure 1. DPPA3 gene is overexpressed in slow-cycling cancer cells (SCCCs)

(A) Pulse-chase experimental design to evaluate SCCCs. After a doxycycline (DOX) treatment, the accumulated H2BeGFP signal in cells was diluted (dil) upon cell divisions revealing label-retaining cells (SCCCs).

(B) Gene set enrichment analysis (GSEA) plot showing enrichment of a custom GERM_CELL gene set in SCCC versus RCCC expression profiles from two CRC models grown in 3D (ArrayExpress: E-MTAB-4004).

(C and D) qRT-PCR analysis showing *DPPA3* expression in SCCCs (C and D), RCCCs (C and D), and super-rapid cycling cancer cells (sRCCCs) (D) obtained from a DOX pulse-chase sorting experiment of each indicated CRC model. Mean ± SD of triplicates.

(E and F) Experimental design (E) and colonies formation capacity evaluation (F) of control (shCTRL) and *DPPA3* knockdown (sh*DPPA3*) RCCCs and SCCCs sorted from CRC SW1222-H2BeGFP pool organoids. Dots indicate the percentage of organoids grown embedded in each single Matrigel.

(G) Representative pictures (left) and percentage (right) of colonies (<400 μm) and megacolonies (≥400 μm) generated from SCCCs in the self-renewal assay shown in (F). Scale bar, 100 μm.

(H) Schematic representation of paired CRC primary tumor (pT) and liver metastasis (Met) biopsies collection.

(I) Representative pictures (left) and immunohistochemical staining quantification (right) of DPPA3 in paired pTs and liver metastases (Met) of CRC patients. Patients were categorized according to the percentage of nuclear DPPA3-positive cells into DPPA3-High (>20%, n = 41) and DPPA3-Low (≤20%, n = 24). Scale bar, 250 μm; high-magnification scale bar, 25 μm.

(J) Percentage of nuclear DPPA3-positive cells in metachronous metastases analyzed in (I) according to the early (≤2 years) or late (>2 years) relapse time of CRC patients. Mean ± SEM.

(C, D, F, G, I, and J) *p ≤ 0.05, ***p ≤ 0.001,****p ≤ 0.0001, unpaired t test (C and J), one-way ANOVA (D), two-way ANOVA (F), chi-square exact test (G), and paired t test (I). NES, normalized enrichment score; *p*, one-way ANOVA p value; RCCCs, rapid-cycling cancer cells. See also Figure S1 and Table S1.

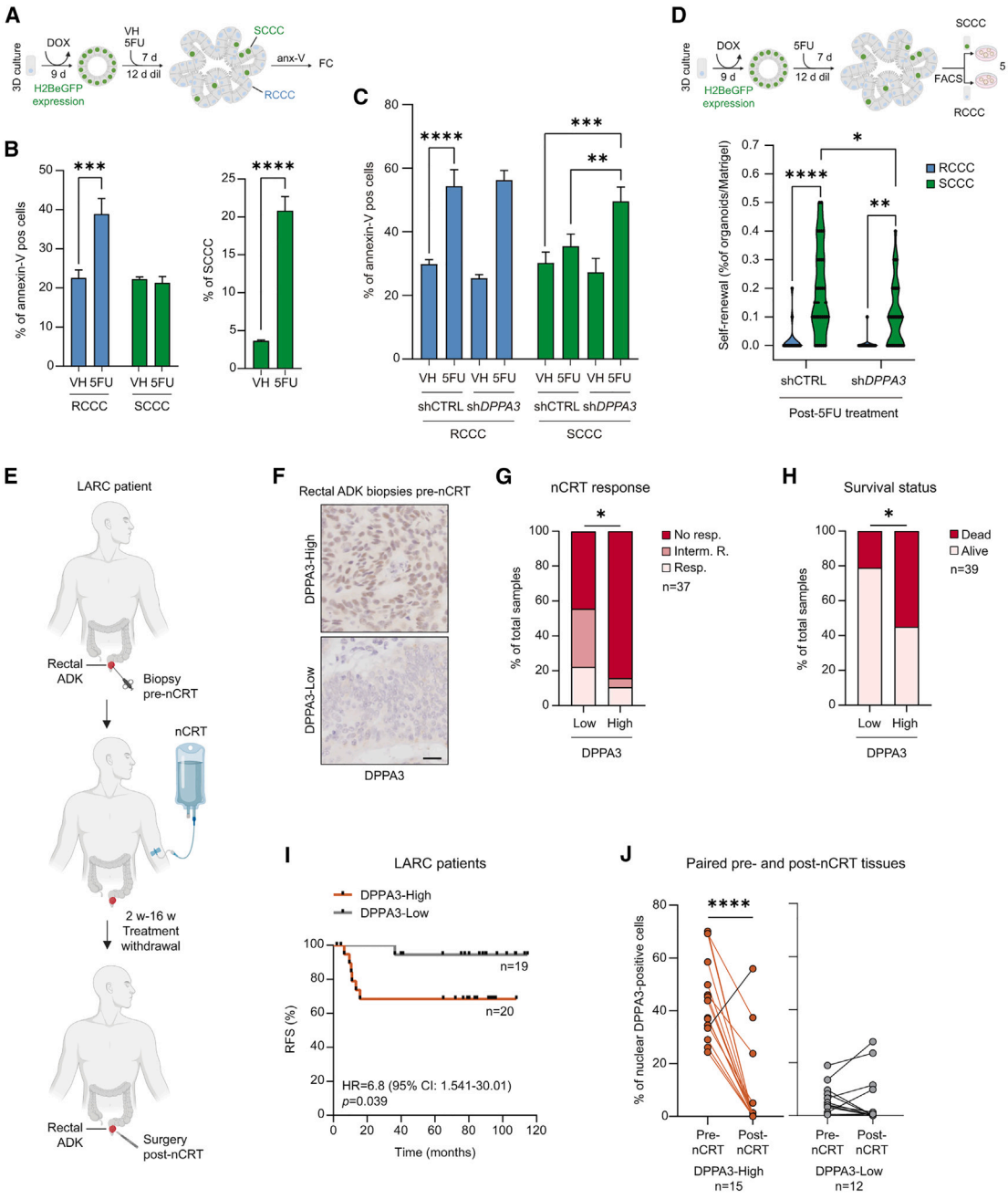


Figure 2. DPPA3 overexpression identifies CRC patients with high risk of disease progression after treatment

(A) Pulse-chase experimental design to evaluate the chemoresistance of SCCCs in CRC cells growing in 3D cultures. After a DOX treatment, the accumulated H2BeGFP signal in cells was diluted upon cell divisions revealing label-retaining cells (SCCCs). Vehicle (VH) or 5-fluorouracil (5FU) was added, after which annexin-V (anx-V) staining was performed and subsequently analyzed by flow cytometry (FC).

(B) Percentage of apoptotic (left) SW1222 RCCCs and SCCCs and proportion of SCCCs (right) upon 5FU exposure. Mean ± SD of triplicates.

(C) Percentage of apoptotic *DPPA3* knockdown (sh*DPPA3*) and control (shCTRL) RCCCs and SCCCs upon 5FU exposure. Mean ± SD of triplicates.

(D) Experimental design (top) and colonies formation capacity evaluation (bottom) of control (shCTRL) and *DPPA3* knockdown (sh*DPPA3*) RCCCs and SCCCs sorted from CRC SW1222-H2BeGFP organoids after 5FU exposure. Dots indicate the percentage of organoids grown embedded in each single Matrigel. Mean ± SD.

(E) Schematic representation of samples collection workflow.

(F) Representative pictures of indicated samples immunostained with DPPA3. Scale bar, 20 μm.

(G and H) Bar plots showing percentage distributions of no responder (No resp.), intermediate responder (Interm. R.) and responder (Resp.) (G), and dead and alive (H) LARC patients according to the percentage of nuclear DPPA3-positive cells in naive biopsies assessed in (F).

(legend continued on next page)

Prompted by these results, we explored the clinical relevance of DPPA3 in CRC patients' chemoresistance. The analysis of *DPPA3* transcript levels in 53 baseline stage II and III CRC tumors showed that high expression of *DPPA3* (*DPPA3*-High, patients with a *DPPA3* mRNA expression value above the overall mean expression) predicted a worse progression-free survival (PFS) in 5FU-treated CRC patients (Figures S1G and S1H). Next, we interrogated human tissue rectal adenocarcinoma biopsies prior to neoadjuvant chemoradiotherapy (pre-nCRT) from a locally advanced rectal cancer (LARC) patient cohort[23] (Figure 2E; Table S2). The DPPA3 analysis revealed a higher proportion of patients with worse outcomes, such as treatment failure and poor disease survival, in the DPPA3-High group (>20% DPPA3-positive cells) compared with the DPPA3-Low group (Figures 2F–2H and S1I). Accordingly, the relapse-free survival (RFS) and PFS curves showed that DPPA3-High patients presented worse outcomes than those categorized as DPPA3-Low (Figures 2I and S1J). Importantly, RFS and PFS rates in patients with a poor response to nCRT showed an equivalent predictive outcome than those presenting high content of DPPA3-positive cells (Figures S1K and S1L), indicating its value as early predictive biomarker of chemoradiotherapy resistance. Altogether, this suggests that the presence of a high proportion DPPA3-positive cells in naive CRC tumors results in disease relapses because of their therapy-resistant nature. Interestingly, the proportion of DPPA3-positive cells was lower in tumors post-treatment than in their corresponding paired pre-treated LARC tumor biopsies (Figure 2J). This result indicates again that tumors need to reduce the proportion of DPPA3-positive cells to resume tumor growth and result in early relapse.

DPPA3 restrains cell proliferation by repressing *FOXM1* expression through DNA methylation

To decipher the underlying molecular mechanism by which DPPA3 regulates the slow-cycling phenotype, we first generated different cellular models. By comparing basal *DPPA3* mRNA expression levels across different CRC cell lines with the embryonal carcinoma cell line NTERA2 (NT2) as positive control, we selected two CRC cell lines, HT29 and SW1222, to generate DPPA3 gain-of-function models (DOX-inducible overexpression) (Figures S2A–S2D). Note that despite high *DPPA3* mRNA expression levels observed in both overexpressing cell lines (Figure S2C), exogenous DPPA3 protein levels were like those of NT2 cells (Figure S2E), indicating that gain-of-function models preserve DPPA3 physiological levels.

Next, we analyzed the gene-expression profile of SW1222-DPPA3 and CTRL cells after 4 h (short-term) and 5 days (long-term) of DPPA3-induction (DOX) in 2D and 3D cultures, respectively (Figure 3A; Table S3). Focusing the GSEA in negatively regulated biological processes in DPPA3 versus CTRL cells, we observed a downregulation of processes related to differentiation, extracellular matrix, GPCR, and ion signaling in the short-term DPPA3-induction condition (Figure 3A). In contrast, long-term DPPA3-induction condition analysis was negatively enriched in cell-cycle, DNA-repair, DNA-replication, translation, and mRNA-processing processes (Figure 3A). Furthermore, the gene-expression profile of long-term DPPA3-induction condition showed a negative enrichment of genes downregulated in the Pan-Cancer SCCC signature,[8] recapitulating the downregulated SCCC transcriptome (Figure 3B). In this line, a DPPA3-dependent cell-cycle gene set (DPPA3_CELL_CYCLE_DW) generated from a leading-edge analysis of cell-cycle-related gene sets negatively enriched in DPPA3 cells was repressed in SCCCs (Figure 3C; Table S4). Interestingly, target genes of the DREAM complex, which has been largely described as a global repressor of cell-cycle genes during quiescence,[24] were downregulated in DPPA3 and SCCCs (Figure 3D). These results suggest that DPPA3 represses the cell-cycle transcriptional program and defines an SCCC identity.

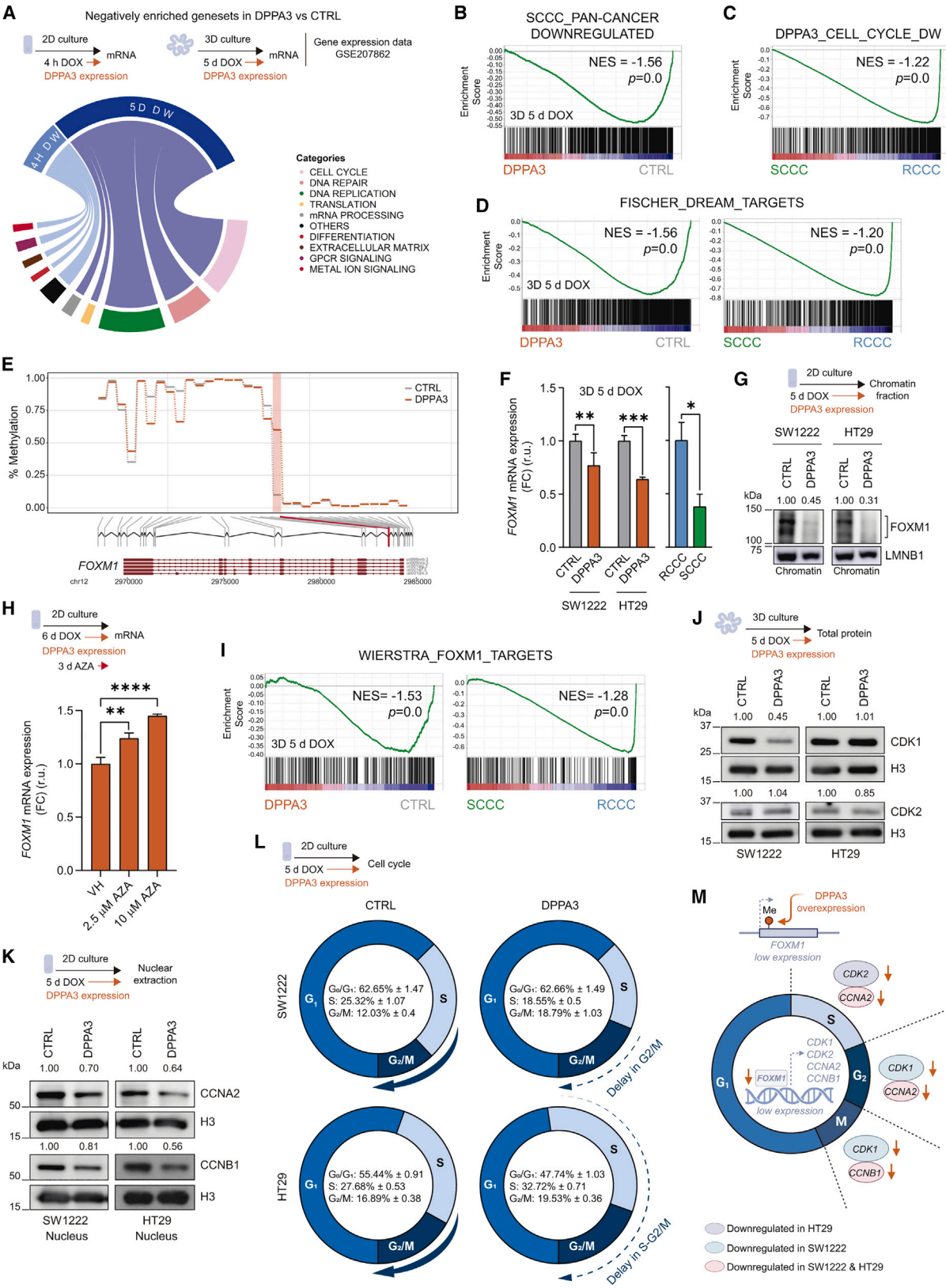
The repression of gene sets related to cell-cycle and DNA replication observed only in the long-term DPPA3-induced condition suggested that time is needed to downregulate proliferation, raising the hypothesis of an epigenetic reprogramming involving DNA methylation. A genome-wide methylome analysis revealed a global demethylation phenotype in DPPA3 long-term overexpressing cells because of decreased chromatin DNMT1 levels, as previously described for DPPA3[15] (Figures S2F–S2H). In addition, DPPA3 overexpression promoted an enrichment of both hypermethylated and hypomethylated CpG sites in distal intergenic regions, but an impairment at promoters (Figure S2I). An integrative analysis between methylation and gene expression data in DPPA3 versus CTRL cells showed a slightly stronger and consistent negative correlation trend between methylation and gene expression in the context of the DPPA3_CELL_CYCLE_DW gene set compared with all other significantly downregulated genes (Figure S2J). A detailed analysis of overall 5mC levels of the *FOXM1* gene, known as a cell-cycle progression stimulator, revealed a significantly hypermethylated probe in its ±1-Kb promoter region when DPPA3 was overexpressed (Figure 3E). Consequently, *FOXM1* mRNA expression was downregulated in SCCCs and SW1222 and HT29 DPPA3 cells, resulting in decreased chromatin FOXM1 levels (Figures 3F and 3G). Finally, DPPA3 cells treated with the DNMT1 inhibitor, 5-azacytidine, confirmed that *FOXM1* repression was methylation dependent (Figure 3H).

Accordingly with these results, FOXM1 target genes[25] were negatively enriched in DPPA3 and SCCC cells (Figures 3I and S2K). Among these genes, we confirmed a downregulation of CDK1 and CDK2 in SW1222 and HT29 DPPA3 cells, respectively, and a reduction of nuclear CCNA2 and CCNB1 levels in

(I) Relapse-free survival (RFS) curves of LARC patients according to the percentage of nuclear DPPA3-positive cells in naive biopsies assessed in (F). Cox proportional hazards model.

(J) Immunohistochemical staining quantification of DPPA3 in paired pre-nCRT biopsies and post-nCRT tissues of LARC patients.

(G–J) Patients were categorized according to the percentage of nuclear DPPA3-positive cells into DPPA3-High (>20%) and DPPA3-Low (≤20%) groups. (B–D, G, H, and J) *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001, two-way ANOVA (B, left, C, D), unpaired t test (B, right), Fisher's (G) and chi-square (H) exact tests, and paired t test (J). ADK, adenocarcinoma; CI, confidence interval; dil, dilution; HR, hazard ratio; LARC, locally advanced rectal cancer; nCRT, neoadjuvant chemoradiotherapy; w, week. See also Figure S1 and Table S2.



(legend on next page)

both DPPA3 cell lines (Figures 3J, 3K, S2L, and S2M). Consequently, SW1222 DPPA3 cells showed a delay in the G2/M phase of the cell cycle, whereas HT29 DPPA3 cells were delayed in the S-G2/M phases (Figures 3L and S2N). Finally, we concluded that the minor difference observed in cell-cycle phase delay between the two cell lines was probably due to the difference in CDK2 and CDK1 levels, which mainly regulate the S and G2/M phases, respectively (Figure 3M). Finally, unaltered H2AX phosphorylation (γ-H2AX) levels upon DPPA3 overexpression discarded a delay in G2/M as a result of a DNA-damage response induction (Figures S2O and S2P).

In summary, DPPA3-overexpressing cells maintain a *FOXM1* hypermethylation, leading to reduced expression of factors required for cell-cycle progression at late phases (Figure 3M).

DPPA3 xenografts recapitulate chemoresistance of CRC patient's tumors with high DPPA3 content

Next, we evaluated the ability of DPPA3 xenograft models to replicate the chemoresistance observed in human CRC tumors with high content of DPPA3-positive cells. First, we confirmed that subcutaneous DPPA3 xenografts continued to express the exogenous factor throughout tumor growth (Figure S3A). Subsequently, we conducted immunohistochemistry analysis to validate the categorization of DPPA3 xenografts as DPPA3-High and CTRL xenografts as DPPA3-Low. This was achieved by comparing their DPPA3-positive cell content with that of primary tumors from previously analyzed CRC patient cohorts (Figures 1, 2, and S3B). By using heterotopic and orthotopic mouse models, we observed that DPPA3 overexpression in SW1222 and HT29 xenografts significantly reduced tumor and lung metastases growth (Figures 4A–4D). DPPA3 overexpression increased the proportion of small versus large metastases without significantly affecting the metastatic capacity of the cells (Figures 4C, 4D, and S3C). Accordingly with *in vitro* results, the percentage of nuclear CCNB1-positive cells was reduced in SW1222 and HT29 DPPA3 xenograft models, explaining the tumor growth differences observed without significant changes in proliferation (Ki67) and apoptosis (caspase-3) marker levels (Figures 4E, 4F, and S3D–S3G).

Based on the findings indicating reduced DPPA3-positive cell content in post-nCRT rectal adenocarcinoma samples (Figure 2J), we opted to explore the impact of 5FU on DPPA3 levels. 5FU treatment increased both exogenous and endogenous DPPA3 protein levels in both cell lines despite no concurrent upregulation of mRNA levels (Figures 4G–4I and S3H). Using the maximum chemotherapy doses tolerated by NOD-SCID mice, we observed that whereas 5FU treatment delayed tumor xenograft growth of both CTRL cell models, DPPA3 tumors were resistant to treatment, consistent with data from CRC patients (Figure 4J). Again, histological analysis confirmed no significant differences in proliferation (Ki67) and apoptosis (cleaved-caspase-3) marker levels between CTRL and DPPA3 tumors in the vehicle condition (Figures S3I and S3J). Interestingly, although no significant differences in proliferation (Ki67) were observed between the chemotreated and vehicle CTRL and DPPA3 xenografts, a slight trend toward increased apoptosis was observed in partially sensitive CTRL tumors in response to 5FU. This trend did not reach statistical significance, probably because of the histological assessment at the endpoint of the experiment (Figures S3I and S3J). In contrast, we detected increased endogenous DPPA3 protein levels in chemotreated CTRL tumors (Figures 4K and 4L), confirming prior *in vitro* observations. Although these results apparently contradict data obtained from CRC patients, we must consider that DPPA3 analysis performed on human CRC samples was done on tissues whose treatment had already been withdrawn (Figure 2E). Therefore, these data suggest that, in response to 5FU, DPPA3 accumulates in cancer cells, leading to an increase of DPPA3-positive cells in xenografts that would be eliminated upon treatment withdraw.

Finally, we investigated whether DPPA3 was involved in protecting cells from chemotherapy-induced death as observed in SCCC. We examined the survival dynamics of SW1222 cells exposed to 5FU for 24 h and analyzed their apoptosis over a recovery period by assessing cleaved-caspase-3 levels (Figure 4M). Induction of cleaved-caspase-3 peaked at 72 h after 5FU treatment in CTRL cells, whereas DPPA3 cells showed a

Figure 3. DPPA3 sustains a slow-cycling phenotype by promoting a G2/M-phase delay

(A) Experimental design (top) and chord diagram (bottom) showing the significant biological functions negatively enriched in SW1222 cells overexpressing DPPA3 for 4 h (4 H DW) or 5 days (5 D DW), respectively, compared with CTRL cells. Biological functions were determined from GSEA results based on curated gene sets. Sample conditions are shown on the top and biological functions on the bottom of the chord diagram. Top-bottom connections indicate biological functions in each sample condition.

(B–D) GSEA plots showing enrichment of the indicated gene sets in DPPA3 versus CTRL cells (B and D, left) and SCCCs versus RCCCs (C and D, right).

(E) Line plot showing the methylation levels in the different probes (HumanMethylationEPIC chip) for the *FOXM1* gene in SW1222 DPPA3 or control (CTRL) cells. Significant differentially methylated position (DMP) (adjusted p value [adj.pval] < 0.05 and absolute difference in β values > 0.25) is indicated with a red line and box.

(F) qRT-PCR analysis showing the expression of *FOXM1* in indicated cells.

(G) Experimental design (top) and western blots (bottom) showing FOXM1 levels in indicated CRC cell lines.

(H) Experimental design (top) and qRT-PCR analysis (bottom) showing the expression of *FOXM1* in CRC SW1222-DPPA3 cells treated with vehicle (VH) or 5-azacytidine (AZA) at indicated concentrations.

(I) GSEA plots showing enrichment of the WIERSTRA_FOXM1_TARGETS gene set in DPPA3 versus CTRL cells (left) and SCCCs versus RCCCs (right).

(J and K) Experimental design (top) and western blots (bottom) showing CDK1 (J), CDK2 (J), CCNA2 (K), and CCNB1 (K) levels in indicated CRC cell lines.

(L) Experimental design (top) and cell-cycle analysis by flow cytometry (bottom) of indicated CRC cell lines. Donut charts show percentage distribution of the cell-cycle phases. Dashed arrows indicate the delayed cell-cycle phase.

(M) Schematic diagram summarizing the molecular mechanism by which DPPA3 overexpression delays the cell cycle in S-G2/M phases promoting a slow-cycling phenotype.

(F, H, and L) Mean ± SD of triplicates. (F and H) *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001, unpaired t test (F) and one-way ANOVA (H). (G, J, and K) Numbers show quantification by densitometry, which were normalized by their corresponding loading control. FC, fold change; Me, CpG methylation; r.u., relative units. See also Figure S2 and Tables S3 and S4.

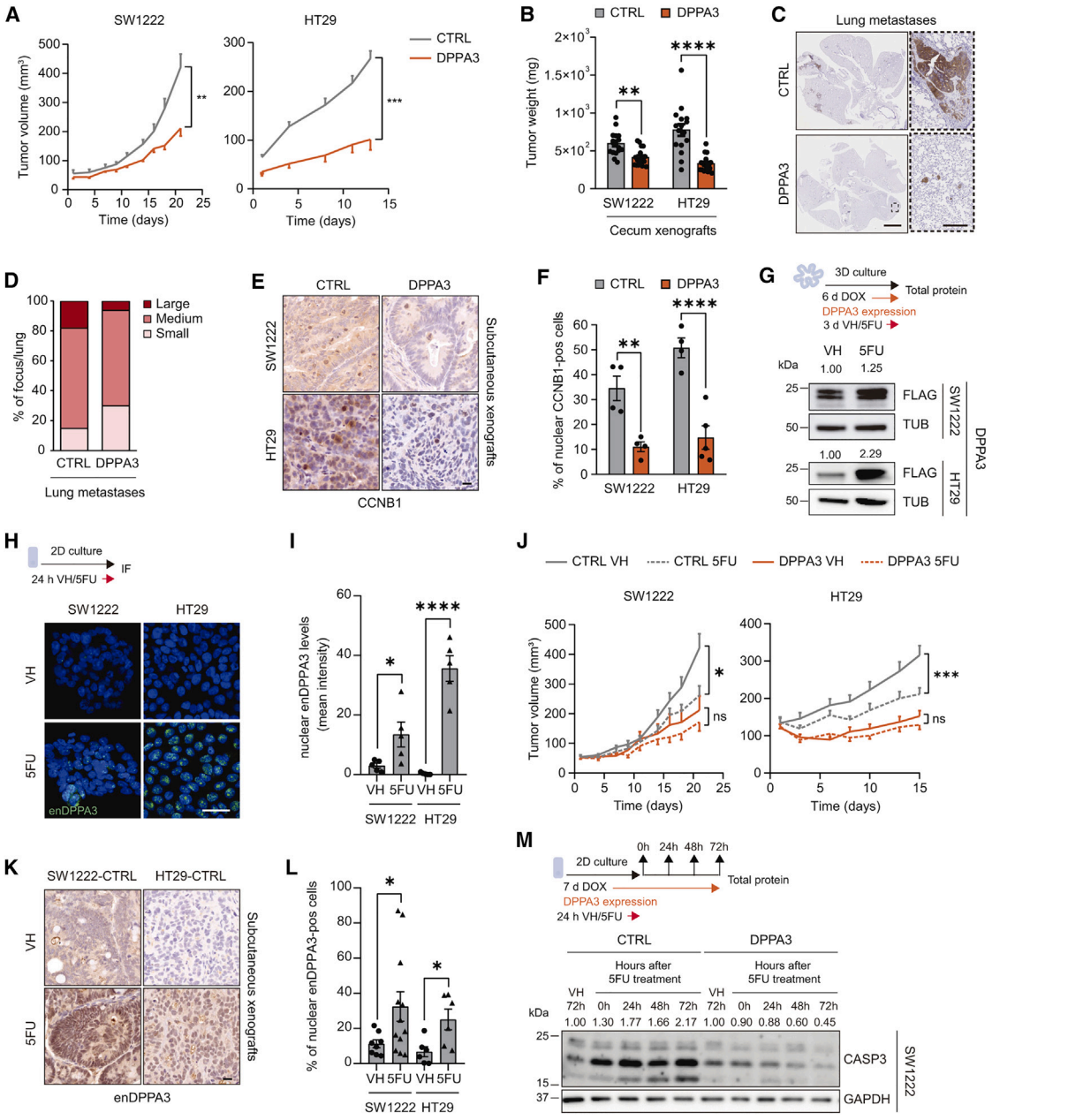


Figure 4. CRC xenografts overexpressing DPPA3 are chemoresistant

(A) Tumor growth curves of the indicated subcutaneous xenografts (n = 4–6 xenografts per condition).

(B) Tumor weight of orthotopic xenografts injected in cecum mice (n = 14–17 tumors per condition).

(C and D) Representative pictures (C) and metastases quantification evaluated by CK20 immunohistochemical staining (D) of lungs from mice orthotopically injected with HT29 CTRL and DPPA3 cells. Bar graphs show percentage of large (>30 cells), medium (5–30 cells), and small (≤4 cells) metastatic CK20-positive foci per lung. Scale bar, 2.5 mm; high-magnification scale bar, 250 μm.

(E and F) Representative pictures (E) and immunohistochemical staining quantification (F) of CCNB1 in indicated subcutaneous xenografts from (A). Scale bar, 20 μm.

(G) Experimental design (top) and western blots (bottom) showing exogen DPPA3 levels (FLAG) in indicated CRC cell lines.

(H and I) Representative immunofluorescence pictures (H) and quantification (I) of nuclear endogenous DPPA3 (enDPPA3) levels in indicated cell lines. Scale bar, 20 μm.

(J) Tumor growth curves of the indicated subcutaneous xenografts from vehicle- (VH) or 5FU-treated mice (n = 8–12 xenografts per condition).

(K and L) Representative pictures (K) and immunohistochemical staining quantification (L) of nuclear endogenous DPPA3 in indicated subcutaneous xenografts from (J). Scale bar, 20 μm.

(legend continued on next page)

decrease during the recovery period. In contrast, SW1222 sh*DPPA3* showed an opposite effect (Figure S3K). However, the impact of 5FU on SW1222 sh*DPPA3* cell viability did not translate into increased chemotherapy efficacy *in vivo* (Figure S3L), probably because of the small proportion of DPPA3-positive cells in SW1222 xenografts (Figure S3B).

From these experiments, we concluded that xenografts overexpressing DPPA3 reproduce the chemoresistant phenotype observed in CRC patients carrying tumors with a high proportion of DPPA3-positive cells.

DPPA3 mediates a hypoxia response in human CRC tumors

Structural prediction suggests that DPPA3 lacks tertiary structure and is predominantly composed of intrinsically disordered regions.[26,27] As such, it lacks any pocket that could be targeted by small-molecule drugs. Therefore, given the "undruggability" of DPPA3, we investigated potential vulnerabilities of DPPA3-overexpressing cells. To this end, we turned again to the gene-expression profile of SW1222-DPPA3 and CTRL cells, but this time we focused the GSEA in positively regulated biological processes (Figure 5A). Processes related to RTK signaling, protein maturation, membrane trafficking, tumor malignancy, and lysosomal activity were positively enriched in the short-term and long-term DPPA3-induced conditions versus control cells (Figure 5A). In addition, a positive enrichment of the ubiquitin system, mitochondria, and metabolism-related processes was observed exclusively in the short-term DPPA3-induced condition, whereas hypoxia was positively enriched only in the long-term DPPA3-induced condition. Furthermore, the gene-expression profile of the long-term DPPA3-induced condition showed that genes upregulated in the Pan-Cancer SCCC signature[8] were positively enriched in DPPA3 cells, recapitulating the upregulated SCCC transcriptome as we observed with downregulated genes (Figure S4A).

Among the different biological processes upregulated by DPPA3, hypoxia was identified as a potential vulnerability of DPPA3 cells, because it has been linked to chemoresistance.[9] Consistent with the upregulation of a hypoxia-responsive gene program in SCCC,[8] immunofluorescence assays revealed that SCCC areas were enriched in the CA9 hypoxia response marker (Figures 5B and 5C). Interestingly, downregulation of *DPPA3* reduced CA9 levels in areas enriched in SCCC, indicating that DPPA3 is required to maintain the SCCC hypoxia response (Figures 5B and 5C). Furthermore, a custom DPPA3-dependent hypoxia gene set (DPPA3_HYPOXIA_UP) generated from a leading-edge analysis of hypoxia-related gene sets enriched in DPPA3 cells was positively enriched in SCCCs, suggesting a role of DPPA3 in mediating the SCCC hypoxia response (Figure 5D; Table S5).

To assess the clinical relevance of these findings, we explored the relationship between DPPA3, hypoxia, and chemoresistance in human CRC samples. We observed a positive correlation between the percentages of CA9- and DPPA3-positive cells in equivalent-sized CRC-patient-derived xenografts (PDXs) (Figures 5E and 5F), suggesting that hypoxia might be an intrinsic feature of CRC tumors with high content of DPPA3-positive cells. We next generated a distinctive robust gene expression signature of DPPA3 (DPPA3 Sig) by an integrative analysis of gene-expression profiles obtained from our SW1222 gain- and loss-of-function cell models (Figure S4B and S4C; Table S6). The DPPA3 signature was significantly enriched in tumors from CRC patients subtyped as mesenchymal (CMS4), which has been previously associated with hypoxia and poor prognosis[28,29] (Figure 5G). Consequently, DPPA3-signature enrichment in naive tumors from this CRC cohort predicted significantly shorter PFS in the subset of patients who underwent 5FU-based adjuvant chemotherapy but was not predictive in the subset of non-chemotreated patients (Figures 5H and S4D). Notably, the predictive value of the DPPA3 signature was independent of other patient factors, including age, sex, tumor stage, and location (Figure S4E). Furthermore, the DPPA3_HYPOXIA_UP gene set was positively enriched in DPPA3-signature-enriched tumors among chemotreated patients (Figure 5I). These data are evidence that enrichment in the DPPA3 signature co-occurs with the upregulation of a hypoxia response in naive CRC chemoresistant tumors, conditioning disease progression.

Prompted by these findings, we evaluated the role of DPPA3 in the hypoxia response regulation in a chemotherapeutic setting. To this end, DPPA3 gain- and loss-of-function cell models were used, and HT29-sh*DPPA3* and shCTRL cell lines were included (Figure S4F). We observed that, compared with control cells, DPPA3 overexpression enhanced hypoxia response in SW1222 and HT29 cell lines in vehicle and 5FU conditions (Figure 5J). In contrast, in vehicle condition, hypoxia response was impaired in both SW1222 DPPA3 loss-of-function models, but not in HT29 sh*DPPA3* cells, probably because of the very low amount of endogenous DPPA3 in this experimental setting (Figure 5J). Interestingly, the hypoxia response was clearly impaired in all DPPA3 loss-of-function cell models upon chemotherapy (Figure 5J). Altogether, we could conclude that DPPA3 induction by chemotherapy (Figures 4 and S3) is required to a proper induction of the hypoxia response upon treatment in CRC (Figure 5).

Immunohistochemical analysis of post-nCRT rectal adenocarcinomas (LARC patients) revealed a higher percentage of CA9-positive cells in post-treatment samples that presented a high proportion of DPPA3-positive cells in naive biopsies (pre-nCRT) (Figure 5K). This result also suggests that DPPA3 mediates a sustained hypoxia response in tumors, which results in a post-hypoxia memory following chemotherapy that may contribute to disease progression in cancer patients.

(M) Experimental design (top) and western blots (bottom) showing cleaved-caspase-3 (CASP3) levels in indicated cells after 24 h of 5FU treatment and different recovery time points.

(A, B, F, I, J, and L) Mean ± SEM. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$ unpaired t test for each cell line (B, F, I, and L), unpaired t test at experiment endpoint (A), and two-way ANOVA at experiment endpoint (J). (G and M) Numbers show quantification by densitometry, which were normalized by their corresponding loading control. See also Figure S3.

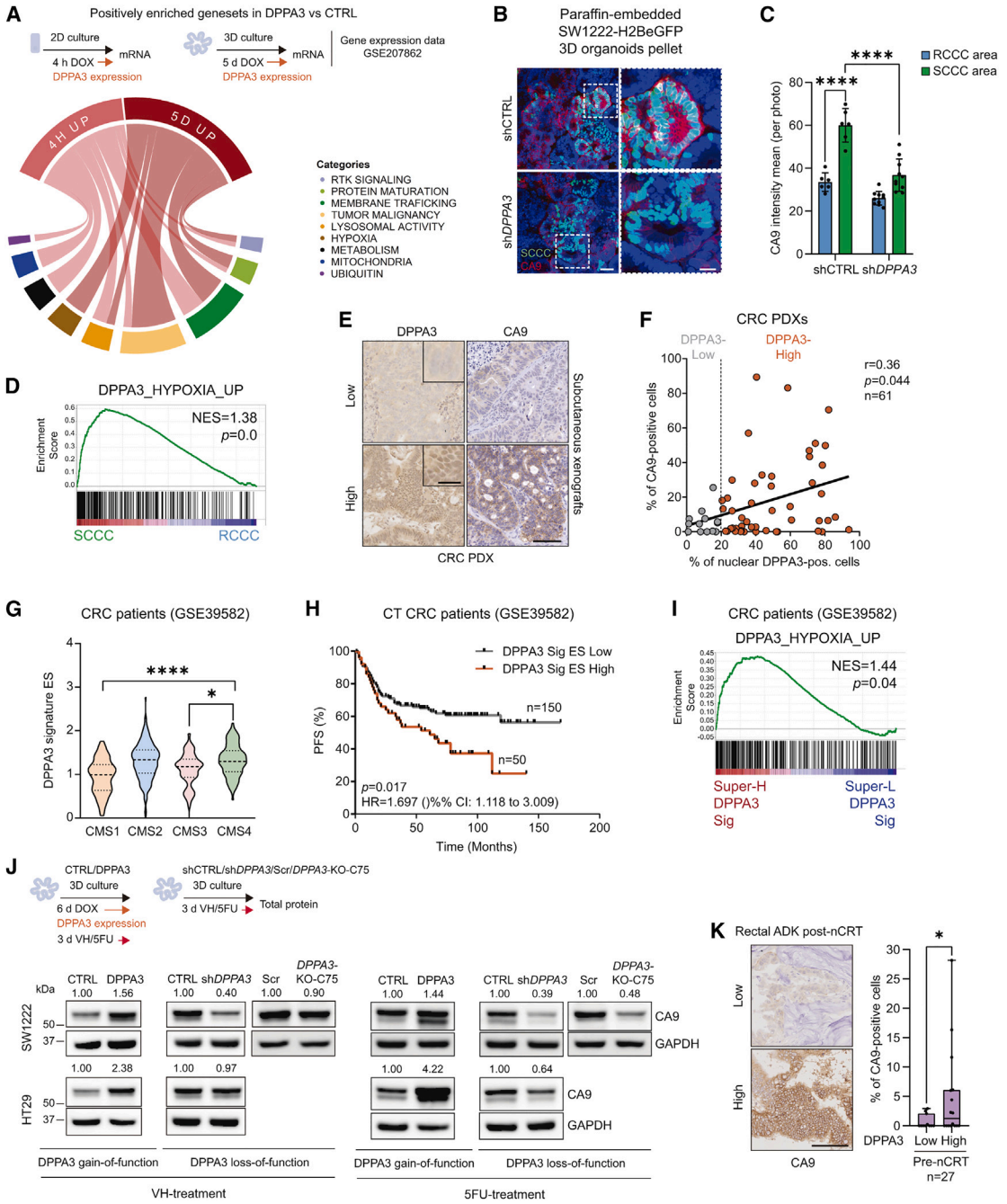


Figure 5. DPPA3 overexpression and hypoxia response program are concomitant in chemoresistant CRC tumors

(A) Experimental design (top) and chord diagram (bottom) showing the significant biological functions positively enriched in SW1222 cells overexpressing DPPA3 for 4 h (4 H DW) or 5 days (5 D DW), respectively, compared with CTRL cells. Biological functions were determined from GSEA results based on curated gene sets. Sample conditions are shown on the top and biological functions on the bottom of the chord diagram. Top-bottom connections indicate biological functions in each sample condition.

(B and C) Representative pictures (B) and quantification (C) of double-GFP (SCCC) and CA9 immunofluorescence staining of paraffin-embedded colonies and megacolonies pellets grown from indicated cell lines. Experimental design is shown in Figure 1A, diluting H2GeGFP for 12 days. Hoechst was used as counterstain. Scale bar, 50 μm; high-magnification scale bar, 20 μm.

(D) GSEA plot showing enrichment of our hypoxia-customized gene set (DPPA3_HYPOXIA_UP) in SCCCs versus RCCCs.

(E and F) Representative pictures (E) and immunohistochemical staining quantification (F) of DPPA3 and CA9 in CRC-patient-derived xenografts (PDXs). Scatterplot shows the correlation between the percentages of CA9-positive versus DPPA3-positive cells in samples analyzed in (E). Scale bar, 100 μm; high-magnification scale bar, 20 μm.

(legend continued on next page)

HIF1α downregulation partially rescues the cell-cycle progression by reducing DPPA3 levels

We next sought to understand how DPPA3 regulates the hypoxia-response pathway. No changes in *HIF1A* expression levels were detected in SW1222 DPPA3 cells (Figure S5A). Instead, we observed a drastic reduction of HIF1α ubiquitinated forms, without major changes in the HIF1α hydroxylation status (Figures 6A and S5B). Furthermore, SW1222 DPPA3 loss-of-function cell models showed decreased HIF1α levels and increased ubiquitinated forms, confirming requirement of DPPA3 for the proper regulation of HIF1α post-translational modifications and protein degradation (Figure S5C). In contrast, no changes in HIF1α levels or its ubiquitinated forms were observed in HT29 sh*DPPA3* cells, again probably because of their low levels of endogenous DPPA3 expression (Figure S5C). Importantly, DPPA3 overexpression promoted HIF1α protein accumulation even under normoxia (Figure 6B), confirming that DPPA3 controls HIF1α protein stabilization.

Next, we evaluated the impact of DPPA3 on the expression of HIF1α/hypoxia-regulated target genes. DPPA3 overexpression significantly increased *CA9* promoter activity in hypoxia (Figure S5D). In contrast, no promoter activity was observed when the hypoxia-response element (HRE) sequence was mutated, supporting that DPPA3 mediates the expression of genes containing HIF1α binding sites (HREs) (Figure S5D). Accordingly, a volcano plot analysis of genes with promoters containing HIF1α-binding motifs revealed specific canonical HIF1α target genes upregulated in SW1222 DPPA3 cells, including *LDHA*, *SLC2A1*, *CA9*, and *VEGFA* (Figure 6C; Table S7). Indeed, we confirmed by qRT-PCR that these four canonical HIF1α-target genes were upregulated by DPPA3 overexpression in both SW1222 3D cultures and tumor xenografts (Figures 6D and S5E). Furthermore, we observed that DPPA3 was required for a full increase of *SLC2A1*, *CA9* and *VEGFA* expression levels (Figure 6D), indicating its direct role in regulating the hypoxia-responsive gene program. These data reveal that DPPA3 contributes to the efficient response of CRC tumor cells to hypoxia by regulating HIF1α stabilization and function.

Because our findings evidenced that DPPA3 mediates the hypoxia response in SCCCs (Figure 5), we sought to study the impact of hypoxia on the regulation of DPPA3 expression. Interestingly, knockdown *HIF1A* in the SW1222 DPPA3 cell line decreased exogenous DPPA3 protein levels in both normoxia and hypoxia (Figures 6E and S5F), confirming that HIF1α plays a role in modulating DPPA3 levels. The analysis of *DPPA3* promoter activity, which contains a putative HRE, showed no difference between normoxia and hypoxia, discarding its transcriptional regulation by hypoxia (Figure S5G). In contrast, we observed an increased stabilization of the exogenous *DPPA3* mRNA levels upon hypoxia, indicating a post-transcriptional regulation (Figure S5H). Accordingly, exogenous DPPA3 mRNA and protein levels increased (Figures S5I and S5J). Importantly, downregulation of *HIF1A* in SW1222 DPPA3 cells led to a partial rescue of chromatin FOXM1 levels, which in turn caused a partial recovery of CDK1, CCNA2, and CCNB1 protein levels (Figures 6F–6H). Consistent with these results, a partial rescue of cell-cycle progression from G2/M to the G1 phase was observed when *HIF1A* was knocked down in SW1222 DPPA3 cells (Figures 6I and S5K).

In summary, our results uncover a DPPA3-HIF1α feedback loop regulation that mediates the slow-cycling phenotype in CRC cells and highlight HIF1α as a potential vulnerability of chemoresistant DPPA3-overexpressing cancer cells.

HIF1α depletion sensitizes DPPA3-overexpressing cancer cells to chemotherapy

To validate HIF1α as a potential therapeutic target to overcome chemoresistance linked to the slow-cycling phenotype, we first interrogated SCCCs. Pulse-chase experiments demonstrated that *HIF1A* reduction sensitized SCCCs to chemotherapy (Figures 7A and S6A). Next, we moved on to *in vivo* experiments. We observed an increase in SW1222 DPPA3 tumor growth when *HIF1A* was knocked down in vehicle condition (Figure 7B).

Importantly, these faster-growing tumors (DPPA3-sh*HIF1A*) restored sensitivity to chemotherapy, whereas control xenografts (DPPA3-shCTRL) maintained resistance.

Accordingly, reduction of *HIF1A* increased the percentage of nuclear CCNB1-positive cells in DPPA3 xenografts compared with DPPA3 shCTRL ones (Figures 7C and 7D). Consequently, *HIF1A* reduction sensitized DPPA3 chemoresistant tumors to 5FU treatment by inducing cell death, although elimination of cancer cells was incomplete (Figures 7B, 7E, and 7F). Note that *HIF1A* reduction failed to fully revert the slow tumor proliferation phenotype observed in DPPA3 xenografts, because tumors did not reach the size of SW1222 CTRL (shCTRL and sh*HIF1A*) tumors, confirming our previous results of partially

(G) Violin plot showing the DPPA3-signature enrichment score (ES) of the GEO: GSE39582 CRC cohort (n = 451) according to its molecular subtypes. CMS1, MSI, immune; CMS2, canonical; CMS3, metabolic; CMS4, mesenchymal.

(H) Progression-free survival (PFS) percentage curves of chemotreated (CT) CRC patients (GEO: GSE39582) according to the DPPA3-signature ES. Samples with ES in the upper quartile were classified as "high." Cox proportional hazards model.

(I) GSEA plot showing the positive enrichment of DPPA3_HYPOXIA_UP gene set in patients with a Super-High DPPA3 (upper 20%, n = 40) versus a Super-Low DPPA3 (lower 20%, n = 40) signature ES from the GEO: GSE39582 CT CRC cohort.

(J) Experimental design (top) and western blots (bottom) showing CA9 levels in indicated cell lines treated with vehicle (VH) or 5FU. Numbers show quantification by densitometry, which were normalized by their corresponding loading control.

(K) Representative pictures (left) and immunohistochemical staining quantification (right) of CA9 in indicated LARC patient samples post-nCRT. Bar plots show the percentage of CA9-positive cells according to the low or high content of nuclear DPPA3-positive cells in naive biopsies (pre-nCRT) analyzed in Figure 2. Scale bar, 100 μm.

(C, G, and K) *p ≤ 0.05, ****p ≤ 0.0001, two-way ANOVA (C), one-way ANOVA (G), and unpaired t test (K). (C and K) Mean ± SEM. (F and K) Patients were categorized according to the percentage of nuclear DPPA3-positive cells into DPPA3-High (>20%) and DPPA3-Low (≤20%). n, number of samples; r, correlation coefficient. See also Figure S4 and Tables S5 and S6.

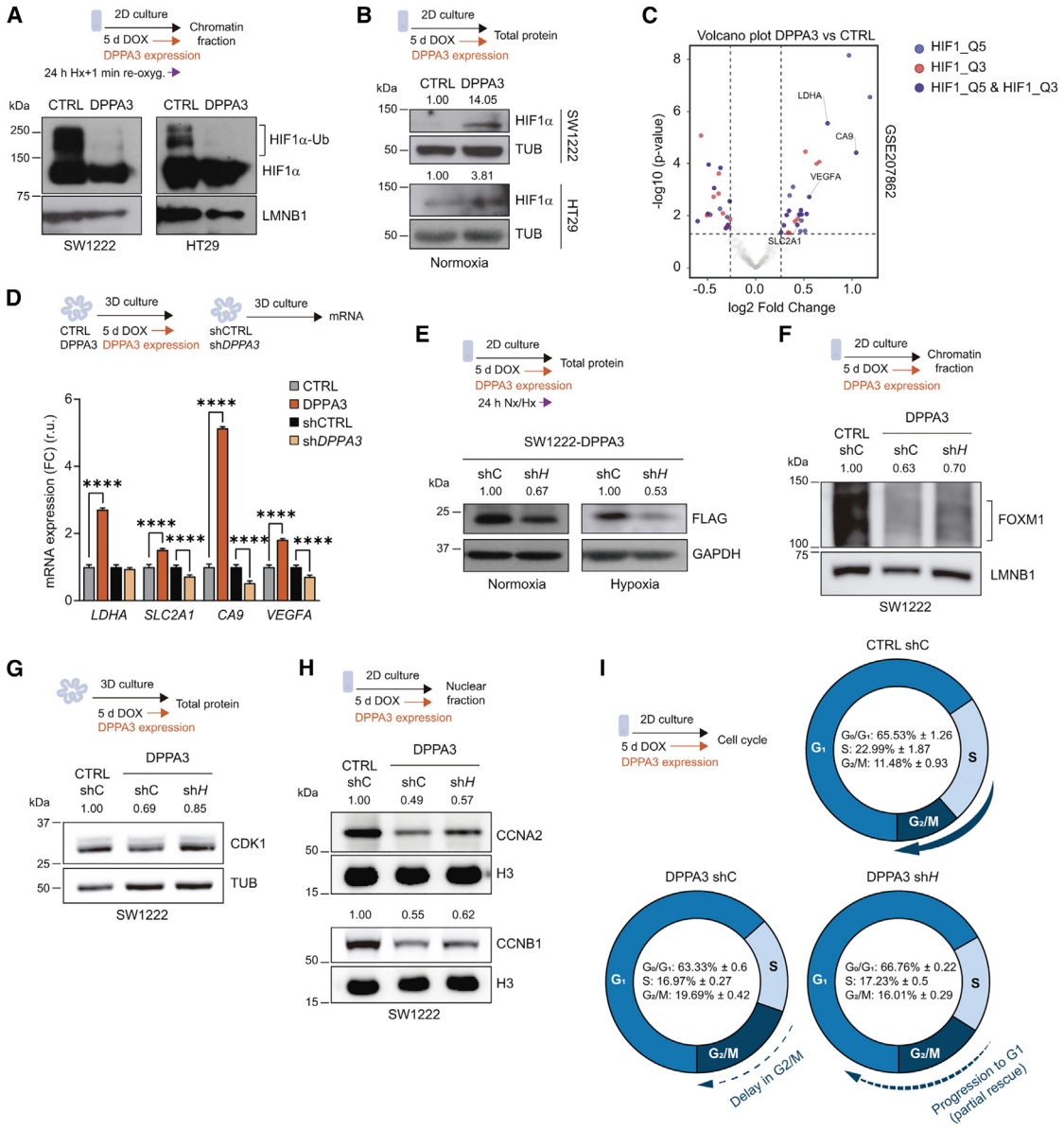


Figure 6. DPPA3 regulates HIF1α stabilization and is required for an efficient hypoxia response

(A) Experimental design (top) and western blots (bottom) showing HIF1α levels and its ubiquitinated forms (HIF1α-Ub) in indicated cell lines.

(B) Experimental design (top) and western blots (bottom) showing HIF1α total levels in indicated cell lines.

(C) Volcano plot showing HIF1 target genes differentially expressed in SW1222 DPPA3 versus control (CTRL) cells (x axis) against their significance (p value, y axis).

(D) Experimental design (top) and expression of the indicated genes (bottom) evaluated by qRT-PCR in indicated cell lines cultured in 3D. Mean ± SD of triplicates. ****$p \leq 0.0001$, one-way ANOVA for each gene evaluated.

(E–H) Experimental design (top) and western blots (bottom) showing indicated protein levels.

(I) Experimental design (left) and cell-cycle analysis by flow cytometry (right and bottom) of indicated CRC cell lines. Donut charts show percentage distribution of the cell-cycle phases. Dashed arrows indicate the delayed cell-cycle phase. Mean ± SD of triplicates.

(B and E–H) Numbers show quantification by densitometry, which were normalized by their corresponding loading control. See also Figure S5 and Table S7.

rescued phenotype when *HIF1A* was reduced in DPPA3 cells (Figures S6B, S6C, and S6D). Using a very restrictive threshold in immunohistochemical analysis, we observed that 5FU treatment increased DPPA3 protein levels in DPPA3 shCTRL xenografts but did not upon *HIF1A* reduction, supporting the positive feedback loop regulation between DPPA3 and HIF1α observed *in vitro* (Figures 7G and 7H). Consequently, such induction of DPPA3 upon 5FU treatment was accompanied by an accumulation of CA9-positive cells, again impaired by *HIF1A* reduction (Figures 7I and 7J). These results unveil

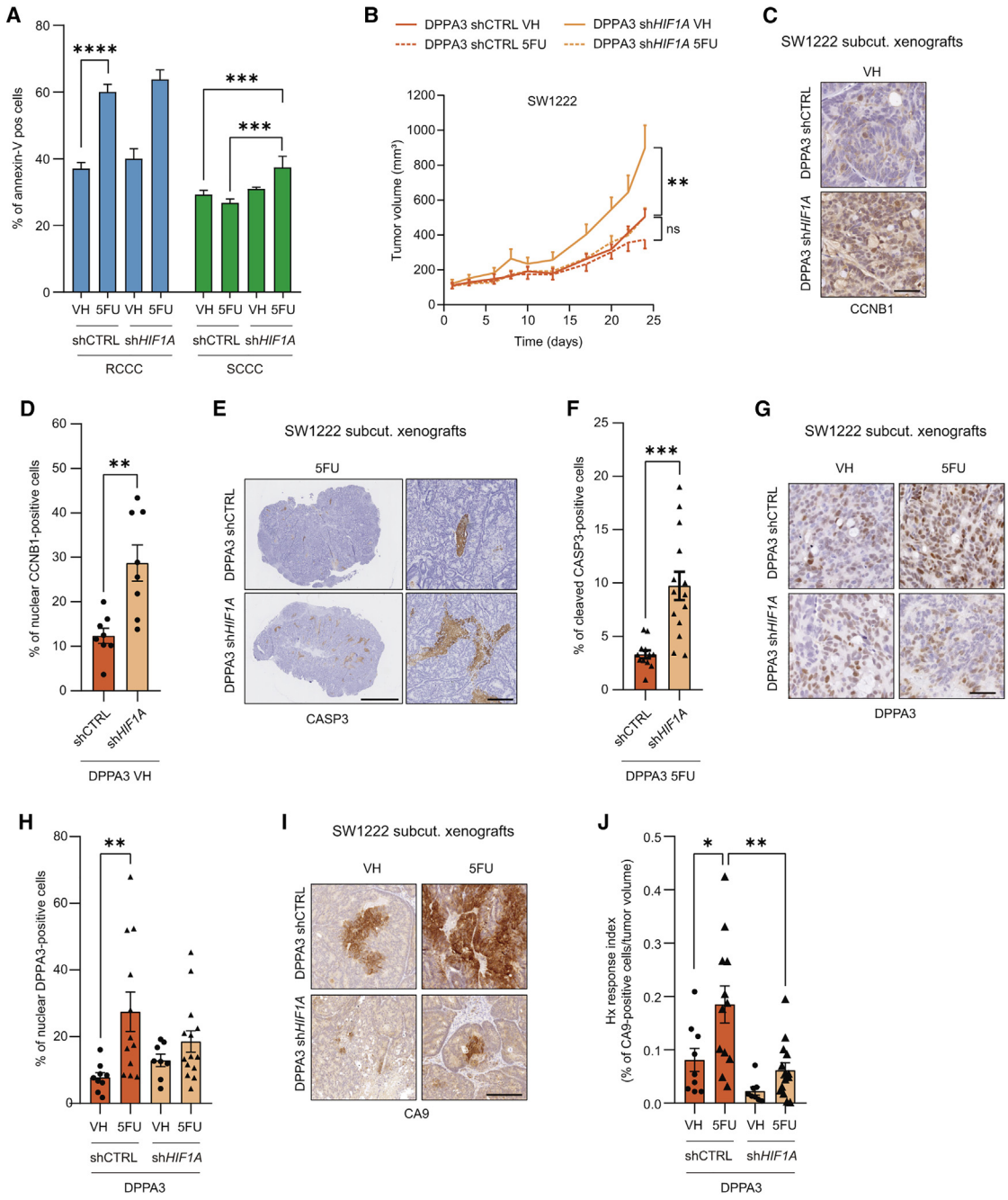


Figure 7. HIF1α depletion partially recovers a chemosensitive proliferative phenotype in DPPA3 cells

(A) Percentage of apoptotic *HIF1A* knockdown (sh*HIF1A*) and control (shCTRL) RCCCs and SCCCs upon 5FU exposure. Mean ± SD of triplicates.
(B) Tumor growth curves of the indicated subcutaneous xenografts from vehicle (VH) or 5FU-treated mice. Mean ± SEM (8–14 xenografts per condition).
(C–J) Representative pictures (C, E, G, and I) and immunohistochemical staining quantifications of CCNB1 (D), CASP3 (F), DPPA3 (H), and hypoxia response (J) in indicated subcutaneous xenografts from (B) (n = 8–14 xenografts per condition). Scale bars, 20 μm (C and G), 2.5 mm (E), and 250 μm (E, high magnification, and I).
(A, B, D, F, H, and J) *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001, two-way ANOVA (A, H, and J), two-way ANOVA at experiment endpoint (B), unpaired t test (D and F). (D, F, H, and J) Mean ± SEM. See also Figure S6.

HIF1α as a molecular vulnerability in DPPA3-overexpressing SCCCs, presenting a potential therapeutic opportunity to target HIF1α in CRC patients with DPPA3-mediated chemoresistant tumors.

DISCUSSION

This study provides a piece in the research puzzle concerning the underexplored area of non-genetic mechanisms of therapeutic

resistance in cancer. Conventional chemotherapies and radiotherapies target proliferating cells and require active cycling for apoptosis induction. The "slow-cycliness" in cancer cells is therefore an inherent mechanism for resistance and cell survival to anticancer therapies.[7,8,30,31] Here we demonstrated that DPPA3 is required to promote and sustain "slow-cycliness" in CRC cells, leading to a chemoresistant phenotype. We showed that chemoresistant SCCCs accumulate upon 5FU treatment, suggesting that these cells would exist prior to treatment. Supporting this notion, Ohta et al.[32] used tracked cellular lineages to conclude that a quiescent subset of CRC cells pre-exists as dormant entities responsible for tumor regrowth after chemotherapy ceases. However, we also observed that 5FU upregulates nuclear endogenous DPPA3 levels, suggesting that cancer cells might be prone to enter a slow-cycling state in response to therapy. Sharma et al.[2] observed that SCCCs were able to resume proliferation and restore drug sensitivity upon its withdrawal, suggesting a dynamically reversible and regulated phenotype of the "slow-cycliness." In line with this, our results showed that isolated SCCCs retain colony-forming capacity, indicating their reversible phenotypic plasticity. Therefore, given the complexity of tumor heterogeneity, it is highly probable that both phenomena, the pre-existing and the chemotherapy-induced slow-cycling phenotypes, are not mutually exclusive but rather concomitant.

Interestingly, *DPPA3* downregulation increased the ability of SCCCs to form new and bigger colonies, indicating that its reduction is required to resume cancer cell proliferation. This is crucial in cancer patients because it would directly impact disease relapse and progression. Following cancer cell dissemination, some tumor types display early metastasis with high mortality rates, whereas others need time to initiate and expand a lesion.[33] This latency period is explained when a slow-cycling or dormant phenotype is transposed to the metastatic cells.[33] In this context, our patient data showed that the content of DPPA3-positive cells is reduced in detectable metastases, supporting our results on self-renewal assays. Moreover, a high percentage of DPPA3-positive cells in metastases correlated with longer time to relapse, suggesting that DPPA3 must be turned off to facilitate metastatic outgrowth. More importantly, our results also revealed that patients with naive tumors carrying a high proportion of DPPA3-positive cells presented a reduced proportion of these cells in the residual disease after anticancer therapy, thus leading to a shorter time to relapse. In addition, we demonstrated that DPPA3 is required to maintain SCCC viability after chemotherapy exposure. Overall, our data support that DPPA3 protects cancer cells under anti-cancer treatment, but once the insult is terminated, it must be cleared from the cells to resume growth. The role of DPPA3 on cancer therapy resistance is also supported by the observations of Yan et al.,[19] who reported the chemoresistance of hepatocellular carcinoma cell lines overexpressing DPPA3 in *in vitro* assays. However, our study demonstrates the role of DPPA3 in regulating chemoresistance in cancer patient samples.

Slowed proliferation has been reported as one of the factors causing hypoxia-related chemoresistance in solid tumors.[34] Our results demonstrated that HIF1α represses cell proliferation by upregulating DPPA3 levels, which in turn downregulates *FOXM1* in a DNA-methylation-dependent manner. Therefore, DPPA3-overexpressing CRC cells may be prone to delay their proliferation in response to hypoxia, fostering SCCCs emergence. These results are supported by multiple evidences that have demonstrated a deep influence of hypoxia on cancer dormancy and "slow-cycliness."[20,35,36] Consistent with the activation of a DPPA3-mediated hypoxia-responsive gene program observed in SCCCs, our data demonstrated that DPPA3 is required to preserve the hypoxia response in SCCC. These results are in line with the data presented by Fluegen et al.[37] suggesting that hypoxic microenvironments within the primary tumor induce a dormant program that gives rise to a subpopulation of chemoresistant latent tumor cells. In this line, we demonstrated that a high DPPA3-positive cell content in naive tumors correlates with a post-treatment hypoxia-response memory in the residual disease of CRC patients with poor prognosis. Interestingly, downmodulation of *HIF1A* in SCCCs and DPPA3-overexpressing xenografts partially sensitized cancer cells to chemotherapy, highlighting HIF1α as a potential target to eliminate DPPA3-induced chemoresistant cells. However, there are some contradictory reports indicating that susceptibility of tumor cells to chemotherapy can be enhanced by HIF1α overexpression.[38] These contradictory observations suggest that dichotomous functions of HIF1α may depend on its distinctive interactions with other factors in different tumor cell types. Therefore, the DPPA3-HIF1α axis observed in CRC tumors might not happen in other tumor types.

Overall, our data indicate that cancer cells in hypoxic niches would enter in a slow-cycling phenotype more frequently in DPPA3-High CRC tumors. When anticancer therapies impact on hypoxia-response cancer cells, a vicious DPPA3-HIF1α-positive feedback loop would be engaged, promoting "slow-cycliness" and sustaining cancer persistence.

Limitations of the study

We clearly demonstrated that DPPA3 overexpression in cancer cells drives a chemoresistant slow-cycling phenotype associated with CRC poor prognosis. However, this current study has several limitations. First, the poor lung implantation efficiency of SW1222 cells upon tail-vein injection experiments, together with a very low proportion of SCCCs in xenografts, prevented the evaluation of their distinctive metastasis initiation capacity *in vivo*. Second, although we proved that DPPA3 is essential for SCCC chemoresistance, its requirement was not able to be demonstrated in xenograft models generated from our parental cell lines because of their low endogenous DPPA3 expression, so further studies using DPPA3-High PDX models will be warranted. Validation of HIF1α as a therapeutic target in DPPA3-overexpressing xenografts was not fully completed. When available, effective and safe HIF1α inhibitors should be validated in our DPPA3-High PDX models.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112927.

ACKNOWLEDGMENTS

We thank Dr. María Abad for providing the Ntera2 cell line and Dr. Stephan Tenbaum for providing the TetON-PGK-rtTA2-P2A-Cherry plasmid. We thank Dr. Xavier Barril for analyzing the DPPA3 druggability. We thank the Cytomics, Microscopy and Genomics Platforms from the High Technology Unit at Vall d'Hebron Research Institute (VHIR). We very much thank Amanda Wren for her assistance preparing the English manuscript, Javier Carmona for revising the entire paper, and Fundación CELLEX for its institutional support. The graphical abstract and schematic representations were created thanks to the BioRender platform. A.M.-E. received the support of a fellowship from "la Caixa" Foundation (100010434; code fellowship: LCF/BQ/DR20/11790017). H.G.P. was granted with a Miguel Servet contract (MSII14/00037). I.P. was supported by Fundación Científica Asociación Española Contra el Cáncer (FcAECC; code fellowships: INVES191PORT and INVES211004PUIG). The Vall d'Hebron Institute of Oncology (VHIO) would like to acknowledge the Cellex Foundation for providing research facilities and equipment, the Fero Foundation for their funding support, and the Centro de Investigación Biomédica en Red Cáncer (CIBERONC). The authors acknowledge financial support from Agencia Estatal de Investigación (CEX2020-001024-S/AEI/10.13039/501100011033), the Instituto de Salud Carlos III (grants PI08/1356, PI20/00895, PI20/00897, and PI20/00968), TRANSCAN-TACTIC (grant AC15/00064), the European Union's Horizon 2020 Research and Innovation program (EDIReX #731105), the EU/EFPIA Innovative Medicines Initiative (IMI) program (Horizon 2020-JTI-IMI2-2020-20-two-stage), "la Caixa" Foundation, and the Olga Torres Foundation (FOT).

AUTHOR CONTRIBUTIONS

E.C.-B., C.S., H.G.P., and I.P. conceived the study and coordinated experiments. E.C.-B. and C.S. designed and conducted the molecular, biochemical, and functional studies. E.C.-B. contributed to revising the manuscript. O.A. analyzed the microarray data and directed bioinformatics analyses. I.C., L.R., and L.C. provided support with *in vivo* experiments. J.M.-Q. and A.M.-E. generated the DPPA3 CRISPR-engineered SW1222 cell line. J.R.T. and M.F.F. contributed to methylome data and bioinformatics analyses. A.G.-A., J.H., O.M., E.E., J.T., P.N., and J.C. provided human samples. O.A., J.M.-Q., and A.M.-E. provided support with figures and manuscript writing. E.C.-B., C.S., and I.P. summarized and interpreted the data. H.G.P. and I.P. jointly supervised the study and wrote the manuscript.

DECLARATION OF INTERESTS

I.C., L.R., L.C., J.M.-Q., and H.G.P. report receiving commercial research grants from Blueprint medicines, Cyclacel, Grifols, Ikena Oncology, and Novartis. A.G.-A. declares speakers' bureau for Angelini and travel, accommodations, and expenses from Pfizer, Ipsen, and Eisai. J.H. reports receiving honoraria from speakers' bureaus from Adacap, Angelini, Bayer, Eisai, Ipsen, Leo Pharma, Novartis, and Roche. E.E. discloses personal financial interests receiving honoraria for advisory roles, travel grants, and research grants from Amgen, Bayer, Hoffman-La Roche, Merck Serono, Sanofi, Pierre Fabre, MSD, Organon, Novartis, and Servier. J.T. reports personal financial interests in the form of a scientific consultancy role for Array Biopharma, AstraZeneca, Bayer, Boehringer Ingelheim, Cardiff Oncology, Chugai, Daiichi Sankyo, F. Hoffmann-La Roche Ltd, Genentech Inc., HalioDX SAS, Hutchison MediPharma International, Ikena Oncology, Inspirna Inc., IQVIA, Lilly, Menarini, Merck Serono, Merus, MSD, Mirati, Neophore, Novartis, Ona Therapeutics, Orion Biotechnology, Peptomyc, Pfizer, Pierre Fabre, Samsung Bioepis, Sanofi, Scandion Oncology, Scorpion Therapeutics, Seattle Genetics, Servier, Sotio Biotech, Taiho, Tessa Therapeutics, TheraMyc, and Tolremo Therapeutics; stocks of Oniria Therapeutics; and also educational collaboration with Imedex/HMP, Medscape Education, MJH Life Sciences, PeerView Institute for Medical Education, and Physicians Education Resource (PER). P.N. discloses personal financial interests of receiving honoraria or consultation fees from Novartis, Bayer, and MSD Oncology and receiving travel and accommodations paid or reimbursed by Novartis. J.C. declares receiving honoraria for serving on speakers' bureaus and scientific advisory from Advanced Accelerator Applications (AAA), Amgen, Bayer, Eisai, Exelixis, Ipsen, Lilly, Merck Serono, Novartis, Pfizer, and Sanofi and receiving research grants from AAA, AstraZeneca, Bayer, Eisai, Novartis, and Pfizer. H.G.P. is a scientific co-founder, Chief Scientific Officer (CSO), executive board member, and equity owner of Oniria Therapeutics; he does not receive financial compensation as a CSO. I.P. is a scientific co-founder, scientific advisory board member, and equity owner of Oniria Therapeutics; she does not receive financial compensation as a consultant.

Received: December 29, 2022
Revised: June 22, 2023
Accepted: July 17, 2023
Published: August 1, 2023

REFERENCES

1. Brauner, A., Fridman, O., Gefen, O., and Balaban, N.Q. (2016). Distinguishing between resistance, tolerance and persistence to antibiotic treatment.

Nat. Rev. Microbiol. *14*, 320–330. https://doi.org/10.1038/nrmicro.2016.34.

2. Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A., et al. (2010). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell *141*, 69–80. https://doi.org/10.1016/j.cell.2010.02.027.
3. Shen, S., Vagner, S., and Robert, C. (2020). Persistent Cancer Cells: The Deadly Survivors. Cell *183*, 860–874. https://doi.org/10.1016/j.cell.2020.10.027.
4. Rehman, S.K., Haynes, J., Collignon, E., Brown, K.R., Wang, Y., Nixon, A.M.L., Bruce, J.P., Wintersinger, J.A., Singh Mer, A., Lo, E.B.L., et al. (2021). Colorectal Cancer Cells Enter a Diapause-like DTP State to Survive Chemotherapy. Cell *184*, 226–242.e21. https://doi.org/10.1016/j.cell.2020.11.018.
5. Dhimolea, E., de Matos Simoes, R., Kansara, D., Al'Khafaji, A., Bouyssou, J., Weng, X., Sharma, S., Raja, J., Awate, P., Shirasaki, R., et al. (2021). An Embryonic Diapause-like Adaptation with Suppressed Myc Activity Enables Tumor Treatment Persistence. Cancer Cell *39*, 240–256.e11. https://doi.org/10.1016/j.ccell.2020.12.002.
6. Hong, S.P., Chan, T.E., Lombardo, Y., Corleone, G., Rotmensz, N., Bravaccini, S., Rocca, A., Pruneri, G., McEwen, K.R., Coombes, R.C., et al. (2019). Single-cell transcriptomics reveals multi-step adaptations to endocrine therapy. Nat. Commun. *10*, 3840. https://doi.org/10.1038/s41467-019-11721-9.
7. Roesch, A., Fukunaga-Kalabis, M., Schmidt, E.C., Zabierowski, S.E., Brafford, P.A., Vultur, A., Basu, D., Gimotty, P., Vogt, T., and Herlyn, M. (2010). A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. Cell *141*, 583–594. https://doi.org/10.1016/j.cell.2010.04.020.
8. Puig, I., Tenbaum, S.P., Chicote, I., Arqués, O., Martínez-Quintanilla, J., Cuesta-Borrás, E., Ramírez, L., Gonzalo, P., Soto, A., Aguilar, S., et al. (2018). TET2 controls chemoresistant slow-cycling cancer cell survival and tumor recurrence. J. Clin. Invest. *128*, 3887–3905. https://doi.org/10.1172/JCI96393.
9. Qiu, G.Z., Jin, M.Z., Dai, J.X., Sun, W., Feng, J.H., and Jin, W.L. (2017). Reprogramming of the Tumor in the Hypoxic Niche: The Emerging Concept and Associated Therapeutic Strategies. Trends Pharmacol. Sci. *38*, 669–686. https://doi.org/10.1016/j.tips.2017.05.002.
10. Fukuyama, M., Rougvie, A.E., and Rothman, J.H. (2006). C. elegans DAF-18/PTEN mediates nutrient-dependent arrest of cell cycle and growth in the germline. Curr. Biol. *16*, 773–779. https://doi.org/10.1016/j.cub.2006.02.073.
11. Saitou, M., Kagiwada, S., and Kurimoto, K. (2012). Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. Development *139*, 15–31. https://doi.org/10.1242/dev.050849.
12. Tumbar, T., Guasch, G., Greco, V., Blanpain, C., Lowry, W.E., Rendl, M., and Fuchs, E. (2004). Defining the epithelial stem cell niche in skin. Science *303*, 359–363. https://doi.org/10.1126/science.1092436.
13. Harris, R.S., Petersen-Mahrt, S.K., and Neuberger, M.S. (2002). RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. Mol. Cell. *10*, 1247–1253. https://doi.org/10.1016/s1097-2765(02)00742-6.
14. Nakashima, H., Kimura, T., Kaga, Y., Nakatani, T., Seki, Y., Nakamura, T., and Nakano, T. (2013). Effects of dppa3 on DNA methylation dynamics during primordial germ cell development in mice. Biol. Reprod. *88*, 125. https://doi.org/10.1095/biolreprod.112.105932.
15. Li, Y., Zhang, Z., Chen, J., Liu, W., Lai, W., Liu, B., Li, X., Liu, L., Xu, S., Dong, Q., et al. (2018). Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. Nature *564*, 136–140. https://doi.org/10.1038/s41586-018-0751-5.
16. Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M., et al. (2007). PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat. Cell Biol. *9*, 64–71. https://doi.org/10.1038/ncb1519.
17. Nakamura, T., Liu, Y.J., Nakashima, H., Umehara, H., Inoue, K., Matoba, S., Tachibana, M., Ogura, A., Shinkai, Y., and Nakano, T. (2012). PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. Nature *486*, 415–419. https://doi.org/10.1038/nature11093.
18. Funaki, S., Nakamura, T., Nakatani, T., Umehara, H., Nakashima, H., and Nakano, T. (2014). Inhibition of maintenance DNA methylation by Stella. Biochem. Biophys. Res. Commun. *453*, 455–460. https://doi.org/10.1016/j.bbrc.2014.09.101.
19. Yan, Q., Zhang, Y., Fang, X., Liu, B., Wong, T.L., Gong, L., Liu, S., Yu, D., Liu, M., Jiang, L., et al. (2021). PGC7 promotes tumor oncogenic dedifferentiation through remodeling DNA methylation pattern for key developmental transcription factors. Cell Death Differ. *28*, 1955–1970. https://doi.org/10.1038/s41418-020-00726-3.
20. Aguirre-Ghiso, J.A. (2007). Models, mechanisms and clinical evidence for cancer dormancy. Nat. Rev. Cancer *7*, 834–846. https://doi.org/10.1038/nrc2256.
21. Burrell, R.A., McGranahan, N., Bartek, J., and Swanton, C. (2013). The causes and consequences of genetic heterogeneity in cancer evolution. Nature *501*, 338–345. https://doi.org/10.1038/nature12625.
22. Talukdar, S., Bhoopathi, P., Emdad, L., Das, S., Sarkar, D., and Fisher, P.B. (2019). Dormancy and cancer stem cells: An enigma for cancer therapeutic targeting. Adv. Cancer Res. *141*, 43–84. https://doi.org/10.1016/bs.acr.2018.12.002.
23. Serna, G., Ruiz-Pace, F., Hernando, J., Alonso, L., Fasani, R., Landolfi, S., Comas, R., Jimenez, J., Elez, E., Bullman, S., et al. (2020). Fusobacterium nucleatum persistence and risk of recurrence after preoperative treatment in locally advanced rectal cancer. Ann. Oncol. *31*, 1366–1375. https://doi.org/10.1016/j.annonc.2020.06.003.
24. Fischer, M., Schade, A.E., Branigan, T.B., Müller, G.A., and DeCaprio, J.A. (2022). Coordinating gene expression during the cell cycle. Trends Biochem. Sci. *47*, 1009–1022. https://doi.org/10.1016/j.tibs.2022.06.007.
25. Wierstra, I. (2013). The transcription factor FOXM1 (Forkhead box M1): proliferation-specific expression, transcription factor function, target genes, mouse models, and normal biological roles. Adv. Cancer Res. *118*, 97–398. https://doi.org/10.1016/B978-0-12-407173-5.00004-2.
26. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.
27. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. *50*, D439–D444. https://doi.org/10.1093/nar/gkab1061.
28. Zhong, M.E., Huang, Z.P., Wang, X., Cai, D., Li, C.H., Gao, F., Wu, X.J., and Wang, W. (2022). A Transcription Factor Signature Can Identify the CMS4 Subtype and Stratify the Prognostic Risk of Colorectal Cancer. Front. Oncol. *12*, 902974. https://doi.org/10.3389/fonc.2022.902974.
29. Guinney, J., Dienstmann, R., Wang, X., de Reyniès, A., Schlicker, A., Soneson, C., Marisa, L., Roepman, P., Nyamundanda, G., Angelino, P., et al. (2015). The consensus molecular subtypes of colorectal cancer. Nat. Med. *21*, 1350–1356. https://doi.org/10.1038/nm.3967.
30. Dembinski, J.L., and Krauss, S. (2009). Characterization and functional analysis of a slow cycling stem cell-like subpopulation in pancreas adenocarcinoma. Clin. Exp. Metastasis *26*, 611–623. https://doi.org/10.1007/s10585-009-9260-0.
31. Fillmore, C.M., and Kuperwasser, C. (2008). Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. Breast Cancer Res. *10*, R25. https://doi.org/10.1186/bcr1982.

32. Ohta, Y., Fujii, M., Takahashi, S., Takano, A., Nanki, K., Matano, M., Hanyu, H., Saito, M., Shimokawa, M., Nishikori, S., et al. (2022). Cell-matrix interface regulates dormancy in human colon cancer stem cells. Nature *608*, 784–794. https://doi.org/10.1038/s41586-022-05043-y.

33. Massagué, J., and Obenauf, A.C. (2016). Metastatic colonization by circulating tumour cells. Nature *529*, 298–306. https://doi.org/10.1038/nature17038.

34. Brown, J.M. (2000). Exploiting the hypoxic cancer cell: mechanisms and therapeutic strategies. Mol. Med. Today *6*, 157–162. https://doi.org/10.1016/s1357-4310(00)01677-4.

35. Carcereri de Prati, A., Butturini, E., Rigo, A., Oppici, E., Rossin, M., Boriero, D., and Mariotto, S. (2017). Metastatic Breast Cancer Cells Enter Into Dormant State and Express Cancer Stem Cells Phenotype Under Chronic Hypoxia. J. Cell. Biochem. *118*, 3237–3248. https://doi.org/10.1002/jcb.25972.

36. Gatenby, R.A., Coia, L.R., Richter, M.P., Katz, H., Moldofsky, P.J., Engstrom, P., Brown, D.Q., Brookland, R., and Broder, G.J. (1985). Oxygen tension in human tumors: in vivo mapping using CT-guided probes. Radiology *156*, 211–214. https://doi.org/10.1148/radiology.156.1.4001408.

37. Fluegen, G., Avivar-Valderas, A., Wang, Y., Padgen, M.R., Williams, J.K., Nobre, A.R., Calvo, V., Cheung, J.F., Bravo-Cordero, J.J., Entenberg, D., et al. (2017). Phenotypic heterogeneity of disseminated tumour cells is preset by primary tumour hypoxic microenvironments. Nat. Cell Biol. *19*, 120–132. https://doi.org/10.1038/ncb3465.

38. Rohwer, N., and Cramer, T. (2011). Hypoxia-mediated drug resistance: novel insights on the functional interaction of HIFs and cell death pathways. Drug Resist. Updates *14*, 191–201. https://doi.org/10.1016/j.drup.2011.03.001.

39. Marisa, L., de Reyniès, A., Duval, A., Selves, J., Gaub, M.P., Vescovo, L., Etienne-Grimaldi, M.C., Schiappa, R., Guenot, D., Ayadi, M., et al. (2013). Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. PLoS Med. *10*, e1001453. https://doi.org/10.1371/journal.pmed.1001453.

40. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308. https://doi.org/10.1038/nprot.2013.143.

41. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

42. Grabmaier, K., A de Weijert, M.C., Verhaegh, G.W., Schalken, J.A., and Oosterwijk, E. (2004). Strict regulation of CAIX(G250/MN) by HIF-1alpha in clear cell renal cell carcinoma. Oncogene *23*, 5624–5631. https://doi.org/10.1038/sj.onc.1207764.

43. Giuliano, C.J., Lin, A., Girish, V., and Sheltzer, J.M. (2019). Generating Single Cell-Derived Knockout Clones in Mammalian Cells with CRISPR/Cas9. Curr. Protoc. Mol. Biol. *128*, e100. https://doi.org/10.1002/cpmb.100.

44. Puig, I., Chicote, I., Tenbaum, S.P., Arqués, O., Herance, J.R., Gispert, J.D., Jimenez, J., Landolfi, S., Caci, K., Allende, H., et al. (2013). A personalized preclinical model to evaluate the metastatic potential of patient-derived colon cancer initiating cells. Clin. Cancer Res. *19*, 6787–6801. https://doi.org/10.1158/1078-0432.CCR-12-1740.

45. Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. *34*, 267–273. https://doi.org/10.1038/ng1180.

46. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.

47. Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. Nat. Protoc. *3*, 1101–1108. https://doi.org/10.1038/nprot.2008.73.

48. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. *3*, RESEARCH0034. https://doi.org/10.1186/gb-2002-3-7-research0034.

49. Céspedes, M.V., Espina, C., García-Cabezas, M.A., Trias, M., Boluda, A., Gómez del Pulgar, M.T., Sancho, F.J., Nistal, M., Lacal, J.C., and Mangues, R. (2007). Orthotopic microinjection of human colon cancer cells in nude mice induces tumor foci in all clinically relevant metastatic sites. Am. J. Pathol. *170*, 1077–1085. https://doi.org/10.2353/ajpath.2007.060773.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-Carbonic Anhydrase IX/CA9	Novus Biologicals	Cat#NB100-417; RRID: AB_10003398
Rabbit monoclonal anti-CCNA2 [Y193]	Abcam	Cat#ab32386; RRID: AB_2244193
Mouse monoclonal anti-CCNB1 [V152]	Abcam	Cat#ab72; RRID: AB_305751
Rabbit monoclonal anti-CCNB1 [Y106]	Abcam	Cat#ab32053; RRID: AB_731779
Mouse monoclonal anti-CDK1 [A17]	Abcam	Cat#ab18; RRID: AB_2074906
Mouse monoclonal anti-CDK2 [D12]	Santa Cruz Biotechnology	Cat#sc-6248; RRID: AB_627238
Mouse monoclonal anti-CK20 [Ks20.8]	Agilent	Cat#M7019; RRID: AB_2133718
Rabbit polyclonal anti-Cleaved Caspase-3 (CASP3) (Asp175)	Cell Signaling Technology	Cat#9661; RRID: AB_2341188
Rabbit monoclonal anti-Cleaved Caspase-3 (CASP3) (Asp175) [5A1E]	Cell Signaling Technology	Cat#9664; RRID: AB_2070042
Rabbit polyclonal anti-DNMT1	EpiGentek	Cat#A1700
Mouse monoclonal anti-Stella (DPPA3) [3H5.2]	Millipore	Cat#MAB4388; RRID: AB_2094156
Mouse monoclonal anti-Stella (DPPA3) [D-5]	Santa Cruz Biotechnology	Cat#sc-376862
Rabbit polyclonal anti-FLAG	Sigma-Aldrich	Cat#F7425; RRID: AB_439687
Mouse monoclonal anti-FLAG [M2]	Sigma-Aldrich	Cat#F1804; RRID: AB_262044
Rabbit monoclonal anti-FOXM1 [EPR17379]	Abcam	Cat#ab207298
Mouse monoclonal anti-GAPDH [FF26A/F9]	BioLegend	Cat#649202; RRID: AB_10612752
Mouse monoclonal anti-GFP [B2]	Santa Cruz Biotechnology	Cat#sc-9996; RRID: AB_627695
Rabbit polyclonal anti-H3	Abcam	Cat#ab1791; RRID: AB_302613
Mouse monoclonal anti-HIF1a [54]	BD Biosciences	Cat#610958; RRID: AB_398272
Rabbit monoclonal anti-Hydroxy-HIF1a (Pro564) [D43B5]	Cell Signaling Technology	Cat#3434; RRID: AB_2116958
Mouse monoclonal anti-KI67 [MIB-1]	Agilent	Cat#M7240; RRID: AB_2142367
Rabbit polyclonal anti-Lamin B1	Abcam	Cat#ab16048; RRID: AB_443298
Rabbit polyclonal anti-Phospho-Histone H2A.X	Cell Signaling Technology	Cat#2577; RRID: AB_2118010
Mouse monoclonal anti-a-Tubulin (TUB) [DM1A]	Sigma-Aldrich	Cat#T9026; RRID: AB_477593
Goat polyclonal secondary anti-Mouse-HRP	Thermo Fisher Scientific	Cat#G-21040: RRID: AB_2536527
Donkey polyclonal secondary anti-Rabbit-HRP	Jackson ImmunoResearch Labs	Cat#711-035-152; RRID: AB_10015282
Goat secondary anti-Mouse SignalStain Boost IHC Detection Reagent (HRP)	Cell Signaling Technology	Cat#8125; RRID: AB_10547893
Goat polyclonal secondary anti-Mouse-Alexa Fluor 488	Thermo Fisher Scientific	Cat#A-11001; RRID: AB_2534069
Goat polyclonal secondary anti-Rabbit-Alexa Fluor 488	Thermo Fisher Scientific	Cat#A-11008; RRID: AB_143165

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Goat polyclonal secondary anti-Rabbit-Alexa Fluor 555	Thermo Fisher Scientific	Cat#A-21428; RRID: AB_2535849
Biological samples		
Patient-derived xenografts (PDX)	Vall d'Hebron Institute of Oncology (Puig et al.)[8]	N/A
Locally advanced rectal cancer (LARC)	Vall d'Hebron Institute of Oncology (Serna et al.)[23]	N/A
Paired primary CRC tumors and liver metastasis	Vall d'Hebron Institute of Oncology	N/A
Chemicals, peptides, and recombinant proteins		
Doxycycline hyclate (DOX)	Sigma-Aldrich	Cat#D9891; CAS: 24390-14-5
5-azacytidine (AZA)	MedChemExpress	Cat#HY-10586; CAS: 320-67-2
5-fluorouracil (5FU)	Sigma-Aldrich	Cat#F6627; CAS: 51-21-8
Corning® Matrigel® Basement Membrane Matrix	Cultek	Cat#45354277
Corning® Cell Recovery Solution	Cultek	Cat#45354253
Polyethylenimine, linear, MW 25000, transfection Grade (PEI)	Polysciences, Inc.	Cat#23966-1; CAS: 9002-98-6, 26913-06-4
Sucrose	Sigma-Aldrich	Cat#84100; CAS: 57-50-1
Penicillin-Streptomycin Gibco™ (P/S)	Thermo Fisher Scientific	Cat#15140122; CAS: 69-57-8/3810-74-0
Amphotericin B Gibco™ (Fungizone)	Thermo Fisher Scientific	Cat#15290026; CAS: 1397-89-3
Kanamycin Sulfate Gibco™	Thermo Fisher Scientific	Cat#11578876; CAS: 25389-94-0
Gentamicin Gibco™	Thermo Fisher Scientific	Cat#11520506; CAS: 1403-66-3
Nystatin	Sigma-Aldrich	Cat#N4014; CAS: 1400-61-9
Collagenase from *Clostridium histolyticum*	Sigma-Aldrich	Cat#C0130; CAS: 9001-12-1
Deoxyribonuclease I from bovine pancreas	Sigma-Aldrich	Cat#D4263; CAS: 9003-98-9
Protease inhibitor cocktail Roche	Sigma-Aldrich	Cat#11836170001
SDS	Sigma-Aldrich	Cat#436143; CAS: 151-21-3
EDTA	Merck Millipore	Cat#324503; CAS: 6381-92-6
HEPES	Sigma-Aldrich	Cat#H0887; CAS: 7365-45-9
KCl	Sigma-Aldrich	Cat#P3911; CAS: 7447-40-7
$MgCl_2$	Sigma-Aldrich	Cat#M8266; CAS: 7786-30-3
DTT	PanReac AppliChem	Cat#A1101; CAS: 3483-12-3
Triton X-100	Sigma-Aldrich	Cat#T9284; CAS: 9036-19-5
Glycerol	Sigma-Aldrich	Cat#G7893; CAS: 56-81-5
NaCl	Sigma-Aldrich	Cat#S9888; CAS: 7647-14-5
Tween 20	Merck Millipore	Cat#817072; CAS: 9005-64-5
Bovine Serum Albumin (BSA)	Sigma-Aldrich	Cat#A3059; CAS: 9048-46-8
4% Paraformaldehyde (PFA)	Santa Cruz Biotechnology	Cat#sc-281692; CAS: 30525-89-4
Hoechst 33342	Sigma-Aldrich	Cat#B2261; CAS: 875756-97-1
Tris-base	Sigma-Aldrich	Cat#T1503; CAS: 77-86-1
Sodium citrate	Merck Millipore	Cat#106448; CAS: 6132-04-3
Ribonuclease A	PanReac AppliChem	Cat#A3832; CAS: 9001-99-4
Propidium iodide	Sigma-Aldrich	Cat#P4170; CAS: 25535-16-4
Trizol Reagent	Thermo Fisher Scientific	Cat#15596018
Glutamine Gibco™	Thermo Fisher Scientific	Cat#11539876; CAS: 56-85-9
Y-27632	Sigma-Aldrich	Cat#Y0503; CAS: 129830-38-2
Puromycin dihydrochloride	Thermo Fisher Scientific	Cat#A111803; CAS: 58-58-2
TrypLE Express Enzyme	Thermo Fisher Scientific	Cat#12605036;

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
Dako Liquid DAB+ Substrate Chromogen System	Agilent	Cat#K3468
Annexin V-APC	eBioscience	Cat#BMS306APC
QuickChange Lightning Multi Site-Directed Mutagenesis Kit	Agilent	Cat#210515
Dual Luciferase reporter Assay System	Promega	Cat#E1910
PerfeCTa SYBR Green FastMix	Quanta Biosciences	Cat#733-1387
FastStart Universal SYBR Green Master Mix	Roche	Cat#4913850001
Pierce BCA Protein Assay Kit	Thermo Fisher Scientific	Cat#23227
SuperSignal West Pico PLUS Chemiluminescent Substrate	Thermo Fisher Scientific	Cat#34580
GeneJET RNA Purification Kit	Thermo Fisher Scientific	Cat#K0732
Roche High Pure FFPE Micro Kit	Merck	Cat#4823125001
SuperScript III Reverse Transcriptase	Thermo Fisher Scientific	Cat#18080093
Deposited data		
Microarray expression data of CRC patient samples	Marisa et al.[39]	GEO: GSE39582
Microarray expression data of SCCCs and RCCCs	Puig et al.[8]	ArrayExpress: E-MTAB-4004
Microarray expression data DPPA3 models	This paper	GEO: GSE207862
MethylationEPIC array data DPPA3 models	This paper	ArrayExpress: E-MTAB-12892
Experimental models: Cell lines		
Human: SW1222	Laboratory of Prof. Meenhard Herlyn	RRID: CVCL_3886
Human: NTERA-2	Laboratory of Dr. Maria Abad	RRID: CVCL_0034
Human: DLD1	ATCC	Cat#CCL-221; RRID: CVCL_0248
Human: HCA7	ECACC	Cat#06061902; RRID: CVCL_0289
Human: HCT116	ATCC	Cat#CCL-247; RRID: CVCL_0291
Human: HT29	ATCC	Cat#HTB-38; RRID: CVCL_0320
Human: RKO	ATCC	Cat#CRL-2577; RRID: CVCL_0504
Human: SW480	ATCC	Cat#CCL-228; RRID: CVCL_0546
Human: SW620	ATCC	Cat#CCL-227; RRID: CVCL_0547
Human: HEK293T	Laboratory of Dr. Sandra Peiró	RRID: CVCL_0063
Experimental models: Organisms/strains		
Mouse: NOD-SCID (NOD.CB17-*Prkdc*^{scid}/NcrCrl)	Charles River	Strain code: 394
Mouse: NOD-SCID (NOD.CB17-*Prkdc*$^{scid/scid}$/Rj)	Janvier Laboratories	N/A
Oligonucleotides		
CRISPR sgRNA DPPA3 Forward: CACCGTTTAATCCAACCTACATCCC	Thermo Fisher Scientific	http://crispr.mit.edu
CRISPR sgRNA DPPA3 Reverse: AAACGGGATGTAGGTTGGATTAAAC	Thermo Fisher Scientific	http://crispr.mit.edu
CRISPR sgRNA Scramble Forward: CACCGGCACTACCAGAGCTAACTCA	OriGene	Cat#GE100003
CRISPR sgRNA Scramble Reverse: AAACTGAGTTAGCTCTGGTAGTGCC	OriGene	Cat#GE100003

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CRISPR DPPA3 exon 1 screening Forward primer: AGATATGACATCCGACAGATTAT	Thermo Fisher Scientific	Custom oligos
CRISPR DPPA3 exon 1 screening Reverse primer: CCCCTATAGTCAGATAGTCAAGAA	Thermo Fisher Scientific	Custom oligos
h*DPPA3* promoter cloning Forward primer: 5′-CCGCTCGAGGTGTTGGGACA TGATCTATGC	Thermo Fisher Scientific	Custom oligos
h*DPPA3* promoter cloning Reverse primer: 5′-CCGAAGCTTGCGTCTTGA CACAACACTAGG	Thermo Fisher Scientific	Custom oligos
hDPPA3 cloning Forward primer: 5′-TTTGCTAGCTACCGAGGAGATCTGCC	Thermo Fisher Scientific	Custom oligos
hDPPA3 cloning Reverse primer: 5′-GGTTAATTAAACAGCTATGACCGCGG	Thermo Fisher Scientific	Custom oligos
CTRL shRNA targeting sequence: CCGGCAACAAGATGAAGAGCACCA ACTCGAGTTGGTGCTCTTCATC TTGTTGTTTTT	Sigma-Aldrich	Cat#SHC202
DPPA3 shRNA targeting sequence: CCGGCCATGTGTCTTAGAAGCCCA ACTCGAGTTGGGCTTCTAAGACAC ATGGTTTTTTG	Sigma-Aldrich	Cat#TRCN0000172771
HIF1a shRNA targeting sequence: CCGGCGGCGAAGTAAAGAATCT GAACTCGAGTTCAGATTCTTTAC TTCGCCGTTTTT	Sigma-Aldrich	Cat#TRCN0000003810
h*CA9* promoter cloning Forward primer: 5′-CCGCTCGAGACCTGC CCCTCACTCCAC	Thermo Fisher Scientific	Custom oligos
h*CA9* promoter cloning Reverse primer: 5′-GGGAAGCTTGGCTGA CTGTGGGGTGTC	Thermo Fisher Scientific	Custom oligos
CA9 promoter HRE mutagenesis oligonucleotide: GCTCTCGTTTCCAATGCTGGTA CAGCCCGTACACAC	Thermo Fisher Scientific	https://www.agilent.com/store/primerDesignProgram.jsp
CA9 PCR Forward primer: 5′-GGCTACAGCTGAACTTCCGA	Thermo Fisher Scientific	Custom oligos
CA9 PCR Reverse primer: 5′-AATTCAGCTGGACTGGCTCA	Thermo Fisher Scientific	Custom oligos
DPPA3 PCR Forward primer: 5′-CTACATCCCAGGGTCTCC	Thermo Fisher Scientific	Custom oligos
DPPA3 PCR Reverse primer: 5′-TGATAGTCAAGTTACTAAGGTTCT	Thermo Fisher Scientific	Custom oligos
exDPPA3-FLAG PCR Forward primer: 5′-CAAGACACCAAGCCACTTC	Thermo Fisher Scientific	Custom oligos
exDPPA3-FLAG PCR Reverse primer: 5′-CGTCGTCATCCTTGTAATCC	Thermo Fisher Scientific	Custom oligos

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FOXM1 PCR Forward primer: 5′-TGCAGCTAGGGATGTGAATCTTC	Thermo Fisher Scientific	Custom oligos
FOXM1 PCR Reverse primer: 5′-GGAGCCCAGTCCATCAGAACT	Thermo Fisher Scientific	Custom oligos
HIF1A PCR Forward primer: 5′-TTCCTTCTCTTCTCCGCGTG	Thermo Fisher Scientific	Custom oligos
HIF1A PCR Reverse primer: 5′-ACTTATCTTTTTCTTGTCGTTCGC	Thermo Fisher Scientific	Custom oligos
LDHA PCR Forward primer: 5′-ACGTGCATTCCCGATTCCTT	Thermo Fisher Scientific	Custom oligos
LDHA PCR Reverse primer: 5′-GGGGGTCTGTTCTTCCTTTAGA	Thermo Fisher Scientific	Custom oligos
PPIA PCR Forward primer: 5′-CTCCTTTGAGCTGTTTGCAG	Thermo Fisher Scientific	Custom oligos
PPIA PCR Reverse primer: 5′-CACCACATGCTTGCCATCC	Thermo Fisher Scientific	Custom oligos
SF3A1 PCR Forward primer: 5′-TAAGCCAGTTGTGGGGAT	Thermo Fisher Scientific	Custom oligos
SF3A1 PCR Reverse primer: 5′-GGGGTTGTTGATCTCGTT	Thermo Fisher Scientific	Custom oligos
SLC2A1 PCR Forward primer: 5′-TGGCATCAACGCTGTCTTCT	Thermo Fisher Scientific	Custom oligos
SLC2A1 PCR Reverse primer: 5′-AGCCAATGGTGGCATACACA	Thermo Fisher Scientific	Custom oligos
TBP PCR Forward primer: 5′-CGGCTGTTTAACTTCGCTTC	Thermo Fisher Scientific	Custom oligos
TBP PCR Reverse primer: 5′-CACACGCCAAGAAACAGTGA	Thermo Fisher Scientific	Custom oligos
VEGFA PCR Forward primer: 5′-CATTGGAGCCTTGCCTTG	Thermo Fisher Scientific	Custom oligos
VEGFA PCR Reverse primer: 5′-CGCTGATAGACATCCATG	Thermo Fisher Scientific	Custom oligos
Recombinant DNA		
Plasmid: pSpCas9(BB)-2A-GFP (px458)	Ran et al.[40]	Addgene plasmid #48138
Plasmid: TetON-PGK-rtTA2-P2A-Cherry	From Dr. Stephan Tenbaum	N/A
Plasmid: TetON-*DPPA3-FLAG*-PGK-rtTA2-P2A-Cherry	This paper	N/A
Plasmid: pCMV6-Entry	Origene	Cat#PS100001
Plasmid: pCMV6-DPPA3	Origene	Cat#RC214676
shRNA lentiviral: pLKO1-puro-shCTRL (non-silencing)	Sigma-Aldrich	Cat#SHC202
shRNA lentiviral: pLKO1-puro-sh*DPPA3*	Sigma-Aldrich	Cat#TRCN0000172771
shRNA lentiviral: pLKO1-puro-sh*HIF1A*	Sigma-Aldrich	Cat#TRCN0000003810
pSIN-TRE-H2BeGFP-rtTA2	(Puig et al.)[8]	Addgene plasmid #165494
psPAX2	From Dr. Didier Trono laboratory	Addgene plasmid #12260
pMD2.G	From Dr. Didier Trono laboratory	Addgene plasmid #12259
Vector: pCR4-TOPO	Thermo Fisher Scientific	Cat#450071
Plasmid: pEZX-PF02.1	Genecopoeia	Cat#HPRM38943-PF02
Luciferase reporter vector: pGL3 basic	Promega	Cat#E1751
Renilla Luciferase control reporter vector: pRL-TK	Promega	Cat#E2241

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
GSEA software	Broad Institute	https://www.gsea-msigdb.org/gsea/index.jsp; RRID: SCR_007073
R	The R Foundation	https://www.r-project.org/; RRID: SCR_001905
GraphPad Prism 8	GraphPad Software Inc.	https://www.graphpad.com/scientific-software/prism/; RRID: SCR_002798
7900 SDS 2.4.1 software	Applied Biosystems	Applied Biosystems 7900HT Fast Real-Time PCR System | Thermo Fisher Scientific - ES
QuantStudio 6 Flex Real-Time PCR System software	Applied Biosystems	QuantStudio 6 and 7 Flex Real-Time PCR System Software | Thermo Fisher Scientific - ES; RRID: SCR_020239
NDP.view 2	Hamamatsu	https://www.hamamatsu.com/eu/en/product/life-science-and-medical-systems/digital-slide-scanner/U12388-01.html
QuPath	University of Edinburgh	https://qupath.github.io/; RRID: SCR_018257
FIJI ImageJ	Schindelin et al.[41]	https://imagej.net/software/fiji/downloads; RRID: SCR_002285
BioRender: Scientific image and illustration Software	Shiz Aoki, CEO and Co-founder	https://www.biorender.com/; RRID: SCR_018361
NIS-Elements Advanced Research	Nikon	https://www.microscope.healthcare.nikon.com/products/software/nis-elements/nis-elements-advanced-research; RRID: SCR_002776
FCS 5 express software	*De Novo* Software	https://denovosoftware.com/; RRID: SCR_016431
nSolver 2.0	Nanostring	https://nanostring.com/products/analysis-solutions/nsolver-advanced-analysis-software/; RRID: SCR_003420
Partek Genomics Suite 6.6	Partek inc	https://www.partek.com/partek-genomics-suite/; RRID: SCR_011860

RESOURCE AVAILABILITY

Lead contact

Further information and reasonable requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Isabel Puig (ipuig@vhio.net).

Materials availability

All unique cell models and plasmids generated and described in this study are available from the lead contact with a completed materials transfer agreement upon reasonable request.

Data and code availability

- Microarray expression and methylationEPIC array data that support the findings of this study have been deposited at Gene Expression Omnibus (GEO) under accession code GEO: GSE207862 and ArrayExpress under accession code ArrayExpress: E-MTAB-12892, respectively, and are publicly available as of the date of publication. Previously published microarray expression data of SCCC s and RCCCs that was reanalyzed here is available at ArrayExpress under accession code ArrayExpress: E-MTAB-4004. Previously published microarray expression data of CRC patient samples reanalyzed here is available at GEO under accession code GEO: GSE39582.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT PARTICIPANT DETAILS

Cell lines

Human SW1222 colorectal cancer (CRC) cells (RRID: CVCL_3886) were kindly provided by Prof. Meenhard Herlyn, NTERA-2 (NT2; RRID: CVCL_0034) testicular embryonal carcinoma cell line by Dr. Maria Abad and HEK293T (RRID: CVCL_0063) cell line by Dr. Sandra Peiró. Other human CRC cell lines used were DLD1 (#CCL-221; RRID: CVCL_0248), HCT116 (#CCL-247; RRID: CVCL_0291), HT29 (#HTB-38; RRID: CVCL_0320), RKO (#CRL-2577; RRID: CVCL_0504), SW480 (#CCL-228; RRID: CVCL_0546), and SW620 (#CCL-227; RRID: CVCL_0547) CRC purchased from American Type Culture Collection (ATCC). HCA7 (#06061902; RRID: CVCL_0289) cell line was obtained from the European Collection of Authenticated Cell Cultures (ECACC). Cell lines were maintained in Dulbecco's modified eagle medium (DMEM; Biowest; #L0102) supplemented with 10% Fetal bovine serum (FBS; Gibco; #10270-106) (complete DMEM) and 1% penicillin/streptomycin (P/S; Life Technologies; #15140122) at 37°C in a humidified incubator with 5% CO_2. Cells were used for experiments after culturing for fewer than 20 passages.

Mice

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, 17/15 CEEA, 18/15 CEEA and 12/18 CEEA). All animal used in this study are female NOD-SCID (NOD.CB17-*Prkdc*^{scid}/NcrCrl) or (NOD.CB17-*Prkdc*$^{scid/scid}$/Rj) *aged between 6-8 weeks (21-24g) and were purchased from* Charles River (strain code: 394) and Janvier Laboratories. Mice were randomized by tumor size, and those that died before the end of the experiments were excluded. Experiments were not performed in a blinded fashion. All mice were closely monitored by authors, facility technicians and by a veterinary scientist responsible for animal welfare. Mice were maintained in a specific-pathogen-free (SPF) facility under controlled temperature and humidity and given *ad libitum* access to standard diet and water.

Patient-derived xenograft (PDX) establishment

Human colorectal carcinoma tissues were obtained upon surgery in accordance with the ethical standards of the institutional committee on human experimentation. Histological diagnosis was based on microscopic features of carcinoma cells determining the histological type and grade. Cancer tissues were washed 3 times in cold PBS solution and incubated overnight in DMEM/F12 (Gibco; Thermo Fisher Scientific) 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; #C0130) and DNase I (20 μg/mL; Sigma-Aldrich; #D4263) 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)) during 1 h at 37°C with intermittent pipetting every 15 min to disperse cells. The dissociated sample was then filtered (100 μm pore size) and washed with fresh medium. Red blood cells were lysed by 10 min exposure to ammonium chloride and the sample was washed again. Finally, $1x10^5$ cells were resuspended in 50 μL of PBS and mixed with 50 μL of Matrigel, and subcutaneously injected into NOD-SCID flank mice to generate subcutaneous tumors further used for histological and RNA analysis.

Clinical cohorts

The study for patients' VHIO samples was approved by the Vall d'Hebron University Hospital institutional ethical review board according to the guidelines of the European Network of Research Ethics Committees, following European, national, and local laws. Written informed consent was signed by all patients.

CRC primary tumor vs. liver metastases-VHIO cohort (private)

The percentage of nuclear DPPA3-positive cells was evaluated by immunohistochemistry on a tissue microarray containing 130 paired patient samples (''Method Details'' section). Primary tumors presenting percentages of nuclear DPPA3-positive cells above 20% were considered ''DPPA3-High'' cases, and equal or under 20%, ''DPPA3-Low'' cases (Figure S1E). We compared percentage of nuclear DPPA3 levels of primary tumor samples with their matched liver metastasis from the same patient using paired *t* test (n = 65) (Figure 1I; Table S1). We then restricted the cohort to those patients with metachronous metastases (n = 27) to analyze the percentage of nuclear DPPA3-positive cells in two patient groups according to time to relapse (Figure 1J). Patients who relapsed 2 years or earlier from diagnosis were considered ''early relapse'' cases ($\leq$2y), and after 2 years, ''late relapse'' cases (>2y).

CRC patients-VHIO cohort (private)

DPPA3 expression levels was evaluated by nCounter (''Method Details'' section) on a high-risk stage II or stage III CRC patient cohort that received adjuvant chemotherapy and that had outcome data available (n = 53). Samples presenting DPPA3 mRNA expression values above the overall mean expression were considered ''DPPA3-High'' cases, and equal or under the overall mean expression, ''DPPA3-Low'' cases (Figure S1G). We performed Cox Proportional Hazards univariable modeling using GraphPad Prism 8 software (Figure S1H). Progression-free survival (PFS) was defined as the time from diagnosis to relapse or death due any cause.

LARC patients-VHIO cohort (private)

The percentage of nuclear DPPA3 and CA9-positive cells was evaluated by immunohistochemistry on FFPE as described below (''Method Details'' section). Samples presenting percentages of nuclear DPPA3-positive cells above 20% were considered

"DPPA3-High" cases, and equal or under 20%, "DPPA3-Low" cases (Figure S1I). The percentage of nuclear DPPA3-positive cells was evaluated on 39 biopsies of rectal adenocarcinomas (ADK) from a cohort of locally advanced rectal cancer (LARC) patients[23] mostly of them treated with chemoradiotherapy (CRT) (Table S2). Treatment response to CRT was classified based on the tumor regression system. We calculated the proportion of LARC patients who showed complete response (absence of residual invasive tumor in the surgical resection specimen (GR0) or near complete response (GR1)), intermediate response (GR2), and no response (GR3 to GR5) according to DPPA3 percentages (n = 37) (Figure 2G). We also calculated the proportion of LARC patients who were alive or dead according to DPPA3 percentages (n = 39) (Figure 2H). We performed Cox Proportional Hazards univariable modeling using GraphPad Prism 8 software (Figure 2I and S1J-S1L). Relapse-free survival (RFS) was defined as the time from diagnosis to relapse, and progression-free survival (PFS) was defined as the time from diagnosis to relapse or death due any cause. We compared the percentage of nuclear DPPA3-positive cells of rectal adenocarcinoma biopsies before neoadjuvant chemoradiotherapy (pre-nCRT) with their matched resected sample after treatment (post-nCRT) (n = 27) (Figure 2J). Finally, we analyzed the percentage of CA9-positive cells on rectal adenocarcinoma resected samples (post-nCRT) in two patient groups according to DPPA3 percenatges (n = 27) (Figure 5K).

PDX-VHIO collection (private)

The percentage of CA9-and nuclear DPPA3-positive cells were evaluated on a collection of 61 CRC patient-derived xenografts (PDXs) by immunohistochemistry on a tissue microarray as described below ("Method Details" section). The percentage of nuclear DPPA3-postive cells per CRC PDX was correlated with their corresponding CA9 staining (Figure 5F).

GSE39582 (public)[39]

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 signature enrichment score (Sig ES) was calculated and samples at the upper quartile were classified as "DPPA3 Sig ES High" while the rest were classified as "DPPA3 Sig ES Low". We calculated the DPPA3 Sig ES by subtracting the median intensity of the signature downregulated genes from the median intensity of the signature upregulated genes across all samples and compared with consensus CRC subtyping label[29] with 1-way ANOVA test (Figure 5G). 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 or death due to any cause. We performed Cox Proportional Hazards univariable and multivariable modeling using GraphPad Prism 8 software and R package *survival*, respectively (Figures 5H, S4D, and S4E).

METHOD DETAILS

Plasmids

All oligonucleotide and primer sequences used to generate all plasmids are described in the key resources table. For generation of *DPPA3*-KO cell lines, a DPPA3-specific sgRNA oligo was cloned into the pSpCas9(BB)-2A-GFP (px458)[40] 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 (Figure S4B), 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. For generation of the lentiviral vector TetON-DPPA3-FLAG-PGK-rtTA2-P2A-Cherry, the coding sequence of human *DPPA3* fused to *CMYC* and *FLAG* sequences was amplified from the pCMV6-*DPPA3* plasmid (Origene plasmid ID: RC214676) by PCR with specific primers containing NheI and PacI restriction sites. The resulting PCR product was then subcloned into the NheI-PacI cut TetON-PGK-rtTA2-P2A-Cherry plasmid, kindly provided by Dr. Stephan Tenbaum, to obtain the final TetON-*DPPA3-FLAG*-PGK-rtTA2-P2A-Cherry construct (Figure S2B). shRNA knockdown control (shCTRL), against *DPPA3* transcript (sh*DPPA3*) (NM_199286.2) and against *HIF1α* (sh *HIF1α*) transcript (NM_001530) were performed using the pLKO1-puro TRC1, 1.5 and TRC2 Mission shRNA Lentiviral vector systems (Sigma Aldrich; #SHC202, #TRCN0000172771 and #TRCN0000003810, respectively). Human *CA9* promoter region (−1476/+42bp) and *DPPA3* (NM_199286.4) promoter region (−977/+76bp) were amplified by PCR from the pEZX-PF02.1 plasmid (Genecopoeia; HPRM38943-PF02) and genomic DNA extracted from SW1222 cells, respectively, using specific primers containing XhoI and HindIII restriction sites. Amplified products were subcloned upstream of a *Firefly* luciferase open reading frame into a Luc reporter pGL3 basic vector (Promega; #E1751). Site-directed mutagenesis at core HRE sequence of human *CA9* promoter has been previously used[42] and was performed in the *CA9*promoter-pGL3 reporter construct using the QuickChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent Technologies; #210515) following manufacturer's instructions and verified by Sanger sequencing.

Generation of DPPA3 knockout cell lines

Primer sequences used to generate *DPPA3* knockout cell lines are described in the key resources table. *DPPA3* knockout (*DPPA3* KO) and control (Scramble) cell lines were generated using CRISPR-Cas9 system. SW1222 cells were transfected with pSpCas9-sgRNA*DPPA3*guide2-2A-GFP or pSpCas9-sgRNAscramble-2A-GFP constructs using linear polyethylenimine (PEI 25000, Polysciences, Inc.; #23966). Forty-eight hours post-transfection, cells were pelleted and sorted using FACS Aria II cell sorter (BD BioSciences). Single cells were expanded to obtain individual clones. Given that DPPA3 protein levels are not detected in the

SW1222 cell line (Figure S2E), we used a PCR-based screening strategy to validate our SW1222 *DPPA3*-KO models.[43] Briefly, 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 and products were analyzed by Sanger sequencing. To validate biallelic mutation in each clone, genomic DNA was isolated from edited clones and non-edited SW1222 control cells. Exon 1 of *DPPA3* was amplified by PCR and products were A-tailed and cloned into pCR4-TOPO vector (Thermo Fisher Scientific; #450071) following the manufacturer's instructions. Individually cloned PCR products were then analyzed by Sanger sequencing (Figure S4B).

Generation of DPPA3-overexpressing and knockdown cell lines

In all cases, lentiviruses were produced in HEK293T cells using standard procedures and psPAX2 and pMD2.G (Addgene, plasmids ID: 12260 and 12259, respectively; Dr. Didier Trono's Laboratory) packaging vectors. Forty-eight after transfection, the supernatant was collected and filtered. This supernatant was then used to transduce cells directly. To obtain SW1222 Histone2B fused to the eGFP (H2BeGFP) clone, cells were transduced with lentiviruses expressing H2BeGFP protein[8] (pSIN-TRE-H2BeGFP-rtTA2; Addgene, plasmid ID: 165494). Afterward, single cells were expanded to obtain individual clones. For generation of the doxycycline (DOX)-inducible DPPA3 overexpressing (DPPA3) cells, SW1222 and HT29 cells were transduced with lentiviruses expressing human DPPA3 protein fused with the sequences encoding CMYC and FLAG tags (TetON-*DPPA3-FLAG*-PGK-rtTA2-P2A-Cherry) (Figures S2B–S2E). TetON-*DPPA3-FLAG*-PGK-rtTA2-P2A-Cherry plasmid generation is described above. SW1222 and HT29 cells infected with the empty lentiviral vector (TetON-PGK-rtTA2-P2A-Cherry) were used as control cells. After transduction, cells were sorted using fluorescence-activated cell sorting (FACS) with a FACS Aria II cell sorter (BD BioSciences). Individual DPPA3 clones were selected from expanded single cells. Clones generated from DPPA3-overexpressing SW1222 or HT29 cells were used in all overexpression experiments unless otherwise specified. *DPPA3* knockdown SW1222-H2BeGFP and HT29 cells and *HIF1α* knockdown SW1222-H2BeGFP, SW1222-DPPA3 and SW1222-CTRL cells were generated by expressing non-targeting shRNA (shCTRL) and DPPA3 TRC-sh*RNA* (sh*DPPA3*) or HIF1α TRC-shRNA (sh*HIF1α*) and selected with Puromycin (1 μg/mL; Thermo Fisher Scientific; #A111803) (Figures S1D, S4F, S5F, S6A, and S6B). TRC-sh*RNA* plasmids are described above.

Two-dimensional (2D) cell culture and hypoxia treatment

Cell lines were cultured in two dimensions and maintained in Dulbecco's modified eagle medium (DMEM; Biowest; #L0102) supplemented with 10% Fetal bovine serum (FBS; Gibco; #10270-106) (complete DMEM) and 1% penicillin/streptomycin (P/S; Life Technologies; #15140122). Normoxic cultured cells were grown at atmospheric O_2 concentrations (21%) with 5% CO_2. Hypoxia treatment of cells followed normoxic culture and consisted in the incubation of cells during 24 h 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 hypoxia treatment, protein or RNA extraction was performed. To induce DPPA3 expression, DPPA3 and control SW1222 and HT29 cells were treated with 5 μg/mL of DOX (Sigma-Aldrich; #D9891). DOX treatment schedule for each experiment is schematically depicted in Figure panels.

5-Azacytadine treatments *in vitro*

To assess the recovery of *FOXM1* expression in SW1222-DPPA3 cells, these were seeded in 2D and pretreated for 72h with 5 μg/mL of DOX (Sigma-Aldrich; #D9891). On the day of treatment with 5-azacytadine (AZA; MedChemExpress; #HY-10586), cells were refreshed with DOX and then treated with AZA at 2.5 μm and 5 μm for 72h when assayed the RNA expression.

Three-dimensional (3D) cell culture

SW1222-H2BeGFP pool and clone, DPPA3, sh*DPPA3*, *DPPA3* KO and respective control SW1222 cells were maintained in complete DMEM and 1% P/S. Three-dimensional (3D) organoids embedded in Corning Matrigel Basement Membrane Matrix (Matrigel; Cultek; #45354277) from SW1222 cells were generated from single-cells resuspended in complete DMEM and mixed 1:1 with Matrigel. Drops of 25 μL of the mixture were seeded in culture plates and incubated for 30 min at 37°C without medium to allow the matrix solidification. Then, complete DMEM was added and changed twice weekly. Cells were cultured until they reached a well-differentiated heterogeneous organoids (3 weeks of culture) and were used for further analyses. For H2BeGFP and DPPA3 overexpression induction, 3D-organoids were treated with 5 μg/mL of DOX. DOX treatment schedule for each experiment is schematically depicted in Figure panels.

Paraffin-embedded of 3D organoids pellet

DOX pulse-chased SW1222-H2BeGFP pool 3D organoids were removed from Matrigel using the ice-cold Corning Cell Recovery Solution (Cultek; #45354253) following the manufacturer's protocol. Three-D organoids were rinsed three times with cold PBS and centrifuged at 830g for 5 min without disturbing the pellet. The resulting mix of organoids pellet was fixed with 4% paraformaldehyde (PFA; Santa Cruz Biotechnology; #sc-281692) at 4°C overnight. Then, the fixing solution was replaced with 70% ethanol and organoids pellet was dehydrated and embedded in paraffin blocks. Finally, paraffin-embedded pellet sections were cut 4 μm thick and analyzed by fluorescent immunohistochemistry (IHC-F).

SCCC chemoresistant assays

Parental and shCTRL/sh*DPPA3* or sh*HIF1A* SW1222-H2BeGFP cells were grown in 3D-cultures as is described above. After a DOX (5 μg/mL) pulse-chase treatment (9 days), cells were treated with 50 μg/mL of 5-fluorouracil (5FU) for an additional 7 days before apoptosis was measured. Detailed treatments schedule for each experiment are schematically depicted in Figure panels. After treatments, 3D organoids were recovered from Matrigel using the ice-cold Corning Cell Recovery Solution (Cultek; #45354253) following the manufacturer's protocol. 3D organoids were filtered through a 100 μm cell strainer (Corning) to purify megacolonies and colonies, which are collected from the cell strainer surface and dissociated using Tryple Express Enzyme (ThermoFisher Scientific; #12605036) to obtain a single cell suspension. Finally, the proportion of apoptotic cells was determined using the Annexin V-APC kit (eBioscience). Cells were analyzed by flow cytometer using a Navios Cytometer (Beckman Coulter).

SCCC self-renewal assays

shCTRL and sh*DPPA3* SW1222-H2BeGFP cells were grown in 3D-cultures and dissociated as is described in "SCCC chemoresistant assays" method. Detailed treatments schedule for each experiment are schematically depicted in Figure panels. Single-cells were resuspended in sorting medium: 4 mM Glutamine, 20% FBS, 1% penicillin/streptomycin, 10 μM Y-27632 (Calbiochem) diluted in DMEM. DAPI was added to exclude dead cells in all cell suspensions. Live SCCC and RCCC were sorted using FACSAria II cell sorter (BD Biosciences). The population of cells retaining an H2BeGFP signal equivalent to that observed in cells continuously exposed to DOX, were considered and sorted as SCCC (3–5%, Figure S1B). The RCCC fraction comprised around 10–20% of all cells with an H2BeGFP signal one order of magnitude lower than the SCCC fraction. SCCC and RCCC were collected into sorting medium and re-embedded in Matrigel as described in "3D cell culture" method. Matrigel-embedded 3D cultures were maintained for up to 3–5 weeks. Organoids were counted under microscope according to their size: megacolonies ($\geq$400 μm), colonies (399-50 μm), and minicolonies (<50 μm). 3D organoids forming units (%) were estimated according to the formula: number of structures/number of live cells plated x 100.

Luciferase reporter assay

CA9 promoter activity was analyzed by co-transfecting HEK293T cells with either pCMV6-Entry or pCMV6-*DPPA3* (NM_199286) plasmids (Origene; ID: PS100001 and RC214676, respectively) together with the mutated (HRE mut) and non-mutated (WT) *CA9*promoter-pGL3 reporter constructs. *DPPA3* promoter activity was analyzed by transfecting *DPPA3*promoter-pGL3 construct in NT2 cells. In all cases, pRL-TK plasmid was co-transfected as reference (Promega; #E2241). Forty-eight hours after transfection, cells underwent hypoxia treatment (0.5% O_2) for 24h when indicated. Luc and RLuc activities were separately measured with the Dual-Luciferase Reporter Assay System (Promega; #E1910) according to the manufacturer's instructions. Luc activity was normalized to RLuc. Experiments were performed in triplicates.

Protein extraction

Total cell extracts from cells were homogenized in SDS lysis buffer (1% SDS, 10 mM EDTA pH 8) containing protease inhibitors cocktail (Complete, Mini, EDTA-free protease inhibitor cocktail, Roche; Sigma-Aldrich; #11836170001). For subcellular fractionation, cells were lysed in buffer A (10 mM HEPES pH 7.8, 10 mM KCl, 1.5 mM $MgCl_2$, and 0.5 mM DTT) containing protease inhibitors cocktail for 10 min at 4°C. Then, 1/3 volume of 10% Triton X-100 was added. Samples were centrifuged at 15,000 g at 4°C for 1 min 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 min at 4°C. Samples were centrifuged at 15,000 g at 4°C for 15 min 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 min at 4°C. Then, samples were centrifuged at 15000 g for 15 min at 4°C. Pelleted chromatin was washed twice in buffer C and resuspended in SDS lysis buffer (1% SDS, 10 mM EDTA pH 8). Chromatin was sonicated for homogenization. Following protein extraction, samples were quantified using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific; #23227). Three-D organoids were removed from Matrigel prior protein extraction using Corning Cell Recovery Solution (Cultek; #45354253) following manufacturer's instructions.

Western blot analyses

Western blots were performed according to standard procedures. Briefly, samples were boiled at 95°C for 5 min. Lysates were separated by SDS-PAGE and protein was transferred onto a nitrocellulose (BioRad; #1620115) or PVDF (Millipore; #IPVH00010) membranes. Membranes were blocked with TBS, 0.1% Tween 20, and 5% non-fat milk or BSA (Sigma.Aldrich; #A3059), incubated overnight at 4°C with corresponding primary antibodies (CA9 1/4000; CCNA2 1/1000; CCNB1 #ab32053 1/2000; CDK1 1/1000; CDK2 1/1000; cleaved-CASP3 #9664 1/1000; DNMT1 1/4000; DPPA3 sc-376862 1/100; FLAG 1/2000; FOXM1 1/1000; GAPDH 1/10000; H3 1/20000; HIF1α 1/1000; HIF1α-OH 1/10000; LMNB1 1/10000; and TUB 1/10000) and 1h at RT with HRP-conjugated secondary antibodies (1/1000-1/10000) diluted in blocking solution, revealed using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific; #34580), and visualized using X-ray films (FUJIFILM; #47410 19289) and Curix 60 processor (AGFA) or Amersham Imager 600 (GE Healthcare).

Fluorescent immunocytochemistry (IC-IF) staining

SW1222 and HT29 CTRL or DPPA3 cells were grown on coverslips. For DPPA3 induction, cells were treated with 5 μg/mL of DOX for 5 days prior to fixation. For hypoxia treatment, cells were subjected to hypoxia 24h prior to fixation. Cells were washed with PBS and fixed in 4% PFA (Santa Cruz Biotechnology; #sc-281692) for 15 min at RT. Cells were washed in 0.1% Tween 20 in PBS (PBST), followed by permeabilization with 1% Triton X-100 in PBST for 1 h and 30 min at RT. Cells were blocked overnight in 4% donkey serum in PBST at 4°C. Corresponding primary antibody (CCNB1 #ab72 1/200; DPPA3 #MAB4388 1/100; p-H2A.X 1/100) was added in 1% donkey serum for 4 h at RT. Cells were washed in PBST and incubated with corresponding Alexa Fluor-conjugated secondary antibody (1/200) in 1% serum together with Hoechst 33342 (5 μg/mL; Sigma-Aldrich; #B2261) as counterstaining for 30 min at RT. Samples were washed with PBST and mounted using VECTASHIELD Mounting Medium (VECTOR laboratories; #H1000-10). A Nikon C2+ Confocal Microscope was used to visualize fluorescence and images acquired using NIS-Elements Advanced Research software (Nikon).

Immunohistochemical (IHC) staining

Formalin-fixed paraffin-embedded (FFPE) tissues, tissue-microarrays (TMA) and paraffin-embedded megacolonies/colonies pellet sections were routine deparaffinated, rehydrated and treated with 1mM EDTA, 10mM Tris-base, 0.05% Tween 20, pH 9 buffer for GFP, CA9 and DPPA3 stainings or with 10 mM sodium citrate pH 6 buffer, for Cleaved-Caspase 3, KI67, CK20, and CCNB1. For all antibodies, after blocking endogenous peroxidase activity, slides were permeabilized with 1% Tween 20 in PBS for 15 min. Then, tissue specimens were blocked for 1 h with 3% BSA in PBS and incubated with corresponding primary antibodies (CA9 1/1000; CCNB1 #ab72 1/200; CK20 1/100; cleaved-CASP3 #9661 1/100; DPPA3 #sc-376862 1/100 and KI67 1/100) diluted in blocking solution at 4°C overnight. For chromogenic detection, after washing, sections were incubated with corresponding HRP-conjugated secondary antibodies (1/250) for 1 h at RT. For DPPA3 detection, a SignalStain Boost IHC Detection Reagent (Cell Signaling Technology, #8125) was used. After washing, Dako Liquid DAB+ Substrate Chromogen System (Agilent; #K3468) was added onto the slides and incubated up to 10 min. Finally, the slides were counterstained with hematoxylin, dehydrated, and mounted. NanoZoomer 2.0-HT Digital slide scanner C9600 (Hamamatsu Photonics K.K.; RRID: SCR_021658) was used to acquire a high-resolution whole slide scanner of the immunostainings.

For fluorescent detection, after primary antibodies (CA9 1/1000; GFP 1/100) incubation and washes, slides were incubated with the corresponding Alexa Fluor-conjugated secondary antibody (1/200; Thermo Fisher Scientific) for 1 h at RT. Hoechst 33342 (5 μg/mL; Sigma-Aldrich; #B2261) was used as counterstaining to detect cell nuclei in all samples. After washing with PBS, slides were mounted using VECTASHIELD Mounting Medium (VECTOR laboratories; #H1000-10). An Olympus FV1000 Confocal Microscope (Olympus) was used to visualize fluorescence and images acquired using FluoView 4.2 (Olympus).

Cell cycle analyses

DPPA3 and control SW1222 and HT29 cells and DPPA3 sh*HIF1A,* DPPA3 shCTRL and control SW1222 cells were cultured and treated with 5 μg/mL of DOX (Sigma-Aldrich; #D9891) during 5 days to induce DPPA3 expression. Then, $1x10^6$ cells were trypsinized and resuspended in 0.5 mL of cold PBS followed by fixation with 1.5 mL of ice-cold absolute ethanol. Cells were washed with ice-cold PBS and resuspended in 1 mL of ice-cold PBS. 100 μL of ribonuclease A (100 μg/mL, DNase free, PanReac AppliChem; #A3832) were added to the samples at RT for 5 min. Then, 400 μL of propidium iodide (50 μg/mL in PBS, Sigma-Aldrich; #P4170) was added and incubated at 4°C for 1 h. Finally, cells were pelleted and resuspended in 500 μL of ice-cold PBS. Cell cycle profile was analyzed by flow cytometry using a Navios Flow Cytometer (Beckman Coulter) and FCS 5 express software (*De Novo* Software; RRID: SCR_016431). Experiments were performed in triplicates.

DPPA3 profiling of CRC PDX FFPE samples (nCounter)

The evaluation of *DPPA3* expression in 53 CRC patient-derived xenografts (PDX-VHIO collection) of similar sizes was analyzed using nCounter platform[44] (Nanostring Technologies; RRID: SCR_021712). Hematoxylin and eosin staining was performed in each FFPE xenograft tissue. Areas enriched in tumor tissue were identified. A minimum of two 1-mm FFPE tumor tissue corres were collected. RNA was purified using HighPure FFPE Micro Kit (Roche; #06650775001), and approximately 100 ng of total RNA was used to measure expression of *DPPA3* using the NanoString Technologies nCounter Platform. Fluorescently labeled probe is designed for *DPPA3* gene and allowed to hybridize to target RNA, and then captured and individual RNA molecules counted using color-coded probe pairs. Raw data was log base 2 transformed and normalized using five house-keeping transcripts in the nSolver 2.0 software (Nanostring; RRID: SCR_003420). *DPPA3* expression values in the upper quartile were classified as High expression, while expression values in the lower quartile were classified as Low.

RNA extraction

For RNA extraction of 2D cultures, cells were washed twice with PBS and lysed with 1 mL of Trizol reagent (Thermo Fisher Scientific; #15596018). For RNA extraction of SW1222 3D-organoids embedded in Matrigel, cells were harvested using Corning Cell Recovery Solution (Cultek; #45354253) following manufacturer's instructions. Then, pelleted cells were collected into 1 mL of Trizol reagent. For RNA extraction of tumor xenografts, snap-frozen samples were homogenized using a pestle in 1 mL of Trizol reagent. Cell lysates from *in vitro* or *in vivo* experiments were mixed with 200 μl of RNase-free chloroform and incubated at RT for 2 min. The solution

was centrifuged at 15,900g for 15 min at 4°C and the supernatant was transferred to a new sterile 1.5 mL microtube and mixed with 500 μl of RNase-free isopropanol and incubated at RT for 10 min. RNA was precipitated by centrifugation at 15,900g for 30 min at 4°C. Pellets were washed once with 800 μl of 75% RNase-free ethanol and centrifuged at 15,900g for 5 min at 4°C. After evaporation of ethanol traces, pellet was gently resuspended in DEPC-water and quantified. Alternatively, RNA was extracted with GeneJET RNA Purification Kit (Thermo Fisher Scientific; #K0732) according to manufacturer's instructions.

RNA extraction from SCCC, RCCC and sRCCC

SCCC and RCCC RNA analyzed in this study was extracted previously[8] (Figure S1B) and is detailed as following. Matrigel-embedded SW1222-H2BeGFP clone 3D-organoids were maintained on 5 μg/mL of DOX (Sigma-Aldrich; #D9891) medium for two weeks (pulse), and then grown on DOX-free medium for additional 7 days (chase) (Figure S1A). Then, 3D-organoids were harvested using ice-cold Corning Cell Recovery Solution (Cultek; #45354253) following manufacturer's protocol and 3D colonies and megacolonies were collected from the 100 μm cell strainer surface. After 3D colonies/megacolonies dissociation using trypsin-EDTA, single-cell pellet was resuspended in sorting medium: 4 mM Glutamine, 20% FBS, 1% penicillin/streptomycin, 10 μM Y-27632 (Sigma-Aldrich; #Y0503) diluted in CO2-independent medium. Finally, live SCCC, RCCC and sRCCC were sorted using FACSAria II cell sorter (BD Biosciences). The population of cells retaining an H2BeGFP signal equivalent to that observed in cells continuously exposed to DOX, were considered and sorted as SCCC (Figure S1C). The RCCC fraction comprised around 10–20% of all cells with an H2BeGFP signal one order of magnitude lower than the SCCC fraction. The gate of DOX-untreated cells was used to identify and sort the sRCCC fraction. Finally, SCCC, RCCC and sRCCC were collected directly into Trizol reagent (Thermo Fisher Scientific; #15596018) not exceeding 10% of its volume for RNA extraction. Total RNA of biological replicates was isolated following manufacturer's instructions. RNA integrity was confirmed in an Agilent 2100 Bioanalyzer using RNA Nano- or Pico-chips. Total RNA was linearly amplified using Ovation Pico WTA System V2 (Nugen).

Gene expression profiles and analyses

Microarray gene expression analysis was performed on DPPA3, sh*DPPA3* and *DPPA3* KO (clones C66 and C75) SW1222 CRC cells and respective controls. For DPPA3 long-term induction and loss-of-function models, cells were grown as 3D megacolonies/colonies embedded in Matrigel. For DPPA3 short-term induction, cells were grown in 2D cultures. In the case of DPPA3 and respective controls, cells were treated with 5 μg/mL of DOX (Sigma-Aldrich; #D9891) during the last 5 days to induce DPPA3 expression (Figure S4C). DPPA3 short-induction condition and control cells cultured in 2D were treated with 5 μg/mL of DOX for 4 h.

Total RNA of biological triplicates was isolated by Trizol (Thermo Fisher Scientific; #15596018) extraction as described above. Transcriptomes were determined on a genome wide GeneChip Human Gene 1.0 ST Array (HuGene-1_0-st-v1; Affymetrix; #901087). RNA integrity was confirmed in an Agilent 2100 Bioanalyzer using RNA Nano-chips. Hybridization data was acquired using the Affymetrix GeneChip/GeneTitan platforms. We used Partek Genomics Suite 6.6 software (Partek Inc.; RRID: SCR_011860) 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 versus control (2D and 3D conditions), sh*DPPA3* versus shCTRL, 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. Microarray expression data that support the findings of this study have been deposited at Gene Expression Omnibus (GEO) under accession code GSE207862. Gene set enrichment analyses (GSEA) and leading-edge analyses[45,46] were performed with GSEA software from the Broad Institute (http://www.broadinstitute.org/gsea; RRID: SCR_003199) and also using the clusterProfiler R package (Bioconductor; RRID: SCR_016884). We used custom and publicly available gene sets (Molecular Signatures Databases v6.2; http://www.broadinstitute.org/gsea; RRID: SCR_016863).

Custom gene sets used in Figure 3B (SCCC_PAN_CANCER_DOWNREGULATED), S4A (SCCC_PAN_CANCER_UPREGULATED), 3C (DPPA3_CELL_CYCLE_DW), and 5D and 5I (DPPA3_HYPOXIA_UP) are detailed in Puig et al.[8] and Tables S4 and S5, respectively. Gene expression analysis of DPPA3 SW1222 (2D and 3D), SW1222 SCCC and RCCC previously generated in the laboratory[8] (ArrayExpress: E-MTAB-4004), and GEO: GSE39582 CRC patients' cohort were used to generate GSEA plots shown. Visualization of the results were done using the Circlize and ggplot2 R packages. This includes enrichment scores in chord diagrams as well as a volcano plot representing the differential expression of genes having at least one occurrence of the HIF1α binding motif in their transcription starting sites (HIF_Q3 and HIF_Q5 gene sets from Molecular Signatures Databases v6.2; http://www.broadinstitute.org/gsea; RRID: SCR_016863) in SW1222 DPPA3 versus control cells. DPPA3_CELL_CYCLE_DW and DPPA3_HYPOXIA_UP gene sets were created grouping the genes present in the core enrichment of the leading edge analysis of the gene sets indicated in Tables S4 and S5, respectively.

Quantitative RT-PCR analyses

To analyze the expression of selected genes in SCCC and RCCC, the resulting single-stranded DNA obtained from the total RNA linearly amplified after RNA extraction was used as input (2 ng) for quantitative PCR (qPCR). To analyze expression genes from cell lines, 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 (Thermo Fisher Scientific; #18080093) following manufacturer's instructions. Analyses were carried out in cDNA triplicates using PerfeCTa SYBR Green FastMix (Quanta Biosciences; VWR; #733–1387) or FastStart Universal SYBR Green Master Mix (Roche Merck; #4913850001) in a 7900HT qPCR System (Applied Biosystems) or Quant Studio 6 Flex (Applied Biosystems) cyclers. Specific pairs of primers were designed (https://www.ncbi.nlm.nih.gov/gene/) and used to detect the indicated transcripts, all sequences are described in the key resources table. Relative gene expression was determined by the comparative C_T method.[47] We applied geNorm algorithms.[47,48] We applied geNorm algorithms[48] to select TATA binding protein (*TBP*), peptidyl-prolyl isomerase A (Cyclophilin A, *PPIA*) and Splicing Factor 3a Subunit 1 (*SF3A1*) 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 analyzed. Experiments were performed in triplicates.

DPPA3 gene expression signature

Based on the microarray gene expression analysis performed in DPPA3 (short-term (2D) and long-term (3D) induction), sh*DPPA3* and *DPPA3* KO (clones C66 and C75 (3D)) SW1222 CRC, we derived the DPPA3 gene expression signature (DPPA3 Sig) (Figure S4C). We considered as DPPA3-induced genes (50 in total) those positively regulated ($\geq$1.2 FC between conditions, p value <0.05) in DPPA3 cells that were common with those negatively regulated ($\leq$-1.2 FC between conditions, p value <0.05) in sh*DPPA3* or *DPPA3* KO (C66 and C75) cells in at least 3 of the 4 different conditions (sh*DPPA3* and *DPPA3* KO (C66 and C75), and 3D and 2D-cultured DPPA3 cells) (Figure S4C). We considered as DPPA3-repressed genes (32 in total) those negatively regulated ($\leq$-1.2 FC between conditions, p value <0.05) in DPPA3 cells that were common with those positively regulated ($\geq$1.2 FC between conditions, p value <0.05) in sh*DPPA3* or *DPPA3* KO cells in at least 3 out of the 4 different conditions (Figure S4C). All genes included in DPPA3 Sig are listed in Table S6. To apply the DPPA3 Sig on the clinical cohort used (GEO: 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).

DNA extraction

Genomic DNA was purified from control or DPPA3 overexpressing SW1222 cells cultured in 3D conditions (Figure S2F) using 500 μl of 0.5% SDS Lysis buffer (100 mM NaCl, 20 mM Tris pH: 8.0, 25 mM EDTA pH: 8.0, 0.5% SDS) and 50 μl (1 mg/mL) Proteinase K (Roche, #03115801001) for 4 h at 56°C in gentle agitation to lysate the samples. A standard phenol-chloroform protocol for phase separation and ethanol extraction as the precipitation step was performed for further sample processing. Briefly, 500 μl of the phenol – chloroform – isoamyl alcohol mixture (49.5:49.5:1) (Sigma-Aldrich, #77618) was added to the previous lysate and mixed until the sample becomes homogeneous. The resulting solution was centrifuged at 15,900G for 5 min at 4°C and the supernatant was transferred to a new sterile 1.5 microtube and mixed with 1/10 volumes of 3M sodium acetate and 2 volumes of DNAse-free ethanol. Samples were freezed for 1 h at −80°C and centrifuged at 15,900G for 30 min at 4°C. The resulting supernatant was discarded and pellets were washed once with 500 μl of 70% DNAse-free ethanol. A final centrifugation step at 15,900G for 5 min at 4°C was performed and supernatants were discarded. Pellets were dried at 37°C for 15 min and were resuspended in DEPC-water.

DNA methylation analyses

DNA samples (500 ng each) were hybridized to the Illumina HumanMethylationEPIC BeadChip array platform following the Illumina Infinium HD methylation protocol (Illumina, # 15019519) at the Spanish Centro Nacional de Genotipado (CEGEN-ISCIII, www.cegen.org). Raw IDAT files were processed using the R/Bioconductor package minfi (version 1.22.1, RRID:SCR_012830), probe signals were normalized using the ssNOOB algorithm with the default parameters (offset = 15, dyeCorr = TRUE and dyeMethod = "single") and probe type distribution was corrected using the BMIQ approach implemented in the R/Bioconductor package ChAMP (version 2.20.1, RRID:SCR_12891). Probes located in chromosomes X and Y, probes overlapping genetic variants (SNP137Common track from UCSC genome browser), cross-reactive and multimapping probes and probes with at least one sample with a detection p value >0.01 were discarded for downstream analyses. M-values were used for differential methylation analyses while B-values were used for graphical visualization and for correlation analyses between DNA methylation and gene expression. Differentially methylated probes (DMPs) were identified using the moderated t test implemented in the R/Bioconductor package limma (v.3.38.3, RRID:SCR_010943). Resulting p values were corrected for multiple testing using the Benjamini-Hochberg method (FDR). An FDR threshold of 0.05, and a minimum absolute difference of 0.25 between mean DNA methylation B-values of DPPA3 overexpressing and control SW1222 cells was used as cut-off to determine significant DMPs.

DMPs were assigned to their corresponding genomic location using the R/Bioconductor packages IlluminaHumanMethylationEPIC.anno.ilm10b4.hg19 (v.0.6.0) and ChIPseeker (v.1.18.0, RRID:SCR_021322). The set of filtered probes (783,556) from the HumanMethylationEPIC platform is represented as the background universe from the experiment. In addition, for each gene, the average methylation value of all probes assigned to their promoter region was calculated for DNA methylation/gene expression integration purposes.

Subcutaneous tumor xenografts

$1x10^6$ DPPA3 and control HT29 or SW1222 cells, DPPA3 and control shC/sh*HIF1A* SW1222 cells, and sh*DPPA3* and shCTRL SW1222 cells were resuspended in 50 μL of PBS, mixed with 50 μL of Corning Matrigel Basement Membrane Matrix (Cultek; #45354277) and injected subcutaneously into both mouse flanks. When tumors were detectable (within two weeks after the injection), mice were treated with 2 mg/mL of doxycycline (DOX; Sigma-Aldrich; #D9891) *ad libitum* in drinking water containing 5% sucrose (Sigma-Aldrich; #84100) to induce DPPA3 expression until experimental endpoint. In the chemoresistance experiments, chemotherapy treatment together with DOX administration started when tumors were detectable. Experimental groups were treated with either vehicle or 40 mpk 5-fluorouracil (5FU; Sigma-Aldrich; #F6627) once per week by intraperitoneal injection. When matching endpoint criteria, mice were euthanized, and parts of the xenograft tumors were snap-frozen for RNA extraction or fixed in paraformaldehyde for histological analysis. Tumor 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 tumor diameter and width represents the perpendicular tumor diameter.

Methodology used to generate cell lines (lentiviral transduction and antibiotic selection) and different Matrigel batches may result in slightly differences in tumor growth rates between different experiments.

Orthotopic tumor xenografts

Female NOD-SCID mice were anesthetized with isoflurane, exteriorizing their cecum by a laparotomy.[49] $1x10^6$ DPPA3 and control HT29 cells or $2x10^6$ DPPA3 and control SW1222 cells were suspended in 50 μL PBS and placed into the syringe. We slowly injected the cell suspension, under a binocular lens and its 30G needle introduced 5 mm into the cecal wall. Then, we applied a slight pressure with a cotton stick. We pulled the needle out and cleaned the area around the injection with saline solution. After injection, the gut was returned to the abdominal cavity and closed with surgical suture.

To induce DPPA3 expression, 5 days after injection, mice were treated with 2 mg/mL of DOX (Sigma-Aldrich; #D9891) *ad libitum* in drinking water containing 5% sucrose (Sigma-Aldrich; #84100) until experimental endpoint. When matching endpoint criteria, mice were euthanized and the cecum containing the primary tumor was extracted, weighed, and fixed for histological analysis. Lungs were also extracted and fixed for histological analyses.

Image analyses

All representative fluorescence images shown in figures were chosen randomly from sets of images. For fluorescent immunocytochemistry staining, image quantification was performed using QuPath open-source software (University of Edinburgh; https://qupath.github.io/) Between five and six standard confocal images were taken for each coverslip. For CCNB1, DPPA3 and γH2A.X quantification, cell nuclei were selected and the mean intensity of CCNB1 and DPPA3 nuclear levels were analyzed. Then, an average value was calculated from all nuclei analyzed in each image and represented as column scatterplots. For fluorescent immunohistochemistry staining, image quantification was done using FIJI ImageJ open-source software[41] (https://imagej.net/software/fiji/downloads; RRID: SCR_002285). Between six to ten standard confocal images were taken for each paraffin-section. For CA9 quantification, GFP-positive (SCCC) and GFP-negative (RCCC) areas were selected as independent ROI (Region-Of-Interest). For each image, red channel mean intensity (CA9) was measured for each selected ROI (SCCC and RCCC) and represented as column scatterplots.

Immunohistochemical images were obtained using the NDP.view 2 software (Hamamatsu) from whole digital slide scanner acquired with NanoZoomer 2.0-HT (Hamamatsu Photonics K.K.; RRID: SCR_021658). Image quantification in chromogenic immunohistochemical analyses was performed using the QuPath open-source software (University of Edinburgh; https://qupath.github.io/; RRID: SCR_018257). The percentage of tumor cells positive for CA9, DPPA3, Cleaved-CASP3, KI67, and CCNB1 staining of each tumoral section was used for quantification. Same DAB intensity detection threshold was used to quantify all conditions within a single experiment. To compare the percentage of DPPA3-positive cells between the different tumor models of this study (Figure S3B), a random selection of cell line xenografts and patient tumors was quantified using the same DAB intensity detection threshold. For CK20 quantification in orthotopic tumor xenograft models, CK20-positive tumor cells foci were counted in each lung section. Foci containing $\leq$4 tumor cells were considered small, those having 5–30 tumor cells were considered medium whereas foci with >30 tumor cells were considered large. To calculate the metastasis index, the number of metastatic foci per lung was divided by weight of the primary tumor. Hypoxia response index was calculated by dividing the percentage of CA9-positive cells by the tumor size. Western blot quantifications were performed using FIJI ImageJ version 1.53q[41] (https://imagej.net/software/fiji/downloads; RRID: SCR_002285). The signal intensity for each protein band of interest analyzed was normalized to its corresponding loading control. Fold change was calculated related to the corresponding control cell line or vehicle condition.

QUANTIFICATION AND STATISTICAL ANALYSIS

Only data that we were able to replicate in at least two independent experiments were included in the manuscript. For *in vivo* experiments, animals were excluded only if they died or has killed according to protocols approved by the animal experimental committees. No statistical method was used to predetermine sample size. Data collection and analyses were not performed blinded to the conditions of the experiments. Animals were randomized and treated as indicated in each experiment. Image from all xenografts

were quantified in each experiment and each condition. For *in vitro* experiments, all samples were analyzed equally with no subsampling. Data distribution was assumed to be normal but not formally tested.

The statistical significance of different groups of values obtained in several experiments across the work is described in each corresponding figure legend. In summary, statistical analyses were performed in GraphPad Prism 8 (GraphPad Software Inc.; RRID: SCR_002798) using unpaired or paired 2-tailed Student's *t* test to compare differences between 2 groups and 1- or two-way ANOVA with multiple-comparisons test to compare 3 or more groups with one or two independent variables. Contingency tables were statistically evaluated using Chi-squared or Fischer's exact test. Statistical tests used for the analyses of transcriptomes (microarrays) were performed in Partek Genomic Suite software (Partek Inc.). The Cox proportional hazards model was used for analyzing significance in progression- and relapse-free survivals. Codes with asterisks were used to indicate different levels of statistical significance: $^{*}p \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $^{****}p \leq 0.0001$.