npj | breast cancer

www.nature.com/npjbcancer

ARTICLE OPEN

Lipid exposure activates gene expression changes associated with estrogen receptor negative breast cancer

Shivangi Yadav[1], Ranya Virk[2], Carolina H. Chung[3], Mariana Bustamante Eduardo[1], David VanDerway[2], Duojiao Chen[4], Kirsten Burdett[5], Hongyu Gao[4], Zexian Zeng[6], Manish Ranjan[1], Gannon Cottone[1], Xiaoling Xuei[4], Sriram Chandrasekaran[3,7,8,9], Vadim Backman[2], Robert Chatterton[10], Seema Ahsan Khan[1]✉ and Susan E. Clare[1]✉

Improved understanding of local breast biology that favors the development of estrogen receptor negative (ER−) breast cancer (BC) would foster better prevention strategies. We have previously shown that overexpression of specific lipid metabolism genes is associated with the development of ER− BC. We now report results of exposure of MCF-10A and MCF-12A cells, and mammary organoids to representative medium- and long-chain polyunsaturated fatty acids. This exposure caused a dynamic and profound change in gene expression, accompanied by changes in chromatin packing density, chromatin accessibility, and histone posttranslational modifications (PTMs). We identified 38 metabolic reactions that showed significantly increased activity, including reactions related to one-carbon metabolism. Among these reactions are those that produce S-adenosyl-L-methionine for histone PTMs. Utilizing both an in-vitro model and samples from women at high risk for ER− BC, we show that lipid exposure engenders gene expression, signaling pathway activation, and histone marks associated with the development of ER− BC.

npj Breast Cancer (2022)8:59 ; https://doi.org/10.1038/s41523-022-00422-0

INTRODUCTION

Breast cancer is a heterogeneous disease with different molecular subtypes that are characterized, at a minimum, by the expression of the estrogen receptor (ER), progesterone receptor (PR), and Human epidermal growth factor receptor 2 (HER2)/neu[1]. Although multiple statistical tools have been developed to quantify breast cancer risk[2], they do not predict breast cancer subtypes. Current breast cancer prevention with selective estrogen receptor modulators (SERM) and aromatase inhibitors decreases the risk of estrogen-receptor (ER) positive breast cancer sub-types, but not those without ER expression[3–5]. Thus, determining the etiologic/biologic factors that favor the development of ER-negative breast cancer will potentially enable the development of both strategies to identify women at risk for ER-negative disease as well as targeted preventive and therapeutic agents.

Given the poor understanding of the genesis of sporadic ER-negative breast cancer, we set out to study this using the contralateral, unaffected breast of patients with unilateral breast cancer as a model. Studies of metachronous contralateral breast cancer show a similarity in the ER status of the contralateral cancer to the index primary[6–8]. Therefore, the contralateral unaffected breast (CUB) of women undergoing surgical therapy for newly diagnosed unilateral breast cancer can be employed as a model to discover potential markers of subtype-specific risk. In a previous study, we performed Illumina expression arrays on epithelial cells from the CUB of breast cancer patients and identified a lipid metabolism (LiMe) gene signature which was enriched in the CUBs of women with ER- breast cancer[9]. Among these are genes that control critical steps in lipid and energy metabolism. We validated this signature in an independent set of 36 human samples and re-confirmed the above results in fresh frozen tissues obtained from a new set of ER+ and ER− breast cancer patients, each time using laser capture microdissection (LCM) to obtain epithelial cells from tumor and CUB samples[10]. Again, we found significantly higher expression of LiMe genes in CUBs from women with ER− breast cancer, compared to both CUBS from women with ER+ breast cancer, and breast epithelium from a control group of women undergoing reduction mammoplasty. However, the specific genes comprising this overexpressed set had no specific function or group of functions in common and did not suggest specific mechanistic explanations as to why lipid metabolism pathways would aid ER− breast cancer development. In the present study, we address possible mechanistic explanations for our previous observations.

Major reprogramming of cellular energetics is one of two emerging hallmarks of cancer[11]. Altered lipid metabolism is posited to be a driver of carcinogenesis in various cancers, including ovarian[12], prostate[13,14], liver[15] and triple negative breast cancer[16,17]. Increased lipid metabolism has also been shown to serve as a survival signal that enables tumor recurrence and has been suggested as an Achilles heel for combating breast cancer progression[18]. Despite this recognition of the importance of fatty acid metabolism, its role in the transformation of a normal cell to the malignant state is largely unknown. Metabolomic studies of the concentrations of several free fatty acids in primary breast tumors, including linoleate, palmitate, and oleate, as a function of breast cancer subtype have revealed significant differences across the subtypes, with the highest concentrations in basal-like breast cancer[19]. Conjugation of long-chain fatty acids to carnitine for

[1]Department of Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA. [2]Department of Biomedical Engineering, Northwestern University, Evanston, IL 60208-2850, USA. [3]Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109, USA. [4]Center of for Medical Genomics, Indiana University School of Medicine, Indianapolis, IN 46202, USA. [5]Department of Preventive Medicine, Northwestern University, Chicago, IL 60611, USA. [6]Department of Data Sciences, Dana Farber Cancer Institute, Harvard T.H. Chan School of Public Health, Boston, MA 02215, USA. [7]Program in Chemical Biology, University of Michigan, Ann Arbor, MI 48109, USA. [8]Rogel Cancer Center, University of Michigan Medical School, Ann Arbor, MI 48109, USA. [9]Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA. [10]Department of Obstetrics and Gynecology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA.
✉email: s-khan2@northwestern.edu; susan.clare@northwestern.edu

transport into the mitochondria and subsequent fatty acid oxidation (FAO) was observed to be highest in basal-like breast cancers, followed by luminal B ~HER2-enriched, with luminal A tumors displaying the lowest levels[19]. Another study, which utilized Raman spectroscopy to interrogate tissue, revealed that histologically normal breast tissue centimeters removed from the breast malignancy have significantly higher polyunsaturated fatty acid levels compared with normal tissue from cancer-free subjects[20].

The kinetic and thermodynamic properties of the chromatin modification reactions are commensurate with the dynamic range of the physiological concentrations of the corresponding intermediates in metabolism[21]. Therefore, we sought to determine if the LiMe signature we observed in the CUBs of ER- patients is associated with chromatin modifications and histone PTMs secondary to changes in metabolism fostered by exposure to medium and long-chain fatty acids.

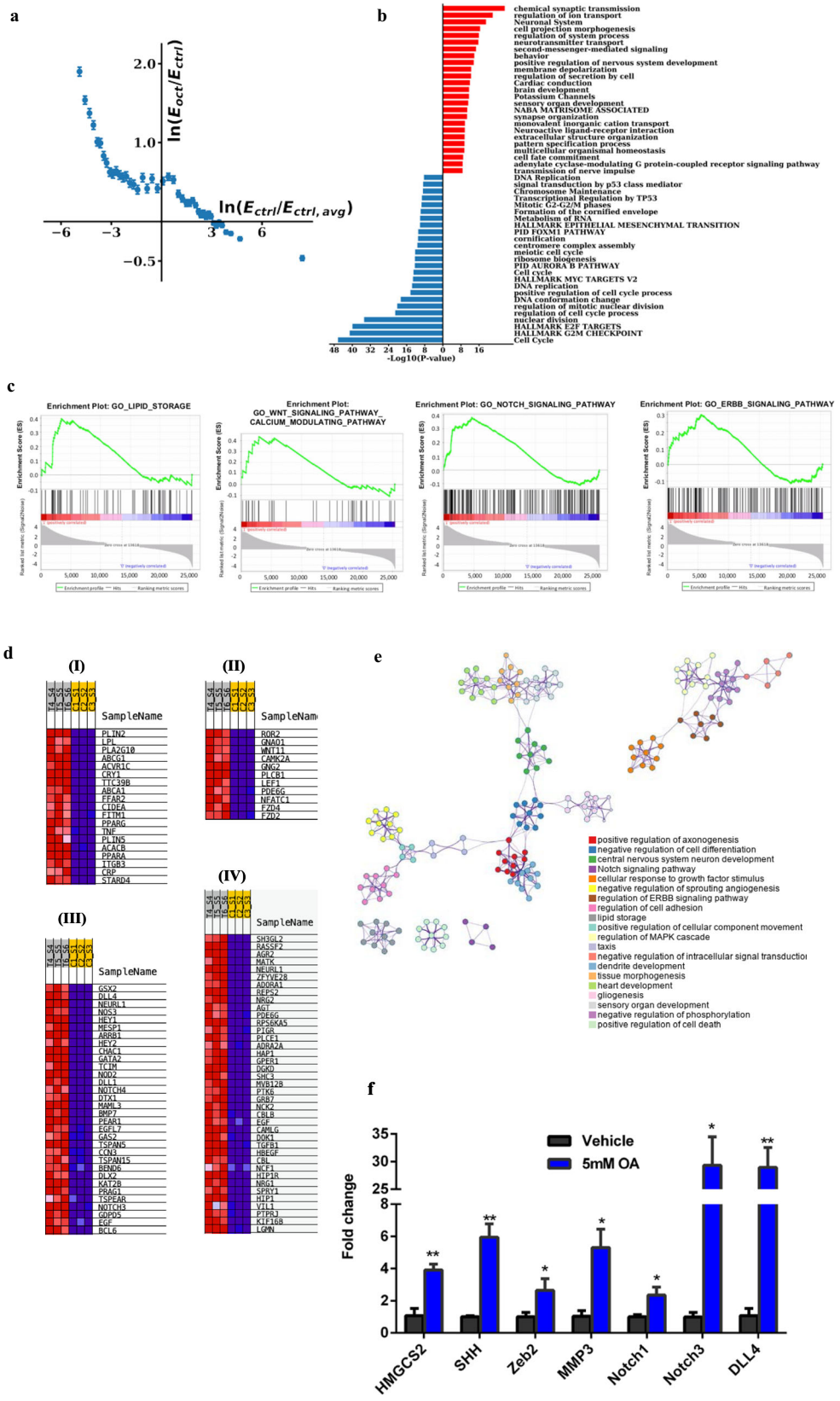
RESULTS

Lipid facilitates transcriptional reprogramming in non-transformed mammary cells

We established an in vitro model by exposing estrogen and progesterone receptor (PR) negative MCF-10A cells to octanoate (OA), a medium chain eight-carbon fatty acid. Due to its small size and lipophilic nature octanoate does not depend on fatty acid transport proteins to traverse cell membranes and is readily oxidized in the mitochondria to form acetyl-CoA[22,23]. We performed RNA-seq to determine the effects of octanoate treatment on gene expression in the MCF-10A cells. RNA-seq analysis revealed that 24 h of octanoate treatment produces a transcriptional profile that is completely distinct from vehicle-treated controls (Fig. 1a, Supplementary Fig. 1A, B). Genes with initially low expression (negative values of $\ln(E_{ctrl}/E_{ctrl,avg})$) are upregulated (corresponding to positive values of $\ln(E_{oct}/E_{ctrl})$) while genes with initially high expression (positive values of $\ln(E_{ctrl}/E_{ctrl,avg})$) are downregulated upon octanoate treatment (corresponding to negative values of $\ln(E_{oct}/E_{ctrl})$)[24]. More specifically, there is a clear trend for initially highly expressed genes in the control condition to be downregulated upon octanoate treatment while genes with initial low expression in the control condition were upregulated. Differential expression analysis performed using DESeq2 revealed a total of 2132 upregulated and 632 downregulated genes (FDR = 0.01, |logFC| >1) in the octanoate treated cells (Supplementary Fig. 1C). Pathway enrichment analysis of the differentially expressed genes induced by the 5 mM octanoate treatment was performed and the top 25 upregulated and downregulated pathways are shown in Fig. 1b. Specifically, this analysis revealed that among the top altered biological processes are second messenger mediated signaling, the Notch signaling pathway, adenylate cyclase-activating adrenergic receptor signaling, cell morphogenesis, and differentiation. In contrast, downregulated genes are involved in cell cycle processes, transcriptional regulation of tumor suppressor genes such as p53, and cell cycle checkpoints (Fig. 1b). Additional gene set enrichment analysis (GSEA) investigating top pathways with coordinated upregulation or downregulation of genes demonstrated that the top pathways associated with octanoate treatment included positive regulation of cell morphogenesis, a process involved in differentiation, as well as several oncogenic pathways associated with breast tumorigenesis, including *ERBB*, *WNT*, and *NOTCH* signaling pathways (Fig. 1c). Subsequent leading-edge analysis of these top upregulated signaling pathways- Lipid storage pathways (I), Wnt pathway (II), Notch signaling (III) and ERBB pathway (IV) shows clear association of core enrichment genes with octanoate treatment across replicates (Fig. 1d). Network analysis of octanoate-associated pathways identified by GSEA analysis revealed linked clusters involved with the nervous system and a second, separate group of linked clusters involved with growth factor stimulation, regulation of the *MAPK* cascade, and *ERBB* signaling (Fig. 1e). We validated the expression of a number of genes that GSEA analysis determined were significantly upregulated in MCF-10A with octanoate treatment using real-time qPCR (Fig. 1f). In order to validate our findings in a second cell line, we chose MCF-12A cells. Sweeney et al. provide the history of the establishment of this cell line as well as a demonstration that the cells are non-responsive to estrogen[25]. 1645 genes were upregulated and 330 downregulated (FDR = 0.01) in the octanoate treated MCF-12A. Comparison of octanoate treated MCF-10A and MCF-12A GSEA reveals considerable overlap for Gene Ontology Biological Processes (GOBP, Supplementary Fig. 2A), Reactome gene sets (R, Supplementary Fig. 2B), and KEGG gene sets (K, Supplementary Fig. 2C). Similar to the linked clusters involved with the nervous system seen in the MCF-10As, the octanoate-treated MCF-12A are enriched for gene sets of nerve development, synapses and neurotransmitters, and axons. Additional overlap includes: adenylate cyclase pathways and cell fate specification (upregulated genes, GOBP), cell cycle and cell cycle checkpoints (downregulated genes, GOBP), cardiac conduction, and muscle contraction (R), and MAPK and HEGDEHOG signaling (K). Examination of individual genes in the NOTCH and Wnt pathways listed in Fig. 1d reveals that in MCF-12A cells exposure to octanoic acid increased the expression of DLL4 by 25.4-fold ($p = 1.93E{-}21$, FDR 7.00E−21), that of HEY1 2.07-fold ($p = 7.49E{-}29$, FDR 3.60E−86), NOTCH3 4.75-fold ($p = 1.34E{-}87$, FDR 3.04E−86), WNT11 4.85-fold ($p = 4.56E{-}38$, FDR 2.96E−37) and FZD4 2.29-fold ($p = 3.36E{-}26$, FDR 1.24E−25). Thus, treatment with medium chain fatty acids induces significant changes in transcription.

Evaluating the lipid composition of the serum of ER− and ER+ BC patients

Next, we investigated whether dietary lipids, which are mainly long chain fatty acids (LCFAs), have a similar effect on the gene transcriptional profile to that of MCF-10A cells. In order to determine the specific lipid(s) to evaluate experimentally, we sought to determine the differences in the percent composition of lipid species as a function of ER expression in serum from patients who had donated CUB samples for our original studies[9,10]. A comprehensive lipid profile of these serum samples was performed by the Northwest Metabolic Research Center at University of Washington, with measurement of more than 700 lipids. For each of the measurements, the association between the measured value and ER status was evaluated using regression models, adjusting for BMI, age, and menopausal status. ER was a categorical variable used to describe subjects having ER+ or ER− cancers, or controls undergoing reduction mammoplasty. As the purpose of this experiment was to identify a lipid for ensuing experiments, lipid species were ranked for effect size comparing serum from patients subjects with ER- disease to those with ER+ disease (Supplementary Table 1). There were 28 serum samples from donors with ER− disease and 28 from ER+ donors. Three of the top four lipid species with the largest effect size were noted to contain linoleic acid: cholesterol ester (CE) 18:2, phosphatidyl choline (PC)16:0/18:2, and triacylglycerol (TAG) 54:6-FA18:2 (Table 1). Linoleic acid as a free fatty acid ranked 11th in the analysis. Linoleic acid is the most highly consumed polyunsaturated fatty acid in the human diet[26], its presence in serum CE has been strongly correlated with intake[27], and its concentration in adult adipose tissue has more than doubled in the past half century[28]. Therefore, linoleic acid (LA) was included in subsequent studies.



Octanoic acid and Linoleic acid influence chromatin packing behavior

The state of chromatin is intimately linked with the regulation of gene transcription, undergoing dynamic changes between transcriptionally active and inactive states. Thus, our next step was to explore the changes in chromatin structure of fatty acid treated MCF-10A cells by employing partial wave spectroscopic (PWS) microscopy, which quantifies chromatin packing scaling (*D*)

Fig. 1 Lipid-rich environment enables transcriptional reprogramming in mammary epithelial cells. a Twenty-four hour treatment of MCF-10A cells with 5 mM octanoate results in a completely distinct transcriptional profile compared to untreated controls. E_{ctrl} is the expression of genes in the control condition across all 3 control replicates, $E_{ctrl,avg}$ is the average expression for the control condition across all genes and replicates, E_{oct} is the expression of genes across all 3 octanoate replicates. $E_{ctrl}/E_{ctrl,avg}$ represents the ratio of expression of a particular gene to the average expression across all control cells. Thus, a positive value of $\ln\left(\frac{E_{ctrl}}{E_{ctrl,avg}}\right)$ corresponds to genes that are highly expressed in the control conditions while a negative value of $\ln\left(\frac{E_{ctrl}}{E_{ctrl,avg}}\right)$ corresponds to genes that have an initial lower expression in the control condition. E_{oct}/E_{ctrl} represents the ratio of expression of a particular gene for octanoate-treated versus vehicle control-treated cells. Genes with initially low expression are upregulated while genes with initially high expression are downregulated upon octanoate treatment. **b** Gene ontology analysis of differentially expressed genes induced by octanoate treatment. Upregulated and downregulated genes were first identified using DESeq2 (FDR < 0.01, |logFC| > 1) for 5 mM octanoate treated cells compared to vehicle-treated control cells. Pathway enrichment analysis was performed on identified differentially expressed genes with annotations from online pathway databases (KEGG, Hallmark, Canonical Pathways, Reactome, BioCarta) and Gene Ontology Biological Processes. Pathway enrichment was ranked by *p*-value on a $-\text{Log}_{10}$ scale and a selection from the top 25 pathways associated with upregulated genes (in red) and downregulated genes (in blue) are shown. **c** GSEA analysis of Gene Ontology Biological Processes showing top pathways associated with octanoate treatment with FDR < 0.1 related to differentiation, cell signaling, and metabolic processes. **d** List of core enrichment genes differentially expressed in treated replicates-T4, T5, T6 versus control replicates- C1, C2, C3: (I) Lipid storage pathways (II) Wnt pathway (III) Notch pathway (IV) ERBB pathway, each pathway as identified by GSEA leading edge analysis. Expression values are represented as colors and range from red (high expression) to dark blue (lowest expression). **e** Network analysis of pathways associated with the octanoate phenotype in GSEA analysis of Gene Ontology Biological Processes. **f** qPCR analysis of genes associated with the NOTCH pathway (mean ± s.d.). Two genes, NOTCH3 and DLL4 show remarkable upregulation upon 5 mM octanoate treatment compared to other identified genes such as NOTCH1. Statistical significance was determined by the unpaired t-test with Welch's correction (**$P < 0.01$, *$P < 0.05$).

Table 1. Lipid species in the serum of CUB patients ranked by effect size, ER negative compared to ER positive.

Order	Variable	ER_Negaitve_Effect_compared_to_ER_Positive
1	CE(18:2)	97.76
2	PC(16:0/18:2)	51.96
3	SM(20:0)	42.09
4	TAG54:6-FA18:2	28.78
5	SM(22:0)	23.33
6	PC(16:0/18:1)	23.13
7	CE(16:1)	20.32
8	TAG54:5-FA18:2	16.57
9	SM(22:1)	16.01
10	CE(18:3)	15.51
11	FFA(18:2)	15.37
12	PC(18:0/18:2)	13.77
14	PC(18:1/18:2)	12.38
14	TAG52:4-FA18:2	11.37
15	TAG54:5-FA18:1	10.83

Abbreviations: *CE* cholesterol ester, *FFA* free fatty acids, *PC* phosphatidylcholine, *SM* sphingomyelin, *TAG* triacylglycerol. TAGs have 3 acyl chains, but it is only possible to measure the length and number of double bonds of 1 of them. For example, TAG54:6-FA18:2 has 1 chain that is an 18:2 FA and the other two have a total of 54 carbons and 6 double bonds. The whole TAG has 72 carbons in the chain (plus 3 from the glycerol backbone) and a total of 8 double bonds.

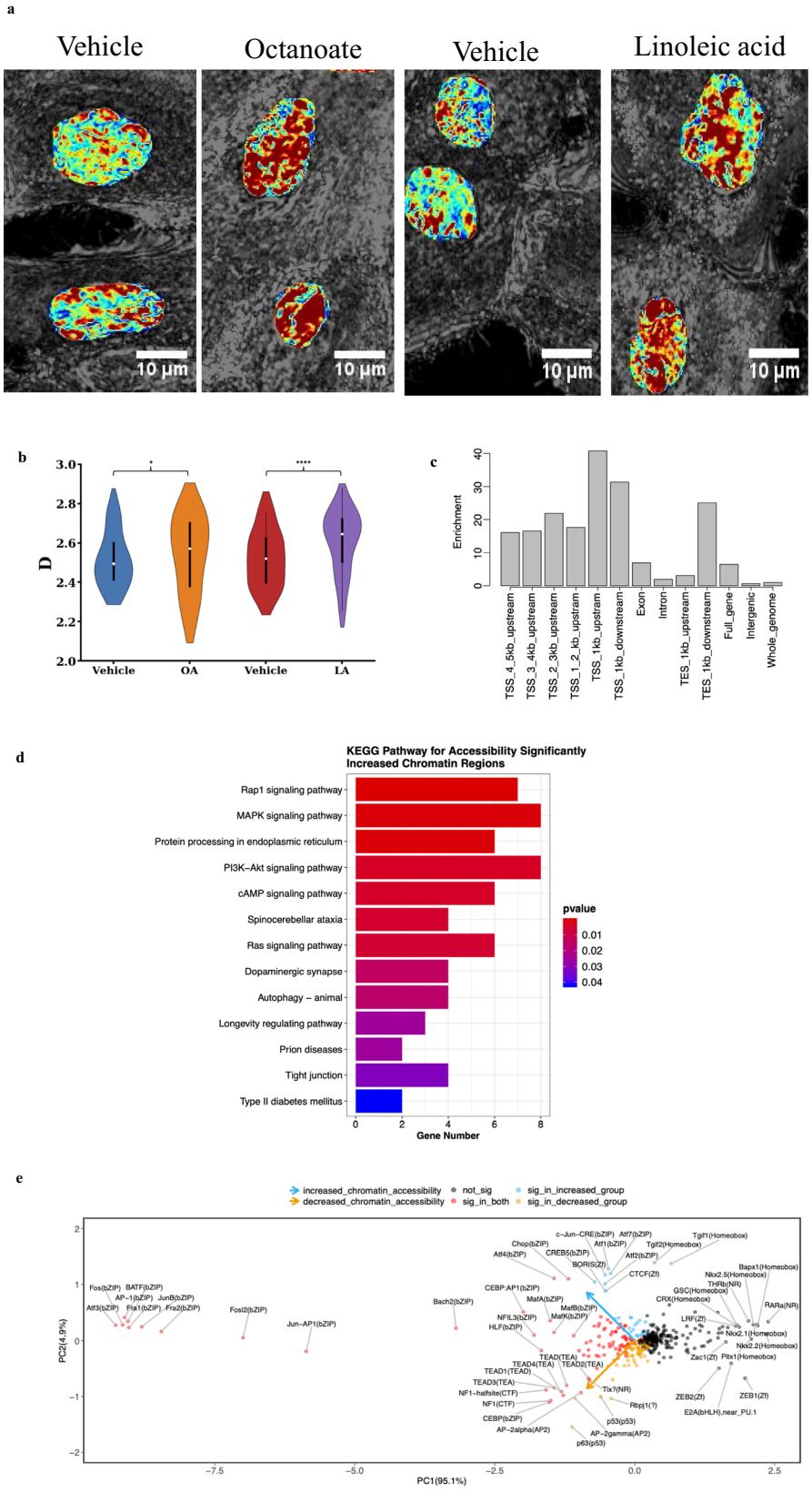
in live cells[29]. *D* represents the power-law scaling relationship between the 1D size of the chromatin polymer i.e., the number of nucleotides and the 3D space the chromatin polymer occupies. Recent evidence indicates that higher chromatin packing scaling is associated with increased intercellular and intra-network transcriptional heterogeneity as well as increased malignancy and chemoresistance in cancer cells[24,30,31]. PWS was used to evaluate the effect of OA and LA on chromatin packing scaling in live MCF-10A cells. Images were obtained every 6 h over a 24 h period. Our results showed significant increases in chromatin packing scaling upon exposure to lipids suggesting that there is an increase in the dynamic range of gene expression and transcriptional gene network heterogeneity (Fig. 2a, b). These significant changes in chromatin packing behavior also indicate significant changes in chromatin accessibility, which is directly associated with chromatin structure[32].

ATAC sequencing reveals increased chromatin accessibility in regulatory regions of genes in the MAPK and cAMP signaling pathways in lipid treated mammary cells

To acquire more detailed insight into the specific regions of open chromatin that were made accessible by LA treatment, we proceeded with ATAC sequencing on LA-treated MCF-10A cells. We examined the genomic locations of ATAC-seq peaks, representing open chromatin sites, and discovered 1704 open chromatin sites. Open chromatin regions were overrepresented within 1 kb of transcription start sites (TSSs) by 40-fold relative to the whole genome (Fig. 2c). Further, KEGG pathway analysis revealed 326 open chromatin regions with a log fold change >= 1.5 and FDR < 0.05 compared to vehicle treated cells. Among the top pathways that were upregulated significantly upon LA treatment are MAPK signaling pathway, PI3K-AKT signaling pathway, and the cAMP adenylate cyclase pathway (Fig. 2d). Additionally, motif analysis conducted using 'HOMER'[33] showed that chromatin regions made accessible/inaccessible by LA treatment have binding motifs for a number of transcription factors (Fig. 2e). These data reveal that linoleic acid affects chromatin heterogeneity and increases/decreases the accessibility of specific regions that include transcription factor binding sites.

Notch pathway genes are overexpressed in patients at high risk of ER- disease

Next, we sought to determine whether the genes, or sets of genes/pathways that we identified in our in vitro study were also differentially expressed in vivo in tissue of patients at risk for ER− and ER+ breast cancer. We took advantage of RNA from the CUB of breast cancer cases utilized in our previous studies, which revealed the association of LiMe genes in the CUBs of women with unilateral ER- breast cancer[9,10]. We combined the data from the RNA and ATAC sequencing experiments and collated a list of 44 genes of interest and 3 housekeeping genes. The list consists of the genes from the *HEDGEHOG*, *NOTCH*, *WNT*, *EMT*, *PPARγ*, and



adenylate cyclase pathways (Supplementary File 1). TaqMan low density arrays were utilized to measure the expression of these genes in CUBs of ER− and ER+ cases compared with the reduction controls. The study population included 84 women, with participants comprised of 28 matched triplets of women with ER-positive breast cancer, ER-negative breast cancer, and reduction mammoplasty controls. The three groups were matched by age, race, and menopausal status as shown in Supplementary

Fig. 2 Linoleic acid alters large-scale chromatin packing behavior in MCF-10A cells. a Representative PWS microscopy images of MCF-10A cell nuclei at 24 h after treatment with vehicle controls and lipids—octanoate and linoleic acid. Scale bars, 10 μm. Chromatin packing scaling *(D)* map of nuclei shows an increase in chromatin packing scaling upon lipid treatment as demonstrated by an increase in red regions. **b** Changes in average chromatin packing scaling among MCF-10A cells upon treatment with vehicle controls and lipids compared to untreated cells. Significance was determined using unpaired Kolmogorov–Smirnov *t*-test (****$P < 0.0001$, *$P < 0.05$). Bar graphs show the mean change in intranuclear *D* across cell populations for $N = 88$ cells PBS (vehicle for octanoate), $N = 110$ cells Octanoate (OA), $N = 103$ cells BSA (vehicle for linoleic acid), and $N = 94$ Linoleic acid (LA). **c** Enrichment of genomic locations for 1704 open chromatin regions (FDR < 0.05, logFC > 1) in LA treated MCF-10A cells. The enrichment of peaks in each type of genomic region relative to the whole genome is shown on the *y*-axis. Two ATAC-seq libraries were used for the analysis. **d** Pathway analysis for the regions with increased chromatin accessibility in linoleic acid-treated cells identified using the KEGG database. **e** Biplot showing changes in chromatin accessibility for specific regions identified by HOMER analysis. Motifs with a significant increase in the chromatin accessibility are shown in blue and those with a significant decrease in accessibility are shown in yellow (FDR < 0.05 and |logFC| > 1).

Fig. 3A. As noted in our original publication, ANOVA revealed a significant difference in BMI across the three groups with BMI in the reduction mammoplasty control group (30.0 ± 5.8) notably higher than in ER-negative cases (25.3 ± 6.3, $p = 0.015$), but not significantly higher than in the ER-positive group (26.7 ± 5.5, $p = 0.136$)[10]. There was no significant difference in HER2 status between ER-positive and ER-negative cases. The majority of the selected genes had higher expression in high-risk CUB specimens than the controls, irrespective of the ER status of the index tumor (Supplementary Fig. 3B). The comparison between the ER− and ER+ CUBs revealed that in the ER− CUBS there is increased expression of genes that function in the Notch pathway: *NOTCH1* (1.7-fold, $p = 0.002$, BH_adjP = 0.07), *NOTCH4* (1.5-fold, $p = 0.04$, BH_adjP = 0.3), *DLL1* (1.2-fold, $p = 0.07$, BH_adjP = 0.4) and *HEY 1* (1.5-fold, $p = 0.05$, BH_adjP = 0.3), in addition to the *SMO* gene (1.5-fold, $p = 0.05$, BH_adjP = 0.3), which is a key component of the hedgehog signaling pathway (Fig. 3). Comparing ER− to control, increased expression was observed for GPR161 (1.7-fold, $p = 0.05$, BH_adjP = 0.7), which plays a role in the Hedgehog pathway via cAMP signaling, and *IGF2* (2.8-fold, $p = 0.07$, BH_adjP = 0.7), which signals via both the MAPK and PI3K-AKT pathways. Altogether, these data reveal upregulation in *NOTCH* signaling in benign breast tissue samples from women at risk for ER− disease, suggesting that dysregulation of these pathways may play a role in the early stages of ER- cancer development.

LA increases the expression of Notch pathway genes and specific genes involved in fatty acid oxidation in vitro

The increased expression of Notch pathway genes we discovered in the ER- CUBs, along with the similar findings in MCF-10A cells exposed to octanoate (described above), led us to test the hypothesis that long chain fatty acids have similar effects on gene expression. We, therefore, investigated whether an increased LA environment influences the expression of Notch pathway genes and specific genes involved with FAO in vitro. We treated MCF-10A cells and mammary organoids from reduction mammoplasty patient samples with LA for 24 h and then quantified changes in gene expression using RT-qPCR. To begin with, we assayed the genes involved in the activation of FAO. Upon entering cells, free fatty acids are converted into fatty acyl-CoA molecules by the enzymes of the acyl-CoA synthetase (ACS) family[34]. Notably, acyl-CoA synthetase long chain (*ACSL3*) is one of the LiME genes found to be upregulated in high-risk ER- CUBs samples. Generation of acetyl-CoA occurs through a cyclical series of reactions in which a fatty acid is shortened by two carbons per cycle, eventually generating acetyl co-A. Acetyl co-A is a substrate for ketogenesis, which is initiated by the mitochondrial enzyme 3-hydroxy-3-methylglutaryl-CoA synthase 2 (*HMGCS2*), another of the previously identified LiMe genes. The mechanism for LCFAs oxidation is slightly more complex than for MCFAs, as this is regulated primarily via the enzyme carnitine palmitoyltransferase 1 (*CPT1*), the rate-limiting enzyme of FAO which enables transport into the mitochondria. As shown in Fig. 4a, the expression of *HMGCS2*, *ACSL3*, and *CPT1B* were increased by LA exposure in MCF-10A cells and mammary organoids. Additionally, we observed a significant increase in *DLL4* expression followed by *HEY1*, *HEY2,* and *NOTCH1* in the lipid-treated mammary cells (Fig. 4a). We revisited the ATAC sequencing data to examine the effect of LA on chromatin architecture near key genes in the *DLL4*/*NOTCH* signaling pathway and observed increased accessibility around the transcription start sites of *DLL4*, *NOTCH1,* and *HEY1* showing significant lowered chromatin density with *p*-values of 1.62e−17, 0.017 and 0.03 respectively (Fig. 4b, c).

The NOTCH signaling pathway is activated in vitro by octanoic acid treatment

Intracellular Notch binds to the transcriptional repressor RBP-Jk in the nucleus, thereby converting it into an activator and inducing the expression of downstream target genes. Therefore, to determine if the NOTCH pathway is activated by OA, we transfected a RBP-Jk reporter construct into MCF-10A cells. The LUC/REN ratio is increased 2-fold by exposure to OA (Fig. 5) indicating that the NOTCH pathway is functionally activated by the lipid.

Fatty acids drive flux through metabolic reactions resulting in increased histone methylation

While most of the experiments reported by McDonnell et al. were performed in AML 12 liver cells, these investigators also demonstrated increased H3K9 acetylation in octanoate-exposed MCF7 and MDAMB-231 breast cancer cells[35]. Therefore, we sought to determine if these same experimental conditions would lead to H3K9 acetylation in a non-malignant MCF-10A cells. We exposed MCF-10A non-transformed ER - breast epithelial cell line to 5 mM octanoate (OA) for 24 h in medium containing both glucose (1.441 g/L) and glutamine (0.292 g/L). Western blot analysis demonstrated that octanoate exposure of MCF-10As resulted in increased acetylation at both H3K9 and H3K14 (Fig. 6a). To demonstrate that this was a fatty acid-specific effect, we treated the cells with 1,4-Cyclohexanedimethanol (1,4-CHDM), an alcohol with the same formula as octanoate; no acetylation was observed consequent to the alcohol exposure (Supplementary Fig. 4A). To validate the specificity of the antibody against the acetylated histone lysines, we treated MCF-10A cells with sodium butyrate, a histone deacetylase (HDAC) inhibitor. Sodium butyrate treatment increased the acetylation of H3K9 and H3K14 as shown in Supplementary Fig. 4B.

To exhaustively explore the impact of octanoate treatment on metabolic pathways, we used flux balance analysis (FBA)[36]. FBA makes use of genome-scale metabolic network models that contain all known metabolic reactions in a cell or tissue based on evidence from the published literature[37]. Genome-scale metabolic models have been widely used to predict the metabolic behavior of various mammalian cell types[38–42]. Here we used the Recon1 human network model that maps the relationship between 3744 reactions, 2766 metabolites, 1496 metabolic genes, and 2004

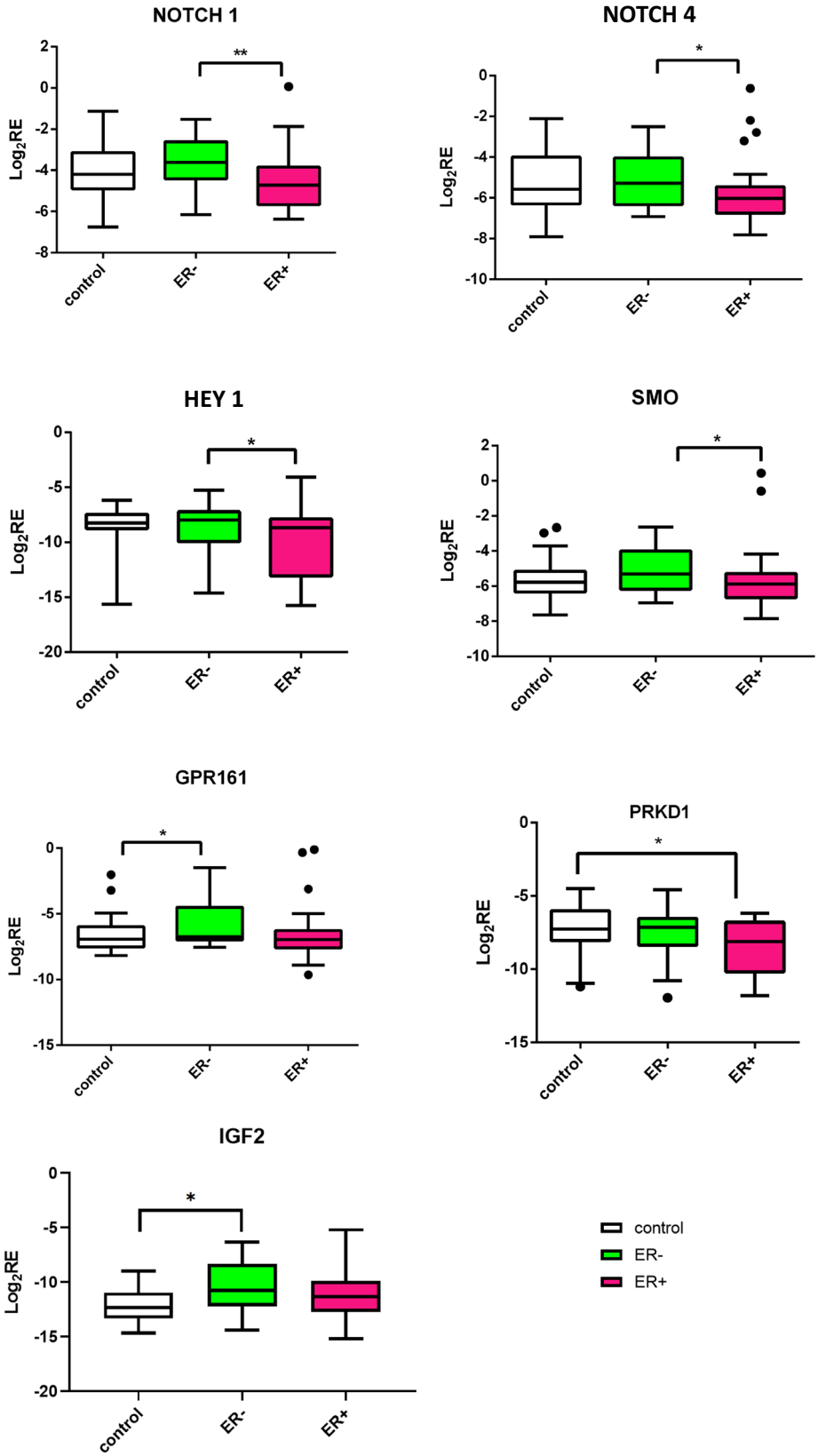
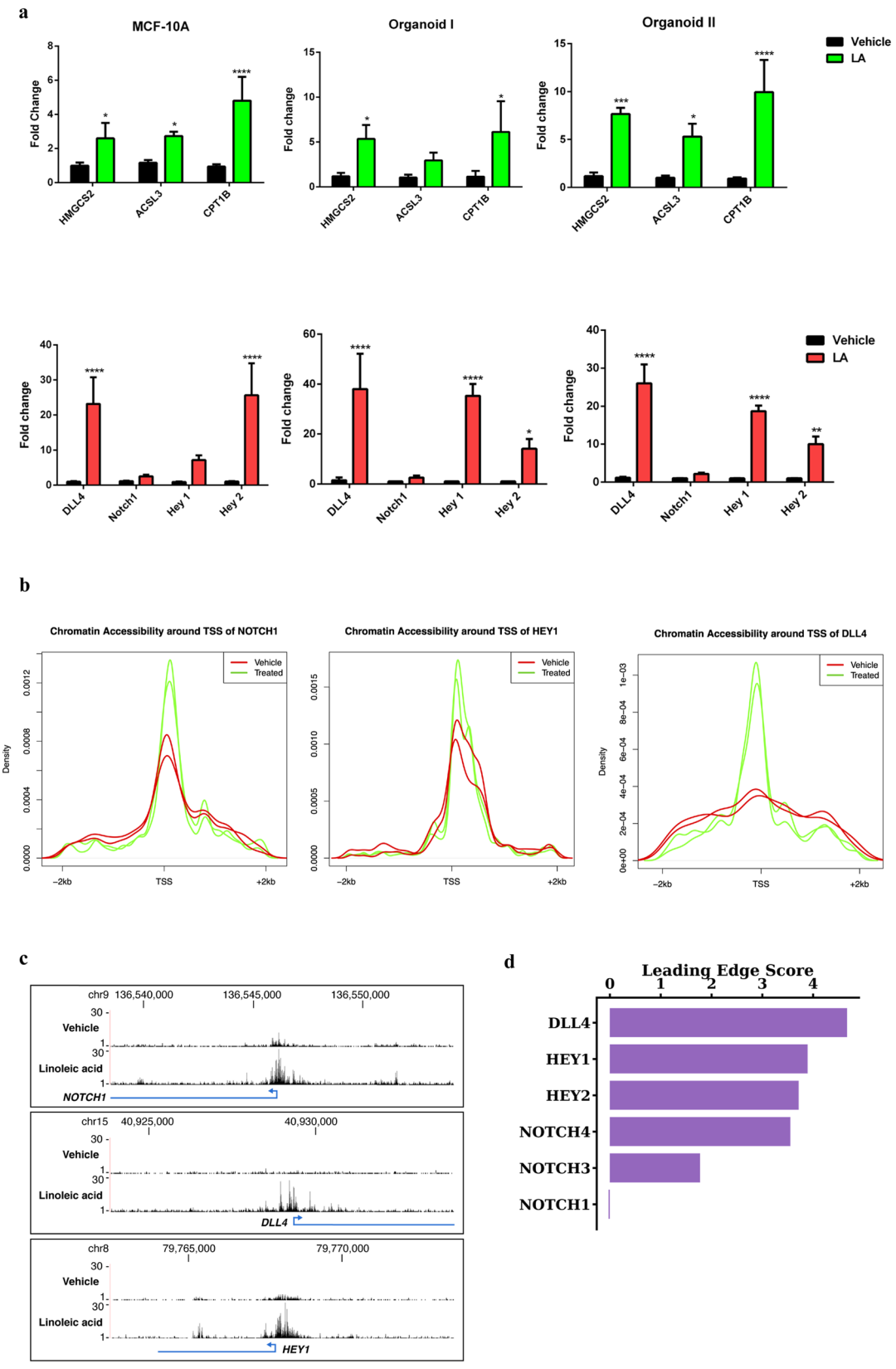


Fig. 3 Notch pathway is overexpressed in CUB samples of patients at high risk of ER− disease. Expression of genes from various pathways in matching CUBs from ER-negative, ER-positive patients, and controls. The log2-transformed relative (log2RE) amounts of mRNA expression normalized to the housekeeping gene and expressed as $\log_2 2^{-(\mathrm{CtX}-\mathrm{CtGAPDH})} = -(\mathrm{CtX} - \mathrm{Ct\ GAPDH})$ where Ct is threshold cycle and X is gene of interest. IGF2 and GPR161 were significantly higher in ER-negative versus control. Genes from the Notch pathway were significantly higher in ER negative CUBs in comparison to ER positive patients. Mann–Whitney test was used to test the pairwise differences between the samples (ER+, ER−, Control) * $P < 0.05$; ** $P < 0.01$. Boxplots show mean and SEM with whiskers indicating 1–99th percentile.



metabolic enzymes[43]. This model was augmented with biochemical reactions corresponding to histone acetylation and methylation[38,44], allowing us to predict the consequences of octanoate-induced metabolic changes on histone modifications by tracking the flux through the substrates for the histone modifications. These models were previously used to predict bulk histone acetylation levels in various cell lines based on the nuclear flux of acetyl-coA directed towards histone acetylation[44]. Similarly, bulk

Fig. 4 Increased DLL4/Notch signaling is associated with the stimulated fatty acid oxidation. a qPCR data showing increase in lipid metabolism genes (green) and Notch pathway genes (red) after 24 h linoleate treatment in MCF-10A and mammary organoids (mean ± s.d.). Organoid I was donated by a postmenopausal 61-year-old with a BMI of 22 and Organoid II by a premenopausal 28-year-old with a BMI of 31. Statistical significance was determined by the unpaired t-test with Welch's correction (****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$). **b** Chromatin accessibility in the lipid-treated cells around the transcription start site (TSS) of NOTCH1, HEY1, and DLL4 (FDR < 0.001). **c** Gene tracks and increase in peaks for the Notch genes in LA treated cells with the exact location on the chromosome. **d** Leading edge scores for genes of interest associated with the NOTCH signaling pathway as determined by GSEA leading edge analysis. DLL4, HEY1, HEY2, NOTCH3, and NOTCH4 were identified as core enrichment genes in the NOTCH pathway.

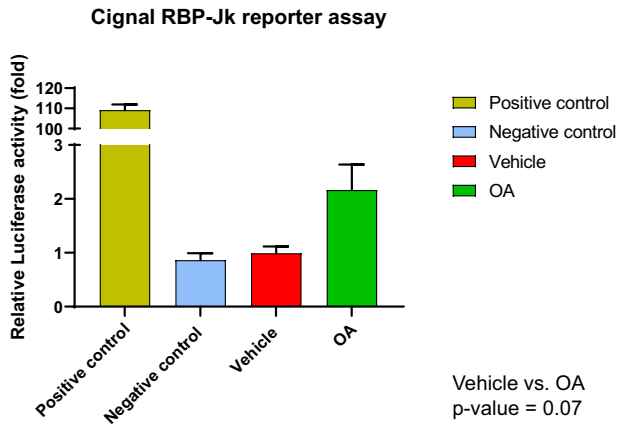


Fig. 5 Effects of OA on Notch signaling. NOTCH transcriptional activity was measured using the Cignal RBP-Jk reporter assay following exposure of MCF-10A cells to 5 mM octanoic acid (OA) for 24 h. Luciferase levels were normalized to Renilla luciferase. The results were plotted as fold change with respect to the untreated. ($n = 3$, mean ± SEM). The p-value was calculated by unpaired t-test.

histone methylation levels can be predicted based on the nuclear flux of S-adenosyl-L-methionine (SAM)[38]. The model predicted octanoate treatment would result in increased histone methylation levels, with a more modest increase in histone acetylation levels (Fig. 6c). As a comparison, we repeated this analysis with immortalized hepatocyte cells used by McDonnel et al.; they found a significant increase in histone acetylation after octanoate treatment[35]. We calculated metabolic flux in these hepatocytes using the transcriptomics data from McDonnel et al and found a much larger increase in histone acetylation after octanoate treatment (Supplementary Fig. 4E). These results suggest that the impact of metabolic alterations on histone acetylation is cell-type specific, as observed in prior studies[45,46]. Overall, out of the 3759 reactions in the model, we identified 38 that showed significant increased activity after octanoate treatment (p-value < 0.01; Supplementary Fig. 4C). As expected, reactions involved in lipid and fatty acid metabolism, specifically triacyl glycerol synthesis and glycerophospholipid metabolism were upregulated. Interestingly, among the upregulated reactions were several reactions related to the one-carbon metabolic pathway, which links folate, SAM, methionine, glycine, and serine metabolism (Fig. 6f). The reactions catalyzed by methionine adenosyltransferase, methionine synthase, adenosyl homocysteinase, 5,10-methylene-tetrahydrofolatereductase, glycine N-methyltransferase, and formyltetrahydrofolate dehydrogenase were all predicted to have increased activity after treatment (p-value < 0.01). These reactions likely support increased histone methylation by providing one carbon units. Examining the reaction fluxes/activities in the OA-treated MCF-12A cells (Fig. 6f), we see differences in one carbon metabolism and glutathione metabolism similar to what we observed in the MCF-10A cells.

Lipid exposure eventuates in histone methylation. In order to profile the specific histone marks significantly changed by the octanoic and linoleic acid treatment, we performed liquid chromatography/mass spectrometry on tryptic peptides isolated from the nuclei of treated and control MCF-10A cells. Increased methylation in the OA treated cells was observed in various histone proteins including H3K9me1/2/3, H3.1K27me2/3, H3.3K36me2/3, H3K79me1/2, and H3K4 (Fig. 6e) together with increased acetylation of H3K14 and H4K16 (Fig. 6d). Similarly increased methylation was observed for H3.1: K27me1, H3.3: K23me1, H3.1: K36me3, H3.3: K36me2, and H3.3: K36me3 in LA treated cells (Fig. 6e). Notably, the GSEA analysis showed a significant correlation of H3K27 methylation (NES = 2.47, FDR q-value = 0.05) and H3K4 methylation (NES = 1.24, FDR q-value = 0.1) with octanoate treatment (Supplementary Fig. 4D) suggesting this lipid-rich environment eventuates in histone methylation in mammary epithelial cells.

DISCUSSION

The known determinants of risk for ER-negative breast cancer are genetic (either specific racial inheritance, germline mutations in genes such as BRCA1) or systemic/behavioral factors (premenopausal obesity[47], absence of a breastfeeding[48]). In contrast, few if any local factors in the breast environment serve to identify women at risk for ER-negative tumors. Local in-breast factors are of great interest, however, since they may be more specifically targetable for breast cancer prevention than systemic factors. Of note, the two strongest risk factors for breast cancer overall (other than high penetrance germline mutations) are local: atypical proliferative lesions, and[49] extremely dense breast tissue[50]. This reasoning motivated us to investigate the local breast biology that may promote the development of ER-negative rather than ER-positive breast cancer, using the CUB of women undergoing surgery for a unilateral primary breast cancer as a model for ER-specific breast cancer risk[7,51]. In our initial study, we identified a highly correlated lipid metabolism (LiMe) gene signature, which was enriched in the CUBs of women with ER- breast cancer.

To explain the biologic basis for this association, we developed an in vitro model wherein we exposed MCF-10A and MCF-12A, ER-negative, non-tumorigenic epithelial cell lines, or breast organoids derived from reduction mammoplasty samples to an extracellular milieu rich in medium or long chain fatty acids. This model system has now enabled us to demonstrate that the exposure of breast epithelial cells to these fatty acids results in a dynamic and profound change in gene expression, accompanied by changes in chromatin packing density, chromatin accessibility, and histone PTMs. The histone modifications, in turn, are the result of both the lipid-engendered increased expression of the requisite enzymes and the increased production of their substrates. Our metabolic flux analysis revealed the upregulation of several reactions related to the one-carbon metabolic pathway, which links folate, SAM, methionine, glycine, and serine metabolism. This insight was not evident upon analysis of differential gene expression, which is not surprising as gene expression changes often do not reflect the flux of metabolic reactions[38]. The substrates for histone methylation and acetylation reactions often have cellular concentrations that are commensurate with enzyme Km values, and thus these reactions are sensitive and responsive to changes in metabolism[21].

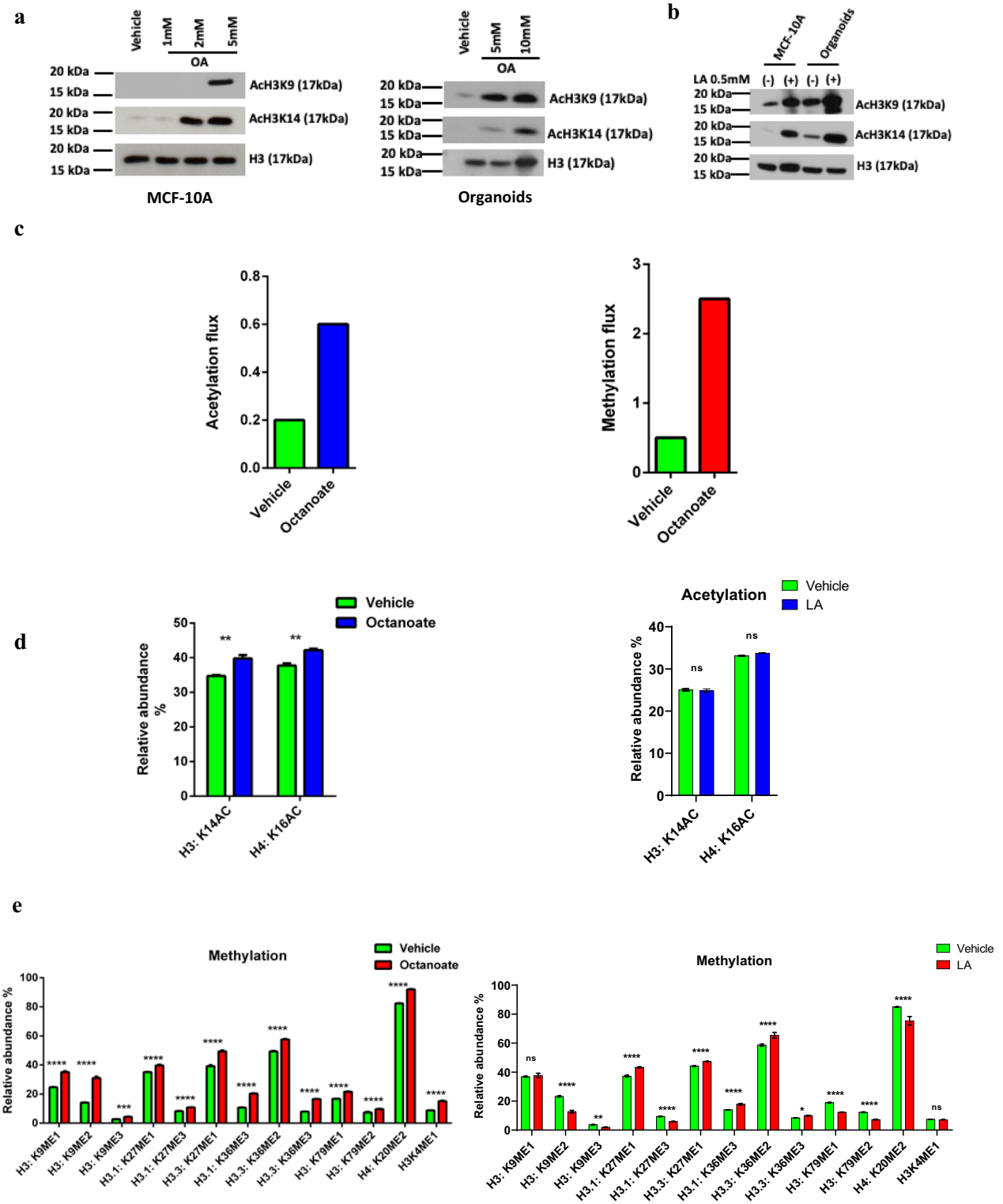


Fig. 6 Fatty acids drive histone modifications and metabolic flux. Western blot of histone acetylation at H3 lysine K9 and K14 in MCF-10A cells and organoids treated with **a** octanoate and **b** linoleic acid. **c** The effect of octanoate treatment on histone acetylation and methylation flux in MCF-10A cells predicted using genome-scale metabolic modeling. Proteomic acetylation (**d**) and methylation (**e**) profiling measured by mass spectrometry of MCF-10A cells treated in triplicate with 5 mM octanoate for 24 h in a complete media compared to vehicle (left) and 0.5 mM linoleate for 24 h in complete media compared to vehicle (right) (mean ± s.d.). Two-way ANOVA was performed to determine the statistical significance and corrected for multiple comparisons using Sidak test (****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, ns not significant). **f** Heatmap of reaction flux differences predicted by metabolic modeling to be differentially active (p-value < 0.01) between control and treatment (increased flux in red, decreased flux in blue). The corresponding pathways (subsystem) that each reaction belongs to is listed in the legend.

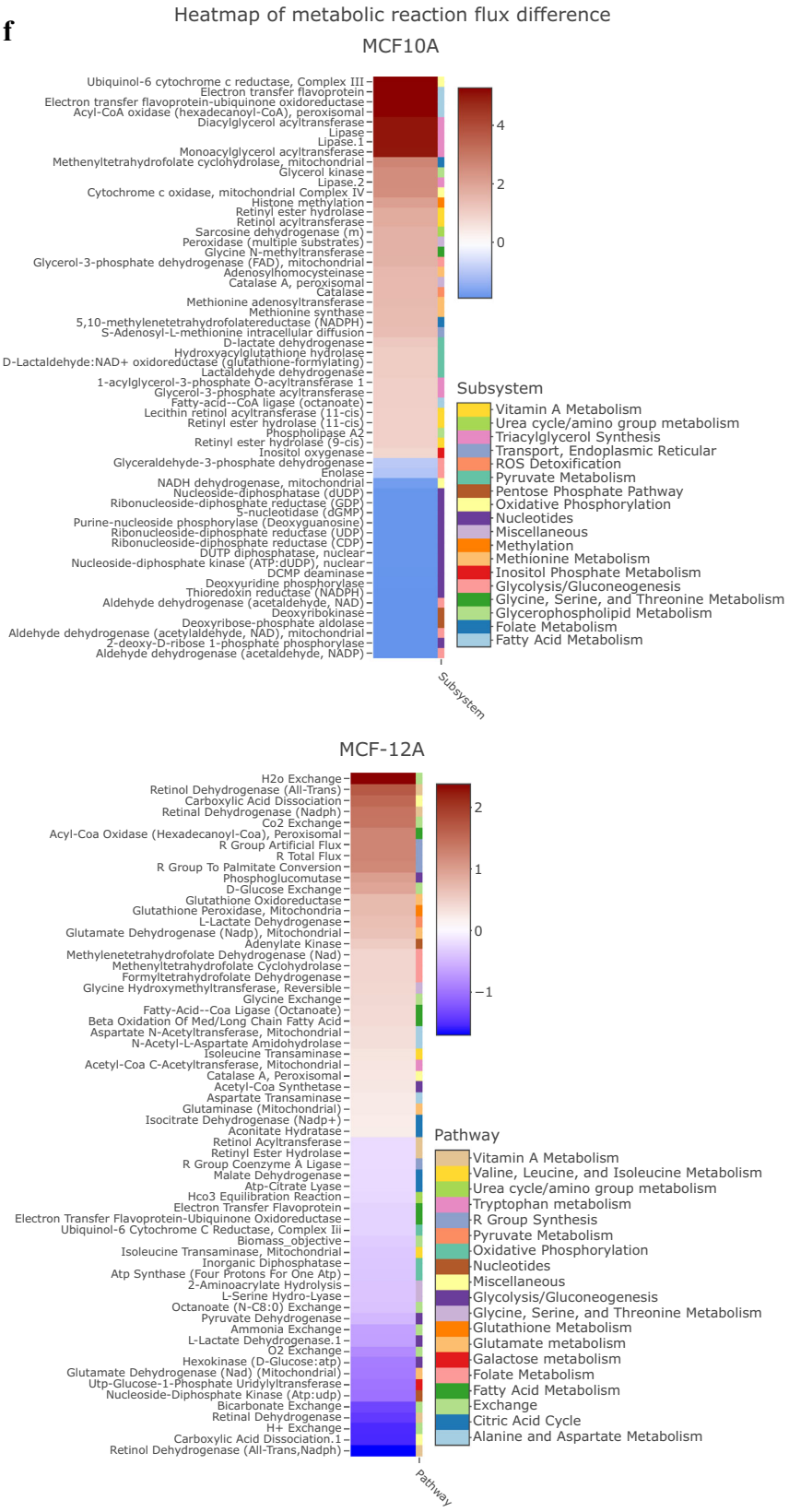


Fig. 6 Continue.

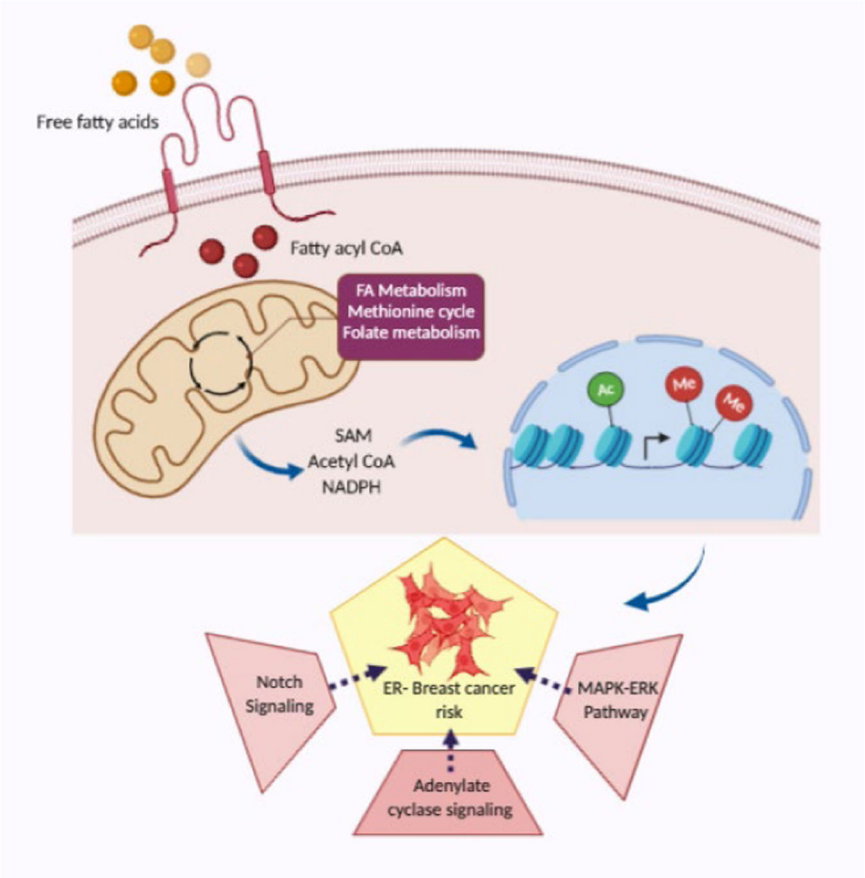


Fig. 7 Proposed model illustrating the orchestration of lipid-induced molecular changes. *Sensors*: Senses the fatty acid-rich environment and perturb cellular metabolism providing the essential substrate for histone modifications and thereby turning on the *Mediators*- histone PTMs, which consequently activates the *Effectors*- Notch, adenylate cyclase, and MAPK-ERK the key protein signaling associated with ER− breast cancer.

Increasing substrate concentration can increase the product of the reaction even if there is no increase in the expression of the enzyme. Our proteomics data reveal increased methylation at H3K27me1 and H3K36me2/3 in cells treated with octanoate in both MCF-10A and MCF12 A, and H3K4me1 in MCF-10A cells; GSEA analysis showed that genes with ontologies related to histone methylation at H3K27 and H3K4 exhibit changes in expression in the lipid-treated cells.

The goal of our investigation was to develop specific mechanistic explanations as to why lipid metabolism pathways would aid ER− breast cancer development. The data have revealed a number of possibilities, all of which will have to be explored further. Mammary stem cell differentiation is a hierarchical organization, and lineage tracing experiments have determined that *NOTCH1* expression exclusively generates ER− luminal cells[52]. A subsequent study by these investigators revealed that during mammary embryogenesis Notch signaling prevents the generation of basal precursors, and cells expressing active *NOTCH1* exclusively give rise to the ER− (Sca1-/CD133-) lineage at any developmental stage from mouse embryonic day 13.5 to postpartum day 3[53]. Even more interesting given our focus on the origins of ER-negative breast cancer was their observation that pubertal cells retain plasticity. Ectopic activation of Notch1 in basal cells at puberty was able to completely switch their identity to ER-negative luminal cells.

Additional clues regarding the association of our experimental findings with ER-negative breast cancer comes from GWAS data. A study that included 21,468 ER-negative cases and 100,594 controls identified independent associations of ten single nucleotide polymorphisms (SNPs) with the development of ER− breast cancer[54]. Pathway analysis was performed by mapping each SNP to the nearest gene. This identified several pathways implicated in susceptibility to ER-negative, but not ER+ breast cancer. Included among these was the adenylate cyclase (AC) activating pathway. One of the significantly altered biologic processes that we identified by RNA sequencing of the octanoic acid-treated cells is adenylate cyclase-activating adrenergic receptor signaling. Adenylate cyclase signals via cyclic AMP. Regions of chromatin with increased accessibility are associated with increased gene expression; our ATAC-Seq results show that linoleic acid exposure significantly increased accessibility to genes in the cAMP signaling pathway. In their discussion of ER− GWAS results, Milne et al. suggest that stimulation of the beta 2 adrenergic-adenylate cyclase-cAMP-β-arrestin–Src–ERK pathway may play a role in the genesis of ER− breast cancer. MetaCore analysis of our RNA-sequencing data reveals similar pathway activation, however, it is the beta1 adrenergic receptor that demonstrates increased expression in the octanoate treated cells. In addition, our ATAC-seq data showed increased RAP1 signaling pathway accessibility. Adenylate cyclase signaling also functions via Epac-Rap1-B-raf-MEK-ERK, with this signaling shown to be responsible for

sustained *ERK* activation that occurs 10–30 min after cAMP activation[55]. The MAPK (ERK) pathway can be stimulated by means other than adrenergic receptor-ligand binding. Activation of this pathway by overexpression of *EGFR* + *EGF*, *c-erbB-2*, *RAF1*, or *MEK* in MCF7 cells leads to estrogen-independent growth and downregulation of ERα expression[56]. These results suggest that hyperactivation of the MAPK(ERK) pathway plays a role in the generation of the ER− phenotype in breast cancer. We observed *MAPK* activation in our analysis of differentially expressed genes, i.e., "positive regulation of the *MAPK* cascade," and in the analysis of regions of chromatin with significantly increased chromatin.

Using stratified LD score regression, a statistical method for identifying functional enrichment from GWAS summary statistics, SNPs associated with the H3K4me3 histone mark were determined to be contributing to the heritability of ER-negative breast cancer, (2.4-fold, $P = 0.0005$)[54]. Increased activity of the one-carbon pathway is associated with increased H3K4 trimethylation in stem cells and cancer cell lines[38,57]. Restriction of methionine with consequent modulation of SAM and S-Adenosyl-L-homocysteine (SAH) levels affects methylation at H3K4me3, H3K27me3, and H3K9me3, with H3K4me3 exhibiting the largest changes (45). Interestingly, this restriction leads to loss of H3K4me3 at the promoters of colorectal cancer (CRC)-associated genes, with resulting decreased expression ($p = 0.02$, Fisher's exact test). A computational model developed to identify the direct influences on methionine concentrations in humans suggests that dietary intake explains about 30% of the variation in methionine concentration, and fats (arachidic acid in this model) are among the foods contributing to higher methionine levels[57].

In conclusion, we have demonstrated in the present study that exposure of breast epithelial cells in vitro to fatty acids results in epigenetic effects that produce dynamic and profound changes in the expression of genes that have been associated with the development of ER- breast cancer (Fig. 7). Next steps include demonstrating that these same changes are observed in vivo. As mentioned in the introduction, polyunsaturated fatty acids are present in normal breast tissue. Although we measured lipid species in the serum of the donors of the CUB specimens, fatty acids can also be mobilized from adjacent adipose tissue; adipocytes have been shown to be a reservoir of lipids for breast cancer stem cells[58]. We hypothesize that the expression of genes associated with the development of ER- breast cancer is consequent to lipid stimulation of one-carbon metabolism with resultant changes in histone methylation. Important roles for glycolysis, glutaminolysis, lipogenesis, and mitochondrial activity have been demonstrated in oncogenesis; the one-carbon pathway has comparatively received less attention and the insights we provide here generate new questions regarding lipid metabolism and ER-negative breast cancer, to be pursued in future investigations.

MATERIALS AND METHODS

Cell culture

MCF-10A and MCF-12A cell lines were obtained from American Type Culture Collection (ATCC) and cultured in mammary epithelial cell growth basal medium with single quots supplements and growth factors (#Lonza CC-4136). Cells were treated with the medium-chain fatty acid sodium octanoate (OA; Sigma # C5038) dissolved in PBS; and long-chain fatty acid Linoleic acid (LA; Sigma # L8134) complexed with fatty acid-free BSA (Roche 10775835001). Alternatively, Linoleic Acid bound to BSA (LA-BSA; Sigma # L9530) was used. PBS and BSA were used as the vehicle control in experiments containing OA and LA, respectively. Cells were counted using an Invitrogen Countess automated cell counter using the Trypan blue exclusion method and seeded at the indicated densities. All experiments were done in complete MEBM media with fatty acids or vehicle.

CUB samples

Patients diagnosed with unilateral breast cancer and undergoing contralateral prophylactic mastectomy at Prentice Women's Hospital of Northwestern Medicine were recruited under an approved protocol (NU11B04), with exclusions for neoadjuvant treatment, prior endocrine therapy, or pregnancy/lactation during the prior 2 years. A group of reduction mammoplasty (RM) patients were also recruited as standard risk controls. All participants provided written informed consent. The fresh tissues were frozen and stored in liquid nitrogen. Tissue samples from 56 bilateral mastectomy cases (28 ER+ and 28 ER−) and 28 healthy RM controls were used in this study. The ER+ cases, ER− cases, and controls were matched by age, race, and menopausal status.

Mammary organoids preparation

Tissues were collected from women admitted for reduction mammoplasty, who were recruited under an approved IRB protocol (NU15B07). All participants provided written informed consent. Breast tissue to be processed is transferred into a sterile petri dish and chopped into small pieces using a scalpel. The minced tissue was transferred to a sterile 50 ml tube and 30 ml of Kaighn's Modification media (Gibco #21127022) containing collagenase from Clostridium histolyticum (Sigma Aldrich, #C0130) was added, final collagenase concentration is 1 mg/mL. Media containing collagenase is filtered using a 0.22 µm filter. The Falcon tube is sealed with parafilm and tissue is gently dissociated on a shaker at 100 rpm and 37 °C, overnight (16 h). The following day, organoids are collected by the centrifugation of the suspension at 114 × *g* for 5 min. The supernatant is discarded, and the organoid pellet washed two-three times with PBS. Organoids with a size between 40 and 100 µm are collected and resuspended in fresh media (3 mL) and added to a six-well plate (Ultra-Low Attachment Surface plate, Corning # CLS3471). Organoids are allowed to stabilize for 24 h before use in the experiments.

Fatty acid preparation

Sodium octanoate (OA) was dissolved in PBS. To bind linoleic acid (Sigma # L8134) to BSA, it was initially dissolved in water to yield a 50 mM final concentration. 0.12 g of BSA was dissolved in 1.2 ml of water resulting a 10% BSA solution. A 0.2 ml aliquot of the Na linoleate solution was combined with the 10% BSA solution. After 15 min of slow stirring at 37 °C, 0.6 ml of water was added to bring the final concentration of Na linoleate to 5 mMol/L[59]. Linoleic acid bound to BSA (Sigma # L9530) was dissolved in water.

Lipid analysis

LC-MS grade methanol, dichloromethane, and ammonium acetate were purchased from Fisher Scientific (Pittsburgh, PA) and HPLC grade 1-propanol from Sigma-Aldrich (Saint Louis, MO). Milli-Q water was obtained from an in-house Ultrapure Water System by EMD Millipore (Billerica, MA). The Lipidyzer isotope labeled internal standards mixture consisting of 54 isotopes from 13 lipid classes was purchased from Sciex (Framingham, MA).

Sample preparation

Frozen plasma samples were thawed at room temperature (25 °C) for 30 min, vortexed; 25 µL of plasma was transferred to a borosilicate glass culture tube (16 × 100 mm). Next, 0.475 mL of water, 1.45 mL of 1:0.45 methanol:dichloromethane, and 25 µL of the isotope labeled internal standards mixture were added to the tube. The mixture was vortexed for 5 s and incubated at room temperature for 30 min. Next, another 0.5 mL of water and 0.45 mL of dichloromethane were added to the tube, followed by gentle vortexing for 5 s, and centrifugation at 2500 × *g* at 15 °C for 10 min. The bottom organic layer was transferred to a new tube and 0.9 mL of dichloromethane was added to the original tube for a second extraction. The combined extracts were concentrated under nitrogen and reconstituted in 0.25 mL of the mobile phase (10 mM ammonium acetate in 50:50 methanol:dichloromethane).

Mass spectrometry

Quantitative lipidomics was performed with the Sciex Lipidyzer platform consisting of Shimadzu Nexera X2 LC-30AD pumps, a Shimadzu Nexera X2 SIL-30AC autosampler, and a Sciex QTRAP® 5500 mass spectrometer equipped with SelexION® for differential mobility spectrometry (DMS).

1-propanol was used as the chemical modifier for the DMS. Samples were introduced to the mass spectrometer by flow injection analysis at 8 µL/min. Each sample was injected twice, once with the DMS on (PC/PE/LPC/LPE/SM), and once with the DMS off (CE/CER/DAG/DCER/FFA/HCER/LCER/TAG). The lipid molecular species were measured using multiple reaction monitoring (MRM) and positive/negative polarity switching. Positive ion mode detected lipid classes SM/DAG/CE/CER/DCER/HCER/DCER/TAG and negative ion mode detected lipid classes LPE/LPC/PC/PE/FFA. A total of 1070 lipids and fatty acids were targeted in the analysis.

Data processing

Data were acquired and processed using Analyst 1.6.3 and Lipidomics Workflow Manager 1.0.5.0. For statistical analysis, we evaluated the lipid species enrichments in the ER+, ER−, and control groups. The different groups were compared in pair-wise and the log-fold changes of lipid enrichment were derived, along with the effect sizes and *p*-values inferred from the regression models using the lipid measurement as an input variable and group information as the output variable.

Library preparation and RNA sequencing

MCF-10A: RNA was isolated with Qiagen RNeasy Plus Mini Kit (# 74134). The concentration and quality of total RNA in samples were assessed using the Agilent 2100 Bioanalyzer. RNA Integrity Number (RIN) of the vehicle and octanoate sample was 9.9 and 9.8, respectively. Sequencing libraries were prepared from a total of 100 ng of RNA using KAPA RNA HyperPrep Kit. Single-Indexed adapters were obtained from KAPA (catalog# KK8701). Library quality was assessed using the KAPA Library Assay kit. Each indexed library was quantified and its quality accessed by Qubit and Agilent Bioanalyzer, and 6 libraries were pooled in equal molarity. 5 µL of 4 nM pooled libraries were denatured, neutralized and a final concentration of 1.5 pM of pooled libraries was loaded to Illumina NextSeq 500 for 75 bp single-read sequencing. Approximately 80 M filtered reads per library was generated. A Phred quality score (Q score) was used to measure the quality of the sequencing. More than 88% of the sequencing reads reached Q30 (99.9% base call accuracy). Single-end FASTQ reads from RNA-seq measurements were aligned and mapped to hg38 ENSEMBL genome using STAR alignment[60]. Transcriptions per million (TPM) from mapped reads were estimated using RSEM from the STAR aligned reads[61]. The DESeq2 Bioconductor R package[62] was employed to determine differentially expressed genes for the octanoate treatment group compared to the vehicle-treated controls with FDR cutoff = 0.01 and $|\log_2 FC| \geq 2$ to identify a reasonable number of differentially expressed genes, on the order of several thousands of genes total, for subsequent analysis

MCF-12A: RNA isolation and library preparation as above. The Illumina NovaSeq 6000 platform was employed for 100 bp paired-end sequencing. The sequence reads were mapped to the hg38 reference geneome using STAR (Spliced Transcripts Alignment to a Reference)[60]. To evaluate the quality of the RNA-seq data, the number of reads that fall into different annotated regions (exonic, intronic, splicing junction, intergenic, promoter, UTR, etc.) of the reference genes was determined with bamUtils[63]. More than 83% of the sequencing reads reached Q30. Low quality mapped reads (including reads mapped to multiple positions) were excluded and featureCounts[64] was used to quantify the gene level expression. Differential gene expression analysis was performed with edgeR[65]. In this workflow, the statistical methodology applied uses negative binomial generalized linear models with likelihood ratio tests.

Gene ontology analysis of differentially expressed genes

Gene ontology pathway analysis for biological processes was performed on each set of differentially expressed genes using *Metascape*[66].

GSEA analysis

Raw counts were first estimated using *HTSeq* from STAR-aligned reads[67]. Next, replicates for control cells and treated cells were merged and normalized using modules from the GenePattern software package[68]. GSEA[69,70] was performed on these DESeq-normalized reads using annotations from online databases, including KEGG, Hallmark, Reactome, BioCarta, and Canonical Pathways. The normalized enrichment score (NES) of these top 20 pathways associated with the control and the octanoate-treated condition is shown with nominal *p*-value = 0.0. *Metascape* was employed to perform network analysis on these top 20 pathways associated with each treatment condition.

ATAC Seq Library preparation and sequencing

1×10^6 cells were pelleted and lysed in ATAC-resuspension buffer[71]. Extracted nuclei were processed for TN-5 mediated tagmentation using the Illumina Tagment DNA Enzyme and buffer kit (Nextera Illumina # 20034210): Transposon reaction mix as 2X TD Buffer-25 µl, Tn5 Transposase-2.5 µl, 1X PBS containing nuclei-16.5 µl, 10% Tween-20–0.5 µl (Sigma # P9416), 1% Digitonin-0.5 µl (Promega # G9441) and water at 37 °C, on a thermomixer at 1000 rpm for 30 min. Tagmented DNA was isolated by Nucleospin PCR clean-up (Takara Bio USA, Inc # 740609.250). Libraries were amplified for 8 cycles and purified using AMPure XP (Agencourt # A63880). Fragment sizes were determined using 106 LabChip GXII Touch HT (PerkinElmer), and 2×50 paired-end sequencing performed on NovaSeq S1 6000 flow cell (Illumina) to yield 100M reads per sample.

ATAC-seq data sequencing and peak calling

Illumina adapter sequences and low-quality base calls were trimmed off the paired end reads with Trim Galore v0.4.3. Sequence reads were aligned to human reference genome hg38 using bowtie2 with default settings. Duplicate reads were discarded with Picard. Reads mapped to mitochondrial DNA together with low mapping quality reads were excluded from further analysis. MACS2 was used to identify the peak regions with options -f BAMPE -g hs –keep-dup all -B -q 0.01[72]. Peaks for samples in the same condition were merged using the function 'merge' of bedtools and peaks for samples in different conditions were intersected using the function of 'intersect' of bedtools[73].

Differential chromatin accessibility analysis

The number of cutting sites of each sample was counted using the script dnase_cut_counter.py of pyDNase (version 0.2.4)[74]. The raw count matrix was normalized by CPM. R package edgeR (version 3.16.5) was used to conduct the differential accessibility analysis for all 66,853 common peaks. Significantly different accessible chromatin regions under different conditions were defined as the threshold 0.05 for FDR. With the cutoff 1 for the absolute value of fold change, comparing the treatment group with vehicle control group, we obtained 1704 significantly increased peaks and 3340 significantly decreased peaks.

Motif analysis

Motif analysis was conducted for significantly changed chromatin regions using 'findMotifsGenome.pl' script of HOMER (version 4.9) with default settings[33]. The principal component analysis was conducted to detect the important motifs using the relative enrichment of motifs calculated from HOMER reports. Biplot was used to visualize the principal component analysis results.

Genomic distribution of open chromatin regions

We calculated the overall genomic distribution of open chromatin regions, comparing the treatment to the vehicle[75]. We used the hg38 refseq genes annotation from UCSC Genome Browser to define the genomic features. All TSSs were considered in the analysis if a gene had multiple TSSs. The formula for reported enrichment is $(a/b)/(c/d)$. a is the number of peaks overlapping a given genomic feature, b is the number of total peaks, *c* is the number of regions corresponding to the feature, and d is the estimated number of discrete regions in the genome where the peaks and feature could overlap. Specifically, d is equal to (genome size)/ (mean peak size + mean feature size), following the implementation in the bedtools fisher (version 2.26.0).

Pathway analysis for open chromatin regions

For the 326 open chromatin regions with logFC ≥ 1.5 and FDR < 0.05 comparing the treatment with the vehicle, we extracted the target genes of the 326 chromatin regions. The function 'enrichKEGG' from the R package 'clusterProfile'[76] (version 3.6.0) was used to conduct KEGG pathway analysis with organism = "hsa" and adjusted.pval=0.05.

Validation of candidate genes qRT-PCR

Treated cells and organoids were washed with PBS and RNA was isolated with Qiagen RNeasy plus mini Kit (# 74134). cDNA was synthesized using the SuperScript VILO cDNA synthesis kit (ThermoFisher #11755250).

Real-time qPCR was performed using Applied biosystem QuantStudio 5 real time PCR System (Thermo Scientific). Expression data of the studied genes was normalized to RPLP1 to control the variability in expression levels and were analyzed using the $2^{-\Delta\Delta CT}$ method described by Livak and Schmittgen[77]. TaqMan gene expression assays and TaqMan fast advanced master mix (# 4444556) were purchased from ThermoFisher Scientific and the list of the assays is provided in Supplementary File 1.

qRT-PCR based TaqMan low density array assays

Based on histological diagnosis, benign breast epithelium was identified and captured by laser capture microdissection (LCM). RNA was isolated with Qiagen RNeasy plus mini Kit (# 74134). RNA quality was checked for integrity using Bioanalyzer 2100 (Agilent). 100 ng RNA was reverse transcribed using High Capacity RNA-to-cDNA Master Mix (ThermosFisher #4388950) and preamplified for 14 cycles using TaqMan PreAmp Master Mix 2X (ThermoFisher #4488593) and pooled assay mix for the genes in which we were interested. Pre-amplified cDNA were diluted to 1:20 with 1X TE buffer and mixed with Fast advanced master mix (ThermoFisher # 4444965) Each sample was loaded in duplicate in 384-well microfluidic cards customized with 47 genes of interest including three housekeeping genes (GAPDH, RPLP0, and RPLP1). TaqMan assays with best coverage attribution were used for the TLDA study as recommended by the manufacturer. A list of the genes and the Assay ID for the primers obtained from ThermoFisher is provided in Supplementary File 2. Real Time PCR reactions were carried out in QuantStudio 7 Flex system for 40 cycles using comparative Ct (ΔΔCt) method. Results were analyzed using Expression-Suite software.

Statistical analysis

Prior to performing the analyses, the log2-transformed relative (log2RE) amounts of mRNA expression were normalized to GAPDH and expressed as $\log_2 2^{-(CtX-CtGAPDH)} = -(CtX - CtGAPDH)$, where Ct is threshold cycle. The Mann–Whitney test was performed to identify genes with pairwise differences between ER+ and ER− samples. The analyses were adjusted for multiple testing, 34 genes, using the Benjamini–Hochberg (BH) adjustment in order to control the false discovery rate at the two-sided 0.05 level. Boxplots were used to visualize differences in log2RE by group. The log2RE analyses were conducted using the R statistical environment [R] version 3.5.1.

Live cell PWS imaging

Before treatment and imaging, MCF-10A cells were seeded in 6 well black culture plates, at least 35% confluency, and allowed to adhere overnight before the treatment with 500 µM LA and 5 mM Octanoate. We based the concentration of LA used in the experiment on the range in human plasma: 0.2–5.0 mmol/L[78]. To determine the chromatin packing behavior of MCF-10A cells under varying treatment conditions, live-cell PWS images were acquired at 37 °C and 5% CO_2 conditions. Imaging was performed using the commercial inverted microscope (Leica DMIRB) Hamamatsu Image-EM CCD camera C9100-13 coupled to a liquid crystal tunable filter (LCTF; CRi Woburn, MA) to acquire mono-chromatic spectrally resolved images that range from 500 to 700 nm at 1 nm intervals produced by a broad band illumination provided by an Xcite-120 LED Lamp (Excelitas, Waltham, MA) as previously described[79,80]. Briefly, PWS measures the standard deviation of internal optical scattering originating from chromatin in the nucleus, which is related to variations in the refractive index distribution (Σ). To obtain the interference signal directly related to refractive index fluctuations in the cell, we normalized measurements to an independent reference measurement acquired in an area of the plate without cells. These normalized spatial variations of refractive index are linearly proportional to nuclear mass density fluctuations, according to the Gladstone–Dale relation, and are characterized by chromatin packing scaling, D, the power-law relationship between the mass M of the chromatin polymer and the three-dimensional space it occupies R, i.e., $M \sim R^D$ [80]. The measured change in chromatin packing scaling between treatment conditions was quantified by first averaging D within each cell's nucleus and then averaging nuclei from over 100 cells per condition.

Notch reporter assay

Notch pathway function was analyzed by measuring the transcriptional activity of its downstream component RBP-jk using a Cignal RBP-Jk Dual Luciferase Reporter assay (Qiagen, Germantown, MD; # 336841). MCF-10A cells were transfected with Transcription Factor Reporter, Negative Control reporter or Positive Control constructs using the Neon Transfection System 10 µl kit (Invitrogen, Walthan, MA; # MPK1025). Twenty-four hours after transfection, cells were exposed to Sodium octanoate (OA, 5 mM) dissolved in PBS or PBS for 24 h. Cells were then lysed using Passive Lysis Buffer (Promega, Madison, WI) and transferred to a 96-well white flat-bottom plate (Corning, Tewksbury, MA). Luciferase activity was measured with the Dual-Luciferase Reporter Activity system (Promega; # E1910) using a Biotek Cytation 3 multiwell reader. Firefly luciferase activity was normalized to Renilla luciferase activity. All transfections were performed in triplicate. Luciferase measurement for each biological replicate was performed in three technical replicates. Values are expressed as mean ± SEM. The p-value was calculated by unpaired t-test.

Flux based analysis (FBA)

We calculated the relative activity of reactions in MCF-10A and MCF-12A cells by interpreting gene expression data using the Recon1 human metabolic model augmented with histone modifications[44,81]. We then identified a metabolic flux state that is most consistent with gene expression data in control and octanoate treatment. This was achieved by maximizing the activity of reactions that are associated with upregulated genes and minimizing flux through reactions that are downregulated in a condition, while simultaneously satisfying the stoichiometric and thermodynamic constraints embedded in the model using linear optimization[44,81]. The glucose, fatty acid, and glutamine levels in the simulations were adjusted based on the growth media used for culturing the cells. All p-values were corrected for multiple comparisons.

LC/MS based post-translation histone modification quantitation

MCF-10A were treated with 5 mM octanoate or 500 µM LA. After 24 h, cells were snap frozen for nuclei extraction. Histones were acid-extracted from 100% nuclei, derivatized via propionylation reaction and digested with trypsin. Each sample was resuspended in 50 µL of 0.1% TFA/mH2O and 2 µl was injected with 3 technical replicates. Multi-reaction monitoring (MRM) technology was used for histone analysis using a triple-quad mass spectrometer, which is programmed to fragment only specific precursor peptides and measure the intensity of specific product ions. Final results show changes of relative abundances of histone mark modification. Error bars are +/− one standard deviation obtained from sample technical replicate intensities.

Western blotting

Cells and organoids were plated and allowed to stabilize overnight and then treated the next day for the indicated times. At the end of the treatment, cells were collected, washed with PBS, and lysed in radio immunoprecipitation assay (RIPA) buffer (ThermoFisher Scientific # 89900) including protease inhibitors (ThermoFisher Scientific # 78430). Protein is estimated using the BCA protein assay kit (ThermoFisher Scientific # 23227) and loaded on 4–12% Bis Tris acetate gel using MES buffer, blotted on a polyvinylidene fluoride (PVDF) membrane (Invitrolon 0.45 µM) and blocked with blocking buffer (10% skimmed milk) for 1 h at room temperature. Primary antibodies were purchased from Cell Signaling Technologies – AcH3K9 (rabbit mAb Cell Signaling #9649) at 1:250 dilution, AcH3K14 (rabbit mAb Cell Signaling #7627) at 1:250 dilution and H3 (Rabbit mAb Cell Signaling #9715) at 1:1000 dilution. Membranes were incubated in primary antibody overnight at 4 degrees Celsius with shaking. Blots were washed with PBS + 0.1% Tween 20, three times 5 min each, and probed with secondary antibodies (Anti-rabbit IgG, HRP-linked Antibody, Cell Signaling #7074) at a concentration of 1:10,000 for 1 h at room temperature. Blots/lysates were derived from the same experiment and were processed concurrently. Uncropped images of the original Western blots are provided in Supplementary File 3.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

The datasets generated and analyzed during the current study are publicly available in the Gene Expression Omnibus: accession numbers GSE126799 (RNA-seq, MCF-10A), GSE190572 (RNA-seq, MCF-12A), and GSE190573 (ATAC-seq).

CODE AVAILABILITY

No custom codes were written to analyze the data presented in this manuscript. The following software versions were used: Trim Galore v0.4.3.; pyDNase version 0.2.4; R package edgeR version 3.16.5; HOMER version 4.9 with default settings; bedtools fisher version 2.26.0; 'enrichKEGG' from the R package 'clusterProfile' version 3.6.0; R statistical environment [R] version 3.5.1.

Received: 26 December 2020; Accepted: 31 March 2022;
Published online: 04 May 2022

REFERENCES

1. Weigelt, B. & Reis-Filho, J. S. Histological and molecular types of breast cancer: Is there a unifying taxonomy? *Nat. Rev. Clin. Oncol.* **6**, 718–730 (2009).
2. Gail, M. H. et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J. Natl Cancer Inst.* **81**, 1879–1886 (1989).
3. Fisher, B. et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J. Natl Cancer Inst.* **90**, 1371–1388 (1998).
4. Cuzick, J. et al. Long-term results of tamoxifen prophylaxis for breast cancer–96-month follow-up of the randomized IBIS-I trial. *J. Natl Cancer Inst.* **99**, 272–282 (2007).
5. Cuzick, J. et al. Use of anastrozole for breast cancer prevention (IBIS-II): Long-term results of a randomised controlled trial. *Lancet* **395**, 117–122 (2020).
6. Bouchardy, C. et al. Risk of second breast cancer according to estrogen receptor status and family history. *Breast Cancer Res. Treat.* **127**, 233–241 (2011).
7. Kurian, A. W. et al. Second primary breast cancer occurrence according to hormone receptor status. *J. Natl Cancer Inst.* **101**, 1058–1065 (2009).
8. Swain, S. M. et al. Estrogen receptor status of primary breast cancer is predictive of estrogen receptor status of contralateral breast cancer. *J. Natl Cancer Inst.* **96**, 516–523 (2004).
9. Wang, J. et al. Lipid metabolism genes in contralateral unaffected breast and estrogen receptor status of breast cancer. *Cancer Prev. Res.* **6**, 321–330 (2013).
10. Wang, J. et al. Overexpression of lipid metabolism genes and PBX1 in the contralateral breasts of women with estrogen receptor-negative breast cancer. *Int. J. Cancer* **140**, 2484–2497 (2017).
11. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011).
12. Nieman, K. M. et al. Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth. *Nat. Med.* **17**, 1498–1503 (2011).
13. Mitra, R., Le, T. T., Gorjala, P. & Goodman, O. B. Jr. Positive regulation of prostate cancer cell growth by lipid droplet forming and processing enzymes DGAT1 and ABHD5. *BMC Cancer* **17**, 631 (2017).
14. Singh, K. B. et al. Prostate cancer chemoprevention by sulforaphane in a preclinical mouse model is associated with inhibition of fatty acid metabolism. *Carcinogenesis* **39**, 826–837 (2018).
15. Huang, Li,T. et al. HIF-1-mediated suppression of acyl-CoA dehydrogenases and fatty acid oxidation is critical for cancer progression. *Cell Rep.* **8**, 1930–1942 (2014).
16. Camarda, R. et al. Inhibition of fatty acid oxidation as a therapy for MYC-overexpressing triple-negative breast cancer. *Nat. Med.* **22**, 427–432 (2016).
17. Park, J. H. et al. Fatty acid oxidation-driven Src links mitochondrial energy reprogramming and oncogenic properties in triple-negative breast cancer. *Cell Rep.* **14**, 2154–2165 (2016).
18. Havas, K. M. et al. Metabolic shifts in residual breast cancer drive tumor recurrence. *J. Clin. Invest.* **127**, 2091–2105 (2017).
19. Cappelletti, V. et al. Metabolic footprints and molecular subtypes in breast cancer. *Dis. Markers* **2017**, 7687851 (2017).
20. You, S. et al. Raman spectroscopic analysis reveals abnormal fatty acid composition in tumor micro- and macroenvironments in human breast and rat mammary cancer. *Sci. Rep.* **6**, 32922 (2016).
21. Reid, M. A., Dai, Z. & Locasale, J. W. The impact of cellular metabolism on chromatin dynamics and epigenetics. *Nat. Cell Biol.* **19**, 1298–1306 (2017).
22. Grabacka, M., Pierzchalska, M., Dean, M. & Reiss, K. Regulation of ketone body metabolism and the role of PPARalpha. *Int. J. Mol. Sci.* https://doi.org/10.3390/ijms17122093 (2016).
23. Schönfeld, P. & Wojtczak, L. Short- and medium-chain fatty acids in energy metabolism: The cellular perspective. *J. Lipid Res.* **57**, 943–954 (2016).
24. Virk, R. K. A. et al. Disordered chromatin packing regulates phenotypic plasticity. *Sci. Adv.* **6**, eaax6232 (2020).
25. Sweeney, M. F., Sonnenschein, C. & Soto, A. M. Characterization of MCF-12A cell phenotype, response to estrogens, and growth in 3D. *Cancer Cell Int.* **18**, 43 (2018).
26. Whelan, J. & Fritsche, K. Linoleic acid. *Adv. Nutr.* **4**, 311–312 (2013).
27. Ahel, D. et al. Poly(ADP-ribose)-dependent regulation of DNA repair by the chromatin remodeling enzyme ALC1. *Science* **325**, 1240–1243 (2009).
28. Guyenet, S. J. & Carlson, S. E. Increase in adipose tissue linoleic acid of US adults in the last half century. *Adv. Nutr.* **6**, 660–664 (2015).
29. Almassalha, L. M. et al. Macrogenomic engineering via modulation of the scaling of chromatin packing density. *Nat. Biomed. Eng.* **1**, 902–913 (2017).
30. Almassalha, L. M. et al. The global relationship between chromatin physical topology, fractal structure, and gene expression. *Sci. Rep.* **7**, 41061–41061 (2017).
31. Almassalha, L. M. et al. The greater genomic landscape: The heterogeneous evolution of cancer. *Cancer Res.* **76**, 5605–5609 (2016).
32. Yadav, S. et al. Abstract P3-02-07: Fatty acid metabolism is associated with chromatin remodeling in mammary epithelial cells. *Cancer Res.* https://doi.org/10.1158/1538-7445.SABCS19-P3-02-07 (2020).
33. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
34. Steensels, S. & Ersoy, B. A. Fatty acid activation in thermogenic adipose tissue. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1864**, 79–90 (2019).
35. McDonnell, E. et al. Lipids reprogram metabolism to become a major carbon source for histone acetylation. *Cell Rep.* **17**, 1463–1472 (2016).
36. O'Brien, E. J., Monk, J. M. & Palsson, B. O. Using genome-scale models to predict biological capabilities. *Cell* **161**, 971–987 (2015).
37. Orth, J. D., Thiele, I. & Palsson, B. O. What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248 (2010).
38. Chandrasekaran, S. et al. Comprehensive mapping of pluripotent stem cell metabolism using dynamic genome-scale network modeling. *Cell Rep.* **21**, 2965–2977 (2017).
39. Frezza, C. et al. Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase. *Nature* **477**, 225–228 (2011).
40. Mardinoglu, A. et al. Personal model-assisted identification of NAD(+) and glutathione metabolism as intervention target in NAFLD. *Mol. Syst. Biol.* **13**, 916 (2017).
41. Shlomi, T., Cabili, M. N., Herrgard, M. J., Palsson, B. O. & Ruppin, E. Network-based prediction of human tissue-specific metabolism. *Nat. Biotechnol.* **26**, 1003–1010 (2008).
42. Yizhak, K., Chaneton, B., Gottlieb, E. & Ruppin, E. Modeling cancer metabolism on a genome scale. *Mol. Syst. Biol.* **11**, 817 (2015).
43. Duarte, N. C. et al. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl Acad. Sci. USA* **104**, 1777–1782 (2007).
44. Shen, F., Boccuto, L., Pauly, R., Srikanth, S. & Chandrasekaran, S. Genome-scale network model of metabolism and histone acetylation reveals metabolic dependencies of histone deacetylase inhibitors. *Genome Biol.* **20**, 49 (2019).
45. Carrer, A. et al. Impact of a high-fat diet on tissue acyl-CoA and histone acetylation levels. *J. Biol. Chem.* **292**, 3312–3322 (2017).
46. Li, J. T. et al. BCAT2-mediated BCAA catabolism is critical for development of pancreatic ductal adenocarcinoma. *Nat. Cell Biol.* **22**, 167–174 (2020).
47. Kerlikowske, K. et al. Risk factors that increase risk of estrogen receptor-positive and -negative breast cancer. *J. Natl Cancer Inst.* https://doi.org/10.1093/jnci/djw276 (2017).
48. Li, K. et al. Risk prediction for estrogen receptor-specific breast cancers in two large prospective cohorts. *Breast Cancer Res.* **20**, 147 (2018).
49. Worsham, M. J. et al. Multiplicity of benign breast lesions is a risk factor for progression to breast cancer. *Clin. Cancer Res.* **13**, 5474–5479 (2007).
50. Alonzo-Proulx, O., Mainprize, J. G., Harvey, J. A. & Yaffe, M. J. Investigating the feasibility of stratified breast cancer screening using a masking risk predictor. *Breast Cancer Res.* **21**, 91 (2019).
51. Reiner, A. S. et al. Hormone receptor status of a first primary breast cancer predicts contralateral breast cancer risk in the WECARE study population. *Breast Cancer Res.* **19**, 83 (2017).
52. Rodilla, V. et al. Luminal progenitors restrict their lineage potential during mammary gland development. *PLoS Biol.* **13**, e1002069 (2015).
53. Lilja, A. M. et al. Clonal analysis of Notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland. *Nat. Cell Biol.* **20**, 677–687 (2018).
54. Milne, R. L. et al. Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. *Nat. Genet.* **49**, 1767–1778 (2017).

55. Li, Y., Dillon, T. J., Takahashi, M., Earley, K. T. & Stork, P. J. Protein kinase A-independent ras protein activation cooperates with Rap1 protein to mediate activation of the extracellular signal-regulated kinases (ERK) by cAMP. *J. Biol. Chem.* **291**, 21584–21595 (2016).
56. Creighton, C. J. et al. Activation of mitogen-activated protein kinase in estrogen receptor alpha-positive breast cancer cells in vitro induces an in vivo molecular phenotype of estrogen receptor alpha-negative human breast tumors. *Cancer Res.* **66**, 3903–3911 (2006).
57. Mentch, S. J. et al. Histone methylation dynamics and gene regulation occur through the sensing of one-carbon metabolism. *Cell Metab.* **22**, 861–873 (2015).
58. Wang, T. et al. JAK/STAT3-regulated fatty acid beta-oxidation is critical for breast cancer stem cell self-renewal and chemoresistance. *Cell Metab.* **27**, 136–150 e135 (2018).
59. Pappas, A., Anthonavage, M. & Gordon, J. S. Metabolic fate and selective utilization of major fatty acids in human sebaceous gland. *J. Invest. Dermatol.* **118**, 164–171 (2002).
60. Dobin, A. et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
61. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **12**, 323 (2011).
62. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
63. Breese, M. R. & Liu, Y. NGSUtils: A software suite for analyzing and manipulating next-generation sequencing datasets. *Bioinformatics* **29**, 494–496 (2013).
64. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
65. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
66. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
67. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
68. Reich, M. et al. GenePattern 2.0. *Nat. Genet.* **38**, 500–501 (2006).
69. Mootha, V. K. et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
70. Subramanian, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545 (2005).
71. Ackermann, A. ATAC-seqProtocolOmni-KaestnerLab.pdf. (2019).
72. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
73. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
74. Piper, J. et al. Wellington: A novel method for the accurate identification of digital genomic footprints from DNase-seq data. *Nucleic Acids Res.* **41**, e201 (2013).
75. Miyamoto, K. et al. Chromatin accessibility impacts transcriptional reprogramming in Oocytes. *Cell Rep.* **24**, 304–311 (2018).
76. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
77. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* **25**, 402–408 (2001).
78. Abdelmagid, S. A. et al. Comprehensive profiling of plasma fatty acid concentrations in young healthy Canadian adults. *PLoS One* **10**, e0116195 (2015).
79. Almassalha, L. M. et al. Label-free imaging of the native, living cellular nanoarchitecture using partial-wave spectroscopic microscopy. *Proc. Natl Acad. Sci. USA* **113**, E6372–E6381 (2016).
80. Eid, A. et al. Characterizing chromatin packing scaling in whole nuclei using interferometric microscopy. *Opt. Lett.* **45**, 4810–4813 (2020).
81. Shen, F., Cheek, C. & Chandrasekaran, S. Dynamic network modeling of stem cell metabolism. *Methods Mol. Biol.* **1975**, 305–320 (2019).

ACKNOWLEDGEMENTS

We would like to thank Professors Matthew D Hirschey and Neil Kelleher for advice regarding histone proteomics and, Jeannie Camarillo and the Northwestern Proteomics Core for conducting the histone proteomic analysis; The Northwest Metabolic Research Center (NW-MRC) at University of Washington for performing the lipidomics analysis; The Center for Medical Genomics at the Indiana University School of Medicine for RNA library preparation and RNA sequencing, and ATAC sequencing; The NU Seq Core facility for providing The QuantStudio 7 Flex system; Natalie Pulliam for consenting patients and collecting tissue, and our many lab colleagues for feedback. This work was supported by the Breast Cancer Research Foundation (S.A.K.), Bramsen-Hamill Foundation (S.A.K.), Otzen Family Foundation (S.E.C.), Ulla and Bertil Brunk Foundation (S.E.C.), R35 GM13779501 from The National Institute of General Medical Sciences, NIH (S.C.). National Institutes of Health (award number R01CA228272).

AUTHOR CONTRIBUTIONS

Conceptualization and project design: S.Y., S.E.C., and S.A.K.; experiments: S.Y. and M.B.E.; sequencing: X.X and H.G; RNA-seq analysis: R.K.A.V., Z.Z., G.C, and M.R.; ATAC Seq analysis: D.C.; PWS microscopy: D.V.D.; metabolic flux analysis: C.H.C. and S.C.; statistical analysis: K.L.B.; writing (original draft): S.Y., S.E.C., and S.A.K.; writing (review and editing): V.B., S.C., and R.C. The authors approved the completed version.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41523-022-00422-0.

Correspondence and requests for materials should be addressed to Seema Ahsan Khan or Susan E. Clare.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.