Collagen prolyl hydroxylation-dependent metabolic perturbation governs epigenetic remodeling and mesenchymal transition in pluripotent and cancer cells

Cristina D'Aniello[1,4], Federica Cermola[1,4], Andrea Palamidessi[2], Luca G. Wanderlingh[3,†], Miriam Gagliardi[4], Agnese Migliaccio[1], Francesca Varrone[3], Laura Casalino[1,4], Maria R. Matarazzo[4], Dario De Cesare[1,4], Giorgio Scita[2,5], Eduardo J. Patriarca[1,4]*, Gabriella Minchiotti[1,4]*

[1] Stem Cell Fate Laboratory, Institute of Genetics and Biophysics, 'A. Buzzati-Traverso', CNR, 80131 Naples, Italy.

[2] IFOM, the FIRC Institute of Molecular Oncology, 20139 Milan, Italy

[3] IRBM, 80131 Naples, Italy.

[4] Institute of Genetics and Biophysics, 'A. Buzzati-Traverso', CNR, 80131 Naples, Italy.

[5] University of Milan, Department of Oncology and Hemato-Oncology.

†Current address: Telethon Institute of Genetics and Medicine (TIGEM), Pozzuoli, Naples, Italy.

*Correspondence to:

Gabriella Minchiotti: Via Pietro Castellino 111, 80131 Naples, Italy; +390816132357; gabriella.minchiotti@igb.cnr.it

Eduardo J. Patriarca: Via Pietro Castellino 111, 80131 Naples, Italy; +390816132431; eduardo.patriarca@igb.cnr.it

Running title: Metabolic control of cell state transition dynamics

Conflict of Interest: "The authors declare no potential conflicts of interest"

Abstract

Collagen prolyl hydroxylation (CPH), which is catalyzed by prolyl 4-hydroxylase (P4H), is the most prevalent posttranslational modification in humans and requires Vitamin C (VitC). Here we demonstrate that CPH acts as an epigenetic modulator of cell plasticity. Increased CPH induced global DNA/histone methylation in pluripotent stem and tumor cells and promoted cell state transition (CST). Interfering with CPH by either genetic ablation of P4H subunit alpha-2 (P4HA2) or pharmacologic treatment reverted epigenetic changes and antagonized CST. Mechanistically, we suggest that CPH modifies the epigenetic landscape by reducing VitC for DNA and histone demethylases. Repurposed drugs targeting CPH-mediated metabolic perturbation, such as the antiasthmatic Budesonide, blocked metastatic dissemination of breast cancer cells in vivo by preventing mesenchymal transition. Our study provides mechanistic insights into how metabolic cues and epigenetic factors integrate to control cell state transition and paves the way for the development of novel antimetastatic strategies

Significance: A phenotype-based HTS reveals unforeseen metabolic control of cell plasticity and identifies Budesonide as a drug candidate for metastatic cancer

Keywords: P4HA2, Vitamin C, Fe^{+2}/αKG -dependent dioxygenases, DNA methylation, pluripotent stem cells, cancer cells, cell state transition, metastasis, drugs repurposing, Budesonide

Introduction

Cell plasticity is the ability of cells to switch between different phenotypic states in a reversible manner. Besides its crucial role in the development of multicellular organisms and tissue homeostasis, cell plasticity contributes to the progression of devastating pathologies, including metastatic cancer and organ fibrosis (1-3). The primary source of cell plasticity is the Epithelial-Mesenchymal Transition (EMT). EMT is not an all-or-none, irreversible process, as cohesive epithelial cells and free motile/migrating mesenchymal cells represent only the extreme of a continuum of intermediate/metastable phenotypes. How cell state transition is controlled and what are the molecular determinants that enable cells to reside stably in one or another phenotypic state is a major unresolved issue (2, 4). Transient metabolic perturbations that modify the availability of epigenetic enzymes required metabolites (EEMs) are emerging as candidate determinants of cell plasticity (5). Typical examples are the ascorbic acid [Vitamin C (VitC)] and the alpha-ketoglutarate (αKG) that are cofactors/enhancers for the Fe^{+2}/αKG -dependent dioxygenases superfamily (6), which include the Ten-Eleven Translocation (TET) DNA demethylases and the Jumonji (JMJ) domain–containing histone demethylases. VitC, for instance, by enhancing the activity of the TET and JMJ demethylases, improves the efficiency of induced pluripotent stem cell (iPSC) generation (7-10), drives embryonic stem cells (ESCs) towards a naive state of pluripotency (11, 12) and regulates the balance between self-renewal and differentiation of hematopoietic stem cells (13, 14). Thus, fluctuation of EEMs' concentrations is considered an additional layer of epigenetic regulation (15). Yet, how this is regulated is still unknown. Here, we reveal that VitC -consuming collagen prolyl hydroxylases (P4Hs), which catalyze the hydroxylation of prolyl residues on nascent collagens, are key players in this complex regulatory scenario. Specifically, through an unbiased phenotype-based screening of

1.200 FDA-approved drugs, we identify compounds that are used to gain insight into unexplored molecular mechanisms of cell plasticity, and as drugs to target diseases associated with cell state transition, most and foremost cancer progression and metastatic dissemination.

Materials and Methods

Cell lines, culture conditions and treatments

SUM159 cells were from Asterand (2015), and A549 cells were from NIH NCI-60 cell line panel (2011) and their identity and mycoplasma free-state is routinely (annually) tested by DNA fingerprinting using Short tandem repeat (STR) profiling, and by PCR-based, MycoAlertTM Mycoplasma Detection Kit (Lonza, Bioscience). Each cell lines were used within passage 4/5 since their thawing from originally frozen vials.

A549 cells were cultured in high glucose DMEM (Invitrogen, Life Technologies), 10% FBS (Euroclone), 2 mM glutamine and penicillin/streptomycin (100 U/ml, GIBCO). SUM-159 cells were cultured in F12 medium with inactivated calf serum (5%), 10 mM Hepes, 1 ug/ml hydrocortisone (HYC), insulin (5ug/ml), and penicillin/streptomycin. For quantification of collagen and Collagen-Hydroxyproline (C-HyP), and histone methylation, WT SUM159 cells were starved of HYC for 48hrs and then treated with HYC ± Budesonide (25μM) for 48 hours.

Wild-type TBV2 (129/SvP) were kindly provided by Dr. Stefania Filosa in 2008, $P4ha2^{KO}$ and $P4ha2^{KO}/P4ha1^{KD}$ TBV2 mouse ESCs were maintained as reported (16); mycoplasma free-state is routinely (twice/year) tested by PCR-based assay. All the experiments were performed between the 5th and the 20th cell passage.

L-Pro-induced esMT was performed as described (16). Briefly, ESCs (250 cells/cm^2) were plated onto gelatin-coated plates in ESC medium (DMEM/10%FBS/LIF) ± L-Pro (500 μM - 1 mM) for 5 days. Medium was renewed at day 3 with addition of fresh L-Pro.

PiC to ESC reverse transition (MesT) was performed as described (16). VitC (Sigma-Aldrich) was dissolved in water. CHIR99021 was dissolved in DMSO and used at 3 μM (Millipore).

EsMT inhibitors (Sigma-Aldrich) were dissolved in DMSO at 10 mM and used at concentration ranging from 5 to 20 μM.

The Alkaline phosphatase (AP) assay was performed using AP staining kit (System Biosciences) following manufacturer's instructions.

For Collagen and C-HyP quantification, ESCs ($15x10^3$/cm2) or PiCs ($15x10^3$/cm2) were plated in gelatin and treated ± L-Pro and ± Budesonide and VitC. After 48hrs in culture, cells were collected and analyzed either by WB or immunofluorescence (see below for details).

High-throughput screening

The Prestwick Chemical Library® 1200 compounds (ww.prestwickchemical.com/libraries-screening-lib-pcl.html) was used for the screening. The screening was performed using the *Cell*^{maker} robotic platform (16). TBV2 ESCs were plated at 250 cells/cm2 in gelatin-coated plates (96-multiwell) in DMEM/10%FBS/LIF. Five hours after plating, cells were supplemented with L-Pro (250 μM) or left untreated as control, and the compounds (Prestwick chemical library®) were singularly added at a final concentration of 10 μM. At day 4 after treatment, the resulted colonies were fixed/stained in PBS1x/6% glutaraldehyde/0.15% crystal violet for 30 min at RT. The plates were scanned/imaged and the cell colonies (if any) were analyzed using the ImageJ 1.46r software (http://image.nih.gov/ij).

Chemical structure information for Budesonide are available at [Pubchem (https://pubchem.ncbi.nlm.nih.gov/compound/5281004),DrugBank(https://www.drugbank.ca/drugs/DB01222) and Therapeutic Target Database (https://db.idrblab.org/ttd/drug/d0y7iu).

RNA extraction and quantitative RT-PCR

Total RNAs were extracted using the RNeasy kit (Quiagen) and reverse transcribed with the QuantiTect Reverse Transcription kit (Qiagen). qPCR was performed using SYBR Green PCR master mix (FluoCycle IITM SYBR, EuroClone). Primers are listed in Table S1.

Cell proliferation assays

Cell proliferation was evaluated using FACS- based EdU incorporation (Click-iT® EdU Flow Cytometry Assay kit; Molecular Probes), and the CCK-8 (Dojindo Laboratories, Kumamoto, Japan) assays following manufacturer instructions.

Western blot

Total proteins were extracted in 100 mM Tris pH 8, 140 mM NaCl, 20 mM EDTA, 0,2% SDS, 1% Nonidet P-40 lysis buffer, resolved on SDS-PAGE gels and transferred onto PVDF membranes (iBlot dry Transfer System; Life Technologies).

Histones proteins were prepared as follows: cell pellets were resuspended in Triton Extraction Buffer (TEB: PBS containing 0.5% Triton X 100 (v/v), 2mM phenylmethylsulfonylfluoride (PMSF), 0.02% (w/v) NaN3) at a cell density of 10^7 cells per ml. Then, the pellets were lysed overnight in 0,2N HCl at a density of $4x10^7$ cells/ml.

Primary Antibodies (listed in Table S2) were used overnight at 4°C followed by the appropriate HRP-conjugated secondary antibodies. Detection was performed with ECL reagents (Pierce, Thermo Scientific). ImageJ software was used for the densitometric analysis.

Immunofluorescence analysis and quantification

Cells were fixed (4% PFA) and permeabilized (0.1% Triton X-100) for 10 minutes at RT and incubated with blocking solution (0.1% Triton X-100/5% BSA) for 1 hrs. Primary Antibodies (listed in Table S2) were incubated overnight (4°C) followed by the relative secondary antibodies (Alexa Fluor Molecular Probes)

Images were obtained using the DMI6000B microscope equipped with DFC 350FX B/W digital camera (Leica Microsystems). Confocal images were obtained on a Nikon A1 microscope. The AF6000 (Leica Microsystems) and NIS Element C (Nikon, Tokyo) software were used for image acquisition/elaboration.

For cytospin samples, dissociated cells were resuspended in 15% FBS/1x PBS and centrifuged at 800 rpm for 8 min using a Thermo Shandon Cytocentrifuge (CytoSpinTM 4), and fixed for further analysis.

C-HyP positive area (pixel/cm^2) was measured over the total nuclei area (100-300 cells/conditions) using ImageJ software (https://imagej.nih.gov/ij/docs/guide/146-30.html#toc-Subsection-30.2)

Generation of *P4ha2*^{KO} and *P4ha2*^{KO}/*P4ha1*^{KD} ESCs and tumor cells

For CRISPR/Cas9 mediated *P4ha2* knock-out, TBV2 ESCs were infected with lentiviral particles (1 MOI) carrying the guide RNA (gRNA), which targets *P4ha2* exon 6 (ATCCGGACACGATTTCCAGA) followed by the Cas9 and the Puromycin resistance [transEDIT™ Lentiviral gRNA plus Cas9 (pCLIP-All) Target Gene; transOMIC technologies]. Following Puromycin selection (1 µg/ml; 6 days), resistant cells were dissociated and single cells were FACS-sorted (FACS Aria, Becton Dickinson) on the basis of FSC/SSC parameters and

seeded into 96-well plates. Genomic DNA was analyzed by PCR using primers designed on intron 5-6 (forward) and 6-7 (reverse) sequences that amplify a product of 735 bp. IN/DEL mutations were identified by digesting PCR products with the T7 endonuclease and confirmed by sequencing. The absence of the wild type allele was confirmed by RT-PCR (see Table S1).

P4ha2^{KO}/*P4ha1*^{KD} ESCs were obtained by transfecting *P4ha2*^{KO} ESCs (Clone #1) with short hairpin RNAs (shRNAs) plasmids targeting different non-overlapping *P4ha1* mRNA sequences followed by a GFP reporter (shERWOOD UltramiR Lentiviral shRNA, pZIP-mEF1a, TransOMIC). 48hrs after electroporation single GFP positive cells were FACS sorted into 96-well plates, and analyzed as above.

Stable *P4HA2*^{KD} A549 and SUM159 cells were generated by lentiviral infection of shRNAs (PLKO, SIGMA) followed by Puromycin selection (1 μg/ml; 10 days).

Quantification of 5hmC

Gelatin plated ESCs (15x10^3/cm2) were pre-treated with VitC (50 μM; 24 hrs) before addition of L-Pro ± Budesonide for 24 hrs. For ELISA- based quantification, genomic DNA was extracted (Wizard^R Genomic DNA Purification kit; Promega) and 5hmC levels were measured using MethylFlashTM Global DNA Hydroxymethylation (5hmC) ELISA Easy Kit (Epigentek) and normalized to DNA content, following the manufacturer instructions.

For immunofluorescence, cells were fixed (4% PFA) and permeabilized (0.4% Triton X-100; 15min at RT). After a denaturation step (2N HCl, 15min at RT) and a neutralization step (100 mM Tris-HCl pH 8.5, 10min at RT), cells were incubated with 0.1% Triton X-100/5% BSA for 1 hr. Anti 5-hmC monoclonal Antibody was incubated overnight at 4°C (1:200, HMC/4D9 A1018, Epigentek) followed by anti-mouse 594 secondary antibody (Alexa Fluor Molecular Probes).

Images were obtained using the DM6000B microscope equipped with DFC 350FX B/W digital camera (Leica Microsystems). The AF6000 (Leica Microsystems) software was used for image acquisition/elaboration.

DNA methylation data analysis

DNA methylation comparison between L-Pro treated- and *Tet* TKO mESCs was performed as described (17). Briefly, the *Tet* TKO hyper-DMRs (18) in which at least 3 CpGs were covered by L-Pro RRBS experiments (6) were selected for the analysis. For each region, the ratio between the number of hypermethylated CpGs ($\Delta_{(\text{L-Pro } vs \text{ VitC})} \geq 20\%$) and total CpGs was calculated. DMRs with a ratio greater than 0.3 were defined hypermethylated. The statistical significance of enrichment was determined using the one tail hypergeometric test available in R.

Time-Lapse

Images from ESCs ± L-Pro (500 μM) ± esMTi and A459 tumor spheroids ± Budesonide were captured (20x) at the indicated time point, every 5 min for about 24 hrs on a Leica DMI6000B equipped with a microscope incubator (Okolab).

2D migration assays

Boyden chamber assay was performed using polycarbonate transwells (8 μm pore, Costar). Briefly, A549 cells pre-treated ± Budesonide or DMSO in complete medium. After 24 hrs, cells were dissociated by trypsin-EDTA, seeded ($4*10^5$ cells/per insert) in DMEM/1% FBS ± Budesonide and allowed to migrate towards FBS gradient (1%-15%) for 6 hrs. Nuclei were stained and migrated cells were counted (4 random fields/insert, 10x).

For A549 tumor spheroid assay, 2.5 x10^3 FACS sorted cells were plated/well in V-shaped low attachment 96-multiwell (Corning Costar) and allowed to aggregate for 4 days. Spheroids were plated on gelatin-coated wells in complete medium without growth factors. After 48 hrs, tumor spheroids were fixed and stained (PBS1x/6% glutaraldehyde/0.15% crystal violet) for 30min at RT. The percentage of spreading spheroids was determined by quantifying the number of spheroids with a crown of mesenchymal/migrating cells over the total number of spheroids plated.

3D Organotypic culture

For the analysis of SUM-159-derived spheroids invasive outgrowth, single SUM-159 cells were plated in 8-well chamber slides (IBIDI) onto a base layer of Matrigel and collagen I (5 mg/ml and 2.3 mg/ml, respectively; BD) and supplemented with a 2% Matrigel/growth medium mixture. Cultures were grown for 6 days. Medium was replenished every 2 days. SUM-159 spheroids were imaged for development of invasive outgrowths by differential interference contrast (DIC) imaging using a 20X objective lens. Colonies circularity and area parameters were calculated using ImageJ software.

Xenograft experiments

All animal experiments were approved by the OPBA (Organisms for the well-being of the animal) of IFOM and Cogentech. All experiments complied with national guidelines and legislation for animal experimentation. All mice were bred and maintained under specific pathogen-free conditions in our animal facilities at Cogentech Consortium at the FIRC Institute of Molecular Oncology Foundation and at the European Institute of Oncology in Milan, under

the authorization from the Italian Ministry of Health (Autorizzazione N° 604-2016). SUM-159 cells, expressing pLenti CMV-Puro-LUC construct (Plasmid #17477, ADDGENE), were trypsin detached, washed twice, and resuspended in PBS to a final concentration of $2.5*10^6$ cells/13 µl. The cell suspension was then mixed with 5 µl growth factor–reduced Matrigel (BD) and 2 µl Trypan blue solution and maintained on ice until injection. Aseptic conditions under a laminar flow hood were used throughout the surgical procedure. Female NOD.Cg-PrkdcscidIl2rgtm1Wjl/SzJ (commonly known as the NOD SCID gamma; NSG) mice (6–9 week old) were anesthetized with 375 mg/Kg Avertin, laid on their backs, and injected with a 20 µl cell suspension directly in the fourth mammary fad pad. After 4 days, mice were administered by intraperitoneal injection (I.P.) either Budesonide (3 mg/Kg) or the same volume of carrier as control, every day for 3 consecutive weeks. For bioluminescence imaging, 100 µl of D-Luciferin (30 mg/ml, Perkin Elmer) was injected I.P. before placing mice under 2% inhaled isofluorane anesthesia. Bioluminescence signal was monitored every 5 days by measuring photon flux (15 min from I.P. injection) using IVIS LUMINA III imaging System (Perkin Elmer). Data were analyzed using average counts in the ROIs and normalized to acquisition time. At the endpoint, mice were sacrificed and tumors were harvested. Tumor volume was measured with digital caliper and calculated according to the formula: $L \times W^2/2 = mm^3$. For histological analysis, primary tumors and lung metastases were fixed in 4% phosphate-buffered formalin and paraffin-embedded. The Scan Scope XT device and the Aperio Digital pathology system software (Aperio) were used to collect pictures and detect metastases.

Immunohistochemistry

Formalin-fixed, paraffin-embedded (FFPE) mouse tumor tissue and lungs sections (3 μm) were incubated with 3% H2O2 for 5 min and then blocked with 2% goat serum in PBS for 1 hr. Samples were incubated with primary antibodies (listed in Table S2) for 2 hrs at RT in 2% goat serum. HRP-conjugated secondary antibodies were used and signals were developed with DAB. Samples were counterstained with hematoxylin (Merck) or stained with Picrosirius Red (Polysciences Inc.). Cell circularity (stained with anti Vimentin) and collagen fiber distribution (observed in polarized light microscopy images) parameters were calculated using ImageJ software.

Bioinformatic analysis of PH4A2 in cancer cohorts

Survival analysis was performed exploiting the METABRIC breast cancer dataset (19), downloaded from the TCGA database. Expression data were extracted from the dataset (1980 Breast Cancer samples) and three Illumina probesets were available for P4HA2; ILMN_1795778, ILMN_2381697, ILMN_2280135. mRNA expression data from the three probesets for each sample were averaged and samples were grouped into Low and High P4HA2 expression based on the mRNA levels with respect to the mean. Data analysis was performed using JMP 10.0 statistical software (SAS Institute, Inc);

Statistical analysis

Statistical significance was determined by a two-tailed paired Student's *t*-test. *p*-values ≤ 0.05 were considered as statistically significant. Error bars represent SEM or SD as indicated.

Results

High-throughput screen identifies inhibitors of embryonic stem to mesenchymal-like transition (esMT)

We exploited the highly dynamic embryonic stem to mesenchymal-like transition (esMT) as a molecular and functional paradigm for reversible cell state transition (20, 21). To search for compounds able to block esMT we developed a fully automated unbiased high-throughput screening *(*HTS) in a 96-well format based on a robust cell colony phenotype assay (CFA) (**Fig. 1A and B**). The CFA was optimized to achieve a high colony forming efficiency (72±9 colonies/well) and high efficiency of domed-to-flat transition (**Fig. 1A**) after L-Pro supplementation (85±7% flat *vs* 12±6% domed colonies) (**Fig. 1C**; Supplementary Fig. S1A). The feasibility of the assay was verified using two potent esMT antagonists, *i.e.* L-Ascorbic acid (Vitamin C, VitC) and the GSK3 inhibitor CHIR99021 (16, 20), which robustly reduced the fraction of flat-shaped colonies to Control levels (~90% domed *vs* ~10% flat colonies) (**Fig. 1C**; Supplementary Fig. S1A). We employed this automated CFA to screen a Prestwick chemical library® containing 1.200 FDA-approved small molecules. ESCs were plated in the presence of L-Pro and exposed to either drugs or solvent (DMSO). At day four after plating, colonies were fixed and stained with a glutaraldehyde/crystal violet mix and subjected to phenotypic analysis (**Fig. 1D**). A substantial number of the drugs uniformly distributed throughout the plates (137, 11.4%) completely inhibited cell colony formation (Table S3). In depth analysis of these ESC inhibitors (ESCi) revealed that: i) the large majority (80%) have anticancer activities (Table S4); ii) ~50% belongs to 5 therapeutic classes of the 135 covered by the library (**Fig. 1E and F**); and iii) their activities rely on specific structural constraints (**Fig. 1G**; Supplementary Fig. S1B and S1C). We next restricted the analysis to the drugs able to inhibit esMT, *i.e.* to significantly

reduce the fraction of L-Pro-induced atypical/flat colonies and increase the fraction of round/domed-shaped colonies. Fourteen esMT inhibitors [(esMTi) 1.2%; Table S5] were selected and validated in secondary CFAs (**Fig. 1H**; Supplementary Fig. S1D). Consistent with their antagonistic effect on esMT, the three most active drugs, Budesonide, Spiramycin and Propafenone (Supplementary Fig. S1E) prevented cell motility and induction of esMT markers (**Fig. 2A**; Supplementary Fig. S2A). Notably, Budesonide, Spiramycin and Propafenone prevented neither L-Pro induced down-regulation of Aminoacid stress response (AAR) pathway markers (**Fig. 2B**) nor cell proliferation (**Fig. 2C**), which are the earliest ESC responses to L-Pro supplementation and depend on L-Pro -tRNA loading (21). Finally, we tested their ability to promote the reversal mesenchymal-like-to-embryonic-stem-cell transition [MesT; (20)]. To this end, L-Pro-induced mesenchymal-like cells (PiCs) were plated at low density and treated with L-Pro ± esMTi or VitC as control (**Fig. 2D**). Cells reverted back to the ESC phenotype in the presence of esMTi, with efficiency comparable to that observed for VitC (**Fig. 2D**). These results indicate that esMTi specifically block esMT acting downstream of L-Pro -tRNA loading, and induce MesT even in a highly rich L-Pro environment.

Collagen accumulation underpins esMT

To obtain mechanistic insights into esMT, we classified the 14 esMTi identified in the screening on the basis of their chemical properties. Despite belonging to 9 different classes (Tanimoto coefficient T>0.75) (Supplementary Fig. S2B), nearly 50% of the esMTi, including Budesonide, are inhibitors of collagen accumulation/fibrosis *in vitro* and/or *in vivo* (Table S6), pointing to a link between collagen accumulation and esMT. Notably, we recently showed that L-Pro -induced esMT depends on protein synthesis and that ESCs are specifically and intrinsically starved of L-

Pro (21). Based on our findings and these considerations, and given that L-Pro and 4-*trans*-hydroxyproline (HyP) are the major components of collagen, we hypothesized that L-Pro supplementation would increase collagen synthesis/hydroxylation in ESCs, eventually inducing esMT. Accordingly, Collagen (Col1a1) and collagen hydroxyproline (C-Hyp) levels both increased in L-Pro treated ESCs, accumulating primarily in the endoplasmic reticulum (ER; **Fig. 2 E and F**; Supplementary Fig. S2C). Conversely, the esMTi Budesonide prevented *Col1a1* expression at RNA and protein levels, as well as C-Hyp in L-Pro-treated cells (**Fig 2G-J**; Supplementary Fig. S2D). Budesonide is a 21-hydroxysteroid glucocorticoid; however, none of the other twenty-four 21-hydroxysteroids derivatives, including the glucocorticoid Dexamethasone, was identified as a positive hit in the primary screening (Supplementary Fig. S2E) and confirmed in the secondary CFA (Supplementary Fig. S2F), thus revealing a highly specific effect of Budesonide. Of note, among the 25 21-hydroxysteroids included in the library, only Budesonide is classified as a high-affinity Glucocorticoid Receptor (GR) antagonist in different databases, including [Pubchem (https://pubchem.ncbi.nlm.nih.gov/compound/5281004), DrugBank (https://www.drugbank.ca/drugs/DB01222) and Therapeutic Target Database (https://db.idrblab.org/ttd/drug/d0y7iu). This raised the possibility that Budesonide specificity might be due, at least in part, to its activity as GR antagonist. Interestingly, addition of another established antagonist of GR, Mifepristone, was able to block esMT similarly to Budesonide (Supplementary Fig. S2F). Furthermore, Budesonide but not Dexamethasone reduced the levels of the well known GR target Alkaline Phosphatase (22), while keeping ESC undifferentiated (Supplementary Fig. S2G). Finally, Budesonide fully counteracted Dexamethasone- induced overexpression of the GR target genes *Fkbp5* and *Tfcp2l1* (Supplementary Fig. S2H).

Glucocorticoids, including Dexamethasone, are commonly used to induce collagens in cultured cells. Of note, we found that Budesonide but not Dexamethasone fully counteracted L-Pro -dependent increase of Col1a and C-HyP (**Fig. 2G-J**; Supplementary Figure S2I).

Together these findings supported the hypothesis that L-Pro -induced esMT depends on L-Pro -mediated induction of collagen synthesis/hydroxylation, and indicated that Budesonide counteracts this effect (**Fig. 2K**).

Genetic and pharmacological inhibition of collagen hydroxylation prevents esMT

To directly test our hypothesis, and given that it is not feasible to knock-out all the members of the collagen family, we decided to interfere with collagen hydroxylation by targeting the Prolyl-4-hydroxylases (P4Hs). P4H enzymes belong to the family of VitC/Fe^{+2}/αKG-dependent dioxygenases, and form tetrameric (α2:β2) complexes able to catalyze the hydroxylation of L-Pro residues in nascent collagens in the ER. Two out of the three genes coding for the alpha subunit of P4H, *i.e. P4ha1* and *2,* were expressed at similar levels in ESCs, whereas *P4ha3* expression was nearly undetectable (Supplementary Fig. S3A). Of note, *P4HA2* has been recently associated with collagen production and tumor progression of human breast cancer (23). We thus inactivated *P4ha2* in ESCs by CRISPR/Cas9 and selected two independent clones (*P4ha2*^{KO} #1 and #2) that carry different genomic deletions at the boundary between exon 6 and intron 7, leading to a frame-shift and the formation of a premature stop codon (**Fig. 3A**; Supplementary Fig. S3B). Both *P4ha2*^{KO} ESCs preserved the typical domed morphology, were highly AP positive (**Fig. 3B**), and retained the expression of key pluripotency markers (Supplementary Fig. S3C and S3D). To avoid the possibility that *P4ha1* may, at least partially, compensate for the lack of *P4ha2* we knocked down *P4ha1* in *P4ha2*^{KO} ESCs using two shRNAs, which target non-

overlapping *P4ha1* mRNA sequences (**Fig. 3C**; Supplementary Fig. S3E). *P4ha2*^{KO}/*P4ha1*^{KD} ESCs retained AP activity and expression of pluripotency genes (**Fig. 3B**; Supplementary Fig. S3C and S3D). We then examined the effect of *P4ha2*^{KO} and *P4ha2*^{KO}/*P4ha1*^{KD} on L-Pro-induced esMT. Control ESCs (NT) retained high frequency of atypical colonies in response to L-Pro, while *P4ha2*^{KO} ESCs mostly developed typical domed colonies lacking the crown of mesenchymal-like cells scattered around the colony core (**Fig. 3D**). Remarkably, this phenotype was robustly enhanced in two independent *P4ha2*^{KO}/*P4ha1*^{KD} clones (**Fig. 3D**), indicating that *P4ha1* partially compensates the lack of *P4ha2*. Consistent with the inhibitory effect on L-Pro-induced phenotypic transition (domed-flat), cell motility was also impaired in P4H deficient ESCs (Supplementary Fig. S3F). Conversely, L-Pro-induced proliferation (**Fig. 3E**) or alleviation of AAR pathway (Supplementary Fig. S3G) was preserved, providing evidence that P4H activity is specifically required for L-Pro- induced esMT. Finally, as expected, C-Hyp accumulation was strongly reduced in P4H deficient ESCs (**Fig. 3F**). Together, these findings indicate that *P4ha* genetic ablation prevents collagen hydroxylation and antagonizes esMT mimicking the phenotypic and molecular effects of Budesonide, thus supporting the concept that esMT is functionally linked to collagen hydroxylation.

Inhibition of collagen hydroxylation enhances VitC-dependent DNA and histone hydroxylation/demethylation in pluripotent stem cells

Collagen hydroxylation requires VitC and our findings indicate that this crucial post-translational modification of collagen underpins L-Pro -induced esMT. However, exogenously added VitC antagonizes rather than induces esMT (Comes et al., 2013; **Fig. 1C**; Supplementary Fig. S1A) without blocking C-HyP accumulation (**Fig. 2I and J**). This apparent contradiction could be

reconciled by assuming the existence of a VitC-dependent mechanism that overcomes the agonistic effect of collagen hydroxylation on esMT (**Fig. 2K**). Of note, besides P4H, VitC is essential for the activity of different enzymes belonging to the family of Fe^{+2}/αKG-dependent dioxygenases with different substrate specificity, including DNA (TET) and Histones (JMJ), subcellular localization (cytoplasm, ER and nucleus), expression levels and kinetics parameters. Hence, we hypothesized that a sudden increase of collagen hydroxylation might channel VitC to the ER for the activity of collagen hydroxylases (P4Hs), concomitantly provoking its stoichiometric reduction in other cellular compartment, which translates into reduced activity of VitC-dependent dioxygenases, ultimately inducing esMT. Of note, we expect that such a substantial perturbation of VitC homeostasis might occur only in a VitC-limiting environment when collagen synthesis is induced.

To test our hypothesis we asked whether a compartmentalized VitC-dependent reaction, i.e collagen hydroxylation in the ER, might impact on the activity of VitC/Fe^{+2}/αKG-dependent TET and JMJ dioxygenases and thus on DNA and histone methylation levels. We first compared the DNA methylation profiles of L-Pro- treated ESCs (6) and the *Tet* triple KO (TKO) ESCs (18) and found that 54% of the TET target regions were hypermethylated in L-Pro-treated cells at early time point (day 2), which increased up to 89% at day 5 (**Fig. 3G**). Prompted by these results, we evaluated the impact of *P4h* KO on the global levels of DNA hydroxymethylation (5hmC), which depends on the activity of TET enzymes. In a high-L-Pro/low-VitC condition, 5hmC levels rapidly decreased in control but not in P4Hs deficient ESCs (**Fig. 3 H and I**; Supplementary Fig. S3H). Similarly, Budesonide counteracted the reduction of 5hmC in wild-type ESCs (Supplementary Fig. S3I). We then evaluated the impact of *P4h* KO on JMJ dioxygenases, which target removal of H3K9/36 methylation and assessed whether VitC

supplementation blunted the effect of L-Pro. In a high-L-Pro/low-VitC environment the global levels of both H3K9me3 and H3K36me3 increased in Control but not in *P4ha2*^{KO} and *P4ha2*^{KO}/*P4ha1*^{KD} ESCs (**Fig. 3J**; Supplementary Fig. S3J). Notably, H3K9me3 levels were even lower in the *P4ha2*^{KO}/*P4ha1*^{KD} double mutant compared to *P4ha2*^{KO} (**Fig. 3J**), consistent with the idea that VitC availability may progressively increase in the nucleus as C-Hyp levels decrease in the ER (**Fig. 3F**). In keeping with the idea that VitC is a limiting metabolite in ESCs (12), its addition not only prevented L-Pro- dependent increase of H3K9me3 levels but also further reduced H3K9me3 global levels in a low L-Pro environment (**Fig. 3K**). Furthermore, Budesonide mimicked the antagonistic effect of VitC, preventing both L-Pro -dependent esMT (**Fig. 1 C and H**; **Fig. 2A**; Supplementary Fig. S1A) and increase of H3K9me3 levels (Supplementary Fig. S3K). Of note, neither L-Pro, nor Budesonide modified the expression of the specific DNA/histone demethylases, including *Tet1*, *2*, *3* and *Jmjd2c* (H3K9me2/me3 and H3K36me2/me3 demethylation; Supplementary Fig. S3L) ruling out the possibility the observed epigenetic changes reflects altered expression of the enzymes.

Collectively, these findings indicate that increased collagen hydroxylation causes DNA and histones methylation, and suggest that this occurs by altering VitC homeostasis and reducing the activity of nuclear dioxygenases (**Fig. 3L**). Budesonide and P4H deficiency interferes with this metabolic perturbation by reducing collagen hydroxylation with different mechanisms. Specifically, Budesonide blocks collagen synthesis at RNA and protein levels, and as a consequence reduces collagen hydroxylation; whereas, P4H deficiency directly interferes with hydroxylation of nascent collagens within the ER.

Collagen hydroxylation influences the epigenetic landscape of tumor cells and induces cell state transition

The above results raise the possibility that this mechanism might be general applicable to mesenchymal transition processes and critically important in pathological situations, such as cancer progression and metastatic dissemination. Consistent with the inhibitory effects on L-Pro - induced esMT in ESCs, Budesonide robustly inhibited human lung cancer A549 cell migration *in vitro*, in tumor spheroid and Boyden chambers assays (**Fig. 4 A** and **B**; Supplementary Fig. S4A-C), without affecting cell proliferation (Supplementary Fig. S4D). Furthermore, similar to what observed in pluripotent stem cells, Budesonide reduced collagen accumulation in A549 at RNA and protein level (**Fig. 4C**; Supplementary Fig. S4E). We thus went on and assessed the effect of Budesonide on histone methylation, by analyzing the global levels of different methylation marks including H3K9me2/me3, H3K36me3, H3K27me3 and H3K4me3, which are the known substrates of the dioxygenases involved in histone demethylation. Remarkably, Budesonide significantly reduced global histone methylation in A549 cells (**Fig. 4D**) suggesting that an antagonistic interplay between collagen and histone hydroxylation occurs in cancer cells. To get further mechanistic insights into this phenomenon, we knocked down *P4HA2* in A549 (Supplementary Fig. S4F-G) and confirmed that C-Hyp accumulation was strongly reduced compared to NT control cells (**Fig. 4E**). Consistent with what observed in ESCs, $P4HA2^{KD}$ strongly reduced A549 cell migration (**Fig. 4F**; Supplementary Fig. S4H) but not proliferation (Supplementary Fig. S4I), and significantly reduced histone methylation (**Fig. 4G**).

Prompted by these results and to validate the model in a different tumor type, we focused on breast cancer. Interestingly, elevated expression of P4HA2 has been recently associated with poor prognosis in a variety of breast cancers (24). We first extended and corroborated the latter observations by analyzing P4HA2 expression levels in an independent Breast Cancer dataset, encompassing nearly 2000 patients with complete clinical-pathological follow up. We found that

elevated level of P4HA2 is significantly associated with decrease survival, and, more relevantly, is an independent predictor of disease outcome with respect to a set of standard clinicopathological parameters in multivariate analysis (**Fig. 5A**).

To further explore this mechanism, we investigated the effect of *P4HA2*^{KD} on the triple negative breast cancer (TNBC) cell line SUM159, which are characterized by a high degree of heterogeneity/plasticity and recapitulates the key features of collective invasion typical of human breast cancers (25). Interestingly, we found that interfering with *P4HA2* expression (**Fig. 5B**) converted SUM159 cells to a more epithelial-like state (**Fig. 5 C and D**), without affecting cell proliferation (**Fig. 5E**), and reduces C-Hyp levels (**Fig. 5F**). Of note, *P4HA2*^{KD} reduced the global levels of different methylation marks including H3K27me3, H3K36me3, H3K4me3, H3K9me2/me3 and (**Fig. 5G**). To investigate this phenotype further, we assessed the effect of Budesonide and first confirmed that it reduced Col1a1 and C-HyP accumulation in SUM159 cells (**Fig. 6 A** and **B**). Remarkably, Budesonide treatment concomitantly led to reduction of H3K9me2/3, H3K36me3 and H3K27me3 but not H3K4me3 levels (**Fig. 6C**). Of note, expression of the specific dioxygenases involved in histone demethylation, including *JMJD3* (H3K27me3), *JMJD2A* (H3K9me2/me3 and H3K36me2/me3) and *JARID1A* (H3K4me3) was not affected (Supplementary Fig. S5A), thus further supporting the idea that the global reduction of histone methylation was due to increased activity rather than expression of these enzymes. Complementary to these findings, we showed that VitC supplementation significantly reduced H3K9me2/3, H3K27me3 and H3K4me3 in a dose- and histone mark- dependent manner (**Fig. 6D**), indicating that VitC is a limiting factor for proper function of JMJ demethylases in SUM159 cells, and further supporting the idea that variations in the levels and/or compartmentalization of VitC impacts on the activity of nuclear dioxygenases in cancer cells.

All together these findings indicate that collagen synthesis/hydroxylation underpins cell state transition dynamics, and influences the epigenetic status of normal and tumor cells.

Budesonide blocks metastatic dissemination of TNBC cells by preventing collagen deposition and mesenchymal transition

Prompted by these results and the potential therapeutic relevance of targeting mesenchymal transition as a strategy to prevent tumor cell dissemination, we further investigated the effect of Budesonide on SUM159 cells. First, we tested the impact of Budesonide on SUM159 cell invasion using a 3D organotypic culture (**Fig. 7A**). Treatment with Budesonide robustly reduced the collective invasive ability of SUM159 spheroids and significantly increased colony compaction and circularity in a dose dependent fashion (**Fig. 7B**; Supplementary Fig. S5B), without affecting cell proliferation (Supplementary Fig. S5C). We further corroborated these *in vitro* findings using an experimental metastatic murine model. To this end, we orthotopically injected luciferase-expressing SUM159 cells into the mammary fat pad of immunocompromised mice. Budesonide treatment altered significantly the cell shape and stromal collagenolytic organization of primary tumors, which were composed of cuboidal, epitheliod cells with a reduced and less organized collagen-surrounding matrix (**Fig. 7C and D).** On the contrary, in control animal primary tumors were primarily composed of elongated, spindle shaped cells surrounded by a thick collagenous stromal desmoplastic reaction (**Fig. 7C and D).** Consistently, Budesonide treatment reduced primary tumors volume (Supplementary Fig. S5D and E), but had no effect on the number of KI67 and Caspase positive cells (Supplementary Fig. S5F). Remarkably, instead, it completely abrogated the development of metastatic foci into the lung (**Fig. 7 E** and **F**). Collectively, these findings indicate that Budesonide-mediated interference

with collagen production/deposition hampers mesenchymal transition state, ultimately impairing metastatic dissemination.

Discussion

Collagen hydroxylation and DNA/histone hydroxylation/demethylation are catalyzed by dioxygenases superfamily enzymes, namely the prolyl 4-hydroxylases (P4Hs) in the endoplasmic reticulum (ER), and Ten-Eleven Translocation (TET) and Jumonji (JMJ) demethylases in the nucleus, respectively. Here, we demonstrate that collagen and DNA/histone hydroxylation show an antagonistic interplay. Mechanistically, we propose that this previously unexplored interplay relies on the availability VitC, which is an essential cofactor for the activity of both P4Hs and TET/JMJ. This idea is in line with emerging evidence that reduced availability of dioxygenases substrates/cofactors, such as O_2 and VitC, promotes cell state transition and metastatic dissemination by reducing the activity of TET and JMJ demethylases with a concomitant increase of DNA and histone methylation (26, 27). Here we propose that collagen prolyl hydroxylation is a previously unexplored key component of this mechanism. Several lines of evidence support our model. First, genome wide analyses in embryonic stem cells (ESCs) reveal that a large fraction of TET1-3 target DNA regions are methylated after exposure to a high L-Pro regimen, which results in increased collagen synthesis/hydroxylation. Second, increased collagen synthesis/hydroxylation under VitC limiting conditions rapidly reduces 5hmC and increases H3K9/K36 histone methylation, which is in agreement with a reduced activity of TET and JMJ, respectively. Finally, interfering with collagen synthesis/hydroxylation, by either pharmacological treatment with the glucocorticoid Budesonide (collagen synthesis) or genetic ablation of P4HA2 (collagen hydroxylation), blocks cell state transition in pluripotent stem cells and tumor cells.

Of note, the latter findings are highly relevant for cancer biology as high P4HA2 expression is a poor prognostic marker in breast cancer patients and correlates with the metastatic progression of

triple negative breast cancer (TNBC) cells (24) and our findings). Accordingly, we demonstrate that P4HA2 knockdown reduces global H3K9, H3K36, H3K4 and H3K27 methylation levels and reverts the mesenchymal phenotype of TNBC SUM159 cells to a more epithelial-like phenotype. Of note, the TNBC SUM159 cells are highly heterogeneous and can be sorted into low and high tumor-initiating populations based on the expression of the basal epithelial marker integrin-β4 (ITGB) (28). Interestingly, we noticed that P4HA2 expression is significantly higher in the fully mesenchymal $ITGB^{low}$ subpopulation *vs* the $ITGB4^{high}$ subpopulation that show hybrid epithelial/mesenchymal phenotype (28). Although further studies are necessary, we speculate that this collagen hydroxylation/P4HA2 -dependent remodeling of the epigenetic landscape may underlie, at least in part, the molecular and functional heterogeneity of TNBC cells.

In line with our idea, pharmacological targeting of this collagen-epigenetic axis with Budesonide, identified in this study, inhibits the metastatic dissemination of SUM159 cells by preventing mesenchymal-transition. Interestingly, Budesonide, has been very recently identified in a combined connectivity mapping and pharmacoepidemiological approach as a drug with potential breast cancer-preventing properties (29). Moreover, Budesonide is universally employed to treat asthma, and asthma was recently associated with reduced risk of developing pancreatic ductal adenocarcinoma (PDAC) (30). Furthermore, Budesonide is also used to treat the collagenous colitis, a rare inflammatory disorder associated with collagen accumulation in the colorectal mucosa (31, 32), and its therapeutic efficacy is ascribed to the anti-inflammatory activity of GCs (33). Thus, our unexpected findings that Budesonide blocks mesenchymal transition interfering with collagen accumulation provide novel insights into its therapeutic effects. Of note, the mechanism behind the effects of Budesonide remains to be fully clarified. Indeed, although our findings raised the possibility of potential involvement of the GR, the

mechanism may prove more complex and its complete elucidation will require further investigations. Specifically, we cannot rule out the possibility that off-target effects of Budesonide at the concentration used, including for instance inhibition of P4H activity, could be responsible of the observed phenotypes.

Epigenetic mechanisms may govern the accumulation of collagen by reducing promoter methylation and increasing expression of collagen genes (34, 35). Disruption of this epigenetic->collagen axis is considered as the main cause of the deregulation of collagen deposition during aging and in pathological conditions, including tumorigenesis (36). Here we reveal the existence of a reverse collagen->epigenetic functional axis associated with cell state transition in normal and pathological conditions.

In conclusion, our findings provide evidence of a previously unexplored functional interplay between collagen and DNA/Histone hydroxylation and identify P4HA2 as a key component of this mechanism. Interestingly, it has been recently reported that expression of another member of P4H family, P4ha1, reduces prolyl-hydroxylation of HIF-1α by modulating the levels of ketoglutarate (KG) and succinate in the cytoplasm (37). Of note, hydroxylation of HIF-1α is catalyzed by the Prolyl hydroxylases (PHDs), which are VitC-dependent dioxygenase belonging to the same family of TET and JMJ. The evidences from this paper of a collagen hydroxylation (ER) -> HIF hydroxylation (cytoplasm) axis strongly support our findings and can be well explained with the model proposed herein.

Finally, we identify Budesonide as a potent inhibitor of this mechanism and a novel drug candidate for the treatment of a broad spectrum of diseases in which cell fate transition and fibrosis are involved, including metastatic cancer.

Acknowledgments

We are grateful to members of the Integrated Microscopy and FACS Facilities of IGB-ABT, CNR, and to Gabriele Di Napoli for excellent technical assistance. We are indebted to Stefano Confalonieri at European Institute of Oncology, Milan, for the bioinformatic analysis of PH4A2 in cancer cohorts. We are most grateful to Prof. Francisco X. Real for discussion and helpful suggestions. This study was supported by AIRC (IG 20736), MIUR Epigenomics Flagship Project (EPIGEN) and POR Campania FESR 2014/2020 (Project SATIN) to G.M.

References:

1. Chaffer CL, San Juan BP, Lim E, Weinberg RA. EMT, cell plasticity and metastasis. Cancer Metastasis Rev. 2016;35:645-54.
2. Nieto MA, Huang RY, Jackson RA, Thiery JP. Emt: 2016. Cell. 2016;166:21-45.
3. Varga J, Greten FR. Cell plasticity in epithelial homeostasis and tumorigenesis. Nat Cell Biol. 2017;19:1133-41.
4. Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. Cell. 2011;147:275-92.
5. Reid MA, Dai Z, Locasale JW. The impact of cellular metabolism on chromatin dynamics and epigenetics. Nat Cell Biol. 2017;19:1298-306.
6. D'Aniello C, Cermola F, Patriarca EJ, Minchiotti G. Vitamin C in Stem Cell Biology: Impact on Extracellular Matrix Homeostasis and Epigenetics. Stem Cells Int. 2017;2017:8936156.
7. Esteban MA, Pei D. Vitamin C improves the quality of somatic cell reprogramming. Nature genetics. 2012;44:366-7.
8. Chen J, Guo L, Zhang L, Wu H, Yang J, Liu H, et al. Vitamin C modulates TET1 function during somatic cell reprogramming. Nature genetics. 2013;45:1504-9.
9. Wang T, Chen K, Zeng X, Yang J, Wu Y, Shi X, et al. The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. Cell Stem Cell. 2011;9:575-87.
10. Hore TA, von Meyenn F, Ravichandran M, Bachman M, Ficz G, Oxley D, et al. Retinol and ascorbate drive erasure of epigenetic memory and enhance reprogramming to naive pluripotency by complementary mechanisms. Proc Natl Acad Sci U S A. 2016;113:12202-7.
11. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martinez JA, Goyal P, Mahapatra S, et al. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature. 2013.
12. D'Aniello C, Habibi E, Cermola F, Paris D, Russo F, Fiorenzano A, et al. Vitamin C and l-Proline Antagonistic Effects Capture Alternative States in the Pluripotency Continuum. Stem Cell Reports. 2017;8:1-10.
13. Agathocleous M, Meacham CE, Burgess RJ, Piskounova E, Zhao Z, Crane GM, et al. Ascorbate regulates haematopoietic stem cell function and leukaemogenesis. Nature. 2017;549:476-81.
14. Cimmino L, Dolgalev I, Wang Y, Yoshimi A, Martin GH, Wang J, et al. Restoration of TET2 Function Blocks Aberrant Self-Renewal and Leukemia Progression. Cell. 2017;170:1079-95 e20.
15. Zhao Z, Wang L, Di LJ. Compartmentation of metabolites in regulating epigenome of cancer. Mol Med. 2016;22.
16. Casalino L, Comes S, Lambazzi G, De Stefano B, Filosa S, De Falco S, et al. Control of embryonic stem cell metastability by L-proline catabolism. J Mol Cell Biol. 2011;3:108-22.
17. Gatto S, Gagliardi M, Franzese M, Leppert S, Papa M, Cammisa M, et al. ICF-specific DNMT3B dysfunction interferes with intragenic regulation of mRNA transcription and alternative splicing. Nucleic Acids Res. 2017;45:5739-56.
18. Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y. Role of Tet proteins in enhancer activity and telomere elongation. Genes Dev. 2014;28:2103-19.

19. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486:346-52.
20. Comes S, Gagliardi M, Laprano N, Fico A, Cimmino A, Palamidessi A, et al. L-Proline induces a mesenchymal-like invasive program in embryonic stem cells by remodeling H3K9 and H3K36 methylation. Stem cell reports. 2013;1:307-21.
21. D'Aniello C, Fico A, Casalino L, Guardiola O, Di Napoli G, Cermola F, et al. A novel autoregulatory loop between the Gcn2-Atf4 pathway and l-Proline metabolism controls stem cell identity. Cell Death Differ. 2015;22:1234.
22. Canalis E. Effect of glucocorticoids on type I collagen synthesis, alkaline phosphatase activity, and deoxyribonucleic acid content in cultured rat calvariae. Endocrinology. 1983;112:931-9.
23. Xiong G, Deng L, Zhu J, Rychahou PG, Xu R. Prolyl-4-hydroxylase alpha subunit 2 promotes breast cancer progression and metastasis by regulating collagen deposition. BMC Cancer. 2014;14:1.
24. Gilkes DM, Chaturvedi P, Bajpai S, Wong CC, Wei H, Pitcairn S, et al. Collagen prolyl hydroxylases are essential for breast cancer metastasis. Cancer Res. 2013;73:3285-96.
25. Westcott JM, Prechtl AM, Maine EA, Dang TT, Esparza MA, Sun H, et al. An epigenetically distinct breast cancer cell subpopulation promotes collective invasion. J Clin Invest. 2015;125:1927-43.
26. Thienpont B, Steinbacher J, Zhao H, D'Anna F, Kuchnio A, Ploumakis A, et al. Tumour hypoxia causes DNA hypermethylation by reducing TET activity. Nature. 2016;537:63-8.
27. Shenoy N, Bhagat T, Nieves E, Stenson M, Lawson J, Choudhary GS, et al. Upregulation of TET activity with ascorbic acid induces epigenetic modulation of lymphoma cells. Blood Cancer J. 2017;7:e587.
28. Bierie B, Pierce SE, Kroeger C, Stover DG, Pattabiraman DR, Thiru P, et al. Integrin-beta4 identifies cancer stem cell-enriched populations of partially mesenchymal carcinoma cells. Proc Natl Acad Sci U S A. 2017;114:E2337-E46.
29. Busby J, Murray L, Mills K, Zhang SD, Liberante F, Cardwell CR. A combined connectivity mapping and pharmacoepidemiology approach to identify existing medications with breast cancer causing or preventing properties. Pharmacoepidemiol Drug Saf. 2018;27:78-86.
30. Gomez-Rubio P, Zock JP, Rava M, Marquez M, Sharp L, Hidalgo M, et al. Reduced risk of pancreatic cancer associated with asthma and nasal allergies. Gut. 2017;66:314-22.
31. Delarive J, Saraga E, Dorta G, Blum A. Budesonide in the treatment of collagenous colitis. Digestion. 1998;59:364-6.
32. Tromm A, Griga T, Mollmann HW, May B, Muller KM, Fisseler-Eckhoff A. Budesonide for the treatment of collagenous colitis: first results of a pilot trial. Am J Gastroenterol. 1999;94:1871-5.
33. Bonderup OK, Hansen JB, Birket-Smith L, Vestergaard V, Teglbjaerg PS, Fallingborg J. Budesonide treatment of collagenous colitis: a randomised, double blind, placebo controlled trial with morphometric analysis. Gut. 2003;52:248-51.
34. O'Reilly S. Epigenetics in fibrosis. Mol Aspects Med. 2017;54:89-102.
35. Zhang X, Hu M, Lyu X, Li C, Thannickal VJ, Sanders YY. DNA Methylation Regulated Gene Expression in Organ Fibrosis. Biochim Biophys Acta. 2017.
36. Lu P, Weaver VM, Werb Z. The extracellular matrix: a dynamic niche in cancer progression. J Cell Biol. 2012;196:395-406.

37. Xiong G, Stewart RL, Chen J, Gao T, Scott TL, Samayoa LM, et al. Collagen prolyl 4-hydroxylase 1 is essential for HIF-1alpha stabilization and TNBC chemoresistance. Nat Commun. 2018;9:4456.

Figure Legends

Figure 1. HTS-based identification of drugs that block esMT. (**A**) Photomicrographs of colony types. (**B**) Miniaturization of the colony phenotype-based drug-screening method. (**C**) Colony type quantification (domed *vs* flat) of ESCs treated with L-Pro (250µM) ± VitC (150µM) or CHIR99021 (CHIR, 3µM) or left untreated (day5). Data are mean±SD (*$p \leq 0.001$; ~100 colonies scored/condition n=5). (**D**) Representative pictures of the screening results. Red and green circles indicate ESC inhibitors (no colonies) and esMT inhibitors (domed-shaped colonies), respectively. (**E**) Enrichment of ESC inhibitors in specific therapeutic classes (Z-score test, $p<0.0001$). (**F**) Molecular target(s) of ESC inhibitors from PubChem/PubMed databases. (**G**) Structure-activity relationship of drugs belonging to the Ca^{2+}-channel blockers 1,4-dihydropyridines. (**H**) Colony type quantification of ESCs treated ± L-Pro ± esMT inhibitors (day5). DMSO was used as control. VitC and CHIR99021 were used as positive controls. Data are mean±SD (*$p \leq 0.001$; ~100 colonies scored/condition n=5).

Figure 2. Molecular and functional characterization of esMT inhibitors. (**A-C**) Effect of Budesonide, Spiramycin, Propafenone on L-Pro -induced cell motility (**A**), down-regulation of AAR markers (**B**) and cell proliferation (**C**). DMSO is used as control. Representative frames of time-lapse series from untreated (-L-Pro) and L-Pro ± esMTi treated ESCs. Images were captured from day 3 to 4 post plating (**A**). qPCR analysis of *Atf4, Trb3* and *Nupr1* expression. Data are fold change *vs* untreated after normalization to *Gapdh* and are mean±SEM (*$p \leq 0.01$; n=3) (**B**). FACS analysis of EdU incorporation. Data are shown as fold change *vs* untreated and are mean±SEM (*$p \leq 0.01$; n=3) (**C**). (**D**) Effect of Budesonide, Spiramycin, Propafenone on the reversed mesenchymal-like to embryonic stem cell transition (MesT). Freshly generated L-Pro-

treated ESCs (mesenchymal-like flat colonies) were plated with L-Pro ± the indicated compounds and analyzed at day 5. Colony type quantification (domed vs flat; *left*). Data are mean±SD (*$p \leq 0.01$; ~100 colonies scored/condition; n=3). Representative pictures of L-Pro ± esMTi or VitC -treated ESCs (*middle*). Effect of L-Pro, VitC and Budesonide on esMT and MesT (*right*). (**E**) Western blot analysis of Col1a1 (*top*) and Collagen Hydroxyproline (C-HyP; *bottom*) in ESCs ± L-Pro (1mM). Densitometric analysis (ADU) is shown as fold change *vs* untreated ESCs. Data are mean±SEM (*$p \leq 0.01$; n=3) after normalization to Gapdh. (**F**) Confocal images of C-HyP (red) and Protein disulphide isomerase (PDI; green) immunofluorescence on ESCs ± L-Pro (*left*). Half circles indicate Endoplasmic reticulum (ER). Quantification of C-HyP intensity (*right*). Data are mean±SD (*$p \leq 0.001$; n=3). (**G**) Western blot analysis of Col1a1 in ESCs + L-Pro (1mM) ± Budesonide (10μM). Densitometric analysis (ADU) is shown as fold change *vs* ESCs + L-Pro. Data are mean±SEM (*$p \leq 0.01$; n=3) after normalization to Gapdh. (**H**) Fluorescence photomicrographs of Col1a1 (green) and confocal images of C-HyP (red) immunofluorescence on L-Pro-treated cells ± Budesonide. Nuclei are stained with DAPI. (**I-J**) Western blot analysis of C-HyP in ESCs (**I**) and PiCs (**J**) treated with L-Pro (1mM) ± Budesonide (10μM) or VitC (500μM). ADU is shown as C-HyP/Gapdh. Data are mean±SEM (*$p \leq 0.01$; n=3). (**K**) Schematic representation of L-Pro, VitC and Budesonide effects on C-HyP.

Figure 3. Generation and functional characterization of *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESCs. (**A**) RT-PCR of *P4ha2* from CRISPR-Cas9 Non Targeting (NT) and *P4ha2*^{KO} ESCs. (**B**) Representative pictures of Alkaline phosphatase (AP) staining of NT, *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESC colonies. (**C**) Western blot analysis of P4ha1 levels in Control (NT), *P4ha2*^{KO}/*a1*^{KD} and *P4ha2*^{KO} ESCs. Anti-Gapdh antibody was used as loading control. (**D**)

Representative photomicrographs (*left*) and colony type quantification (*right*) of colonies generated from NT, *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro (500μM) at day 5. Data are mean±SD (*p<0.01, ~100 colonies scored/condition; n=5). (**E**) L-Pro -induced proliferation of NT, *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro (500μM) analyzed by CCK-8 assay. Data are shown as fold change vs NT minus L-Pro. Data are mean±SEM (*p<0.01; n=3). (**F**) Confocal images of C-HyP immunofluorescence on NT, *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESCs treated with L-Pro (500μM) (*top*). Nuclei are stained with DAPI. C-Hyp intensity quantification (*bottom*). Data are mean±SD (*$p\leq$0.01; n=3). (**G**) Pie graphs comparing the DNA methylation profiles of L-Pro-treated ESCs and the *Tet* triple KO (TKO) ESCs. Blue slices indicate the percentage of hyper-DMRs in TKO that are hypermethylated by L-Pro at day 2 (*top)* and day 5 (*bottom*); p-value<10^{-16}. (**H-I**) ELISA-based quantification (**H**) and immunofluorescence (**I**) of 5hmC in Control (NT), *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro (500 μM; 24 hrs). (**H**) 5hmC levels are shown as fold change *vs* NT minus L-Pro. Data are mean±SEM (*$p\leq$0.01; n=3). (**I**) Fluorescence photomicrographs of 5hmC (red). Nuclei are stained with DAPI. (**J-K**) Western blot analysis of H3K9me3 in NT, *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro (500μM) (**J**), and in wild-type ESCs ± L-Pro (500μM) ± VitC (**K**). ADU is shown as fold change *vs* untreated (minus L-Pro) Controls after normalization to total histone H3. Data are mean±SEM (*$p\leq$0.01; n=3). (**L**) Schematic representation of VitC-dependent collagen hydroxylation (P4H) and DNA/Histone hydroxylation (TET/JMJ) in the different subcellular compartments.

Figure 4. Collagen synthesis/hydroxylation induces mesenchymal transition in lung epithelial tumor cells (A549) modifying the epigenetic landscape. (**A**) Schematic representation of 2D cell migration assay from A549 lung tumor spheroid. (**B**) Representative photomicrographs of crystal

violet stained A549 spheroids treated with Budesonide (10 and 20 μM) or vehicle (DMSO). (**C**) Representative pictures of Col1a1 immunofluorescence on A549 cells treated with Budesonide (10μM) or DMSO for 48 hrs. Nuclei are stained with DAPI. (**D**) Western blot analysis of H3K9me2/me3, H3K27me3, H3K36me3 and H3K4me3 in wild-type A549 cells treated with Budesonide (10 and 20μM) or DMSO as control, for 48 hrs. ADU is shown as fold change *vs* DMSO after normalization to total histone H3. Data are mean±SEM (*$p \leq 0.01$; n=3). (**E-G**) Generation and functional characterization of *P4HA2*^{KD} A549 cells. (**E**) Confocal images of C-HyP immunofluorescence on NT and *P4HA2*^{KD} A549 cells (*top*). Nuclei are stained with DAPI. C-Hyp intensity quantification (*bottom*). Data are mean±SE (*$p \leq 0.01$; n=2). (**F**) Representative photomicrographs of NT and *P4HA2*^{KD} A549 derived spheroids stained with crystal violet (*top*). Quantification of spreading spheroids (*bottom*). Data are mean±SE (*$p \leq 0.01$; n=3). (**G**) Western blot analysis of H3K9me2 and H3K36me3 in NT and *P4HA2*^{KD} A549 cells. ADU is shown as fold change *vs* NT after normalization to total histone H3. Data are mean±SEM (*$p \leq 0.01$; n=3).

Figure 5. Collagen prolyl hydroxylation influences the mesenchymal phenotype of triple-negative breast cancer (TNBC) cells modifying the epigenetic landscape. (**A**) P4HA2 expression is an independent predictor of poor prognosis in Breast Cancer. Left panel: prognostic significance of P4HA2 expression determined by Kaplan-Meier survival analysis in the METABRIC datasets. Right panels: The risk ratios (High vs. Low P4HA2 expression) were estimated with a Cox proportional hazards in univariate and multivariable model, adjusted for standard clinicopathological parameters. (**B**) Western blot analysis of P4HA2 in Control non-targeted (NT) and *P4HA2*^{KD} SUM159 cells. Anti-GAPDH antibody was used as loading control. (**C**) Representative pictures of NT and *P4HA2*^{KD} SUM159 cell morphology. (**D**) Western blot

analysis of FIBRONECTIN1 (FB1) and VIMENTIN levels in NT and *P4HA2*^{KD} SUM159 cells. Anti-GAPDH antibody was used as loading control. (**E**) CCK-8 based proliferation assay of NT and *P4HA2*^{KD} SUM159 cells. Data are shown as fold change vs NT and are mean±SEM (n=3). (**F**) Confocal images of C-HyP immunofluorescence (*left*) and C-Hyp intensity quantification (*right*) of NT and *P4HA2*^{KD} SUM159 cells. Nuclei are stained with DAPI. Data are mean±SE (*$p \leq 0.01$; n=2). (**G**) Western blot analysis of H3K27me3, H3K36me3, H3K4me3 and H3K9me2/me3 in NT and *P4HA2*^{KD} SUM159 cells. ADU is shown as fold change *vs* NT after normalization to total histone H3. Data are mean±SEM (*$p \leq 0.01$; n=3).

Figure 6. (**A-B**) Confocal images of Col1a1 (**A**) and C-HyP (**B**) immunofluorescence and signal intensity quantification of wild-type SUM159 cells ± Budesonide (25 μM) for 48 hrs. Nuclei are stained with DAPI. Data are mean±SE (*p≤0.01; n=3). (**C**) Western blot analysis of H3K9me2/me3, H3K36me3, and H3K27me3 and H3K4me3 in SUM159 cells ± Budesonide (25 μM) for 48 hrs. ADU is shown as fold change vs DMSO after normalization to total histone H3. Data are mean ± SEM (*$p \leq 0.01$; n=3). (**D**) Western blot analysis of H3K9me2/me3, H3K27me3, H3K4me3 and H3K36me3 in SUM159 cells ± VitC (10, 20 and 50 μM). ADU is shown as fold change vs untreated cells after normalization to total histone H3. Data are mean±SEM (*$p \leq 0.01$; n=3).

Figure 7. Budesonide blocks triple-negative breast cancer cell migration and invasion *in vitro* and *in vivo* (**A**) Schematic representation of SUM-159 3D organotypic culture assay. (**B**) Representative pictures (*left*) and circularity index (*right*) of SUM-159-derived spheroids ± Budesonide. The colony circularity was measured by using the formula {4π*area/perimeter^2}.

A value of 1.0 indicates a perfect circle. As the value approaches to 0.0, it indicates an increasingly elongated shape. Data are mean±SD (*$p \leq 0.01$; 50 colonies/group; n=3). (**C**) Representative pictures of H&E and VIMENTIN staining on sections from SUM-159 Luciferase-derived tumors (*left*). Cell circularity (*right*) is shown as mean±SD (*$p \leq 0.01$; 50 cells/tumors, n=5 mice/group). (**D**) Representative pictures of Picrosirius red staining and polarized light microscopy of tumor sections (*left*). Collagen quantification is shown as percentage of collagen area/field (*right*). Data are mean±SD (*$p \leq 0.01$; n=30 fields/group). (**E**) Dissemination to lung and distal lymph nodes by *in vivo* imaging in mice ± Budesonide (*top*). A dark screen was placed on top of the primary tumor to visualize the lower signal of distal foci. The Relative Luciferase counts/mouse (*bottom*) are shown as mean±SD (*$p < 0.01$; n=5 mice/group). (**F**) Representative pictures of Human Nuclei staining on lung sections from mice ± Budesonide (*left*). Quantification of disseminated tumor cells is shown as number of micrometastasis/lung (*right*). Data are the mean±SD (*$p < 0.01$; n=5 mice/group).

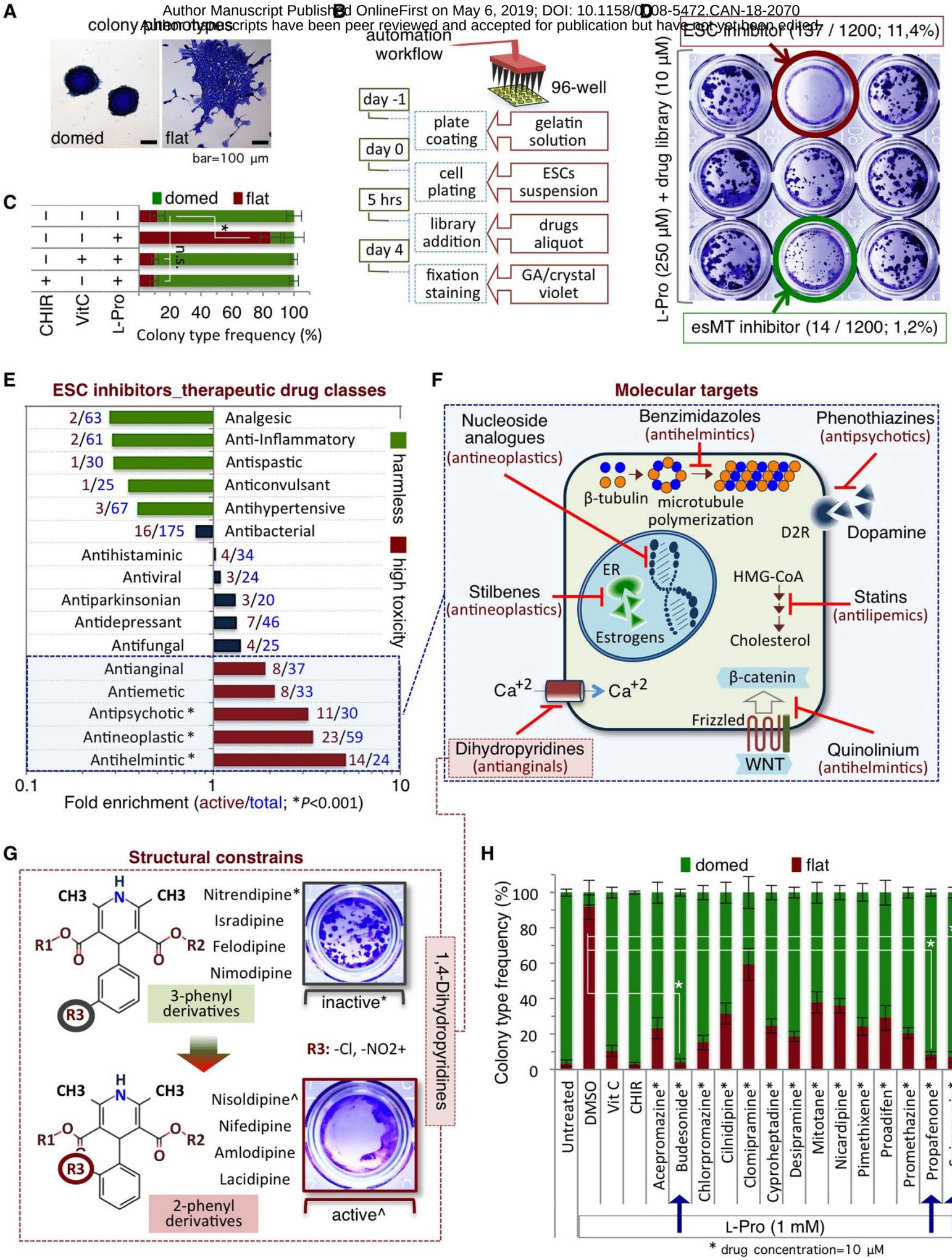
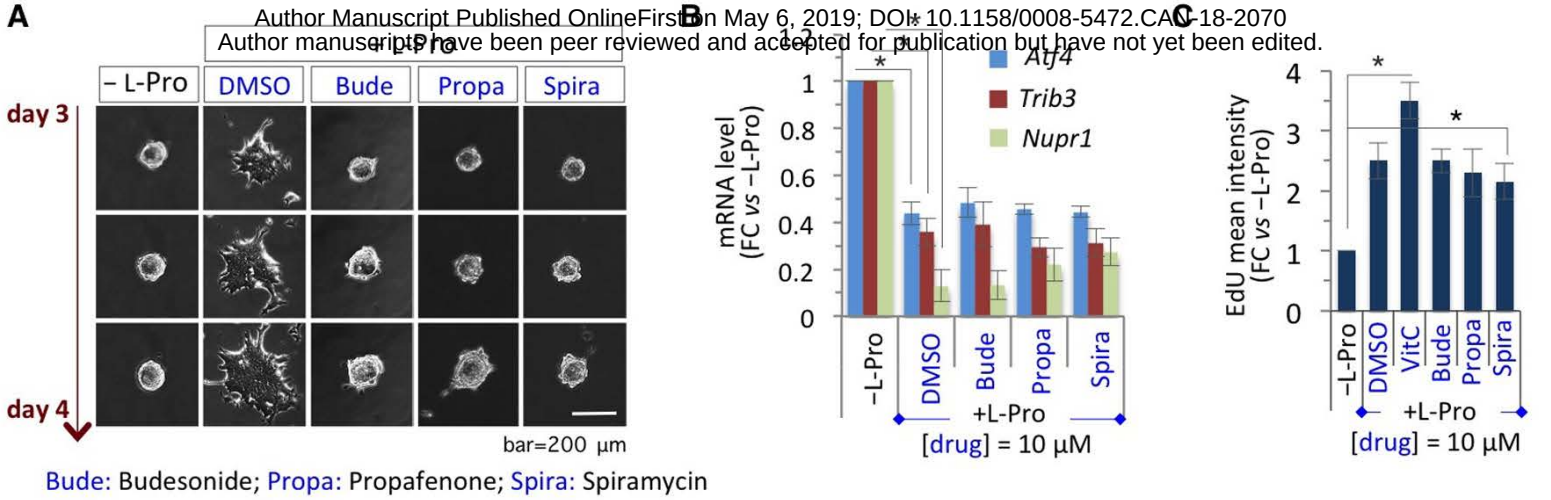
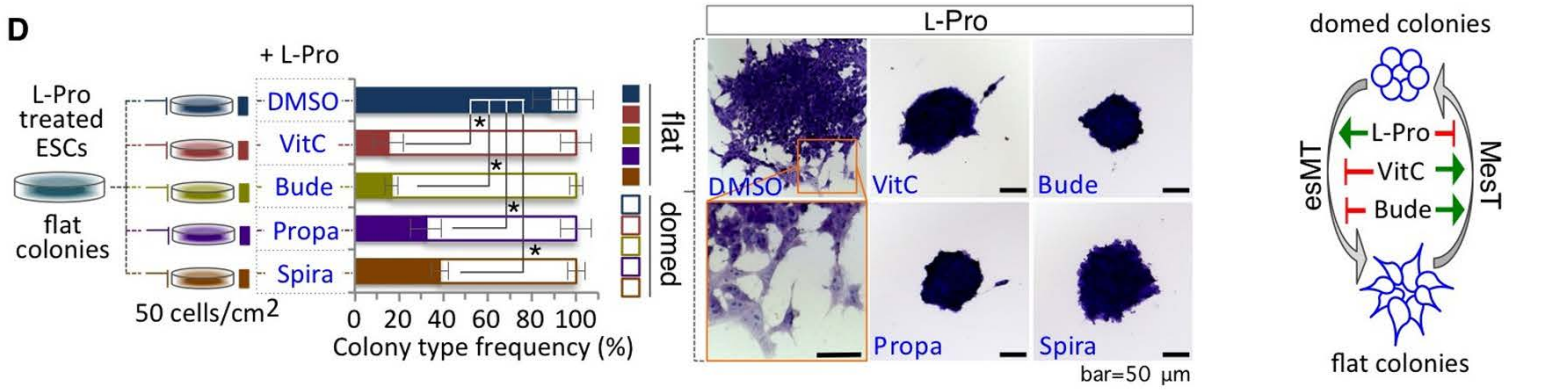
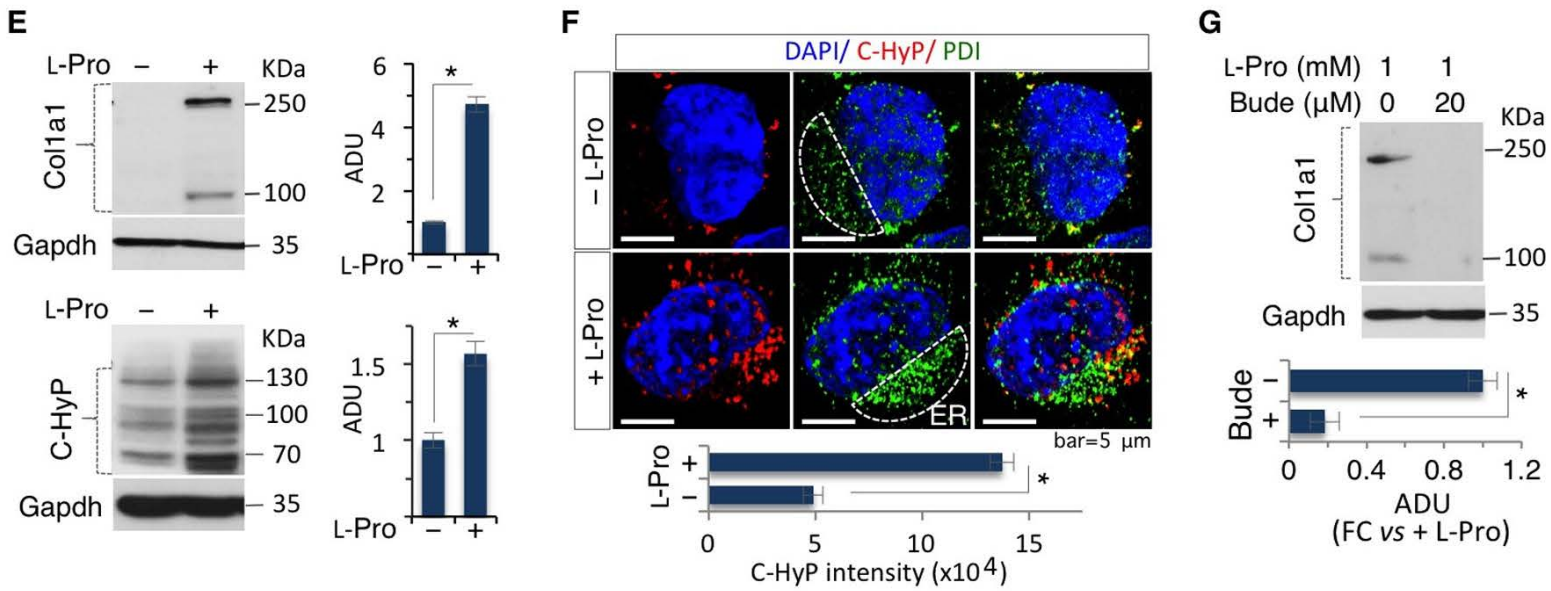


Figure 1



Bude: Budesonide; Propa: Propafenone; Spira: Spiramycin





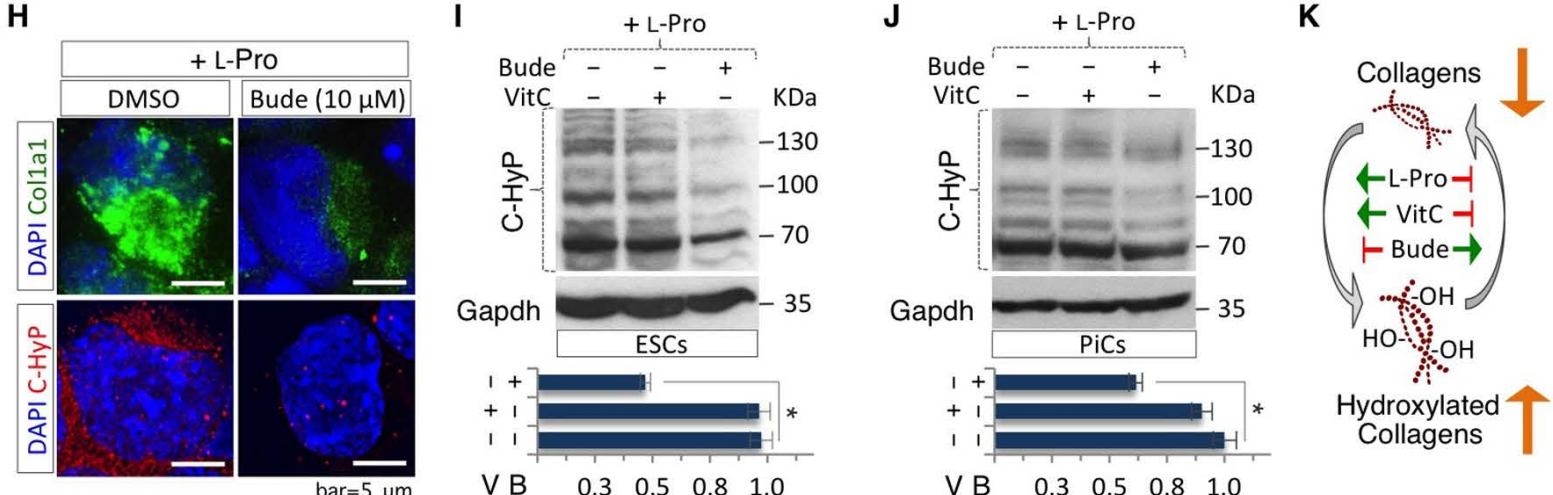


Figure 2

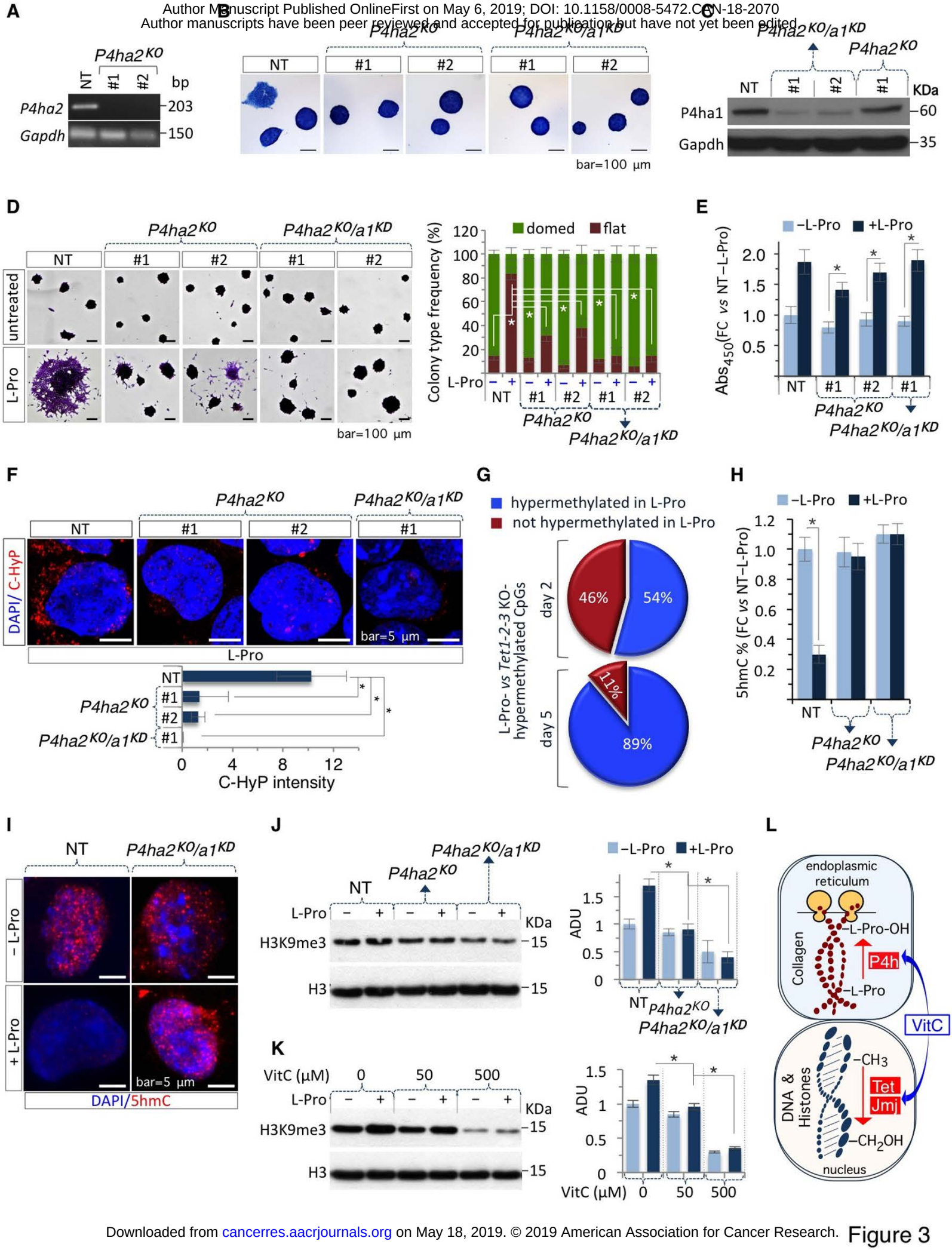


Figure 3

Author Manuscript Published OnlineFirst on May 6, 2019; DOI: 10.1158/0008-5472.CAN-18-2070
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

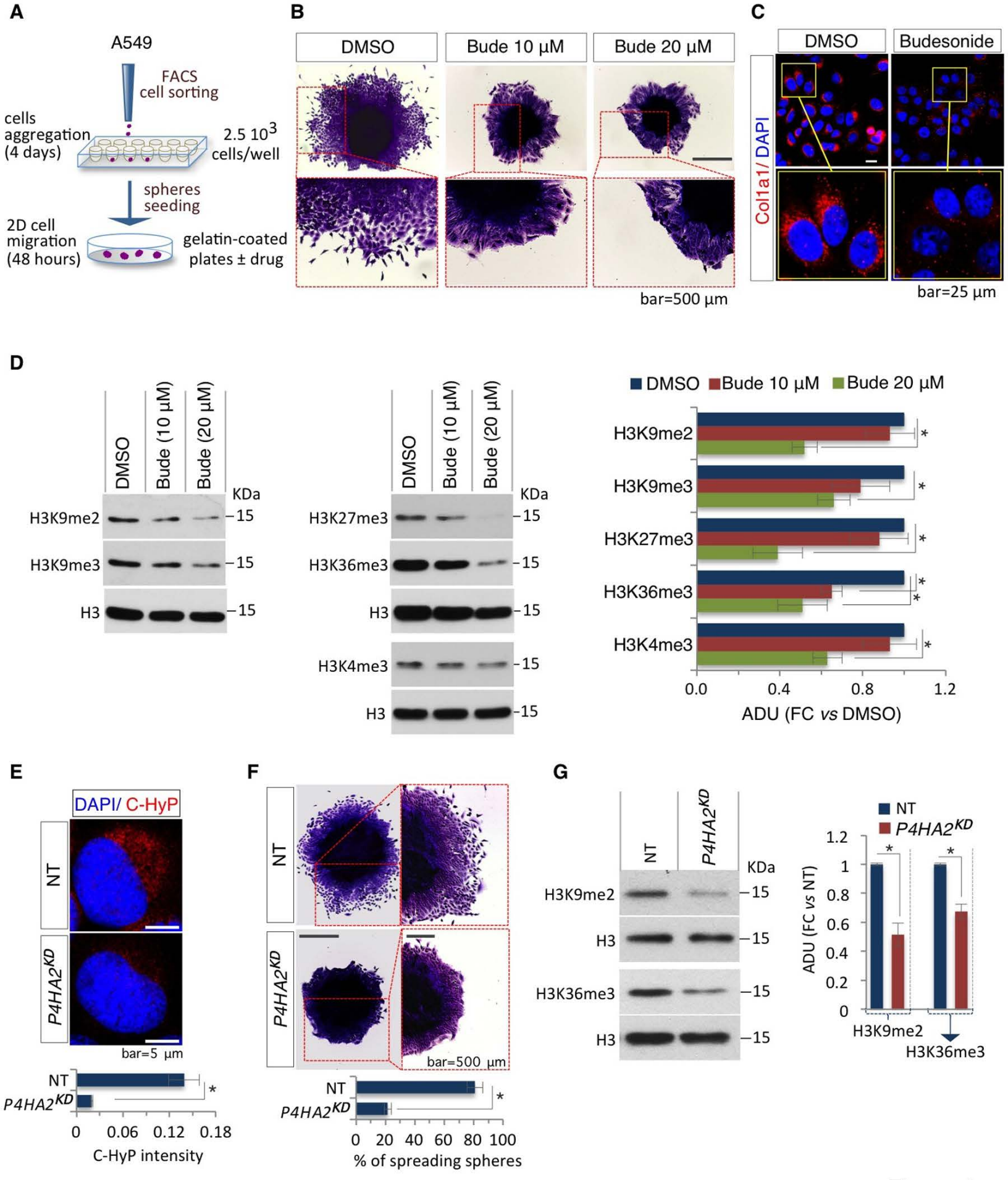
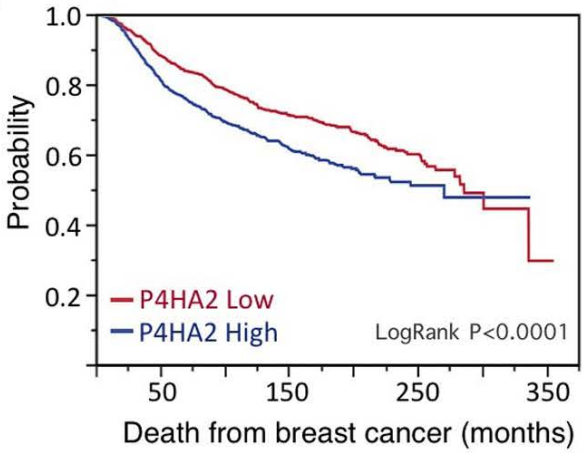


Figure 4

Author Manuscript Published OnlineFirst on May 6, 2019; DOI: 10.1158/0008-5472.CAN-18-2070
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

A



P4HA2 Univariate DFBC (Death from breast cancer)

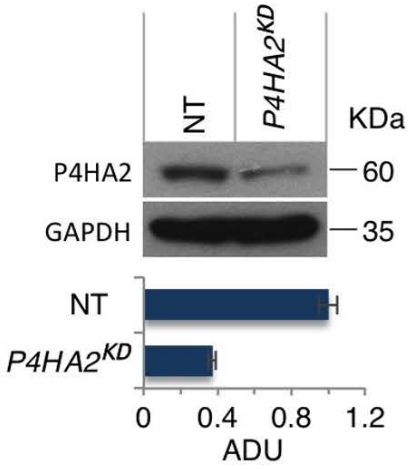
Level1	Level2	Risk Ratio	Lower 95%	Upper 95%	Prob<Chisq
High*	Low*	1.415	1.212	1.652	<.0001

P4HA2 Multivariate DFBC (Death from breast cancer)

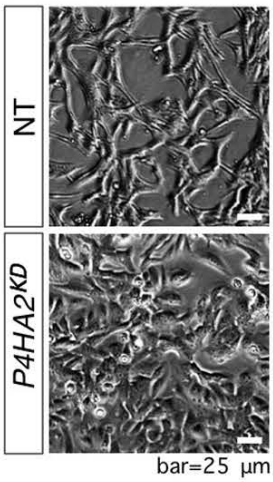
Level1	Level2	Risk Ratio	Lower 95%	Upper 95%	Prob<Chisq
High*	Low*	1.379	1.177	1.161	<.0001
G2-3	G1	2.217	1.504	3.438	<.0001
Age<50	Age≥50	0.946	0.779	1.143	0.572
pT2-3	pT1	1.740	1.447	2.106	<.0001
ER-	ER+	1.528	1.273	1.827	<.0001

*P4HA2 expression

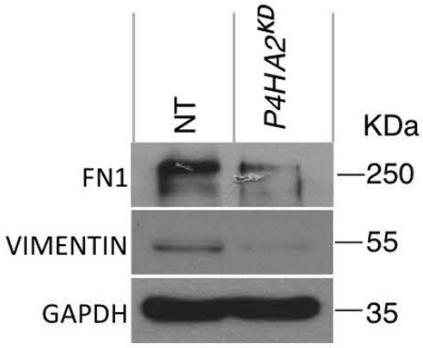
B



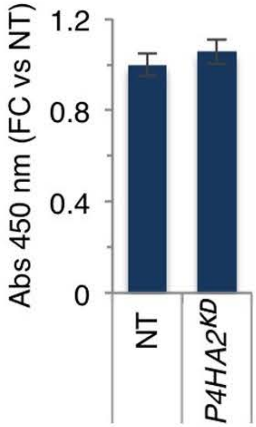
C



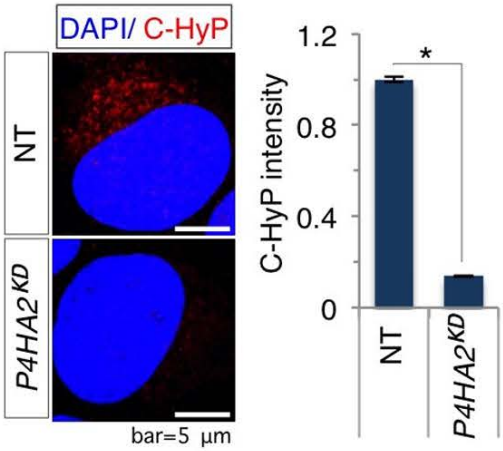
D



E



F



G

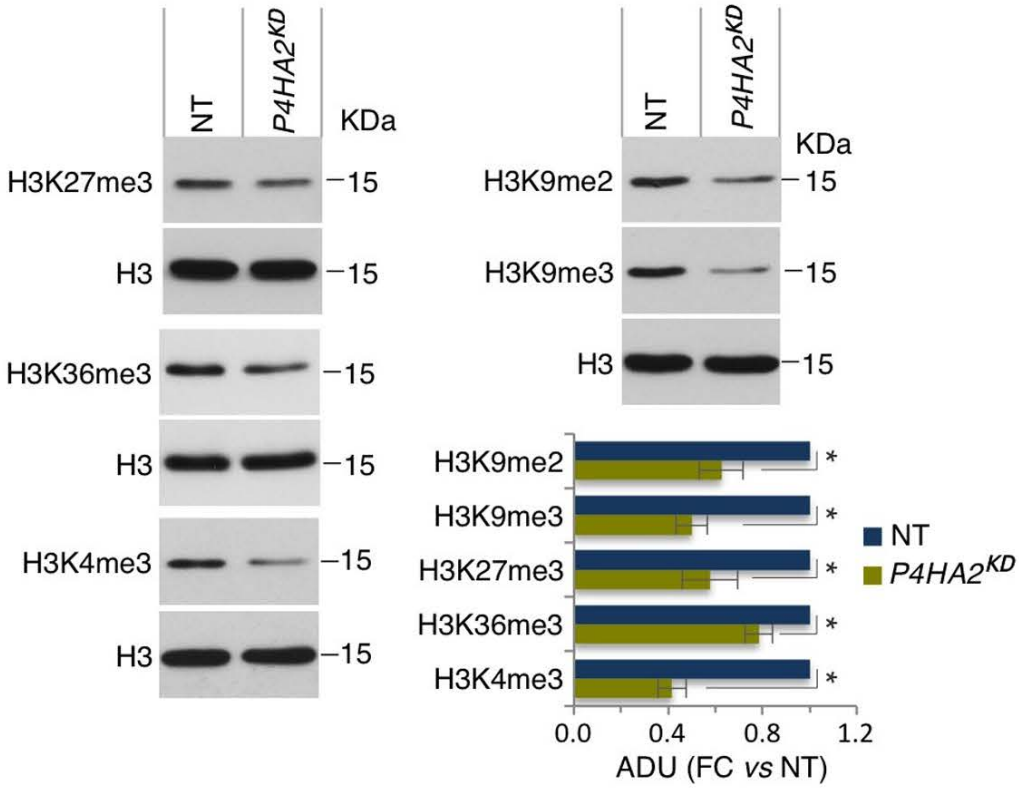


Figure 5

Author Manuscript Published OnlineFirst on May 6, 2019; DOI: 10.1158/0008-5472.CAN-18-2070
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

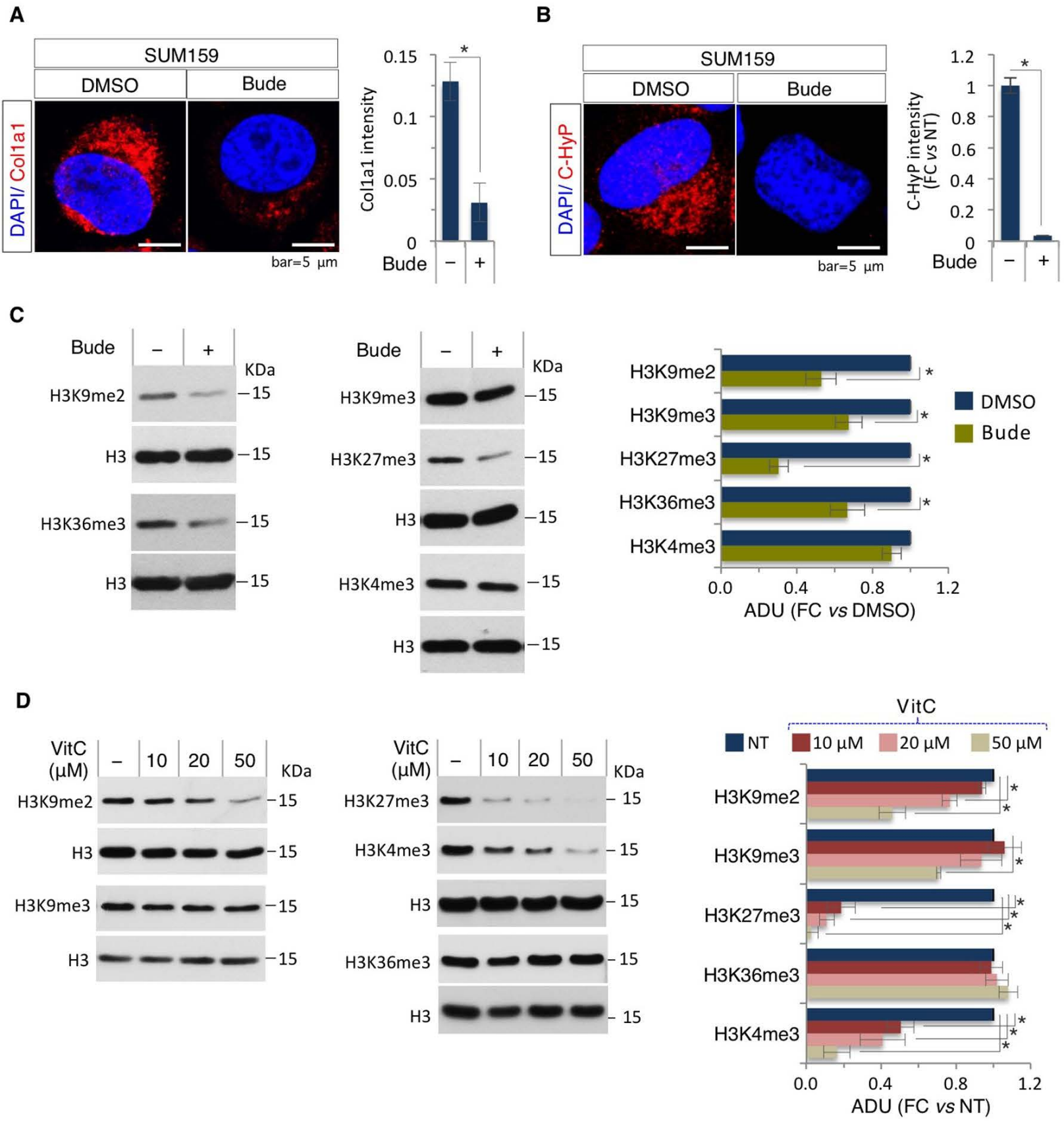


Figure 6

Author Manuscript Published OnlineFirst on May 6, 2019; DOI: 10.1158/0008-5472.CAN-18-2070
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

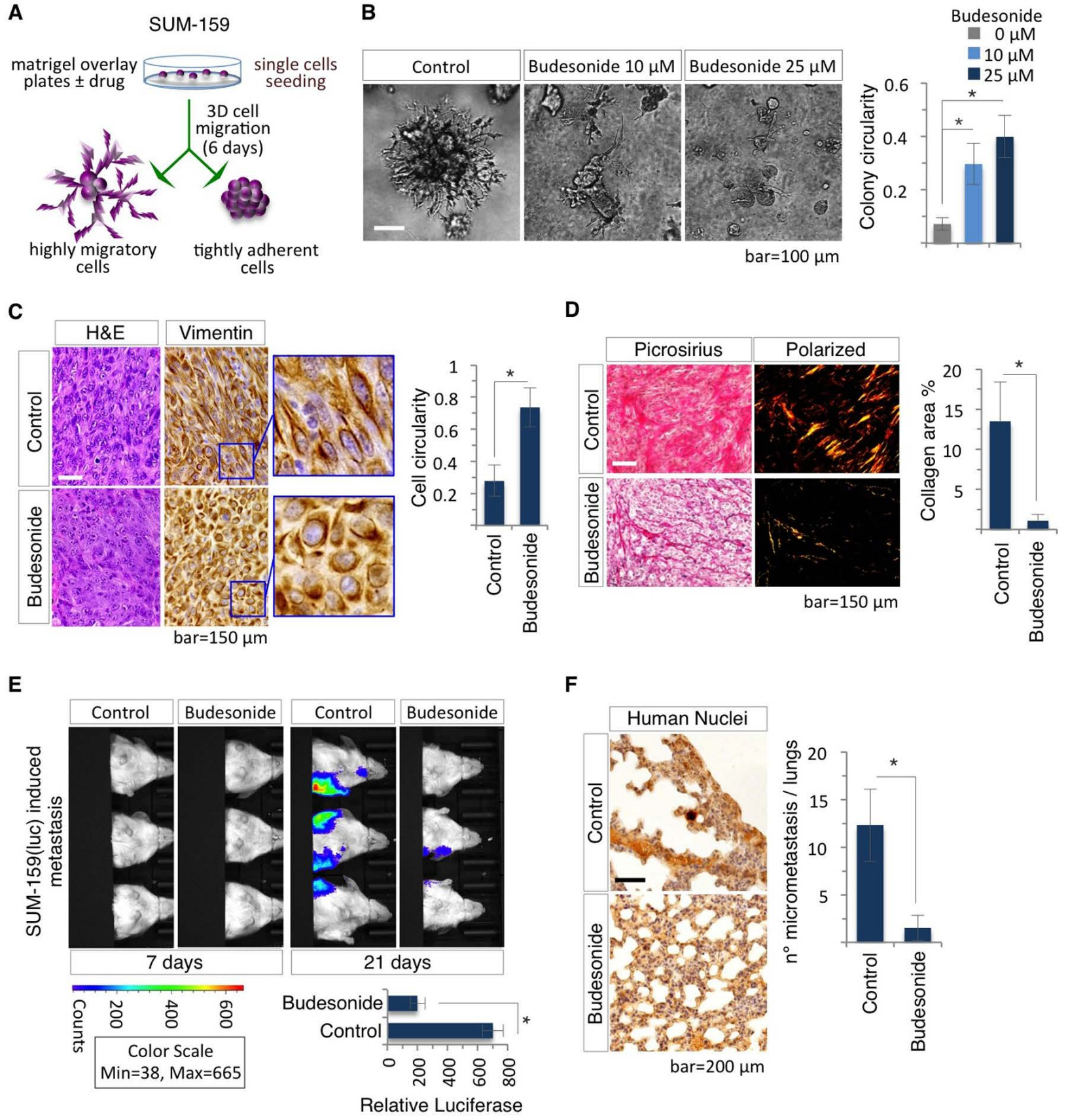


Figure 7

Cancer Research
The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

AACR American Association for Cancer Research

Collagen prolyl hydroxylation-dependent metabolic perturbation governs epigenetic remodeling and mesenchymal transition in pluripotent and cancer cells

Cristina D'Aniello, Federica Cermola, Andrea Palamidessi, et al.

Cancer Res Published OnlineFirst May 6, 2019.