C/EBPB-dependent Adaptation to Palmitic Acid Promotes Stemness in Hormone Receptor Negative Breast Cancer

Xiao-Zheng Liu[1,7], Anastasiia Rulina[1,7], Man Hung Choi[2,3], Line Pedersen[1], Johanna Lepland[1], Noelly Madeleine[1], Stacey D'mello Peters[1], Cara Ellen Wogsland[1], Sturla Magnus Grøndal[1], James B Lorens[1], Hani Goodarzi[4], Anders Molven[2,3], Per E Lønning[5,6], Stian Knappskog[5,6], Nils Halberg[1,*]

[1]Department of Biomedicine, University of Bergen, N-5020 Bergen, Norway

[2]Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, N-5020 Bergen, Norway

[3]Department of Pathology, Haukeland University Hospital, N-5021 Bergen, Norway

[4]Department of Biophysics and Biochemistry, University of California San Francisco, San Francisco, CA 94158, USA

[5]Department of Clinical Science, Faculty of Medicine, University of Bergen, N-5020 Bergen, Norway

[6]Department of Oncology, Haukeland University Hospital, N-5021 Bergen, Norway

[7]These authors contributed equally

*Correspondence:

Nils Halberg

Department of Biomedicine

University of Bergen

Jonas Lies vei 91

5020 Bergen, Norway

Phone: +47 5558 6442

Email: nils.halberg@uib.no

Abstract

Epidemiological studies have established a positive association between obesity and the incidence of postmenopausal (PM) breast cancer. However, the molecular mechanisms underlying this correlation are not well defined. A central phenotypic characteristic of obese individuals is increased circulating and interstitial abundance of free fatty acids. Here we demonstrate that long-term exposure to palmitic acid (PA) drives cancer cell dedifferentiation towards a cancer stem-like phenotype and enhanced tumor formation capacity. We demonstrate that this process is governed epigenetically through increased chromatin occupancy of CCAAT/enhancer-binding protein beta (C/EBPB). C/EBPB regulates cancer stem-like properties by modulating the expression of key downstream regulators of the extracellular matrix (ECM) including *SERPINB2* and *LCN2*. Collectively, our findings demonstrate that C/EBPB plays a critical role in the initiation of cancer cells in obesity.

Key words:

C/EBPB, obesity, breast cancer, palmitic acid, stemness

Statement of Significance:

Cellular adaptation to obesity-induced palmitic acid drives tumor initiation through activation of a C/EBPB-dependent transcriptional network. This highlights a mechanistic connection between obesity and postmenopausal hormone receptor negative breast cancer.

INTRODUCTION

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer-related death amongst women. Risk factors for breast cancer include the non-modifiable factors such as age, genetics and reproductive history, as well as modifiable factors as obesity, alcohol consumption, and tobacco smoking. The majority of cancers can be attributed to a combination of several of such genetic, hormonal, and environmental factors. As an independent risk factor, PM obesity accounts for up to 20% higher risk of developing breast cancer, and every 5-unit increase in BMI is associated with a 12% increase in breast cancer risk (1). Whereas obesity in

PM individuals has been consistently linked to enhanced risk of developing estrogen receptor (ER) positive breast cancer, there has been more debate on the effect in ER negative breast cancer (2). These discrepancies likely relate to challenges to align study designs across heterogeneous studies, highlighting the need for a mechanistic understanding on the interactions between the obese state and breast cancer risk. In addition to the reported effects on breast cancer incidence, several meta-analyses have shown that overweight and obesity are associated with worse overall survival and metastasis-free survival independent of their menopause or hormone receptor status (3,4). In ER positive breast cancers, the link to obesity has been attributed to increased estrogen signaling (5,6). However, in ER negative breast cancers the molecular mechanisms of this connection are largely unknown – particularly for obesity-induced tumor initiation. Thus far, proposed mechanisms includes obesity-induced chronic inflammation (7,8), altered insulin signaling (9,10), deregulation of estrogen (11), rewiring of cancer metabolism (12) and secreted adipokines (13).

Here we aim to determine the molecular mechanisms that link breast cancer and obesity. We demonstrate that obesity had adverse effects on patient survival in PM, ER/progesterone receptor (PR) negative breast cancers compared to other subtypes. Using single cell-based analysis we show that long-term exposure to high concentrations of palmitic acid (PA) led to dedifferentiation towards cancer stem cell-like properties across human and mouse cell models of ER/PR negative breast cancer cell lines as determined by higher expression of CD44, CD133 and Axl. Using patient tissue microarray (TMA) we show that the frequency of cancer cells expressing the stem cell marker CD133 were more abundant in samples from overweight and obese patients compared to normal weight ER/PR negative PM patients. ATACseq and Cut&Run coupled with transcriptomic analysis of cells adapted to high PA levels followed by loss-of-function and gain-of-function studies, identified CCAAT/enhancer-binding protein beta (C/EBPB) as a required transcriptional regulator of PA-induced cancer stem-like properties. We further demonstrate that C/EBPB induced stemness through the modulation of the ECM proteins SERPINB2 and LCN2. Taken together, our findings indicate that C/EBPB plays a critical role in the initiation of PM/ER$^-$/PR$^-$ breast cancer cells in obesity.

RESULTS

Transcriptional changes induced by long-term culture in palmitic acid overlap with obesity-dependent transcriptional changes in hormone receptor negative patients.

To set the framework for our mechanistic studies of the connection between obesity and breast cancer, we first sought to identify a group of patients affected by the obese state. To that end, we performed Cox regression survival analyses of 115 PM (defined by age of >50) breast cancer patients using BMI and hormone (estrogen and progesterone) receptor status as variables in a highly controlled in-house dataset (14). Despite equal BMI distribution (Figure S1A), overweight and obesity (BMI>25) were associated with significantly reduced survival rates in hormone receptor negative patients as compared to non-obese patients (Figure 1A) whereas no effects of BMI were observed in the hormone receptor positive patients (Figure S1B). Importantly, within the PM/ER$^-$/PR$^-$ patient group there were no differences between the high and low BMI groups in patient age (Figure S1C), tumor size (Figure S1D) or tumor stage (all included patients were stage 3) at the time of diagnosis. Obesity leads to the production of reactive oxygen species in adipose tissue (15,16). Given the abundant adipose tissue in the mammary gland and the association between reactive oxygen species and mutagenesis (17) we next performed high-coverage sequencing of 360 known cancer genes (18) in tumor samples collected from the PM/ER$^-$/PR$^-$ patient group at the time of diagnosis. Based on this analysis we were not able to detect any mutations correlating to obesity across this panel (Supplementary Table S1). Combined, this suggests that in PM/ER$^-$/PR$^-$ breast cancer patients, obesity potently modulates the tumor biology towards a more aggressive disease independent of genetic changes. The obesity-programmed environment is complex, and includes an altered immune status and deregulated abundancies of circulating hormones and metabolites. To understand how cancer cells evolving in such an environment impacts patient survival, we focused on the interaction between obesity-induced circulating levels of free fatty acids (19-21) and cancer phenotypes. In particular, we were interested in PA as this is the most abundant fatty acid in circulation, and has been reported to be epidemiologically associated with a higher risk of developing PM breast cancer (22). We exposed hormone receptor negative breast cancer lines to increasing PA concentrations over a period of 2 months to enable cellular growth in PA concentrations corresponding to serum levels in obese patients (Figure 1B). Human hormone receptor negative (MDA-MB-231 and HCC1806) and mouse (E0771 and TeLimet) breast cancer cells consistently adapted to acquire resistance to PA-induced apoptosis to enable persistent growth even in the high PA environment (Figure 1C, D). In MDA-MB-231, HCC1806, and TeLimet cells, the acquired resistance was accompanied by a reduction in growth rate, whereas E0771 cells maintained its growth rate even after adaptation to high levels of PA (Figure 1E). To ascertain how such adaptation resembles what is observed in obese breast cancer patients, we compared the transcriptional alterations observed during cellular adaptations to PA to the transcriptional changes induced by obesity in PM hormone

negative breast cancers patients. To this end, we applied iPAGE, an information-theoretic framework (23), to query how genes induced or repressed in obesity were changed upon adaptation to PA in the *in vitro* model. For this analysis, genes were first ordered based on their expression changes between MDA-MB-231 parental and adapted cells (termed MDApar and MDAapa, respectively) and were subsequently divided into 10 equally populated bins. We then assessed the distribution of obesity-associated genes across these bins. As shown in Figure 1F, we observed a significant depletion/enrichment pattern (MI=0.006 bits, z-score=21.14). We specifically noted a significant overlap between genes that were induced by the obesogenic state in patients and those up-regulated through *in vitro* adaptation to PA (Figure 1F). This shared reprogramming of the gene expression landscape suggested that the *in vitro* long-term adaptation to high abundancies of PA provides clinically relevant information on the molecular drivers of obesity-induced hormone receptor negative breast cancers.

Long-term adaptation to palmitic acid in culture facilitates cellular dedifferentiation towards cancer stem cell-like properties

Cellular adaptation to new extracellular environments is a central mechanism underlying disease progression and therapy resistance (24). To understand the cellular phenotypes enriched in breast cancer cells subjected to long-term exposure to PA, we performed a single cell mass cytometry analysis using an antibody panel targeting 27 markers of cellular differentiation states and signaling pathways (Figure S2A). The distribution of cellular subpopulations during adaptation to the PA-rich environments was analyzed using the X-shift algorithm (25) for both MDA-MB-231 (Figure 2A,B) and HCC1806 cells (Figure 2D,E). This analysis revealed an enhanced expression of cancer stem cell markers CD44 (26,27), CD133 (28) and Axl (29) in the PA-induced cell clusters (CL1059 for MDAapa and CL731 for HCCC1806apa) of the adapted cells (Figure 2C, F). This suggested that long-term adaptation to PA induced a consistent cellular dedifferentiation towards a cancer stem-cell like phenotype in MDA-MB-231 and HCC1806 cells. Increased frequency of $CD133^{+}$ cell populations were validated using flow cytometry (Figure 2G). To test if the tumor stem cell-like phenotype was associated with tumor formation capacity *in vitro*, we subjected parental and adapted cells to a tumor spheroid formation assay. All four adapted cell lines were consistently able to form spheroids at significantly higher frequencies compared to their parental counterparts (Figure 2H). Combined, these findings suggested that cellular adaptation to long-term exposure of fatty acids induces a stem cell-like cancer phenotype.

Obesity is associated with increased frequency of stem cell-like cancer cells in PM/ER$^-$/PR$^-$ breast cancer patients and mouse models of breast cancer

Having shown that the adapted cell lines were associated with enhanced tumor spheroid formation *in vitro*, we then wondered if this was translatable to in the clinical setting. To that end, we obtained tumor tissue microarrays (TMA) from the PM/ER$^-$/PR$^-$ patients used in the initial survival analysis (Figure 1A) and immunostained the cores for CD133. The image analysis platform QuPath (30) was used to segment the images, differentiate between stromal and cancer cells and to quantify CD133$^+$ cell frequencies. Consistent with our *in vitro* observations, PM/ER$^-$/PR$^-$ breast cancer patients with a BMI above 25 displayed higher CD133$^+$ cancer cell frequencies as compared to the normal BMI patients (Figure 3A). To confirm this obesity-induced cancer stem cell-like phenotype *in vivo*, we orthotopically implanted E0771 and TeLi cells at limiting dilutions in a C57BL/6J diet-induced obesity model and measured tumor formation. Following 10 weeks of high-fat diet (HFD) feeding the mice gained more weight (Figure 3B) and displayed multiple hallmarks of obesity-induced comorbidities such as liver steatosis, hyperinsulinemia, hyperglycemia and reduced glucose clearance compared to the regular chow fed mice, suggesting that the diet-induced obesity recapitulated the systemic obese environment found in humans (Figure 3C, D, E). Following mammary gland implantation of limiting number of E0771 and TeLi cells we found that the high-fat environment consistently promoted tumor formation with a 6 to 10-fold enrichment in stem cell frequencies (Figure 3F). To test the robustness of the induced stem cell-like phenotype *in vivo*, we next dissociated tumors formed in the obesogenic and non-obesogenic environments, allowed them to grow in standard 2D culture for 4 days and subjected them to an *ex vivo* spheroid formation assay. Interestingly, both E0771 and TeLi cancer cells adapted to the *in vivo* obesogenic environment maintained their enhanced ability to form spheroids (Figure 3G). Consistent with the overlap between transcriptional signatures between obese patients and the PA-adapted cell lines, E0771apa cells formed tumors earlier than E0771par cells (Figure 3H). Combined, these findings suggested that an obesogenic environment promoted a cancer stem-cell phenotype in breast cancer cells in patients, in *in vivo* obese breast cancer mouse models and in *ex vivo* cellular models upon long-term adaptation to PA.

Adaptation to a PA-rich environment induces metabolic reprogramming towards fatty acid oxidation

Metabolic reprogramming has been linked to stem-cell behavior in breast cancer (31,32). To understand the mechanism(s) by which PA adaptation drives differentiation towards a stem-cell like phenotype, we first metabolically characterized parental and adapted MDA-MB-231 and HCC1806 cells using radiolabeled [1-^{14}C]PA and D-[^{14}C(U)]glucose as tracers. After adaptation to a PA-rich environment, both MDA-MB-231 and HCC1806 cells displayed higher reliance on free fatty acid oxidation (Figure 4A and B). In contrast, glucose oxidation was unchanged in MDA-MB-231 cells and slightly decreased in adapted HCC-1806 cells (Figure 4A and B). Consistent with the *in vitro* system, E0771 cells isolated from tumors grown in obese and non-obese mice also displayed increased PA oxidation and reduced glucose oxidation *ex vivo* (Figure 4C). In addition to deregulated PA oxidation, we also found increased lipid droplet formation in adapted cells as assessed by incorporation of fluorescent dye BODIPY, suggesting that adaptation to PA enhanced increased fatty acid oxidation and storage (Figure 4D). To functionally test if such metabolic reprogramming towards fatty acid oxidation might be linked to the observed stem cell-like properties, we treated cells with etomoxir (an inhibitor of fatty acid oxidation by inhibiting free fatty acid uptake into the mitochondria) and subjected them to the spheroid formation assay. However, at concentrations that robustly reduced cellular respiration (Figure 4E), we were unable to detect any effects on spheroid formation capacity in either HCC1806apa or E0771apa cells, suggesting that the shift in substrate oxidation was a correlative event and not the main driver of the PA-induced stem cell-like phenotype (Figure 4F).

Adaptation to a PA-rich environment induces open chromatin linked with C/EBPB occupancy

Deregulation of metabolic intermediates has recently been tightly linked to epigenetic remodeling and cell fates (33). We therefore next assessed chromatin accessibility by ATAC sequencing (ATACseq) of parental and PA-adapted cells. Interestingly, adaptation to PA led to a substantial gain and loss in chromatin accessibility in 42336 and 48991 regions across the genome, respectively (Figure 5A and Figure S5A, S5B). As expected, chromatin accessibility changes in promoter regions were correlated with concordant changes in transcription of the downstream genes as determined by RNA sequencing (RNAseq) (Figures 5B, C and Figure S5C, S5D). To identify potential regulators associated with the observed chromatin alterations upon PA-adaptation, we next aggregated changes in chromatin accessibility near putative binding motifs to infer differential motif activity and occupancy of transcription factors (34). This analysis identified the C/EBPB transcription factor as the topmost hit associated with the more accessible

chromatin in the PA-adapted cells (Figure 5D, E). This suggested that C/EBPB acts as a transcriptional regulator of obesity-induced tumor initiation capacity in obese breast cancer patients. Functional depletion of C/EBPB by independent short hairpin RNAs (shRNAs) in MDAapa, HCC1806apa and E0771apa cells (Figure S5E-G) led to a significant reduction in spheroid formation capacity *in vitro* (Figure 5F). In contrast, depletion of RUNX1 and C/EBPBA did not affect spheroid formation capacity (Figure S5H-K). Further, upon transplantation into the mammary fat pad, depletion of C/EBPB significantly delayed tumor formation in the diet-induced obese setting, while the knockdown had no effect in the non-obese setting (Figure 5G). All together, these experiments support a model wherein C/EBPB is associated with transcriptionally active chromatin and is required for the cancer stem-like phenotype in obesity.

C/EBPB is encoded by an intron-less gene that is expressed in three isoforms; LAP1, LAP2 and LIP by alternative use of transcription start sites (35,36). Both LAP1 and LAP2 isoforms contain a dimerization and a transcriptional regulation domain and functions as dimers (35). LIP lacks the DNA binding domain and has been suggested to function as a competitive inhibitor of LAP1/2 (35). We next asked whether C/EBPB expression itself was sufficient to confer stem-like properties. To this end, we overexpressed either LAP2 or LIP in MDA-MB-231 cells (Figure S5L). The LAP2 construct contains a conservative mutation ATG to ATC (Met to Ile) that removes the LIP translational start site and thereby preventing the co-expression of LIP (35). MDApar cells fail to form spheroids under normal culture conditions, however ectopic overexpression of LAP2, but not the LIP isoform, enabled spheroid formation in all biological replicates (Figure 5H). Interestingly, overexpression of LAP2 in MDAapa cells further enhanced spheroid formation after 1 day of culture as opposed to 5 days in control conditions (Figure 5I). In contrast and in line with its suggested dominant negative effect, LIP overexpression in the MDAapa cells abrogated spheroid formation (Figure 5I). Moreover, LAP2 overexpression increased the fraction of $CD133^{+}$ cells, and this effect was more pronounced in adapted than in parental cells (Figure 5J), consistent with the higher chromatin accessibility to genomic regions containing C/EBPB binding motifs in the adapted cells. These findings collectively suggested that the C/EBPB isoform LAP2 is a key regulator of cancer stem stem-like properties.

Unexpectedly, we found that the protein levels of C/EBPB isoforms did not differ between adapted and parental cells (Figure S5M-O). Neither did we detect any differences in C/EBPB nuclear localization (Figure S5P). The lack of difference in protein levels or localization suggested that the activity of C/EBPB might be regulated post-translationally, and that its effect on cancer

stemness could be associated with epigenetically determined increased occupancy in chromatin regions with gains in target accessibility.

Differential C/EBPB occupancy regulates the expression of extracellular matrix proteins

Having shown that C/EBPB is required and sufficient for spheroid formation capacity, we next applied Cut&Run to confirm its genome-wide occupancy and to identify its putative downstream transcriptional targets. Cut&Run uses micrococcal nuclease tethered to DNA-bound proteins to generate short DNA cleavage fragments and thus enables base-resolution digital footprints that reflect precise protein-DNA binding sites (37). We enumerated the ends of every Cut&Run fragment (≤ 120 bp) for each base of the genome and detected significant footprints de novo based on the footprint occupancy score (38). As expected, motif enrichment analysis identified C/EBPB as the topmost enriched motif in significant Cut&Run footprints, confirming successful genome-wide profiling of C/EBPB under PA adaptation (Figure 6A, Figure S6A and S6B). Similar results in motif enrichment were confirmed using a peak-based approach (Figure S6C).

In line with the ATACseq data, Cut&Run confirmed increased C/EBPB occupancy, in the same chromatin regions which had increased accessibility in the adapted cells as compared with the parental cells (Figure 6B). We linked distal and proximal gains in C/EBPB occupancy and chromatin accessibility in PA adaptation to genes whose expression correspondingly increased, and/or based on high-confidence enhancer-gene associations identified cross-platform in GeneHancer (39) (e.g. *LCN2*; Figure 6C). Pathway analysis of these regions revealed a significant enrichment in processes involved in extracellular matrix (Figure 6D), suggesting a potential link between ECM remodeling and cancer stemness.

To translate our data derived from the *in vitro* PA adaptation system into the *in vivo* and clinical settings, we subsequently focused on the subset of the putative C/EBPB target genes whose expression was significantly elevated in the obese as compared to the lean PM/ER$^-$/PR$^-$ patients. This analysis identified nine genes, namely, *SERPINB2*, *LCN2*, *SERPINB7*, *NELL2*, *MMP9*, *CLDN1*, *LYPD6B*, *CRISPLD1* and *CHST4* (Figure 6E). Interestingly, all of these nine genes had elevated expression in E0771 cells analyzed *ex vivo* after having been grown in obese as compared with non-obese mice; whereas the expression of C/EBPB was unchanged as expected (Figure 6F). In short, these data supported a model wherein obesity induced C/EBPB chromatin binding, activating a transcriptional network involved in ECM processes.

C/EBPB target genes SERPINB2 and LCN2 are required for cancer stem cell-like capabilities.

To determine the functional importance of the nine genes in CEPBP-dependent cancer stemness, we next assessed the levels of the nine genes in cells where C/EBPB was overexpressed. Ectopic overexpression of the LAP2 isoform of C/EBPB in MDAapa cells led to the induction of five of these nine genes (Figure 7A); whereas ectopic expression of LIP did not affect the expression level of the genes (Figure 7A). Interestingly, LAP2 overexpression particularly augmented the expression of SERPINB2, LCN2 and CLDN1, which paralleled the differential expression patterns observed in cells adapted to obese and non-obese environment (Figure 6E). We therefore functionally tested the role of SERPINB2, LCN2 and CLDN1 in epistatic spheroid formation assay and found that SERPINB2 and LCN2 were required for LAP2 induced spheroid formation capacity (Figure 7B and Figure S7A-C). The combined depletion of SERPINB2, LCN2 and CLDN1 also prevented LAP2-induced spheroid formation (Figure 7B). Further, individual depletion of SERPINB2 and LCN2 (Figure S7D, E) phenocopied C/EBPB knockdown and significantly reduced tumor spheroid formation capacity (Figure 7C, D). This suggested that SERPINB2 and LCN2 were the main downstream mediators of C/EBPB. To assess the clinical impact of these findings, we stratified PM/ER$^-$/PR$^-$ patients according to their SERPINB2 and LCN2 expression and determined survival outcomes in our in-house as well as an independent dataset (GSE25066). Consistent with C/EBPB driving a more aggressive cancer phenotype in PM/ER$^-$/PR$^-$ breast cancer patients, these survival analyses demonstrated that high expression of SERPINB2 and LCN2 were associated with worse survival outcomes (Figure 7E-F). Combined, these findings support a model wherein the obese environment epigenetically activates C/EBPB transcriptional activity that is required and sufficient for tumor formation capacity through the regulation of its target genes SERPINB2 and LCN2 (Figure 7G).

DISCUSSION

Obesity is a complex pathological condition that conceivably affects the formation and development of cancers through multiple avenues. Here we have demonstrated that cancer cell adapted to high levels of PA is one such potent mechanism through which obesity drives enhanced tumor formation capacity in PM/ER$^-$/PR$^-$ breast cancer. We find that adaptation to PA

governed dedifferentiation of cancer cells towards a tumor stem cell-like phenotype leading to augmented tumor formation capacity. Clinically this manifest in a higher cancer cell frequency of $CD133^+$ cancer stem cells and shorter disease-specific survival in obese and overweight $PM/ER^-/PR^-$ breast cancer patient compared to normal weight patients. This is corroborated epidemiologically by the association of obesity with higher cancer risk (40) and poor prognosis (4) of $PM/ER^-/PR^-$ breast cancer patients. Our findings thus provide a cancer cell autonomous mechanism for the increased appreciation that obese environments lead to enhanced tumor formation capacity in breast cancer (7,11,41-43). Our findings further expand on the molecular mechanisms of PA-induced stemness, by demonstrating that that obesity-induced stemness is mediated through the epigenetic activation of a C/EBPB dependent transcriptional network. At the genome-wide level we demonstrate that the obese environment facilitates a widespread deregulation of chromatin accessibility. Chromatin regions with increased accessibility in cells adapted to high levels of PA were enriched for C/EBPB binding motifs. Through complementary sets of *in vitro* and *in vivo* experiments, we show that C/EBPB is required for obesity-induced tumor formation. Conversely, ectopic overexpression of C/EBPB enhanced the frequency of cancer stem cells. Previous reports observed that C/EBPB is required for stem cell maintenance in the developing breast (44) and that expression of the LAP2 isoform of C/EBPB can transform a non-cancerous cell line MCF10A (45), lending further support to the functional role for C/EBPB-dependent cancer stem cell-like properties.

Our unbiased Cut&Run analysis of direct C/EBPB target genes suggested that C/EBPB regulates stemness features through regulation of the surrounding ECM. Cancer cell-autonomous regulation of the ECM is intrinsically linked to cancer stemness through manipulation of mechanical properties and signaling molecules (46,47). Consistent with our findings, obesity-induced alterations in the ECM mechanics has been reported to support tumorigenesis (48). Interestingly, a total of 9 C/EBPB target genes were also induced in obese $PM/ER^-/PR^-$ breast cancer patients. Of these 9 genes, depletion of ECM proteins SERPINB2 and LCN2 phenocopied C/EBPB knockdown and were epistatically required for C/EBPB induced spheroid formation capacity suggesting that these engender the downstream effects of C/EBPB. Both of these factors have previously been implicated in the regulation of cancer stem cell-like properties. SERPINB2, also known as plasminogen activator inhibitor type 2 (PAI-2), is widely described as an extracellular urokinase inhibitor that is upregulated in many inflammatory states. In cancer biology, SERPINB2 was observed to be induced in brain metastatic breast cancer cell subpopulations (49) and was recently suggested to be broad marker for cancer stemness in multiple cell culture models (50). Although the mechanism linking SERPINB2 to stemness,

particularly in the context of an obese environment, is currently unknown, our work suggests a link to extracellular plasmin homeostasis. Interestingly, our unbiased analysis of C/EBPB dependent drivers of obesity-induced stemness in breast cancer additionally highlighted a functional role of the small extracellular protein LCN2. LCN2 is induced in adipose tissue of obese individuals (51) and were previously described to reduce inflammation and fibrosis and in an obesity-driven pancreatic ductal adenocarcinoma model (52). In breast cancer, LCN2 has been linked to cellular differentiation through modulation of the epithelial to mesenchymal transition (53). While SERPINB2 and LCN2 factors have been suggested to be involved cancer stemness, future work is needed to establish the mechanistic basis of their actions – especially in the context of obese environments.

Aberrant lipid metabolism is a hallmark of deregulated cancer metabolism (54). It has been widely reported that cancer cells augment their *de novo* lipid biosynthesis for energy production, synthesis of new membranes, to regulate membrane structures that coordinate signal transduction, and for the biosynthesis of lipid signaling molecules such as phosphatidylinositol-3,4,5-trisphosphate (55). In addition, cancer cells can stimulate the release of fatty acids from surrounding adipocytes to provide energy for tumor growth (56). Here we demonstrate that an obese environment governs a metabolic switch towards higher fatty acid oxidation. However, this reprogramming was not causally linked to obesity-induced stemness. In support of a link between fatty acids and stemness, is the observation that slow-cycling metastasis-initiating cells are dependent on the lipid uptake protein CD36 (57). While we did not observe any direct involvement of CD36 in our studies of obesity-induced breast cancer, both studies describe a critical role for fatty acid metabolism in cancer stemness.

Our findings furthermore identify a critical link between adaptation to obese environments and genome-wide changes in chromatin accessibility. This is analogous to recently observations that high fat feeding leads to alterations in chromatin interactions to drive adaptive networks (58). These interactions likely reflect diet-induced alterations in metabolic intermediates that are intimately connected to epigenetic control of gene transcription (59,60). Interestingly, lipid-derived acetyl-CoA has been suggested to be the source of up to 90% of acetylation modifications of certain histone lysine's (61).

Combined, our analysis of cellular adaptations to obese environments has revealed changes of cellular phenotypes, driven by the combined modulation of C/EBPB transcriptional activity. In the context of personalized medicine, this suggest that obese cancer patients might benefit from specific targeted therapies rather than generic treatment regiments.

METHODS

Breast Cancer Patient Cohort

This study enrolled a total of 223 patients with primary stage III breast cancers. Out of these 115 patients were PM patients (defined by age > 50 years). Recruitment period was between November 24, 1997 and December 16, 2003. The median age was 51 years (range 25–70). Patient's BMI, age, hormone status at the time of diagnosis as well as patient survival times (overall survival and disease specific survival) were documented. The study was approved by the regional committees for medical and health research of Western Norway (REK-Vest; approval number 273/96-82.96). More details about the study cohort can be found in the following report (14).

Animal Models

All animal experiments were approved by the Norwegian Animal Research Authority and conducted according to the European Convention for the Protection of Vertebrates Used for Scientific Purposes, Norway. The Animal Care and Use Programs at University of Bergen are accredited by AAALAC international. The laboratory animal facility at University of Bergen was used for the housing and care of all mice. C57BL/6J mice were obtained from Jackson Laboratories and bred on site. Female mice were kept in IVC-II cages (SealsafeÒ IVC Blue Line 1284L, Tecniplast, Buguggiate, Italy); 5-6 mice were housed together and maintained under standard housing conditions at 21°C ± 0.5°C, 55% ± 5% humidity, and 12h artificial light-dark cycle (150 lux). Mice were provided with standard rodent chow (Special Diet Services, RM1 801151, Scanbur BK, Oslo Norway) and water *ab libitium*.

To mimic both obese and non-obese environments, 6 weeks old female littermates were randomly assigned to chow and HFD groups and fed either standard chow diet (75% kcal from fat, 17.5% from proteins and 75% from carbohydrates, Special Diet Services RM1, 801151) or high fat containing diets (60% kcal from fat, 20% from protein and 20% from carbohydrates, Research Diets, D12492) for 10 weeks prior to tumor cell implantations. Body weight was monitored every week. The respective diets were maintained throughout the experiment.

Cell Lines and Culture

MDA-MB-231 (TNBC, human), HCC1806 (TNBC, human) and HEK293T cell lines were purchased from the American Type Culture Collection (ATCC). E0771 (TNBC, mouse) cell line was purchased from the CH3 BioSystems. TeLi (basal breast cancer, mouse) cells were originally derived from a tumor formed in MMTV-Wnt1 transgenic mouse and then propagated *in vivo* for four generations through mammary fat pad injections before being passaged *in vitro*. Tumors was dissociated using Mouse tumor dissociation kit (Miltenyi Biotec, 130-096-730) according to manufacturer's instructions. Dissociated tumor cells were cultured *in vitro* for two months to obtain pure tumor cells. The *in vivo* passaged MMTV-Wnt cells were kindly provided by Stein-Ove Døskeland, University of Bergen. The TeLimet cell line was generated in house by dissociating lung metastasis derived from tail vein injected TeLi cells that were stably transfected with a reporter plasmid containing green fluorescence protein (GFP) and luciferase. MDA-MB-231, E0771, TeLi and TeLimet cells were cultured at 37°C, 5% CO_2 in high-glucose DMEM (Sigma, D5671) supplemented with 10% FBS (Sigma, F-7524), 100U/mL penicillin and 100 µg/mL streptomycin (Sigma, P-0781) and 2 mM L-glutamine (Sigma, G-7513). HCC1806 cells were cultured in RPMI1640 (Sigma, R8758) supplemented with 10% FBS and 100U/mL penicillin/ and 100 µg/mL streptomycin.

For cell line authentication, MDA-MB-231 cells were harvested for genomic DNA extraction using Genomic DNA isolation kit (Norgen Biotek, 24700). Isolated genomic DNA was analyzed by Eurofins Genomics laboratory and the cell line authenticated based on genetic fingerprinting and short tandem repeat (STR) profiling.

Patient Tissue Microarray and Transcriptomic Analysis

Tissue Microarray

Tissue specimens were from the human breast cancer patient cohort described above (14). At the time of diagnosis, each patient from the study cohort had an incisional tumor biopsy. All tissue samples were fixed in formaldehyde for paraffin embedding, in the operating theatre immediately on removal. Paraffin embedded tissue were subject to tissue microarray (TMA) construction. From each tumor, 4 cores of 1.2 mm diameter from tumor rich areas were punched out using Manual Tissue Arrayer Punchers (MP10; Beecher Instruments). The patient cores were embedded into

ten 8 x 10 array blocks plus 1 to 2 liver control cores for orientation. Microtome sectioned slides were stored at 4°C until ready for use.

Immunohistochemical staining was done as described previously (28). In short, slides were dried at 58°C over two days and deparaffinization was performed using xylene, rehydrated with ethanol and dH_2O. Target retrieval was done in Tris/EDTA buffer, pH 9 (Dako, S2367) in a microwave for 25 min. Slides with buffers were cooled down at room temperature for 15 min, followed by rinsing with cold dH_2O. Samples were then blocked in the Peroxidase Blocking solution (Dako REAL, S2023) for 8 min, rinsed with water and then blocked in a serum-free protein block buffer for 8 min (Dako, X0909). Primary CD133 antibody (Miltenyi Biotec, 130-090-422) was diluted 1:25 in Antibody Diluent with Background Reducing Components (Dako, S3022). 200 µl of antibody solution was put on each slide to cover all TMA specimens and incubated overnight at 4°C.

The following day, slides were washed twice with Dako Wash Buffer (S3006). Primary antibody detection was performed using MACH3 mouse probe (Biocare Medical) followed by MACH3 HRP polymer (Biocare Medical, BC-M3M530H) and the signal was developed with diamino-benzidine DAB+ (Dako, K3468). Finally, the slides were counterstained with hematoxylin (Dako, S3301), dehydrated in alcohol solutions and xylene, and mounted in Pertex Mount Agent (Histolab, 00801).

Transcriptomics

mRNA expression levels were extracted from previously reported microarray analyses (62). In brief, these analyses were performed on a Human HT-12-v4 BeadChip (Illumina) after labeling (Ambion; Aros Applied Biotechnology). Illumina BeadArray Reader (Illumina) and the Bead Scan Software (Illumina) were used to scan BeadChips. Expression signals from the beads were normalized and further processed as previously described (63). The data set was re-annotated using illuminaHumanv4.db from AnnotationDbi package, built under Bioconductor 3.3 in R (64), to select only probes with “perfect” annotation (65). The probes represented 21043 identified and unique genes.

Sequencing of 360 cancer related genes

Targeted sequencing of 360 cancer genes, was performed and described previously (18). In brief, native, genomic DNA from tumor, was fragmented and subjected to Illumina DNA sequencing

library preparation. Libraries were then hybridized to custom RNA baits according to the Agilent SureSelect protocol. Paired-end, 75bp sequence reads were generated. Sequencing coverage for the targeted regions (average per bp) within each sample was >120x for all samples (mean 439x). Supplemental Table S1 lists the included 360 genes.

Proliferation assay

Cell proliferation assay was determined by high-content imaging using the IncuCyte Zoom (Essen Bioscience) according to the manufacturer's instructions. In all experiments, cells were seeded into a 96-well culture plate and for each well four fields were imaged under 10x magnification every 2 h. The IncuCyte Zoom (v2018A) software was used to calculate confluency values.

Glucose and insulin measurements

For glucose and insulin measurements, mice were fasted overnight (9 hours) with free access to water. Blood glucose concentrations were determined using Accu-Check Aviva glucometer (Roche). For insulin measurements, blood was collected from the tail using EDTA coated capillary tubes (Fisher Scientific, 11383994), stored on ice before centrifuged at 2000 g, 4 °C for 10 min. Plasma insulin concentrations was determined in duplicates using the Ultra Sensitive Mouse Insulin ELISA Kit (Crystal Chem, 90080) following the manufactures instructions for wide range measures.

Glucose tolerance test

For glucose tolerance test, mice fed a HFD or chow-diet for 10 weeks were fasted overnight (15 hours) with free access to water. Glucose (2.5 g/kg) was administered by gavage, and blood glucose concentrations were determined by using Accu-Check Aviva glucometer (Roche).

Mammary Fat Pad Implantations

E0771 or TeLi cells were prepared in PBS and mixed 1:1 by volume with Matrigel (Corning, 356231) and orthotopically implanted into the 4th inguinal mammary fat pad of chow and HFD fed mice in a total volume of 50µL. Tumor diameters (width and length) were measured 2-3 times per

week with caliper. Tumor volumes were calculated using formula Tumor volume (mm^3) = Width x $Length^2$ x π/6. Tumors were considered established when the volumes were larger than $50mm^3$.

Cellular Adaptation to Palmitic Acid

Cells were seeded on 10 cm culture dishes so that the confluency at the starting day of selection was 80-90%. To start selection, all media was removed and replaced by growth media supplemented with 200 μM palmitic acid (PA) (Sigma, P5585). After cells acquired resistance to this concentration, the concentration of palmitic acid was increased to 400 μM. For E0771 and TeLi cell lines we finally increased concentration to 500 μM and 600 μM, respectively, due to high intrinsic PA resistance of these cells. For HCC1806 cells the concentration were reveresed to 200 μM due to the fragility of the cells and their inability to survive in 400 μM PA after the standard cryopreservation in FBS/10%DMSO. Parental cells were cultured in parallel using growth media supplemented with 1% fatty acid free BSA (Sigma, A7030). PA adapted cells were cultured in growth media supplemented with 1% fatty acid free BSA and indicated concentration of PA. For PA supplemented media, PA was first dissolved in absolute ethanol to obtain a 50mM stock. To prepare the working concentrations, certain volumes of PA stock were added into 1%BSA growth media and incubated at 37°C for 1hour. PA stock was stored at 4°C and used for no longer than 2 weeks.

Generation of knockdown and overexpressing cell lines

Short hairpin RNAs (shRNA) for target genes and scramble (shCtrl) were purchased from Sigma as bacterial glycerol stocks (#1864). pBabe-puro plasmids containing human C/EBPB LAP2 and LIP isoforms were from Addgene (Cat.# 15712 and 15713).

For production of virus, HEK293T cells were seeded onto 10 cm plates to reach 80% confluency on the following day. For retroviral overexpression, 12μg of Gag/Pol plasmid, 6μg of VSVG plasmid and 12 μg of pBabe-puro plasmid containing C/EBPB isoforms were respectively co-transfected into the HEK293T cells using 60 μL Lipofectamine 2000 according to manufacturer's protocol. For lentiviral-mediated depletion of target genes, cells were transfected with 12μg Gag/Pol plasmid, 6μg envelope plasmid and 12μg shRNA containing plasmid (pLKO).

6 hours following transfection, the media was replaced with fresh media. The virus was harvested 48hours post transfection by spinning the collected culture media for 5 mins at 1200 rpm and then

filtered through a 0.22 µm filter to completely remove cell debris. The virus was then stored at -20°C for several days or at -80°C for several months.

To infect target cells, 5mL of the appropriate virus was used to infect a subconfluent 10 cm cell culture dish in the presence of 10 µg/mL of polybrene overnight. 48 hours after infection, puromycin was added to select for successfully infected cells: 4 µg/mL for TeLi, 2 µg/mL for MDA-MB-231 and E0771 and 1.33 µg/mL for HCC1806 cells. Uninfected control cells were processed the same way to determine the endpoint of selection. Typically, selection took 2-3 days for all cell lines. After the end of selection cells were released from puromycin for at least 1 day before starting experiments.

Spheroid Formation Assay

Cells were harvested using Trypsin, resuspended in the corresponding medium, and seeded onto ultra-low attachment U-bottom plates (Corning® Costar® Ultra-Low Attachment Multiple Well Plate, CLS7007-24EA) at a concentration of 2000 cells/100 µL/well for HCC1806, E0771 and TeLi cells or 4000 cells/100 µL/well for MDA-MB-231 cells. After 5 days MDA-MB-231 (2 days for E0771, TeLimet and 1 day for HCC1806) spheroids were imaged using the Nikon TE2000 fluorescence microscope. Spheroids were considered formed when cells were tightly adhered to each other, hindering the recognition of individual cells and formed a round sphere structure with clear boundary. Spheroid formation was quantified independently and blinded by two investigators.

Apoptosis

Analysis of apoptosis was performed using Alexa Fluor™ 488 conjugate Annexin V (Thermo Fisher, A13201) and propidium iodide (PI) according to the manufacturer's instructions. Shortly, cells and their culture media were harvested and washed once in cold PBS. Cells were then resuspended in Annexin binding buffer (10 mM HEPES, 140 mM NaCl, and 2.5 mM $CaCl_2$, pH 7.4) in a concentration of 1×10^6 cells/mL. To each 100 µL of cell suspension 5 µL of the Annexin V and 2 µL PI (at final concentration 2 µg/mL) was added. Cells were incubated in the dark at room temperature for 15 min. After the incubation period, 400 µL of Annexin binding buffer was added and cells were analyzed by flow cytometry (BD LSR Fortessa).

Flow cytometry analysis

For immunostaining for flow cytometry, cells were collected using Accutase (Sigma, A6964) and washed once in PBS. 1×10^6 cells per sample were stained with 0.6 µl of APC conjugated CD133 antibodies (key resource table) in 100 µl of PBS +1%BSA solution and incubated in dark for 20 min at room temperature. After incubation, cells were washed once with 5 ml of PBS/1% BSA and analyzed on flow cytometry (BD LSR Fortessa).

Immunofluorescent analysis

Cells were seeded in 24-well plates on Poly-L-lysin treated cover slips at 75 000 cells per well one day before the staining. On the day of the analysis, culture media was removed and 4% paraformaldehyde (PFA) in Distilled-PBS (DPBS) was added to fix cells for 20 minutes. Then PFA was removed and cells were permeabilized in 0.4% Tween/DPBS for 10 min at RT. This was followed by 3 washes in DPBS. Blocking was performed in 3%BSA/0.2% Tween/DPBS for 90 min. Slides were shortly washed in staining media containing DPBS/0.2% Tween/1.5% BSA. Then slides were covered by 500 µl of staining media with C/EBPB antibodies (1:100 dilution) and incubated overnight at 4°C on rocking platform.

Next day, slides were washed in DPBS 3 x 5 min and incubated with secondary antibodies (1:500 dilution) for 2 hours. This was followed by 5 min wash in DPBS, then 5 min incubation with DAPI (1:500 in DPBS) and then another wash in DPBS.

Further, slides were rinsed in distilled water and mounted with ProLong™ Diamond Antifade Mountant. Slides were dried overnight and imaged using Leica SP5 with 63x magnification.

Image quantification was performed using Fiji software. The nucleus and whole cell were demarcated based on DAPI and bright field, respectively. % nuclear C/EBPB were calculated by diving the nuclear signal by whole cell signal multiplied by 100.

Fatty Acid and Glucose Oxidation Assay

Fatty acid and glucose oxidation were assessed by providing ^{14}C-labled palmitic acid or glucose to the cells, with subsequent capture of the released $^{14}CO_2$; a technique previously described (66). In brief, cells were plated into 96-well tissue culture plates (MDA-MB-231, 45000 cells/well; HCC1806, 45000 cells/well; dissociated E0771, 25000 cells/well) in corresponding growth

medium and incubated overnight to allow proper attachment. Radiolabeled [1-^{14}C]palmitic acid (1 µCi/ml) and D-[^{14}C(U)]glucose (1 µCi/ml) were given in PBS supplemented with 10mM HEPES and 1mM L-carnitine. Respective amounts of non-radiolabeled substrate were added to obtain final concentrations of D-glucose (5 mM) and BSA-conjugated palmitic acid (100 µM). Etomoxir (40 µM) was added to certain wells during palmitic acid oxidation, to monitor the non-mitochondrial CO_2 production. An UniFilter®-96w GF/B microplate was activated for capture of CO_2 by the addition of 1M NaOH (25 µL/well) and sealed to the top of the 96-well tissue culture plates and incubated for the indicated period of time at 37 °C. Subsequently, 30 µL scintillation liquid (MicroScint PS PerkinElmer) was added to the filters and the filter plate was sealed with a TopSealA (PerkinElmer). Radioactivity was measured using MicroBeta2 Microplate Counter (PerkinElmer). Protein measurement was performed for data normalization. The cells were washed twice with PBS, lysed by 0.1 M NaOH, and protein was measured using Pierce® BCA Protein Assay Kit.

Staining of Lipid Droplets

Cells were seeded on tissue culture plates 24 hours before the analysis. BODIPY (Thermo Fisher, D3922) was diluted in DPBS (Gibco, 14040-133) to a final concentration of 2 µM. Cells were washed in DBPS once and stained with BODIPY/DPBS mix for 15 min at 37°C. Further, cells were harvested from tissue culture plates, washed once in PBS and analysed on Flow cytometer BD LSR Fortessa. Data analysis was performed using FlowJo software.

RNA extraction, RT-PCR, and qPCR

Total RNA was extracted with a Total RNA purification Kit (NORGEN Biotek, 37500) according to the manufacturer's protocol. cDNA was synthesized from 1 µg total RNA template with oligo-dT primers using a SuperScript® III First-Strand Synthesis kit (ThermoFisher Scientific, 18080-051) according to the manufacturer's protocol. qPCR was carried out in quadruplicates with a LightCycler® 480 SYBR Green I Master Mix (Roche, 04887352001) using a LightCycler® 480 Instrument II (Roche, 05015243001). The results were calculated by ΔΔCt method using human HPRT (hHPRT) for human genes and mouse actin (mActin) for mouse genes. Primer sequences are listed in materials sources table (Supplementary Table S2).

Transfection of siRNA duplexes

One day before transfection cells were plated on T25 flasks at the density 250 000 cells/flask. After overnight incubation, cells were transfected using Lipofectamine® RNAiMAX according to the manufacturer's protocol with some modifications: we used 3 µl of Lipofectamine per flask and final concentration of siRNAs was 20 nM. After 48 hours incubation cells were harvested and seeded for the analysis on ultra-low attachment U-bottom plates. After seeding, cells were infected for the second time using 0,075µl of Lipofectamine RNAiMAX per well with 20 nM siRNA.

Western blotting

Cells were lysed in RIPA lysis buffer (Thermo Scientific, 89901) complemented with protease inhibitor cocktail (cOmplete ULTRA Tablets, MINI, EDTA-free, EASYpack, 05892 791001) and phosphatase inhibitor cocktail (PhosStop, 04 906 837 001). After quantification with a BCA protein assay kit (Pierce, 23225), equal amounts of protein (typically 20-50 µg of protein per lane) were separated by electrophoresis on a NuPAGE 10% Bis-Tris Gel (Invitrogen, NP0315BOX) in NuPAGE™ MOPS SDS Running Buffer (20X, Invitrogen, NP000102) and then transferred to an activated Immobilon-P PVDF Membrane (Merck Millipore Ltd, IPVH00010 PORE SIZE: 0,45 µm). The membranes were blocked using 5% nonfat dry milk in PBS/0.1% Tween20 for 1h at RT, incubated with indicated primary antibodies for overnight at 4°C. This step was followed by an incubation with secondary IRDye-conjugated antibodies (Leicor, P/N 925-68070, P/N 926-32213). Detection and quantification were performed on Amersham Typhoon Gel and Blot Imaging Systems. A list of antibodies is given in the key resources table.

RNA sequencing

MDA-MB-231 parental and selected cells were plated at 1×10^5 cells/mL into 6-well plates in corresponding medium. After three days, cells were harvested, and RNA extraction was performed according to the manufacturer's protocol. Potential DNA contaminations were removed by applying the RNA Clean & concentrator with DnaseI kit (Zymo, R1013). RNA sequencing libraries were prepared at the Genomic Core Facility at University of Bergen using Illumina TruSeq Stranded mRNA sample preparation kit according to the manufacturer's instructions and sequenced on the same lane on a HiSeq 4000 sequencer with pair-end 75bp reads.

ATACseq library construction

ATACseq libraries were constructed as previously described (67). In brief, $5x10^4$ cells were washed once with ice-cold PBS and pelleted by centrifugation. Cells were lysed in 50 µl RSB buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl and 3 mM $MgCl_2$) containing 0.1% NP-40, 0.1% Tween-20 and 0.01% digitonin, and incubated on ice for 3 minutes for permeabilization. After incubation, samples were washed in 1 mL RSB containing 0.1% Tween-20 and pelleted at 500 g for 10 minutes at 4°C. Samples were then resuspended on ice in 50 µl transposition reaction mix containing 2.5 µl Tn5 transposase, 1x TD buffer (both Illumina FC-121-1030), 1x PBS, 0.1% Tween-20 and 0.01% digitonin, and incubated at 37°C for 30 minutes with agitation. Tagmented DNA was purified using Zymo DNA Clean and Concentrator-5 kit (Zymo D4014). The resulting DNA was amplified for 12-13 cycles. The libraries were purified with AMPure XP beads (Beckman A63880), quality-checked on Bioanalyzer (Agilent) and 75 bp paired-end sequenced on Illumina Hiseq 4000 at Genomic Core Facility at University of Bergen.

Cut&Run and library construction

Cut&Run was performed as described with minor modifications (37). Briefly, $5x10^5$ cells were washed and bound to concanavalin A-coated magnetic beads (Bangs Laboratories, BP531). The cells were then permeabilized with Wash Buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM spermidine and 1x Roche Complete Protease Inhibitor, EDTA-free) containing 0.025% digitonin (Digitonin Buffer) and 2 mM EDTA and incubated with primary antibody (anti-C/EBPB or IgG isotype control) overnight at 4°C. The cell-bead slurry was washed twice with Digitonin Buffer and incubated with 1x Protein-A/G-MNase (pAG-MNase; Epicypher) in Digitonin Buffer for 10 minutes at room temperature. The slurry was washed twice with Digitonin Buffer and incubated in Digitonin Buffer containing 2 mM $CaCl_2$ for 2 hours at 4°C to activate pAG-MNase digestion. The digestion was stopped by addition of 2x Stop Buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 50 µg/mL RNase A, 50 µg/mL GlycoBlue and 300 pg/mL in-house MNase-digested yeast spike-in chromatin) and the sample was incubated for 10 minutes at 37°C to release chromatin to the supernatant and degrade RNA. The supernatant was recovered, and DNA was isolated through phenol-chloroform extraction and ethanol precipitation. Libraries were constructed to enrich for sub-nucleosomal fragments using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina as

described (NEB, E7645S). The libraries were size-selected and purified with AMPure XP beads, quality-checked on Tapestation (Agilent) and 100 bp paired-end sequenced on MiSeq at Genomic Core Facility at University of Bergen.

Mass Cytometry

Cells were plated in 10cm plates in triplicates to reach a confluency of 80% after 48 hours. For the analysis, cells were collected using TrypLE Express (Gibco 12604-021). 1 x 10^6 cells per condition were included. Cells were resuspended in cell culture media and treated with 0.25 µM Cisplatin for 5 min at RT. Further, cells were fixed in 1 mL of 1.6 % PFA in PBS for 10 min at RT. Cells were pelleted by centrifugation for 5 min at 900g and the pellets were stored at -80°C until staining with CyTOF antibodies. On the day of staining, samples were thawed on ice, resuspended in 500 µl of DPBS (Gibco, 14040-133) and incubated for 10 min at RT in DNases (Sigma, DN25)/DPBS solution. Further, cells were washed in D-WASH solution (DPBS + 1% FA-free BSA + 0,02% NaN3 + DNase) and barcoded (Fluidigm, 201060) according to the manufacturer's protocol. Cells were then washed twice in the Cell Staining Buffer (Fluidigm, 201068), all samples were combined and labeled with surface antibody cocktail (Figure S2A, extracellular) for 30 min at RT. Further, cells were pelleted by centrifugation and incubated in 4 mL of DPBS/DNase solution for 10 min at RT. After this step cells were washed in PBS-EDTA and fixed in 2% PFA/PBS (filtered through a 0.22 µm filter) for 30 min RT, followed by wash in Cell Staining Buffer and permeabilization in cold methanol (-20°C) for 10 min. After incubation, cells were washed in once PBS, once in D-WASH and labeled with intracellular antibody cocktail (Figure S2A, intracellular) for 30 min at RT. This was followed by incubation of cells in D-WASH for 10 min at RT and double wash in D-WASH. Then cells were incubated in 2% PFA/PBS with iridium cell tracker at 4°C overnight. The samples were spun down the following day and incubated in D-WASH for 10 min at RT, washed once in PBS/EDTA once, 3 times in di water (Fluidigm, 201069), resuspended in EQ beads (Fluidigm, 201078) diluted 1:9 in water, and analyzed on Helios - Mass Cytometer.

QUANTIFICATION AND STATISTICAL ANALYSIS

CyTOF data development using X-shift

Raw FCS-files were normalized, concatenated and debarcoded in R using Cytometry dATa anALYSis Tools; CATALYST (68). X-shift algorithm from VorteX clustering and visualization environment (25) was applied on CyTOF data of parental or adapted MDA-MB-231 and HCC1806 cell lines. To enable comparisons, cells from the different conditions were downsampled (85 000 and 100 000 per sample for HCC and MDA respectively) and combined for each cell line prior to clustering. All markers were used except Keratin 7 and the cell cycle marker p-HisH3. The following parameters were used in Vortex: numerical transformation: arcsinh (x/f), f = 5.0, noise threshold = 1.0, distance measure: angular distance, clustering algorithm: X-shift (gradient assignment), density estimate: N nearest neighbors, number of neighbors for mode finding (N): determine automatically. The resulting clusters yielded elbow points of k= 27 and k= 35 for HCC and MDA cell lines respectively. For the visualization, a maximum of 20 000 events extracted from each cluster were plotted in a force-directed graph layout using VorteX and Gephi Toolkit 0.8.7 (https://gephi.org/toolkit/). For the analysis of the two more representative clusters in each cell line, the log2 fold change has been calculated from the X-shift scores converted to the normalized raw intensities values.

Survival Analysis

Patients were stratified into two groups by BMI 25. Disease-Specific survival (DSS) Kaplan-Meier curves were generated using GraphPad Prism software and statistical significance was calculated using Log-rank (Mantel-Cox) test.

The combined effect of LCN2 and SERPINB2 to patient survival was analyzed on PM/ER-/PR- patients from GSE25066. Patients were stratified by median of average normalized expression of LCN2 and SERPINB2. Kaplan-Meier curves and statistics were performed in the same way.

Mutual Information

Mutual information was calculated as described in Goodarzi 2009 (23)

Tissue Microarray Analysis

The ten CD133-stained TMA slides were scanned with an Aperio Scanscope CS Slide Scanner. The breast cancer cores were 1.2 mm in diameter with up to 4 cores per patient. Full analysis

was performed on valid cores for patients 50 years and older with ER and PR negative status. Cores with too few cells, poor quality, excessive tearing, or folding were not considered valid and were omitted from analysis.

QuPath Version: 0.2.0-m5 was used to dearray the TMAs, segment cells, and classify cell types. The following detection steps and parameters were applied to all TMA slides. Simple tissue detection was used to find the approximate tissue borders within each dearrayed TMA core. A threshold of 229 (default 127), requested pixel size of 1 μm (default 20 μm), and checking the box for Expand boundaries were found to be the most important parameter setting changes for accurate tissue detection.

Watershed cell detection was used to create cell masks within the detected tissue of each valid core. The watershed parameters were optimized to detect large weakly hematoxylin stained cancer cells, to minimize false positive cell detection from areas of high background signal, and to reduce the creation of cell masks that spanned multiple cells. The watershed parameter changes deemed most important for accurate cell mask creation were: nucleus background radius of 10 μm (default 8 μm), nucleus minimum area of 24 μm^2 (default 10 μm^2,), nucleus maximum area of 230 μm^2 (default 400 μm^2), intensity parameters for threshold and max background both set to 0.07, and exclusion of DAB staining (as was recommended for membrane staining markers). Additionally, the cell expansion was set to 10 μm, 5 μm larger than the default setting, in order to capture the CD133 membrane staining on the large cancer cells.

Annotation objects were drawn around easily defined areas that contained primarily cancer cells, non-cancer cells, or platelets/RBCs and labeled as the classes tumor, stroma, or ignore, respectively. Platelets/RBCs were ignored because they appeared brown even before staining and show up as falsely positive for CD133. 9039 cells from the annotation objects drawn across 5 of the 10 slides were used to train the random forest (trees) classifier in QuPath. DAB specific measurements were excluded from the classifier selected features. The intensity feature used to identify CD133 positive cells was Cell: DAB OD max at a threshold of 0.45. With these parameters, the detection classifier created 7 classification groups of cells: total (base) tumor cells, total stroma cells, $CD133^+$ tumor cells, $CD133^-$ tumor cells, $CD133^+$ stroma cells, $CD133^-$ stroma cells, and ignored cells. Cell masks from cores with partial low quality due to folding or poor imaging were removed to prevent false positive cells. All cores were visually inspected for false positive cancer cell masks and false positive masks were removed. Mean $CD133^+$ cancer cell percentage was calculated for each patient for all valid tumor cores by QuPath and exported

to MS Excel. Patients with greater than 2% CD133 positive cancer cells were considered to have CD133 positive tumors. Statistical analysis was performed in GraphPad Prism.

Student's t-test

Statistical analysis of flow cytomentry data was performed using student's t-test on GraphPad Prism 8 software. 3 replicates per condition were performed and the experiment was performed 3 times.

Fisher's exact test

Statistical analysis of spheroids formation experiments was performed using Fisher's exact test.

Limiting dilution analysis

The frequency of tumor initiating cells was calculated using the Extreme Limiting Dilution Analysis (ELDA) (http://bioinf.wehi.edu.au/software/elda/index.html) (69)

RNA sequencing data analysis

Sequenced reads were quality checked with FastQC and aligned to the UCSC hg19 reference genome with Hisat2. Aligned reads were counted and summarized for the annotated genes using featureCounts. Differential gene expression analysis was performed by DESeq2. For visualization, normalized read counts were regularized log transformed (rlog).

ATACseq data analysis

ATACseq reads were quality-checked with FastQC (70) before and after adapter trimming with Trimmomatic (71). The trimmed reads were aligned to the UCSC hg19 reference genome using Bowtie2 (72) with the parameters --phred33 --end-to-end --very-sensitive -X 2000. Reads were then removed if they were mapped to the mitochondria and non-assembled contigs, had a mapping quality score below 10 and were PCR duplicates. Read start sites were adjusted for Tn5

insertion by offsetting +stand by +4 bp and -strand by -5 bp as previously described (67). For peak calling, MACS2 (73) was used with the parameters -q 0.01 --nomodel. Peaks residing in the ENCODE blacklisted regions were removed for further downstream analysis. deepTools (74) was used to generate 1x normalized bigwig files for visualization.

Analysis of differential accessible peaks was performed using DiffBind (75) with default settings, and annotated genome-wide with respect to the closest transcription start site with ChIPseeker (76). Peaks with a mean peak count ≥ 10 were kept for further analaysis. Annotated differential peaks were checked for phastCons conservation scores (77) for placental mammals against random noncoding background regions generated using bedtools shuffle. To infer differential transcription factor binding motif activity, diffTF (34) was used. Input transcription factor binding sites for 640 human transcription factors were generated as described using the HOCOMOCO database and PWMscan (cutoff p-value - 0.00001, background base composition - 0.29;0.21;0.21;0.29).

Cut&Run data analysis

Cut&Run reads were quality-checked with FastQC before and after adapter trimming with Trimmomatic. The trimmed reads were separately aligned to the UCSC hg19 and sacCer3 reference genomes using Bowtie2 with the parameters --local --very-sensitive-local --no-unal --no-mixed --no-discordant --phred33 -I 10 -X 700 and --local --very-sensitive-local --no-unal --no-mixed --no-discordant --phred33 -I 10 -X 700 --no-overlap --no-dovetail, respectively. Reads were then removed if they were mapped to the mitochondria and non-assembled contigs and had a mapping quality score below 10. Mapped reads were converted to paired-end BED files containing coordinates for the termini of each read pair and the fragment length, and calibrated to the yeast spike-in using spike_in_calibration.csh (https://github.com/Henikoff/Cut-and-Run/) in bedgraph formats for visualization. Peaks were called with SEACR (78) with respect to the IgG control using the norm and stringent mode. Peaks overlapping with the ENCODE blacklisted regions were removed for further downstream analysis. Consensus peaksets across samples were generated using DiffBind. Raw counts of the consensus peaksets across samples were input to DESeq2 with the inverse of the spike-in calibration factors as sizeFactors to perform differential analysis. Differential peaks were annotated respect to the closest transcription start site with ChIPseeker.

For motif discovery within peaks, EChO (79) was run to identify direct binding sites in the foci mode. Sites with the mean fragment length ≤ 120 bp were retained, extended to a 100-bp window and converted to BED files. HOMER (80) was used for motif enrichment analysis using the position weight matrices (PWMs) from the HOCOMOCO database.

To identify enriched motif sequences protected by transcription factor binding independent of the peak calling algorithm, pA/G-MNase cutting footprints were detected. Ends of all CUT&RUN fragments ≤ 120 bp were enumerated to determine the precise single base pair cut sites and sorted. Footprints were detected using Footprint Occupancy Score (FOS) (38). Significant footprints with FOS ≤ 1 were analyzed for enriched motif sequences with HOMER.

Acknowledgements

We thank Erik Løkkevik, Bjørn Østenstad, Steinar Lundgren, Terje Risberg, and Ingvil Mjaaland for providing clinical samples. We thank the genomic score facility (GSF) at the University of Bergen, which is a part of the NorSeq consortium, provided services on RNAseq, ATACseq and Cut&Run. GSF is supported by grants from the Research Council of Norway (245979/F50) and the Trond Mohn Foundation (BFS2016-genom). The flow cytometry and mass cytometry were performed at the Flow Cytometry Core Facility, Department of Clinical Science, University of Bergen. Helios Mass Cytometer was supported by the Trond Mohn Foundation. We thank Ingeborg Winge from Department of Pathology, Haukeland University Hospital, Bergen, for provided training and help with TMA Immunohistochemistry. Hani Goodarzi is supported by R01CA240984 and R01GM123977. NH was funded by a Starter Grant from the Trond Mohn Foundation and the Norwegian Research Council (275250).

Authors Contributions

Conceptualization, N.H; Methodology, N.H., X.L., A.R., S.G.M. and L.P.; Software, S.G.M., C.E.W., X.L. Validation: T.L., A.R., Formal Analysis, C.E.W., N.M., P.E.L., S.K., H.G., S.D.P., X.L., A.R., M.H.C.; Investigation, A.R., X.L.; Resources, N.H., S.D.P., S.M.G., J.L., S.K., P.E.L., S.K., A.M.; Writing – Original Draft, N.H., X.L., A.R.; Visualization, N.H., A.R., X.L., M.H.C., C.E.W.; Supervision, N.H.; Funding Acquisition, N.H.

Declaration of interests

The authors declare no competing interests.

CONTACT FOR REAGENT FOR AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Nils Halberg (nils.halberg@uib.no). This study did not generate new unique reagents.

REFERENCES

1. Munsell MF, Sprague BL, Berry DA, Chisholm G, Trentham-Dietz A. Body mass index and breast cancer risk according to postmenopausal estrogen-progestin use and hormone receptor status. Epidemiol Rev **2014**;36:114-36 doi 10.1093/epirev/mxt010.
2. Jiralerspong S, Goodwin PJ. Obesity and Breast Cancer Prognosis: Evidence, Challenges, and Opportunities. J Clin Oncol **2016**;34(35):4203-16 doi 10.1200/JCO.2016.68.4480.
3. Chan DS, Vieira AR, Aune D, Bandera EV, Greenwood DC, McTiernan A*, et al.* Body mass index and survival in women with breast cancer-systematic literature review and meta-analysis of 82 follow-up studies. Ann Oncol **2014**;25(10):1901-14 doi 10.1093/annonc/mdu042.
4. Niraula S, Ocana A, Ennis M, Goodwin PJ. Body size and breast cancer prognosis in relation to hormone receptor and menopausal status: a meta-analysis. Breast Cancer Res Treat **2012**;134(2):769-81 doi 10.1007/s10549-012-2073-x.
5. Key TJ, Appleby PN, Reeves GK, Roddam A, Dorgan JF, Longcope C*, et al.* Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. J Natl Cancer Inst **2003**;95(16):1218-26 doi 10.1093/jnci/djg022.
6. Key T, Appleby P, Barnes I, Reeves G, Endogenous H, Breast Cancer Collaborative G. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst **2002**;94(8):606-16 doi 10.1093/jnci/94.8.606.
7. Tiwari P, Blank A, Cui C, Schoenfelt KQ, Zhou G, Xu Y*, et al.* Metabolically activated adipose tissue macrophages link obesity to triple-negative breast cancer. J Exp Med **2019**;216(6):1345-58 doi 10.1084/jem.20181616.
8. Incio J, Tam J, Rahbari NN, Suboj P, McManus DT, Chin SM*, et al.* PlGF/VEGFR-1 Signaling Promotes Macrophage Polarization and Accelerated Tumor Progression in Obesity. Clin Cancer Res **2016**;22(12):2993-3004 doi 10.1158/1078-0432.CCR-15-1839.
9. Saxena NK, Taliaferro-Smith L, Knight BB, Merlin D, Anania FA, O'Regan RM*, et al.* Bidirectional crosstalk between leptin and insulin-like growth factor-I signaling promotes invasion and migration of breast cancer cells via transactivation of epidermal growth factor receptor. Cancer Res **2008**;68(23):9712-22 doi 10.1158/0008-5472.CAN-08-1952.
10. Garofalo C, Koda M, Cascio S, Sulkowska M, Kanczuga-Koda L, Golaszewska J*, et al.* Increased expression of leptin and the leptin receptor as a marker of breast cancer

progression: possible role of obesity-related stimuli. Clin Cancer Res **2006**;12(5):1447-53 doi 10.1158/1078-0432.CCR-05-1913.

11. Qureshi R, Picon-Ruiz M, Aurrekoetxea-Rodriguez I, Nunes de Paiva V, D'Amico M, Yoon H*, et al.* The Major Pre- and Postmenopausal Estrogens Play Opposing Roles in Obesity-Driven Mammary Inflammation and Breast Cancer Development. Cell Metab **2020**;31(6):1154-72 e9 doi 10.1016/j.cmet.2020.05.008.
12. Madak-Erdogan Z, Band S, Zhao YC, Smith BP, Kulkoyluoglu-Cotul E, Zuo Q*, et al.* Free Fatty Acids Rewire Cancer Metabolism in Obesity-Associated Breast Cancer via Estrogen Receptor and mTOR Signaling. Cancer Res **2019**;79(10):2494-510 doi 10.1158/0008-5472.CAN-18-2849.
13. Hao J, Zhang Y, Yan X, Yan F, Sun Y, Zeng J*, et al.* Circulating Adipose Fatty Acid Binding Protein Is a New Link Underlying Obesity-Associated Breast/Mammary Tumor Development. Cell Metab **2018**;28(5):689-705 e5 doi 10.1016/j.cmet.2018.07.006.
14. Chrisanthar R, Knappskog S, Lokkevik E, Anker G, Ostenstad B, Lundgren S*, et al.* Predictive and prognostic impact of TP53 mutations and MDM2 promoter genotype in primary breast cancer patients treated with epirubicin or paclitaxel. PLoS One **2011**;6(4):e19249 doi 10.1371/journal.pone.0019249.
15. Furukawa S, Fujita T, Shimabukuro M, Iwaki M, Yamada Y, Nakajima Y*, et al.* Increased oxidative stress in obesity and its impact on metabolic syndrome. J Clin Invest **2004**;114(12):1752-61 doi 10.1172/JCI21625.
16. Urakawa H, Katsuki A, Sumida Y, Gabazza EC, Murashima S, Morioka K*, et al.* Oxidative stress is associated with adiposity and insulin resistance in men. J Clin Endocrinol Metab **2003**;88(10):4673-6 doi 10.1210/jc.2003-030202.
17. Sabharwal SS, Schumacker PT. Mitochondrial ROS in cancer: initiators, amplifiers or an Achilles' heel? Nat Rev Cancer **2014**;14(11):709-21 doi 10.1038/nrc3803.
18. Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P*, et al.* Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat Med **2015**;21(7):751-9 doi 10.1038/nm.3886.
19. Karpe F, Dickmann JR, Frayn KN. Fatty acids, obesity, and insulin resistance: time for a reevaluation. Diabetes **2011**;60(10):2441-9 doi 10.2337/db11-0425.
20. Tremblay AJ, Despres JP, Piche ME, Nadeau A, Bergeron J, Almeras N*, et al.* Associations between the fatty acid content of triglyceride, visceral adipose tissue accumulation, and components of the insulin resistance syndrome. Metabolism **2004**;53(3):310-7 doi 10.1016/j.metabol.2003.10.011.
21. Korbecki J, Bajdak-Rusinek K. The effect of palmitic acid on inflammatory response in macrophages: an overview of molecular mechanisms. Inflamm Res **2019**;68(11):915-32 doi 10.1007/s00011-019-01273-5.
22. Saadatian-Elahi M, Toniolo P, Ferrari P, Goudable J, Akhmedkhanov A, Zeleniuch-Jacquotte A*, et al.* Serum fatty acids and risk of breast cancer in a nested case-control study of the New York University Women's Health Study. Cancer Epidemiol Biomarkers Prev **2002**;11(11):1353-60.
23. Goodarzi H, Elemento O, Tavazoie S. Revealing global regulatory perturbations across human cancers. Mol Cell **2009**;36(5):900-11 doi 10.1016/j.molcel.2009.11.016.
24. Huang S. Genetic and non-genetic instability in tumor progression: link between the fitness landscape and the epigenetic landscape of cancer cells. Cancer Metastasis Rev **2013**;32(3-4):423-48 doi 10.1007/s10555-013-9435-7.
25. Samusik N, Good Z, Spitzer MH, Davis KL, Nolan GP. Automated mapping of phenotype space with single-cell data. Nat Methods **2016**;13(6):493-6 doi 10.1038/nmeth.3863.
26. Draffin JE, McFarlane S, Hill A, Johnston PG, Waugh DJ. CD44 potentiates the adherence of metastatic prostate and breast cancer cells to bone marrow endothelial cells. Cancer Res **2004**;64(16):5702-11 doi 10.1158/0008-5472.CAN-04-0389.

27. Sheridan C, Kishimoto H, Fuchs RK, Mehrotra S, Bhat-Nakshatri P, Turner CH*, et al.* CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. Breast Cancer Res **2006**;8(5):R59 doi 10.1186/bcr1610.
28. Immervoll H, Hoem D, Sakariassen PO, Steffensen OJ, Molven A. Expression of the "stem cell marker" CD133 in pancreas and pancreatic ductal adenocarcinomas. BMC Cancer **2008**;8:48 doi 10.1186/1471-2407-8-48.
29. Asiedu MK, Beauchamp-Perez FD, Ingle JN, Behrens MD, Radisky DC, Knutson KL. AXL induces epithelial-to-mesenchymal transition and regulates the function of breast cancer stem cells. Oncogene **2014**;33(10):1316-24 doi 10.1038/onc.2013.57.
30. Bankhead P, Loughrey MB, Fernandez JA, Dombrowski Y, McArt DG, Dunne PD*, et al.* QuPath: Open source software for digital pathology image analysis. Sci Rep **2017**;7(1):16878 doi 10.1038/s41598-017-17204-5.
31. Dupuy F, Tabaries S, Andrzejewski S, Dong Z, Blagih J, Annis MG*, et al.* PDK1-Dependent Metabolic Reprogramming Dictates Metastatic Potential in Breast Cancer. Cell Metab **2015**;22(4):577-89 doi 10.1016/j.cmet.2015.08.007.
32. O'Flanagan CH, Rossi EL, McDonell SB, Chen X, Tsai YH, Parker JS*, et al.* Metabolic reprogramming underlies metastatic potential in an obesity-responsive murine model of metastatic triple negative breast cancer. NPJ Breast Cancer **2017**;3:26 doi 10.1038/s41523-017-0027-5.
33. Ly CH, Lynch GS, Ryall JG. A Metabolic Roadmap for Somatic Stem Cell Fate. Cell Metab **2020**;31(6):1052-67 doi 10.1016/j.cmet.2020.04.022.
34. Berest I, Arnold C, Reyes-Palomares A, Palla G, Rasmussen KD, Giles H*, et al.* Quantification of Differential Transcription Factor Activity and Multiomics-Based Classification into Activators and Repressors: diffTF. Cell Rep **2019**;29(10):3147-59 e12 doi 10.1016/j.celrep.2019.10.106.
35. Descombes P, Schibler U. A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. Cell **1991**;67(3):569-79 doi 10.1016/0092-8674(91)90531-3.
36. Xiong W, Hsieh CC, Kurtz AJ, Rabek JP, Papaconstantinou J. Regulation of CCAAT/enhancer-binding protein-beta isoform synthesis by alternative translational initiation at multiple AUG start sites. Nucleic Acids Res **2001**;29(14):3087-98 doi 10.1093/nar/29.14.3087.
37. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife **2017**;6 doi 10.7554/eLife.21856.
38. Neph S, Vierstra J, Stergachis AB, Reynolds AP, Haugen E, Vernot B*, et al.* An expansive human regulatory lexicon encoded in transcription factor footprints. Nature **2012**;489(7414):83-90 doi 10.1038/nature11212.
39. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T*, et al.* GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database (Oxford) **2017**;2017 doi 10.1093/database/bax028.
40. Neuhouser ML, Aragaki AK, Prentice RL, Manson JE, Chlebowski R, Carty CL*, et al.* Overweight, Obesity, and Postmenopausal Invasive Breast Cancer Risk: A Secondary Analysis of the Women's Health Initiative Randomized Clinical Trials. JAMA Oncol **2015**;1(5):611-21 doi 10.1001/jamaoncol.2015.1546.
41. Hillers-Ziemer LE, McMahon RQ, Hietpas M, Paderta G, LeBeau J, McCready J*, et al.* Obesity Promotes Cooperation of Cancer Stem-Like Cells and Macrophages to Enhance Mammary Tumor Angiogenesis. Cancers (Basel) **2020**;12(2) doi 10.3390/cancers12020502.
42. Teslow EA, Mitrea C, Bao B, Mohammad RM, Polin LA, Dyson G*, et al.* Obesity-induced MBD2_v2 expression promotes tumor-initiating triple-negative breast cancer stem cells. Mol Oncol **2019**;13(4):894-908 doi 10.1002/1878-0261.12444.

43. Guo L, Cheng X, Chen H, Chen C, Xie S, Zhao M*, et al.* Induction of breast cancer stem cells by M1 macrophages through Lin-28B-let-7-HMGA2 axis. Cancer Lett **2019**;452:213-25 doi 10.1016/j.canlet.2019.03.032.
44. LaMarca HL, Visbal AP, Creighton CJ, Liu H, Zhang Y, Behbod F*, et al.* CCAAT/enhancer binding protein beta regulates stem cell activity and specifies luminal cell fate in the mammary gland. Stem Cells **2010**;28(3):535-44 doi 10.1002/stem.297.
45. Bundy LM, Sealy L. CCAAT/enhancer binding protein beta (C/EBPbeta)-2 transforms normal mammary epithelial cells and induces epithelial to mesenchymal transition in culture. Oncogene **2003**;22(6):869-83 doi 10.1038/sj.onc.1206216.
46. Nallanthighal S, Heiserman JP, Cheon DJ. The Role of the Extracellular Matrix in Cancer Stemness. Front Cell Dev Biol **2019**;7:86 doi 10.3389/fcell.2019.00086.
47. Watt FM, Huck WT. Role of the extracellular matrix in regulating stem cell fate. Nat Rev Mol Cell Biol **2013**;14(8):467-73 doi 10.1038/nrm3620.
48. Seo BR, Bhardwaj P, Choi S, Gonzalez J, Andresen Eguiluz RC, Wang K*, et al.* Obesity-dependent changes in interstitial ECM mechanics promote breast tumorigenesis. Sci Transl Med **2015**;7(301):301ra130 doi 10.1126/scitranslmed.3010467.
49. Valiente M, Obenauf AC, Jin X, Chen Q, Zhang XH, Lee DJ*, et al.* Serpins promote cancer cell survival and vascular co-option in brain metastasis. Cell **2014**;156(5):1002-16 doi 10.1016/j.cell.2014.01.040.
50. Lee NH, Park SR, Lee JW, Lim S, Lee SH, Nam S*, et al.* SERPINB2 Is a Novel Indicator of Cancer Stem Cell Tumorigenicity in Multiple Cancer Types. Cancers (Basel) **2019**;11(4) doi 10.3390/cancers11040499.
51. Catalan V, Gomez-Ambrosi J, Rodriguez A, Ramirez B, Silva C, Rotellar F*, et al.* Increased adipose tissue expression of lipocalin-2 in obesity is related to inflammation and matrix metalloproteinase-2 and metalloproteinase-9 activities in humans. J Mol Med (Berl) **2009**;87(8):803-13 doi 10.1007/s00109-009-0486-8.
52. Gomez-Chou SB, Swidnicka-Siergiejko AK, Badi N, Chavez-Tomar M, Lesinski GB, Bekaii-Saab T*, et al.* Lipocalin-2 Promotes Pancreatic Ductal Adenocarcinoma by Regulating Inflammation in the Tumor Microenvironment. Cancer Res **2017**;77(10):2647-60 doi 10.1158/0008-5472.CAN-16-1986.
53. Yang J, Bielenberg DR, Rodig SJ, Doiron R, Clifton MC, Kung AL*, et al.* Lipocalin 2 promotes breast cancer progression. Proc Natl Acad Sci U S A **2009**;106(10):3913-8 doi 10.1073/pnas.0810617106.
54. Menendez JA, Lupu R. Fatty acid synthase and the lipogenic phenotype in cancer pathogenesis. Nat Rev Cancer **2007**;7(10):763-77 doi 10.1038/nrc2222.
55. DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB. The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. Cell Metab **2008**;7(1):11-20 doi 10.1016/j.cmet.2007.10.002.
56. Nieman KM, Kenny HA, Penicka CV, Ladanyi A, Buell-Gutbrod R, Zillhardt MR*, et al.* Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth. Nat Med **2011**;17(11):1498-503 doi 10.1038/nm.2492.
57. Pascual G, Avgustinova A, Mejetta S, Martin M, Castellanos A, Attolini CS*, et al.* Targeting metastasis-initiating cells through the fatty acid receptor CD36. Nature **2017**;541(7635):41-5 doi 10.1038/nature20791.
58. Qin Y, Grimm SA, Roberts JD, Chrysovergis K, Wade PA. Alterations in promoter interaction landscape and transcriptional network underlying metabolic adaptation to diet. Nat Commun **2020**;11(1):962 doi 10.1038/s41467-020-14796-x.
59. Reid MA, Dai Z, Locasale JW. The impact of cellular metabolism on chromatin dynamics and epigenetics. Nat Cell Biol **2017**;19(11):1298-306 doi 10.1038/ncb3629.

60. Li X, Egervari G, Wang Y, Berger SL, Lu Z. Regulation of chromatin and gene expression by metabolic enzymes and metabolites. Nat Rev Mol Cell Biol **2018**;19(9):563-78 doi 10.1038/s41580-018-0029-7.
61. McDonnell E, Crown SB, Fox DB, Kitir B, Ilkayeva OR, Olsen CA*, et al.* Lipids Reprogram Metabolism to Become a Major Carbon Source for Histone Acetylation. Cell Rep **2016**;17(6):1463-72 doi 10.1016/j.celrep.2016.10.012.
62. Poduval D, Sichmanova Z, Straume AH, Lonning PE, Knappskog S. The novel microRNAs hsa-miR-nov7 and hsa-miR-nov3 are over-expressed in locally advanced breast cancer. PLoS One **2020**;15(4):e0225357 doi 10.1371/journal.pone.0225357.
63. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ*, et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature **2012**;486(7403):346-52 doi 10.1038/nature10983.
64. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res **2015**;43(7):e47 doi 10.1093/nar/gkv007.
65. Barbosa-Morais NL, Dunning MJ, Samarajiwa SA, Darot JF, Ritchie ME, Lynch AG*, et al.* A re-annotation pipeline for Illumina BeadArrays: improving the interpretation of gene expression data. Nucleic Acids Res **2010**;38(3):e17 doi 10.1093/nar/gkp942.
66. Wensaas AJ, Rustan AC, Lovstedt K, Kull B, Wikstrom S, Drevon CA*, et al.* Cell-based multiwell assays for the detection of substrate accumulation and oxidation. J Lipid Res **2007**;48(4):961-7 doi 10.1194/jlr.D600047-JLR200.
67. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods **2013**;10(12):1213-8 doi 10.1038/nmeth.2688.
68. Chevrier S, Crowell HL, Zanotelli VRT, Engler S, Robinson MD, Bodenmiller B. Compensation of Signal Spillover in Suspension and Imaging Mass Cytometry. Cell Syst **2018**;6(5):612-20 e5 doi 10.1016/j.cels.2018.02.010.
69. Hu Y, Smyth GK. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J Immunol Methods **2009**;347(1-2):70-8 doi 10.1016/j.jim.2009.06.008.
70. Andrews S. 2010 FastQC: A Quality Control Tool for High Throughput Sequence Data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
71. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **2014**;30(15):2114-20 doi 10.1093/bioinformatics/btu170.
72. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods **2012**;9(4):357-9 doi 10.1038/nmeth.1923.
73. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE*, et al.* Model-based analysis of ChIP-Seq (MACS). Genome Biol **2008**;9(9):R137 doi 10.1186/gb-2008-9-9-r137.
74. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res **2014**;42(Web Server issue):W187-91 doi 10.1093/nar/gku365.
75. Stark R BG. 2011 DiffBind: differential binding analysis of ChIP-Seq peak data. <http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf>.
76. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics **2015**;31(14):2382-3 doi 10.1093/bioinformatics/btv145.
77. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K*, et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res **2005**;15(8):1034-50 doi 10.1101/gr.3715005.

78. Meers MP, Tenenbaum D, Henikoff S. Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling. Epigenetics Chromatin **2019**;12(1):42 doi 10.1186/s13072-019-0287-4.
79. Meers MP, Janssens DH, Henikoff S. Pioneer Factor-Nucleosome Binding Events during Differentiation Are Motif Encoded. Mol Cell **2019**;75(3):562-75 e5 doi 10.1016/j.molcel.2019.05.025.
80. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P*, et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell **2010**;38(4):576-89 doi 10.1016/j.molcel.2010.05.004.

FIGURE 1

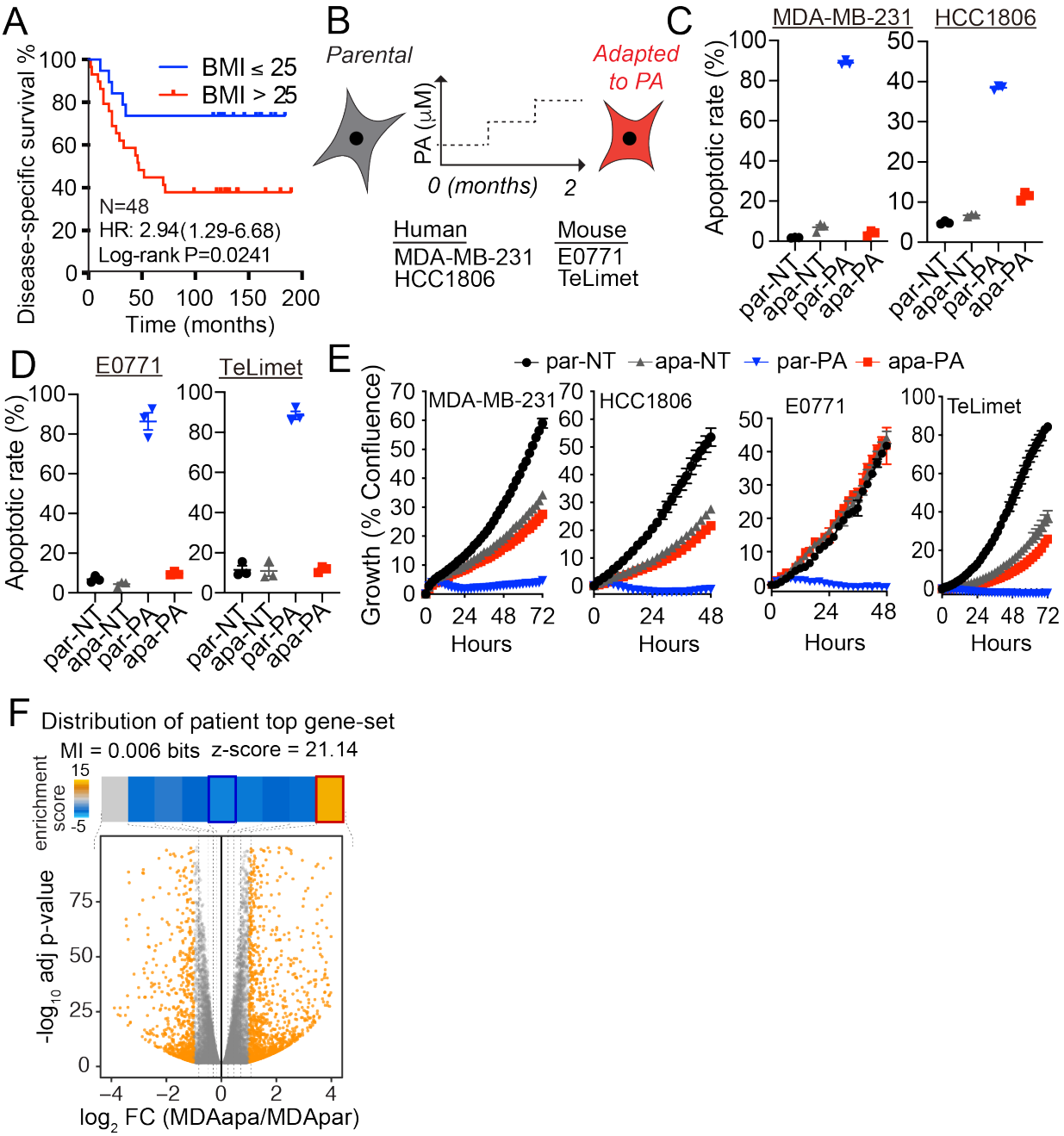


Figure 1. Transcriptional changes induced by long-term culture in palmitic acid overlap with obesity-dependent transcriptional changes in hormone receptor negative patients

(A) Kaplan-Meier curves show disease specific survival for postmenopausal and hormone receptor negative patients (N=48) with high (Red, BMI > 25) or low (Blue, BMI ≤ 25) BMI. Log-rank (Mantel-Cox) P value is denoted for difference in disease specific survival.

(B) Schematic representation of establishment of palmitic acid (PA) adaptation cell lines. Cells were first cultured in 200 μM PA supplemented media, after adaptation, PA concentration increased to 400 μM for HCC1806 and MDA-MB-231 cell lines. For TeLimet and E0771 cell lines, PA concentrations were increased to 600 μM and 500 μM respectively.

(C) Apoptotic rate of parental and adapted MDA-MB-231 and HCC1806 cells that were treated with 400 μM PA and vehicle (NT) for 48hrs. Data are represented as mean ± SEM of 3 replicates.

(D) Apoptotic rate of parental and adapted E0771 and TeLimet cells that were treated with 600 μM PA and vehicle (NT) for 48hrs. Data are represented as mean ± SEM of 3 replicates.

(E) Time-dependent proliferation assay of parental and adapted MDA-MB-231, HCC1806, E0771 and TeLimet cells following 48-72hrs. Cells were exposed to 400 μM (for MDA-MB-231 and HCC1806 cells) or 600 μM (for E0771 and TeLimet cells) PA and vehicle (NT). Cell growth was determined by high content imaging and represented as % confluence normalized to t=0. For each time point, data are represented as mean ± SEM of 4-8 replicates.

(F) The distribution of genes induced by obesity in PM hormone negative breast cancers patients among the gene expression changes observed in PA-adapted cell lines. The ~300 genes were used to perform gene-set enrichment analysis using iPAGE. Shown here is the volcano plot showing gene expression changes in PA-adapted cells relative to their parental line. iPAGE divided the spectrum of log-fold changes into equally populated bins (dotted line) and used mutual information to assess the non-random distribution of the query gene-set among these bins. We have included the mutual information value (MI) and its associated z-score reported by iPAGE. For visualization, the enrichment/depletion of the query gene-set was determined using the hyper-geometric test and the resulting p-value was used to define an enrichment score that is shown as a heatmap across the expression bins. The obesity-induced genes were significantly enrichment in the top-most bin. The red and blue borders in the heatmap denoted statistical significance for enrichment and depletion respectively.

Figure 2

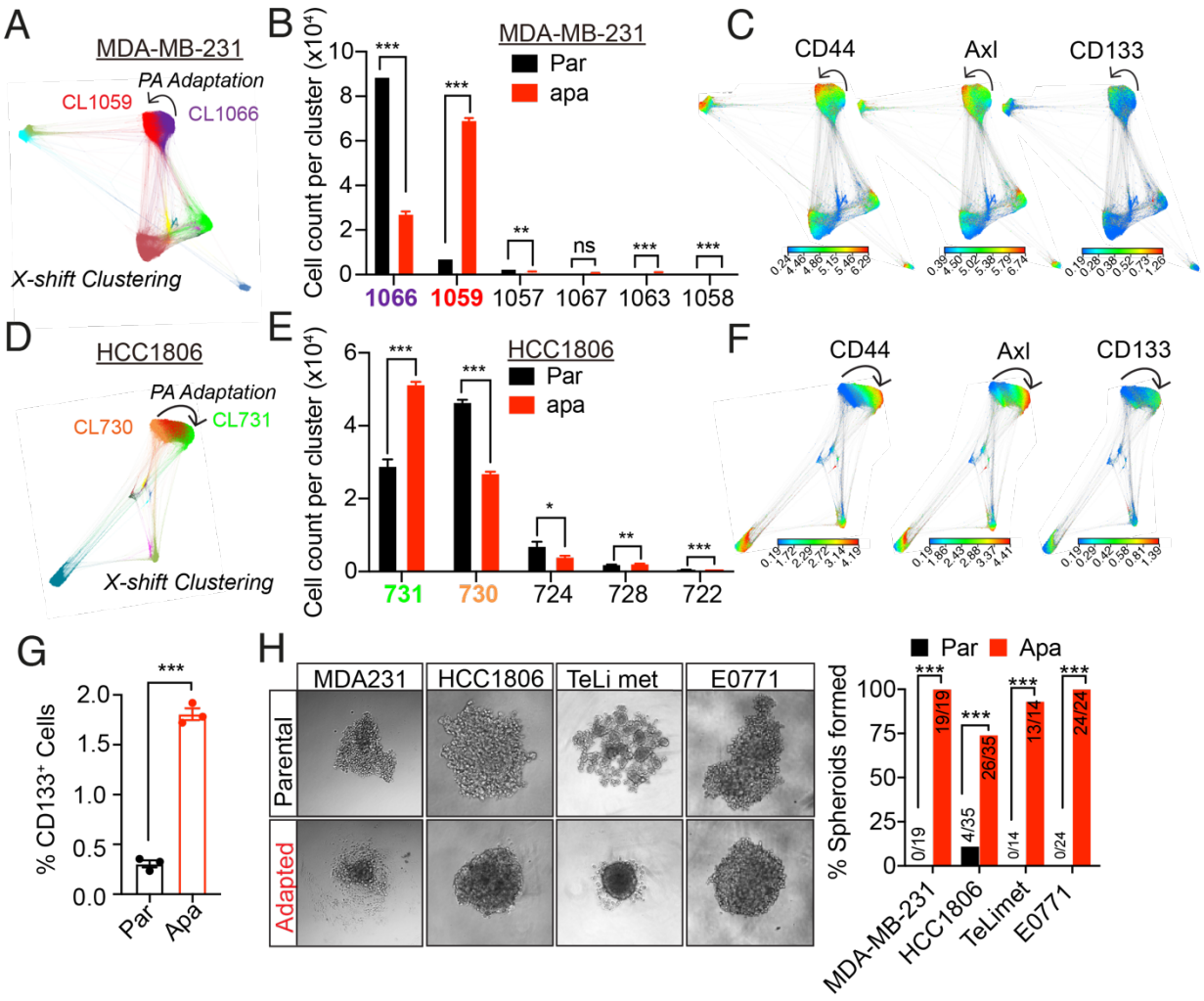


Figure 2. Long-term adaptation to palmitic acid in culture facilitates cellular dedifferentiation towards cancer stem cell-like properties

(A). X-shift clustering analysis of mass cytometry (CyTOF) data of parental and adapted MDA-MB-231 cells. Total number of analyzed cells per cell line is equal to 100 000 cells x 3 replicates. X-shift clustering algorithm is based on intensity of markers in cell population: cells with similar profiles of expression will end up in the same cluster (X-shift K = 35, color-coded for 12 clusters). Proximity between clusters shows how similar clusters are by markers expression. More than 95% of cells are located in clusters 1059 and 1066, while all other clusters together contain less than 5% of cell population. On the graph, each cluster is represented by a maximum of 1000 cells.

(B) Parental and adapted MDA-MB-231 cell count per cluster. Clusters for both parental and adapted cells containing less than 500 cells were not shown in figure.

(C) Cellular expression level of CD44, Axl and CD133 was visualized in cells from the X-shift clusters shown in (A). Color code represents expression levels of indicated markers.

(D) X-shift clustering analysis of mass cytometry (CyTOF) data of parental and adapted HCC1806 cells. Total number of analysed cells per cell line is equal to 85 000 cells x 3 replicates. X-shift clustering algorithm is based on intensity of markers in cell population: cells with similar profiles of expression will end up in the same cluster (X-shift K = 27, color-coded for 12 clusters). Proximity between clusters shows how similar clusters are by markers expression. More than 85% of cells are located in clusters 730 and 731, while all other clusters together contain less than 15% of cell population. On the graph, each cluster is represented by a maximum of 1000 cells.

(E) Parental and adapted HCC1806 cells count per cluster. Clusters for both parental and adapted cells contain less than 500 cells were not shown in figure.

(F) Cellular expression level of CD44, Axl and CD133 was visualized in cells from the X-shift clusters shown in (D)

(G) $CD133^{+}$ cells population in parental and adapted MDA-MB-231 cells. Cells were stained by CD133-APC antibody and measured by flow cytometry. Data shown as mean ± SEM of 3 replicates.

(H) Representative spheroid formation of parental and adapted MDA-MD-231 (n=19), HCC1806 (n=35), TeLimet (n=14) and E0771(n=24) cells (left panel). Quantification of spheroid formation (right panel).

For B and E, data are represented as mean ± SEM and multiple t tests was performed to assess statistical significance. For G, significance was determined with unpaired, two-tailed Student t test. For H, significance was determined with Fisher's exact test. (ns, P value ≥ 0.05; *, P value < 0.05; **, P value < 0.01; ***, P value < 0.001)

Figure 3

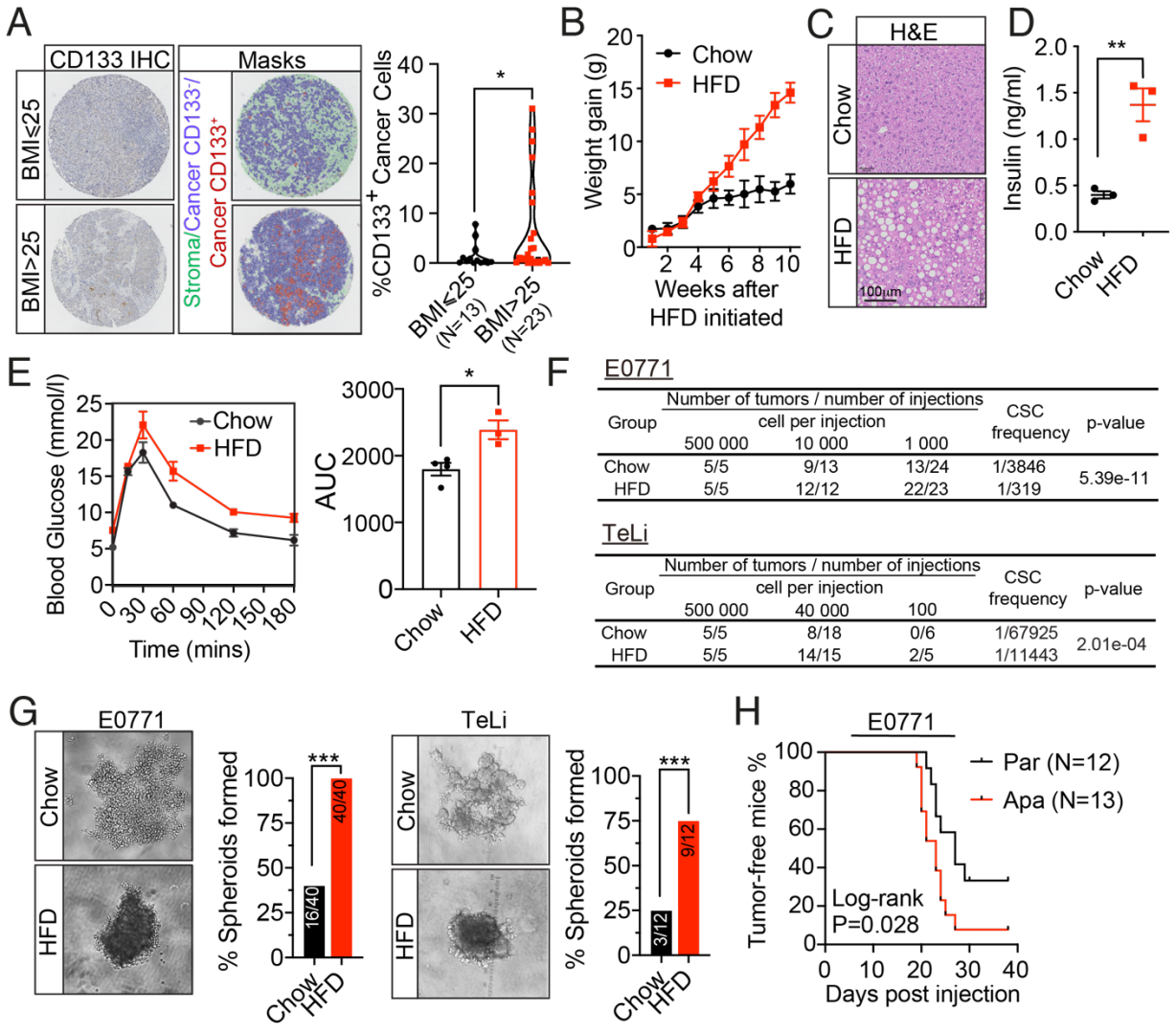


Figure 3. Obesity is associated with increased frequency of stem cell-like cancer cells in PM/ER$^-$/PR$^-$ breast cancer patients and mouse models of breast cancer

(A) Representative tissue microarray and QuPath analysis mask pictures of CD133 staining in high (BMI > 25, N = 23) or low (BMI ≤ 25, N = 13) BMI PM/ER-/PR- patients' tumor samples, stroma is marked in green, cancer CD133- cells are marked in blue and cancer CD133+ cells are marked in red (Left panel). Quantification of CD133+ cancer cells in the tumor samples based on CD133 staining (Right panel).

(B) Six-weeks old female C57BL/6J mice were fed HFD or standard chow diet (n=4 per group) for ten weeks prior to tumor implantation. Animal weight was recorded weekly, and weight gain calculated by normalizing to body weight measured at 6 weeks of age.

(C) After ten weeks of HFD or chow diet feeding, female C57BL/6J mouse liver sections were stained using H&E. Histological analysis showed increased liver steatosis in mice from the HFD group compared to mice from the chow group.

(D) Fasting plasma insulin concentrations were determined by ELISA using blood samples collected from female C57BL/6J mice fed an HFD or chow diet for ten weeks (n=3 per group).

(E) Blood glucose clearance was determined in mice fed an HFD (n=3) or chow (n=4) diet for ten weeks. Blood glucose concentrations were measured at 0min, 15mins, 30mins, 60mins, 120mins and 180mins following glucose administration by oral gavage. For each time point, data is represented as mean ± SEM. AUC = area under the curve.

(F) Tumor incidence following orthotopic implantation of the indicated number of cells. The frequency of cancer stem-like cells was calculated by the extreme limiting dilution analysis.

(G) Spheroid formation of E0771 and TeLi cells isolated from chow diet or HFD fed mice.

(H) Tumor-free survival curves of HFD fed mice implanted with parental (N=12) or adapted (N=13) E0771 cells. A palpable mass with a volume ≥ 50 mm^3 was considered a tumor. Log-rank (Mantel-Cox) P value is denoted for the difference.

For A, D and E, statistical significance determined with unpaired, two-tailed Student t test. For G, significance determined with Fisher's exact test. (*, P value < 0.05; **, P value < 0.01; ***, P value < 0.001).

Figure 4

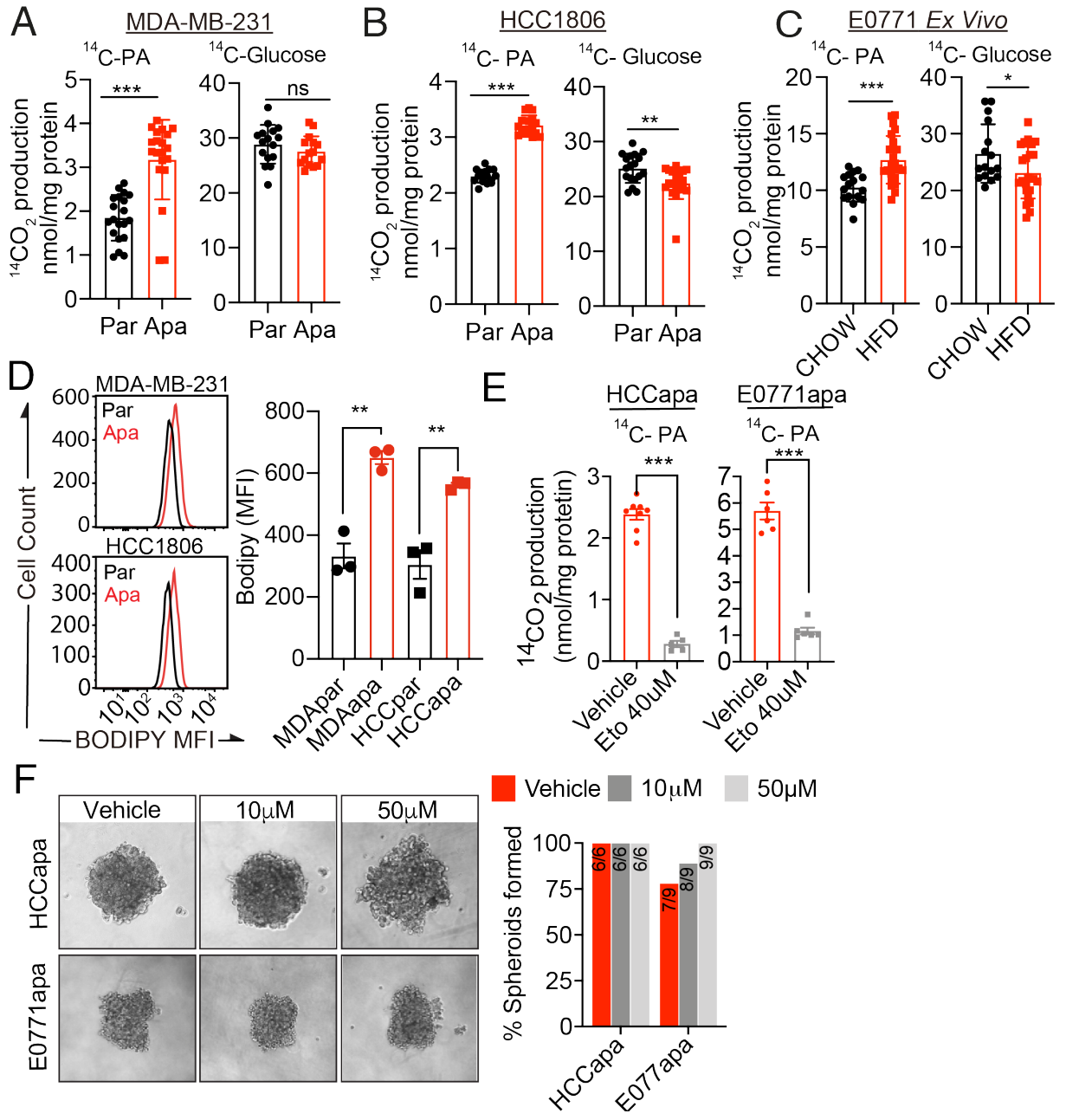


Figure 4. Adaptation to a PA-rich environment induces metabolic reprogramming towards fatty acid oxidation

(A-B) Fatty acid oxidation on parental and adapted MDA-MB-231 (A) and HCC1806 (B) cells was measured by cumulative $^{14}CO_2$-production during incubation with radio-labeled [1-^{14}C] palmitic acid (left panel). Glucose oxidation was shown by cumulative $^{14}CO_2$-production during incubation with radio-labeled D-[^{14}C(U)] glucose (right panel).

(C) Fatty acid oxidation on E0771 cells isolated from chow diet and HFD fed mice was shown by cumulative $^{14}CO_2$-production during incubation with radio-labeled [1-^{14}C] palmitic acid (left panel). Glucose oxidation was shown by cumulative $^{14}CO_2$-production during incubation with radio-labeled D-[^{14}C(U)] glucose (right panel). The oxidation data are normalized to cell protein content.

(D) Neutral lipids in parental and adapted MDA-MB-231 (upper) and HCC1806 (lower) cells stained by BODIPY and measured by flow cytometry. Representative histograms shown the mean fluorescence intensity (MFI) of BODIPY in parental and adapted cells (left panel). Quantifications (right panel) are represented as mean ± SEM of 3 replicates for each condition.

(E) Fatty acid oxidation on adapted HCC1806 (left panel) and E0771 (right panel) cells treated with vehicle control or 40 μM etomoxir. The oxidation data are normalized to cell protein content.

(F) Spheroids formation of adapted HCC1806 (n=6) and E0771 (n=9) cells in the presence of vehicle control or indicated concentrations of etomoxir.

For A-E, statistical significance determined with unpaired, two-tailed Student t test. (ns, P value > 0.05; *, P value < 0.05; **, P value < 0.01; ***, P value < 0.001)

Figure 5

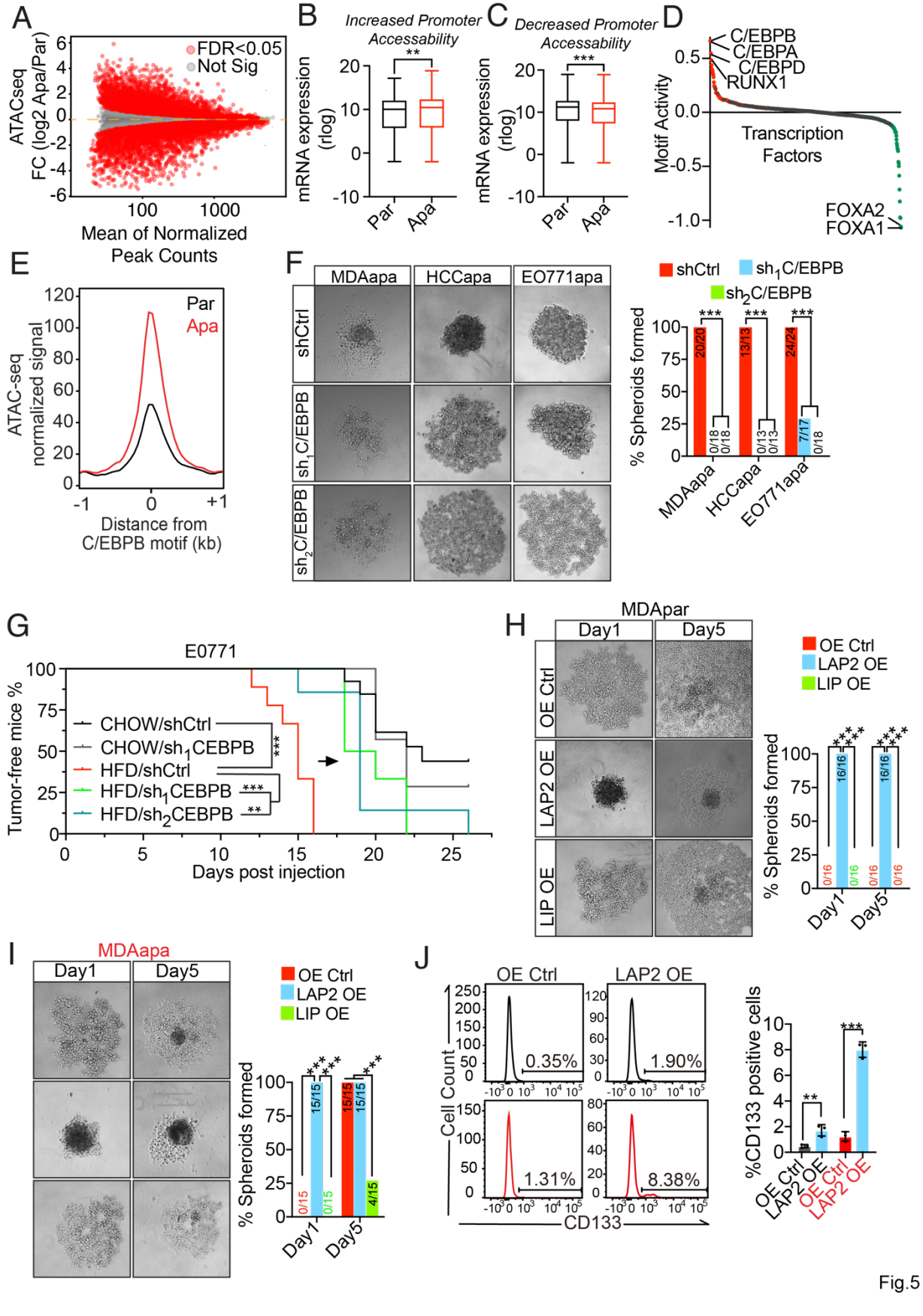


Fig.5

Figure 5. Adaptation to a PA-rich environment induces open chromatin linked with C/EBPB occupancy

(A) MA plot showing genome-wide changes in chromatin accessibility as a result of long-term adaptation to a PA-rich environment in three biological replicates of MDA-MB-231 cells as determined by ATACseq. Differentially accessible chromatin peaks are denoted in red (FDR < 0.05). There are n = 42336 increased peaks and n = 48991 decreased peaks.

(B-C) Comparison of mean mRNA expression level of genes with increased and decreased promoter accessibility, respectively, upon adaptation to a PA-rich environment in three biological replicates of parental or adapted MDA-MB-231 cells.

(D) Representation of changes in motif activity of 640 transcription factors as inferred from the degree of differential chromatin accessibility in three biological replicates of adapted versus parental MDA-MB-231 cells in (A) using diffTF. Transcription factors with differential motif activity are represented as red or green dots (FDR < 0.05). C/EBPB is identified the topmost active factor in adapted relative to parental cells.

(E) Metagene representation of the mean ATACseq signal across more accessible C/EBPB motif regions in MDA-MB-231 cells as a result of adaptation to a PA-rich environment. The mean signal of three adapted or parental MDA-MB-231 biological replicates was determined by averaging signals of 1 kb around the center of C/EBPB DNA-binding motifs.

(F) Spheroid formation of C/EBPB knockdown (sh_1C/EBPB, sh_2C/EBPB) and control (shCtrl) adapted MDA-MB-231, HCC1806 and E0771 cells.

(G) Tumor-free survival curves of chow diet and HFD fed mice orthotopically implanted with E0771 control and C/EBPB knockdown cells. Tumor volume was measured every 2-3days and tumor formation were recorded when reached a volume 50 mm^3.

(H-I) Spheroid formation of parental (H) and adapted (I) MDA-MB-231 cells after the overexpression of C/EBPB LAP2 or LIP isoforms. Spheroid formation was observed at day 1 and day 5 after seeding.

(J) Changes in the fraction of $CD133^+$ cells upon C/EBPB isoform LAP2 overexpression in parental and adapted MDA-MB-231 cells. Cells were stained by CD133-APC antibody and measured by flow cytometry. Data shown as mean ± SEM of 3 replicates.

For B and C, Mann-Whitney U test was used for statistical testing. For F, H and I, Fisher's exact test was used for statistical testing. For G, Log-rank (Mantel-Cox) test was used for statistical testing. For J, statistical significance determined with unpaired, two-tailed Student t test. (*, P value < 0.05; **, P value < 0.01; ***, P value < 0.001)

Figure 6

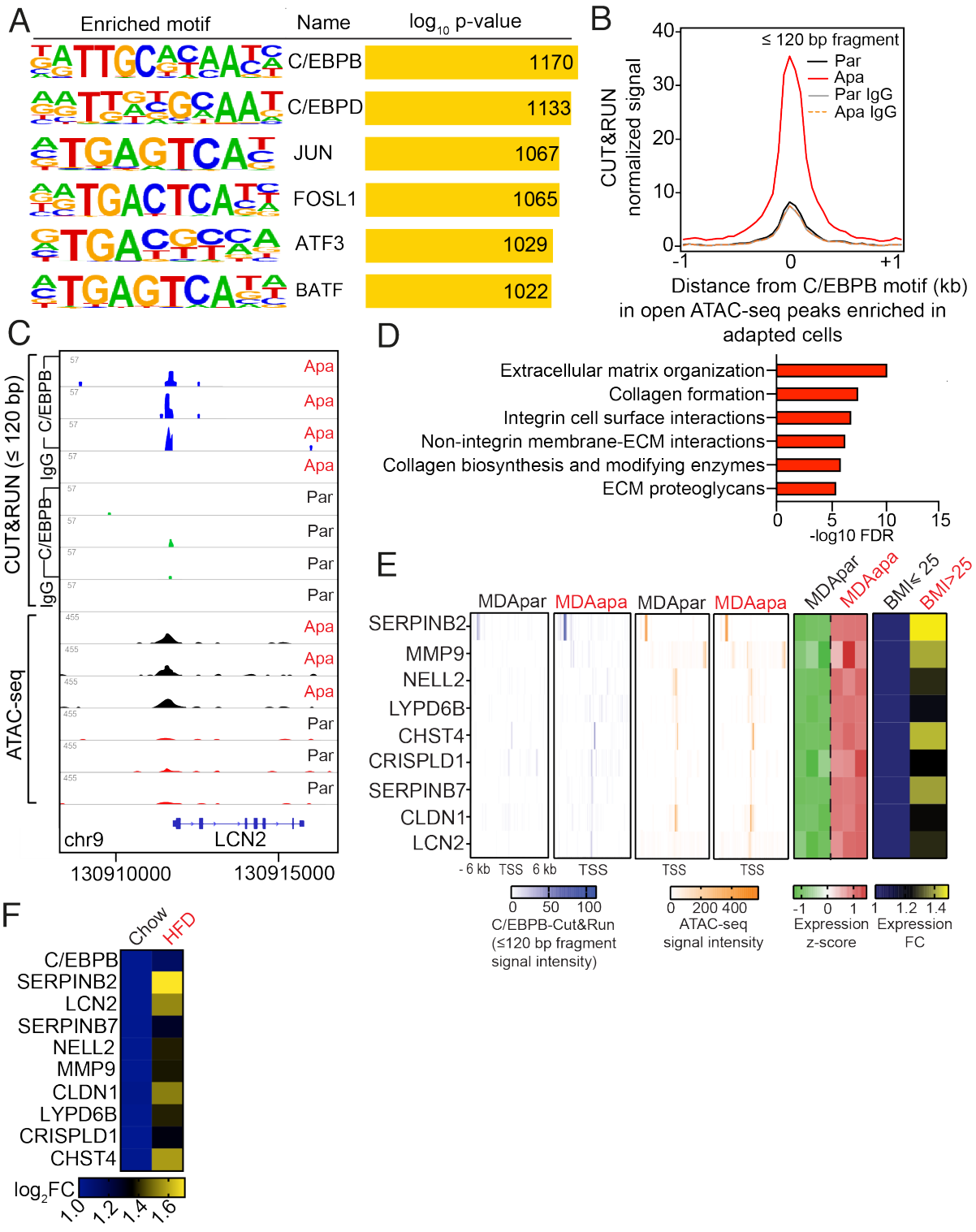


Figure 6. Differential C/EBPB occupancy regulates the expression of extracellular matrix proteins

(A) Motifs enriched in C/EBPB Cut&Run footprints in adapted MDA-MB-231 cells. The p-values shown in the figure were reported by HOMER using HOCOMOCO motifs.

(B) Metagene representation of the mean C/EBPB Cut&Run signal (fragment length ≤ 120 bp) across the same chromatin regions as in (5E) from three biological replicates of adapted or parental MDA-MB-231 cells. Control IgG Cut&Run experiment in adapted or parental cells was included for comparison.

(C) Representative genome browser tracks of normalized C/EBPB and IgG Cut&Run and ATACseq profiles around the LCN2 locus in biological replicates of adapted or parental MDA-MB-231 cells.

(D) Reactome pathway analysis of genes containing gained chromatin accessibility to C/EBPB.

(E) Heatmaps showing average Cut&Run and ATACseq signal intensity centered around the transcription start site (TSS) of the nine putative C/EBPB target genes, and the corresponding mRNA expression of the same genes in three biological replicates of MDApar and MDAapa cells (panels 1-5). Heatmap of expression fold change of the same genes in obese and overweight compared to lean patients was also shown (panel 6).

(F) Heat map showing mRNA expression of potential C/EBPB targets in E0771 cells isolated from chow diet and HFD fed mice. mRNA expression was measured by RT-qPCR with cells isolated from N=2 chow tumors and N=3 HFD tumors.

Figure 7

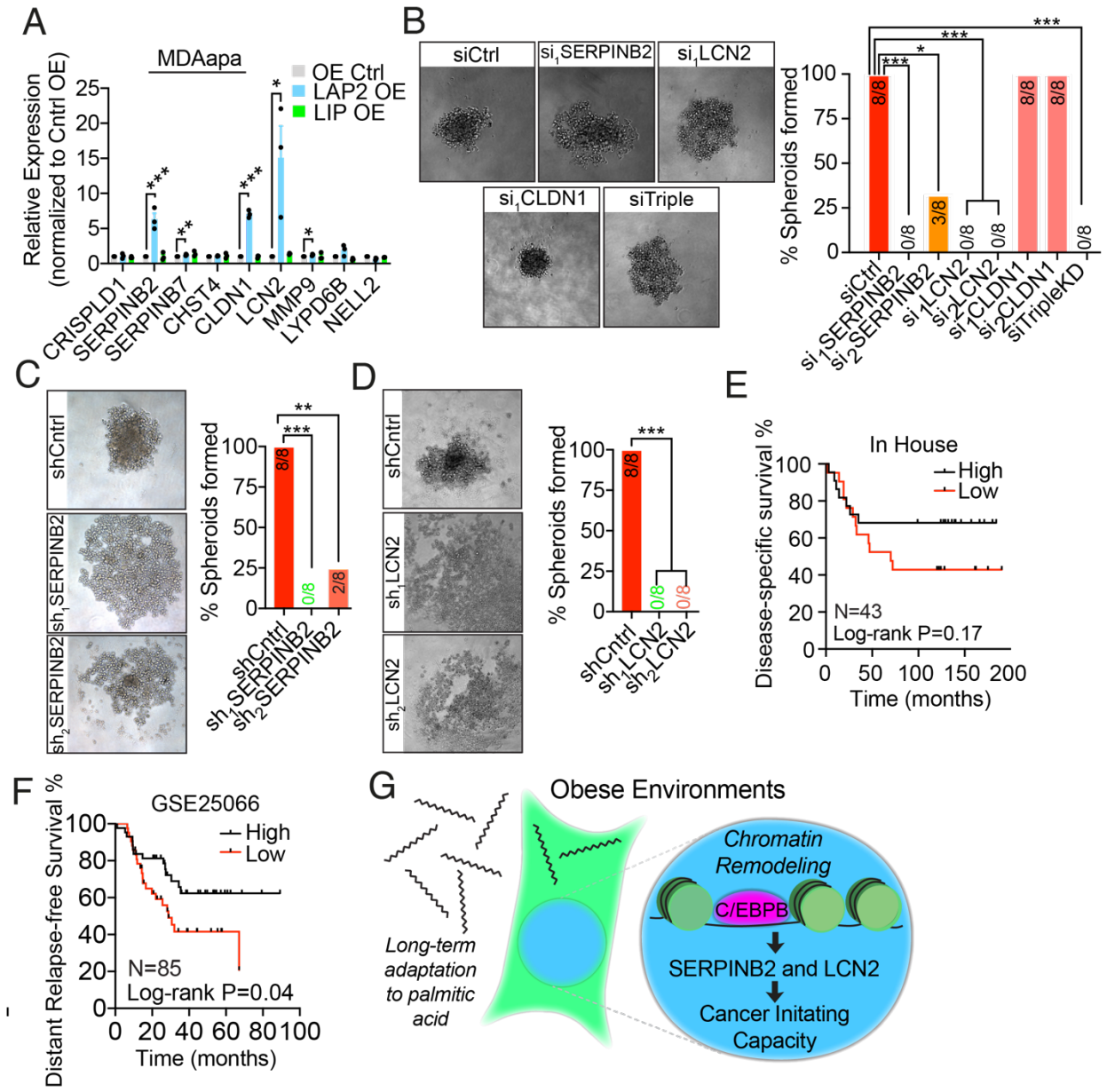


Figure 7. C/EBPB target genes SERPINB2 and LCN2 are required for cancer stem cell-like capabilities

(A) RT-qPCR was used to measure changes in the expression of C/EBPB potential target genes upon the overexpression of C/EBPB LIP or LAP2 isoforms on adapted MDA-MB-231 cells. The expression of target genes is shown as relative fold change over Control OE. Data shown as mean ± SEM of 3 independently repeated experiments.

(B) The changes of spheroids formation upon knockdown of SERPINB2, LCN2, CLDN1 and triple knockdown with siRNAs on adapted MDA-MB-231 LAP2 overexpression cells. The knockdown was performed by using two independent siRNAs for each gene.

(C-D) Changes in spheroids formation upon stable knockdown SERPINB2 (C) or LCN2 (D) on adapted MDA-MB-231 cells. The knockdown was performed by using two independent shRNAs for each gene.

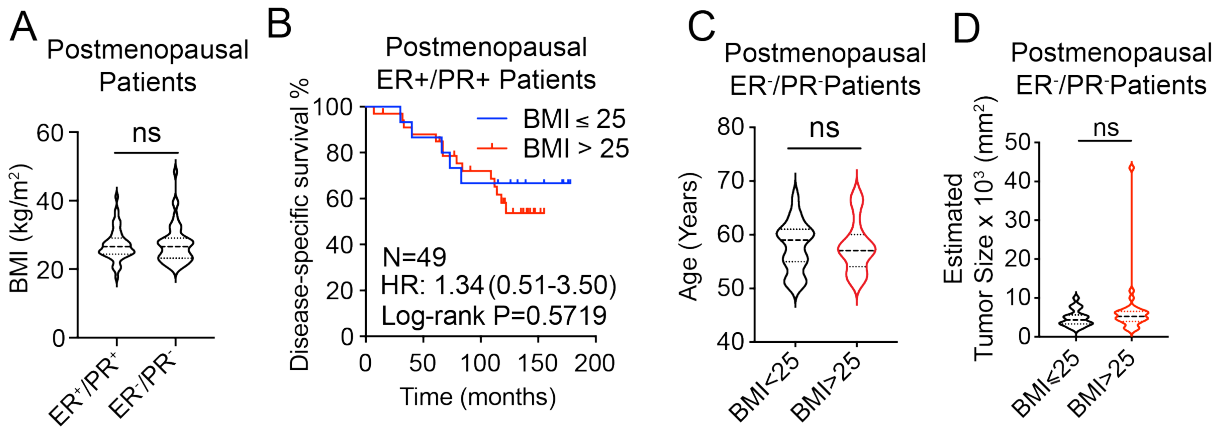
(E-F) Kaplan-Meier curves showing the effects of expression of LCN2 and SERPINB2 in PM/ER-/PR- breast cancer patients' survival. Patients were stratified into low and high expression groups by median of average expression of LCN2 and SERPINB2 (normalized gene expression) in each individual dataset, in house dataset (E) and GSE25066 (F).

(G) A schematic model of obese environment on breast cancer initiating capacity. Long-term adaptation of breast cancer cells to palmitic acid promotes initiating capacity through increased accessibility of C/EBPB binding motifs, which induces the expression of C/EBPB targets SERPINB2 and LCN2.

For A, multiple t tests were performed to assess statistical significance. For B, C and D, Fisher's exact test was used for statistical testing. For E and F, P values were determined with Log-rank (Mantel-Cox) test. (*, P value < 0.05; **, P value < 0.01; ***, P value < 0.001)

SUPPLEMENTAL INFORMATION

Supplementary Figure 1.



(A) Distributions of BMI in postmenopausal ER+/PR+ and ER-/PR- patients.

(B) Kaplan-Meier curves show disease specific survival for postmenopausal and ER+/PR+ patients (N=49) with high (Red, BMI > 25) or low (Blue, BMI ≤ 25) BMI. Log-rank (Mantel-Cox) P value is denoted for difference in disease specific survival.

(C-D) Distributions of postmenopausal ER-/PR- patients' age (C) and estimated tumor size (D) in high (BMI > 25) and low (BMI ≤ 25) BMI groups. The estimated tumor size was calculated by multiplying the largest diameter by its perpendicular.

For A, C, D, statistical significance determined with unpaired, two-tailed Student t test (ns, P value > 0.05).

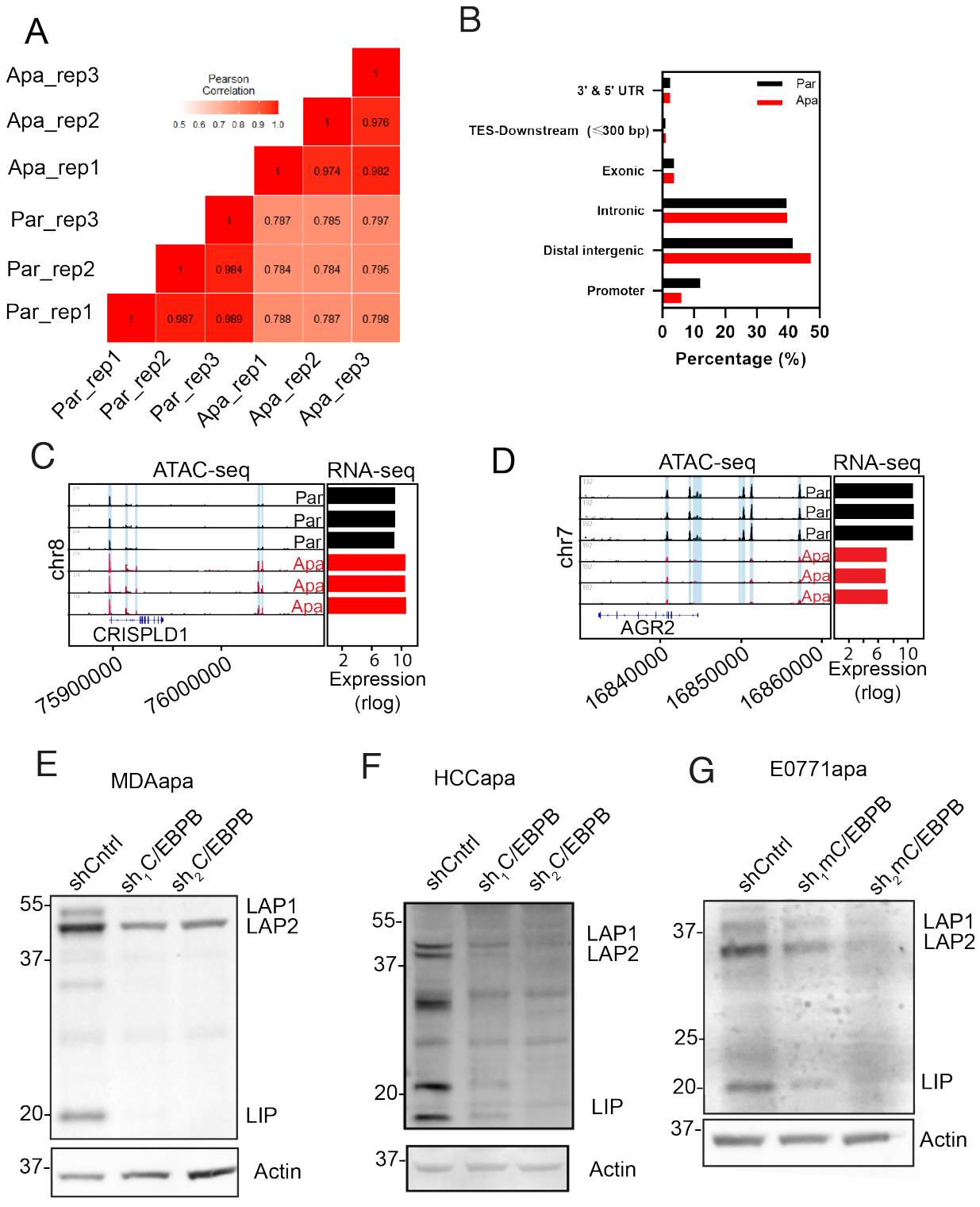
Supplementary Figure 2.

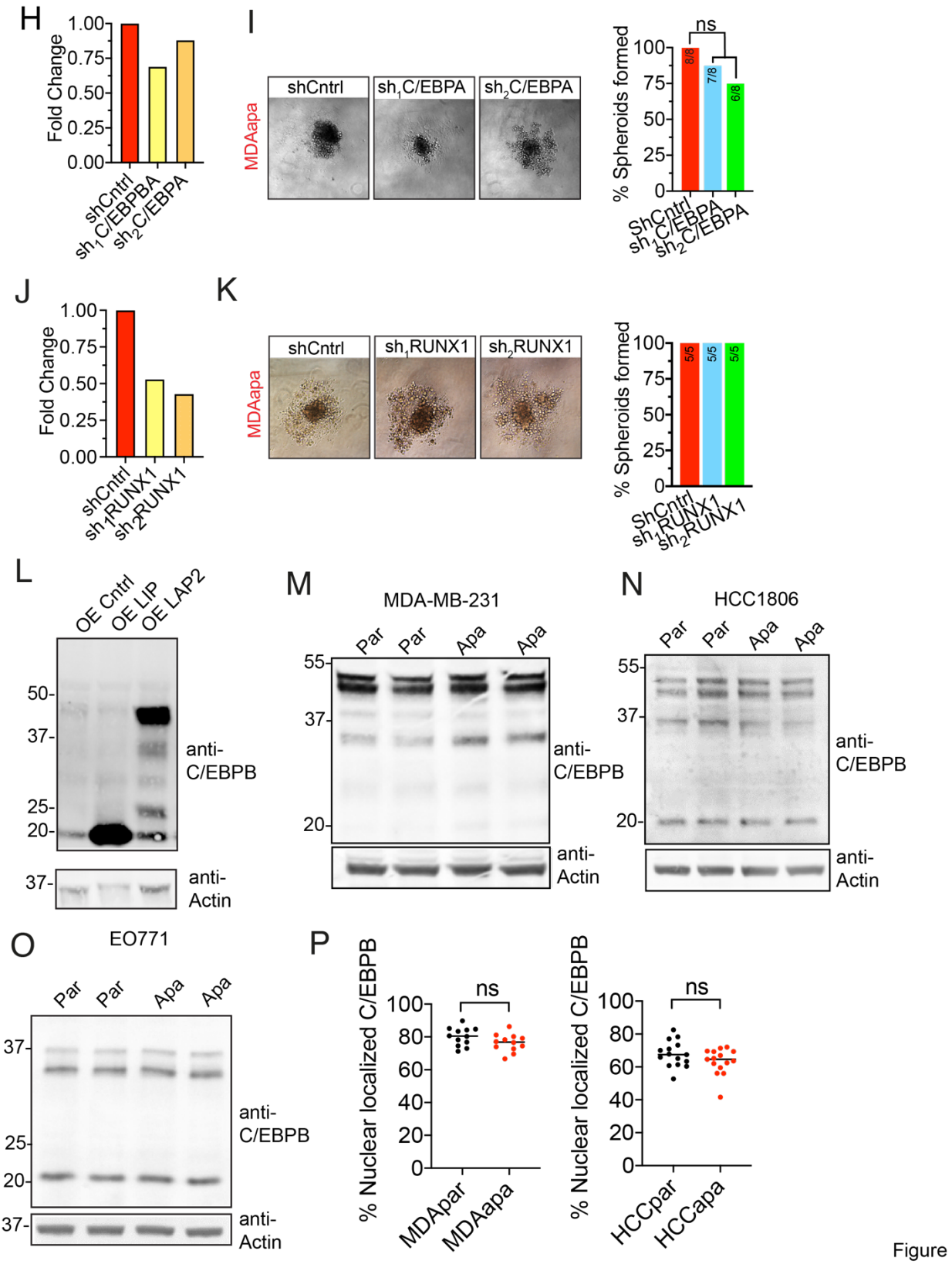
A

Isotope	Antigen	Cell Location	Epitope	Phenotype	Clone
168Er	Axl	Extracellular	Total	Stemness/EMT	1H12
160Gd	CD133	Extracellular	Total, Epitope 1	Stemness	AC133
173Yb	CD44	Extracellular	Total, Surface	Stemness	IM7
158Gd	E-cadherin	Extracellular	CD324/E-Cadherin	Epithelial	"24E10"
170Er	EGFR	Extracellular	Total EGFR	Epithelial organs/ initiates MAPk, Akt and JNk signalling	AY13
143Nd	N-cadherin	Extracellular	CD325/N-Cadherin	Stemness/mesenchymal	8C11
156Gd	p38	Intracellular	p38 [T180/Y182]	MAPK for stress response	D3F9
152Sm	pAkt	Intracellular	pAkt [S473]	PI3K pathway	D9E
176Yb	pCreb	Intracellular	pCREB [S133]	Transcription factor- stress and growth	87G3
151Eu	pEGFR	Intracellular	pEGFR [Y1068]	Activated EGFR	Y38
154Sm	pErk1/2	Intracellular	pT202/pY204	Branch of MAPK-Mek pathway	20A
175Lu	pHistone H3	Intracellular	pHistone H2A.X [Ser139]	Metaphase. Activated dwnstream of p38 or Erk1/2	HTA28
159Tb	pMAPKAPK2	Intracellular	pMAPKAPK2 [T334]	ERk1/2 activated protein downstream of p38. Response to stress	27B7
166Er	pNFKB	Intracellular	pNF-kB p65 [S529]	Transcription Factor mediator of inflammatory and immune responses	K10-895.12.50
162Dy	pPLCgamma 2	Intracellular	pPLCg2 [pY759]	Mediator of inflammatory and immune responses	K86-689.37
150Nd	pRb	Intracellular	pRb [S807/811]	G1 to S cell cycle phase	J112-906
172Yb	pS6	Intracellular	pS6 [S235/S236]	Protein translation	N7-548
141Pr	pSHP2	Intracellular	Y580	RTK phosphatase promotes signaling of JAK/STAT, PI3K/Akt Ras/MAPK pathway	D66F10
153Eu	pStat1	Intracellular	Y704		4a
145Nd	pStat3	Intracellular	pY705		4/p
146Nd	pStat5	Intracellular	pY694		00047
149Sm	pStat6	Intracellular	Y641		18/P-stat6
163Dy	TGFbeta	Intracellular	Total		TW4-6H10
154Sm	Vimentin	Intracellular	Total	Mesenchymal	D21H3
167Er	YAP	Intracellular	CTD 379-407	Stemness Hippo	H9
172YT	CC3	Intracellular	Cleavage at D175	Apoptosis	5A1E
164DY	CK7	Intracellular	Total	Luminal marker	RCK105

(A) Information of antibody panel that used in mass cytometry analysis.

Supplementary Figure 5.





Figure

(A) Pearson's correlation heatmap of ATACseq signal among biological replicates of adapted and parental MDA-MB-231 cells.

(B) Bar chart showing proportions of differential ATACseq peaks with respect to their genomic position using ChIPseeker. Promoters are defined as -1000 and + 100 bp from the nearest transcription start site (TSS), distal intergenic regions as beyond -1000 bp from this start site, and TES-downstream as ≤ +300 bp from the nearest transcription end site.

(C-D) Representative genome browser tracks of normalized ATACseq profiles around the CRISPLD1 and AGR2 loci, respectively, and the corresponding mRNA expression of the same genes, in biological replicates of adapted and parental MDA-MB-231 cells.

(E-G) Western blot of C/EBPB and Actin in whole cell lysates extracted from control (shCtrl) and two independent CEBP/B knockdown (sh_1C/EBPB and sh_2C/EBPB) adapted MDA-MB-231 (E), HCC1806 (F) and E0771 (G) cells. Actin was used for the normalization.

(H) C/EBPA mRNA expression fold changes upon C/EBPBA KD using two independent shRNAs on adapted MDA-MB-231 cells.

(I) Spheroids formation on C/EBPA knockdown and control adapted MDA-MB-231 cells (left panel) and its quantification (right panel).

(J) RUNX1 mRNA expression fold change in cells with C/EBPBA KD using two independent shRNAs on adapted MDA-MB-231 cells.

(K) Spheroids formation on RUNX1 knockdown and control adapted MDA-MB-231 cells (left panel) and its quantification (right panel).

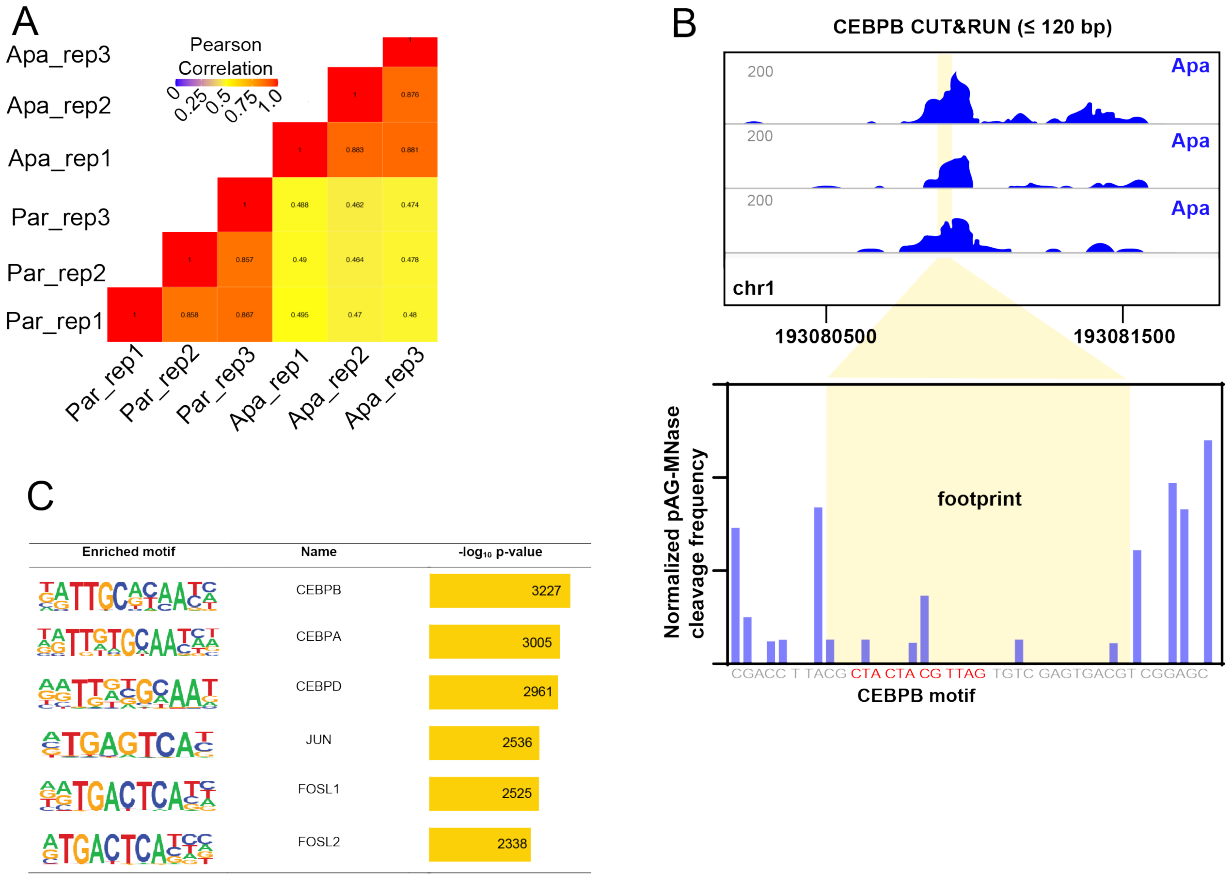
(L) Western blot against C/EBPB in cells with OE of LAP2 and LIP isoforms of C/EBPB and control MDA-MB231 PA-adapted cell line. Actin was used for the normalization.

(M-O) Western Blot against C/EBPB in parental and adapted MDA-MB231 (M), HCC1806 (N) and E0771 (O) cell lines. Actin was used for the normalization.

(P) Percentage of C/EBPB localized in the nucleus for MDA-MB231 and HCC1806 parental and PA-adapted cell lines, compared to the total C/EBPB. Quantification was based on the images of C/EBPB-immunofluorescent staining of the cell lines.

For I, Fisher's exact test was used for statistical testing. For P, statistical significance determined with unpaired, two-tailed Student t test. (ns, P value > 0.05)

Supplementary Figure 6.

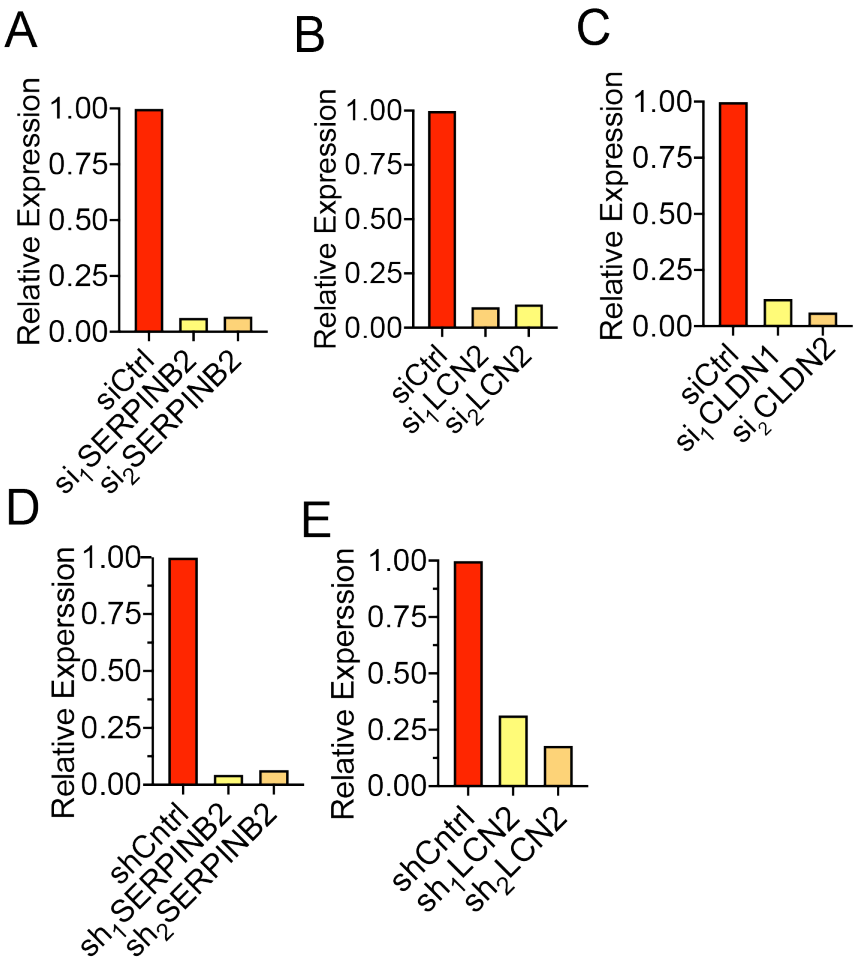


(A) Pearson's correlation heatmap of C/EBPB Cut&Run signal among biological replicates of adapted and parental MDA-MB-231 cells.

(B) Single locus footprint analysis of C/EBPB Cut&Run experiments in biological replicates of adapted MDA-MB-231 cells. Upper panel shows representative genome browser tracks of C/EBPB Cut&Run signal in the specified region in chromosome 1. Lower panel shows the total normalized pA/G-MNase cut frequency of the three biological replicates at each nucleotide around the C/EBPB motif within the identified footprint in the specified region.

(C) Motifs enriched in C/EBPB Cut&Run peaks in adapted MDA-MB-231 cells. The p-values shown in figure were reported by HOMER using HOCOMOCO motifs.

Supplementary Figure 7.



(A-C) RT-qPCR was used to measure efficiency of SERPINB2 (A), LCN2(B), CLDN1 (C) knockdown in adapted MDA-MB-231 LAP2 OE cells. Knockdown was performed by using two independent siRNAs for each gene.

(D-E) RT-qPCR was used to measure efficiency of SERPINB2 (D) and LCN2 (E) knockdown in adapted MDA-MB-231 cells. Knockdown was performed by using two independent shRNAs for each gene.

Supplementary Table S1. List of genes included in the targeted sequencing of PM/ER$^-$/PR$^-$

List of genes included in the targeted sequencing panel

ABL1	ERCC2	MAPK10	ROS1
ABL2	ERCC3	MAPK7	RPS6KB1
ACVR2A	ERCC4	MAPK8	RPTOR
AKT1	ERCC5	MAPK9	RRM2B
AKT2	ESR1	MCL1	RSPO2
AKT3	ETV1	MDM2	RSPO3
ALK	EZH2	MDM4	RUNX1
APC	FADD	MED12	SETD2
AR	FAM123B	MED12L	SF3B1
ARAF	FANCA	MED13	SFTPA1
ARFRP1	FANCC	MED29	SHC1
ARID1A	FANCD2	MEN1	SKP2
ARID1B	FANCE	MET	SLIT2
ARID2	FANCF	MITF	SMAD2
ASXL1	FANCG	MLH1	SMAD3
ATM	FAS	MLL	SMAD4
ATR	FBXO11	MLL2	SMARCA4
ATRX	FBXW7	MLL3	SMARCB1
AURKA	FGFR1	MPL	SMO
AURKB	FGFR2	MRAS	SMURF1
AXIN1	FGFR3	MRE11A	SOCS1
BAG4	FGFR4	MSH2	SOX10
BAP1	FH	MSH6	SOX2
BCL11A	FLT1	MST1	SOX9
BCL2	FLT3	MTDH	SPOP
BCL2A1	FLT4	MTOR	SRC
BCL2L1	FOXA1	MUTYH	SRSF2
BCL2L2	FOXL2	MYB	STAT3
BCL6	FOXO1	MYC	STK11
BCOR	FOXP4	MYCL1	SUFU
BIRC2	GAB2	MYCN	TBX22
BIRC7	GABRG1	MYD88	TBX3
BLM	GATA1	MYO3A	TERT
BPTF	GATA2	MYO5B	TET2
BRAF	GATA3	MYOC	TGFBR2
BRCA1	GATA6	NBN	TNFAIP3
BRCA2	GNA11	NCOA2	TOP1
BRIP1	GNAQ	NCOA3	TP53
BUB1B	GNAS	NF1	TP63
C11orf30	GPC5	NF2	TP73
CARD11	GPR124	NFE2L2	TRAF2
CASP8	GRB2	NGFR	TSC1
CBL	GRB7	NKX2-1	TSC2
CCND1	GRID1	NOTCH1	TSHR
CCND2	GUCY1A2	NOTCH2	U2AF1
CCND3	H3F3A	NOTCH3	USP9X
CCNE1	HIST1H3B	NOTCH4	VEGFA

Supplementary Table S2. The list of materials sources

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
C/EBPB	Santa Cruz	sc-7962
CD133 (application – TMA staining)	Miltenyi Biotec	130-090-422
Beta-actin	Invitrogen	PA1-183
E-cadherin-158Gd, Extracellular	Fluidigm	3158021A
Cleaved caspase 3-142Nd, Intracellular	Cell Signalling technology	Clone SA1E
CD44-173Yb, Extracellular	Fluidigm	3150018B
CD133-160Gd, Extra- and intracellular, conjugated in the lab	Miltenyi Biotec	130-090-422
Axl-168Er, Extracellular	BGB/creative biolabs	HPAB-0110-LS
CD133-APC	Invitrogen	17-1331-81
CD44-FITC	BioLegend	338803
PEGFR-151Eu, Intracellular, conjugated in the lab	abcam	ab32430
PCreb-176Yb, Intracellular	Fluidigm	3176005A
PAkt-152Sm, Intracellular	Fluidigm	3156002A
P38-156Gd, Intracellular	Fluidigm	3156002A
N-cadherin-143Nd, Extracellular	Fluidigm	3143016B
Keratin7-164Dy, Extracellular	BD	ab9021
EGFR-170Er, Extra- and intracellular	Fluidigm	3170009B

YAP-167Er, Intracellular, conjugated in the lab	Santa Cruz	sc-271134
Vimentin-154Sm, Intracellular	Fluidigm	3154014A
TGFβ-163Dy, Extra- and intracellular	Fluidigm	3163010B
PStat5-147Sm, Intracellular, conjugated in the lab	BD	562077
PStat3-145Nd, Intracellular, conjugated in the lab	BD	624084
PStat1-153Eu, Intracellular	Fluidigm	3153005A
PSHP2-141Pr, Intracellular	Fluidigm	3141002A
PS6-172Yb, Intracellular	Fluidigm	3172008A
PRb-150Nd, Intracellular	Fluidigm	3150013A
PNFKB-166Er, Intracellular	Fluidigm	3166006A
PMAPKAPK2-159Tb, Intracellular	Fluidigm	3159010A
pHistone H3-175Lu, Intracellular	Fluidigm	3175012A
pErk1/2-171Yb, Intracellular, conjugated in the lab	BD	624084
AF647 goat anti-mouse	Life Technologies	A21238
IRDye® 800CW Donkey anti-Rabbit IgG (H + L), 0.1 mg	Leicor	[P/N 926-32213], 0,1 mg
IRDye® 680RD Goat anti-Mouse IgG (H + L), 0.1 mg	Leicor	[P/N 925-68070], 0,1 mg
Mouse IgG isotype	Merck-Millipore	12-371
Bacterial and Virus Strains		
MAX Efficiency™ DH5α™ Competent Cells	Thermofisher Scientific	18258012
Biological Samples		

TMAs of breast cancer tissues	Haaukeland University Hospital	(Chrisanthar et al., 2011)
Chemicals, Peptides, and Recombinant Proteins		
Penicillin/streptomycin	Sigma	P-0781
Fetal bovine serum	Sigma	F-7524
DMEM	Sigma	D5671
RPMI1640	Sigma	R8758
BSA, fatty-acids free	Sigma	A7030
Palmitic Acid	Sigma	P5585
Phusion polymerase	NEB	M0530S
QIAquick Gel Extraction Kit	Qiagen	28704
QIAGEN Plasmid Plus Maxi Kit	Qiagen	12965
Opti-MEM	Thermo Fisher	31985070
Lipofectamine 2000	Invitrogen	11668019
Polybrene Infection / Transfection Reagent	Sigma	TR-1003-G (1 ML)
Puromycin dihydrochloride from Streptomyces alboniger	Sigma	P8833-100MG
Annexin V, Alexa Fluor™ 488 conjugate	Thermo Fisher	A13201
Propidium Iodide	Sigma	P4864
Trypsin	Sigma	T4049
BSA	Sigma	A9647
Accutase	Sigma	A6964
Bodipy reagent	Thermo Fisher	D3922
DPBS	Gibco	14040-133
DAPI	Sigma	D9542

ProLong™ Diamond Antifade Mountant	Invitrogen	P36970
NuPAGE™ MOPS SDS Running Buffer (20X)	Invitrogen	NP000102
HEPES solution	Sigma	H0887-100ML
Tween 20	Sigma	P9616-100ML
PBS	Thermo Fischer Scientific	14040133
Tris/EDTA buffer, pH 9	Dako	S2367
Dako Real Peroxidase Blocking solution	Dako	S2023
Protein Block, Serum-free	Dako	X0909
Antibody Diluent with Background Reducing Components	Dako	S3022
Dako Wash Buffer	Dako	S3006
DAB+	Dako	K3468
Hematoxylin	Dako	S3301
Pertex	Histolab	00801
TrypLE Express	Gibco	12604-021
DNase	Sigma	DN25
CyTOF barcodes	Fluidigm	201060
Permeabilization buffer	Fluidigm	201057
Cell Staining Buffer	Fluidigm	201068
iridium cell tracker	Fluidigm	201192A
CyTOF water	Fluidigm	201069
NP-40	New England Biolabs	B2704
5% digitonin	Invitrogen	BN2006
AMPure XP beads	Beckman	A63880

BioMag® Plus Concanavalin A	Bangs Laboratories	BP531
Spermidine	Sigma	S2501
Roche Complete Protease Inhibitor, EDTA-free	Roche	05 892 791 001
PhosSTOP™	Roche	04 906 837 001
Protein-A/G-MNase	Epicypher	15-1016
EDTA	Sigma	03690
EGTA	Boston BioProducts	BM723
RNase A	Sigma	R4642
GlycoBlue	ThermoFisher	AM9515
Matrigel	Corning	356231
L-Glutamine	Sigma	G-7513
Poly-L-Lysine	Sigma	P4832
Paraformaldehyde Aqueous Solution (PFA)	Electron Microscopy Sciences	15710
L-carnitine	Sigma	C-0283
D-[^{14}C(u)]-Glucose	Perkin Elmer	NEC042A001MC
[^{14}C]-Palmitic Acid	PerkinElmer	NEC075H25OUC
Etomoxir	Sigma	E-1905-5M
NaF	Sigma	S-6776
$NaVO_4$	Aldrich	450243
Blotting-Grade Blocker	BioRad	1706404
Tris-HCl	Sigma	T2194-1L
NaCl	Sigma	S5150
$MgCl_2$	Sigma	8266

CaCl2	Sigma	C4830
Lipofectamine™ RNAiMAX Transfection Reagent	Invitrogen	13778075
Critical Commercial Assays		
PI/RNase staining kit	BD Pharmigen	550825
Senescence β-Galactosidase Staining Kit	Cell Signalling	9860S
Total RNA purification Kit	NORGEN Biotek	37500
SuperScript® III First-Strand Synthesis Kit	ThermoFisher Scientific	18080-051
LightCycler® 480 SYBR Green I Master Mix	Roche	04887352001
MACH3 mouse probe	Biocare Medical	BC-M3M530H
RNA Clean & concentrator with DnaseI kit	Biosite	R1013
Nextera DNA Library Prep kit	Illumina	FC-121-1030
DNA Clean and Concentrator-5 kit	Zymo	D4014
NEBNext® Ultra™ II DNA Library Prep Kit	New Englands Biolabs	E7645
MicroScintPS	PerkinElmer	6013631
Deposited Data		
GSE25066		
Experimental Models: Cell Lines		
MDA-MB-231	ATCC	RRID:CVCL_0062
HCC1806	ATCC	PRID:CVCL_1258
E0771	CH3 BioSystems	SKU: 94A001
TeLi	This article	N/A
TeLimet	This article	N/A

HEK293T	ATCC	PRID:CVCL_0063
Experimental Models: Organisms/Strains		
C57BL/6J	The Jackson Laboratory	000664
Oligonucleotides		
Mouse Actin Forward: TACCACAGGCATTGTGATGG Reverse: TTTGATGTCACGCACGATTT Application: qPCR	IDT	N/A
Mouse C/EBPB Forward: GGTTTCGGGACTTGATGCA Reverse: CAACAACCCCGCAGGAAC Application: qPCR	IDT	N/A
Human HPRT Forward: CCTGACCAAGGAAAGCAAAG Reverse: GACCAGTCAACAGGGGACAT Application: qPCR	IDT	N/A
Human C/EBPB Forward:	IDT	N/A

TCGCAGGTCAAGAGCAAGG Reverse: TACTCGTCGCTGTGCTTGTC Application: qPCR		
Human RUNX1 Forward: CTGCTCCGTGCTGCCTAC Reverse: AGCCATCACAGTGACCAGAGT Application: qPCR	IDT	N/A
Human C/EBPA Forward: GGAGCTGAGATCCCGACA Reverse: TTCTAAGGACAGGCGTGGAG Application: qPCR	IDT	N/A
Human SERPINB2 Forward: CATGGAGCATCTCGTCCAC Reverse: ACTGCATTGGCTCCCACTT Application: qPCR	IDT	N/A
Human NELL2 Forward: TAAGGGTATAATGCAAGATGTCCAATT Reverse: AGATCTGGGCACTGAGCAATAAA Application: qPCR	IDT	N/A

Human CLADN1 Forward: GAAGTGCTTGGAAGACGATG Reverse: GAGCCTGACCAAATTCGTAC Application: qPCR	IDT	N/A
Human SERPINB7 Forward: CACTGGTGACTTGACCCTTCCT Reverse: GGTGAGACACATGGTGGTAGAATG Application: qPCR	IDT	N/A
Human CHST4 Forward: TGGCCATCTTGGCTCTATTC Reverse: CTGCTTGAAGGTCATCCACA Application: qPCR	IDT	N/A
Human MMP9 Forward: GTTCGACGTGAAGGCGCAG Reverse: TAGTGTGGTGTCTCACGAAGG Application: qPCR	IDT	N/A
Human LCN2 Forward: TCACCCTCTACGGGAGAACC	IDT	N/A

Reverse: GGGACAGGGAAGACGATGTG Application: qPCR		
Human LYPD6B Forward: TGCAAACCTTTTCACTGTTCCA Reverse: GAGAGCGTGACAGAGGAGCAG Application: qPCR	IDT	N/A
Human CRISPLD1 Forward: TGCCCAAGAGTATACTGTCCT Reverse: GATTTCGAACCACTCCAGCA Application: qPCR	IDT	N/A
LCN2 siRNA	IDT	hs.Ri.LCN2.13.1, hs.Ri.LCN2.13.3
SERPINB2 siRNA	IDT	hs.Ri.SERPINB2.13.1, hs.Ri.SERPINB2.13.2
CLDN1 siRNA	IDT	hs.Ri.CLDN1.13.1, hs.Ri.CLDN1.13.2, hs.Ri.CLDN1.13.3
Negative Control DsiRNA	IDT	51-01-14-04
Recombinant DNA		
mC/EBPB-shRNA-1	Sigma	SHCLNG-NM_005194, TRCN0000364533

mC/EBPB-shRNA-2	Sigma	SHCLNG-NM_009883, TRCN0000231409
hC/EBPB-shRNA-1	Sigma	SHCLNG-NM_005194, TRCN0000364533
hC/EBPB-shRNA-2	Sigma	SHCLNG-NM_005194, TRCN0000007443
scramble shRNA	Sigma	1864
hC/EBPB-LAP2	Addgene	15712
hC/EBPB-LIP	Addgene	15713
pBABE-puro	Addgene	1764
Software and Algorithms		
GraphPad Prism	GraphPad	Version 8.4.1
FlowJo	BD Biosciences	version 10.7.0
X-Shift	VorteX	(Samusik et al., 2016)
R		Version 3.5.0
Other		
Rat and Mouse No.1 Maintenance	Special Diet Services	RM1 (P) 801151
Rodent Diet With 60 kcal% Fat	Research Diets	D12492