Metabolic control of tumour extracellular matrix production in cancer-associated fibroblasts

Emily J Kay[1,2], Karla Paterson[3], David Sumpton[1], Ekaterina Stepanova[4], Claudia Boldrini[1], Juan R Hernandez-Fernaud[1], Sandeep Dhayade[1], Enio Gjerga[5,6], Robin Shaw[1], Lisa J Neilson[1], Grigorios Koulouras[1], Grace McGregor[1,2], Sergio Lilla[1], Craig Jamieson[7], Ann Hedley[1], Radia Marie Johnson[8], Morag Park[8-11]; Crispin Miller[1,2], Jurre J Kamphorst[1,2], Fabricio Loayza-Puch[4], Julio Saez-Rodriguez[5,6], Karen Blyth[1,2], Michele Zagnoni[3], Sara Zanivan[1,2,*]

[1]Cancer Research UK Beatson Institute, Glasgow G611BD, UK; [2]Institute of Cancer Sciences, University of Glasgow, Glasgow, G611QH, UK; [3]Centre for Microsystems and Photonics, EEE Department, University of Strathclyde, Glasgow, UK; [4]Translational Control and Metabolism, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; [5]Heidelberg University, Faculty of Medicine, Institute for Computational Biomedicine, Bioquant, INF 267, 69120 Heidelberg, Germany; [6]RWTH Aachen University, Faculty of Medicine, Joint Research Centre for Computational Biomedicine (JRC-COMBINE), Aachen, Germany; [7]Department of Pure and Applied Chemistry, Thomas Graham Building, University of Strathclyde, Glasgow G1 1XL, UK; [8]Rosalind and Morris Goodman Cancer Research Centre, McGill University, Montreal, QC H3A 1A3, Canada; [9]Department of Biochemistry, McGill University, Montreal, QC H3A 1A3, Canada; [10]Department of Medicine, McGill University, Montreal, QC H3A 1A3, Canada; [11]Department of Oncology, McGill University, Montreal, QC H3A 1A3, Canada.

Correspondence: s.zanivan@beatson.gla.ac.uk

Abstract

The extracellular matrix (ECM) is the central driver of the desmoplastic reaction that fosters cancer aggressiveness. Cancer associated fibroblasts (CAFs) are the major source of ECM in tumours, thus being the optimal target to limit deposition of pro-tumourigenic ECM to oppose cancer. CAFs are metabolically active cells, however, how they support the biosynthetic requirements of producing ECM, and whether this can be targeted to influence tumour progression has not been investigated. We found that the pyruvate dehydrogenase kinase 2 (PDK2), a major inhibitor of the pyruvate dehydrogenase complex (PDC), is highly downregulated in CAFs and in the tumour stroma, when compared to normal fibroblasts. As consequence, PDC is more activated and generates acetyl-CoA, which elicits an epigenetic reprogramming through the histone acetyl transferase P300/CBP. This epigenetic reprogramming drives increased ECM production through increasing transcription of collagen genes and proline synthesis. We found that increased proline availability is necessary to support the biosynthetic requirements that follow the epigenetic reprogramming, for the translation of collagen to make abundant ECM. Targeting the rate-limiting enzyme for proline synthesis, pyrroline-5-carboxylate reductase 1 (PYCR1), in CAFs was sufficient to limit collagen deposition and hamper tumour growth. In conclusion, ECM production in CAFs is under strict metabolic control, and our results warrant considering targeting proline synthesis to normalise ECM production in tumours and possibly other diseases involving collagen production, such as fibrosis.

Keywords: CAFs, metabolism, ECM, collagen, cancer, pyruvate dehydrogenase, proline, epigenetics

Introduction

The extracellular matrix (ECM) actively regulates tumour initiation, growth, metastasis and response to therapy [1-4]. Cancer-associated fibroblasts (CAFs) are mesenchymal cells abundant in the stroma of solid tumours and are the major producers of ECM [5]. CAFs can originate from normal resident fibroblasts (NFs) that have undergone activation and the production of vast amounts of ECM is a hallmark event of their activation [6-8]. Pathways that support the biosynthetic requirements for ECM production in CAFs may offer opportunities for therapeutic intervention to block the pro-tumorigenic signals of the ECM.

Collagen is the most abundant component of the tumour ECM, and accumulation of collagen-rich ECM accompanies the development of solid tumours and is an indicator of a higher risk of invasive cancer [9] and resistance to therapeutic treatment [10]. The tumour ECM promotes cancer cells' growth, survival and invasion, and actively modifies the behaviour of stromal cells, including vascular and immune cells [1-4]. Experimental models have demonstrated that high levels of collagen enhance the formation, growth and progression of tumorigenic lesions and increase incidence of metastasis [11-14]. A collagen-rich desmoplastic stroma also contributes to impeding effective delivery of therapeutics, as well as intratumoural recruitment of immune cells, because it hampers the growth of a functional tumour vasculature [15-19]. Moreover, collagens can act as metabolic fuels, because cancer cells can exploit their unusually high content of proline residues to survive under nutrient limited conditions [20]. It is clear, therefore, that collagens have a central role in cancer and that inhibiting their production may have a favourable impact on cancer patients.

Metabolic reprogramming is a hallmark of cancer [21, 22] and, similarly to cancer cells, CAFs have an active metabolism. Most of the studies on CAF metabolism have focussed on secreted metabolites and how they fuel cancer cell growth and therapy response [23-28], while how specific metabolic pathways affect ECM production has been completely overlooked so far.

Metabolism is tightly intertwined with epigenetics [29, 30]. A recent study on ovarian CAFs has shown that targeting metabolism to increase methyl group availability promotes gene silencing through histone methylation, and that this rewires CAFs to being less pro-tumorigenic and pro-metastatic [31]. Histone acetylation is also closely linked to cell metabolism via acetyl-CoA, which, in addition to being a central metabolic intermediate for the TCA cycle and for lipid synthesis, is a second messenger and acetyl donor for the acetylation of proteins, including epigenetic regulators and histones [32, 33].

Here we show that acetyl-CoA in CAFs is central for coordinating the production of collagens through the regulation of the epigenetic regulator histone acetyl transferase P300/CBP, and that proline

synthesis through PYCR1 is essential to support the biosynthetic requirements activated by the epigenetic reprogramming to produce abundant collagen-rich ECM, as collagens are exceptionally rich in proline residues. These data establish a critical role for metabolism to influence hallmarks of CAFs important for cancer pathology, and that PYCR1 is a potentially targetable vulnerability of tumour ECM production.

Results

The pyruvate dehydrogenase complex is highly active in mammary CAFs

To identify pathways that underpin CAF functions acquired upon activation, we used a model of immortalised human mammary NFs and CAFs (iCAFs and iNFs), where iNFs had been activated into iCAFs by human breast cancer cells *in-vivo* [34]. We and others have shown that iCAFs have a myofibroblast-like phenotype, because they express high levels of alpha smooth muscle actin (ACTA2/αSMA) and produce abundant ECM (**Extended Data Figure 1a,b** and [35]). Moreover, iCAFs promote tumour growth, invasion and angiogenesis [34-36]. Computational analysis of mass spectrometry (MS)-phosphoproteomic data of iCAFs and iNFs (**Supplemental Data S1**) predicted altered kinase activities in iCAFs, in particular, pyruvate dehydrogenase kinase 2 (PDK2) was predicted to be the most de-activated kinase (**Figure 1a**). Supporting this prediction, the phosphorylation levels of the pyruvate dehydrogenase E1 component subunit alpha A1 (PDHA1) at serine 293, which is a known substrate of PDK2 [37], was strongly reduced in iCAFs and CAFs isolated from breast cancer patients (pCAFs) in comparison to their NF counterparts (**Figure 1b,c**. For iCAF see also **Supplemental Data S1**). Similarly to the iCAFs, the pCAFs showed myofibroblast-like features, because they expressed higher levels of αSMA compared to their NF counterpart (**Extended Data Figure 1a**). The phosphorylation of PDHA1 at serine 293 inhibits the activity of the pyruvate dehydrogenase complex (PDC), which converts pyruvate to acetyl-CoA. Concordantly, PDC activity was higher in CAFs, as measured with an enzymatic assay (**Figure 1d**), and CAFs had higher levels of intracellular acetyl-CoA than NFs (**Figure 1e**). RT-qPCR analysis showed that CAFs express less *PDK2* than NFs (**Figure 1f**). Highlighting the relevance of this finding in patients, gene expression data of microdissected stroma from normal breast and triple negative breast cancers [38] showed that *PDK2* is downregulated in the stroma in patients with advanced disease (**Figure 1g**). *PDK2* was also downregulated in the stroma of high grade serous ovarian cancers [39] (**Figure 1g**), suggesting that this regulation is not unique to breast cancer.

Silencing PDK2 in NFs was sufficient to decrease PDHA1 phosphorylation similarly to the levels measured in iCAFs, and overexpression of PDK2 wild type ($PDK2^{WT}$) in CAFs, but not a mutant

enzymatically inactive form [40] ($PDK2^{N255A}$), increased PDHA1 phosphorylation (**Figure 1h** and **Extended Data Figure 1c,d**). Showing that PDHA1 phosphorylation levels was an indicator of PDC activity, acetyl-CoA levels increased upon PDK2 silencing in iNFs while they decreased upon PDK2 overexpressing in iCAFs (**Figure 1i**). Hence, PDK2 is a major regulator of PDC activity in mammary fibroblasts and is downregulated in the stroma of breast and ovarian tumours.

Kay et al. Figure 1

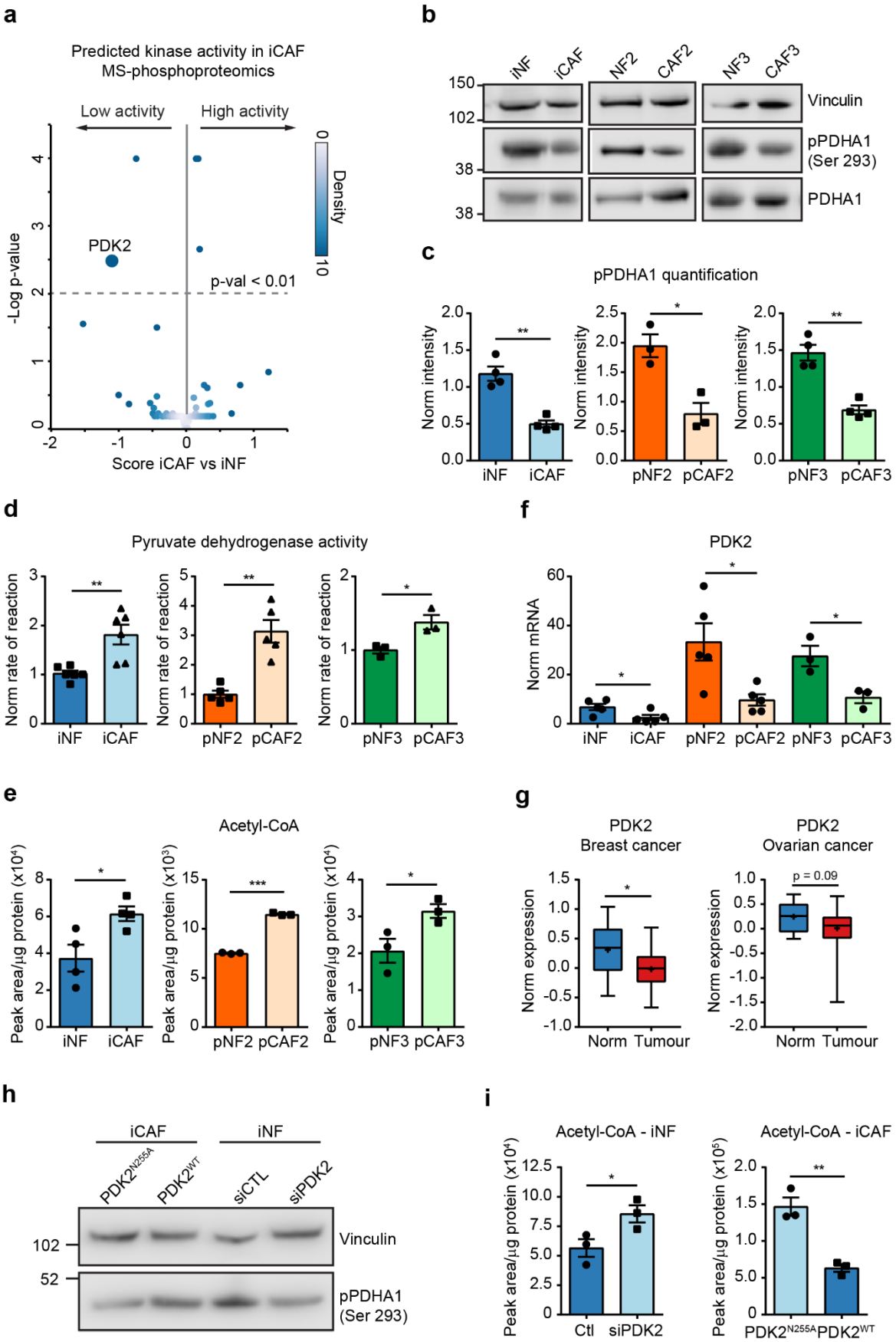
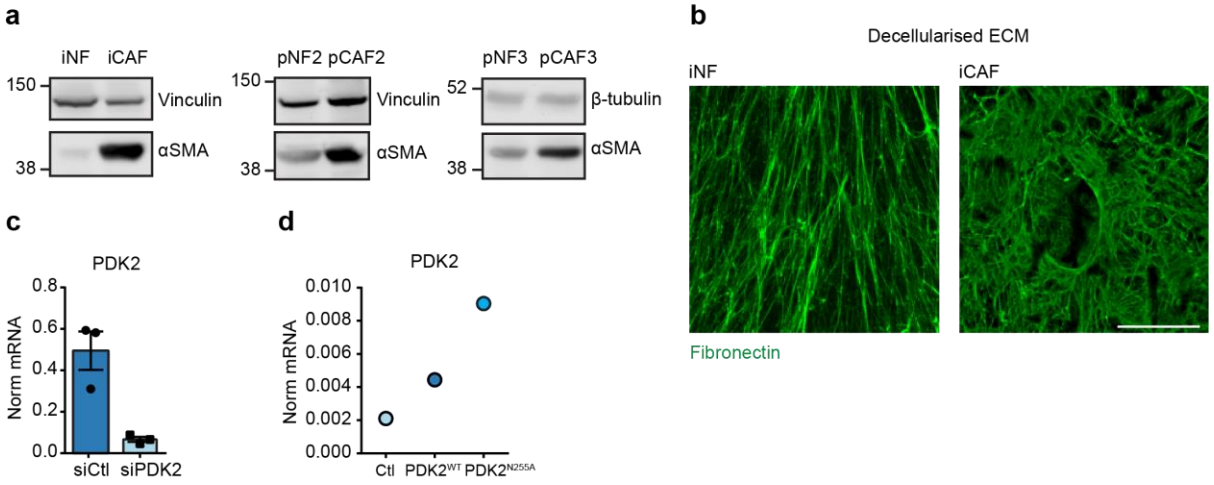


Figure 1. PDK2 is downregulated in the tumour stroma and increases PDC activity in mammary CAFs

a. Predicted kinase activity in iCAFs compared to iNFs based on the modelling of MS phosphoproteomic data. **b.** Representative western blots showing the phosphorylation levels of PDHA1 regulatory site (Ser 293) in total lysate from NFs and CAFs. Vinculin was used as a loading control. **c.** Quantification of the phosphorylation levels of PDHA1 regulatory site Ser 293 in NFs and CAFs. N = 3 or 4 independent experiments. PDHA1 phosphorylation was normalised to total PDHA1 levels. **d.** Pyruvate dehydrogenase activity of NFs and CAFs measured as the rate of NAD+ reduction *in vitro*. N ≥ 3 independent experiments. Data was normalised to the average rate of NAD+ reduction of the NFs for each pair. **e.** Total intracellular acetyl-CoA levels in NFs and CAFs measured by MS. N = 3 independent experiments. **f.** mRNA levels of *PDK2* in NFs and CAFs measured by RT-qPCR. N ≥ 3 independent experiments. Data was normalised to *18S* levels. **g.** *PDK2* gene expression in microdissected sections of normal and triple negative breast cancer or high grade serous ovarian cancer stroma. Data are from gene expression datasets GSE90505 and GSE40595. **h.** Representative western blot showing PDHA1 phosphorylation levels in total lysate from iNFs transfected with siCtl or siPDK2 and iCAFs transfected with pGC-PDK2^{N255A} or pGC-PDK2^{WT} for 48 **h.** Vinculin was used as a loading control. **i.** Total intracellular acetyl-CoA measured by MS in iNFs transfected with siCtl or siPDK2 and iCAFs transfected with pGC-PDK2^{N255A} or pGC-PDK2^{WT} for 48 h. N = 3 independent experiments.
Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. A two-tailed unpaired t-test with Welch's correction was used to determine the p-value.

Kay et al. Extended Data Figure 1



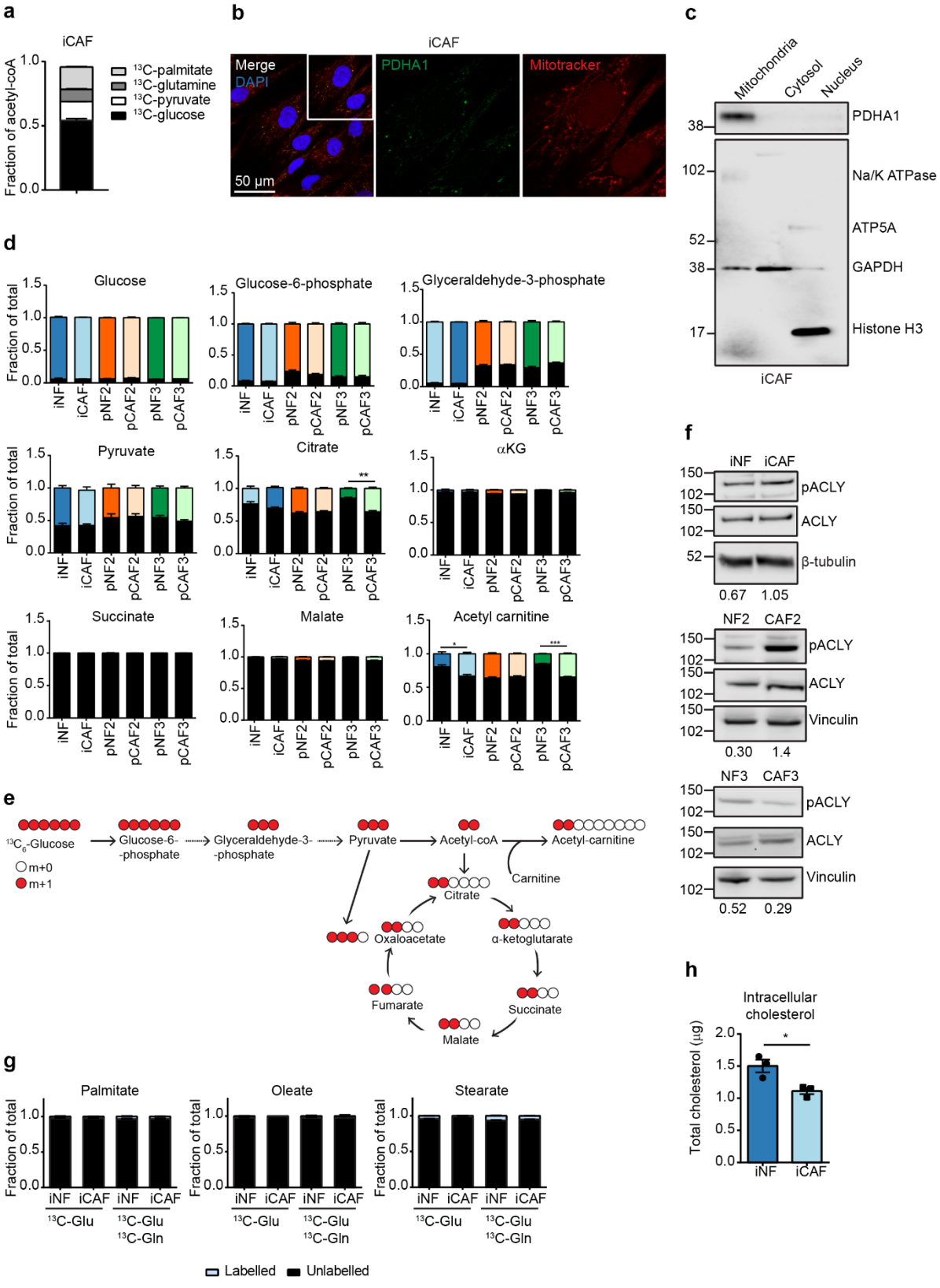
Extended Data Figure 1. Modulation of PDK2 expression in mammary CAFs and NFs

a. Western blot for αSMA to show its levels in the total lysate from CAFs and corresponding NFs. Vinculin and β-tubulin were used as loading controls. **b.** Representative confocal microscope images of decellularised ECM generated by iCAFs or iNFs grown for 10 days in culture, which was stained for fibronectin. **c.** *PDK2* mRNA expression in iNFs transfected with siCtl or siPDK2, as measured by RT-qPCR. N = 3 independent experiments. **d.** Representative RT-qPCR data showing *PDK2* expression in iCAFs transfected with control vector, pGC-PDK2^{N255A} or pGC-PDK2^{WT}. Data was normalised to *18S* expression.
Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. A two-tailed unpaired t-test with Welch's correction was used to determine the p-value.

Increased pyruvate dehydrogenase activity does not impact TCA cycle or lipid metabolism

PDC activity was the major source of acetyl-CoA in CAFs. MS tracing experiments using uniformly ^{13}C-labelled glucose, glutamine, palmitate or pyruvate showed that over 70% of acetyl-CoA was derived from glucose or pyruvate, both of which are converted to acetyl-CoA via PDC (**Figure 2a** and **Extended Data Figure Sa**). There is evidence that PDC can translocate into the nucleus to synthesise acetyl-CoA *in loco* [41, 42], but we excluded this mechanism in CAFs because we detected the majority of PDHA1 in the mitochondria, by immunofluorescence staining of fixed cells and western blot analysis of mitochondria (**Extended Data Figure 2b,c**). Tracing experiments with $^{13}C_6$-glucose additionally indicated that PDC-derived acetyl-CoA was likely to be transported into the cytosol, because only a minimal portion of ^{13}C-labelled citrate was funnelled into the TCA cycle (**Extended Data Figure 2d,e**). Moreover, two out of three CAF lines had increased levels of ATP citrate synthase (ACLY) phosphorylated at serine 455 (active [43, 44]) compared to their NF counterpart (**Extended Data Figure 2f**). Next, we asked whether the excess of acetyl-CoA produced by CAFs was exploited to increase lipid synthesis. We excluded this possibility, because iCAFs and iNFs in culture had similar levels of cholesterol and synthesised only minimal amounts of fatty acids from $^{13}C_6$-glucose and $^{13}C_5$-glutamine (**Extended Data Figure 2g,h**). Thus, increased PDC activity in cultured mammary CAFs does not increase TCA cycle activity or lipid synthesis.

Kay et al. Extended Data Figure 2



Extended Data Figure 2. Increased PDH activity does not fuel the TCA cycle or lipid synthesis

a. Fraction of acetyl-CoA labelled by $^{13}C_6$-glucose, $^{13}C_5$-glutamine, $^{13}C_3$-pyruvate or $^{13}C_{16}$-palmitate after 48 h of labelling in iCAFs, as measured by MS. N = 3 independent experiments for each labelling condition. **b.** Representative confocal images of iCAFs stained with MitoTracker (red), PDHA1 (green) and DAPI (blue). Scale bar = 50 µm. **c.** Western blot for PDHA1 in nuclear, cytosolic and mitochondrial protein fractions of iCAFs. An organelle antibody cocktail containing markers for each compartment was used to show the purity of the fractionation. **d.** Fraction of $^{13}C_6$-glucose incorporation into metabolites (coloured) involved in glycolysis and the TCA cycle in CAFs and NFs. In black is the unlabelled fraction of the metabolite. N = 3 independent experiments. **e.** Schematic of ^{13}C incorporation in metabolites derived from $^{13}C_6$-glucose and shown in (d). **f.** Western blots of phosphorylated and total ACLY in total lysate from all CAF and NF pairs. Vinculin or β-tubulin was used as a loading control. Quantification of phosphorylated ACLY, which was normalised to total ACLY and the loading control, is shown below the blot. **g.** Fraction of $^{13}C_6$-glucose and $^{13}C_5$-glutamine incorporation into fatty acids in iCAFs and iNFs measured by MS. N = 3 independent experiments. **h.** Total cholesterol in iCAFs and iNFs measured by MS. N = 3 independent experiments.
Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. A two-tailed unpaired t-test with Welch's correction was used to determine the p-value.

PDC activity elicits an epigenetic switch through P300/CBP

Acetyl-CoA is a known second messenger [32, 33] and increased levels of acetyl-CoA in cancer cells promote histone acetylation through the activation of the epigenetic regulator histone acetyl transferase P300/CBP [45, 46]. Our unbiased MS analysis of global protein acetylation pinpointed higher levels of acetylated histones in iCAFs than in iNFs (**Figure 2b** and **Supplemental Data S2**). Among known regulatory sites, we found hyperacetylation of the histone H3 sites K18, K23 and K27 (**Figure 2c**), which are substrates of P300/CBP [47]. Western blot analysis confirmed H3K27 hyperacetylation in CAFs (**Extended Data Figure 3a**). We reasoned that PDC-derived acetyl-CoA in CAFs is an epigenetic rather than a metabolic regulator. To assess this, we reduced acetyl-CoA levels by treating CAFs with the ACLY inhibitor BMS303141 [45, 48] and monitored H3K27 acetylation (H3K27ac). BMS303141 treatment decreased H3K27ac, and this effect was abolished by replenishing the pool of nucleo-cytosolic acetyl-CoA with exogenous acetate (**Figure 2d,e**). Similarly, pharmacological inhibition of P300/CBP with c646 [45, 49] or A-485 [50] reduced H3K27ac in iCAFs (**Figure 2f-i**). H3K27 acetylation relied on PDC activity, because overexpression of PDK2 in CAFs to inactivate PDC reduced H3K27ac, and exogenous acetate restored acetylation levels (**Figure 2j,k**). Moreover, silencing PDK2 in NFs to increase PDC activity was sufficient to enhance H3K27ac, and this effect was blocked when P300/CBP was inhibited pharmacologically with c646 (**Figure 2l,m**). Hence, increased PDC activity in cultured mammary CAFs is an epigenetic regulator.

Kay et al. Figure 2

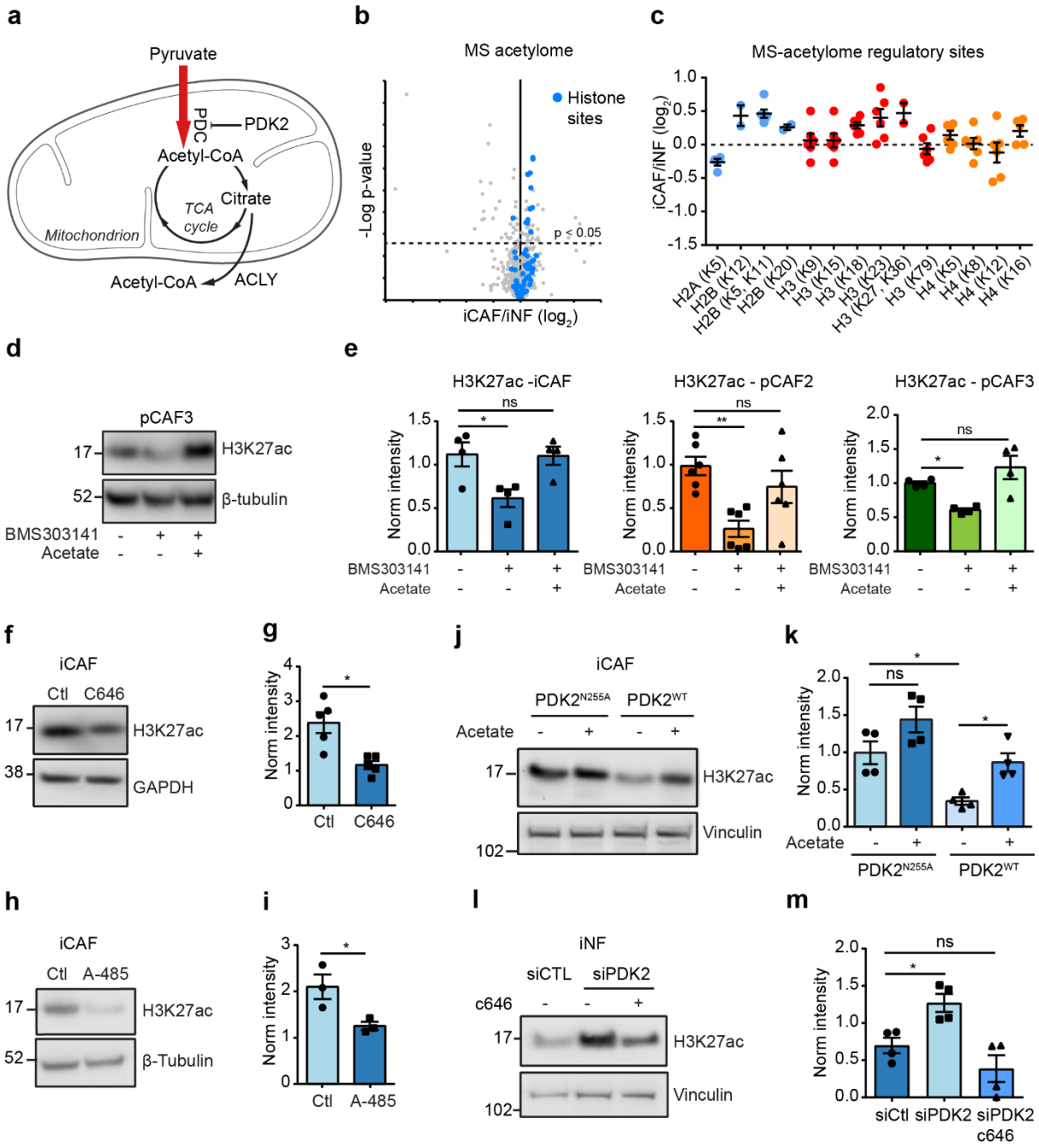
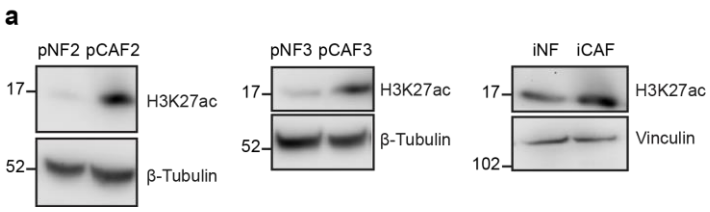


Figure 2. PDH-activity drives an epigenetic regulation that requires P300/CBP

a. Scheme showing that in CAFs the downregulation of PDK2 induces increased PDC activity (red), which in turn produces more acetyl-CoA that could be converted into citrate and either used in the TCA cycle or exported and converted back to acetyl-CoA in the cytosol/nucleus by ACLY. **b.** Volcano plot showing the results of a one sample t-test on the SILAC ratios between iCAFs and iNFs of peptides carrying lysine acetylation sites. N = 5 independent experiments. **c.** SILAC ratios between iCAFs and iNFs of histone acetylation sites with a known regulatory function (based on PhosphoSitePlus [51]) identified in (b). N = 2-5 independent experiments. **d, e.** Representative western blot (d) and quantification (e) showing H3K27ac levels in total lysate from CAFs following treatment with BMS303141 ± acetate for 48 h. Vinculin was used as loading control. N = 4-6 independent experiments. **f,g.** Representative western blot (f) and quantification (g) showing H3K27ac levels in total lysate from iCAFs ± c646 treatment for 48 h. GAPDH was used as loading control. N = 5 independent experiments. **h,i.** Representative western blot (h) and quantification (i) showing H3K27ac levels in total lysate from iCAFs ± A-485 treatment for 48 h. β-tubulin was used as loading control. N = 3 independent experiments. **j,k.** Representative western blot (j) and quantification (k) showing H3K27ac levels in total lysate from iCAFs transfected with pGC-PDK2^{N255A} or pGC-PDK2^{WT} for 48 h. Vinculin was used as loading control. N = 4 independent experiments. **l,m.** Representative western blot and quantification showing H3K27ac in total lysate from iNFs transfected with siCtl or siPDK2 ± c646 for 48 h. Vinculin was used as loading control. N = 4 independent experiments.
Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. For experiments with two conditions, a two-tailed unpaired t-test with Welch's correction was used to determine the p-value. For experiments with more than two conditions, a one way ANOVA test with Dunnett's multiple comparison test was used.

Kay et al. Extended Data Figure 3



Extended Data Figure 3. H3K27 is hyperacetylated in mammary CAFs

a. Representative western blots showing H3K27ac levels in total lysate from all CAF and NF cell lines. β-tubulin or vinculin was used as a loading control.

The PDC-P300/CBP pathway co-regulates ECM production and proline synthesis

Histone hyperacetylation is a hallmark of gene expression activation, and the recruitment of P300/CBP and hyperacetylation of histone H3 at enhancers of pro-fibrotic genes is a landmark event in fibrosis [52-54]. Upon treatment of CAFs with the P300/CBP inhibitor c646, collagens were the most downregulated proteins (**Figure 3a** and **Supplemental Data S3**). This result indicates that P300/CBP may also be a key regulator of fibrosis in the tumour stroma. Our proteomic analysis also pinpointed a strong downregulation of the rate-limiting enzyme for proline synthesis, PYCR1. This was intriguing because many of the highly abundant proteins in the iCAF ECM [35], particularly collagens, contain an above average number of proline residues (**Figure 3b** and **Supplemental Data S4**), and there is no clear link described between ECM production and proline metabolism in cancer. Further supporting a possible connection between these two mechanisms, CAFs had more intracellular proline than NFs (**Figure 3c**), in addition to producing more ECM [35]. Therefore we explored further whether the PDC-P300/CBP pathway is an upstream regulator of both proline and ECM synthesis.

RT-qPCR analysis confirmed that the expression of *COL1A1* and *COL6A1*, which were among the most abundant components of the iCAF ECM (**Figure 3b**) and are fundamental regulators of cancer progression [4, 55], was reduced upon P300/CBP inhibition (**Figure 3d** and **Extended Data Figure 4a**). Concordantly, there was less collagen in the ECM deposited by CAFs treated with c646 (**Figure 3e,f**). Collagen expression was reliant on PDC activity and downstream signalling. Pharmacological inhibition of ACLY in CAFs decreased COL6A1 levels in the ECM, which were rescued with exogenous acetate (**Figure 3g,h**). Moreover, silencing PDK2 in NFs was sufficient to increase *COL6A1* and *COL1A1* expression and collagen deposition in the ECM, and this effect was blocked upon P300/CBP inhibition (**Figure 3i**). Conversely, PDK2 overexpression in CAFs reduced *COL6A1* and *COL1A1* expression and collagen deposition in the ECM, and this was rescued by providing exogenous acetate to the cells (**Figure 3j**). Similarly to the collagen, *PYCR1* expression and proline levels increased with acetyl-CoA availability (**Figure 3k,l** and **Extended Data Figure 4b,c**), and *PYCR1* expression was reliant on P300/CBP and PDC activity (**Figure 3d,i,j**). Thus, the PDK2-PDC-P300/CBP pathway influences both proline and ECM synthesis.

Kay et al. Figure 3

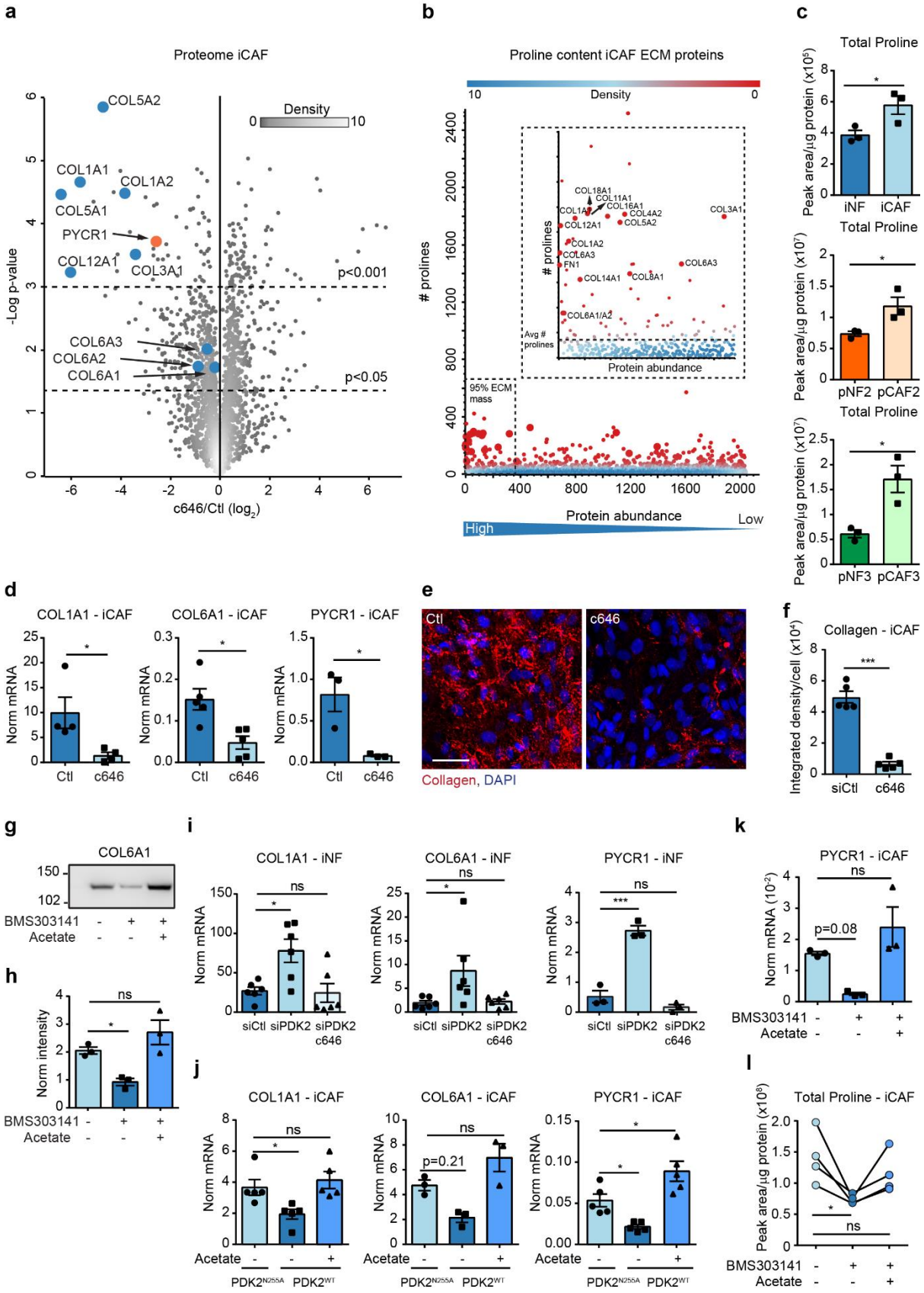
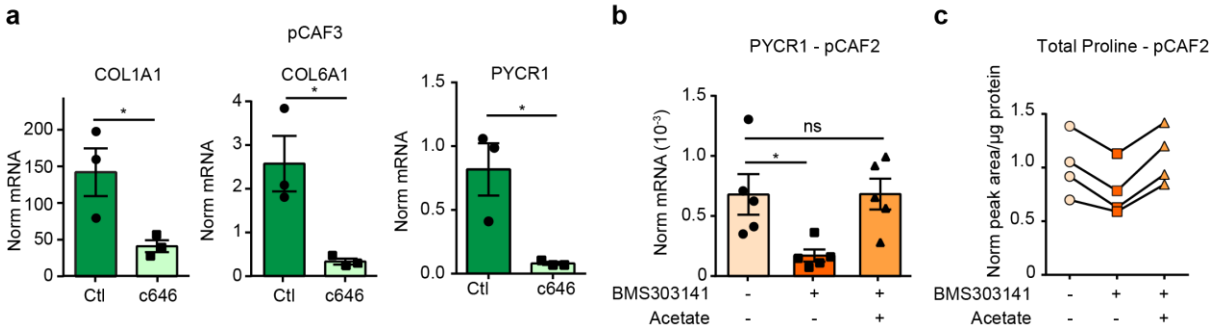


Figure 3. Increased PDH-P300/CBP activity enhances expression of collagens and of the rate-limiting enzyme for proline synthesis PYCR1

a. Volcano plot showing the average log_2 ratios of proteins quantified in the total proteome of iCAFs ± c646 for 24 h. P-values highlight regulated proteins. The identified collagens and PYCR1 are highlighted in colours. N = 3 independent experiments. **b.** Dot plot showing the rank of the proteins identified in the iCAF ECM by MS proteomic analysis in the x axis [35]. Ranking was based on the estimated total abundance of each protein, which was calculated by dividing the intensity value of each protein by its molecular weight. Proteins are ranked from the most (left) to the least (right) abundant. The number of proline residues contained in each protein is indicated on the y axis. **c.** Total intracellular proline levels measured by MS in CAFs and NFs. N = 3 independent experiments. **d.** mRNA expression of *COL1A1*, *COL6A1* and *PYCR1* in iCAFs ± c646 as measured by RT-qPCR. N ≥ 3 independent experiments. Data was normalised to *TBP2* levels. **e, f.** Representative images (e) and quantification (f) of collagen produced by iCAFs ± c646 treatment for 96 h. Collagen was visualised with the collagen binding protein CNA35-mCherry. Nuclei were stained with DAPI. N = 5 independent experiments. **g, h.** Representative western blot (g) and quantification (h) of COL6A1 in decellularised ECM derived from iCAFs following treatment with BMS303141 ± acetate for 96 h. COL6A1 levels were normalised to total protein content in each lane, which was measured with Ponceau staining. N = 3 independent experiments. **i.** mRNA expression of *COL1A1*, *COL6A1* and *PYCR1* in iNFs transfected with siCtl/siPDK2 ± c646 for 48 h, as measured by RT-qPCR. N ≥ 3 independent experiments. Data was normalised to *TBP2* levels. **j.** mRNA expression of *COL1A1*, *COL6A1* and *PYCR1* in iCAFs transfected with pGC-PDK2^{N255A} or pGC^{WT} ± acetate for 48 h, as measured by RT-qPCR. N ≥ 3 independent experiments. Data was normalised to *TBP2* levels. **k.** mRNA expression of *PYCR1* in iCAFs treated with BMS303141 ± acetate for 48 h, as measured by RT-qPCR. N = 3 independent experiments. Data was normalised to *ACTB* levels. **l.** Total intracellular proline levels measured by MS in iCAFs following treatment with BMS303141 ± acetate for 48 h. N = 4 independent experiments.
Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. Scale bar = 50 µm. For experiments with two conditions, a two-tailed unpaired t-test with Welch's correction was used to determine the p-value. For experiments with more than two conditions, a one way ANOVA test with Dunnett's multiple comparison test was used.

Kay et al. Extended Data Figure 4



Extended Data Figure 4. Nucleocytosolic acetyl-coA and P300/CBP promote collagen and proline production

a. *COL1A1*, *COL6A1* and *PYCR1* mRNA expression in pCAF3 ± c646, as measured by RT-qPCR. N = 3 independent experiments. Data was normalised to *TBP2* levels. **b.** *PYCR1* mRNA expression in pCAF2s treated with BMS303141 ± acetate, as measured by RT-qPCR. N = 5 independent experiments. Data was normalised to *ACTB* levels. **c.** Total intracellular proline in pCAF2 treated with BMS303141 ± acetate measured by MS and normalised to total protein content. N = 4 independent experiments. Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. For experiments with two conditions, a two-tailed unpaired t-test with Welch's correction was used to determine the p-value. For experiments with more than two conditions, a one way ANOVA test with Dunnett's multiple comparison test was used.

PYCR1-derived proline is necessary to sustain increased ECM production in CAFs

To further investigate the link between proline synthesis and collagen production, we assessed whether CAFs in culture use PYCR1-derived proline to make collagen. This was the case, because when CAFs were grown in medium containing $^{13}C_5$-glutamine, ^{13}C-labelled proline was detected in collagen-derived peptides by MS-proteomic analysis of their ECM (**Figure 4a,b** and **Supplemental Data S5**). Next, we inhibited PYCR1 genetically, with siRNA or shRNA (**Extended Data Figure 5a-c**), or pharmacologically with a recently developed small molecule inhibitor (PYCR1i) [56], and found that it was sufficient to decrease proline production (**Figure 4c-e**) and collagen deposition in the ECM, which was rescued by providing cells with exogenous proline (**Figure 4f-k**). Similarly, inhibiting the glutaminase enzyme (GLS), which catalyses the conversion of glutamine to glutamate (**Figure 4a**), the first step for proline synthesis from glutamine via PYCR1, with the clinical compound CB-839 [57] reduced proline synthesis and collagen deposition in the ECM (**Extended Data Figure 5d,e**). This further indicates that the pool of glutamine-derived proline is important for collagen production. In our experimental conditions, FBS was the only source of extracellular proline in the culture medium (~4 µM, not shown) and levels were lower than those found in the blood of human adults (50-200 µM, www.serummetabolome.ca). We assessed the impact of physiologically relevant levels of proline on collagen production when targeting PYCR1. CAFs treated with PYCR1i or silenced for PYCR1 became insensitive to the effects of PYCR1i on collagen production only in the presence of above-physiological doses of proline (**Extended Data Figure 5f-i**), indicating that proline produced by PYCR1 is required for collagen synthesis, in addition to exogenous proline.

To determine whether proline shortage had a direct effect on protein translation, we performed differential ribosome codon reading (diricore) analysis [58]. This analysis showed that silencing PYCR1 in CAFs induced ribosome stalling specifically at proline codons, which was rescued with the addition of exogenous proline (**Extended Data Figure 5j**). Among the genes that were identified and found affected by proline levels, there were several collagens, including *COL1A1* (**Figure 4l**). Conversely, *COL1A1* mRNA levels were not modulated by proline levels (**Extended Data Figure 5k**). Thus, PYCR1-derived proline in CAFs is required for the translation of collagen for ECM production.

Finally, reducing PYCR1 levels in NFs silenced for PDK2 inhibited collagen production induced by PDC activation (**Figure 4m,n** and **Extended Data Figure 5l**), indicating that PYCR1 is required to support collagen production upon PDC activation.

Hence, proline availability is a limiting factor for the production of collagen-rich ECM and inhibiting PYCR1 can overwrite the epigenetic reprogramming induced by the activation of the PDK2-PDC-P300/CBP pathway to increase collagen genes expression.

Kay et al. Figure 4

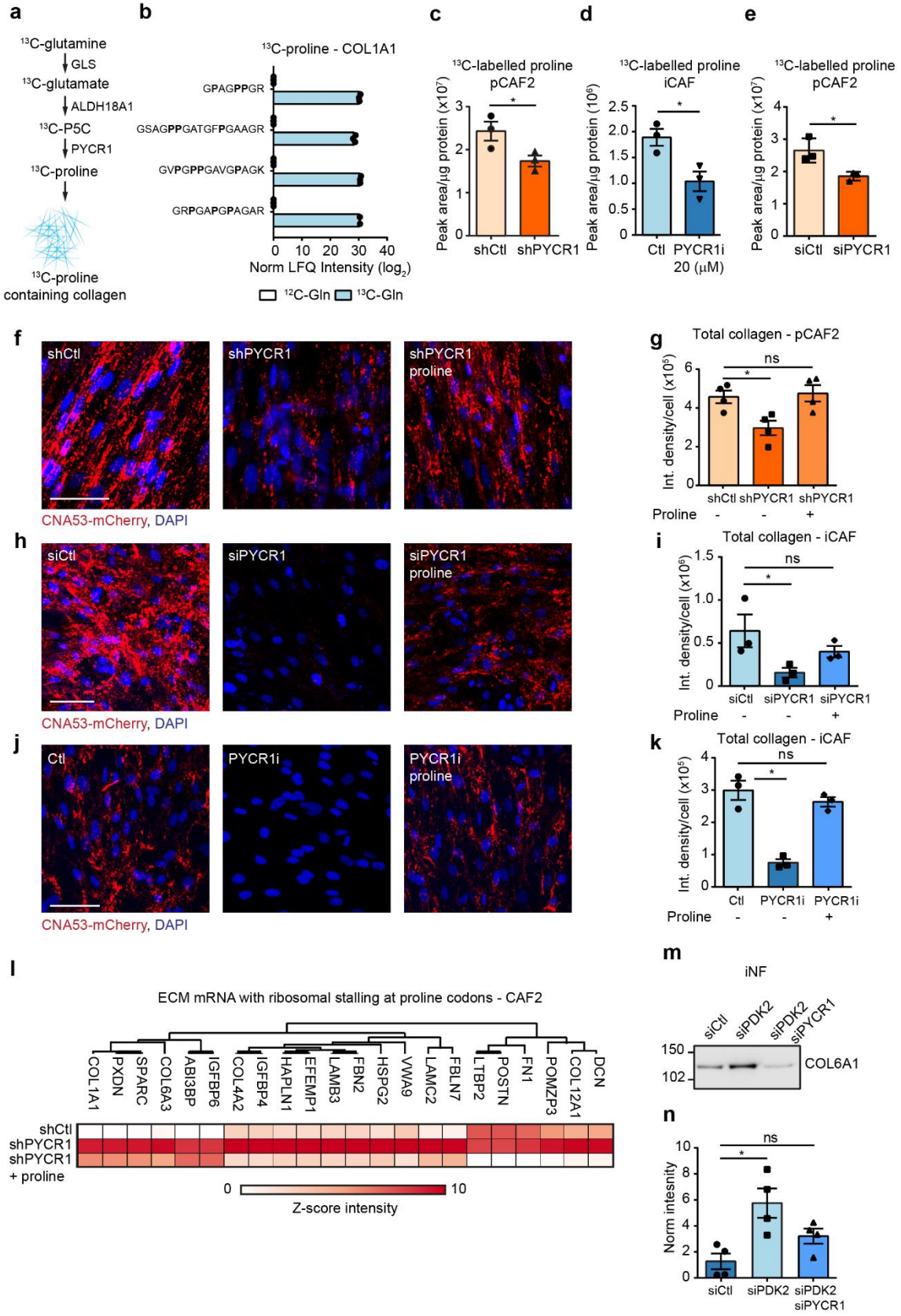
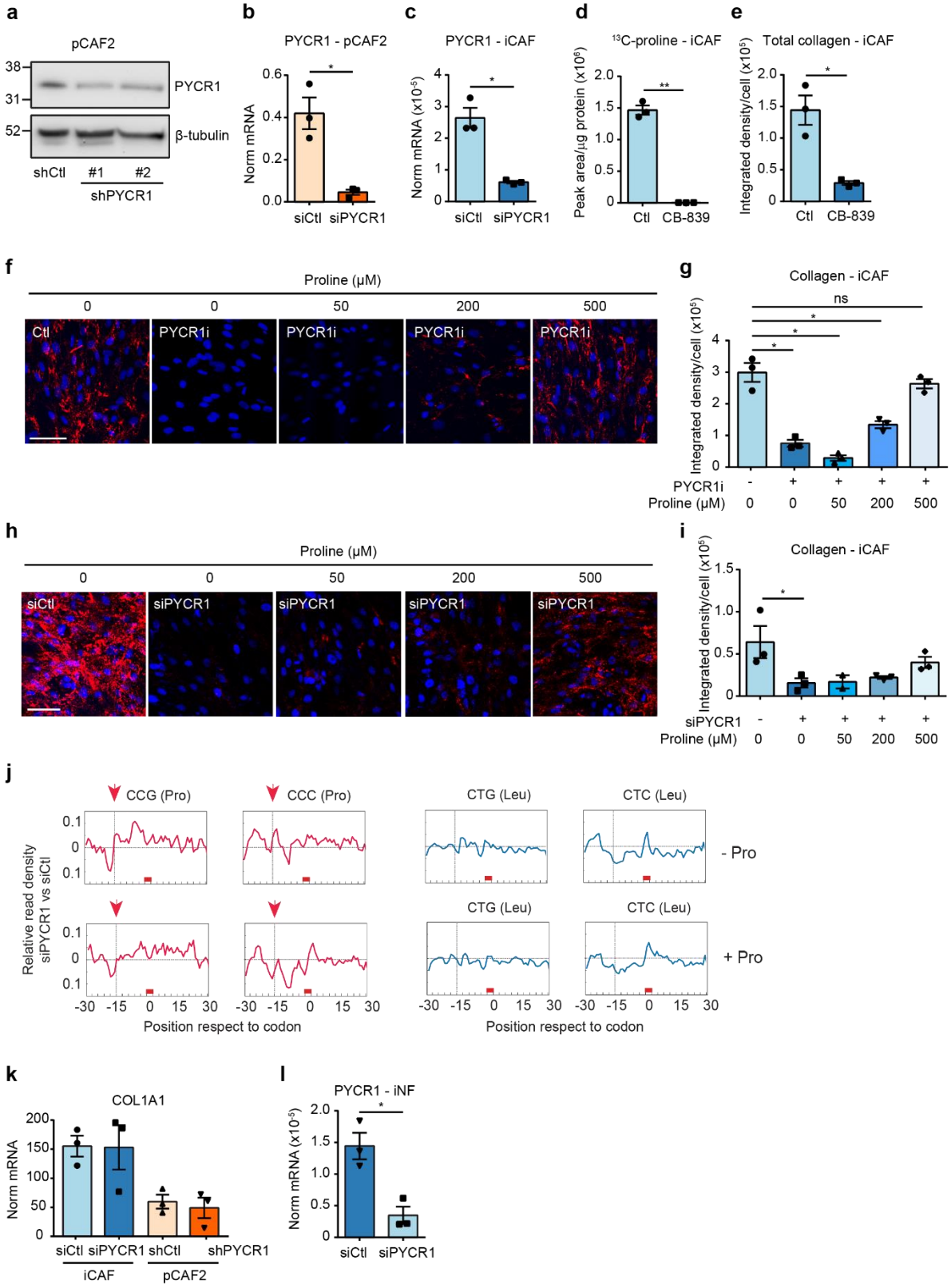


Figure 4: Increased proline synthesis induced by PDH activity is necessary for the production of collagen-rich ECM

a. Scheme showing how ^{13}C is incorporated into proline from $^{13}C_5$-glutamine and used to make collagen in the ECM. **b.** Log_2 LFQ intensity of ^{13}C-proline-labelled COL1A1 peptides in decellularised ECM derived from CAFs cultured in medium with $^{13}C_5$-glutamine for 48 h and analysed by MS. **c-e.** Intracellular ^{13}C-labelled proline measured by MS in CAFs ± PYCR1i or PYCR1 knockdown for 48 h, cultured in medium with $^{13}C_5$-glutamine. N = 3 independent experiments. **f, g.** Representative images (f) and quantification (g) of collagen produced by shCtl and shPYCR1 pCAFs ± proline. N = 4 independent experiments. **h, i.** Representative images (h) and quantification (i) of collagen produced by iCAFs transfected with siCtl or siPYCR1 ± proline. N = 3 independent experiments. **j, k.** Representative images (j) and quantification (k) of collagen produced by iCAFs treated with PYCR1i ± proline. N = 3 independent experiments. **l.** ECM proteins significantly affected by ribosomal stalling at proline codons when PYCR1 is silenced and rescued with exogenous proline. The score indicates the normalized density of ribosomes at the A-site of Pro codons and measures the tendency of ribosomes to spend more time in that position. Original data were Z-scored. **m.** Representative western blot (m) and quantification (n) of COL6A1 in decellularised ECM derived from iNFs transfected with siCtl, siPDK2 or siPDK2 + siPYCR1 for 72 h. COL6A1 levels were normalised to total protein levels in each lane, which was measured with Ponceau staining. N = 4 independent experiments.
Collagen was visualised with the collagen binding-protein CNA35-mCherry. Int. density = Integrated density. Scale bar = 50 µm. Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. For experiments with two conditions, a two-tailed unpaired t-test with Welch's correction was used to determine the p-value. For experiments with more than two conditions, a one way ANOVA test with Dunnett's multiple comparison test was used.

Kay et al. Extended Data Figure 5



Extended Data Figure 5. Increased proline synthesis by PYCR1 is necessary for collagen synthesis

a. Representative western blot of PYCR1 in total lysate from shCtl and shPYCR1 CAFs. β-tubulin was used as a loading control. **b,c.** *PYCR1* mRNA expression in iCAFs and pCAF2 transfected with siCtl or siPYCR1, as measured by RT-qPCR. N = 3 independent experiments. **d.** ^{13}C-labelled proline in CAFs ± glutaminase inhibitor CB-839 cultured in media with $^{13}C_5$-glutamine, measured by MS. N = 3 independent experiments. **e.** Quantification of collagen, based on fluorescence signal, produced by CAFs cultured ± CB-839 for 72 h. **f,g.** Representative images (f) and quantification (g) of collagen produced by CAFs ± PYCR1i and treated with 0-500 µM proline. N = 3 independent experiments. **h,i.** Representative images (h) and quantification (i) of collagen produced by CAFs transfected with siCtl or siPYCR1 for 72 h and treated with 0-500 µM proline. N = 3 independent experiments. **j.** Density plots for proline (Pro) codons (all genes) comparing siPYCR1 and siCtl pCAF2. The peak at position 15 corresponds to the site A of Pro codons. Leucine (Leu) was used as control. **k.** *COL1A1* mRNA expression in iCAFs transfected with siCtl/ siPYCR1 for 48 h and pCAF2 expressing shCtl/shPYCR1, as measured by RT-qPCR. Data was normalised to *TBP2* levels. N = 3 independent experiments. **l.** *PYCR1* mRNA expression in iNFs transfected with siCtl or siPYCR1 for 48 h, as measured by RT-qPCR. N = 3 independent experiments. Data was normalised to *ACTB* levels.
Collagen was visualised with the fluorescent collagen-binding protein CNA35-mCherry. Scale bar = 50 µm. Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. A two-tailed unpaired t-test with Welch's correction was used to determine the p-value.

The PDC-P300/CBP-PYCR1 pathway supports collagen production in co-cultures

Next, we assessed the role of the PDC-P300/CBP-PYCR1 pathway in supporting ECM production in a more physiological context with pharmacological inhibition of the key components of the pathway. To inhibit the pyruvate dehydrogenase complex, we used the clinical compound CPI-613, which is a dehydrogenase inhibitor [59] that reduced histone acetylation and expression of collagen and *PYCR1*, as well as deposition of collagen-rich ECM (**Extended Data Figure 6a-e**), mirroring PDC inhibition by overexpression of PDK2 in CAFs.

CAFs were co-cultured with a primary breast cancer cell line in a 2D or 3D environment. Imaging analysis of the 2D co-cultures showed that the majority of the collagen co-localised with CAFs, indicating that, similarly to *in vivo* [5], CAFs are the major source of collagen , while the contribution of the cancer cells is marginal (**Figure 5a**). Collagen production was strongly reduced upon pharmacological inhibition of PDC, ACLY, P300/CBP or PYCR1 with the corresponding inhibitors, which was rescued by providing exogenous acetate or proline when the cells were treated with ACLYi (BMS303141) and PDCi (CPI-613) or PYCR1i, respectively (**Figures 5b-i**). Notably, each pharmacological treatment, except for PDCi, significantly reduced cancer cell proliferation (**Figure 5j-m**). Similar results were obtained when cells were co-cultured in 3D in microfluidic devices, as spheroids, and the effects of the treatments were dose-dependent (**Figure 5n-r**). To determine whether the reduction in cancer cell growth was a consequence of reduced collagen in the ECM or a direct effect of the drugs on the cancer cells, we measured cancer cell proliferation upon treatment when grown in mono-culture (**Extended Data Figure 6f**). All the treatments reduced cancer cell growth to some extent, but only the reduction induced upon CPI-613 and c646 treatment was statistically significant. Conversely, the inhibition of cancer cell proliferation induced by BMS303141 and PYCR1i treatment was very pronounced only when the cells were co-cultured with CAFs. This indicates that, with those treatments, decreased availability of collagen in the ECM may actively contribute to reduce cancer cell growth. Thus, targeting the PDC-P300/CBP-PYCR1 pathway is a potent inhibitor of collagen-rich ECM production and cancer cell growth.

Kay et al. Figure 5

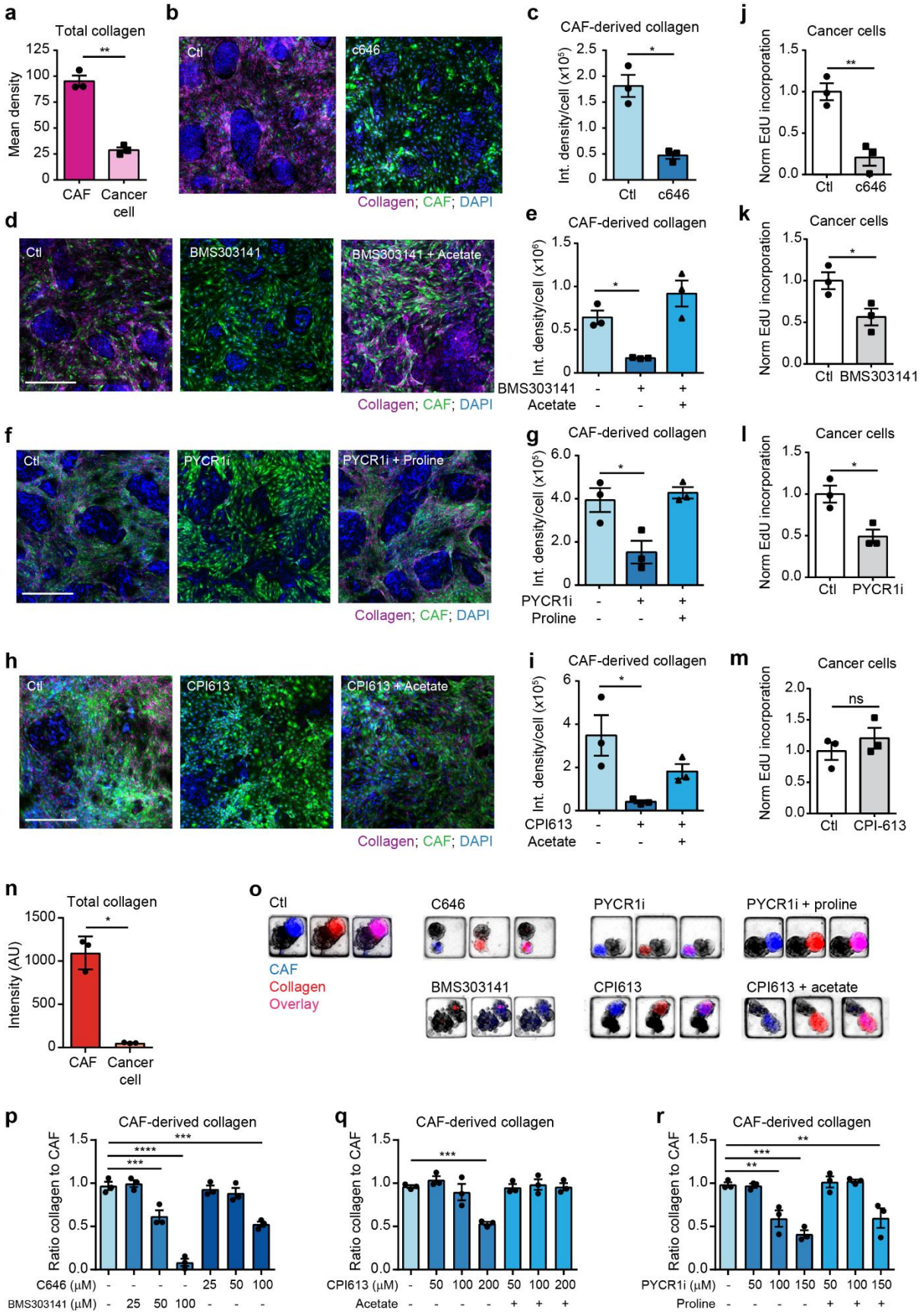
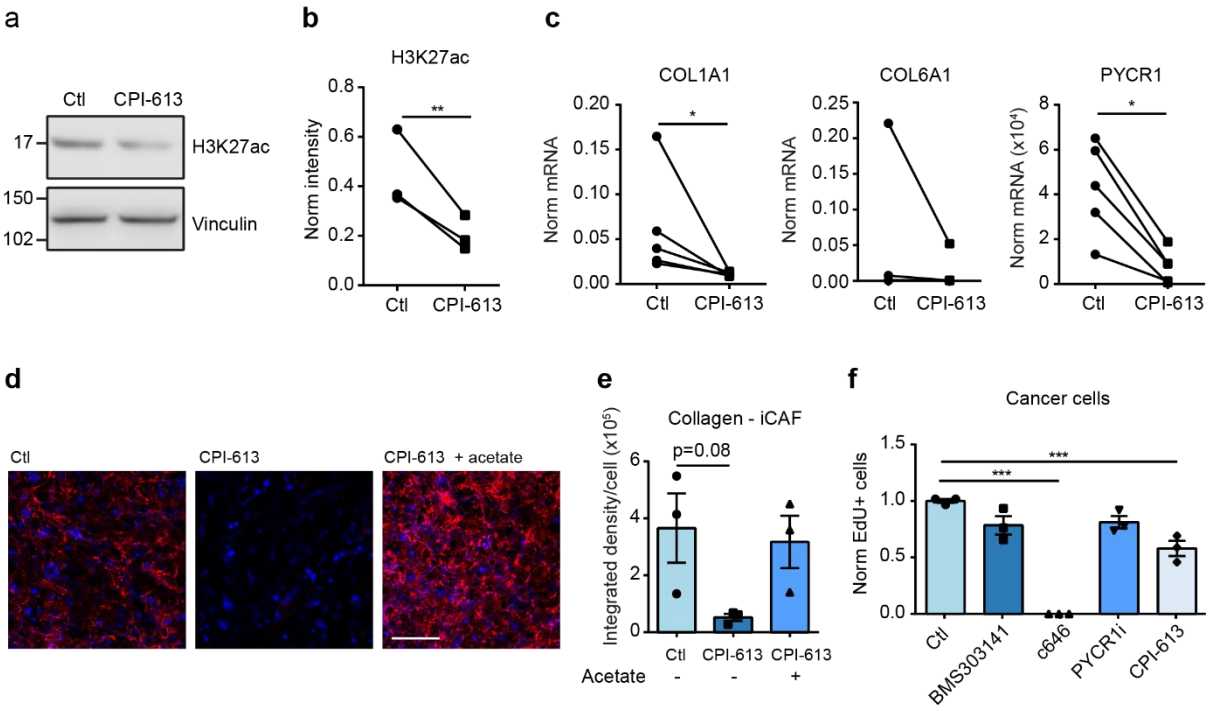


Figure 5: Pharmacological inhibition of the PDH-P300/CBP-PYCR1 pathway in CAF-cancer cell co-cultures strongly reduces ECM production and cancer cell growth

a. Quantification of collagen co-localised with iCAFs and Wood primary breast cancer cells respectively in 2D co-culture. N = 3 independent experiments. **b, c.** Representative images (b) and quantification (c) of collagen (purple) in 2D co-cultures of iCAFs (green) with Wood primary breast cancer cells ± c646 treatment for 96 h. Nuclei were visualised with DAPI. N = 3 independent experiments. **d,e.** Representative images (d) and quantification (e) of collagen (purple) in 2D co-cultures of iCAFs with Wood primary breast cancer cells treated with BMS303141 ± acetate for 96 h. N = 3 independent experiments. **f,g.** Representative images (f) and quantification (g) of collagen (purple) in 2D co-cultures of iCAFs with Wood primary breast cancer cells treated with CPI-613 ± acetate for 96 h. N = 3 independent experiments. **h,i.** Representative images (h) and quantification (i) of collagen (purple) in 2D co-cultures of iCAFs with Wood breast breast cancer cells treated with PYCR1i ± proline for 96 h. N = 3 independent experiments. **j-m.** EdU incorporation in Wood primary breast cancer cells in 2D co-culture with iCAFs control or treated with the indicated drug. N = 3 independent experiments. **n.** Quantification of collagen co-localised with iCAFs and Wood primary breast cancer cells respectively in 3D spheroid co-cultures in microfluidic devices. N = 3 independent experiments. **o.** Representative images of collagen (red) in 3D spheroid co-cultures of CAFs (blue) and Wood primary breast cancer cells treated with c646, PYCR1i ± proline, CPI-613 ± acetate or BMS303141 for 7 days. Images represent the lowest dose giving a significant change in collagen quantification (p-r). **p.** Quantification of CAF-derived collagen (purple signal in (o)) in 3D CAF-Wood co-cultures treated with c646 or BMS303141. N = 3 independent experiments **q.** Quantification of CAF-derived collagen (purple signal in (o)) in 3D CAF-Wood co-cultures treated with CPI-613 ± acetate. N = 3 independent experiments **r.** Quantification of CAF-derived collagen (purple signal in (o)) in 3D CAF-Wood co-cultures treated with PYCR1i ± proline. N = 3 independent experiments.
Collagen was visualised with the collagen-binding protein CNA35-mCherry. Int. density = Integrated density. Scale bar = 200 µm. Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. For experiments with two conditions, a two-tailed unpaired t-test with Welch's correction was used to determine the p-value. For experiments with more than two conditions, a one way ANOVA test with Dunnett's multiple comparison test was used.

Kay et al. Extended Data Figure 6



Extended Data Figure 6. Inhibition of the PDC-P300/CBP-PYCR1 pathway inhibits cancer cell growth

a,b. Representative western blot (a) and quantification (b) for H3K27ac in total lysate from iCAFs treated with CPI-613. Vinculin was used as loading control. N = 3 independent experiments. **c.** *PYCR1, COL1A1*, and *COL6A1* mRNA expression in iCAFs ± CPI-613, as measured by RT-qPCR. N = 3-5 independent experiments. Data was normalised to *18S* levels. **d,e.** Representative images (d) and quantification (e) of collagen produced by iCAFs ± CPI-613 and treated with 1 mM acetate for 72 h. N = 3 independent experiments. **f.** EdU incorporation in Wood breast cancer cells in monoculture after 48 h treatment with c646, BMS303141, PYCR1i or CPI-613. N = 3 independent experiments.
Collagen was visualised with the fluorescent collagen-binding protein CNA35-mCherry. Scale bar = 200 µm. Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. For experiments with two conditions, a two-tailed unpaired t-test with Welch's correction was used to determine the p-value. For experiments with more than two conditions, a one way ANOVA test with Dunnett's multiple comparison test was used.

Targeting PYCR1 in mammary CAFs reduces stromal collagen and tumour growth

There is increasing evidence that proline metabolism is a tumour-specific vulnerability [20, 58, 60, 61], and that targeting PYCR1 in cancer cells or reducing proline availability impedes tumorigenic growth *in vivo* [58, 61]. Our data prompted us to assess whether targeting PYCR1 is also a stromal vulnerability. As a xenograft model we used MCF10DCIS.com breast cancer cells co-transplanted subcutaneously with CAFs, because these tumours have a sizable stroma and CAFs accelerate tumour growth [62]. MCF10DCIS.com cells were co-transplanted subcutaneously with pCAFs expressing normal or reduced levels of PYCR1 (**Extended Data Figure 5a**), and tumours were harvested two weeks after transplantation. At this stage, the tumours had started growing and pCAFs had not been fully replaced with endogenous murine stroma. Microscopy analysis of the tumours showed a clear decrease in collagen deposited around pCAF silenced for PYCR1, while pCAF abundance in the tumours was not affected (**Figure 6a-c**). Notably, reducing stromal PYCR1 reduced tumour growth (**Figure 6d**). Thus, stromal PYCR1 represents a stromal vulnerability for ECM production *in vivo* and can be targeted to hamper tumour growth.

PDK2, PYCR1 and collagens are co-regulated in many cancer types

To assess the extent at which *PDK2*, *PYCR1* and collagen expression are simultaneously deregulated in tumours, we performed a meta-analysis of 4,000 human tumour samples across ten tumour types using data from The Cancer Genome Atlas (TCGA) (**Supplemental Data S6**). *PDK2* downregulation and *PYCR1* and *COL1A1* upregulation co-occurred in a substantial portion of patients in each tumour type (26%-84%) (**Figure 6e**). In addition, *PDK2* expression was downregulated (**Figure 1g**), while the expression of genes encoding enzymes for proline synthesis (*PYCR1*, *PYCR2*, *ALDH18A1*, *OAT*), but not those for proline degradation (*PRODH*), and tumour ECM components (*COL1A1*, *COL6A1* and *FN1*) was significantly upregulated in laser-captured microdissected stroma of triple negative breast cancer and high grade serous ovarian tumours compared to their corresponding normal tissues [38, 39] (**Figure 6f,g**), suggesting that this pathway is deregulated in the tumour stromal compartment. These data support that pyruvate dehydrogenase, through PDK2 downregulation, and proline synthesis may be master metabolic regulators of ECM production across multiple cancer types.

Kay et al. Figure 6

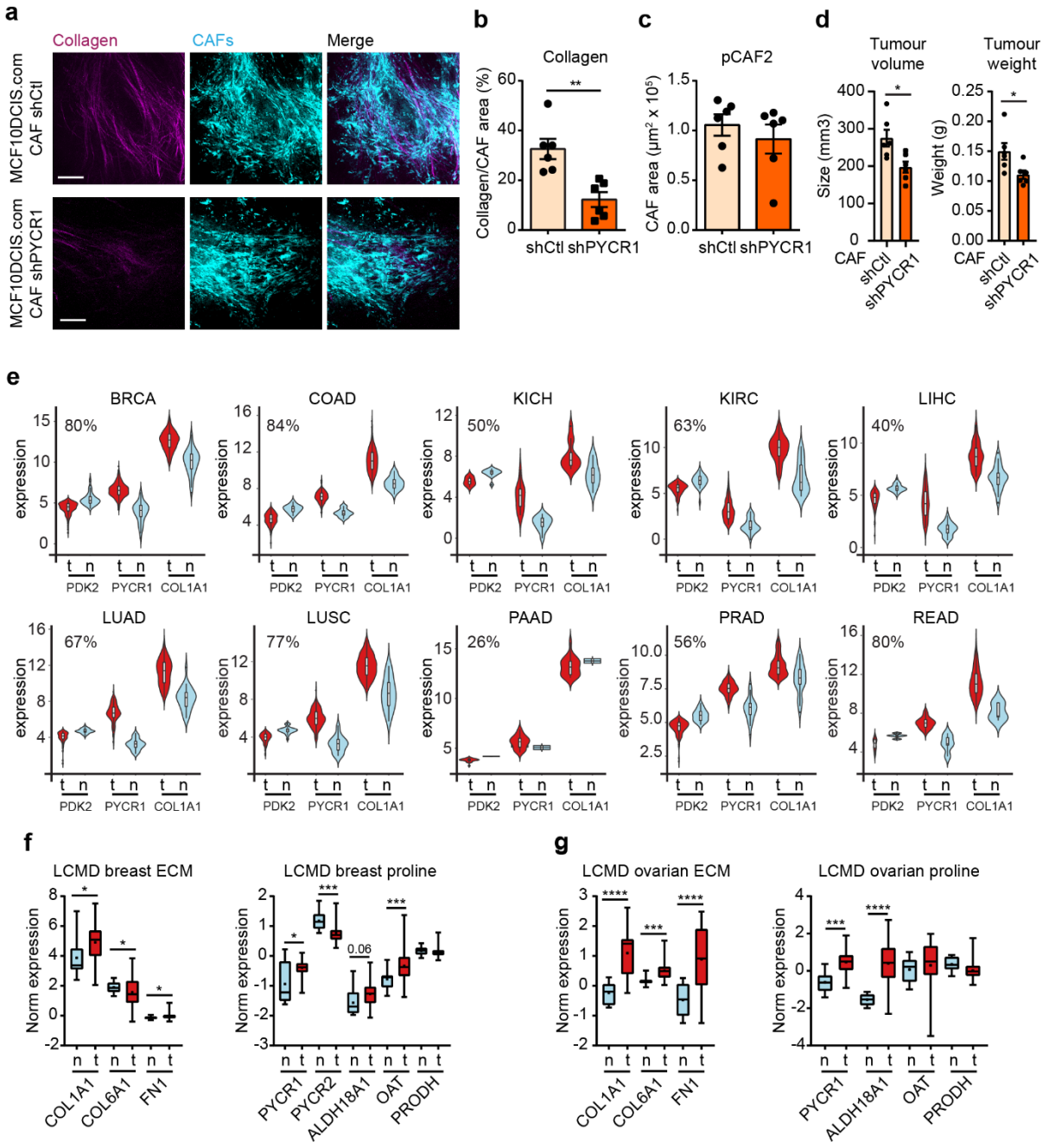


Figure 6: Reducing proline synthesis in CAFs is sufficient to reduce collagen production and reduce tumour growth

a,b. Representative images (a) and quantification (b) of collagen co-localised with GFP^+ signal coming from shCtl or shPYCR1 pCAF2s in sections of MCF10DCIS.com-CAF co-transplantation xenografts. Collagen was detected by SHG. N = 6 mice/group. Data shows average collagen density in each slice of 2 z-stack images spanning a 400 µm thick section for each tumour. **c.** Average GFP^+ pCAF2 area in sections of tumours from (e). N = 6 tumours/group. **d.** Volume (before explant) and weight of MCF10DCIS.com-CAF tumours, which were explanted 2 weeks after transplantation. N = 6 tumours/group. **e.** Violin plots showing the expression levels of PDK2, PYCR1, COL1A1 in normal (n) and tumour (t) tissues in the indicated TCGA cohorts. Tissues represented in the plots comprise those in which PDK2 was downregulated while PYCR1 and COL1A1 were upregulated. The percentage of tumours meeting the selection criteria is indicated at the top left of each plot. **f,g.** Box plot (minimum and maximum limit, line at median, "+" indicates mean) of the expression levels of ECM proteins and enzymes of the proline synthesis pathway in laser-capture microdissected sections of normal (n) and TNBC (f) or HGS ovarian (g) tumour (t) stroma. **h.** Schematic of the metabolic pathways involved in promoting and sustaining ECM production in CAFs. In red are highlighted the genetic and pharmacological approaches used in this work to inhibit key molecules of the pathway.
Error bars indicate mean ± SEM. *p ≤0.05, **p ≤0.01, ***p ≤0.001. A two-tailed unpaired t-test with Welch's correction was used to determine the p-value. Scale bar = 100 µm.

Discussion

The ECM is a central guiding force that influences all stages of cancer pathology [1-4]. Finding ways to reduce ECM production, in particular collagens, is important because it can delay tumour growth, reduce metastasis, and improve tumour perfusion for effective drug delivery and recruitment of immune cells [15-19]. CAFs with myofibroblast-like features extensively populate the stroma of solid tumours and make most of the collagen found in the tumour ECM [5]. Up to 25% of collagen amino acid residues are proline. Notably, proline is an abundant amino acid in the blood circulation (www.serummetabolome.ca) to support physiological production of collagen in connective tissues [63]. In spite of this, we show that the proline synthesised by CAFs through PYCR1 is essential to fulfil the biosynthetic needs for collagen production and to support tumour growth. Our study also identified PDC as a central enzyme that triggers an epigenetic response to enhance the expression of collagen genes, thus showing that cell metabolism is fundamental to coordinate enhanced ECM production in CAFs (**Figure 6h**).

Pyruvate dehydrogenase is the metabolic gatekeeper that links glycolysis to oxidative phosphorylation [64]. In contrast, in cultured CAFs, we show that PDC is an epigenetic regulator important to maintain hallmarks of CAF activation. CAFs had increased PDC activity, which led to an increase in acetyl-CoA. However, CAFs did not use the excess of acetyl-CoA for lipid synthesis or TCA cycle. Instead, the increased acetyl-CoA functioned as a second messenger of epigenetic regulation triggering a

molecular rewiring that required P300/CBP to increase collagen production. Our ^{13}C-tracing data showed that CAFs in culture also stored acetyl-CoA as acetyl-carnitine (**Extended Data Figure 2d**), suggesting that under nutrient constriction, they may be able to support both biosynthetic and epigenetic pathways.

Epigenetics facilitates the acquisition and maintenance of tumour phenotypes [30, 65], including hallmarks of activated CAFs [31, 66]. Because acetylation and methylation both originate from metabolites, and because cell metabolism is a targetable driver of cancer, it is important to understand how metabolism and epigenetics interact with one another in cancer. Tumour and immune cells have been the focus of many studies [29, 30], and only recently has it been shown that targeting metabolism in CAFs to influence histone methylation can reduce tumour growth and metastasis [31]. Our work has identified acetyl-CoA as another epigenetic regulator in CAFs, and showed that targeting PDC to reduce acetyl-CoA availability effectively suppresses ECM production in several in vitro models. Thus, metabolism is a central regulator of epigenetics in CAFs. It is still an open question whether metabolic pathways that control epigenetics through histone acetylation and methylation are working in synergy or alone in a context-specific manner. Notably, simply increasing acetyl-CoA availability via activating PDC in normal fibroblasts was sufficient to trigger ECM production. Thus, other sources of acetyl-CoA in tumours, such as acetate which is converted to acetyl-CoA through ACSS1 and ACSS2, need further investigation to understand whether they also contribute to tumour fibrosis.

Another open question in our study, which we are investigating further, is whether collagens and *PYCR1* genes are direct targets of P300/CBP. Previous studies using cell lines have shown that hyperacetylation of histones, including H3K27 sites, occur at enhancer regions of pro-fibrotic genes, including collagen 1 [52-54], and, in PDAC models, inhibiting bromodomain-containing proteins, which recognise acetylated histone tails and promote transcription by recruiting the transcription factor machinery, attenuated fibrosis [67, 68]. Instead, *PYCR1* expression has been found to be regulated by the amino-acid starvation response (AAR/*Atf4*) [69] and induced by shortage of proline precursors [58]. We can therefore speculate that in CAFs P300/CBP promotes collagen expression, while *PYCR1* expression increases following a reduction in proline availability due to enhanced rate of collagen synthesis.

Proline metabolism in cancer cells has been recently recognised as a tumour vulnerability and potential target for therapy [70]. Cancer cells with low levels of proline synthesis enzymes or under nutrient limiting conditions are vulnerable to proline shortage [58, 61] and proline synthesis is important to maintain mitochondrial redox homeostasis in cancer cells with IDH1 oncogenic mutations [71]. Moreover, proline catabolism is a vulnerability for metastatic cancer cells [60]. But what happens to the tumour stroma if we target proline metabolism? Recent work correlated decreased levels of PYCR1

and proline synthesis with reduced ECM production in a lung adenocarcinoma model. The authors showed that kindlin-2 binds to and stabilises PYCR1 to promote proline production, and that deleting kindlin-2 reduced PYCR1, proline levels and fibrosis [72]. However, the amount of αSMA$^+$ cells (likely to be CAFs) also diminished in kindlin-2 KO mice, making it difficult to conclude whether CAFs produced less collagen. Interestingly, another recent work showed that the delta-1-pyrroline-5-carboxylate synthase (ALDH18A1), which supports the first step of proline synthesis converting glutamate to pyrroline-5-carboxylate (**Figure 4a**), was important for collagen synthesis in TGFβ-activated lung fibroblasts [73]. By genetic and pharmacological suppression of PYCR1 in CAFs in the presence of physiological levels of proline *in vitro*, and reducing PYCR1 levels in CAFs *in vivo*, we demonstrated that inhibiting proline synthesis is sufficient to effectively reduce ECM production and hamper tumour growth.

In addition to providing proline residues for protein synthesis, proline metabolism is important to maintain redox homeostasis between the mitochondria and the cytosol [70]. Previous studies have shown that silencing ALDH18A1 to reduce proline synthesis in TGFβ-activated fibroblasts or loss of PYCR1 in fibroblasts isolated from patients carrying PYCR1 mutations caused mitochondrial damage and redox stress [73, 74]. Our study did not investigate whether targeting PYCR1 in CAFs induced redox stress. However, because silencing PYCR1 specifically affected ribosomal stalling at proline codons of several collagens, including the most abundant ECM component COL1A1, and because we could rescue collagen translation and ECM production by providing exogenous proline to the cells, we propose that a major function of proline metabolism in CAFs is to provide proline residues to maintain collagen production to make ECM.

In conclusion, we have identified PDC as a novel potent epigenetic regulator of collagen-rich ECM production and pinpointed proline metabolism as a critical link between the epigenetic regulation of collagen gene expression and protein production. So far, most of the efforts to reduce ECM production in tumours have focused on targeting pathways that are well known to trigger CAF activation [75] (e.g. TGFβ, hedgehog, PDGF [1-4]). However, these factors are pleiotropic and have a direct impact also on other cell types, raising the possibility of unexpected effects when targeted in clinical settings. Our work offers a new paradigm, whereby CAF metabolism is a major vulnerability of tumour ECM production to limit desmoplastic reaction in cancer. This may have important implications for the development of strategies to reduce the production of pro-tumourigenic ECM and the formation of desmoplastic stroma, and may be relevant to other diseases involving collagen production, such as fibrosis.

Acknowledgements

We thank Saverio Tardito, Henry Däbritz, Leo Carlin, Frederic Fercoq, Ewan McGhee, Lynn Mcgarry and the Cancer Research UK Beatson Institute core research services and advanced technology facilities, including BSU, HiCAR, BAIR, and histology; Clare Orange for histopathology services; NHS Greater Glasgow and Clyde Biorepository for providing patient samples, and the PRIDE team. The results shown in Figure 6e are based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

This work was funded by Cancer Research UK (CRUK Beatson Institute A17196, CRUK Glasgow Centre A18076, and Stand Up to Cancer campaign for Cancer Research UK A29800 (to S.Z.)). E.G. was supported by the European Union's H2020 programme (675585 Marie-Curie ITN ''SymBioSys'') and JRC for Computational Biomedicine, which is partially funded by Bayer, K.P. was supported by AMS Biotechnology (Europe) Ltd and the University of Strathclyde.

Author contributions

Conceptualization: E.J.K. and S.Z.; methodology: E.J.K., D.S., M.Z., S.L., K.B., F.L-P., J.J.K. J.S-R., C.M. and S.Z.; investigation: E.J.K., K.P., E.S., C.B., L.J.N., J.R.H-F., S.L., G.K., S.D., A.H., G.M., E.G., R.S. and S.Z.; resources: C.J., R.M.J. and M.P.; writing original draft, review and editing: E.J.K., K.P, M.Z. and S.Z.; supervision: M.P., C.M., J.J.K., F.L-P., J.S-R., K.B., M.Z. and S.Z.

Declaration of Interests

We declare no competing interests.

Methods

Cell culture

Patient-derived mammary cancer-associated fibroblasts and normal fibroblasts (pCAFs and pNFs) were isolated in house from breast cancer patient samples obtained through NHS Greater Glasgow and Clyde Biorepository. All participants gave specific consent to use their tissue samples for research. The cancer cell-derived, immortalised human mammary CAFs and NFs (iCAFs and iNFs) were kindly provided by Professor Akira Orimo (Juntendo University, Tokyo). The fibroblasts and HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine

serum, 2mM glutamine and 1% penicillin/streptomycin. MCF10DCIS.com cells were cultured in F12 medium supplemented with 5% horse serum, 2mM glutamine, 1% penicillin/streptomycin and 0.1% fungizone. Wood primary breast cancer cells were purchased from AMS Biotechnology Europe Ltd (AMSBIO) and cultured in Renaissance essential tumour medium (RETM, AMSBIO) supplemented with 5% fetal bovine serum and 1% penicillin/streptomycin. For 2D and 3D co-cultures, CAFs and cancer cells were mixed in a 1:1 ratio and cultured in a 1:1 mixture of DMEM and RETM. For SILAC proteomics experiments, iCAFs and iNFs were cultured in SILAC DMEM supplemented with 2% FBS, 8% 10 kDa dialysed FBS (PAA), 2 mM glutamine and 1% penicillin/streptomycin. SILAC DMEM used for the 'light' labelled cells contained 84 mg/l L-arginine and 146 mg/l L-lysine (Sigma), whereas the medium for the 'heavy' labelled cells contained 84 mg/l $^{13}C_6{}^{15}N_4$ L-arginine and 175 mg/l $^{13}C_6{}^{15}N_2$ L-lysine (Cambridge Isotope Laboratories). Cells were regularly tested for mycoplasma and MCF10DCIS.com cells authenticated. The following inhibitors were used to treat cells in culture: c646 (Sigma), A485 (Tocris Bioscience), BMS303141 (Sigma) and CPI-613 (Sigma). The PYCR1i was made as previously described [56].

pCAF and pNF isolation and immortalisation

pCAFs and pNFs were isolated in house from patient samples using previously described methods [34-36] . From each patient, pCAFs were isolated from breast tumour tissue and pNFs from normal, tumour adjacent tissue. Patient samples were obtained complying with ethical regulations through the National Health Service (NHS) Greater Glasgow and Clyde Biorepository. All participants gave specific consent to use their tissue samples for research. pCAF/NF2 were from a ER+, PR+, HER2- breast cancer patient and pCAF/NF3 were from a triple negative breast cancer patient. The pCAFs and pNFs were immortalised using a human telomerase reverse transcriptase (hTERT)-expressing plasmid (pIRES2-hygro), kindly provided by Dr. Fernando Calvo (IBBTEC, Santander). Lentivirus containing the hTERT plasmid was generated in HEK293T cells. Two rounds of viral transduction in fibroblasts were performed on consecutive days. Cells were selected using 50 µg/ml hygromycin.

Western blotting analysis

Cells were lysed in SDS buffer (2% SDS, 100 mM TrisHCl pH 7.4), incubated at 95 °C for 5 min, sonicated using a metal tip (Soniprep 150, MSE) and centrifuged at 16000 x g for 10 min. Protein concentration was determined using Optiblot Bradford reagent (Abcam). 20-25 µg of proteins were separated using 4-12% gradient NuPAGE Novex Bis-Tris gel (Life technologies). Protein transfer was performed on methanol-activated PVDF or Nitrocellulose membrane. Membrane was blocked for 1 hr in 3% BSA (Sigma) in TBST at RT and incubated with primary antibodies overnight at 4°C. The following primary antibodies were used: PDHA1 E1-alpha subunit, phospho S293 (1:2000, abcam ab177461), PDHA1 E1-

alpha subunit (1:1000, abcam ab110334), PYCR1 (1:1000, Proteintech 22150-1-AP), Histone H3, acetyl K27 (1:1000, abcam ab4729), Collagen VI (1:1000, abcam ab182744). Vinculin (1:2000, Sigma V9131), GAPDH (1:1000, Santa Cruz, sc48167), β-tubulin (1:1000, abcam ab179513) were used as loading controls for the experimental antibodies. The membrane was incubated with HRP-conjugated secondary antibodies (1:5000, NEB) for 45 mins at RT. Western blot images were acquired using a myECL Imager (Thermo Scientific).

PDH activity assay

PDH activity was measured using the Pyruvate dehydrogenase Enzyme Activity Microplate Assay Kit (abcam ab109902) according to the manufacturer's protocol.

EDU proliferation assay

Cells were seeded on 13 mm glass coverslips. Following 48 h of drug treatment, cells were incubated in 1 µM EDU for 2 h and fixed in 4% PFA. EDU was fluorescently labelled using the Click-iT™ EdU Cell Proliferation Kit (Life Technologies) according to the manufacturers' protocol, and nuclei were counterstained with DAPI. Images were acquired using a Zeiss 710 confocal microscope and ImageJ was used to count the number of total nuclei and EDU positive nuclei.

ECM preparation

Cells were seeded at 100% confluence on 0.2% gelatine, which was crosslinked using 1% glutaraldehyde, then were cultured for 3-7 days. The ECM was decellularised with extraction buffer (20mM NH_4OH, 0.5% Triton X-100 in PBS) until no intact cells were visible but the ECM remained on the dish. ECM was washed in PBS with Ca^{2+} and Mg^{2+}, collected and lysed in SDS buffer (4% SDS, 0.1 M DTT, Tris-HCl pH 7.4).

Cell transfection and infection

For transient expression or siRNA knockdown, $2x10^6$ fibroblasts were harvested and used in each transfection with a Nucleofector device (Lonza) according to the manufacturer's protocol using the program T-20 and the Amaxa kit R (Lonza). Cells were transfected with 1-3 nM non-targeting siRNA as a control (D-001810-10-05, GE Healthcare Dharmacon) or with siRNAs targeting *PDK2* and *PYCR1* (Dharmacon, pool of 4), or with 5µg pGCA-$PDK2^{N255A}$ or pGCA-$PDK2^{WT}$ (kindly provided by Prof. Angus McQuibban, University of Toronto [47]). Cells were used for experiments 48-72h after transfection.
For stable knock down of PYCR1, shPYCR1 (shPCYR1 #1: CCGGTGAGAAGAA GCTGTCAGCGTTCTCGAGAACGCTGACAGCTTCTTCTCATTTTTG, shPYCR1 #2: CCGGCACAGTTTCTGC TCTCAGGAACTCGAGTTCCTGAGAGCAGAAACTGTGTTTTTG) and shCTL (Sigma, Mission shRNA)

lentivirus was generated in HEK293 cells. Two rounds of viral transduction in pCAF2 were performed on consecutive days. Cells were selected using 2 µg/ml puromycin.

Reverse transcriptase polymerase chain reaction (RT-qPCR)

Primers for RT-qPCR were designed using Primer Blast (NCBI database): *PDK2*: CGGGGACCACAACCAAAGTC (forward) GCTGGATCCGAAGTCCAGAAA (reverse), *PYCR1*: CCCCGCCTACGCATTCACA (forward) GCGCGTTGGAAGTCCCATCT (reverse), *COL1A1*: TGAAGGGACACAGAGGTTTCAG (forward) GTAGCACCATCATTTCCACGA (reverse), *COL6A1*: AGCAAGTGTGCTGCTCCTTC (forward) CTTCCAGGATCTCCGGCTTC (reverse). mRNA levels were normalised to the following housekeeping genes: *TBP2*: AGTGACCCAGCATCACTGTTT (forward) TAAGGTGGCAGGCTGTTGTT (reverse), *18S*: AGGAATTGACGGAAGGGCAC (forward) GGACATCTAAGGGCATCACA (reverse) or *ACTB*: GGCATGGGTCAGAAGGATT (forward) ACATGATCTGGGTCATCTTCTC (reverse). Total RNA was isolated and treated with DNAse using the RNEasy mini kit (Qiagen) according to the manufacturer's instructions. Complementary DNA (cDNA) was synthesised from 1 µg RNA using an iScript kit (BioRad). cDNA was diluted 1:5 and 2 µl was used in each RT-qPCR reaction with 10 µl iTAQ Universal SYBR green supermix (BioRad) and 400 nM primers. Reactions were performed using a Quant Studio 3 PCR machine (Thermo Scientific).

MS-proteomic analysis

For the total proteome, cells were lysed in SDS buffer (2% SDS, 100 mM TrisHCl pH 7.4). For the proteome with c646 treatment, proteins were precipitated with acetone and redissolved in urea buffer (6M urea, 2M thiourea, 10 mM TCEP, 40 mM CAA, 75 mM NaCl, 50 mM Tris-HCl). The proteins were then trypsin digested. For SILAC experiments, equal quantities of heavy and light samples were mixed. For the total proteome that was used to normalise peptide acetylation levels, lysates were either in-gel digested with trypsin or, after trypsin digestion, peptides were fractionated using high pH reverse phase fractionation. The proteins were desalted by C18 StageTip [76] prior to MS analysis.

For phosphorylated peptide enrichment, trypsin-digested peptides were acidified to pH 2.6 and acetonitrile (ACN) was added to a final concentration of 30%. The peptides were fractionated using an Akta system into 6 equal fractions, using an increasing concentration of KCl in 5 mM KH_2PO_4 to a final concentration of 350 mM KCl. Each fraction was then enriched for phosphorylated peptides by incubation with TiO_2 beads (GL Sciences) in the presence of 2,5-dihydroxybenzoic acid [77]. Phosphorylated peptides were eluted with 15% ammonium hydroxide and 40% acetonitrile (ACN), and desalted by C18 StageTip.

For acetylated peptide enrichment, the deacetylase inhibitors nicotinamide (10 mM) and trichostatin A (1 µM) were added to the lysis buffer. Different protocols were used to prepare acetylated peptides.

i) Cells were lysed in RIPA buffer (50 mM TrisHCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.1% sodium deoxycholate). Proteins were precipitated with acetone, redissolved in urea buffer and quantified by Bradford assay. Equal quantities of heavy and light labelled proteins were combined and trypsin digested. Peptides were desalted by C18 SepPak filtration and resuspended in immunoprecipitation (IAP) buffer (50 mM MOPS; pH 7.2, 10 mM Na-phosphate, 50 mM NaCl). Acetylated peptides were enriched using anti-acetyllysine antibody (Acetyl Lysine Antibody, Agarose, ImmuneChem). Up to three consecutive incubations were performed to maximise peptide recovery. Acetylated peptides were eluted with acidified water (0.1% TFA in water). ii) Lysis and protein digestion was performed as in i), but the enrichment for acetylated peptides was performed with PTMScan Acetyl-Lysine Motif Kit (Cell Signalling Technology #13416) according to the manufacturers' protocol. iii) Subcellular fractionation of the cells was performed (Cell Fractionation Kit, Standard, Abcam, ab109719) according to the manufacturers' protocol. Proteins recovered in the three fractions, nuclear, cytosolic and mitochondrial, were digested with trypsin and acetylated peptides enriched using anti-acetyllysine antibody (Acetyl Lysine Antibody, Agarose, ImmuneChem).

For ECM analysis, cell-derived ECM was prepared as above. To analyse glutamine-derived proline incorporation into collagen, the cells were cultured in media containing 2 mM $^{13}C_5$-glutamine for 72 h prior to ECM collection. Each sample was separated on 4–12% gradient NuPAGE Novex Bis-Tris gel (Life Technologies). The gel was sliced into 3 fractions, and each fraction was in-gel digested with trypsin.

Peptides were resuspended in 1% TFA, 0.2% acetic acid or formic acid buffer and injected on an EASY-nLC (Thermo Fisher Scientific) coupled online to a mass spectrometer. Peptides were separated on a 20-cm fused silica emitter (New Objective) packed in-house with reverse-phase Reprosil Pur Basic 1.9 µm (Dr. Maisch GmbH). Peptides were eluted with a flow of 300 nl/min from 5% to 30% of buffer B (80% ACN, 0.1% formic acid) in a 60-min linear gradient. Eluted peptides were injected into an Orbitrap Elite, Q-Exactive HF or Orbitrap Fusion Lumos (Thermo Fisher Scientific) via electrospray ionisation. MS data were acquired using XCalibur software (Thermo Fisher Scientific).

MS-proteomic data analysis

The MS .raw files were processed with MaxQuant software [78] and searched with the Andromeda search engine with the following settings: minimal peptide length 7 amino acids, fixed modification Carbamidomethyl (C) and variable modifications Acetyl (Protein N-term) and Oxidation (M). For the acetylome and phosphoproteome, Acetyl (K) and Phospho (STY) were added as variable modifications respectively. For the tracing experiments with $^{13}C_5$-glutamine, ^{13}C-labelled proline was added as variable modification. Specificity for trypsin cleavage was required and maximum 2 missed cleavages were allowed. For SILAC experiments, multiplicity was set to 2, where the light labels were Arg0 and

Lys0 and the heavy labels were Arg10 and Lys8. For label free (LFQ) experiments, the LFQ setting was enabled. The false discovery rates (FDRs) at the protein and peptide level were set to 1%.

Perseus [79] (version 1.5.0.36 for the phosphoproteome, 1.5.5.1 for the acetylome and corresponding proteome, 1.6.2.2 for total proteome) was used for downstream analysis. The data were filtered to remove potential contaminants, reverse peptides which match a decoy database, and proteins only identified by site. To ensure unambiguous identification, only proteins identified with at least one unique peptide were considered. For SILAC experiments, the SILAC ratio was used for the analysis. Ratios from the 'Reverse' experiment were inverted. Then, SILAC ratios were transformed by log2 and intensities by log10. For the acetylome, peptide acetylation levels were normalised by total protein abundance (**Supplementary Data S2**). For $^{13}C_5$-glutamine tracing experiment only peptides in which all the proline was ^{13}C-labelled were considered for analysis.

Estimation of kinase activities

KinAct [80], is a computational method used to predict kinase-activity scores from MS-based data. It infers an activity score for each protein kinase based on the regulation levels of phosphorylation events catalysed by this specific kinase. The method relies on prior knowledge of kinase/phosphatase-to-substrate (k/p-s) relations and the KSEA (Kinase-Substrate Enrichment Analysis) method [81] for the kinase activity estimation.

Currently, there are multiple freely available databases that collect k/p-s interactions which are experimentally verified or manually curated. Some of them are integrated into OmniPath [82], a comprehensive collection of signalling resources. There are also resources in which the k/p-s relations are inferred based on consensus kinase recognition motifs and other information, such as NetworKIN [83]). The KSEA method integrates the information from such databases with the $\log_2$ transformed fold changes from mass-spectrometry data to compute enrichment scores for each kinase together with a significance value for each score.

In the KSEA method, the score is equal to the mean of the fold changes of each phospho-measurement of the substrate set mS of a specific kinase. The significance of the score, on the other hand, is calculated from a z-statistic as: $z = (mS - mP\sqrt{m})/\delta$, where mP is the mean $\log_2$ fold-change of the complete data-set, m is the size of the substrate set mS (how many substrates for a kinase) and δ is the standard deviation of the $\log_2$ fold-change values for the whole dataset. Significance is then estimated from the z-statistic and the estimated p-values for each kinase are then adjusted via the Benjamini-Hochberg correction method.

KinAct was applied to the phosphoproteomic SILAC-labelled NF (normal fibroblast) and CAF (cancer-associated fibroblast) data which was performed in two independent experiments. The $\log_2$ ratios of

the two experiments were averaged and inputinto the KinAct pipeline. KinAct analysis was performed over the OmniPath and NetwroKIN prior knowledge. For the OmniPath prior knowledge, kinase-substrate relations coming from PhosphoSitePlus [51] and SIGNOR [84] were used. Results from both resources yielded similar results where PDK2 was observed to be significantly downregulated in both cases. Codes of the analysis together with usage documentation are made available in https://github.com/saezlab/iCAF_KinAct.

Metabolites extraction and LC-MS analysis

For tracing experiments, cells were labelled for 24h with $^{13}C_6$-glucose, $^{13}C_5$-glutamine, $^{13}C_3$-pyruvate or $^{13}C_{16}$-palmitate. Cells were washed with ice-cold PBS and intracellular metabolites were extracted with an aqueous solution of 50% methanol and 30% acetonitrile. Cellular extracts were centrifuged at 16,000 x g for 5mins, and the supernatant was analysed using a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific) in combination with a Thermo Ultimate 3000 HPLC system. 5 µl of cell extract was injected and the metabolites were separated over a 15 min mobile phase gradient from an initial ACN content of 80% ACN with 20% ammonium bicarbonate (pH 9.2) decreasing to 20% ACN with a flow rate of 200 µL/min. The metabolites were detected over a period of 25 min using the Q-Exactive mass spectrometer across a mass range of 75-1000 m/z and at a resolution of 35,000 (at 200 m/z). To detect acetyl-coA, a single ion monitoring (SIM) method was employed. The Q-Exactive mass spectrometer was used to monitor the three masses for acetyl-coA labelled +0, +1 or +2 (810.1331, 811.1364 and 812.13976 m/z) with an isolation window of 0.7 m/z for each isotope. Peak identification and area quantification were carried out using TraceFinder software by comparison of the retention time and exact ion mass to that of authentic standards.

Cholesterol extraction and GC-MS analysis

Cells were washed with ice-cold PBS and metabolites were extracted with 1:9 (v/v) water: methanol buffer and 20 µl lathosterol (100 ng/µl) was added as an internal standard. Saponification to obtain the total cholesterol pool and subsequent GC-MS analysis was carried out as in McGregor et al. [85].

Trans-fatty acid extraction and LC-MS analysis

Cells were washed with ice-cold PBS and metabolites were extracted with 1:1 (v/v) PBS: methanol buffer. SPLASH lipidomix internal standard mix (Avanti Polar Lipids) was added at 1 µl/10^5 cells. Subsequent chloroform extraction of lipids and LC-MS analysis were carried out as in McGregor et al. [85].

Ribosome profiling and diricore analysis

$30x10^6$ cells were treated with cycloheximide (100 µg/ml) for 5 minutes and lysed in buffer A (20 mM Tris-HCl, pH 7.8, 100 mM KCl, 10 mM $MgCl_2$, 1% Triton X-100, 2 mM DTT, 100 µg/ml cycloheximide, 1X complete protease inhibitor). Lysates were treated with 2 U/µl of RNase I (Ambion) for 45 min at room temperature. Lysates were fractionated on a linear sucrose gradient and the fractions enriched in monosomes were pooled. Ribosome protected fragments (RPFs) were purified using Trizol reagent (Invitrogen). Library preparation and differential ribosome codon reading (diricore) analysis were performed according to the method previously described [58].

Collagen quantification in mono culture and 2D co-cultures

Cells were seeded at 100% confluence on glass coverslips, either as a CAF monoculture or as a 1:1 coculture of CAFs and Wood primary breast cancer cells. The cells were cultured for 4-7 days to allow accumulation of matrix. Cells were incubated with 1 µM of the native collagen binding protein CNA35 labelled with fluorescent dye mCherry (CNA35-mCherry) [86] for 1 hr to label collagen, then fixed in 4% PFA and counterstained with DAPI. Images were taken on the Zeiss 710 confocal microscope. Regions of CAFs were defined and collagen staining was quantified using ImageJ software. Cell number for normalisation of collagen quantification was calculated by counting the number of DAPI positive nuclei using ImageJ software.

Microfluidic device design and preparation

Microfluidic devices were fabricated using previously established soft lithography methods and used to culture spheroids [87]. Multilayer devices were composed of arrays of microfluidic channels, each of which was connected by two open wells. Within each channel, an array of microwells of dimension 150x150x150 μm^3 were situated below the channel level [87]. In short, polydimethylsiloxane (PDMS) prepolymer (Sylgard 184, Dow Corning) and curing agent were combined in a 1:10 ratio and poured onto patterned silicon wafers. Wafers were placed inside a desiccator for degassing, prior to incubation at 85°C for a minimum duration of 3 hours. Once cured, the PDMS was removed from the wafers and open wells created using a 4mm surgical biopsy punch (Miltex). Devices were cleaned and exposed to an oxygen plasma (Pico plasma cleaner, Diener electronic) to permanently bond the upper and lower PDMS layers together. Devices were incubated with a solution of 1% Synperonic F108 solution (Sigma Aldrich) to achieve ultra-low adhesion conditions.

3D co-culture in microfluidic devices

Cells were seeded at a 1:1 ratio of Wood primary breast cancer cells:CAFs into devices at a concentration of 7×10^6 cells/ml to form spheroids, with each microfluidic channel containing at least 32 spheroids of similar dimension (~80 µm diameter) for analysis. A cell suspension was injected in

the open wells, which flowed into the microfluidic channels until they remained trapped into the microwells, as previously described [87]. Spheroids were formed within 24-48 hours. Cells that were not trapped into the microwells were removed from the device. Cells were cultured in a 1:1 mix of the supplier recommended complete culture media for the two cell types: RETM for the primary breast cancer cells and DMEM for the CAF. Media with and without enzyme inhibitors and rescue agents was exchanged every 48 hours.

Enzyme inhibitor treatment in microfluidic devices

Spheroids were treated with BMS303141 (Concentrations: 25 µM, 50 µM and 100 µM), c646 (Concentrations: 25 µM, 50 µM and 100 µM), PYCR1i (Concentrations: 50 µM, 100 µM and 150 µM) or CPI-613 (Concentrations: 50 µM, 100 µM and 200 µM). Inhibitor action on cells was mitigated with administration of either 1 mM acetate or 500 µM proline, used as rescue agents and applied together with the enzyme inhibitor treatment. Inhibitor treatment was administered every second day for one week beginning 24 hours after cell seeding. Control experiments were performed for each experimental setup. All experiments were performed in triplicates, with each experiments performed on at least 32 spheroids.

Collagen quantification in 3D co-cultures

For visualisation of total collagen, 1 µM CNA35-mCherry was incubated with the cells for a 2-hour period. After this, cells were washed twice with PBS to ensure removal of any residual staining solution. PBS was then added again prior to imaging the devices.

An inverted microscope (Observer A1, Zeiss) connected to an Orca Flash 4.0 camera (Hamamatsu) was used to acquire bright field images of spheroids every 24–48 hours. Epifluorescence microscopy was performed immediately after cell staining and image analysis carried out using ZEN Blue and Fiji. czi files from ZEN Blue were split into separate channels and converted into TIFF files in Fiji. Images from each channel were normalized to the same threshold range. Due to the distinct dissociation of the two cell types after 7 days of co-culture, it was possible to quantify the spheroid areas and perimeters of both CAFs (expressing green fluorescent protein) and the cancer cells (from bright field images). Collagen deposition was estimated from fluorescent images (mCherry) and was plotted as a ratio of collagen area vs CAF spheroid area. Signal intensity was measured in these regions correcting for background signal.

MCF10DCIS.com-CAF xenograft

All mouse procedures were in accordance with ethical approval from University of Glasgow under the revised Animal (Scientific Procedures) Act 1986 and the EU Directive 2010/63/EU authorised through Home Office Approval (Project licence number 70/8645).

$1.5x10^6$ pCAF2s expressing shCtl or shPYCR1 and $5x10^5$ MCF10DCIS.com in 200 µl 50% growth factor reduced phenol red free Matrigel (BD Biosciences) in PBS were injected subcutaneously into the flank of 8 week old female BALB/c nude mice (Charles River). Six mice per group were transplanted. Mice were randomly allocated to the two groups. The mice were sacrificed 14 days after inoculation and the tumours were excised, weighed and fixed in 4% PFA. The tumours were sliced into 400 µm sections and Z-stacks of each section were captured. The collagen was imaged using second harmonic generation (SHG) microscopy in combination with confocal microscopy to detect GFP-expressing fibroblasts. Using an ImageJ macro, regions of human CAFs were defined for each slice of the Z-stack and the area of collagen surrounding the CAFs was quantified.

Gene expression analysis

The breast cancer (GSE90505) and high grade serous ovarian cancer (GSE40595) microarray datasets were downloaded from the Gene Expression Omnibus (GEO) using the R statistical environment, version 3.5.0, and the Bioconductor package GEOquery, version 2.40.0 [88]. Differential gene probe expression analysis was done using the linear models and differential expression for microarray data (Limma) package version 3.29.8 [89].

TCGA data analysis

Harmonised RNA Sequencing Data for breast invasive carcinoma (BRCA), pancreatic adenocarcinoma (PAAD), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), rectum adenocarcinoma (READ), colon adenocarcinoma (COAD), prostate adenocarcinoma (PRAD), kidney chromophobe (KICH), kidney renal clear cell carcinoma (KIRC) was obtained from The Cancer Genome Atlas (TCGA) Research Network: https://www.cancer.gov/tcga. Each cancer type was separately analysed using the R programming environment [90] with RStudio [91]and the data were normalised using Limma [89]. Mean *PDK2*, *PYCR1* and *COL1A1* expression was calculated for the normal tissue samples in each tumour type, and patients stratified according to the expression level of these genes in the tumour samples relative to the normal sample means. Samples plotted in Figure 6 correspond to those in which *PYCR1* and *COL1A1* exceeded the normal-tissue mean, while *PDK2* was below their respective normal-tissue means. Data were visualised using the Tidyverse package [92].

Statistical analysis

GraphPad Prism version 7.0 was used for statistical analysis. For experiments with two conditions, a two-tailed unpaired t-test with Welch's correction was used to determine the P value. For experiments with more than two conditions, a one way ANOVA test with Dunnett's multiple comparison test was used. $P \leq 0.05$ was considered significant. All graphs show the mean ± SEM of at least 3 biological replicates (independent experiments) unless otherwise stated. For MS-proteomic analysis, Perseus software was used for statistical analysis. A one-sample t-test for SILAC experiments or a 2-sample t-test for LFQ experiments was used to determine significantly regulated proteins and enable data visualisation as a volcano plot.

Data availability

The .raw MS files and search/identification files obtained with MaxQuant have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org/cgi/GetDataset) via the PRIDE partner repository [93] with dataset identifier PXD018343. .raw data are associated to Figure 1a, Figure 2b,c and Figure 3a. All unique materials used are readily available from the authors.

References

1. Kai, F., Drain, A.P. & Weaver, V.M. The Extracellular Matrix Modulates the Metastatic Journey. *Dev Cell* **49**, 332-346 (2019).
2. Alexander, J. & Cukierman, E. Stromal dynamic reciprocity in cancer: intricacies of fibroblastic-ECM interactions. *Curr Opin Cell Biol* **42**, 80-93 (2016).
3. Pickup, M.W., Mouw, J.K. & Weaver, V.M. The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep* **15**, 1243-1253 (2014).
4. Nissen, N.I., Karsdal, M. & Willumsen, N. Collagens and Cancer associated fibroblasts in the reactive stroma and its relation to Cancer biology. *J Exp Clin Cancer Res* **38**, 115 (2019).
5. Tian, C. *et al.* Proteomic analyses of ECM during pancreatic ductal adenocarcinoma progression reveal different contributions by tumor and stromal cells. *Proc Natl Acad Sci U S A* **116**, 19609-19618 (2019).
6. Kalluri, R. The biology and function of fibroblasts in cancer. *Nature reviews. Cancer* **16**, 582-598 (2016).
7. Ohlund, D., Elyada, E. & Tuveson, D. Fibroblast heterogeneity in the cancer wound. *J Exp Med* **211**, 1503-1523 (2014).
8. Santi, A., Kugeratski, F.G. & Zanivan, S. Cancer Associated Fibroblasts: The Architects of Stroma Remodelling. *Proteomics* (2017).
9. Gill, J.K., Maskarinec, G., Pagano, I. & Kolonel, L.N. The association of mammographic density with ductal carcinoma in situ of the breast: the Multiethnic Cohort. *Breast Cancer Res* **8**, R30 (2006).
10. Minchinton, A.I. & Tannock, I.F. Drug penetration in solid tumours. *Nature reviews. Cancer* **6**, 583-592 (2006).
11. Barcus, C.E. *et al.* Elevated collagen-I augments tumor progressive signals, intravasation and metastasis of prolactin-induced estrogen receptor alpha positive mammary tumor cells. *Breast Cancer Res* **19**, 9 (2017).

12. Provenzano, P.P. *et al.* Collagen density promotes mammary tumor initiation and progression. *BMC Med* **6**, 11 (2008).
13. Iyengar, P. *et al.* Adipocyte-derived collagen VI affects early mammary tumor progression in vivo, demonstrating a critical interaction in the tumor/stroma microenvironment. *J Clin Invest* **115**, 1163-1176 (2005).
14. Park, J. & Scherer, P.E. Adipocyte-derived endotrophin promotes malignant tumor progression. *J Clin Invest* **122**, 4243-4256 (2012).
15. Liu, J. *et al.* TGF-beta blockade improves the distribution and efficacy of therapeutics in breast carcinoma by normalizing the tumor stroma. *Proc Natl Acad Sci U S A* **109**, 16618-16623 (2012).
16. Olive, K.P. *et al.* Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. *Science* **324**, 1457-1461 (2009).
17. Diop-Frimpong, B., Chauhan, V.P., Krane, S., Boucher, Y. & Jain, R.K. Losartan inhibits collagen I synthesis and improves the distribution and efficacy of nanotherapeutics in tumors. *Proc Natl Acad Sci U S A* **108**, 2909-2914 (2011).
18. Polydorou, C., Mpekris, F., Papageorgis, P., Voutouri, C. & Stylianopoulos, T. Pirfenidone normalizes the tumor microenvironment to improve chemotherapy. *Oncotarget* **8**, 24506-24517 (2017).
19. Takai, K., Le, A., Weaver, V.M. & Werb, Z. Targeting the cancer-associated fibroblasts as a treatment in triple-negative breast cancer. *Oncotarget* **7**, 82889-82901 (2016).
20. Olivares, O. *et al.* Collagen-derived proline promotes pancreatic ductal adenocarcinoma cell survival under nutrient limited conditions. *Nature communications* **8**, 16031 (2017).
21. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
22. Vander Heiden, M.G. & DeBerardinis, R.J. Understanding the Intersections between Metabolism and Cancer Biology. *Cell* **168**, 657-669 (2017).
23. Martinez-Outschoorn, U.E. *et al.* Oxidative stress in cancer associated fibroblasts drives tumor-stroma co-evolution: A new paradigm for understanding tumor metabolism, the field effect and genomic instability in cancer cells. *Cell Cycle* **9**, 3256-3276 (2010).
24. Sousa, C.M. *et al.* Pancreatic stellate cells support tumour metabolism through autophagic alanine secretion. *Nature* **536**, 479-483 (2016).
25. Yang, L. *et al.* Targeting Stromal Glutamine Synthetase in Tumors Disrupts Tumor Microenvironment-Regulated Cancer Cell Growth. *Cell Metab* **24**, 685-700 (2016).
26. Bertero, T. *et al.* Tumor-Stroma Mechanics Coordinate Amino Acid Availability to Sustain Tumor Growth and Malignancy. *Cell Metab* **29**, 124-140 e110 (2019).
27. Zhang, D. *et al.* Metabolic reprogramming of cancer-associated fibroblasts by IDH3alpha downregulation. *Cell Rep* **10**, 1335-1348 (2015).
28. Wang, W. *et al.* Effector T Cells Abrogate Stroma-Mediated Chemoresistance in Ovarian Cancer. *Cell* **165**, 1092-1105 (2016).
29. Etchegaray, J.P. & Mostoslavsky, R. Interplay between Metabolism and Epigenetics: A Nuclear Adaptation to Environmental Changes. *Mol Cell* **62**, 695-711 (2016).
30. Kinnaird, A., Zhao, S., Wellen, K.E. & Michelakis, E.D. Metabolic control of epigenetics in cancer. *Nature reviews. Cancer* **16**, 694-707 (2016).
31. Eckert, M.A. *et al.* Proteomics reveals NNMT as a master metabolic regulator of cancer-associated fibroblasts. *Nature* **569**, 723-728 (2019).
32. Campbell, S.L. & Wellen, K.E. Metabolic Signaling to the Nucleus in Cancer. *Mol Cell* **71**, 398-408 (2018).
33. Pietrocola, F., Galluzzi, L., Bravo-San Pedro, J.M., Madeo, F. & Kroemer, G. Acetyl coenzyme A: a central metabolite and second messenger. *Cell Metab* **21**, 805-821 (2015).

34. Kojima, Y. *et al.* Autocrine TGF-beta and stromal cell-derived factor-1 (SDF-1) signaling drives the evolution of tumor-promoting mammary stromal myofibroblasts. *Proc Natl Acad Sci U S A* **107**, 20009-20014 (2010).
35. Hernandez-Fernaud, J.R. *et al.* Secreted CLIC3 drives cancer progression through its glutathione-dependent oxidoreductase activity. *Nat Commun* **8**, 14206 (2017).
36. Kugeratski, F.G. *et al.* Hypoxic cancer-associated fibroblasts increase NCBP2-AS2/HIAR to promote endothelial sprouting through enhanced VEGF signaling. *Sci Signal* **12** (2019).
37. Saunier, E., Benelli, C. & Bortoli, S. The pyruvate dehydrogenase complex in cancer: An old metabolic gatekeeper regulated by new pathways and pharmacological agents. *Int J Cancer* **138**, 809-817 (2016).
38. Saleh, S.M.I. *et al.* Identification of Interacting Stromal Axes in Triple-Negative Breast Cancer. *Cancer Res* **77**, 4673-4683 (2017).
39. Yeung, T.L. *et al.* TGF-beta modulates ovarian cancer invasion by upregulating CAF-derived versican in the tumor microenvironment. *Cancer Res* **73**, 5016-5028 (2013).
40. Shi, G. & McQuibban, G.A. The Mitochondrial Rhomboid Protease PARL Is Regulated by PDK2 to Integrate Mitochondrial Quality Control and Metabolism. *Cell Rep* **18**, 1458-1472 (2017).
41. Chen, J. *et al.* Compartmentalized activities of the pyruvate dehydrogenase complex sustain lipogenesis in prostate cancer. *Nat Genet* **50**, 219-228 (2018).
42. Sutendra, G. *et al.* A nuclear pyruvate dehydrogenase complex is important for the generation of acetyl-CoA and histone acetylation. *Cell* **158**, 84-97 (2014).
43. Berwick, D.C., Hers, I., Heesom, K.J., Moule, S.K. & Tavare, J.M. The identification of ATP-citrate lyase as a protein kinase B (Akt) substrate in primary adipocytes. *J Biol Chem* **277**, 33895-33900 (2002).
44. Potapova, I.A., El-Maghrabi, M.R., Doronin, S.V. & Benjamin, W.B. Phosphorylation of recombinant human ATP:citrate lyase by cAMP-dependent protein kinase abolishes homotropic allosteric regulation of the enzyme by citrate and increases the enzyme activity. Allosteric activation of ATP:citrate lyase by phosphorylated sugars. *Biochemistry* **39**, 1169-1179 (2000).
45. Lee, J.V. *et al.* Acetyl-CoA promotes glioblastoma cell adhesion and migration through Ca(2+)-NFAT signaling. *Genes Dev* **32**, 497-511 (2018).
46. Wellen, K.E. *et al.* ATP-citrate lyase links cellular metabolism to histone acetylation. *Science* **324**, 1076-1080 (2009).
47. Richters, A. & Koehler, A.N. Epigenetic Modulation Using Small Molecules - Targeting Histone Acetyltransferases in Disease. *Curr Med Chem* **24**, 4121-4150 (2017).
48. Li, J.J. *et al.* 2-hydroxy-N-arylbenzenesulfonamides as ATP-citrate lyase inhibitors. *Bioorg Med Chem Lett* **17**, 3208-3211 (2007).
49. Bowers, E.M. *et al.* Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. *Chem Biol* **17**, 471-482 (2010).
50. Lasko, L.M. *et al.* Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. *Nature* **550**, 128-132 (2017).
51. Hornbeck, P.V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* **43**, D512-520 (2015).
52. Ding, N. *et al.* A vitamin D receptor/SMAD genomic circuit gates hepatic fibrotic response. *Cell* **153**, 601-613 (2013).
53. Massague, J., Seoane, J. & Wotton, D. Smad transcription factors. *Genes Dev* **19**, 2783-2810 (2005).
54. Raisner, R. *et al.* Enhancer Activity Requires CBP/P300 Bromodomain-Dependent Histone H3K27 Acetylation. *Cell Rep* **24**, 1722-1729 (2018).
55. Chen, P., Cescon, M. & Bonaldo, P. Collagen VI in cancer and its biological mechanisms. *Trends Mol Med* **19**, 410-417 (2013).

56. Milne, K. *et al.* A fragment-like approach to PYCR1 inhibition. *Bioorg Med Chem Lett* **29**, 2626-2631 (2019).
57. Gross, M.I. *et al.* Antitumor activity of the glutaminase inhibitor CB-839 in triple-negative breast cancer. *Mol Cancer Ther* **13**, 890-901 (2014).
58. Loayza-Puch, F. *et al.* Tumour-specific proline vulnerability uncovered by differential ribosome codon reading. *Nature* **530**, 490-494 (2016).
59. Zachar, Z. *et al.* Non-redox-active lipoate derivates disrupt cancer cell mitochondrial metabolism and are potent anticancer agents in vivo. *J Mol Med (Berl)* **89**, 1137-1148 (2011).
60. Elia, I. *et al.* Proline metabolism supports metastasis formation and could be inhibited to selectively target metastasizing cancer cells. *Nature communications* **8**, 15267 (2017).
61. Sahu, N. *et al.* Proline Starvation Induces Unresolved ER Stress and Hinders mTORC1-Dependent Tumorigenesis. *Cell Metab* **24**, 753-761 (2016).
62. Hu, M. *et al.* Regulation of in situ to invasive breast carcinoma transition. *Cancer Cell* **13**, 394-406 (2008).
63. Myllyharju, J. & Kivirikko, K.I. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet* **20**, 33-43 (2004).
64. Park, S. *et al.* Role of the Pyruvate Dehydrogenase Complex in Metabolic Remodeling: Differential Pyruvate Dehydrogenase Complex Functions in Metabolism. *Diabetes Metab J* **42**, 270-281 (2018).
65. Esteller, M. Epigenetics in cancer. *N Engl J Med* **358**, 1148-1159 (2008).
66. Albrengues, J. *et al.* Epigenetic switch drives the conversion of fibroblasts into proinvasive cancer-associated fibroblasts. *Nature communications* **6**, 10204 (2015).
67. Kumar, K. *et al.* BET inhibitors block pancreatic stellate cell collagen I production and attenuate fibrosis in vivo. *JCI Insight* **2**, e88032 (2017).
68. Mazur, P.K. *et al.* Combined inhibition of BET family proteins and histone deacetylases as a potential epigenetics-based therapy for pancreatic ductal adenocarcinoma. *Nat Med* **21**, 1163-1171 (2015).
69. Deval, C. *et al.* Amino acid limitation regulates the expression of genes involved in several specific biological processes through GCN2-dependent and GCN2-independent pathways. *FEBS J* **276**, 707-718 (2009).
70. Tanner, J.J., Fendt, S.M. & Becker, D.F. The Proline Cycle As a Potential Cancer Therapy Target. *Biochemistry* **57**, 3433-3444 (2018).
71. Hollinshead, K.E.R. *et al.* Oncogenic IDH1 Mutations Promote Enhanced Proline Synthesis through PYCR1 to Support the Maintenance of Mitochondrial Redox Homeostasis. *Cell Rep* **22**, 3107-3114 (2018).
72. Guo, L. *et al.* Kindlin-2 links mechano-environment to proline synthesis and tumor growth. *Nature communications* **10**, 845 (2019).
73. Schworer, S. *et al.* Proline biosynthesis is a vent for TGFbeta-induced mitochondrial redox stress. *EMBO J*, e103334 (2020).
74. Reversade, B. *et al.* Mutations in PYCR1 cause cutis laxa with progeroid features. *Nat Genet* **41**, 1016-1021 (2009).
75. Sahai, E. *et al.* A framework for advancing our understanding of cancer-associated fibroblasts. *Nature reviews. Cancer* **20**, 174-186 (2020).
76. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* **75**, 663-670 (2003).
77. Larsen, M.R., Thingholm, T.E., Jensen, O.N., Roepstorff, P. & Jorgensen, T.J. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* **4**, 873-886 (2005).
78. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319 (2016).

79. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731-740 (2016).
80. Wirbel, J., Cutillas, P. & Saez-Rodriguez, J. Phosphoproteomics-Based Profiling of Kinase Activities in Cancer Cells. *Methods Mol Biol* **1711**, 103-132 (2018).
81. Casado, P. *et al.* Kinase-substrate enrichment analysis provides insights into the heterogeneity of signaling pathway activation in leukemia cells. *Sci Signal* **6**, rs6 (2013).
82. Turei, D., Korcsmaros, T. & Saez-Rodriguez, J. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nat Methods* **13**, 966-967 (2016).
83. Linding, R. *et al.* NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res* **36**, D695-699 (2008).
84. Perfetto, L. *et al.* SIGNOR: a database of causal relationships between biological entities. *Nucleic Acids Res* **44**, D548-554 (2016).
85. McGregor, G.H. *et al.* Targeting the Metabolic Response to Statin-Mediated Oxidative Stress Produces a Synergistic Antitumor Response. *Cancer Res* **80**, 175-188 (2020).
86. Aper, S.J. *et al.* Colorful protein-based fluorescent probes for collagen imaging. *PLoS One* **9**, e114983 (2014).
87. Mulholland, T. *et al.* Drug screening of biopsy-derived spheroids using a self-generated microfluidic concentration gradient. *Sci Rep* **8**, 14672 (2018).
88. Davis, S. & Meltzer, P.S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* **23**, 1846-1847 (2007).
89. Ritchie, M.E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
90. Team, R.C. R: A language and environment for statistical computing. (2013).
91. Team, R. RStudio: Integrated Development for R. (2018).
92. Wickham H., A.M., Bryan J., Chang W., D'Agostino McGowan L., François R., Grolemund G., Hayes A., Henry L., Hester J., Kuhn M., Pedersen T.L., Miller E., Bache S.M., Müller K., Ooms J., Robinson D., Seidel D.P., Spinu V., Takahashi K., Vaughan K., Wilke C., Woo K., and Yutani H. Welcome to the Tidyverse. *The Journal of Open Source Software* **4** (2019).
93. Vizcaino, J.A. *et al.* The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res* **41**, D1063-1069 (2013).