Stem cell-like breast cancer cells with acquired resistance to metformin are sensitive to inhibitors of NADH-dependent CtBP dimerisation

Stem cell-like breast cancer cells with acquired resistance to metformin are sensitive to inhibitors of NADH-dependent CtBP dimerisation

Running title: CtBPs in metformin-resistant cancer stem cells

Arindam Banerjee[1], Charles N. Birts[1,4], Matthew Darley[1], Rachel Parker[1], Alex H. Mirnezami[1,3], Jonathan J. West[1,4], Ramsey I. Cutress[1,3], Stephen A. Beers[1,2], Matthew J.J. Rose-Zerilli[1,4] and Jeremy P. Blaydes[1,4]*

[1]Cancer Sciences Unit, Faculty of Medicine, University of Southampton, SO16 6YD, UK.
[2]Antibody & Vaccine Group, Cancer Sciences Unit, Faculty of Medicine, University of Southampton, SO16 6YD, UK.
[3]University Hospital Southampton, Faculty of Medicine, University of Southampton, SO16 6YD, UK.
[4]Institute for Life Sciences, University of Southampton, SO17 1BJ, UK

* To whom correspondence should be addressed. Tel: +44 2381 205151; Fax: +44 2381 205152; Email: j.p.blaydes@soton.ac.uk

Correspondence may also be addressed to Arindam Banerjee Email: A.Banerjee@soton.ac.uk

Abstract

Altered flux through major metabolic pathways is a hallmark of cancer cells, and provides opportunities for therapy. Stem cell-like cancer (SCLC) cells can cause metastasis and therapy resistance. They possess metabolic plasticity, theoretically enabling resistance to therapies targeting a specific metabolic state. The CtBP transcriptional regulators are potential therapeutic targets in highly glycolytic cancer cells, as they are activated by the glycolytic coenzyme NADH. However, SCLC cells commonly exist in an oxidative state with low rates of glycolysis. Metformin inhibits complex I of the mitochondrial electron transport chain; it can kill oxidative SCLC cells, and has anti-cancer activity in patients. SCLC cells can acquire resistance to metformin through increased glycolysis. Given the potential for long-term metformin therapy, we have studied acquired metformin resistance in cells from the claudin-low subtype of breast cancer. Cells cultured for 8 weeks in sub-IC_{50} metformin concentration proliferated comparably to untreated cells, and exhibited higher rates of glucose uptake. SCLC cells were enriched in metformin-adapted cultures. These SCLC cells acquired sensitivity to multiple methods of inhibition of CtBP function, including a cyclic peptide inhibitor of NADH-induced CtBP dimerisation. Single cell mRNA sequencing identified a reprogramming of epithelial mesenchymal and stem cell gene expression in the metformin-adapted SCLC cells. These SCLC cells demonstrated an acquired dependency on one of these genes, Tenascin C. Thus, in addition to acquisition of sensitivity to glycolysis-targeting therapeutic strategies, the reprograming of gene expression in the metformin-adapted SCLC cells renders them sensitive to potential therapeutic approaches not directly linked to cell metabolism.

Summary

Acquired resistance to metformin was used to investigate the role of OXPHOS in determining the phenotype, and therapeutic vulnerabilities, of triple negative breast cancer stem-cell like cells.

Introduction

An increasing understanding of the altered metabolism of cancer cells is providing insights into how tumours develop, and novel opportunities for therapeutic intervention (1). Tumours are heterogeneous, and key to the efficacy of any cancer therapeutic is its ability to target cell populations which would otherwise be able to regenerate new tumour. Such cells possess stem-like characteristics (2) and we refer to them as stem cell-like cancer (SCLC) cells. SCLC cells exhibit metabolic plasticity (3-5), and thus pose a particular challenge to therapeutic strategies that target a single metabolic pathway in tumour cells. The archetypal pathway for altered metabolism in cancer cells is increased glycolysis under conditions of sufficient oxygen for effective mitochondrial respiration (aerobic glycolysis, or "The Warburg Effect")(1,6), and multiple potential anti-cancer therapeutic strategies targeting glycolysis have therefore been investigated. Most such approaches target enzymes or transporters directly involved in glucose metabolism (7,8), however metabolite-dependent transcriptional regulators (1) also provide opportunities for therapeutic intervention in glycolytic tumours, as we have demonstrated (9) through the development of inhibitors of the CtBP family of NADH-dependent transcriptional regulators (10,11).

Even within tumours derived from the same tissue, such as breast cancers, SCLC cells can be either highly glycolytic or, conversely, dependent on mitochondrial oxidative phosphorylation (OXPHOS) (4,12,13). This heterogeneity may be due to cells adapting to particular environmental and / or experimental conditions (5,14). It does mean, however, that highly valuable adjuncts for the use of therapies targeting glycolytic metabolism would be strategies to restrict the SCLC cells into a state of glycolytic dependency, as well as biomarkers of cells that are in this state.

Metformin is a useful tool for the manipulation of metabolism in SCLC cells, with the potential for application in the clinic to target cancer (8,15). It inhibits complex I of the mitochondrial electron transport chain, thus limiting the ability of cells to generate ATP by OXPHOS, and increasing their dependence on glycolysis (16,17). Furthermore, in contrast to the (patho)physiological inhibition of OXPHOS by hypoxia, which increases the activity of hypoxia-inducible factors (HIFs) that promote the survival and accumulation of SCLC cells, metformin does not activate HIFs (16). The clinical interest in metformin in cancer stems from initial observations that, in patients with type 2 diabetes, those taking metformin had a reduced risk of developing cancer (18). Subsequent studies have supported this initial observation, albeit with some limitations, and others have indicated it may have a role as an adjuvant therapy (19). In breast cancer, metformin appears not to affect incidence, but has a substantial beneficial effect on mortality in diabetic patients (19,20). Mechanisms whereby metformin exerts its anti-tumour effects may involve either systemic effects on glucose, insulin and IGF-1 concentrations in the serum, or direct effects of metformin on the tumour cells (15,19). A key argument against a direct effect has been that metformin concentrations required *in vitro* to inhibit tumour cell proliferation are substantially higher than those achievable *in vivo* (15). This discrepancy has been demonstrated to be due to the supraphysiological concentrations of nutrients, particularly pyruvate, present in cell culture media (17). Furthermore, the demonstration that a cancer cell line expressing complex I components resistant to metformin is insensitive to metformin when grown as tumours *in vivo* (16) confirms that metformin can exert its effects directly on the tumour cells.

In vitro, oxidative SCLC cells are highly sensitive to metformin (21,22), however the cells can develop resistance (21,23,24). Given the potential for resistance also occurring *in vivo* in metformin-treated patients, we reasoned that the study of metformin-resistant cancer cells would be a clinically relevant experimental system in which to examine the effects of restricting SCLC cells to a less oxidative, more glycolytic, metabolic state. We focussed on triple negative breast cancer (TNBC), in which SCLC cells have been well characterised, and for which new therapeutic strategies are required (25-27). We find that resistance to metformin selects for a population of SCLC cells with markedly altered expression of epithelial mesenchymal differentiation markers. These cells exhibit heightened sensitivity to inhibitors of

glycolytic metabolism, as well as to inhibition of CtBPs. Tenascin C is also identified as a potential biomarker and therapeutic target of metformin-treated SCLC cells in TNBC.

Materials and methods

Cell culture

Cell lines (MDA-MB-231 [ATCC® HTB-26™] and MDA-MB-436 [ATCC® HTB-130™]) were originally obtained from ATCC and re-validated by STR profiling (ECACC, 08/2017) and mycoplasma testing (in house nested PCR of rRNA genes, 05/2016). Multiple stock vials were frozen at the time of mycoplasma testing, and cells were then maintained in culture for up to six months after revival. Cells were cultured in DMEM (#D5030 Sigma, U.K. with 25 mM glucose, 1 mM sodium pyruvate, 2 mM L-glutamine, 3.7 g/L sodium bicarbonate, Penicillin (100 U/mL) / Streptomycin (100 µg/mL) and 10% foetal calf serum). Metformin hydrochloride, bromopyruvic acid (BrP) and α-Keto-γ-(methylthio) butyric acid sodium salt (MTOB) were from Sigma-Aldrich, UK; 2-deoxyglucose (2-DG) was from Thermo Fisher, UK; doxycycline was from Clontech, USA., and CP61-TAT and TAT peptides (9) were synthesised by Chinapeptide Co., Ltd. (Shanghai, China). Transfection with 5 nM siRNA used INTERFERin® (Polyplus-transfection®, France), 48 hours prior to assay. siRNAs were: *CtBP1/2* siRNA (28), *SOX4* Ambion ID: s13301*, TNC* Ambion ID: s7070, or universal negative control (Ambion, Thermo Fisher Scientific, UK). To generate the [IND]CtBP[DD] vector, CtBP[DD] (CtBP2[104-361] (9)) was cloned into pENTR4-V5 (Addgene #17425) and then pLentiCMVtighthygroDEST (Addgene #26433) (29), which was co-transfected into HEK-T cells with pCMVdR8.91 and pMD2.G packaging vectors to generate lentiviral particles. The generated lentiviral particles were transduced into MDA-MB-231, which had previously been transduced with particles generated using pLentiCMVrtTA3 Blast (Addgene #26429) controller vector. Transduced cells were selected with blasticidin (10 µg/mL) and hygromycin (200 µg/mL). All photo-microscopy of cells was performed using an Olympus IX81 microscope-based system equipped with Xcellence Pro software.

Oxygen Consumption Rate (OCR), glucose uptake and ATP

OCR was measured using Seahorse Bioscience XF96 Extracellular Flux Analyzer (Seahorse Bioscience, MA). Analysis was based on manufacturer's protocols and the data were normalised to the total protein per well. Glucose Uptake Cell-Based Assay Kit (Cayman Chemical, USA) and ATPlite Luminescence Assay System (Perkin Elmer, Boston, MA) were used according to manufacturer's protocols.

Cell function assays

3D matrigel cells growth, cell invasion and colony growth were performed as previously described (14). For all assays metformin (2 mM) was present in the medium for the metformin-adapted cells.

SCLC cell assays

Mammosphere culture*:* 2000 cells were seeded in 100 µl DMEM:F12-based mammosphere media (supplemented with 20 ng/mL recombinant human EGF, 20 ng/mL recombinant human basic FGF, B27 supplement, 0.4% FCS, penicillin-streptomycin, L-glutamine (all from Life Technology, USA) and 5 µg/mL bovine insulin (Sigma)) on polyhema coated low attachment 96-well plates. After 8-10 days, representative images were captured by microscopy and mammospheres were measured by alamarBlue. Comparable conditions were also used for harvesting mammospheres for Drop-Seq analysis. ALDEFLUOR assay*:* The ALDEFLUOR assay kit was purchased from Stem Cell Technologies, UK and the assay performed as per the manufacturer's instructions. Briefly, cells were incubated with ALDEFLUOR substrate (BAAA, BODIPYaminoacetaldehyde. Stem Cell Technologies, UK) to define the ALDEFLUOR-positive cells, and a specific inhibitor of ALDH1, diethylaminobenzaldehyde (DEAB), was

used to establish the baseline fluorescence. Flow cytometry plots indicate side scatter (SSC) versus fluorescence intensity.

Protein & RNA quantification

Immunoblotting was performed as previously described (14). Primary antibodies used were αV5 (clone E10/V4RR) (Thermo Fisher, UK), R41 polyclonal αCtBP (30) and rabbit polyclonal anti-β-actin antibody (Sigma, St Louis, MO). Reverse transcription-quantitative PCR (RT-qPCR) was performed as previously described (14). Primers and probes were Roche universal probe library assays (Roche, Germany), *ABCG2*: TGGCTTAGACTCAAGCACAGC/ TCGTCCCTGCTTAGACATCC/#56; *CTNNB1*: GCTTTCAGTTGAGCTGACCA/CAAGTCCAAGATCAGCAGTCTC/#21; *DVL1*: AAGAACGTGCTCAGCAACC/AGCTTGGCATTGTCATCAAA/#63; *BCL-2*: GTACCTGAACCGGCATCTG/GGGGCCATATAGTTCCACAA/#75; *SLC7A5*: GTGGAAAAACAAGCCCAAGT/GCATGAGCTTCTGACACAGG/#25; SAA2: ACCTCTCCCTGGCAGACA/AGAAGCTTCGGCTGCTGAC /#68; *BMI1*: CCATTGAATTCTTTGACCAGAA/CTGCTGGGCATCGTAAGTATC/#63; *SAA1*: GCAGCTCAGCTACAGCACA/AGAAGCTTCGGCTGCTGAC /#68; *CD49f* /*ITGA6*: ATTCTCATGCGAGCCTTCAT/GGAAACACAGTCACTCGAACC/#74; *CXCL1*: TCCTGCATCCCCCATAGTTA/CTTCAGGAACAGCCACCAGT/#52; *ID1*: CCAGAACCGCAAGGTGAG/GGTCCCTGATGTAGTCGATGA/#39; *ID3*: CATCTCCAACGACAAAAGGAG/CTTCCGGCAGGAGAGGTT/#59; *S100A4*: CGCTTCTTCTTTCTTGGTTTG/GAGTACTTGTGGAAGGTGGACA/#81. For *CtBP1* (Assay ID: Hs000972284_m1), *CtBP2* (Assay ID: Hs00949547_g1), *TNC* (Assay ID: Hs01115665_m1), *VIM* (Assay ID:Hs00958111_m1), *CDH11* (Assay ID: Hs00901479_m1) and *SOX4* (Assay ID: Hs00268388_s1) validated Taqman® probes were used (Thermo Fisher, UK). Expression was normalised to *UBC* (Hs00824723_m1) using the ΔΔCt method. n=3 biological repeats. mRNA expression in metformin-adapted cells is shown relative to the expression in non-adapted cells. Statistical analysis (paired t-test) was performed on ΔCt values.

Whole transcriptome single-cell mRNA sequencing (Drop-seq)

Using a custom microfluidic platform https://dropletkitchen.github.io/, experiments were performed according to Macosko *et al.* (31) with the following adjustments: 13 PCR cycles and 500 pg cDNA from 1000 cells per culture condition was used as template for Nextera XT library preparation. Libraries were sequenced on the Illumina NextSeq 500 platform. Raw sequencing reads were aligned to hg19 and binned/collapsed onto the cell barcodes corresponding to individual mRNA capture beads using Drop-seq tools (31) (http://mccarrolllab.com/dropseq). To exclude low quality cells and likely cell doublets, cells with fewer than 2000 genes and greater than 15,000 UMIs were removed. All genes that were not detected in at least 3 cells were discarded, and all mitochondrial DNA-encoded genes were excluded, leaving 14698 genes across 651 cells (381 non-adapted, 270 metformin-adapted). The digital gene expression matrix was library-size normalised, scaled by the total number of transcripts, multiplied by 10,000 and natural-log transformed before further downstream analysis with Seurat (31) (http://satijalab.org/seurat/,). For sub-setting on *ITGA6* moderate/high cells (135 non-adapted, 177 metformin-adapted), non-adapted cells were randomly down-sampled. R Scripts are provide in Supplementary File S1. Gene list enrichment analysis (GLEA) was performed using ToppGene Suite, (32) and Bonferroni-corrected p values reported. The sequencing data have been deposited to the gene expression omnibus database [https://www.ncbi.nlm.nih.gov/geo/] with identifier [GSE110949].

Orthotopic tumour xenograft in nude mice

All animal work was done in accordance with a protocol approved by the University of Southampton Animal Welfare and Ethical Review Body and under Home Office licence P8969333C. Based on our prior experiments with MDA-MB-231 cells (14), 3×10^4 cells of either non-adapted or metformin-adapted MDA-MB-231 [IND]CtBP[DD], which had each been cultured in the absence or presence of 1 μg / mL doxycycline for 24 hours were injected in the mammary gland of NOD/SCID mice on both flanks in three mice per group. Doxycycline (200 μg/mL) and/or metformin (1.75 mg/mL to achieve an estimated 250 mg/kg/day) was included in the drinking water according to the treatment regimen, commencing 5 days prior to injection. Tumour growth was observed twice a week and the volume calculated using the formula (length X width^2)/2.

Data analysis and statistical methods

All experimental data for the cell assays and RT-qPCR were analysed in Microsoft Excel and GraphPad Prism. Unless stated otherwise data is presented as mean ± SEM of biological triplicates (n=3, each with $\geqslant$ 2 technical replicates), and expressed as relative to the means from non-adapted cells. Where used, t-tests are two sample, two sided, paired, unless stated otherwise. For all statistical analysis: *** $P<0.001$, ** $P<0.01$, * $P<0.05$.

Results

Metformin reduces the SCLC cell population in TNBC cell lines

We selected two human TNBC cell lines (MDA-MB-231 and MDA-MB-436) of the claudin low subtype, as this subtype have a gene expression profile resembling mammary stem cells (33). Dose-response analysis on cells cultured as adherent monolayers (2D cultures) (Figure 1A) showed IC_{50} values for 72 h metformin treatment of 33.0 to 51.8 mM (95% confidence intervals) in MDA-MB-231 and 25.1 to 30.7 mM in MDA-MB-436. However, the same treatment at only 2 mM metformin was sufficient to reduce the mammosphere-forming potential of MDA-MB-231 and MDA-MB-436 by 30.2% and 28.4% respectively (Figure 1B). The mammosphere assay quantifies SCLC cells in a population, based on their ability to avoid anoikis under non-adherent growth conditions (34). The percentage of cells positive for a SCLC cell marker, aldehyde dehydrogenase (ALDH) activity (25), was also reduced by 2 mM metformin (Figure 1C). Together, these data are consistent with the conclusions of Hirsch *et al* (22) that metformin selectively targets breast SCLC cells.

Adaptation to metformin promotes the accumulation of breast SCLC cells

To determine the consequences long-term exposure to metformin, MDA-MB-231 and MDA-MB-436 cells were maintained as 2D cultures in media containing 0 or 2 mM metformin for >8 weeks. Following this period, the elongated fibroblastoid morphology of MDA-MB-231 cells was essentially unchanged, whereas the initially more epithelioid MDA-MB-436 cells became somewhat more elongated following metformin adaptation (Figure 2A). Rates of cell proliferation in 2D cultures were comparable between non-adapted and metformin-adapted cells (Figure 2B), and intracellular ATP concentrations were not depleted in the metformin–adapted cells (Figure 2C). Consistent with the known effects of metformin on mitochondrial respiration, in both cell lines, mitochondrial oxygen consumption (OCR) was substantially reduced in the metformin-adapted cells (Figure 2D). There was also a modest, but significant, increase in glucose uptake (Figure 2E), consistent with an increase in glycolysis to compensate for the reduced mitochondrial ATP generation.

When cells were plated at low dilution for 8-10 day colony growth assays (Figure 3A) there was a non-significant trend for the metformin-adapted cells to exhibit more colony growth. In 3D matrigel-based cell

cultures metformin adaptation did result in significantly increased cell growth in both cell lines (Figure 3B) (MDA-MB-231 51.2% increase, MDA-MB-436 52.5% increase). Invasion through matrigel was also significantly increased in metformin-adapted cells (Figure 3C) (MDA-MB-231: 59.0% increase, MDA-MB-436: 66.0% increase). As these effects could potentially result from increased SCLC cell populations in the metformin-adapted cells, we measured SCLC cells using mammosphere (Figure 4A) and ALDH activity (Figure 4B) assays. Significant increases in mammosphere-forming potential were observed in the metformin-adapted variants of both cell lines (MDA-MB-231: 57.8%, MDA-MB-436: 58%). The percentage of cells positive for ALDH activity also increased. We therefore also performed RT-qPCR gene expression analysis with a panel of genes selected (14) as being relevant to the SCLC cell phenotype (Figure 4C). In both cell lines, metformin adaptation resulted in increased expression of *ITGA6,* which encodes the breast cancer stem cell marker CD49f (25,35) as well as the multidrug resistance transporter and stem cell marker *ABCG2* and *DVL1,* a WNT target gene known to be dysregulated in breast cancers. *SLC7A5* was also increased in both cell lines, indicating adaptation of metabolic pathways in response to metformin. *BMI1* and *CTNNB1* were upregulated in MDA-MB-436 only. Together these data demonstrate that long term culture in 2 mM metformin has selected for a population of metformin-resistant SCLC cells.

Metformin adapted breast SCLC cells are sensitised to glycolysis-targeting therapeutic strategies

We next set out to determine whether the metformin-adapted SCLC cells demonstrate an increased dependence on glycolytic metabolism, by treating 2D cultures of MDA-MB-231 or MDA-MB-436 with glycolysis-targeting interventions, prior to assaying for mammosphere forming potential. The hexokinase inhibitors 12.5 µM BrP (36) and 1 mM 2-DG (37) both caused a marked enrichment of SCLC cells in the non-adapted cultures, whereas in the metformin-adapted cells they significantly reduced mammosphere-forming potential (Figure 5Ai and 5Aii). These results are consistent with the SCLC cells switching from a predominantly oxidative to a more glycolytic metabolic profile upon their acquisition of resistance to metformin, with the enrichment of SCLC cells in the non-adapted cells indicating that the oxidative SCLC cells are more resistant to the glycolytic inhibitors than the non SCLC population.

Building on our previous work (9), and because the metformin adaptation is predicted to increase intracellular NADH concentrations (demonstrated in Supplementary Fig. S2A), which promotes CtBP function, we then determined the effects of inhibiting CtBP expression and function on non-adapted and metformin-adapted SCLC cells. Whilst, as we have shown previously (28) siRNA knockdown of *CTBP1* and *CTBP2* markedly inhibits total cell viability by approximately 50% after 48 h, here we find that it has a marked selectivity for SCLC cells only in the metformin-adapted cell lines (Figure 5Bi). As siRNA reduces both NADH-independent and NADH-dependent functions of CtBPs, we then used three different approaches to more selectively target the NADH-dependent functions that are increased in glycolytic cells. MTOB binds CtBPs and inhibits their ability to regulate transcription (38). 4 mM MTOB (38) had comparable effects to the glycolytic inhibitors (Figure 5Bii). CP61-TAT is a cell permeable cyclic peptide that we have developed to inhibit the NADH-induced dimerization of both CtBP1 and CtBP2 (9). 50 µM CP61-TAT (9) significantly inhibited mammosphere forming ability in the metformin-adapted cultures of both cell lines, but not in the non-adapted cells (Figure 5Biii). CtBP[DD] is the isolated dimerization domain of CtBP2, and is dominant-negative inhibitor of CtBP dimerisation (9). Here we have generated an inducible [IND]CtBP[DD] construct, stably incorporated into MDA-MB-231 cells. As shown in Figure 5Ci, the doxycyline inducing agent had no effect on mammosphere-forming potential in either non-adapted, or metformin-adapted cultures of the parental cell line, or non-adapted MDA-MB-231-[IND]CtBP[DD] cells. However, it did cause a significant, 44%, reduction in mammosphere-forming potential in the metformin-adapted MDA-MB-231-[IND]CtBP[DD] cells. Cell invasion was inhibited by [IND]CtBP[DD] in both non-adapted and Met-adapted cells, with the magnitude of the effect was greater in the Met-adapted cells (34.6% vs. 45.7% reduction) (Figure 5Cii). Results from *in vivo* tumour-forming assays reflected those of the

mammosphere experiments (Figure 5D). Together these data demonstrate that metformin-resistant SCLC cells are sensitive to inhibition of NADH-dependent CtBP dimerisation.

Single cell mRNA sequencing identifies reprogramming of gene expression in metformin-resistant SCLC cells

To obtain further insight into how SCLC cells were altered by metformin-adaptation, we analysed genome-wide mRNA expression in single-cells from MDA-MB-231 mammospheres, using Drop-seq (31). Firstly, this identified genes differentially expressed in metformin-adapted versus non-adapted cells (Supplementary Fig. S3Ai and File S2i). GLEA (Supplementary File S4) identified only weak enrichment of defined gene sets, however visual interrogation of the gene lists (Figure 6A, Supplementary Fig. S3Ai and File S2i) identified notable increases in metformin-adapted cells of specific genes including *ITGA6* ($p=4.6x10^{-21}$), which is consistent with the previously described enrichment of SCLC cells, and RT-qPCR analysis of SCLC cell markers in 2D cultures. Also upregulated were *SAA1* ($p=5.2x10^{-30}$), *SAA2* ($p=2.5x10^{-30}$), *CXCL1* ($p=12.x10^{-11}$) and *S100A4* ($p=1.5x10^{-17}$), which are pro-inflammatory genes that promote breast tumorigenesis (39,40). Strikingly, genes most significantly higher in the non-adapted cells include *VIM* ($p=8.6x10^{-42}$), *ACTB* ($p=1.7x10^{-34}$), *ID1* ($p=6.1x10^{-27}$) and *ID3* ($p=33.0x10^{-31}$), indicating that metformin adaptation has resulted in a marked change in genes that define differentiation state.

Machine learning-based clustering analysis (using 'metagenes' of correlated gene sets identified by principle component analysis) on combined data from both sets of cells demonstrated a strong gene expression signature from cell cycle phase (Supplementary Fig. S3Aii-v and File S2ii), so this signature was regressed out of the data before further analysis. This revealed that the metformin-adaptation had a dominant role in determining the transcriptional profile of the cells (Figure 6B & Supplementary Fig. S3Bi). Medium resolution analysis (res.0.75), partitioned cells into five clusters (Figure 6B and Supplementary Fig. S3Bii and File S2iii). Most notable was the appearance of a new cluster, cluster $0^{0.75}$, in the metformin-adapted cells, coincident with the loss of cells from clusters $1^{0.75}$, $2^{0.75}$ and $4^{0.75}$. Positive and negative gene expression markers of cluster $0^{0.75}$ reflect the overall gene expression changes in metformin adapted cells (e.g. increased *SAA1*, *SAA2*, *CXCL1* and *ITGA6*, decreased *ID1*, *ID3* and *VIM*), but also include other positive markers including *DEFB1*, *SOX4*, *CDH11* and *CPVL* (Supplementary Fig. S3Bii, S3Biii and File S2iii). Cells in clusters $2^{0.75}$ and $4^{0.75}$ both have relatively high *VIM* expression, and cluster $2^{0.75}$ cells also have high *ID1* expression. Lower resolution analysis (res.0.5) essentially resulted in clusters $0^{0.75}$ and $1^{0.75}$ being merged into a single cluster, suggesting that these two clusters contain cells with a related transcriptome, modified by the adaptation to metformin (Supplementary Fig. S3Biv and File S2iv). Direct comparison of cluster $0^{0.75}$ (almost exclusively metformin-adapted) with cluster $1^{0.75}$ (mostly non-adapted) showed increases in *CPVL*, *DEFB1*, *SOX4*, *CDH11*, *CXCL1* and *ITGA6* in cluster $0^{0.75}$, and a substantial drop in *ID3* and *ACTB* (Supplementary File S3). Together, these analyses indicate that the non-adapted cells contain multiple clusters of cells with distinct transcriptional profiles, and that the metformin-adapted cells are largely restricted to, a modified form of one of these clusters.

As not all the cells within the mammospheres are SCLC cells, we filtered the original dataset to only include *ITGA6*^med/high cells, and repeated the analysis. Differential gene expression between non-adapted and metformin-adapted *ITGA6*^med/high cells is shown in Supplementary Fig. S3Ci and File S2v). Clustering analysis demonstrated that metformin-adapted *ITGA6*^med/high cells had a clearly distinct gene expression profile to their non-adapted counterparts (Figure 6C and Supplementary Fig. S3Cii). GLEA of the altered genes (Supplementary File S5) again only identified relatively weak gene set enrichment, however one gene set not observed in analysis of the bulk cell population was notable for its increase in metformin-adapted cells "molecular function- NADH dehydrogenase (quinone) activity" ($p=4.1x10^{-3}$). These six genes (Figure 6D) are all components of complex I of the mitochondrial electron transport chain, suggesting a response in *ITGA6*^med/high cells to develop metformin resistance though the upregulation of

its molecular target. Medium resolution clustering analysis (res.0.5) resolved the cells into four clusters (Figure 6E and Supplementary Fig. S3Ciii and File S2vi), two of which, clusters $1^{0.5}$ and $3^{0.5}$, contained both non-adapted and metformin-adapted cells, whereas cluster $0^{0.5}$ consisted primarily of the majority of metformin-adapted cells and cluster $2^{0.5}$ was almost exclusively from the non-adapted populations. Positive gene expression markers (Figure 6F and Supplementary Fig. S3Ciii and File S2vi) of cluster $0^{0.5}$ included *CDH11*, *DEFB1*, *SOX4* and *TNC*, and, for cluster $2^{0.5}$, *ID1* and *VIM*. Clustering at lower resolution (res.0.25) indicates that cells in clusters $0^{0.5}$, $1^{0.5}$ and $3^{0.5}$ co-cluster whereas cluster $2^{0.5}$ remains distinct (Supplementary Fig. S3Civ and File S2vii). Together these data demonstrate that, in non-adapted cell cultures, distinct subgroups of *ITGA6*med/high positive cells exist. Adaptation to metformin strongly selects against one of these subgroups (cluster $2^{0.5}$) and results in the marked expansion of a previously rare subgroup of cells (Cluster $0^{0.5}$) that overexpress genes known to be important for breast cancer progression.

To determine whether the changes in gene expression identified by the Drop-seq analysis of MDA-MB-231 mammospheres could also be observed in a second TNBC cell line, we performed qRT-PCR analysis of mRNA from 2D cultures of MDA-MB-231 and MDA-MB-436 cells (Figure 6G). The majority of genes examined that were identified as differentially regulated in the mammospheres were similarly affected in 2D-cultured MDA-MB-231 cells. The exception to this was *VIM*, which was not downregulated by metformin adaptation. Based on the low ΔCt values of the *VIM* qRT-PCRs (not shown) *VIM* may be expressed on the 2 mM metformin-insensitive bulk cell population, alternatively the *VIM* in the SCLC cells may be specifically upregulated only when cells are cultured as mammospheres, as had been shown previously (41). Of the nine genes examined that did change in MDA-MB-231, only *SAA1, SOX4* and *TNC*, showed a comparable response in MDA-MB-436. Of the remaining genes, based on ΔCt values, *ID1* and *ID3* were already low in non-adapted MDA-MB-436, *S100A4* was already high, whereas *CHD11* and *SAA2* were undetectable. To begin to investigate the potential application of some of the Drop-seq identified transcripts in the biological response of SCLC cells to metformin-adaptation, we examined changes in their expression in response to induction of INDCtBPDD in cells cultured in 2D (Figure 6H). *SAA1,* which was strongly induced by metformin-adaptation in the analysis of mammosphere mRNA, was unaffected by INDCtBPDD in non-adapted cells, but was significantly downregulated by CtBP inhibition in the metformin-adapted cells. This is consistent with *SAA1* being either a marker of the metformin-adapted SCLC cells that are sensitive to INDCtBPDD, or a CtBP-regulated gene that is required for SCLC cell survival.

Finally, we performed siRNA analysis on two of the genes, *SOX4* and *TNC,* identified as upregulated in the metformin-adapted *ITGA6*med/high cells, as both have recently been reported as key drivers of SCLC phenotypes in breast cancer (42,43). *SOX4* was required for optimal mammosphere forming ability in both non-adapted and metformin-adapted cultures on MDA-MB-231 and MDA-MB-436 (Supplementary Fig. S4). In contrast *TNC* knockdown had no effect on mammosphere-forming ability in non-adapted cells, but caused a significant reduction in metformin-adapted MDA-MB-231 by 24%, and by 27% in MDA-MB-436 (Figure 6I). Furthermore, in both cell lines, the negative effect of *TNC* knockdown on invasion was markedly more pronounced in metformin-adapted, compared to non-adapted, cells (Figure 6J). *TNC* siRNA also appeared to partly reverse the effects of metformin adaptation on the morphology of MDA-MB-231 cells (Supplementary Fig. S5B).

Discussion

The literature contains limited examples of acquired resistance to metformin in cancer cell models. Sancho *et al*, (21), showed effective targeting of pancreatic SCLC cells by metformin, and effective treatment of primary cancer-derived xenograft model. However, all the tested tumours eventually

regressed due to the emergence of SCLC cells with an intermediate glycolytic/respiratory phenotype. This metabolic adaptation was dependent on c-MYC, and metformin resistance could be reversed by c-MYC inhibition or knockdown. In breast models, recognising the potential for acquired resistance in patients treated with metformin, Oliveras-Ferraros *et al* (23) and subsequently Scherbakov *et al* (24) showed substantial reprogramming of gene-expression networks following metformin adaptation in the oestrogen receptor positive MCF-7 model.

We have focussed on the triple negative subtype of breast cancer due to the requirement for novel therapeutic strategies for this disease and because, whilst *in vitro* and *in vivo* models have demonstrated single agent treatment with metformin to effectively target TNBC cells and SCLC cells (44,45), current evidence suggests that it has limited benefit for TNBC patients (46). We have firstly confirmed these *in vitro* effects of metformin in TNBC, and then demonstrated that cells adapted to culture in 2 mM metformin proliferate in 2D culture at normal rates, with a less oxidative, more glycolytic metabolism. Notably, as demonstrated by both molecular and phenotypic assays, these cultures become enriched in CSLC cells. These metformin-resistant CSLC cells are metabolically distinct from the non-adapted ones, as demonstrated by both their insensitivity to metformin and their acquisition of sensitivity to two different small molecule inhibitors of glycolytic metabolism. Thus, following adaptation to metformin, the metabolism of SCLC cells may become more similar to that of the more glycolytic non-stem-like population, which has previously been shown to be vulnerable to combined metformin and glycolytic inhibitors (47) and also confirms that SCLC cells can exist in either OXPHOS or glycolysis-dependent states (14).

Single cell mRNA analysis was then used to gain transcriptome-wide gene expression data in CSLC cell-enriched populations without and with adaptation to metformin. We focussed on the *ITGA6*[med/high] population and found these cells were not only enriched in the metformin-adapted cells, but also substantially altered in terms of their gene expression profile. Firstly, the increased expression of a panel of genes encoding components of complex I of the electron transport chain, accompanied by increases in other genes involved in ATP metabolism, is most likely part of the mechanism of acquired resistance to metformin. That this response was not observed when the entire population of cells in the mammospheres were analysed indicates that *ITGA6*[med/high] population may have an increased capacity to adapt their metabolism in response to stress, which would account for their enrichment in metformin-adapted cultures. The lack of any evidence at the transcriptional level for an increase in glycolysis is consistent both with the lack of HIF-activation by metformin (16), and that the cells are already highly glycolytic (6). The increased glucose uptake in the metformin-adapted cells may therefore be due to biochemical regulation of glycolysis.

The most striking observations from the RNA-seq analysis, however, was the marked difference in cell populations with distinct profiles of expression of epithelial mesenchymal and stem cell markers between the non-adapted and metformin-resistant SCLC cells. Breast SCLC cells may exist in distinct states with either epithelial or mesenchymal phenotypes (25,33) and in the SUM149 TNBC model, in which both states are present (33), culture as mammospheres enriches for more mesenchymal-like cells with higher expression of vimentin (*VIM*) (41). Similarly, we found within the non-adapted *ITGA6*[med/high] cells in MDA-MB-231 mammospheres, a large cluster of cells for which the most significant positive expression marker was *VIM*. These cells were almost completely absent in the metformin-adapted mammospheres, a finding consistent with an earlier report that suppression of SCLC cells by metformin is accompanied by downregulation of epithelial-mesenchymal transition markers (48). Instead, in the metformin-resistant SCLC cells the largest cluster of cells was marked by the cadherin *CDH11*. *CDH11* is known to promote invasion and metastasis in this cell model. Its elevated expression correlates with poor survival in breast cancer (49), including the claudin-low subtype (Supplementary Fig. S6). It is a suppressor of vimentin

expression and the mesenchymal phenotype in some tumour types (50). As *CDH11* is critical in lineage commitment in mesenchymal stem cells (51), it may also be involved in defining the characteristic of a breast SCLC cell. *CDH11* is also a regulator of extracellular matrix (ECM) synthesis (52), and *TNC,* which encodes the ECM protein Tenascin C, was also amongst the top markers of the cells in the metformin-adapted cell-specific cluster. Genes upregulated in this cluster of cells, and the proteins they encode, may therefore serve as potential biomarkers for cells in which environmental, or therapeutic, stress has driven breast SCLC cells from an oxidative to a more glycolytic state, though the RT-qPCR data from MDA-MB-436 indicates that *CDH11* induction at least may not be ubiquitous to all claudin-low tumours. Furthermore, autocrine secretion of *TNC* can be critical for metastasis in MDA-MB-231 models, and also supports the formation of mammospheres under conditions of cellular stress (42). Our finding that MDA-MB-231 and MDA-MB-436 mammosphere formation was selectively dependent on *TNC* only in the metformin-adapted cells is consistent with this, and confirms that the altered gene expression profile in the metformin-adapted *ITGA6*[med/high] cells does serve to promote their SCLC cell characteristics, but at the same time introduces novel vulnerabilities to the cells which could serve as therapeutic targets in metformin-treated breast cancer patients.

Our initial aim of this work was to generate a clinically relevant model system in which oxidative breast SCLC cells are forced out of their normal metabolic state into a more glycolytic state, with a specific view to determining the potential of targeting glycolytic metabolism through the inhibition of NADH-dependent dimerization of the CtBP family of transcriptional regulators. Due to their pleiotropic effects on multiple hallmarks of cancer, and that their activity is induced by one of the most ubiquitous features of cancer cells, increased aerobic glycolysis, we and others have identified CtBPs as a potential therapeutic target in cancer (10,11,38), and have begun to develop small molecule inhibitors of CtBP function (9). However, the increasing understanding that SCLC cells exhibit a degree of metabolic plasticity that could enable them to escape such a therapy through the maintenance of a more oxidative, less glycolytic state, invokes the requirement for strategies to overcome this potential escape mechanism. Treatment with biguanides, such as metformin provides an attractive mechanism for doing this, particularly for CtBPs as metformin specifically functions by inhibiting the mitochondrial regeneration of NAD^+ to NADH by the electron transport chain (17), without markedly promoting the cytoplasmic, lactate dehydrogenase A-driven reduction of NADH by the activation of HIFs (16). NADH concentrations thus increase in metformin-treated cells (17). NADH binding to CtBPs promotes their dimerization and consequent functions as transcriptional regulators which includes suppression of genes that promote apoptosis and anoikis (53), resistance to anoikis being a key feature of SCLC cells. Notably, from our single cell RNA-seq analysis, three of the nine such genes that were previously identified from studies of *Ctbp*[-/-]*/Ctbp*[-/-] -MEFs (53) appear as significant positive expression markers of the non-adapted *ITGA6*[med/high] MDA-MB-231 mammosphere cells compared to the Met-adapted ones (*PERP, ID-1* and *IGFBP-7*, Supplementary File S2v), consistent with their having been repressed by CtBPs in the metformin-adapted SCLC cells. Here we have used four mechanistically distinct strategies to inhibit CtBP expression and function, including a cyclic peptide inhibitor of NADH-dependent CtBP dimerization, to demonstrate that chronic exposure to low concentrations of metformin renders breast SCLC cells dependent upon the function of the CtBP family of glycolysis-sensing transcriptional regulators for their survival.

Supplementary material

Supplementary Figures S1 to S6 and Supplementary Files S1 to S5 can be found at http://carcin.oxfordjournals.org/

Funding

References

1.   Pavlova, N.N.*, et al.* (2016) The Emerging Hallmarks of Cancer Metabolism. *Cell Metab*, **23**, 27-47.

2.   Magee, J.A.*, et al.* (2012) Cancer stem cells: impact, heterogeneity, and uncertainty. *Cancer Cell*, **21**, 283-96.

3.   Cuyas, E.*, et al.* (2014) The nutritional phenome of EMT-induced cancer stem-like cells. *Oncotarget*, **5**, 3970-82.

4.   Luo, M.*, et al.* (2015) Metabolic plasticity of cancer stem cells. *Oncotarget*, **6**, 35141-2.

5.   Sancho, P.*, et al.* (2016) Hallmarks of cancer stem cell metabolism. *Br J Cancer*, **114**, 1305-12.

6.   Gatenby, R.A.*, et al.* (2004) Why do cancers have high aerobic glycolysis? *Nat Rev Cancer*, **4**, 891-9.

7.   Yu, L.*, et al.* (2016) The sweet trap in tumors: aerobic glycolysis and potential targets for therapy. *Oncotarget*, **7**, 38908-38926.

8.   Luengo, A.*, et al.* (2017) Targeting Metabolism for Cancer Therapy. *Cell Chem Biol*, **24**, 1161-1180.

9.   Birts, C.N.*, et al.* (2013) A cyclic peptide inhibitor of C-terminal binding protein dimerization links metabolism with mitotic fidelity in breast cancer cells. *Chemical Science*, **4**, 3046-3057.

10.   Bergman, L.M.*, et al.* (2006) C-terminal binding proteins: Emerging roles in cell survival and tumorigenesis. *Apoptosis*, **11**, 879-888.

11.   Dcona, M.M.*, et al.* (2017) CtBP- an emerging oncogene and novel small molecule drug target: advances in the understanding of its oncogenic action and identification of therapeutic inhibitors. *Cancer Biol Ther*, 0.

12.   Peiris-Pages, M.*, et al.* (2016) Cancer stem cell metabolism. *Breast Cancer Res*, **18**, 55.

13.   Vlashi, E.*, et al.* (2015) The metabolic state of cancer stem cells-a valid target for cancer therapy? *Free Radic Biol Med*, **79**, 264-8.

14.   Banerjee, A.*, et al.* (2018) The effects of restricted glycolysis on stem-cell like characteristics of breast cancer cells. *Oncotarget*, **9**, 23274-23288.

15.   Pernicova, I.*, et al.* (2014) Metformin--mode of action and clinical implications for diabetes and cancer. *Nat Rev Endocrinol*, **10**, 143-56.

16. Wheaton, W.W., *et al.* (2014) Metformin inhibits mitochondrial complex I of cancer cells to reduce tumorigenesis. *Elife*, **3**, e02242.

17. Gui, D.Y., *et al.* (2016) Environment Dictates Dependence on Mitochondrial Complex I for NAD+ and Aspartate Production and Determines Cancer Cell Sensitivity to Metformin. *Cell Metab*, **24**, 716-727.

18. Evans, J.M., *et al.* (2005) Metformin and reduced risk of cancer in diabetic patients. *BMJ*, **330**, 1304-5.

19. Morales, D.R., *et al.* (2015) Metformin in cancer treatment and prevention. *Annu Rev Med*, **66**, 17-29.

20. Tang, G.H., *et al.* (2018) Association of metformin with breast cancer incidence and mortality in patients with type 2 diabetes: a GRADE assessed systematic review and meta-analysis. *Cancer Epidemiol Biomarkers Prev*.

21. Sancho, P., *et al.* (2015) MYC/PGC-1alpha Balance Determines the Metabolic Phenotype and Plasticity of Pancreatic Cancer Stem Cells. *Cell Metab*, **22**, 590-605.

22. Hirsch, H.A., *et al.* (2009) Metformin selectively targets cancer stem cells, and acts together with chemotherapy to block tumor growth and prolong remission. *Cancer Res*, **69**, 7507-11.

23. Oliveras-Ferraros, C., *et al.* (2014) Acquired resistance to metformin in breast cancer cells triggers transcriptome reprogramming toward a degradome-related metastatic stem-like profile. *Cell Cycle*, **13**, 1132-44.

24. Scherbakov, A.M., *et al.* (2016) The phenomenon of acquired resistance to metformin in breast cancer cells: The interaction of growth pathways and estrogen receptor signaling. *IUBMB Life*, **68**, 281-92.

25. Liu, S., *et al.* (2014) Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Reports*, **2**, 78-91.

26. Lim, E., *et al.* (2009) Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med*, **15**, 907-13.

27. Prat, A., *et al.* (2013) Characterization of cell lines derived from breast cancers and normal mammary tissues for the study of the intrinsic molecular subtypes. *Breast Cancer Res Treat*, **142**, 237-55.

28. Bergman, L.M., *et al.* (2009) CtBPs promote cell survival through the maintenance of mitotic fidelity. *Mol Cell Biol*, **29**, 4539-4551.

29. Campeau, E., *et al.* (2009) A versatile viral system for expression and depletion of proteins in mammalian cells. *PLoS One*, **4**, e6529.

30. Mirnezami, A.H., *et al.* (2003) Hdm2 recruits a hypoxia sensitive co-repressor to negatively regulate p53-dependent transcription. *Curr Biol*, **13**, 1234-1239.

31. Macosko, E.Z., *et al.* (2015) Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, **161**, 1202-1214.

32. Chen, J., *et al.* (2009) ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res*, **37**, W305-11.

33.   Prat, A., *et al.* (2010) Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res*, **12**, R68.

34.   Dontu, G., *et al.* (2003) In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev*, **17**, 1253-70.

35.   Visvader, J.E., *et al.* (2014) Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev*, **28**, 1143-58.

36.   Wu, L., *et al.* (2014) The reversal effects of 3-bromopyruvate on multidrug resistance in vitro and in vivo derived from human breast MCF-7/ADR cells. *PLoS One*, **9**, e112132.

37.   Ben Sahra, I., *et al.* (2010) Targeting cancer cell metabolism: the combination of metformin and 2-deoxyglucose induces p53-dependent apoptosis in prostate cancer cells. *Cancer Res*, **70**, 2465-75.

38.   Straza, M.W., *et al.* (2010) Therapeutic targeting of C-terminal binding protein in human cancer. *Cell Cycle*, **9**.

39.   Hansen, M.T., *et al.* (2015) A link between inflammation and metastasis: serum amyloid A1 and A3 induce metastasis, and are targets of metastasis-inducing S100A4. *Oncogene*, **34**, 424-35.

40.   Hartman, Z.C., *et al.* (2013) Growth of triple-negative breast cancer cells relies upon coordinate autocrine expression of the proinflammatory cytokines IL-6 and IL-8. *Cancer Res*, **73**, 3470-80.

41.   Wolf, J., *et al.* (2013) A mammosphere formation RNAi screen reveals that ATG4A promotes a breast cancer stem-like phenotype. *Breast Cancer Res*, **15**, R109.

42.   Oskarsson, T., *et al.* (2011) Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nat Med*, **17**, 867-74.

43.   Tiwari, N., *et al.* (2013) Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming. *Cancer Cell*, **23**, 768-83.

44.   Liu, B., *et al.* (2009) Metformin induces unique biological and molecular responses in triple negative breast cancer cells. *Cell Cycle*, **8**, 2031-40.

45.   Shi, P., *et al.* (2017) Metformin suppresses triple-negative breast cancer stem cells by targeting KLF5 for degradation. *Cell Discov*, **3**, 17010.

46.   Bayraktar, S., *et al.* (2012) Effect of metformin on survival outcomes in diabetic patients with triple receptor-negative breast cancer. *Cancer*, **118**, 1202-11.

47.   Rajh, M., *et al.* (2016) Medium Renewal Blocks Anti-Proliferative Effects of Metformin in Cultured MDA-MB-231 Breast Cancer Cells. *PLoS One*, **11**, e0154747.

48.   Vazquez-Martin, A., *et al.* (2010) Metformin regulates breast cancer stem cell ontogeny by transcriptional regulation of the epithelial-mesenchymal transition (EMT) status. *Cell Cycle*, **9**, 3807-14.

49.   Li, Y., *et al.* (2014) HOXC8 promotes breast tumorigenesis by transcriptionally facilitating cadherin-11 expression. *Oncotarget*, **5**, 2596-607.

50.   Li, L., *et al.* (2012) The human cadherin 11 is a pro-apoptotic tumor suppressor modulating cell stemness through Wnt/beta-catenin signaling and silenced in common carcinomas. *Oncogene*, **31**, 3901-12.

51.   Alimperti, S., *et al.* (2015) CDH2 and CDH11 act as regulators of stem cell fate decisions. *Stem Cell Res*, **14**, 270-82.

52.   Row, S., *et al.* (2016) Cadherin-11 is a novel regulator of extracellular matrix synthesis and tissue mechanics. *J Cell Sci*, **129**, 2950-61.

53.   Grooteclaes, M., *et al.* (2003) C-terminal-binding protein corepresses epithelial and proapoptotic gene expression programs. *Proc Natl Acad Sci U S A*, **100**, 4568-4573.

FIGURE LEGENDS

Figure 1. Short-term 2 mM metformin treatment inhibits triple negative breast SCLC cells

(A) MDA-MB-231 and MDA-MB-436 cells were cultured in 2D adherent culture with 0-80 mM metformin for 72 hours. Viable cells were then measured using alamarBlue. (B,C) MDA-MB-231 and MDA-MB-436 cells in 2D adherent culture were treated without/with metformin (2 mM) for 72 h prior to assay. (B) Mammosphere assay. 8-10 days after plating treated cells in non-adherent mammosphere culture conditions (bar 50 μm MDA-MB-231, 100 μm MDA-MB-436) (C) ALDEFLUOR assay for ALDH activity in treated cells from 2D adherent cultures; DEAB, was used to establish the baseline fluorescence. Flow cytometry plots indicate side scatter (SSC) versus fluorescence intensity (full dataset from three biological repeats is shown in Supplementary Fig. S1).

Figure 2. Effect of adaptation to continuous exposure to 2 mM metformin on cellular bio-energetic balance

(A-E) Matched pairs of non-adapted and 2 mM metformin-adapted cells were assayed in 2D adherent culture conditions. (A) Cell morphology (bar 100 μm). (B) 72 h monolayer proliferation assay. (C) Cellular ATP abundance assay. (D) Mitochondrial oxygen consumption rate (OCR) was analysed by a Seahorse Bioscience XF96 Extracellular Flux Analyzer. (n=5 from a representative of 3 independent experiments). (E) 2-NBDG glucose uptake assay.

Figure 3. Phenotypic consequence of continuous exposure to non-cytotoxic concentrations of metformin

(A-C) Matched pairs of non-adapted, or 2 mM metformin-adapted MDA-MB-231 and MDA-MB-436 cells were assayed. Assay media contained 0 mM or 2 mM metformin for non-adapted and metformin-adapted cells respectively (A) 8-10 day colony growth assay (B) 7-8 day 3D matrigel cell growth assay (bar 100 μm) and (C) 24 h cell invasion assay (bar 100 μm).

Figure 4. Continuous exposure to non-cytotoxic concentrations of metformin promotes SCLC-like population in TNBC cells

(A-C) Matched pairs of non-adapted, or 2 mM metformin-adapted MDA-MB-231 and MDA-MB-436 cells in 2D adherent cultures were assayed (A) Mammosphere assay. 8-10 days after plating cells in non-adherent mammosphere culture conditions. (bar 100 µm). (B) ALDEFLUOR assay; DEAB, was used to establish the baseline fluorescence. Flow cytometry plots indicate side scatter (SSC) versus fluorescence intensity (full dataset from three biological repeats is shown in Supplementary Fig. S1). (C) Expression of genes relevant to SCLC cell characteristics. mRNA expression in metformin-adapted cells is shown relative to the expression in non-adapted cells. (black-higher in non-adapted. grey-higher in metformin-adapted. ≥ 25% change, n=3 biological repeats).

Figure 5. Continuous exposure to non-cytotoxic concentrations of metformin selectively sensitises triple negative breast SCLC cells towards inhibitors of glycolytic metabolism or glycolysis-dependent transcriptional regulators

(A,B) Matched pairs of non-adapted, or 2 mM metformin-adapted MDA-MB-231 and MDA-MB-436 cells in 2D adherent cultures were treated as indicated, and then assayed for mammosphere forming capacity (A) (i) 3-bromopyruvate (BrP;12.5 µM) for 24 hours. (ii) 2-deoxyglucose (2-DG; 1 mM) for 24 hours. (B) (i) Transfection with the indicated siRNA 48 h prior to plating for mammosphere assay. Validation of the efficacy of the siRNA is shown in Supplementary Fig. S2B,C (ii) 4-Methylthio 2-oxobutyric acid (MTOB; 4 mM) for 24 hours. (iii) CtBP peptide inhibitor (CP61-TAT; 50 µM) of carrier peptide control (TAT, 50 µM). (C) Matched pairs of non-adapted, or 2 mM metformin-adapted MDA-MB-231 and MDA-MB-231-[IND]CtBP[DD] cells in 2D adherent cultures were cultured in the absence of presence of doxycycline (1 µg/mL) for 24 h, and then assayed for (i) mammosphere forming capacity and (ii) invasion ability. Western blot demonstrating induction of the CtBP[DD] protein is shown in Supplementary Fig. S2D, morphology of the cells in 2D adherent culture in S2E, and images of the invasion assays in S2F. (D) Effect of [IND]CtBP[DD] on tumour-forming ability in MDA-MB-231 orthotopic xenografts. Results are shown at 4 weeks post injection, which was the earliest time point at which tumours were observed at all injection sites for the metformin-adapted (no doxycycline) cells. By 8 weeks post injection, all animals had formed at least one tumour. (A-C) one way ANOVA and Fisher's LSD test.

Figure 6. Single cell gene expression profiling of metformin-resistant SCLC cells

MDA-MB-231 cultured as mammospheres were analysed by Drop-seq. (A) Violin plots of showing the effects of adaptation to 2 mM metformin of the expression of selected differentially expressed genes in the whole cell population. See also Supplementary Fig. S3A and File S2i,ii. (B) t-SNE analysis of the whole cell population after regression of cell cycle signatures, with cell clusters identified at resolution 0.75. Insert is a bar chart showing the number of cells from each culture condition in the clusters. See also Supplementary Fig. S3A and File S2iii,iv for further data. (C-F) Analyses were performed on *ITGA6*[med/high] cells only (expression >1 on the scale in Fig. 6A). See also Supplementary Fig. S3C and File S2v-vii for further data. (C) t-SNE analysis showing the original cell identity. (D) Violin plots of showing the effects of adaptation to 2 mM metformin of the expression of components of complex I of the mitochondrial electron transport chain (E) t-SNE analysis showing the cell clusters at resolution 0.5. Insert is a bar chart showing the number of cells from each culture condition in the clusters. (F) Violin plots showing selected gene markers of clusters 0 and 2. (G) Expression of genes identified by single cell mRNA seq in MDA-MB-231 and MDA-MB-436 cultured in 2D. mRNA expression in metformin-adapted cells is shown relative to the expression in non-adapted cells (blue-higher in non-adapted, orange-higher

in metformin-adapted. ≥ 25% change, n=3 biological repeats). (H) RT-qPCR was performed on RNA extracted from MDA-MB-231-^{IND}CtBP^{DD} cells cultured in 2D in the absence of presence of doxycycline for 24 hours. ((yellow-increased by ^{IND}CtBP^{DD} , blue-decreased by ^{IND}CtBP^{DD}. ≥ 25% change, n=3 biological repeats). The *CtBP2* qPCR confirms induction of the ^{IND}CtBP^{DD} mRNA. (I,J) Cells were transfected with TNC siRNA 48 h prior to plating for (I) mammosphere assay and (J) migration assay. Validation of the efficacy of the siRNA, and images showing its effects on migration as well as morphology in 2D, are shown in Supplementary Fig. S5.
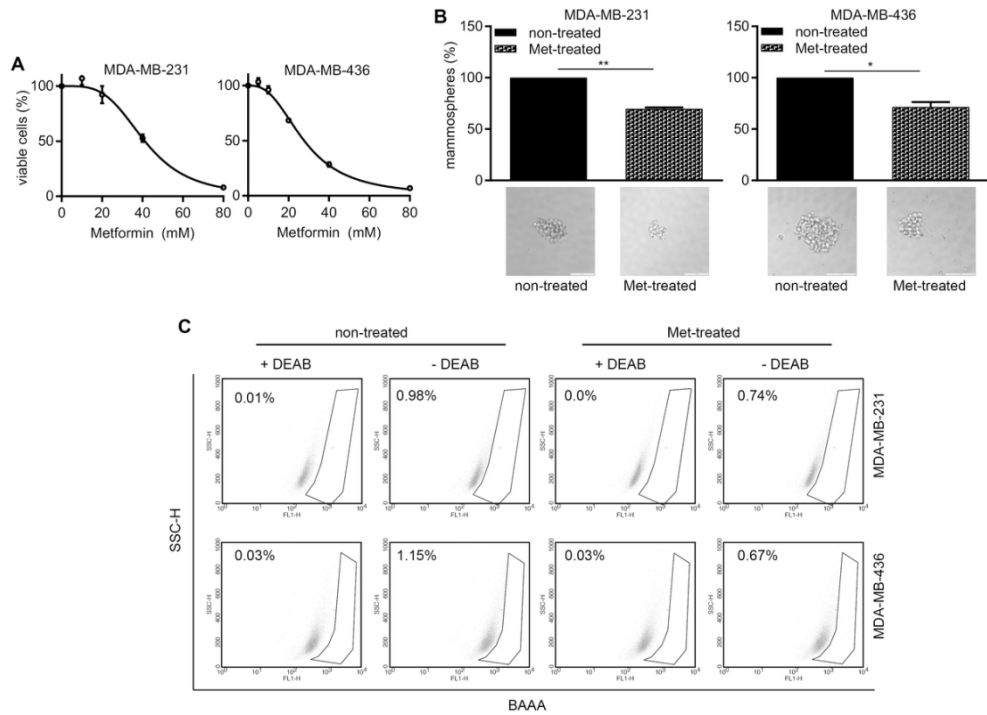
Figure 1

107x77mm (300 x 300 DPI)
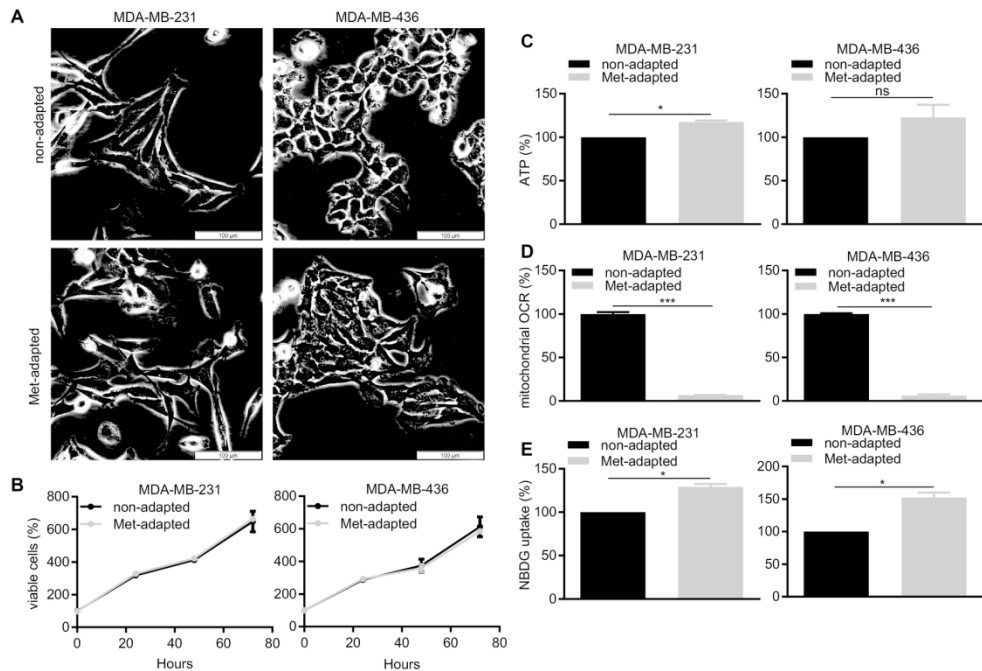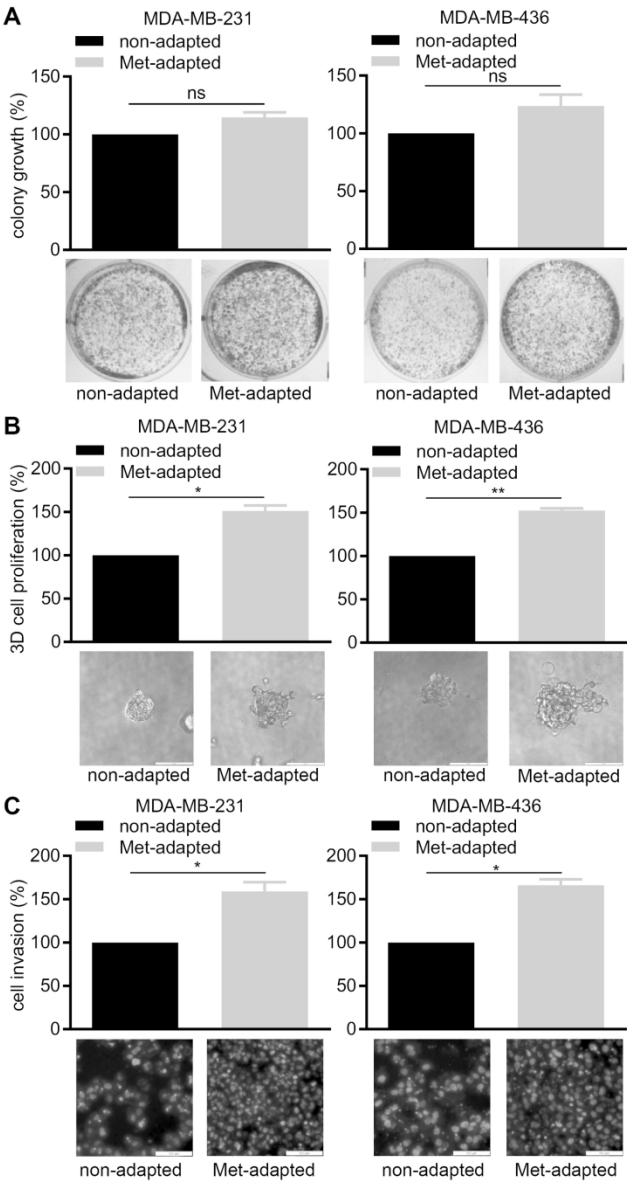
Figure 2

153x104mm (300 x 300 DPI)
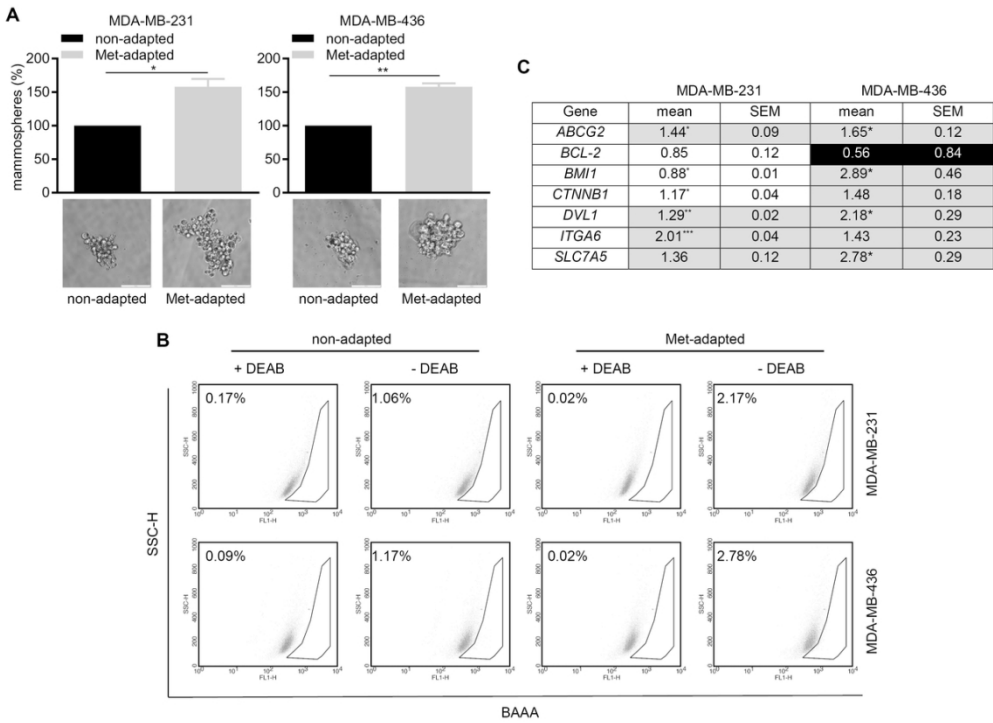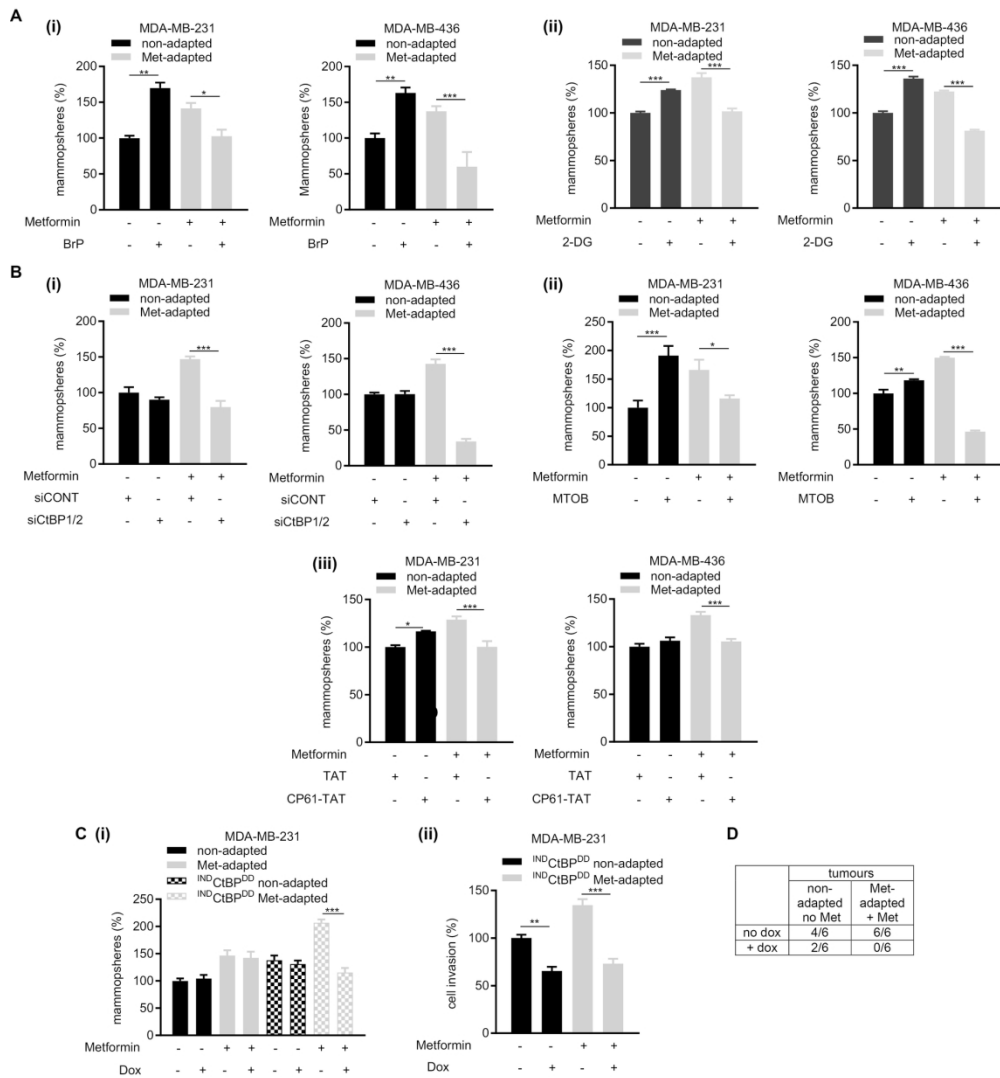
Figure 3

138x254mm (300 x 300 DPI)

Figure 4

108x79mm (300 x 300 DPI)
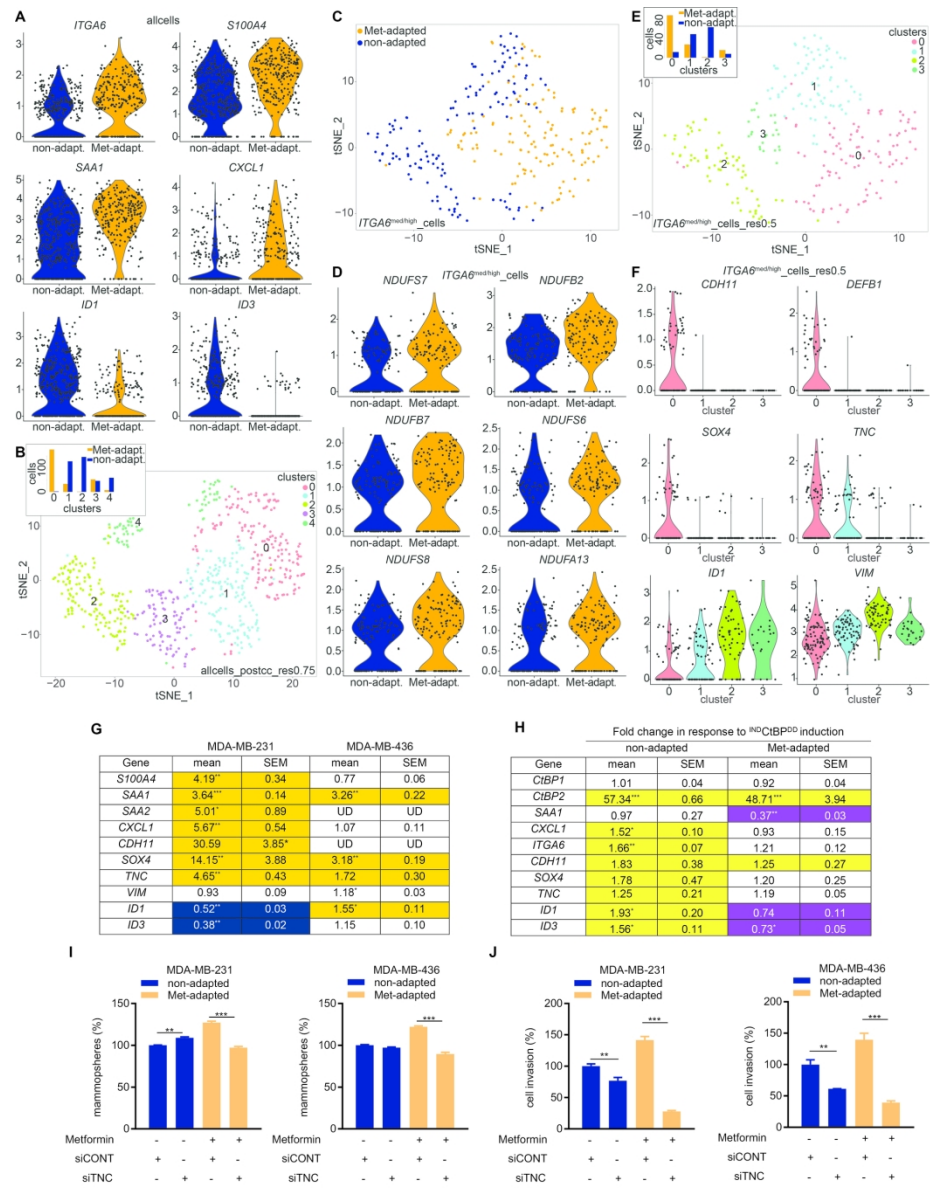
Figure 5

168x181mm (300 x 300 DPI)

Figure 6

175x229mm (300 x 300 DPI)

A

Cell line	Repeat	Treatment	ALDH +	ALDH -	% +ve	P
231	1	non-treated	106	9126	1.15	
231	1	Met-treated	55	9189	0.6	<0.0001
231	2	non-treated	71	7217	0.98	
231	2	Met-treated	53	7177	0.74	0.1253
231	3	non-treated	90	7477	1.19	
231	3	Met-treated	22	4518	0.49	<0.0001
					0	
436	1	non-treated	104	8992	1.15	
436	1	Met-treated	58	8677	0.67	0.0009
436	2	non-treated	104	9109	1.13	
436	2	Met-treated	12	9177	0.14	<0.0001
436	3	non-treated	82	8815	0.93	
436	3	Met-treated	55	7938	0.69	0.1024

B

| Cell line	Repeat	Treatment	ALDH +	ALDH -	% +ve	P
231	1	non-adapted	99	9261	1.06	
231	1	Met-adapted	204	9232	2.17	<0.0001
231	2	non-adapt	83	9619	0.86	
231	2	Met-adapted	360	9306	3.73	<0.0001
231	3	non-adapted	105	9352	1.12	
231	3	Met-adapted	977	8749	10.05	<0.0001
					0	
436	1	non-adapted	109	9230	1.17	
436	1	Met-adapted	260	9105	2.78	<0.0001
436	2	non-adapted	91	7824	1.16	
436	2	Met-adapted	214	8004	2.61	<0.0001
436	3	non-adapted	112	8426	1.32	
436	3	Met-adapted	490	8063	5.74	<0.0001

Supplementary Figure S1: Experimental data for Aldefluor assays. Flow cytometry counts for the minus DEAB samples from the three independent experiments in each of (A) Fig 1C and (B) Fig 4B are shown. The effects of metformin treatment or adaptation are tested by a two-sided Fisher's exact test (GraphPad Prism).
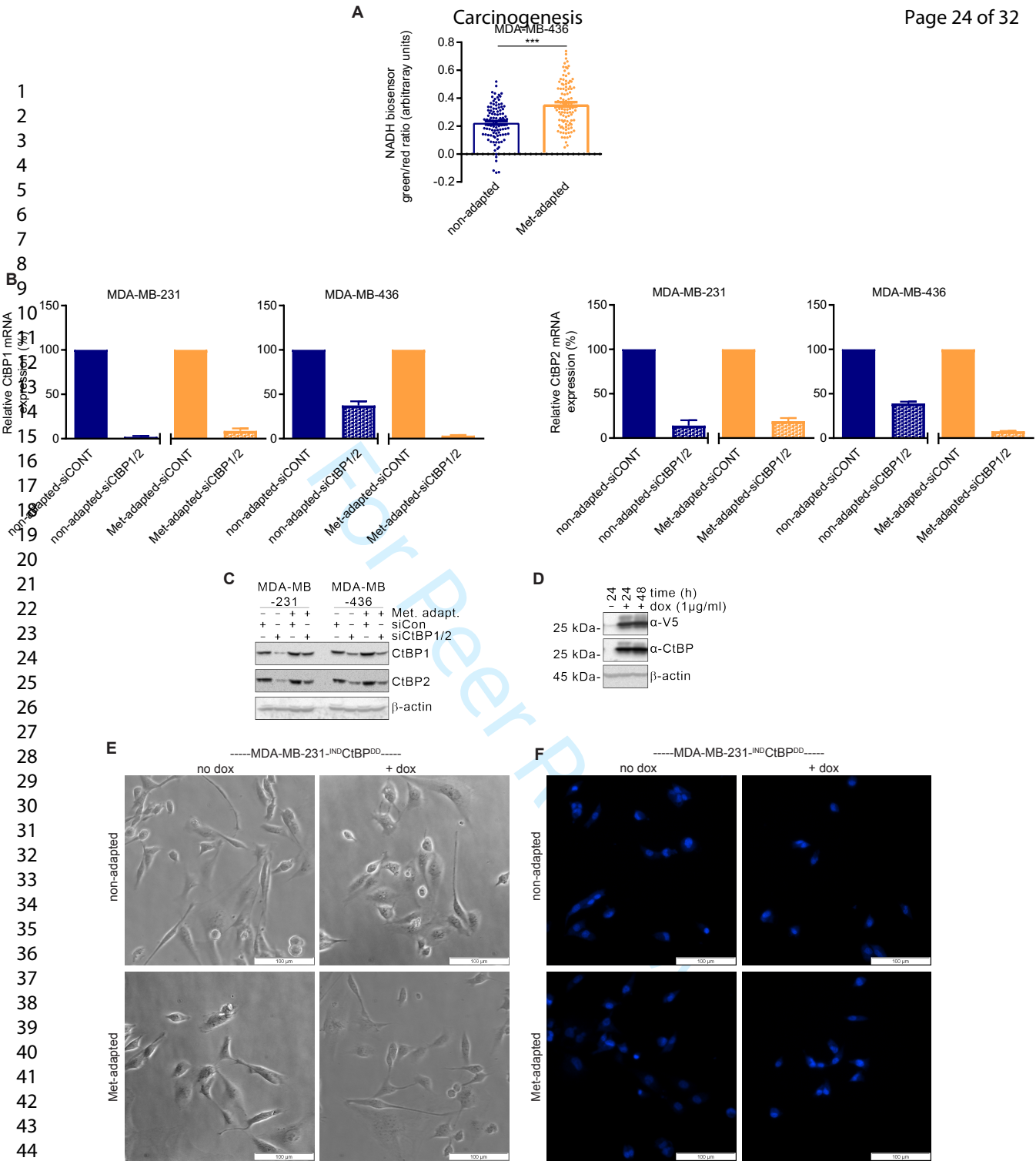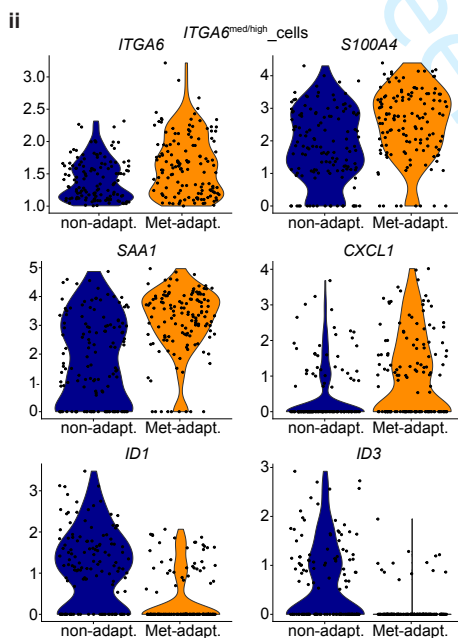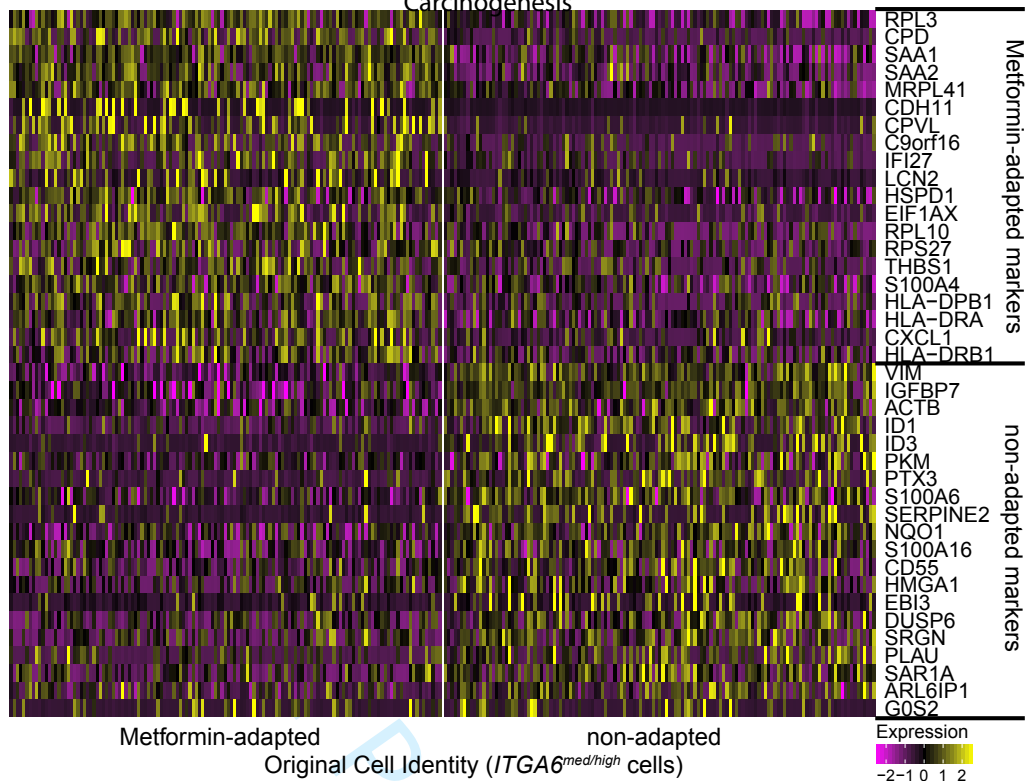
Supplementary Figure S2: Validation of CtBP models used in Figure 5. (A) Effect of metformin adaptation on the NADH levels. MDA-MB-436 were transfected with the nuclear NADH:NAD+ ratio biosensor Peredox-NLS (Y.P. Hung et al, Cell Metab. 2011 14 pp 545-554). 72 hours after transfection cells were transferred to assay medium (DMEM pH 7.4 HEPES, 25 mM glucose, 1 mM pyruvate, 2 mM glutamine, penicillin/streptomycin) without or with 2 mM metformin. Green (NADH sensor) and red (mCherry) fluorescence was captured in > 100 cells per condition from 6 separate wells, using an Olympus IX81 microscope system and background-subtracted green:red ratios quantified using Image J and Microsoft Excel. Statistical analysis is by unpaired t-test (GraphPad Prism (*** P<0.001). (B,C) Validation of efficacy of CtBP1/2 siRNA. (B) mRNA was extracted 48 hours after transfection and quantified by RT-qPCR. Mean ± SEM of n=3 biological replicates (C) protein was extracted 48 hours after transfection and quantified by western blotting. (D) Western blot demonstrating doxycycline induction of CtBP^{DD} protein in MDA-MB-231-^{IND}CtBP^{DD} using antibodies to either anti-CtBP or V5 tag. (E) Phase contrast micrographs of MDA-MB-231-^{IND}CtBP^{DD} 24 h post induction. Bar 100 μm. (F) Phase contrast micrographs of MDA-MB-231-INDCtBP^{DD} from the invasion assays show in Figure 5Cii. Bar 100 μm.
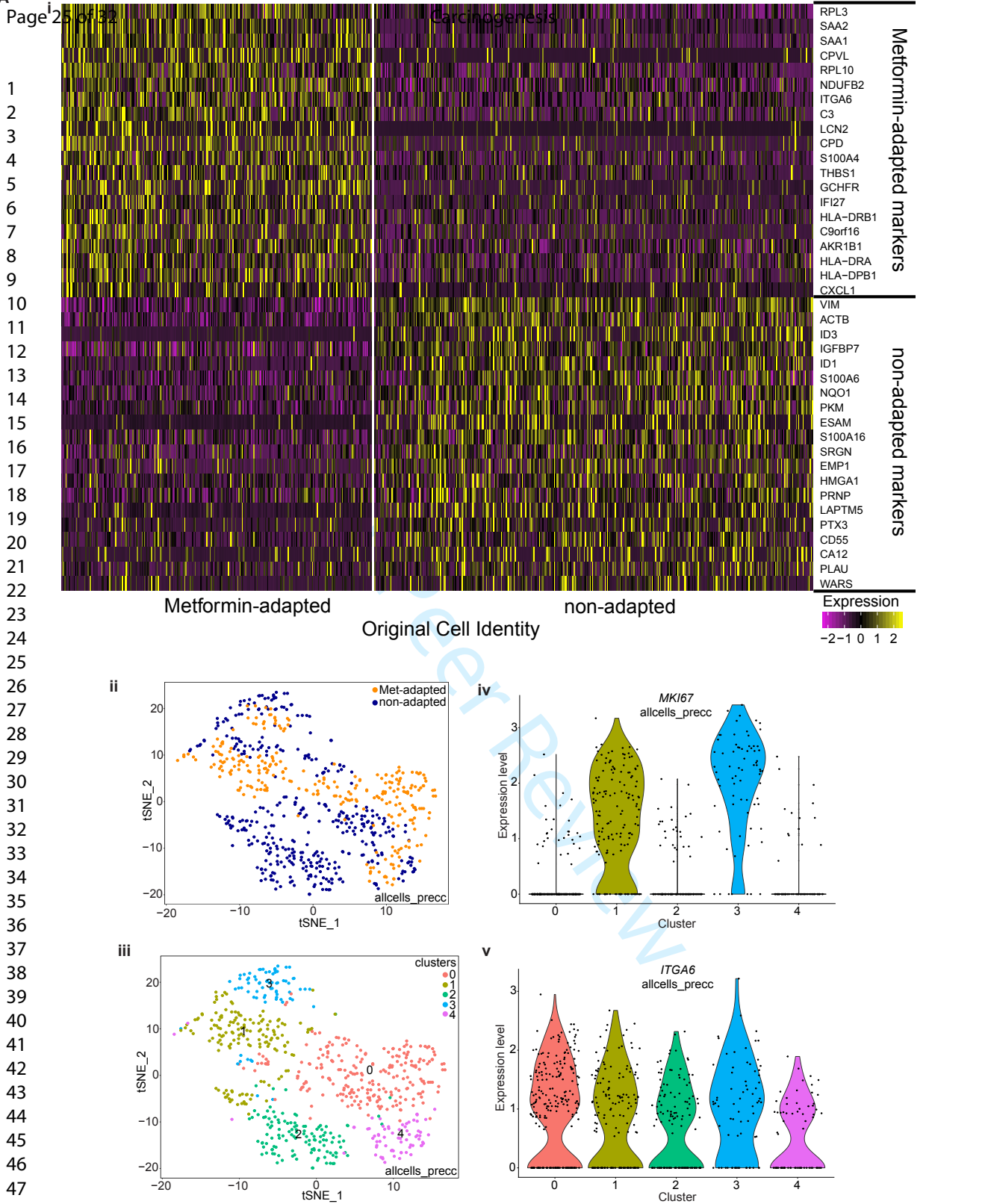
Supplementary Figure S3: Single cell mRNA-seq (Drop-seq) analysis of mammosphere-cultured MDA-MB-231 cells. (A) Drop-seq analysis of MDA-MB-231 mammospheres prior to regressing out of cell cycle scores. Generated by R script Seurat_MET_precc.R (Supplementary File S1). (i) heatmap of variable gene expression between non-adapted and metformin-adapted cells. For full gene list and functional analysis see Supplementary File S2i and S4. (ii) t-SNE plot showing original cell identity. (iii) t-SNE plot showing clustering. (iv) violin plot of MKI67 (Ki67) expression demonstrating the enrichment of cycling cells in clusters 1 and 3. (v) violin plot of *ITGA6* (CD49f) expression demonstrating that this SCLC cell marker is not restricted to cycling cells.
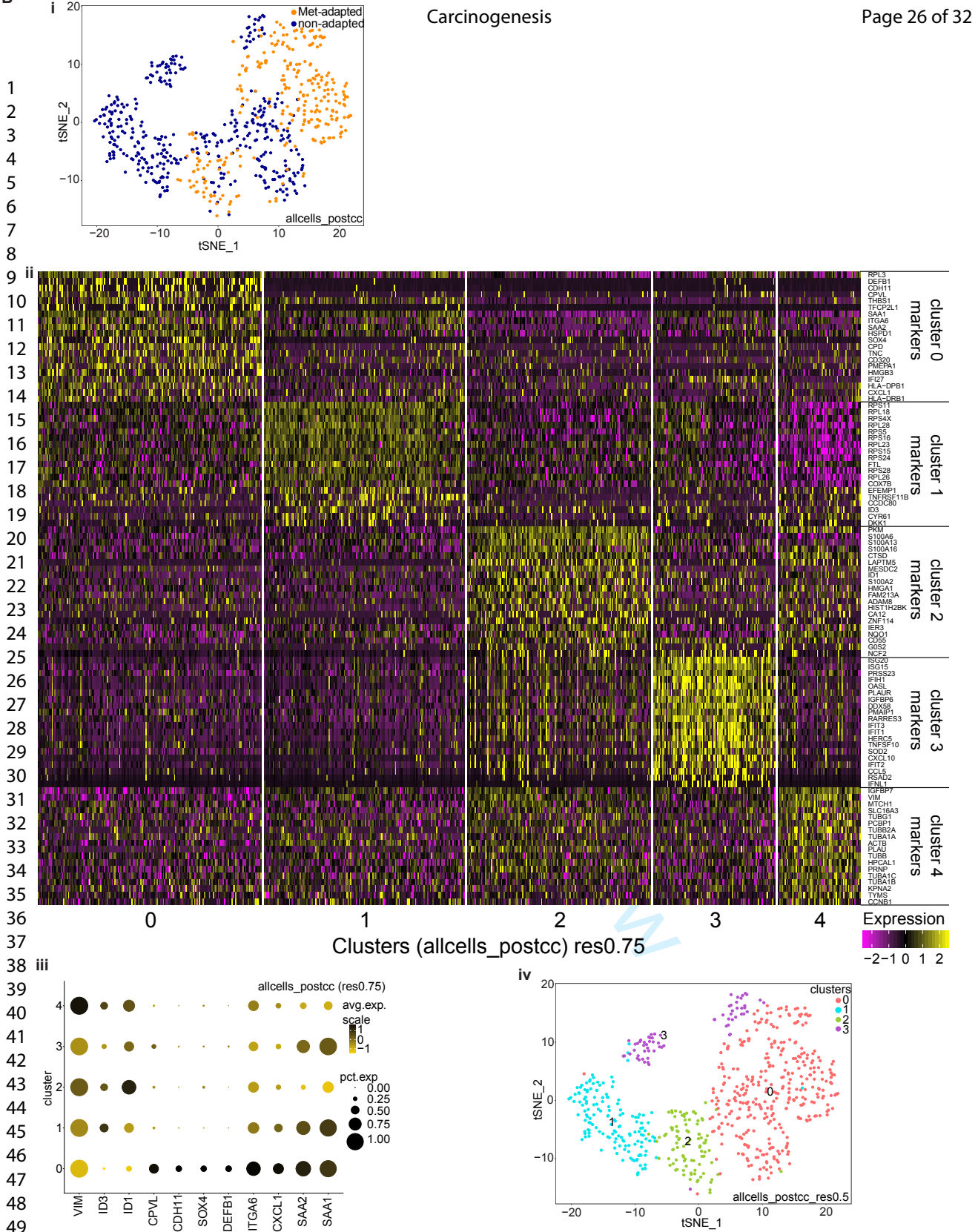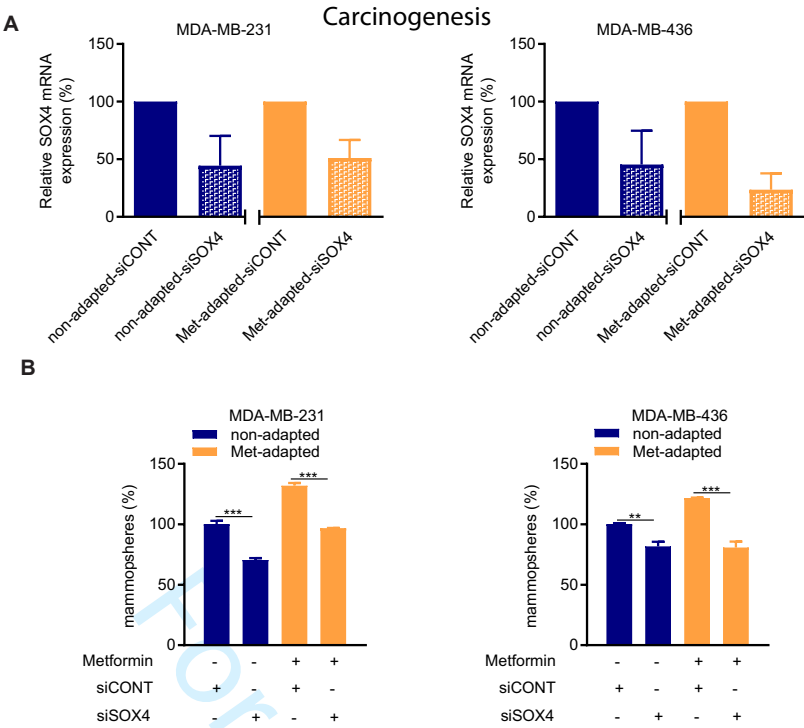
Supplementary Figure S3: Single cell mRNA-seq (Drop-seq) analysis of mammosphere-cultured MDA-MB-231 cells.
(B) Drop-seq analysis of MDA-MB-231 after regression out cell cycle scores. Generated by R script Seurat_MET_postcc.R (Supplementary File S1). (i) t-SNE plot showing original cell identity. (ii) Heatmap of variable gene expression between clusters, with clustering performed at medium resolution (res.0.75). For full gene list see Supplementary File S2iii. (iii) Dotplot showing average expression per cell, and percentage of cells expressing key positive and negative markers of cluster 0 (iv) t-SNE analysis after regression of cell cycle signatures, with cell clusters identified at low resolution (res 0.5). Res 0.5 cluster markers are shown in Supplementary File S2iv.

c Carcinogenesis

Metformin-adapted markers: RPL3, CPD, SAA1, SAA2, MRPL41, CDH11, CPVL, C9orf16, IFI27, LCN2, HSPD1, EIF1AX, RPL10, RPS27, THBS1, S100A4, HLA−DPB1, HLA−DRA, CXCL1, HLA−DRB1

non-adapted markers: VIM, IGFBP7, ACTB, ID1, ID3, PKM, PTX3, S100A6, SERPINE2, NQO1, S100A16, CD55, HMGA1, EBI3, DUSP6, SRGN, PLAU, SAR1A, ARL6IP1, G0S2

Metformin-adapted   non-adapted
Original Cell Identity (*ITGA6*med/high cells)

Expression −2 −1 0 1 2

ii *ITGA6*med/high_cells

ITGA6 | S100A4 | SAA1 | CXCL1 | ID1 | ID3

(non-adapt. / Met-adapt.)

**Supplementary Figure S3: Single cell mRNA-seq (Drop-seq) analysis of mammosphere-cultured MDA-MB-231 cells.** (C) Drop-seq analysis of *ITGA6*med/high MDA-MB-231 cells. Generated by R script Seurat_MET_postcc_ITGA6.R (Supplementary File S1). (i) Heatmap of variable gene expression between non-adapted and metformin-adapted cells. For full gene list and functional analysis see Supplementary File S2v and S5. (ii) Violin plots of showing the effects of adaptation to 2 mM metformin of the expression of the same genes as shown in Figure 6A.

c

iii

Carcinogenesis

Clusters (*ITGA6*^{med/high}cells) res0.5

cluster 0 markers
cluster 1 markers
cluster 2 markers
cluster 3 markers

Expression
−2 −1 0 1 2

iv

Supplementary Figure S3: Single cell mRNA-seq (Drop-seq) analysis of mammosphere-cultured MDA-MB-231 cells.
(C) (iii) Heatmap of variable gene expression between clusters, with clustering performed at resolution 0.5. For full gene list see Supplementary File S2vi. (iv) t-SNE analysis with cell clusters identified at low resolution (0.25). Insert is a bar chart showing the number of cells from each culture condition in the clusters. For full gene list see Supplementary File S2vii.

A

Carcinogenesis

B

Supplementary Figure S4: Effects of SOX4 knockdown on mammosphere-forming ability in non-adapted and Met-adapted cells. Cells were transfected with SOX4 siRNA 48 hours before (A) mRNA was extracted and SOX4 mRNA quantified by RT-qPCR. (B) Mammosphere assay. (A,B) Mean ± SEM of n=3 biological replicates.
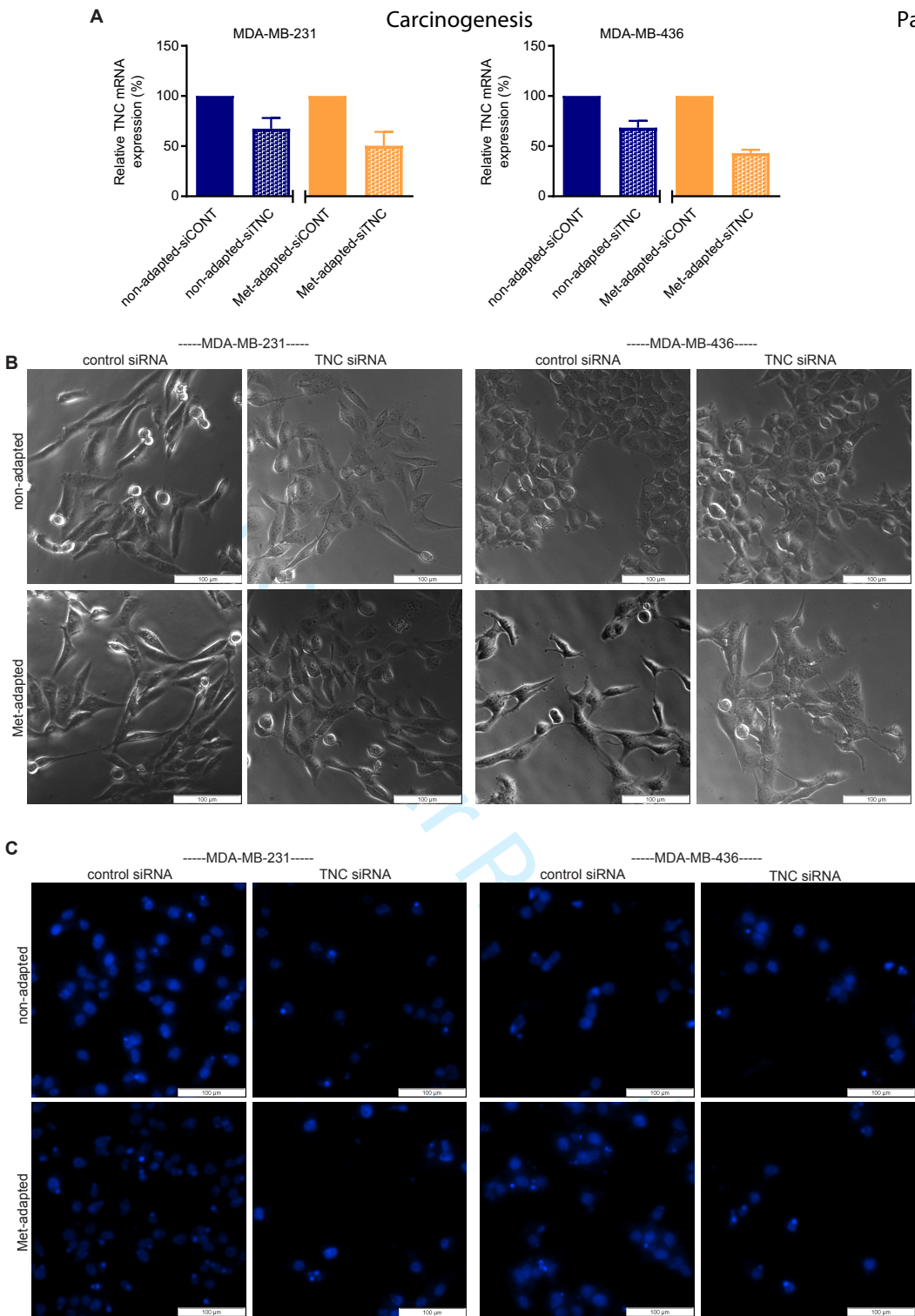
Supplementary Figure S5: Effect of TNC knockdown on non-adapted and Met-adapted cells. Cells were transfected with TNC siRNA 48 hours before (A) mRNA was extracted and TNC mRNA quantified by RT-qPCR. Mean ± SEM of n=3 biological replicates. (B) Photo-microscopy of cells cultured in 2D adherent conditions (bar 100 µm) (C) Invasion assay (representative images from the assays show in in Fig. 6J) (bar 100 µm).

1
2
3
4
5
6
7
8
9
10
11
12
13
14

Claudin low breast cancer (METABRIC)
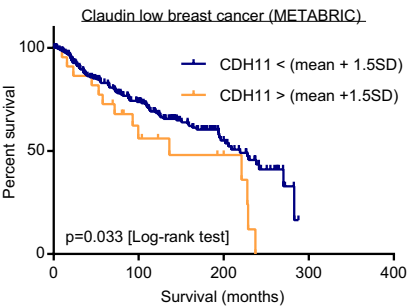
Supplementary Figure S6: High *CDH11* expression correlates with reduced survival in claudin-low breast cancer. The claudin-low tumours from the METABRIC cohort (Curtis, C. et al Nature 486, 346-352 (2012)) were analysed using cBioPortal (http://www.cbioportal.org/) to identify tumours with an mRNA expression (Illumina Human v3 microarray) of *CDH11* of >1.5SD above the mean (22/199 patients). Kaplan-Mayer survival analysis was performed using GraphPad Prism.

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58

LEGENDS TO SUPPLEMENTARY FILES

File S1 (.txt)

Three R scripts used for the analysis of Drop-seq data from MDA-MB-231 cells cultured as mammospheres in media containing either 0 mM or 2 mM metformin.

File S2-S3 Gene lists (.xlsx):

File S2: Markers generated by Function [FindAllMarkers(test.use="bimod")]

(i) allcells_origident

(ii) allcells_precc_clust

(iii) allcells_postcc_clustr0.75

(iv) allcells_postcc_clustr0.5

(v) ITGA6cells_postcc_origident

(vi) ITGA6cells_postcc_clustr0.5

(vii) ITGA6cells_postcc_clustr0.25

File S3: Markers generated by Function [FindMarkers(test.use="roc")]

allcells_postcc_res0.75_0v1

File S4-S5 Gene list enrichment analysis and candidate gene prioritization based on functional annotations and protein interactions network [https://toppgene.cchmc.org/enrichment.jsp] (.xlsx)

File S4: Analysis of S2i

Metformin-adapted UP

Non-adapted UP

File S5: Analysis of S2v

Metformin-adapted UP

Non-adapted UP