ORIGINAL ARTICLE

A partial epithelial–mesenchymal transition signature for highly aggressive colorectal cancer cells that survive under nutrient restriction

Gil A Pastorino[1†], Ilir Sheraj[2], Kerstin Huebner[1†], Giulio Ferrero[3], Philipp Kunze[1†], Arndt Hartmann[1‡], Chuanpit Hampel[1†], Hepsen Hazal Husnugil[2], Arnatchai Maiuthed[4,5], Florian Gebhart[1†], Fynn Schlattmann[1†], Aliye Ezgi Gulec Taskiran[2,6], Goksu Oral[2], Ralph Palmisano[7], Barbara Pardini[8,9], Alessio Naccarati[8,9], Katharina Erlenbach-Wuensch[1], Sreeparna Banerjee[2,10] and Regine Schneider-Stock[1*†,‡,§] (ID)

[1]   *Institute of Pathology, Universitätsklinikum Erlangen, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany*
[2]   *Department of Biological Sciences, Orta Dogu Teknik Universitesi, Ankara, Turkey*
[3]   *Department of Clinical and Biological Sciences, University of Turin, Turin, Italy*
[4]   *Department of Pharmacology, Mahidol University, Bangkok, Thailand*
[5]   *Centre of Biopharmaceutical Science for Healthy Ageing, Faculty of Pharmacy, Mahidol University, Bangkok, Thailand*
[6]   *Department of Molecular Biology and Genetics, Baskent University, Ankara, Turkey*
[7]   *Optical Imaging Competence Centre FAU OICE, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany*
[8]   *Italian Institute for Genomic Medicine (IIGM), c/o FPO-IRCCS Candiolo, Turin, Italy*
[9]   *Candiolo Cancer Institute, FPO-IRCCS, Turin, Italy*
[10] *Cancer Systems Biology Laboratory (CanSyl), Orta Dogu Teknik Universitesi, Ankara, Turkey*

*Correspondence to: R Schneider-Stock, Experimental Tumorpathology, Institute of Pathology, Universitätsklinikum Erlangen, Universitätsstr. 22, 91054 Erlangen, Germany. E-mail: regine.schneider-stock@uk-erlangen.de

†Current address: Experimental Tumorpathology Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany.
‡Comprehensive Cancer Center Erlangen-EMN (CCC ER-EMN), Universitätsklinikum Erlangen, FAU Erlangen-Nürnberg, Erlangen, Germany.
§Adjunct Professor, Faculty of Pharmacy, Mahidol University, Bangkok, Thailand.

Abstract

Partial epithelial–mesenchymal transition (p-EMT) has recently been identified as a hybrid state consisting of cells with both epithelial and mesenchymal characteristics and is associated with the migration, metastasis, and chemoresistance of cancer cells. Here, we describe the induction of p-EMT in starved colorectal cancer (CRC) cells and identify a p-EMT gene signature that can predict prognosis. Functional characterisation of starvation-induced p-EMT in HCT116, DLD1, and HT29 cells showed changes in proliferation, morphology, and drug sensitivity, supported by *in vivo* studies using the chorioallantoic membrane model. An EMT-specific quantitative polymerase chain reaction (qPCR) array was used to screen for deregulated genes, leading to the establishment of an *in silico* gene signature that was correlated with poor disease-free survival in CRC patients along with the CRC consensus molecular subtype CMS4. Among the significantly deregulated p-EMT genes, a triple-gene signature consisting of *SERPINE1*, *SOX10*, and epidermal growth factor receptor (*EGFR*) was identified. Starvation-induced p-EMT was characterised by increased migratory potential and chemoresistance, as well as E-cadherin processing and internalisation. Both gene signature and E-cadherin alterations could be reversed by the proteasomal inhibitor MG132. Spatially resolving EGFR expression with high-resolution immunofluorescence imaging identified a proliferation stop in starved CRC cells caused by EGFR internalisation. In conclusion, we have gained insight into a previously undiscovered EMT mechanism that may become relevant when tumour cells are under nutrient stress, as seen in early stages of metastasis. Targeting this process of tumour cell dissemination might help to prevent EMT and overcome drug resistance.
© 2024 The Authors. *The Journal of Pathology* published by John Wiley & Sons Ltd on behalf of The Pathological Society of Great Britain and Ireland.

Keywords: partial EMT; intermediate EMT; starvation; proteasome; EGFR internalisation; E-cadherin processing; consensus molecular subtype 4; prognosis

Introduction

Metastasis remains the leading cause of tumour-related death in colorectal cancer (CRC) [1]. Although advances in molecular analysis using single-cell or omics approaches have significantly improved our understanding of tumour progression over the past decade, the mechanisms underlying CRC metastasis remain unclear.

A well-known driver of cancer metastasis is the epithelial-to-mesenchymal transition (EMT), a differentiation program by which epithelial cells acquire mesenchymal-like properties that are necessary for physiological processes such as gastrulation, neural crest development, and wound healing [2]. The canonical EMT is characterised by the downregulation of epithelial markers through transcriptional repression by master EMT regulators such as the Snail, Twist, and Zeb transcription factor families [3]. Epithelial markers include membrane proteins, most notably E-cadherin (CDH1) [4], a widely recognised tumour suppressor and calcium-dependent cell adhesion molecule of the cadherin superfamily [5]. Loss of these markers triggers cell de-adhesion and subsequently enables their migratory and invasive potential [6].

EMT research has witnessed a significant resurgence as recent studies have shed light on the existence of hybrid states within the spectrum of epithelial and mesenchymal cells, challenging the traditional binary categorisation. These hybrid states, known as partial EMT (p-EMT), are now recognised as more influential in promoting metastatic cancer cell dissemination. They enable a collective migration of metastatic cells, which contrasts with the relatively ineffective single-cell migration associated with complete EMT [7,8]. While EMT has well-defined endpoints, cell states resulting from p-EMT are vaguely described, as they do not fit neatly into either category [9]. In addition, it is still unclear whether p-EMT truly resembles a subset of EMT or represents a mechanism with distinct endpoints of their own [10]. Although the induction of EMT via pathways such as the TGF-β pathway is well established [11], there is still a lack of knowledge regarding the driving mechanisms of p-EMT.

Cancer cells are well known for their ability to recapitulate and exploit distinct developmental programs, such as EMT, to cope with selective pressures of environmental stress [11]. Recent studies have identified the metabolic reprogramming of cancer cells as a regulator of p-EMT [12–14]. Indeed, cancer cells often rely on aerobic glycolysis and glucose availability rather than oxidative phosphorylation, a phenomenon known as the Warburg effect [15]. Uncontrolled cell growth leads to disproportionate distribution of nutrients along the tumour microenvironment (TME). Therefore, cancer cells often face an imbalance between nutrient supply and demand, especially in the centre of the tumour mass [16]. Consequently, these cells must adapt their metabolism according to the availability of nutrients in the surrounding TME. Although nutrient starvation by glucose and growth factor deprivation are known to induce EMT [17–19], the molecular signature and the nature of transdifferentiation have not yet been described. Recently, the 'EMTome' portal has provided EMT- and p-EMT-related gene signatures in different cancer types [20]. Here, a highly multifaceted nature for p-EMT is suggested, which differs in a stimulus-dependent manner.

In this study, we aimed to gain insight into the molecular mechanisms of p-EMT induced by nutrient deprivation-mediated metabolic stress in different colon cancer cell lines. We functionally characterised the starved cells *in vitro* and *in vivo* and identified a specific triple-gene signature that might identify patients at risk for metastasis. Spatial analysis of these gene signature markers allowed the interpretation of the underlying molecular mechanism.

Materials and methods

Cell culture and reagents

All cell lines (HCT116, DLD1, HT29, HCT116 p21−/−, and LoVo were cultured under standard conditions (37 °C, 5% CO_2). HCT116, DLD1, HT29, and LoVo cells were maintained in RPMI 1640 medium supplemented with 10% FBS and 1% penicillin–streptomycin (P/S, all PAN Biotech, Aidenbach, Germany). HCT116 p21−/− cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco™, Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 10% FBS, 1% P/S, 1% sodium pyruvate solution (Sigma-Aldrich, St. Louis, MO, USA), and 1% MEM-Non-Essential Amino Acids solution (Gibco). All cell lines were regularly tested for mycoplasma contamination and authenticated by Multiplexion (Heidelberg, Germany) [21]. For drug testing, DAPT (Santa Cruz Biotechnology, Dallas, TX, USA), TAPI-2 (Abcam, Cambridge, UK), MG132 (Calbiochem, San Diego, CA, USA), and 5-FU (Sigma-Aldrich) were used. See Supplementary material and methods and supplementary material, Table S1 for further details.

Cell viability assays

To determine the appropriate dose for the treatment of cells with drugs, a crystal violet assay was carried out. For this, cells were treated with various concentrations of DAPT (10–150 μM), TAPI-2 (0.625–80 μM), and MG132 (0.075–5 μM) for 48 h.

To determine the susceptibility of cells to 5-FU, cells were treated with control or starvation medium combined with MG132, 5-FU, or both for 48 h. After treatment, cell viability was determined by crystal violet or 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) assays. Details are given in Supplementary materials and methods and in supplementary material, Table S1.

Cell starvation assay

Cells were incubated in the starvation medium (described in supplementary material, Table S2) for the indicated time points. Details of treatment procedures are given in Supplementary materials and methods.

Cell migration assay

Prestarved cells were seeded into two-well culture inserts (Ibidi, Martinsried, Germany). On the next day, cells were treated with 1.0 μg/ml mitomycin C

(Sigma-Aldrich) for 2 h prior to the removal of the cell culture insert. Migration was documented at 0 and 24 h after insert removal. Details are given in Supplementary materials and methods.

Western blotting

Cell pellets were collected by scraping and lysed in urea lysis buffer, and western blotting was performed as described previously [22]. In brief, 30–60 μg of proteins were separated by SDS-PAGE and transferred onto nitrocellulose membranes (GE Healthcare, Chicago, IL, USA). After blocking, the membranes were probed with primary antibodies overnight. Details are given in Supplementary materials and methods. All antibodies are listed in supplementary material, Table S3.

Immunofluorescence and confocal imaging

Cells were seeded either on 8-well μ-slides (Ibidi) or on cover slips. After treatment, cells were fixed with 4% paraformaldehyde (PFA), permeabilised in 0.1% Triton X-100 (in PBS) and blocked with 1% BSA solution. Cells were incubated with the primary antibodies for 1 h at room temperature (RT) and subsequently probed with secondary antibodies for 30 min at RT. Cells were counterstained with DAPI (Catalogue No. MBD0015, Sigma-Aldrich) and with phalloidin (Catalogue No. PHDH1-A, Cytoskeleton Inc., Denver, CO, USA).

For live cell cytoskeleton staining, cells were seeded on cover slips and treated as indicated. After treatment, cells were stained with Hoechst 33342 (Catalogue No. 14533, Sigma-Aldrich) and CellMask Green Actin (Catalogue No. A57243, Invitrogen™, Waltham, MA, USA) for 3 h. Transporter and pump blockers probenicid (Catalogue No. P36400, Invitrogen™) and verapamil (Catalogue No. CY-SC001, Cytoskeleton Inc.) were added to the staining mixture to prevent any loss of signal due to export of the dyes. After staining, cells were fixed in 4% PFA and imaged.

Images were captured using a Nikon Ti2, CQ1 Confocal Imaging Cytometer (Yokogawa Life Sciences, Tokyo, Japan), Stellaris 8 (confocal, Leica, Wetzlar, Germany), or Zeiss laser scanning microscopy system (confocal, LSMT-PMT Observer ZI, LSM 710; Carl Zeiss Inc., Oberkochen, Germany). Details of antibody dilutions and methods are provided in Supplementary materials and methods.

RT–qPCR

Total RNA was extracted as described previously [22] using the QIAzol® Lysis Reagent in combination with the RNeasy Mini Kit according to the manufacturer's instructions (both Qiagen, Hilden, Germany). A list of primers is provided in supplementary material, Table S4. See Supplementary materials and methods for more details.

RT^2 profiler PCR array

Two micrograms of isolated total RNA were transcribed into cDNA using RT^2 First Strand Kit. RT^2 Profiler PCR Arrays for EMT gene panel (#PAHS-090ZD-0803, Qiagen) were performed using the RT^2 SYBR® qPCR Mastermix (Qiagen) as described in the manufacturer's protocol. See Supplementary materials and methods for more details. PCR results are summarised in supplementary material, Table S5.

Chorioallantoic membrane (CAM) assay

The CAM assay was performed as described previously [23]. In brief, fertilised specific pathogen-free (SPF) chicken eggs (VALO Biomedia, Osterholz-Scharmbeck, Germany) were incubated at 37 °C and 60% humidity in an upright position. For starvation, HCT116 cells were starved for 48 h. On embryonic developmental day 9, 1×10^6 cells of control and prestarved adherent HCT116 cells were resuspended in Matrigel (Corning® Inc., Corning, NY, USA) and transplanted onto the CAM. CAM tumours were resected after 5 days of *in ovo* growth and prepared for immunohistochemical evaluation. See Supplementary materials and methods for more details.

Immunohistochemistry

After deparaffinisation, haematoxylin and eosin (H&E), Picro-Sirius red and immunohistochemistry (IHC) for E-cadherin, EGFR, and SOX10 was performed using a Ventana BenchMark Ultra (Ventana Medical Systems, Inc., Oro Valley, AZ, USA) in combination with an UltraView Universal DAB Detection Kit (Roche, Basel, Switzerland) following the manufacturer's instructions. PAI-1 IHC staining was conducted manually. Details for antibody dilutions and slide scanning are provided in Supplementary materials and methods and supplementary material, Table S6.

In silico analyses

Details for the generation of the three-gene signature on GSE39582 and its validation on three additional datasets (GSE17536 [24], an RNA-Seq dataset [25,26], and the TCGA COAD data set [27,28]), as well as gene signature association with clinical data and the EMTome database are explained in detail in Supplementary materials and methods. Differential gene expression was investigated between high- and low-risk group samples of the GSE39582 dataset using the limma package [29], followed by Gene Set Enrichment Analysis (GSEA). For this, we used the fast GSEA (fGSEA) package with the hallmarks molecular signature database (MSigDB) after converting the microarray probes into their respective ENTREZ gene IDs and listing them by log-fold-change [30–32].

Statistical analyses

Statistical analyses were performed using GraphPad Prism version 8.3.0 (GraphPad, San Diego, CA, USA). Statistical testing, *n*-numbers, and *p* values for statistical significance are described in the figure legends.

Statistical significance was defined as $p < 0.05$. See Supplementary materials and methods for more details.

Results

Nutritional restriction triggers mesenchymal characteristics in CRC cell lines via proteasomal processing of E-cadherin

Nutrient deprivation is a typical metabolic challenge to which tumours are constantly exposed [33]. By mimicking these conditions *in vitro* using a standardised starvation protocol consisting of reduced levels of glucose, glutamine, and serum (supplementary material, Table S2),

we studied three colon cancer cell lines (HCT116, DLD1, HT29) representing three different basal EMT stages according to Tan *et al* [34] (Figure 1A). We observed that upon starvation, all three epithelial cell lines exhibited macroscopic EMT-like phenotypes with impaired intercellular adhesion and reduced cell numbers (Figure 1B). Starved HCT116 and DLD1 cells revealed elongated cell bodies with a few dense adherent junctions and the formation of spike-like structures between neighbouring cells. In HT29 cells, single clusters of tumour cells were visible (Figure 1B). This phenotypic switch was consistently reversed by the subsequent incubation of the cells with nutrient-rich cell culture medium (recovery) (Figure 1B), and our data show that this increase in the mesenchymal population upon
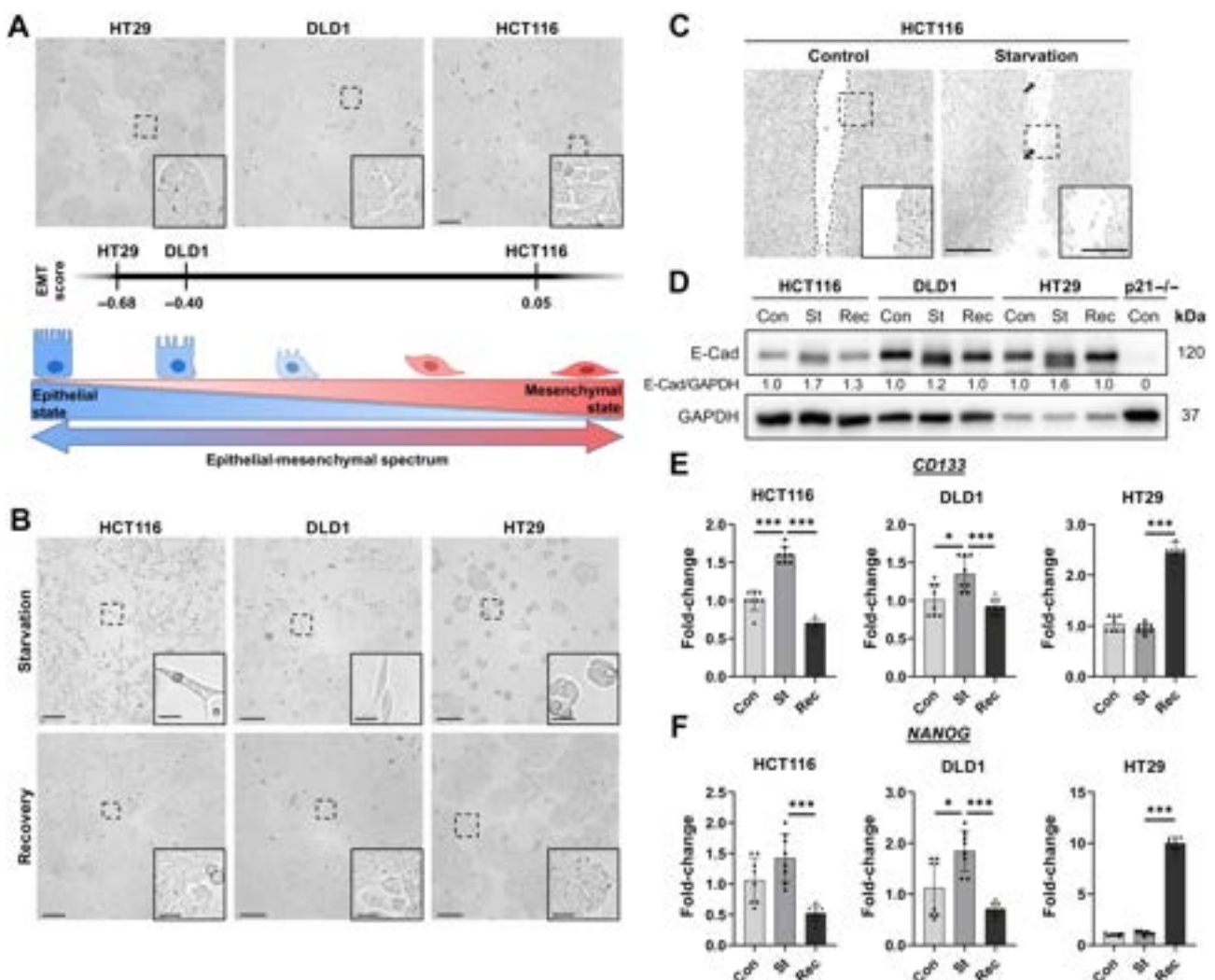

Figure 1. Nutrient restriction triggers partial epithelial-mesenchymal transition (p-EMT) in colorectal cancer (CRC) cells. (A) Upper part: brightfield control images and EMT score of the three colon cancer cell lines HT29, DLD1, and HCT116 used in this study. Overview images scale bar, 100 μm; insert scale bar, 25 μm. Lower part: EMT score was extracted from Tan *et al* [34] and is presented as the Avg. generic EMT score. Created using BioRender.com. (B) Representative images of 48 h starved and recovered (48 h starved plus 24 h recovered) CRC cell lines HCT116, DLD1, and HT29. Overview image scale bar, 100 μm; insert scale bar, 25 μm. (C) Migration assay of HCT116 cells displaying enhanced mobility of 48 h starved HCT116 cells (arrowheads) versus untreated controls (drawn borders of pushing front). Representative images. Overview image scale bar, 500 μm; insert scale bar, 250 μm. (D) Representative western blots of HCT116, DLD1, and HT29 cells with p21-knockout HCT116 cells (p21−/−) as mesenchymal positive control. The E-cadherin band was slightly shifted downward in 48 h starved cells (St) compared to untreated controls (Con) and recovered cells (Rec, 48 h starved plus 24 h recovery). (E and F) RT-qPCR for (E) *CD133* and (F) *NANOG* expressed as fold-change of 48 h starved and recovered (48 h starved plus 24 h recovered) HCT116, DLD1, and HT29 cells compared to non-starved controls (***$p < 0.001$, *$p < 0.05$, Mann–Whitney test).

starvation was not caused by apoptosis (supplementary material, Figure S1A–C).

To evaluate the migration pattern of starved CRC cells, we performed a migration assay of starved HCT116, DLD1, and HT29 cells. Consistent with our hypothesis, nutrient starvation resulted in rough cellular edges at the thrust front with an EMT-like single-cell migration pattern, in contrast to the clustered migration pattern observed in non-starved cells (Figure 1C, and supplementary material, Figure S1D). Next, we examined the two classical canonical EMT markers in CRC cells, E-cadherin and vimentin. Interestingly, a shift of the E-cadherin band from 120 kDa to a smaller band of ~115 kDa was observed in starved HCT116, DLD1, and HT29 cells (Figure 1D). None of the cell lines showed any upregulation of vimentin (data not shown). Next, we performed RT-qPCR for four different stemness markers, *CD133*, *NANOG*, *CD44*, and *OCT4* (Figure 1E,F and supplementary material, Figure S1E,F). We found that *CD133* and *NANOG* were upregulated in HCT116 and DLD1 cells upon starvation with a decrease in the recovery phase partly even below the levels of controls (supplementary material, Figure S1E,F). Interestingly, HT29 cells having higher *per se* basal endogenous CD133 levels (supplementary material, Figure S1G) showed a delay in the upregulation of *CD133* and *NANOG* upon starvation but had significantly higher levels in the recovery phase (Figure 1E,F).

Our findings suggest that nutrient starvation triggers p-EMT in HCT116, DLD1, and HT29 cells that cannot be classified as either fully epithelial or fully mesenchymal. Moreover, this starvation-induced p-EMT seems to be accompanied by an acquisition of stem-cell-like properties.

Starvation-induced E-cadherin expression pattern *in vivo*

Next, we evaluated E-cadherin expression and localisation *in vivo* using the CAM assay. We observed no significant reduction in the volume of tumours generated from starved cells compared to tumours generated from control (non-starved) cells. No other macroscopic differences were observed (supplementary material, Figure S2A,B). H&E staining of tumour sections revealed a highly invasive growth pattern of the ovografts derived from prestarved cells (Figure 2A). These ovografts exhibited a fingertip-like infiltrative growth pattern at the tumour invasion front, whereas tumours generated with control cells invaded the CAM with a clear pushing front (Figure 2A). Consistent with our *in vitro* studies, we observed two major phenomena: first, a loss of E-cadherin membrane association in starved ovografts at the tumour centre (Figure 2B,C), and second, a loss of membranous E-cadherin at the invasion front visible as a loss of sharp linings (Figure 2B). Single cells or small cell clusters showed a complete loss of E-cadherin staining (Figure 2B, magnifications), with the latter effects being more pronounced in prestarved ovografts. We suggest that the loss of membrane-localised

E-cadherin may have led to a loss of intercellular adhesion, facilitating migration into the tumour-adjacent stroma.

E-cadherin processing is controlled by the ubiquitin–proteasome system (UPS)

Given the central role of E-cadherin in intercellular adhesion and EMT and the post-translational processing observed in three different cell lines under starvation conditions, we sought to identify the key regulatory mechanisms involved in this process. E-cadherin is known to be cleaved by the autophagy-associated enzyme γ-secretase complex, an intramembranous protease complex that has been directly linked to EMT [35]. However, inhibition of γ-secretase with DAPT (*N*-[*N*-(3,5-difluorophenacetyl)-l-alanyl]-s-phenylglycine t-butyl ester) did not prevent the macroscopic mesenchymal phenotype observed in 2D cell culture and did not inhibit the formation of the 115 kDa cleaved form of E-cadherin (Figure 2D,E, left panels). We next tested TAPI-2, a broad-spectrum inhibitor of tumour necrosis factor-alpha converting enzyme (TACE) and a potent E-cadherin cleaving protease [36]. However, TAPI-2 also failed to reverse the elongated cell shape and E-cadherin processing observed upon nutrient starvation (Figure 2D,E, middle panels). Finally, we focused on the ubiquitin–proteasome system (UPS), which is strongly implicated in cell survival during nutrient deprivation [37]. To test whether the UPS is involved in starvation-induced p-EMT, we used MG132 to inhibit proteasomal degradation. Interestingly, MG132 treatment of HCT116 cells reverted the elongated cell shape to the cobblestone shape characteristic of epithelial cells and prevented the cleavage of E-cadherin (Figure 2D,E, right panels). All inhibitor concentrations were optimised for HCT116 cells in separate crystal violet assays to ensure that non-toxic concentrations were used for the treatments (supplementary material, Figure S3A–C). Next, we investigated the role of MG132 in E-cadherin trafficking and analysed the subcellular localization of E-cadherin by immunofluorescence. We observed an upregulation and internalisation of E-cadherin in starved HCT116 cells, which was reversed by MG132 treatment (Figure 2F). Taken together, our results strongly suggest a critical role of the UPS in the starvation-induced processing of E-cadherin.

A p-EMT-associated gene signature has prognostic value for CRC patients

Next, we investigated whether starvation resulted in changes in the expression of EMT-related genes. We performed a qPCR array analysis of HCT116 cell RNA, which included 84 EMT-related genes (supplementary material, Figure S4A, and Table S5). After starvation, 21 out of 84 EMT-related genes were found to be significantly deregulated ($p < 0.05$, 1-fold regulation), with 15 of them being associated with a ≥2-fold expression change (Figure 3A,B and supplementary material, Figure S4B).
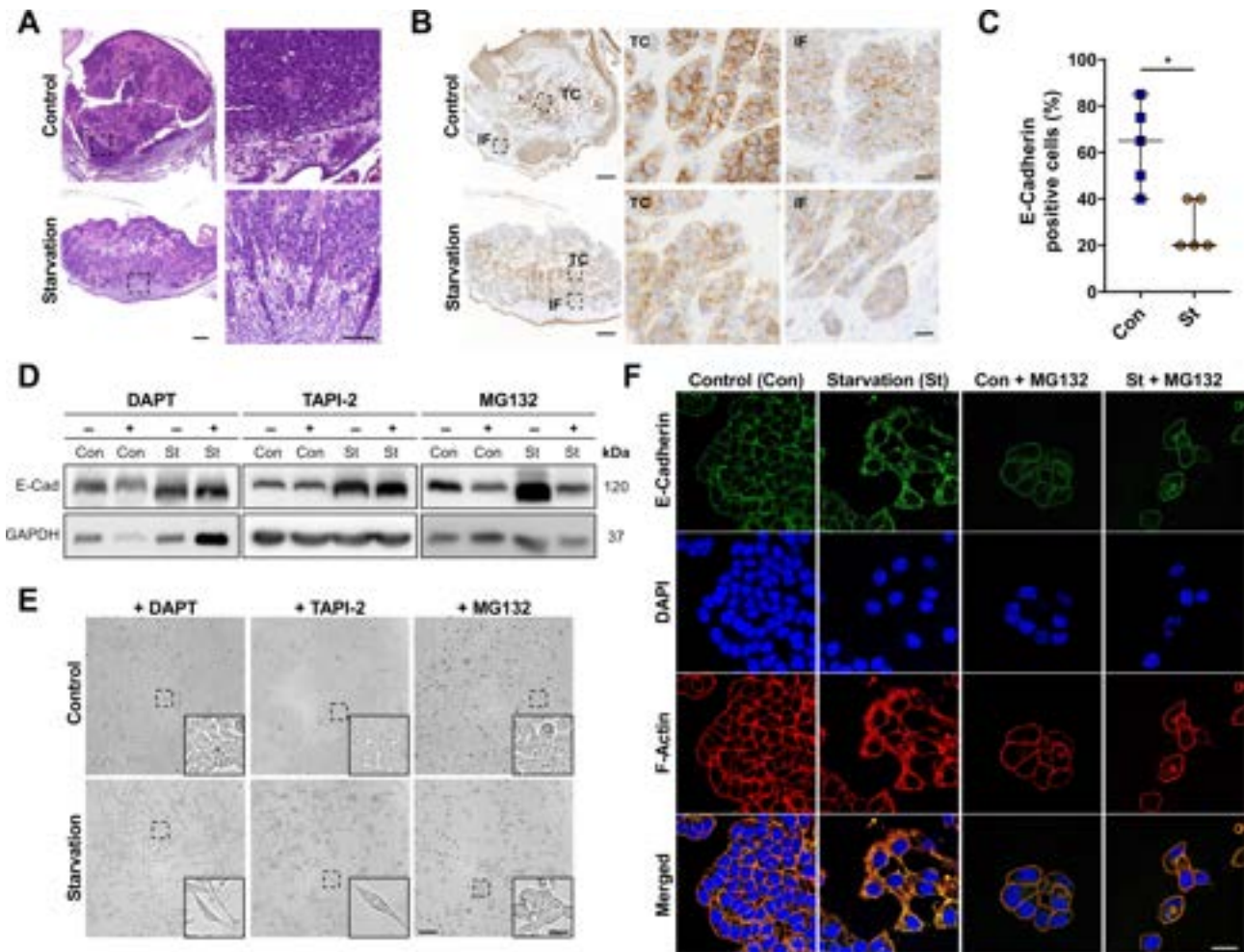
Figure 2. Starvation-induced modifications of E-cadherin expression pattern *in vivo* and *in vitro*. (A) H&E staining of control and starved HCT116-derived CAM ovografts showing collective pushing of control cells and an invasive growth pattern of starved cells (all images 10% brightness, 10% contrast). Overview image scale bar, 200 μm; enlarged section scale bar, 50 μm. (B) Immunohistochemistry (IHC) for E-cadherin in non-starved HCT116-derived control and starved ovografts. Representative sections of ovograft tumour centres (TCs) and invasion fronts (IFs) were enlarged (all images 15% contrast). Overview image scale bar, 200 μm; enlarged sections scale bar 20 μm. (C) Quantification of E-cadherin-positive tumour cells (TC) in control (Con; blue) and starved (St; yellow) ovografts shows a decrease of E-cadherin in starved tumour cells grown on the CAM. Data are presented as median with 95% confidence interval (*$p < 0.05$, Mann–Whitney test). (D) Representative western blots for E-cadherin in HCT116 cells with untreated 48 h control (Con) and 48 h starvation (St). DAPT (25 μM) and TAPI-2 (20 μM) treatment did not prevent the mobility shift of the E-cadherin band in starved cells. MG132 (proteasome inhibitor, 150 nM) prevented the processing of E-cadherin after starvation. (E) Representative images of 48 h starved HCT116 cells and their corresponding controls. DAPT (25 μM) and TAPI-2 (20 μM) treatment did not prevent the starvation-induced p-EMT phenotype. MG132 (proteasome inhibitor, 150 nM) prevented spindle-cell formation and partly retained intercellular adhesion. Overview image scale bar 100 μm; insert scale bar 25 μm. (F) Representative confocal images of HCT116 cells with E-cadherin (green), DAPI (blue), and phalloidin (F-Actin; red) in 48 h control and 48 h starvation condition showing upregulated, intracellular E-cadherin upon starvation which was prevented by MG132 (proteasome inhibitor, 150 nM) treatment (all images +40% brightness). Scale bar, 25 μm. Images were taken using a Zeiss laser scanning microscopy system (×63 objective; confocal, LSMT-PMT Observer ZI, LSM 710).

Consistent with previously described EMT gene expression patterns, *IL1RN* and *TSPAN13* were repressed by starvation, whereas microtubule-associated protein 1B (*MAP1B*), plasminogen activator inhibitor-1, PAI-1 (*SERPINE1*), *EGFR*, and SRY-box transcription factor 10 (*SOX10*) levels were significantly upregulated. Interestingly, the expression of EMT drivers such as *s*ecreted phosphoprotein 1 (*SPP1*), secreted protein acidic and rich in cysteine (*SPARC*), Gem nuclear organelle associated protein 2, *SIP1* (*GEMIN2*), and regulator of G-protein signalling 2 (*RGS2*) were upregulated, whereas the levels of the EMT repressors *CDH1* (E-cadherin), Occludin (*OCLN*), and Tissue

Factor Pathway Inhibitor 2 (*TFPI2*) were also increased. These data further support the induction of a p-EMT phenotype upon starvation in HCT116 cells. The relationship between the deregulated genes is shown as a STRING analysis map (Figure 3C).

We analysed the correlation of the significantly differentially expressed genes identified from the qPCR array with disease-free survival (DFS) in the GSE39582 dataset using Cox regression analysis [38] and then used machine learning [39] to construct a multigene signature (*EGFR*, *SERPINE1*, and *SOX1*0) using regularised Cox regression (supplementary material, Table S7). We used DFS as clinical endpoint with the aim of excluding other
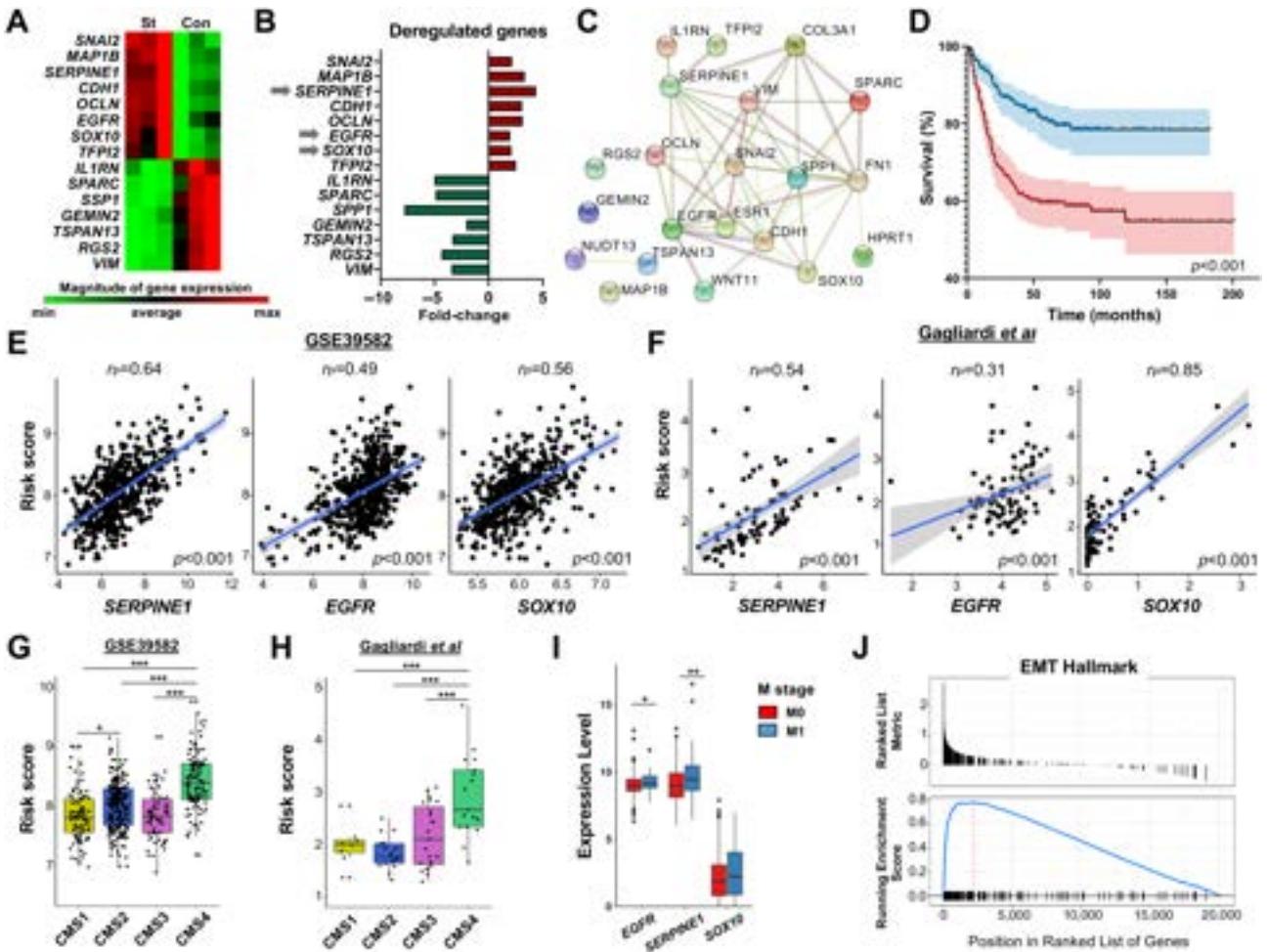
Figure 3. A gene signature consisting of *EGFR*, *SERPINE1*, and *SOX10* that is derived from starvation-induced p-EMT identifies high-risk CRC patients. (A) Heatmap showing significant deregulation of genes following 48 h of starvation (St) of HCT116 cells compared to control condition (Con) in the RT2 Profiler PCR Array. Relative upregulation (red) and downregulation (green) values are expressed as saturation. In total, 84 EMT-related genes were tested. (B) Fold-change values portraying the alteration of gene expression of 48 h starved HCT116 cells in relation to non-starved controls. (C) STRING analysis (version 11.0) displaying the significantly deregulated genes derived from RT2 Profiler PCR Array (fold-regulation ±2) analysis. Housekeeping genes were excluded. (D) Kaplan–Meier plot reporting the difference in disease-free survival (DFS) of patients classified into high-risk ($n = 259$) and low-risk ($n = 260$) groups according to the three-gene signature (red: high risk, blue: low risk, HR: 2.578, log-rank test $p < 0.0001$). (E) Correlations of signature genes with risk score for GSE39582 ($n = 519$). Pearson correlation coefficients (r_p) are shown on title of each plot separately. (F) Correlations of signature genes with risk score for the dataset from Gagliardi *et al* ($n = 76$) [26]. Pearson correlation coefficients (r_p) are shown on title of each plot separately. (G) Risk score values and the corresponding CMS extracted from the GSE39582 dataset ($n = 475$). Overall significance was determined by One-Way ANOVA ($p < 0.001$) followed by Tukey's *post hoc* test (significance bars). (H) Risk score values and the corresponding CMS extracted from the Gagliardi *et al* dataset ($n = 76$) [26]. Overall significance was determined by one-way ANOVA ($p < 0.001$) followed by Tukey's *post hoc* test (significance bars). (I) Expression levels of *EGFR*, *SERPINE1*, and *SOX10* in CRC tissues from patients stratified by metastatic status. Patients were derived from the colon adenocarcinoma (COAD) cohort of the EMTome database ($n = 445$, M0 = 325 and M1 = 65 samples respectively; 56 samples had no M-stage classification). A *t*-test was used to test significance (significance bars). (J) GSEA enrichment plot representing EMT hallmark obtained from GSEA analysis of high- versus low-risk score (supplementary material, Table S9 shows results for all the hallmarks). Significance levels: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

external factors due to the slow progress of colon cancer disease as highlighted by Delgado and Guddati [40]. All three genes were also listed as mesenchymal genes in the EMT portal dataset [41].

The risk score generated by the coefficients of *SERPINE1*, *EGFR*, and *SOX10* allowed us to calculate a risk score for each patient, and survival analysis showed significant DFS difference between high- and low-risk groups based on the median of the risk score (Figure 3D, and supplementary material, Figure S5A–C). A similar survival difference using the same model was

replicated on GSE17536 and TCGA-COAD cohorts (supplementary material, Figure S5D,E), while we saw no survival difference in the fourth dataset (RNA-Seq dataset [31], results not shown). The performance of the machine learning model on the training set, the test set, and the entire GSE39582 dataset is shown in supplementary material, Figure S5F–H, and that of GSE17536 and TCGA-COAD in supplementary material, Figure S5I,J, respectively. We found strong positive correlations (Pearson, $p < 0.05$) between the expression of each signature gene and the risk score on all the

datasets tested (GSE39582, Figure 3E; RNA-Seq dataset [31], Figure 3F; GSE17536; supplementary material, Figure S6A; TCGA COAD, supplementary material, Figure S6B).

Dependency analysis showed that the classification of patients into the two risk groups based on the multigene signature was independent of classical oncogene drivers such as p53, KRAS, or BRAF (supplementary material, Tables S8 and S9), but it was significantly correlated with the consensus molecular subtype (CMS) classification in all patient cohorts, in particular with CMS4 class (Figure 3G,H, and supplementary material, Figure S6C,D). Data from the EMTome portal showed that the expression of two out of the three genes used to build the signature was associated with metastasis, while *SOX10* did not reach significance (Figure 3I).

Differential expression followed by GSEA on the GSE39582 dataset showed EMT, angiogenesis, and hypoxia among the most enriched hallmarks in the high-risk group (Figure 3J and supplementary material, Table S10). Finally, we performed an *in silico* analysis correlating the expression of the four stemness markers *CD44*, *OCT4*, *CD133*, and *NANOG* with the risk groups from the GSE39582 dataset. Interestingly, there were significant but only weak inverse correlations for *CD133* and *NANOG* (*CD133*: $r_P = -0.19$, $p < 0.001$; *NANOG*: $r_P = -0.21$, $p < 0.001$) and no correlation for *CD44* and *OCT4* (*CD44*: $r_P = -0.06$, $p = 0.14$; *OCT4*: $r_P = 0.06$, $p < 0.19$), suggesting that the triple mesenchymal gene signature is a highly dominant driver signature for worse DFS.

Combined *EGFR*, *SERPINE1*, and *SOX10* upregulation upon starvation is a general feature in CRC cell lines and suppressed by UPS inhibition

Given the critical upregulated gene signature in HCT116 cells and its relevance to disease outcome and CMS prediction in CRC patients, we investigated whether the same signature could be observed in other CRC cell lines undergoing starvation-induced p-EMT. We determined by qPCR the expression of *EGFR*, *SERPINE1*, and *SOX10* in non-starved, starved, and replenished HCT116, DLD1, HT29, and LoVo cells (supplementary material, Figure S7A–C). As expected, starvation generally resulted in the upregulation of the signature genes, while subsequent replenishment resulted in downregulation to levels observed in non-starved controls, with only a few exceptions: *SERPINE1* gene expression was not upregulated by starvation in DLD1 cells, and *SOX10* and *EGFR* levels did not return to the non-starved state after replenishment of the starved cells with complete medium in HT29 and HCT116 cells, respectively.

Since we observed a reversal of EMT characteristics in HCT116 cells upon inhibition of UPS with MG132, we next examined the effect of proteasome inhibition on the triple-gene signature by qPCR. We observed that all three genes, *EGFR*, *SERPINE1*, and *SOX10*, were downregulated in starved HCT116 cells treated with MG132. A decreasing trend was observed for *SOX10*;

however, the difference did not reach statistical significance (supplementary material, Figure S8A–C). Thus, proteasome inhibition with MG132 in starved HCT116 cells not only reversed the p-EMT phenotype but also inhibited the expression of the signature genes.

EMT is known to be closely associated with chemoresistance in CRC [42]. Therefore, we further investigated whether the reversal of p-EMT with MG132 could increase the sensitivity of starved cells to 5-FU. As suggested by their phenotypic induction of p-EMT along with reduced proliferative capacity, starved HCT116 and HT29 cells were less sensitive to 5-FU in Crystal Violet and MTT assays (supplementary material, Figure S8D–G). Notably, the combination of 5-FU with MG132 attenuated EMT-associated drug resistance and increased the relative 5-FU toxicity in starved HCT116 and HT29 cells (supplementary material, Figure S8D–G). Thus, inhibition of UPS via MG132 prevents starvation-induced p-EMT initiation, preserves epithelial characteristics, and thus sensitises CRC cells to 5-FU treatment.

The starvation-induced triple-gene signature triggers an invasive tumour growth pattern *in vivo*

Next, we analysed the expression of proteins of our triple-gene signature and their spatial distribution by immunohistochemistry *in vivo* in the CAM model. PAI-1, which is encoded by the *SERPINE1* gene and physiologically involved in extracellular matrix processing [43], was upregulated at the tumour invasion front in control ovografts. Consistent with the increased *SOX10* expression *in vitro*, the nuclear staining of SOX10 was increased in starved ovografts compared to their non-starved counterparts (Figure 4A,B). Moreover, CAM ovografts derived from prestarved HCT116 cells showed strong PAI-1 expression throughout the whole tumour (Figure 4C,D) and showed a remarkable reduction of collagen fibres using Picro-Sirius red staining (Figure 4E). Using western blotting, we also showed a clear upregulation of PAI-1 in all three cell lines upon starvation (Figure 4F). In contrast to the increased expression of *EGFR* observed at the mRNA level *in vitro*, the protein levels of EGFR were found decreased in starved cells *in vivo* in CAM ovografts and *in vitro* using western blotting (Figure 5A–C), along with a loss of membrane association of EGFR (Figure 5A). In control ovografts, the strong EGFR membrane association was visible but reduced at the invasion front (stars), whereas in ovografts derived from prestarved cells, we found a decrease in EGFR signals and an internalisation in the tumour centre and at the tumour margins seen as an accumulation of single dots in the cytosolic compartment (Figure 5D). Notably, a few cell clusters showed a complete loss of EGFR staining (black arrows) (Figure 5D). Similar to our observation in CAM ovografts, confocal microscopy revealed that the EGFR protein was internalised upon starvation in 2D culture of HCT116 and DLD1 cells, whereas it was translocated back to the membrane upon treatment of starved cells with
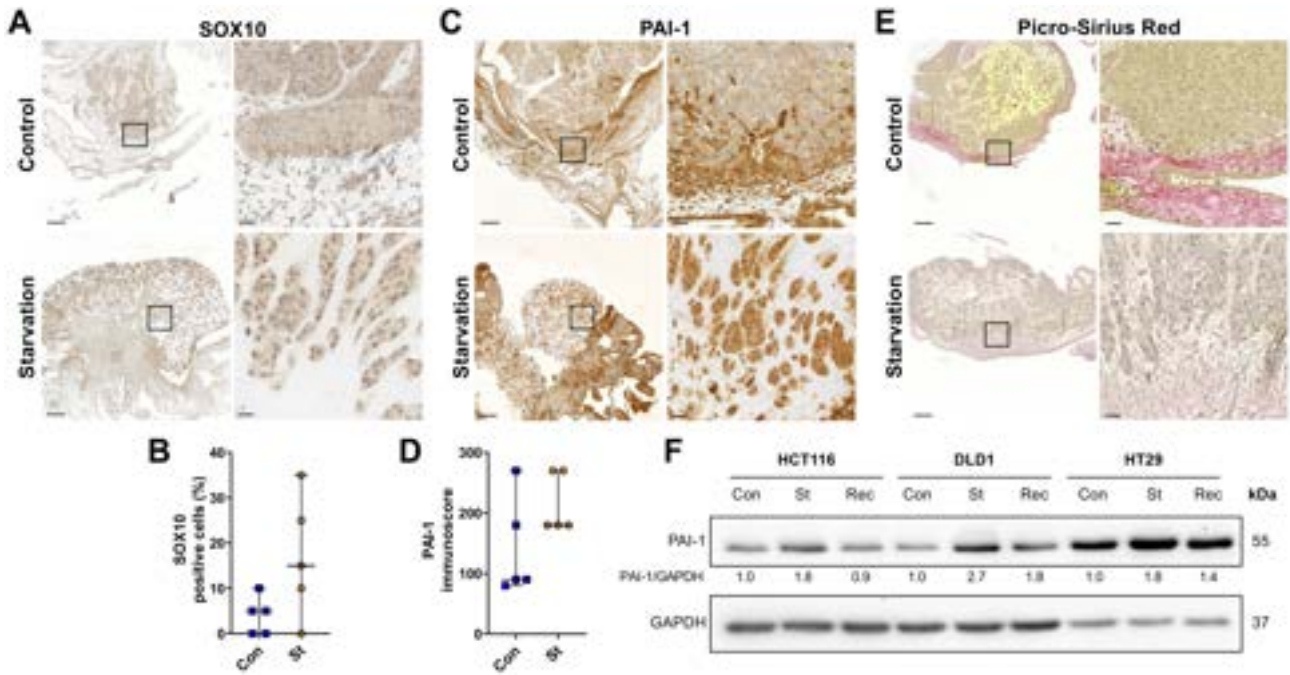
Figure 4. Immunohistochemistry (IHC) staining for the signature proteins SOX10 and PAI-1. (A) SOX10 IHC staining showed increased nuclear staining in starved HCT116-derived CAM ovografts. Overview image scale bar, 200 μm; enlarged section scale bar, 20 μm. (B) Quantification of SOX10-positive tumour cells (tumour centre) in control (blue) and starved (yellow) ovografts shows increase of SOX10 in starved tumour cells grown on the CAM. Data are presented as median with 95% confidence interval (ns., Mann–Whitney test). (C) PAI-1 (encoded by *SERPINE1*) IHC staining showing ubiquitous upregulation of PAI-1 in starved HCT116-derived CAM ovografts compared to margin-accentuated expression in controls. Overview image scale bar, 200 μm; enlarged section scale bar, 20 μm. (D) PAI-1 histoscore of tumour cells (tumour centre) in control (blue) and starved (yellow) ovografts shows increase of PAI-1 in starved tumour cells grown on the CAM. Data are presented as median with 95% confidence interval (ns., Mann–Whitney test). (E) Picro-Sirius red staining reveals degraded extracellular collagen in starved HCT116-derived ovografts. Overview image, scale bar 200 μm; enlarged section, scale bar 20 μm. (F) Representative western blots for PAI-1 in control (Con), starved (St), and recovered (Rec) HCT116, DLD1, and HT29 cells show an upregulation of PAI-1 upon starvation.

MG132 (Figure 5E,F). Thus, increased *EGFR* transcription may ensure the availability of the EGFR protein, although post-translational regulatory mechanisms appear to determine the actual fate of the protein.

Juxtacrine EGFR signalling controls nutrient-efficient proliferation in starvation-induced p-EMT

Since EGFR is a key regulator of cell proliferation [44], we aimed to further investigate its role in starvation-induced p-EMT. First, we examined by western blotting the protein levels of two target genes of the EGFR signalling cascade, total Cyclin D1 and c-Myc. Indeed, both markers were repressed upon starvation (Figure 6A), and their levels increased to the control levels when the cells recovered from nutrient stress in all three cell lines (Figure 6A).

To evaluate the role of EGFR internalisation in starved cells, we monitored the proliferation of these cells by live imaging microscopy in HCT116 and HT29 cells. We observed that the starved cells underwent cell division mostly after cell–cell contact between individual mesenchymal cells (Figure 6B,C and supplementary material, Movies S1–S3). Examples of these 'EMT kisses' after starvation are shown in Figure 6C,D. Interestingly, from time to time in starved HCT116 cells, we observed the phenomenon of asymmetric cell division, which was described as the predominant cell division mode for progenitor and stem cells [45]. Here, one daughter cell showed the mesenchymal phenotype, while the other one was round and non-proliferating (Figure 6E and supplementary material, Movie S4).

Taken together, we suggest that the spatial regulation of EGFR with internalisation and recycling plays an important role during starvation-induced p-EMT and provides nutrient-depleted tumour cells with an extraordinarily high plasticity despite the limited availability of nutrients. The general working model of our findings is shown in Figure 6F.

Discussion

Recently, partial or intermediate EMT phenotypes with co-expression of epithelial and mesenchymal markers have attracted much attention due to their high level of plasticity and importance for metastasis [10,34,46]. However, it is still unclear why tumour cells initiate EMT but stop at a certain stage during the transition from epithelial to mesenchymal phenotype. Both endogenous and extrinsic stromal/extracellular matrix-associated drivers of p-EMT are poorly understood, making
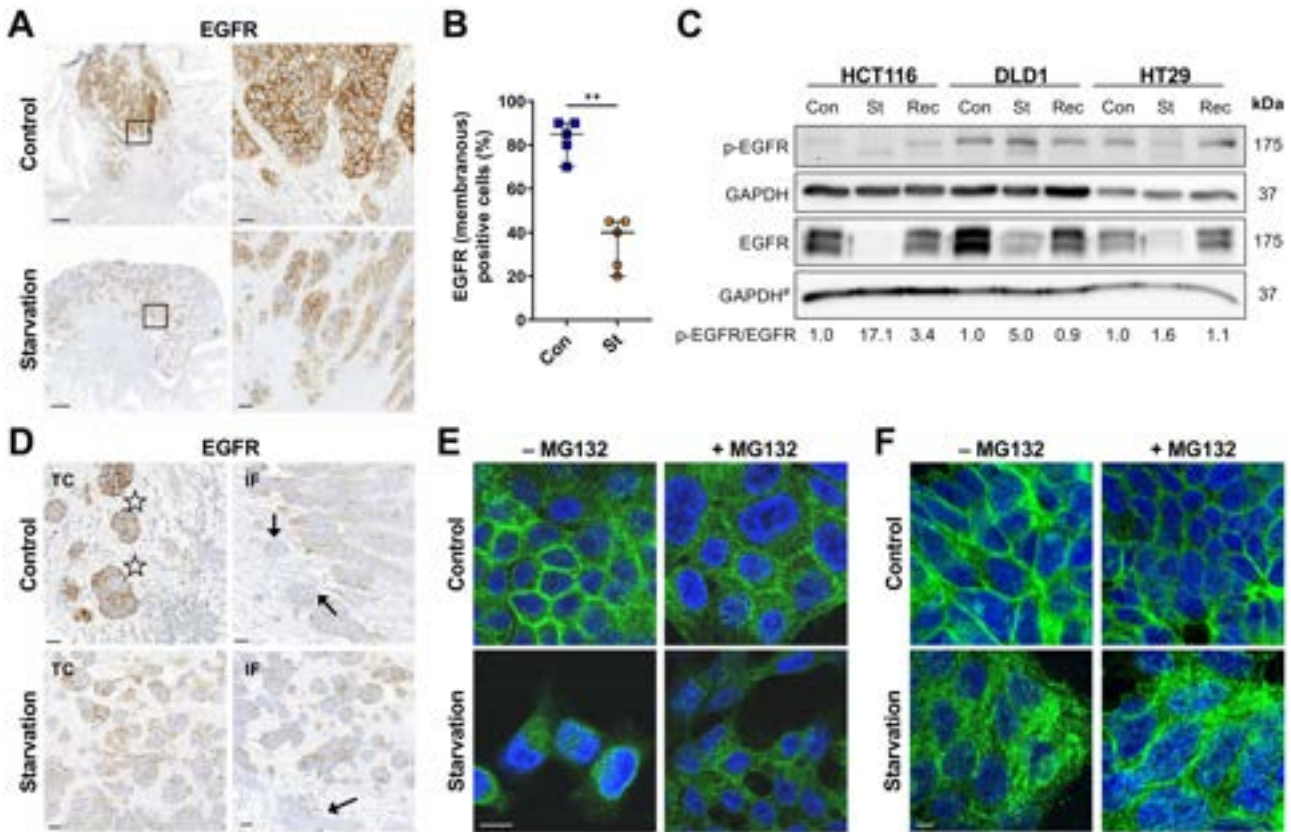
Figure 5. Alteration of EGFR expression and subcellular localisation depending on MG132. (A) EGFR IHC staining showing a loss of EGFR membrane association in starved HCT116-derived CAM ovografts. Overview image scale bar, 200 μm; enlarged section scale bar, 20 μm. (B) Quantification of EGFR-positive tumour cells (tumour centre) in control (blue) and starved (yellow) ovografts shows a decrease of EGFR in starved tumour cells grown on CAM. Data are presented as median with 95% confidence interval (**$p < 0.01$, Mann–Whitney test). (C) Representative western blots for p-EGFR and EGFR in control (Con; untreated 48 h), starved (St; 48 h), and recovered (Rec; 48 h starved plus 24 h recovery) conditions. Western blots depict the degradation/internalisation of EGFR during starvation. #GAPDH of EGFR is the same as used in Figure 6A. (D) Representative sections of centre and invasion front from HCT116-derived ovografts. While EGFR is strongly expressed at the tumour centre of controls (stars) and lost at the front (black arrows), starved tumours show a ubiquitous reduction of EGFR staining. Overview image scale bar, 200 μm; enlarged section scale bar, 20 μm. Black arrows represent cell clusters with complete loss of EGFR staining. (E) Representative confocal images of HCT116 cells with EGFR (green) and DAPI (blue) staining in 48 h control and 48 h starvation with and without MG132 treatment (150 nM) (all images +20% brightness, +25% contrast). Scale bar, 10 μm. Images were taken using a Stellaris 8 confocal microscope (×63 objective, Leica). (F) Representative confocal images of DLD1 cells with EGFR (green) and DAPI (blue) staining in 48 h control and 48 h starvation with and without MG132 treatment (150 nM) (all images +25% brightness, +5% contrast). Scale bar, 10 μm. Images were taken using a CQ1 Confocal Imaging Cytometer (×40 objective; Yokogawa).

therapeutic targeting of tumour cells undergoing p-EMT a major challenge [47].

The rationale of our study was to model tumour cell escape under nutrient-depleted conditions. Such metabolically challenging conditions occur at many stages of tumour development (i.e. when the microenvironment is changing, in the circulation, in the peritoneal cavity, in the unprepared extracellular matrix (ECM) environment of a distant organ, or at the tumour invasion front) [48-50]. Hypothetically, cancer cells exposed to nutrient-deficient conditions could initiate the p-EMT program to enhance mobility and 'escape' from intratumoural nutrient deprivation. Aggressive tumour cells not only need to escape from nutrient deprivation but also have to survive, maintain their cell population, stop proliferating, and adapt, which could lead to the activation of a transdifferentiation program that relies on the high cellular plasticity of cancer cells. Epigenetic alterations seem to be crucial for the reversibility of EMT

phenotypes [51]. We have identified nutrient deprivation in CRC cells as a trigger of p-EMT and aggressive growth behaviour. However, our observation window was limited regarding the time-scale of starvation and the reversibility from p-EMT since cells will die when extending the starvation trigger.

It is hotly debated whether EMT is accompanied by the acquisition of stem-like properties. Cell division is critical for gaining stemness during EMT; stem cell properties can also develop later during the time frame of EMT induction. Similarly, when removing the EMT-inducing trigger, mesenchymal features are known to disappear faster than stem cell properties. This might be one explanation on why HT29 cells showed an increase in stem-like properties with a time delay only during the recovery phase, when the cell cycle had restarted.

In contrast to previous studies linking EMT with the downregulation of the epithelial marker E-cadherin via master EMT transcription factors, we observed
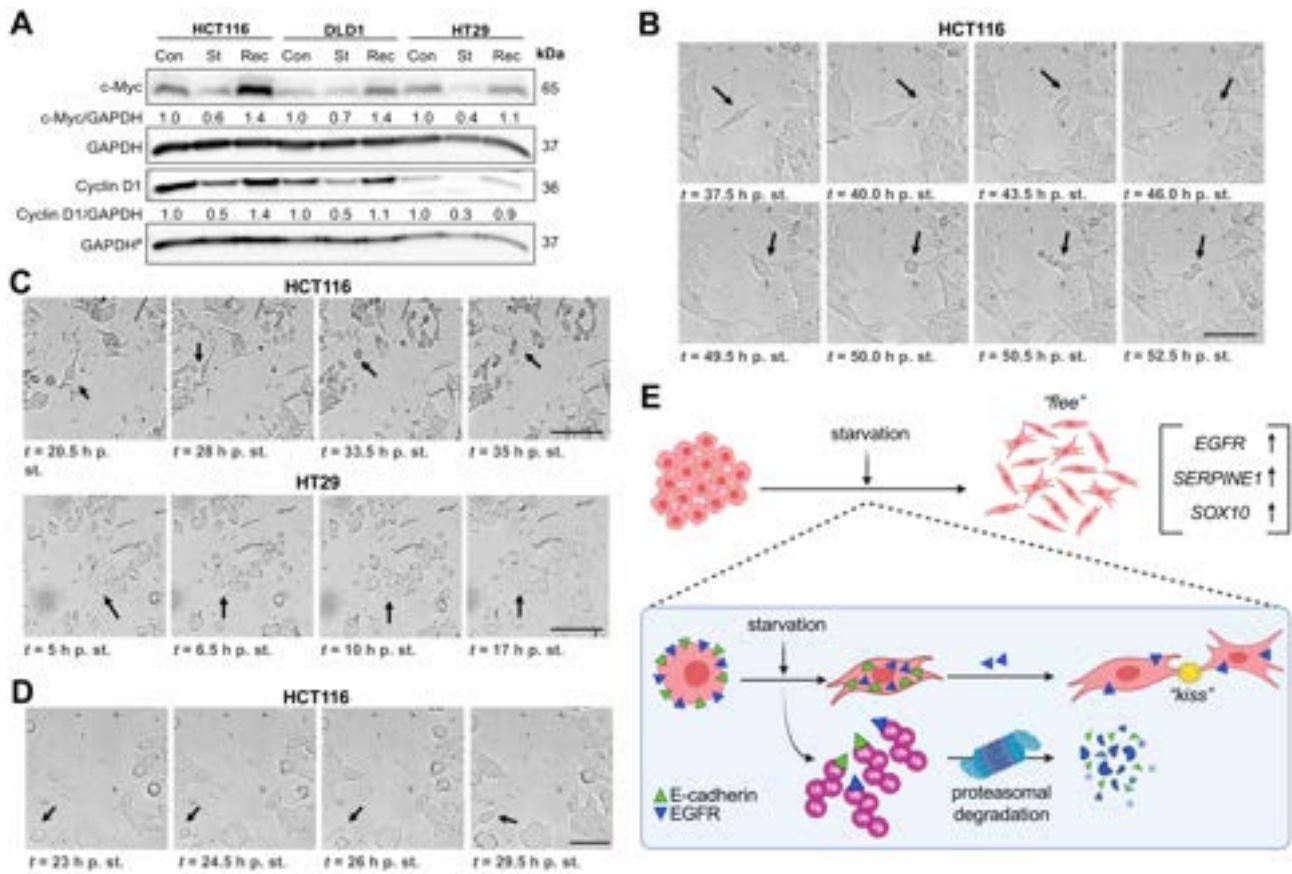
Figure 6. Juxtacrine EGFR signalling controls nutrient-efficient proliferation in starvation-induced p-EMT. (A) Representative western blots for proliferation-associated proteins Cyclin-D1 and c-Myc in HCT116, DLD1, and HT29 cells with untreated 48 h control (Con), 48 h starvation (St), and recovery (Rec; 48 h starved plus 24 h recovery) conditions. #GAPDH of Cyclin D1 is the same as used in Figure 5C. (B) Snapshots of a movie of starved HCT116 cells (supplementary material, Movie S1). Transdifferentiated HCT116 cells divide upon cell contact (black arrows). Scale bar 50 μm. (C) Snapshots of a movie of starved HCT116 (upper panels) and HT29 cells (lower panels; supplementary material, Movies S2 and S3). Transdifferentiated HCT116 and HT29 cells divide upon cell contact. Scale bar, 50 μm. (D) Snapshots of a movie of starved HCT116 cells presenting an asymmetric cell division (supplementary material, Movie S4), which results in one rounded and one elongated spindle-like cell (black arrows). (E) Schematic model of starvation-mediated processing of E-cadherin and EGFR. Under starvation conditions, epithelial tumour cells 'flee' from the nutrient-restricted environment by acquiring a distinct triple-gene signature and a mesenchymal phenotype. Membranous E-cadherin (green triangles) is degraded by the proteasome, resulting in processed cytosolic E-cadherin (green trapezia). Membranous EGFR (blue triangles) is severely degraded by the proteasome, thereby shifting EGFR from the membrane to the cytosol and further decreasing its availability. Upon intercellular contact ('EMT kiss'), juxtacrine EGFR signalling occurs, thereby providing a cell proliferation stimulus. Ub: ubiquitin residues. Created with BioRender.com.

proteasome-dependent E-cadherin processing that was highly dependent on nutrient availability. The identified p-EMT-associated triple-gene signature (including *EGFR*, *SERPINE1*, and *SOX10*) had high prognostic relevance and could be a marker for the identification of a well-adapted and highly aggressive tumour subpopulation.

Many attempts to classify the different EMT transdifferentiation states have been reported. The seminal paper by Tan *et al* described a generic EMT score to assess the extent of EMT for tumours and cancer cell lines [34]. Recently, Vasaikar *et al* released the EMTome portal, a database of EMT signatures derived from multiple sources, such as cell lines, patient-derived xenografts, or patients [41]. In parallel to these broad and useful analytical resource, we report a nutrient-dependent p-EMT signature. Complementary to classical gene expression analysis, our approach of linking p-EMT to a gene signature consisting of *EGFR*, *SERPINE1*, and

SOX10 associates this phenomenon to the prognosis of CRC patients. All three genes have been listed as mesenchymal genes in the EMTome database. In addition, SOX10 is known to be an essential marker for heterogeneity and stemness in melanoma, affecting cell-type specification and lineage differentiation [52,53]; however, its role in CRC is not yet known. Serpine1, also known as plasminogen activator inhibitor 1 (PAI-1), is a 45 kDa glycoprotein with a cleavage site for uPA. Upon docking of uPA to its receptor (uPAR), PAI-1 promotes the internalisation of this complex by binding to the endocytosis machinery [54]. PAI-1 has multiple functions in cancer but has been mostly associated with a poor clinical outcome and chemotherapy resistance [55]. High PAI-1 expression was associated with tumour progression in an azoxymethane dextran sulphate-sodium model of colorectal carcinogenesis [56]. In support of these findings, we observed strong PAI-1 expression at the tumour invasion front in CAM ovografts.

In this study, nutrient deprivation caused an increase in E-cadherin protein levels. However, the function of E-cadherin as a cell adhesion molecule was lost, as evidenced by impaired intercellular adhesion. For the first time, we have identified the proteasome as an essential regulator of starvation-induced p-EMT. Proteasome inhibition by MG132 treatment prevented the processing and cytosolic translocation of E-cadherin and averted phenotypic mesenchymal differentiation. This seems to provide the tumour cells with extremely high plasticity compared to classical EMT, where E-cadherin must be synthesised in the reverse process of mesenchymal-to-epithelial transition. This novel proteasome-mediated early EMT mechanism is proposed to be an effective alternative to the well-known transcriptional repression of E-cadherin [57]. Another accompanying mechanism appears to be the downregulation and internalisation of the EGFR protein. Although we observed an increase in *EGFR* mRNA expression *in vitro* as a member of the triple-gene signature, the *in vivo* staining of EGFR showed a controversial picture. In particular, at the tumour invasion front, EGFR membrane association was decreased, and EGFR appeared to be located in the cytosolic compartment. Previous studies showed that nutrient deprivation leads to a minimization of cellular metabolism in the terms of energy efficiency [58]. Since EGFR confers proliferation and metabolic activity, sustained activation of the receptor during starvation in cancer cells may lead to their death. It was reported that epithelial cells that are not able to stop the cell cycle show massive aneuploidy due to errors in cytokinesis [59]. These mitotic defects were shown to disappear after withdrawal of the EMT trigger transforming growth factor beta (TGFβ), but genomic alterations persisted and further increased tumour aggressiveness [59]. Comaills *et al* (2016) described this phenomenon as 'incompatibility of EMT and cellular proliferation' [59]. Accordingly, they found a higher proportion of multinucleated cells in mesenchymal circulating tumour cells (CTC) compared to epithelial CTCs of metastatic breast cancer. Interestingly, such mesenchymal CTCs were characterised by high PAI-1 expression, a marker we have identified in starvation-induced p-EMT. Partial EMT was also described in other pathological conditions, such as chronic kidney disease, where the partial EMT program seemed to arrest injured damaged tubular epithelial cells in the G2 phase, limiting their repair capacity and regeneration and finally leading to chronic fibrosis [60]. Finally, the observed reduction in EGFR downstream signalling may indicate a transient reduction in proliferative signalling in favour of quiescence, drug resistance, and survival. Increased *EGFR* transcription upon nutrient deprivation may ensure the replenishment of the EGFR protein.

Another interesting, but rather coincidental, observation of the current study was cell division upon intercellular contact. Obviously, this is juxtacrine EGFR signalling, which has been described as an energy-efficient proliferation mechanism [61]. Therefore, we suggest that the spatial regulation of EGFR with internalisation and recycling most likely further stabilises the p-EMT status. While lysosomal degradation is known to attenuate EGFR signalling, recycling of the receptor to the plasma membrane may prolong the signal [62].

We have accidentally observed asymmetric cell division upon starvation. This phenomenon has been described as a self-renewal mechanism during stem cell proliferation [45]. One daughter cell commits to the cell fate to be a mesenchymal cell, while the other daughter cell maintains the stem cell identity [63]. This theory could explain the high phenotypic heterogeneity despite similar gene expression levels [64,65]. Taken together, our data paint a highly complex and multifaceted picture of starvation-induced p-EMT.

Our strategy to define a starvation-induced p-EMT signature could be clinically translated to predict the DFS of CRC patients. The triple-gene signature of upregulated *EGFR*, *SERPINE1*, and *SOX10* identifies cancer cells that undergo p-EMT and are therefore at a high risk of metastasis. This further supports the association between p-EMT and poor disease outcome described in previous studies [47]. In contrast, Zheng *et al* (2015) described in pancreatic cancer that EMT suppression does not change metastasis capability of primary tumours, but it sensitises the tumours to gemcitabine treatment, both *in vitro* and *in vivo*. This effect was suggested to be caused by the increased proliferation after EMT suppression [66]. In our study, we observed an attenuated proliferation during the starvation-induced p-EMT differentiation program. We suggest that the reduced proliferation made the cells less sensitive to 5-FU, a clinically used drug that needs replication forks (i.e. proliferative cells) to be effective. Furthermore, tumours from high-risk CRC patients identified by the proposed gene signature were strongly associated with the CMS4 category, which is characterised by mesenchymal gene expression, therapy resistance, and development of distant metastases [28]. Although the CMS classification is of evident prognostic relevance, its predictive benefit has not yet been translated into clinical practice [67].

Finally, the spatial view of the p-EMT signature revealed a completely novel p-EMT regulatory mechanism involving proteasomal processing and internalisation of E-cadherin and EGFR. Our risk-predictive triple-gene signature may be an important piece of the puzzle in understanding the early steps of CRC metastasis.

Acknowledgements

Author contributions statement

GAP conceived, carried out experiments, analysed data and wrote the first draft of the manuscript. IS and SB conceived and performed prognostic bioinformatic analysis. KH performed experimental work and helped in the generation of figures and with original data archiving. GF, BP and AN provided the second validation cohort and the bioinformatic analysis. PK, FG, FS, AEGT, CH, HHH, AM and GO carried out experiments and analysed data. RP helped with acquiring Leica Stellaris images and post-imaging processing. SB, BP and AN conceived of the experiments and critically reviewed the final version of the manuscript. AH and KE-W helped with immunohistochemical analysis. RS-S conceived and supervised the study and finalised the manuscript. All authors had final approval of the submitted and published versions.

Data availability statement

The qPCR array data are listed fully in the supplementary material. Additional methodological details will be available upon request.

References

1. Chen J, Zheng Y, Wang H, *et al*. Cause of death among patients with colorectal cancer: a population-based study in the United States. *Aging (Albany NY)* 2020; **12:** 22927–22948.

2. Nieto MA, Huang RY, Jackson RA, *et al*. Emt: 2016. *Cell* 2016; **166:** 21–45.

3. Serrano-Gomez SJ, Maziveyi M, Alahari SK. Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications. *Mol Cancer* 2016; **15:** 18–31.

4. Busch EL, McGraw KA, Sandler RS. The potential for markers of epithelial-mesenchymal transition to improve colorectal cancer outcomes: a systematic review. *Cancer Epidemiol Biomarkers Prev* 2014; **23:** 1164–1175.

5. Semb H, Christofori G. The tumor-suppressor function of E-cadherin. *Am J Hum Genet* 1998; **63:** 1588–1593.

6. Na TY, Schecterson L, Mendonsa AM, *et al*. The functional activity of E-cadherin controls tumor cell metastasis at multiple steps. *Proc Natl Acad Sci U S A* 2020; **117:** 5931–5937.

7. Futterman MA, Garcia AJ, Zamir EA. Evidence for partial epithelial-to-mesenchymal transition (pEMT) and recruitment of motile blastoderm edge cells during avian epiboly. *Dev Dyn* 2011; **240:** 1502–1511.

8. Aiello NM, Maddipati R, Norgard RJ, *et al*. EMT subtype influences epithelial plasticity and mode of cell migration. *Dev Cell* 2018; **45:** e684.

9. Pang QY, Tan TZ, Sundararajan V, *et al*. 3D genome organization in the epithelial-mesenchymal transition spectrum. *Genome Biol* 2022; **23:** 121.

10. Aiello NM, Kang Y. Context-dependent EMT programs in cancer metastasis. *J Exp Med* 2019; **216:** 1016–1026.

11. Hao Y, Baker D, Ten Dijke P. TGF-β-mediated epithelial-mesenchymal transition and cancer metastasis. *Int J Mol Sci* 2019; **20:** 2767.

12. Kariya Y, Oyama M, Suzuki T, *et al*. αvβ3 integrin induces partial EMT independent of TGF-β signaling. *Commun Biol* 2021; **4:** 490.

13. Jia D, Park JH, Jung KH, *et al*. Elucidating the metabolic plasticity of cancer: mitochondrial reprogramming and hybrid metabolic states. *Cell* 2018; **7:** 21.

14. Ramesh V, Brabletz T, Ceppi P. Targeting EMT in cancer with repurposed metabolic inhibitors. *Trends Cancer* 2020; **6:** 942–950.

15. Warburg O. Versuche an überlebendem carcinom-gewebe. *Klin Wochenschr* 1923; **2:** 776–777.

16. Martinez AF, McCachren SS 3rd, Lee M, *et al*. Metaboloptics: visualization of the tumor functional landscape via metabolic and vascular imaging. *Sci Rep* 2018; **8:** 4171.

17. Chen G, Liu H, Zhang Y, *et al*. Silencing PFKP inhibits starvation-induced autophagy, glycolysis, and epithelial mesenchymal transition in oral squamous cell carcinoma. *Exp Cell Res* 2018; **370:** 46–57.

18. Gundamaraju R, Lu W, Paul MK, *et al*. Autophagy and EMT in cancer and metastasis: who controls whom? *Biochim Biophys Acta Mol Basis Dis* 2022; **1868:** 166431.

19. Georgakopoulos-Soares I, Chartoumpekis DV, Kyriazopoulou V, *et al*. EMT factors and metabolic pathways in cancer. *Front Oncol* 2020; **10:** 499.

20. Gillette MA, Satpathy S, Cao S, *et al*. Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. *Cell* 2020; **182:** 200–225.e35.

21. Castro F, Dirks WG, Fähnrich S, *et al*. High-throughput SNP-based authentication of human cell lines. *Int J Cancer* 2013; **132:** 308–314.

22. Huebner K, Erlenbach-Wuensch K, Prochazka J, *et al*. ATF2 loss promotes tumor invasion in colorectal cancer cells via upregulation of cancer driver TROP2. *Cell Mol Life Sci* 2022; **79:** 423.

23. Ranjan RA, Muenzner JK, Kunze P, *et al*. The chorioallantoic membrane xenograft assay as a reliable model for investigating the biology of breast cancer. *Cancers (Basel)* 2023; **15:** 1704.

24. Smith JJ, Deane NG, Wu F, *et al*. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology* 2010; **138:** 958–968.

25. Pardini B, Ferrero G, Tarallo S, *et al*. A fecal MicroRNA signature by small RNA sequencing accurately distinguishes colorectal cancers: results from a multicenter study. *Gastroenterology* 2023; **165:** 582–599.e8.

26. Gagliardi A, Francescato G, Ferrero G, *et al*. The 8q24 region hosts miRNAs altered in biospecimens of colorectal and bladder cancer patients. *Cancer Med* 2023; **12:** 5859–5873.

27. Liu J, Lichtenberg T, Hoadley KA, *et al*. An integrated TCGA pancancer clinical data resource to drive high-quality survival outcome analytics. *Cell* 2018; **173:** 400–416.e11.

28. Guinney J, Dienstmann R, Wang X, *et al*. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015; **21:** 1350–1356.

29. Ritchie ME, Phipson B, Wu D, *et al*. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; **43:** e47.

30. Wu T, Hu E, Xu S, *et al*. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb)* 2021; **2:** 100141.

31. Yu G, Wang LG, Yan GR, *et al*. DOSE: an R/bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* 2015; **31:** 608–609.

32. Yu G. Gene ontology semantic similarity analysis using GOSemSim. *Methods Mol Biol* 2020; **2117:** 207–215.

33. Muir A, Vander Heiden MG. The nutrient environment affects therapy. *Science* 2018; **360:** 962–963.

34. Tan TZ, Miow QH, Miki Y, *et al*. Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol Med* 2014; **6:** 1279–1293.

35. Dinicola S, Pasqualato A, Proietti S, *et al*. Paradoxical E-cadherin increase in 5FU-resistant colon cancer is unaffected during mesenchymal-epithelial reversion induced by γ-secretase inhibition. *Life Sci* 2016; **145:** 174–183.

36. Aghababaei M, Hogg K, Perdu S, *et al*. ADAM12-directed ectodomain shedding of E-cadherin potentiates trophoblast fusion. *Cell Death Differ* 2015; **22:** 1970–1984.

37. Park J, Cho J, Song EJ. Ubiquitin-proteasome system (UPS) as a target for anticancer treatment. *Arch Pharm Res* 2020; **43:** 1144–1161.

38. Marisa L, de Reyniès A, Duval A, *et al*. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med* 2013; **10:** e1001453.

39. Simon N, Friedman J, Hastie T, *et al*. Regularization paths for Cox's proportional hazards model via coordinate descent. *J Stat Softw* 2011; **39:** 1–13.

40. Delgado A, Guddati AK. Clinical endpoints in oncology – a primer. *Am J Cancer Res* 2021; **11:** 1121–1131.

41. Vasaikar SV, Deshmukh AP, den Hollander P, *et al*. EMTome: a resource for pan-cancer analysis of epithelial-mesenchymal transition genes and signatures. *Br J Cancer* 2021; **124:** 259–269.

42. Escalante PI, Quiñones LA, Contreras HR. Epithelial-mesenchymal transition and MicroRNAs in colorectal cancer chemoresistance to FOLFOX. *Pharmaceutics* 2021; **13:** 75–93.

43. Milenkovic J, Milojkovic M, Jevtovic Stoimenov T, *et al*. Mechanisms of plasminogen activator inhibitor 1 action in stromal remodeling and related diseases. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* 2017; **161:** 339–347.

44. Maity P, Chatterjee J, Patil KT, *et al*. Targeting the epidermal growth factor receptor with molecular degraders: state-of-the-art and future opportunities. *J Med Chem* 2023; **66:** 3135–3172.

45. Denisov EV, Perelmuter VM. A fixed partial epithelial-mesenchymal transition (EMT) triggers carcinogenesis, whereas asymmetrical division of hybrid EMT cells drives cancer progression. *Hepatology* 2018; **68:** 807–810.

46. Jolly MK, Boareto M, Huang B, *et al*. Implications of the hybrid epithelial/mesenchymal phenotype in metastasis. *Front Oncol* 2015; **5:** 155.

47. Jolly MK, Somarelli JA, Sheth M, *et al*. Hybrid epithelial/mesenchymal phenotypes promote metastasis and therapy resistance across carcinomas. *Pharmacol Ther* 2019; **194:** 161–184.

48. Coloff JL, Brugge JS. Coping with the metabolic stress of leaving home. *Cell Res* 2016; **26:** 757–758.

49. Garg M. Epithelial, mesenchymal and hybrid epithelial/mesenchymal phenotypes and their clinical relevance in cancer metastasis. *Expert Rev Mol Med* 2017; **19:** e3.

50. Recouvreux MV, Moldenhauer MR, Galenkamp KMO, *et al*. Glutamine depletion regulates slug to promote EMT and metastasis in pancreatic cancer. *J Exp Med* 2020; **217:** e20200388.

51. Jain P, Corbo S, Mohammad K, *et al*. Epigenetic memory acquired during long-term EMT induction governs the recovery to the epithelial state. *J R Soc Interface* 2023; **20:** 20220627.

52. Gambi G, Mengus G, Davidson G, *et al*. The LncRNA LENOX interacts with RAP2C to regulate metabolism and promote resistance to MAPK inhibition in melanoma. *Cancer Res* 2022; **82:** 4555–4570.

53. Pingault V, Zerad L, Bertani-Torres W, *et al*. SOX10: 20 years of phenotypic plurality and current understanding of its developmental function. *J Med Genet* 2022; **59:** 105–114.

54. Webb DJ, Nguyen DH, Sankovic M, *et al*. The very low density lipoprotein receptor regulates urokinase receptor catabolism and breast cancer cell motility in vitro. *J Biol Chem* 1999; **274:** 7412–7420.

55. Kubala MH, DeClerck YA. The plasminogen activator inhibitor-1 paradox in cancer: a mechanistic understanding. *Cancer Metastasis Rev* 2019; **38:** 483–492.

56. Öner MG, Rokavec M, Kaller M, *et al*. Combined inactivation of TP53 and MIR34A promotes colorectal cancer development and progression in mice via increasing levels of IL6R and PAI1. *Gastroenterology* 2018; **155:** 1868–1882.

57. Hu Y, Dai M, Zheng Y, *et al*. Epigenetic suppression of E-cadherin expression by Snail2 during the metastasis of colorectal cancer. *Clin Epigenetics* 2018; **10:** 154.

58. Olsen L, Thum E, Rohner N. Lipid metabolism in adaptation to extreme nutritional challenges. *Dev Cell* 2021; **56:** 1417–1429.

59. Comaills V, Kabeche L, Morris R, *et al*. Genomic instability is induced by persistent proliferation of cells undergoing epithelial-to-mesenchymal transition. *Cell Rep* 2016; **17:** 2632–2647.

60. Lovisa S, LeBleu VS, Tampe B, *et al*. Epithelial-to-mesenchymal transition induces cell cycle arrest and parenchymal damage in renal fibrosis. *Nat Med* 2015; **21:** 998–1009.

61. Singh AB, Sugimoto K, Harris RC. Juxtacrine activation of epidermal growth factor (EGF) receptor by membrane-anchored heparin-binding EGF-like growth factor protects epithelial cells from anoikis while maintaining an epithelial phenotype. *J Biol Chem* 2007; **282:** 32890–32901.

62. Singh AB, Harris RC. Autocrine, paracrine and juxtacrine signaling by EGFR ligands. *Cell Signal* 2005; **17:** 1183–1193.

63. Wu M-J, Chen Y-S, Kim MR, *et al*. Epithelial-mesenchymal transition directs stem cell polarity via regulation of mitofusin. *Cell Metab* 2019; **29:** 993–1002.e6.

64. Lenz G, Onzi GR, Lenz LS, *et al*. The origins of phenotypic heterogeneity in cancer. *Cancer Res* 2022; **82:** 3–11.

65. Jain P, Bhatia S, Thompson EW, *et al*. Population dynamics of epithelial-mesenchymal heterogeneity in cancer cells. *Biomolecules* 2022; **12:** 348–367.

66. Zheng X, Carstens JL, Kim J, *et al*. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* 2015; **527:** 525–530.

67. Valenzuela G, Canepa J, Simonetti C, *et al*. Consensus molecular subtypes of colorectal cancer in clinical practice: a translational approach. *World J Clin Oncol* 2021; **12:** 1000–1008.

68. Gautier L, Cope L, Bolstad BM, *et al*. Affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004; **20:** 307–315.

69. Wang J, Li S, Liu Y, *et al*. Metastatic patterns and survival outcomes in patients with stage IV colon cancer: a population-based analysis. *Cancer Med* 2020; **9:** 361–373.

References 68 and 69 are cited only in the supplementary material

SUPPLEMENTARY MATERIAL ONLINE

Supplementary materials and methods

Figure S1. Starvation triggers non-canonical partial epithelial-mesenchymal transition (p-EMT) with procession of membrane proteins

Figure S2. CAM tumour volumes of cells derived from control and starvation treatment groups

Figure S3. Toxicity test for suitable concentration of inhibitors

Figure S4. RT^2 Profiler PCR Array and cross-validation of gene signature overall survival prediction

Figure S5. Machine learning model performance on train data and its external validations

Figure S6. Correlation and CMS associations

Figure S7. Dependency of signature genes on nutrition availability

Figure S8. Effect of MG132 on *SERPINE1*, *SOX10*, and *EGFR* and susceptibility of CRC cells to 5-FU treatment depending on MG132

Table S1. Drug information

Table S2. Starvation medium composition

Table S3. Antibodies used for western blotting

Table S4. PCR primers

Table S5. Complete RT^2 Profiler PCR Array results

Table S6. Antibodies used for immunohistochemistry

Table S7. Independent gene analysis using COX proportional-hazards model

Table S8. Clinicopathological data and risk groups

Table S9. Clinicopathological data and risk groups (TCGA cohort)

Table S10. GSEA hallmark enrichment in high-risk [positive normalised enrichment score (NES)] and low-risk (negative NES) patients from GSE39582 dataset

Movie S1. EMT kiss 1 – HCT116 cells

Movie S2. EMT kiss 2 – HCT116 cells

Movie S3. EMT kiss 3 – HT29 cells

Movie S4. Asymmetric cell division – HCT116 cells