Imperial College London

Faculty of Medicine
Department of Surgery and Cancer

The role of SREBP1 in

hormone-dependent breast cancer

By

Ylenia Perone

A thesis submitted for the degree of Doctor of Philosophy

Supervisor: Dr Luca Magnani

CANCER RESEARCH UK

A Piero, Delia, Claudia, Andrea ed Emilio,
influssi inesorabili della mia vita

Sans passion il n'y a pas d'art
Henri Matisse

<u>Declaration of originality</u>

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy. I declare that the work reported in this thesis is my own and all else is appropriately referenced to my best knowledge. Where results generated by a colleague are included, this is clearly stated and fully acknowledged. Some of the findings presented in this thesis have been published in the following articles:

1. **Y Perone** and L Magnani. **"**Going off the grid: ERα breast cancer beyond estradiol." Journal of Molecular Endocrinology, 2016 doi: 10.1530/JME-16-0062

2. L Magnani, G Frige`, RM Gadaleta, G Corleone, S Fabris, H Kempe, PJ Verschure, I Barozzi, V Vircillo, S Hong, **Y Perone**, M Saini, A Trumpp, G Viale, A Neri, S Ali, MA Colleoni, G Pruneri and S Minucci. "Acquired *CYP19A1* amplification is an early specific mechanism of aromatase inhibitor resistance in ERα metastatic breast cancer" Nature Genetics, 2017 doi: 10.1038/ng.3773

3. DK Patten, GC, B Győrffy, **Y Perone**, N Slaven, I Barozzi, E Erdős, A Saiakhova, K Goddard, A Vingiani, S Shousha, LS Pongor, DJ Hadjiminas, G Schiavon, P Barry, C Palmieri, RC Coombes, P Scacheri, G Pruneri and L Magnani. "Enhancer mapping uncovers phenotypic heterogeneity and evolution in patients with luminal breast cancer" Nature Medicine 2018 doi: 10.1038/s41591-018-0091-x

4. **Y Perone,** AJ Farrugia, A Rodríguez Meira, B Győrffy, C Ion, A Uggetti, A Chronopoulos, P Marrazzo, M Faronato, S Shousha, C Davies, JH Steel, N Patel, A del Rio Hernandez, C Coombes, G Pruneri, A Lim, F Calvo and L Magnani. "SREBP1 drives Keratin-80-dependent cytoskeletal changes and invasive behavior in endocrine- resistant ERα breast cancer" Nature Communications 2019 doi: 10.1038/s41467-019-09676-y

Copyright declaration

Ylenia Perone

5 September 2019

Abstract

Despite significant progress in diagnostics and treatments, such as targeted endocrine therapy, 30% of patients with hormone-dependent breast cancer eventually develop disease recurrence predominantly due to drug resistance. Resistance to hormone deprivation therapy is multifactorial and involves several molecular events. An increasing body of research has identified an emerging hallmark of cancer describing the capability of modifying and reprogramming cellular metabolism in order to fuel neoplastic proliferation. Our group has previously uncovered how the specific type of treatment plays a significant role in this process. In particular, breast cancer (BCa) cells developing resistance to aromatase inhibitors (AI) endogenously trigger cholesterol biosynthesis (CB) through sterol regulatory element binding protein 1 (SREBP1) regulation leading to a sustained oestrogen independent, oestrogen receptor alpha (ERα) activation. Cellular lipid metabolism is controlled by SREBP1. The altered lipid metabolism, also known as "lipogenic phenotype", has been linked with prostate cancer (PCa) pathogenesis: the expression of SREBP1 in prostate cancer is strongly correlated with Gleason grade (pathological grade) and its overexpression is sufficient to increase tumorigenicity and invasion of prostate cancer cells. Furthermore, *de novo* lipid biosynthesis has been associated with cancer progression, poorer prognosis and shorter patient survival. Considering what is known about the pathobiology of lipids in cancer, it is plausible that invading cells evolve mechanisms to bypass the tight homeostatic regulation of intracellular cholesterol adapting to their new environmental conditions. With this idea in mind, we sought to identify the molecular mechanisms of activation of SREBP1 as key regulator of *de novo* cholesterol biosynthesis in hormone-dependent cancers resistant to endocrine therapy.

Firstly, we wanted to investigate SREBP1 regulation in hormone-dependent cancer cells. We found that SREBP1-driven lipogenesis is consistently upregulated after long-term steroid deprivation, thus when cells become hormone independent. *In vivo* immunohistochemistry (IHC) data support the hypothesis that SREBP1 might be pivotal in driving *de novo* cholesterol biosynthesis in endocrine therapy resistant BCa cells. Moreover, the switch of metabolic dependency upon resistance development identified by metabolic profiling, is associated with increased *de novo* cholesterol and fatty acid synthesis.

In order to examine whether cholesterol biosynthesis may be upregulated by modulating SREBP1 signalling, we investigated SREBP1 recruitment to the chromatin. Optimization of

ChIP protocol allowed for a genome-wide profiling of SREBP1 binding in BCa and PCa cancer cell lines. Downstream analysis showed a difference in SREBP1 recruitment between parental and long-term starved derived cell lines in MCF7. Furthermore, SREBP1 binding profiles distinguished cancer cells based on the tissue of origin (breast versus prostate cancer). Our data also confirmed a significant co-occurrence between AR and SREBP1 binding sites in PCa and suggested a possible crosstalk between SREBP1 and the ERα in BCa on chromatin.

Thirdly, we asked what the targets of SREBP1 are and if they can promote invasive potential. ChIP-seq differential binding analysis unexpectedly revealed non-canonical targets for SREBP1. In particular, we showed that cells acquiring resistance to AI undergo active cytoskeleton re-organisation via Keratin 80 (KRT80) and actin remodelling. This process is driven by epigenetic reprogramming at the type II keratin locus dependent on *de novo* SREBP1 binding to a single enhancer that is activated upon chronic AI treatment and leading to KRT80 upregulation. Our data strongly suggest that therapy plays a direct role in shaping the biophysical properties and invasive potential of breast cancer cells, by inducing epigenetic rearrangements leading to KRT80 upregulation and concomitant cytoskeletal reorganization.

In summary, our study investigates the role of SREBP1 as a key player in endogenous cholesterol accumulation and autonomous activation of the nuclear receptor signalling, leading to the hormone independent tumour proliferation and invasion via global cytoskeletal re-arrangements in hormone-dependent cancers.

Acknowledgements

Table of Contents

List of Figures

List of Tables

List of Abbreviations

17ß-HSD 17ß-hydroxysteroid dehydrogenase
25-, 27-HC 25-hydroxycholesterol and 27-hydroxycholesterol
ACC1acetyl-CoA carboxylase 1
ACLY ATP citrate lyase
ADT Androgen deprivation therapy
AI Aromatase inhibitor
AICAR 5-aminoimidazole-4-carboxamide ribonucleotide transformylase
AMP adenosine monophosphate
AMPK AMP-activated protein kinase
APS Ammonium persulphate
AR Androgen receptor
AR-FL Full length AR
AR-Vs Androgen receptor variants
ATF6 Activating transcription factor 6
ATP Adenosine triphosphate
AKR1C3 Aldo-Keto Reductase Family 1 Member C3
B2-M B2-microglobulin
BCS Breast conserving surgery
bp Base pair
BCa Breast cancer
bHLHZip basic Helix-loop-helix zipper region
BRCA1 Breast Cancer Type 1 susceptibility protein
BRCA2 Breast Cancer Type 2 susceptibility protein
BSA Bovine serum albumin
CB Cholesterol biosynthesis
cDNA Complementary DNA
CE cholesteryl ester
ChIP Chromatin immunoprecipitation
ChIP-seq Chromatin immunoprecipitation sequencing
CKD chronic kidney disease
CNA Copy Number Aberration
CNV Copy Number Variation
CREB cAMP-response element binding protein
CREBBP CREB -binding protein
CRE Cyclic adenosine monophosphate-responsive element
CRISPR Clustered regularly interspaced short palindromic repeats
CRPC Castration resistant prostate cancer
CRUK Cancer Research United Kingdom
CRTC 2 CRE binding protein-regulated transcription coactivator
CTRL control
CYP11A1 Cytochrome P450 Family 11 Subfamily A Member 1
CYP17A1 Cytochrome P450 Family 17 Subfamily A Member 1
CYP19A1 Cytochrome P450 Family 19 Subfamily A Member 1
DBD DNA-binding domain
DCIS Ductal carcinoma in situ
DCFCS Charcoal-stripped fetal calf serum
ddH2O Double distilled deionised water

DHEA dehydroepiandrosterone
DHS DNase I hypersensitivity sites
DHT 5α-Dihydrotestosterone
DMEM Dulbecco's Modified Eagle Media
DMSO Dimethyl sulphoxide
DSS Dextran coated charcoal-stripped FCS
DXT Radiotherapy
E1 oestrone
E2 17β oestradiol
E3 oestriol
E4BP4 E4 promoter-binding protein 4
EBCTCG Early Breast Cancer Trialists Collaborative Group
ECM Extracellular Matrix
EDTA Ethylene di-amine tetra-acetic acid
EGFR Epidermal growth factor receptor
ENCODE The Encyclopaedia of DNA Elements project
EMT Epithelial-to-mesenchymal transition
ER endoplasmic reticulum
ER Oestrogen receptor
ER+ Oestrogen receptor positive
ER- Oestrogen receptor negative
ERα Oestrogen receptor alpha
ERβ Oestrogen receptor beta
ERα BCa Breast cancer
ERE Oestrogen response elements
ET endocrine therapy
EtOH Ethanol
ETR endocrine therapy resistant
FA fatty acid
FAs focal adhesions
FAO FA oxidation
FASN Fatty acid synthase
FCS Foetal calf serum
FISH Fluorescence *in situ* hybridization
FPKM Fragments Per Kilobase of transcript per Million mapped reads
FM full media
FSH Follicle-stimulating hormone
G6P Glucose-6-phosphate
GAPDH Glyceraldehyde 3-phosphate dehydrogenase
GC-MS Gas chromatography mass spectrometry
GFP Green fluorescent protein
GGPP geranylgeranyl pyrophosphate
GLUT glucose transporter
GnRH Gonadotrophin releasing hormone
GO gene ontology
GREAT Genomic Regions Enrichment of Annotations Tool
GSEA Gene set enrichment analysis
H3K27ac Acetylation of lysine 27 of histone H3
H3k4me2 Dimethylation of lysine 4 of histone H3
H3k4me1 Monomethylation of lysine 4 of histone H3

HCC hepatocarcinoma
HDL high-density lipoprotein
HER2 Human epidermal growth factor receptor 2
HER2- Human epidermal growth factor receptor 2 negative
HIF hypoxia inducible factor
HMGCR 3-hydroxy-3-methylglu-taryl-CoA reductase
HMGCS Hydroxymethylglutaryl-CoA synthase
HRT Hormonal replacement therapy
HSD3B1 3beta-hydroxysteroid dehydrogenase/delta(5)-delta(4)isomerase type I
HSP90 Heat shock protein 90
ICR Institute of Cancer Research
IDC Invasive ductal carcinoma
IEO European Institute of Oncology
IHC Immunohistochemistry
IL Interleukin
ILC Invasive lobular carcinoma
Indel Insertion/deletion
INSIG insulin-induced gene
K cytokeratin (protein)
KRT cytokeratin (gene, RNA)
kb Kilobases (1000 bp)
kDa Kilodalton (1000 Da)
LBD Ligand binding domain
LDL Low-density lipoprotein
LDLR low-density lipoprotein receptor
LOG2FC Log2 fold change
LOH Loss of heterozygosity
LPIN1 phosphatidate phosphatase lipin-1
LPS lipopolysaccharide stimulating lipogenesis
LTED Long term oestrogen deprived
LTEDF LTED cells resistant to Fulvestrant
LTEDT LTED cells resistant to Tamoxifen
LXRα Liver X receptor alpha
M molar
μM Micromolar
MAPK Mitogen activated protein kinase
Mb Megabase
mCRPC metastatic androgen independent/castration-resistant prostate cancer
Mbase Million bases
MEM Modified eagle's medium
MEK Mitogen-activated protein kinase
MCF7F/MCF7 FulvR MCF7 fulvestrant resistant cells
MCF7T MCF7 TamR MCF7 tamoxifen resistant cells
miRNAs/miRs microRNAs
mM Millimolar
mSREBP1 mature form of SREBP1
mTOR mammalian target of rapamycin
mTORC1 mammalian target of rapamycin complex 1
MUFA Monounsaturated fatty acids
NACT neoadjuvant chemotherapy

NADH Nicotinamide adenine dinucleotide
NAFLD non-alcoholic fatty liver disease
NASH non-alcoholic steatohepatitis
NFY nuclear factor Y
NGS Next-generation sequencing
NHEJ Non-homologous end-joining
NR Nuclear receptor
N-SREBP nuclear form of SREBP
OHT 4-hydroxytamoxifen
OXPHOS Oxidative phosphorylation
O.S./OS Overall survival
PAM Protospacer adjacent motif
PARP Poly (ADP-ribose) polymerase
PBS Phosphate buffered saline
PBST Tween 20 in PBS
PC phosphatidylcoline
PC/PCa Prostate Cancer
PCA Principle component analysis
PCR Polymerase chain reaction
PDX Patient-derived xenograft
PFS Progression free survival
PI3K Phosphoinositide 3-kinase
PI3K/Akt phosphatidylinositol 3-OH-kinase/ proteinkinase B
PKA protein kinase A
PKC protein kinase C
PL phospholipid
PR Progesterone receptor
PSG L-glutamine-penicillin-streptomycin
PSA Prostate-specific antigen
pSREBP1 premature form of SREBP1
PTEN Phosphatase and tensin homologue
PUFA polyunsaturated fatty acids
qPCR Quantitative PCR
Rb Retinoblastoma protein gene
RIN RNA integrity number
RIPA Radioimmune precipitation buffer
RM Red media
RNA Ribonucleic acid
RNA-seq RNA-sequencing
RPKM Reads per kilobase of transcript per million mapped reads
RT-PCR Reverse transcriptase-PCR
RT-qPCR Real time-quantitative PCR
S1P Site-1 Protease
S2P Site-2 Proteases
S6K Hepatic p70 S6 kinase
SCAP SREBP cleavage-activating protein
SCD1 Stearoyl-CoA desaturase
SD standard deviation
SDS Sodium deoxycholate
SEM Standard error of the mean

SERD Selective oestrogen receptor down regulator
SERM Selective oestrogen receptor modulator
SRB Sulphorhodamine B
sgRNA Single guide RNA
siRNA small interference RNA
shRNA Short hairpin RNA
SQLE squalene epoxidase
SM stripped media
SRB Sulforhodamine B
SRD5A1 Steroid 5 Alpha-Reductase 1
SREs sterol response elements
SREBF1 Sterol Regulatory Element Binding factor 1
SREBF2 Sterol Regulatory Element Binding factor 2
SREBP1 Sterol Regulatory Element Binding Protein 1
SREBP2 Sterol Regulatory Element Binding Protein 2
StAR steroidogenic acute regulartory protein
SUMO Small ubiquitin-like modifier
T Testosterone
TAD Topologically associated domain
TAG triacylglyceride
TAZ Transcriptional coactivator with PDZ-binding motif
TCA Trichloroacetic acid
TCA triacylglycerol
TE Tris-EDTA
TF Transcription factor
TMA Tissue microarray
TNBC Triple negative breast cancer
TNM (Tumour, Node, Metastasis) Classification of Malignant Tumours
TSS Transcription start site
Ub ubiquitylated
UFAs unsaturated fatty acids
UPR unfolded protein response
UT Untreated
WM white media
YAP1 Yes-associated protein 1

"The more you know,
the more you know you don't know"
Aristotele

Chapter 1: Introduction

1.1 Cancer

Cancer continues to be a major health issue and still represents the first or second leading cause of death before age 70 years in most countries[1]. Worldwide cancer statistics of 2018 reported that cancers of the lung, female breast, and colorectum are the top three cancer types in terms of incidence and are ranked within the top five in terms of mortality[1] (first, fifth, and second, respectively) (Fig. 1.1). Together, these three cancer types are responsible for one third of the cancer incidence and mortality burden worldwide[1]. Lung, breast, bowel and prostate cancers together also account for more than half of all new cases in the U. K [CRUK 2019, https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence]. Cancer-related mortality is decreasing despite the aforementioned increases in incidence[1] (Fig. 1.1). Earlier diagnosis coupled with improved methods of treatment have largely contributed to the fall in cancer mortality.

Estimated number of incident cases and deaths worldwide, both sexes, all ages
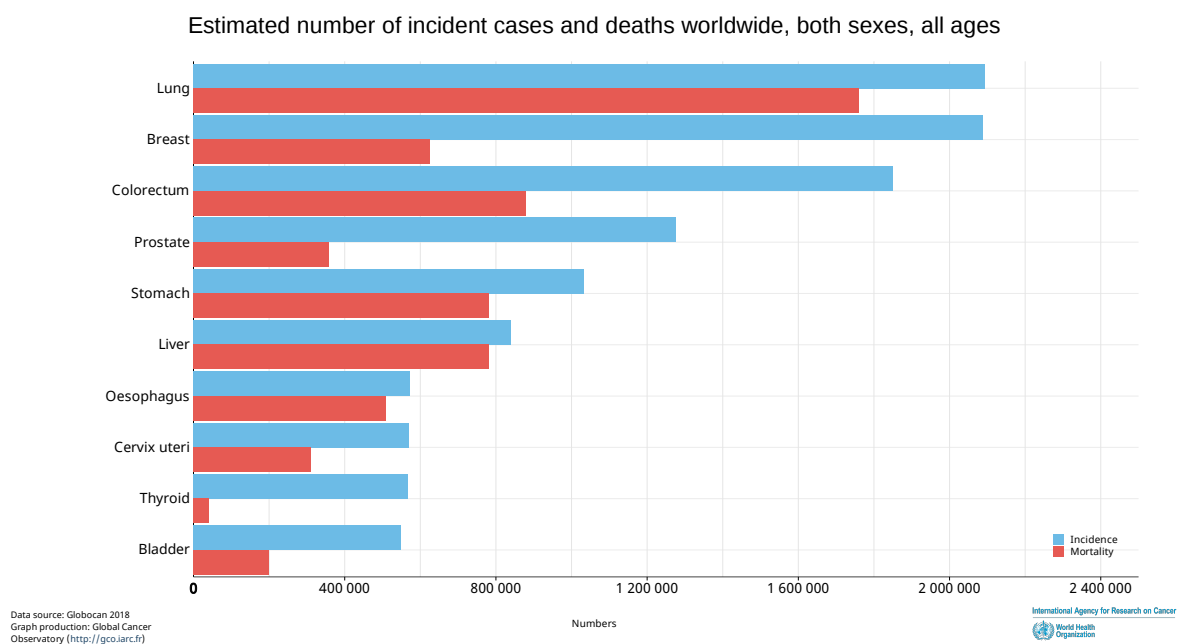
Figure 1.1: Cancer incidence and mortality. Data showing estimated number of incident cases (blue) and deaths (red) of all cancers worldwide, both sexes, all ages (adapted from GLOBOCAN, 2018[1]).

1.1.1 Breast cancer

Breast cancer is the most commonly diagnosed cancer (24.2%, i.e. about one in 4 of all new cancer cases diagnosed in women worldwide are breast cancers) and the leading cause of cancer death in women in over 100 countries[1] (15.0%, Fig.1.2). Breast cancer is also the most common cancer in the UK, accounting for 15% of all new cases (CRUK, 2019 https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive).

Hereditary and genetic factors, including a personal or family history of breast or ovarian cancer and inherited mutations (such as BRCA1, BRCA2, and other breast cancer susceptibility genes), account for only 5% to 10% of breast cancer cases. Conversely, studies of migrants have shown that non-hereditary factors are the major drivers of the observed international and interethnic differences in incidence[1]. Almost 40 years ago, epidemiological studies firstly demonstrated that greater lifetime exposure to oestrogen raises breast cancer risk[2]. Risk factors identified, which would all increase exposure to oestrogen[3,4], include earlier age of menarche[3,4], later onset of menopause[4], long-term oral contraceptive use[4], hormonal replacement therapy (HRT)[5], elevated levels of circulating oestrogens and androgens[3,6,7], obesity and oestrogen receptor-dependent proliferative activity of mammary epithelial cells[8]. Interestingly, an early first full-term pregnancy provides considerable protection against breast cancer development. This has been replicated in mouse and rat models, by supplying levels of oestrogen and progesterone found in pregnancy, which reduced their risk of mammary tumours following carcinogen exposure[9].

Subtypes of breast cancer can be defined by the expression status of key driver genes (intrinsic subtypes), with the PAM50 gene signature sometimes used to further define tumour subtypes or immunohistochemistry (IHC) of tumour biopsies[10,11]. Often times, molecular and histological subtyping have considerable overlap. Seventy percent of cases, characterised as luminal breast cancer, express oestrogen receptor α (ER), and can be further sub-divided into luminal A and B subtypes by the expression of the proliferation marker, Ki67, with greater than 14% expression for luminal B, indicating a worse prognosis[12]. Progesterone receptor (PR) status is also determined by IHC, and PR expression is associated with a better response to therapy in patients with ER-positive breast cancer[13](BCa). ER-negative tumours can either be defined as human epidermal growth factor receptor 2 (HER2) positive or basal/triple negative (TNBC), the latter being tumours that do not express ER, PR or HER2[10,14].
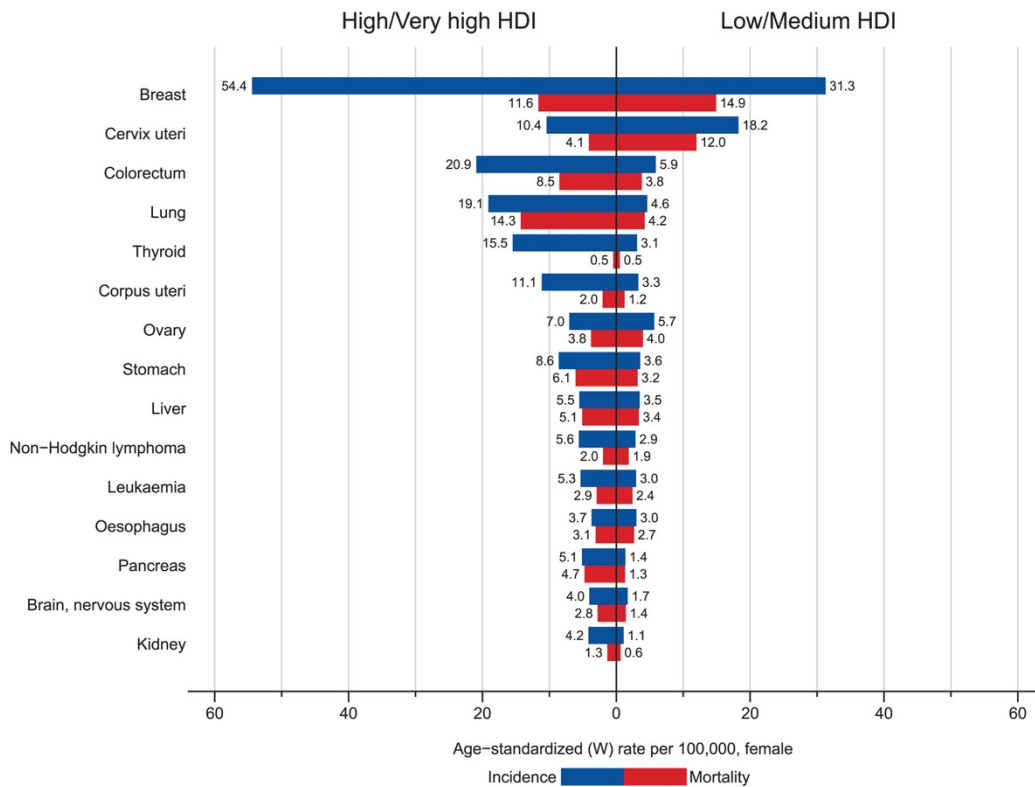
Figure 1.2: Breast cancer incidence and mortality. Age-Standardized Rates of incidence (blue) and mortality (red) in High/Very-High Human Development Index (HDI) Regions Versus Low/Medium HDI Regions Among Women in 2018. The 15 most common cancers world (W) in 2018 are shown in descending order of the overall age-standardized rate (Modified from GLOBOCAN 2018[1]).

Pathological characterisation of BCa has been so far categorised into various molecular subtypes including:

1. Luminal A: oestrogen receptor (ER) positive (+), progesterone receptor (PR) positive and human epidermal growth factor 2 receptor (HER2) negative (-),

2. Luminal B: ER+, PR lowly expressed or absent, HER2 occasionally overexpressed with the tumour exhibiting high proliferation rates,

3. HER2+: ER-, PR- and HER2 overexpressed,

4. Basal-like: this includes the triple negative subtype ER- PR- HER2-,

5. Claudin-low that are also triple negative but resemble cells undergoing epithelial-mesenchymal transition (EMT).

The molecular subtypes of BCa all exhibit variable responses to treatment and also differ in incidence and survival. Luminal A and luminal B (i.e. ER+ positive BCa) molecular subtypes represent over two thirds of all BCa and are amenable to endocrine therapy (ET) which aims to target the ER[15]. Such approaches have proved very successful, and together with screening programs aiding earlier detection[16], have led to substantial reductions in breast cancer mortality[1]. Despite this, many patients eventually relapse with tumours that have developed resistance to endocrine therapy, including treatment with anti-oestrogens[17–19] and aromatase inhibitors[20–22]. It is now clear that intra-tumour heterogeneity should be taken into account since, generally, not all cells in breast cancer share subtypes.

1.1.2 The oestrogen action in physiological human development and in breast cancer

Oestrogens are primarily synthesised in the ovaries in pre-menopausal women, through aromatisation of the precursor steroid hormones androstenedione and testosterone to estrone (E1) and 17 β-oestradiol (E2), respectively, by the P450 aromatase enzyme[23]. Regulation of this process involves the combined actions of the gonadotropin-releasing hormone (GnRH), luteinizing hormone (LH) and follicle-stimulating hormone (FSH). The release of androgen from theca cells and aromatase expression in the granulosa cells of the ovary is stimulated by LH and FSH, respectively; the latter are both positively regulated by GnRH[24]. Following release from the ovaries, oestrogen circulates through the bloodstream to reach distal target sites, including the mammary gland and reproductive organs, which require oestrogens for their development and function. Oestrogens also play important roles in the cardiovascular, skeletal and central nervous systems[25–27]. Oestriol (E3), another endogenous oestrogen primarily found during pregnancy, can be produced in the liver via conversion from E2[23].

Following onset of menopause, the ovaries stop producing E2, and instead E1, produced locally through aromatisation of adrenal androgens at extragonadal sites such as bone, brain and adipose tissue, becomes the dominant form of oestrogen[28]. E1 can be converted to E2 by 17β-hydroxysteroid dehydrogenase (17β-HSD)[29]. Nevertheless, due to the lower levels of protective circulating E2, postmenopausal women are at a greater risk of dementia, osteoporosis and cardiovascular disease[30].

Oestrogens exert their effects primarily through the action of two receptors, ERα and ERβ, members of the nuclear receptor superfamily of small molecule activated transcription factors. ERα and ERβ are encoded within distinct genes on human chromosomes 6 and 14, respectively. E2 is the highest affinity ligand for both receptors[31]. Both genes are expressed in the normal mammary gland, with ERα localised to ductal epithelial cells and ERβ expression being more widespread in myoepithelial, endothelial and stromal cells, as well as in epithelial cells[32]. Mammary gland development is dependent on the action of oestrogens and ERα, which is demonstrated by ERα and CYP19A1 aromatase gene knockout mouse models; impaired mammary duct formation as well as breast tissue development seen beyond the pre-pubertal stage[33,34]. ERα knockout mice also showed effects on the skeletal and cardiovascular systems, with altered behaviour and complete infertility seen in both genders[33]. However, knockout of ERβ had no impact on breast development but did affect optimal ovulation efficiency[35].

Although ERα and ERβ both bind to the same DNA response element sequence, they have differing effects on target gene expression[36,37]. In BCa, it is the ERα subtype, which has been shown to be responsible for driving breast cancer cell growth, whereas ERβ seems to have a tumour-suppressive action[38,39]. Importantly, immunohistochemical assessment of breast cancer biopsies is used to determine ERα status that, in turn, is used to determine if a patient is offered endocrine therapies[40,41]. Moreover, the successful treatment of ER-positive disease with endocrine therapies directed at inhibiting ER activity provides the strongest evidence of the importance of oestrogen and ER activity in breast cancer progression[18].

1.1.3 The treatment of breast cancer

Achieving long-term local disease control coupled with minimal local morbidity, are the goals of local treatment of BCa. Breast conserving surgery (BCS) or mastectomy, with or without axillary node clearance, are the usual surgical modalities employed in BCa treatment. Surgery aims to completely excise the BCa tumour, aiming for at least microscopically disease-free margins[42]. The application of radiotherapy (DXT) following surgery in BCa remains one of the fundamental adjuvant treatment modalities[43]. After BCS, DXT to the conserved breast halves the rate at which the disease recurs and reduces the breast cancer death rate by about a sixth[44]. Since Beatson's observations over a century ago, that the induction of regression in advanced BCa could be propagated by bilateral oophorectomy[45], endocrine therapy (ET) has shown to be one of the most fundamental treatment modalities in cancer medicine[46]. Endocrine therapy forms part of the key treatment modalities in the management of ERα-positive BCa. ET can be

administered in the pre-operative (neoadjuvant), post-operative (adjuvant) and in the advanced (metastatic) disease setting. The aims of current ET are to control or inhibit peripheral oestrogen production or the function of ERα within the BCa cells[46] (Fig. 1.3).

Patients diagnosed with ERα-positive BCa in the pre-menopausal setting are managed with postoperative Tamoxifen (a selective oestrogen receptor modulator (SERM)) for 5 years. Nonetheless, results from a randomised trial showed a further reduction in recurrence and mortality with adjuvant Tamoxifen treatment prolonged to 10 years[49]. Ovarian suppression, which can be achieved surgically or through the use of gonadotrophin releasing hormone (GnRH) analogues, in addition to Tamoxifen, chemotherapy or both, is associated with reduced disease recurrence and death after recurrence in premenopausal patients with ERα-positive BCa[50].
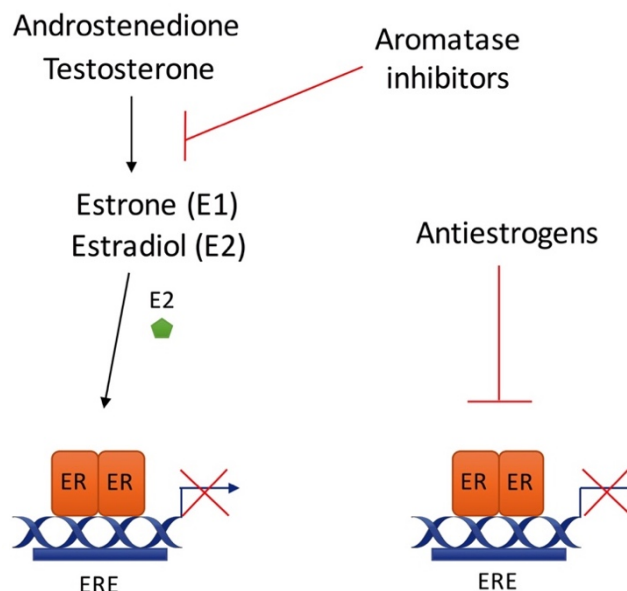
Figure 1.3: Endocrine therapy: mechanisms of action. Aromatase inhibitors and antiestrogens prevent ER activity through different mechanisms. Aromatase inhibitors prevent the aromatisation of peripheral androgens into oestrogen, thus preventing ER activation. Antiestrogens, such as Tamoxifen, directly target the ER competing with oestrogen for binding to the receptor and inhibiting its activity (Figure created by Alison Harrod).

In the postmenopausal setting, BCa patients who have undergone surgery are usually started on either adjuvant Tamoxifen or Aromatase Inhibitors (AIs). Aromatase inhibitors, such as exemestane, letrozole and anastrozole, prevent oestrogen biosynthesis by competing with androgens for the active site of aromatase CYP19A1. AIs have been associated with an increased disease-free survival compared to Tamoxifen alone[51]. The combination of

Tamoxifen and/or AI (i.e. Tamoxifen for 2-3 years followed by AI to complete a total of 5 years of treatment) or monotherapy (i.e. Tamoxifen or AI alone) is administered for 5 years [46]. However, after 5 years of adjuvant ET, breast cancer recurrences continued to occur steadily to up to 20 years with risks ranging from 15 to 49%, depending on nodal status (TN) and tumour grade[52] (Fig. 1.4). Many mechanisms have been considered as potential causes of the eventual resistance to endocrine therapy, which are found to be either intrinsic to the patient or acquired during therapy[22,53,54]. Early detection or prediction of mechanisms leading to resistance could greatly benefit patients, as pre-existing minor clones expand and/or the tumours evolve through genetic change over time, decreasing the likelihood of a cure[54,55]. Neoadjuvant ET is used to decrease tumour size in large operable BCa tumours, in order to apply BCS instead of mastectomy. Although Tamoxifen or AI can be administered in the neoadjuvant setting for ERα-positive BCa, chemotherapy (NACT) remains the first-line therapy to down-stage BCa tumours irrespective of ER positivity[56]. However, recent evidence showed that tumours downsized by neoadjuvant chemotherapy might have higher local recurrence after breast-conserving surgery than tumours of the same dimensions in women who have not received NACT[57].

Despite advances in adjuvant therapy, metastatic BCa remains a major challenge. In some cases, the hormone receptor status between the primary and metastatic disease can be completely different and hence undertaking a biopsy of the recurrence would aid the physician in tailoring more focal treatment[46]. Depending on the metastatic BCa tumour hormone status and the patient's wellbeing, proposed treatments include, ET with Tamoxifen, third generation AIs (which have shown to be more beneficial than Tamoxifen in recurrent BCa), Trastuzumab (a monoclonal antibody to the HER2) usually administered to HER2 as well as hormone receptor positive metastatic BCa patients and Fulvestrant (a 7α- alkylsulphinyl analogue of 17β-oestradiol) which is a selective ER downregulator (SERD) [18,41,46]. Fulvestrant is unique among approved ER therapeutics due to its capacity for full ER antagonism, thought to be achieved through ER degradation. However, it has been recently shown that optimization of ER degradation does not guarantee full ER antagonism in breast cancer cells that, indeed, still display transcriptional activities and anti-proliferative potential[58]. The same group demonstrated that, contrarily to what was thought to be the main mechanism of action, fulvestrant-like antagonists significantly slow ER intra-nuclear mobility rather than eliminate ER; in turn, ER immobilization leads to the increased ER turnover. These findings open new

avenues for the development of novel compounds able to perturbate transcription factor mobility.

A Risk of Distant Recurrence

Distant Recurrence (%) vs Years (0–20)

N4–9: 22 (5), 36 (10), 45 (15), 52 (20)
N1–3: 10 (5), 19 (10), 25 (15), 31 (20)
N0: 6 (5), 11 (10), 16 (15), 22 (20)

No. at Risk					
N4–9	12,333	8,116	2165	259	52
N1–3	31,936	23,576	7250	949	183
N0	29,925	24,081	8571	1982	414

No. of Events — annual rate (%)				
N4–9	2568 (4.8)	969 (4.0)	121 (3.1)	13 (2.2)
N1–3	3126 (2.2)	1421 (1.9)	241 (1.7)	39 (1.8)
N0	1646 (1.2)	835 (1.1)	272 (1.3)	68 (1.4)

B Risk of Death from Breast Cancer

Death from Breast Cancer (%) vs Years (0–20)

N4–9: 12 (5), 29 (10), 40 (15), 49 (20)
N1–3: 5 (5), 14 (10), 21 (15), 28 (20)
N0: 3 (5), 8 (10), 12 (15), 15 (20)

No. at Risk					
N4–9	12,333	9,079	2481	294	57
N1–3	31,936	24,866	7728	1011	197
N0	29,925	24,819	8926	2144	476

No. of Events — annual rate (%)				
N4–9	1463 (2.6)	1154 (4.1)	185 (3.7)	20 (2.3)
N1–3	1600 (1.1)	1506 (1.9)	319 (1.9)	52 (1.8)
N0	826 (0.6)	890 (1.0)	228 (0.8)	77 (1.0)

Figure 1.4: Risk of recurrence and death from breast cancer. Association between Pathological Nodal Status and the Risk of Distant Recurrence (Panel A) or Death from Breast Cancer (Panel B) after 5 years of adjuvant endocrine therapy during the 20-Year Study Period. The risk was calculated according to the patients' pathological nodal status at the time of diagnosis: N0, N1–3, or N4–9 (Modified from Pan H 2017[52]).

1.1.4 Hallmarks of cancer

Cancer is postulated to arise from one single cell. The oncogenic transformation of a normal cell into a malignant one, is a multistage process involving contributory changes which lead to the gradual conversion of normal human cell processes to an oncogenic state[59,60]. Advances in cancer research over the past decades have elucidated the stages of development and progression. Cancer cells are not just thought of as a mass of uncontrolled proliferating cells; they are also composed of various cell types, such as normal cells contributing to the tumour-associated stroma which are involved in initiation, development and finally metastasis. Six common traits outlining the hallmarks of cancer were initially proposed by Hanahan and Weinberg in 2000[59]. These include: 1. Self-sufficiency in growth signals, 2. Insensitivity to anti-growth signals, 3. Evading programmed cell death, 4. Limitless replicative potential, 5. Developing blood vessels and 6. Tissue invasion and metastasis. These aforementioned traits were then revisited in 2011 following which, Hanahan and Weinberg further added 2 hallmarks

with 2 enabling characteristics; 1. Deregulated metabolism, 2. Evading the immune system 3. Genome instability and 4. Inflammation[60] (Fig. 1.5).
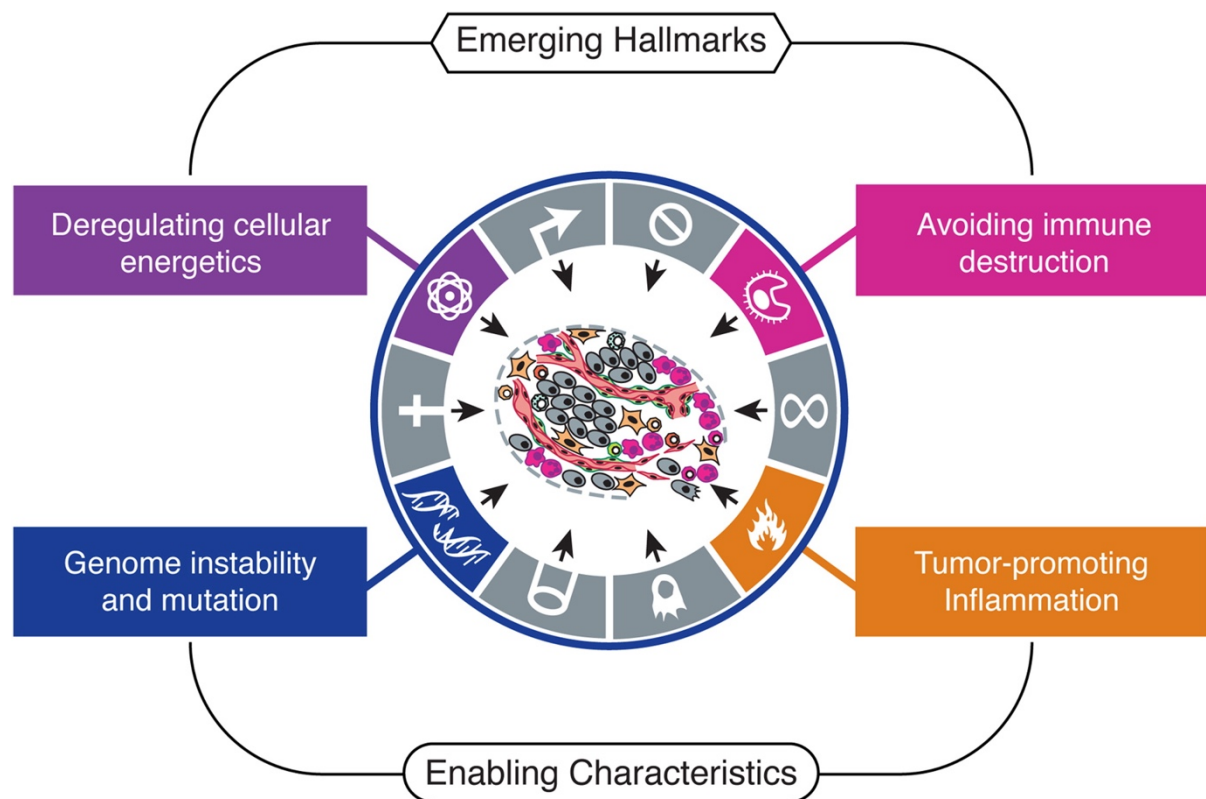
Figure 1.5: Emerging Hallmarks of cancer. An increasing body of research suggests that two additional hallmarks of cancer are involved in cancer pathogenesis: capability to reprogram cellular metabolism in order to most effectively support neoplastic proliferation and evasion of immunological destruction[60] (Modified from Hanahan and Weinberg, 2011).

1.1.5 Mechanisms of metastatic dissemination

Tissue invasion and metastasis, one of the hallmarks of cancer, is the main cause of tumour-related death. Through a complex, multistep process cancer cells detach from the primary site, migrate through surrounding tissues, access and travel through the vasculature or the lymphatic system, arrest at a distant organ, survive in a new environment and start a metastatic colonisation[61,62].

Despite decades of study, the process of tumour metastasis remains controversial with several open questions regarding the nature of the mechanisms of dissemination. Furthermore, therapeutic advances in oncology have not fully translated to the treatment of metastatic

disease, which remains almost always (with the notable exception of testicular cancer) incurable and the main cause of cancer-related deaths. Therefore, the need to understand the cellular and molecular features of the cancer cells acquiring the ability to spread to distant organs.

Although during development and tissue repair epithelial tissues are highly dynamic and migratory, in physiological conditions adult epithelial cells, from which carcinoma arise, are polarised and nonmotile adding a further challenge to the metastatic seeding process[63]. Conventional models support spread of tumours like triple negative breast cancer (TNBC) through single-cell invasion typical of epithelial-to-mesenchymal transition (EMT)[64] (Fig. 1.6). Epithelial cells are characterised by specific intercellular adhesion complexes. These junctions are often thought to impede motility; thus, it is conceptually attractive to think of metastases as involving a transient or permanent loss of epithelial features, through a process such as the EMT[64]. Nonetheless, an increasing body of literature has shown a model for metastatic spread of BCa in which epithelial adhesion complexes play an important role in metastatic tumour cells, a process known as collective invasion [67-69, 276-277, 354]. Cell clusters held together through tight intercellular adhesion molecules can greatly contribute to the dissemination and multi-clonal metastatic seeding to secondary sites [63] (Fig. 1.6). This idea has already been suggested in the 1950s, when some reports showed blood samples from cancer patients containing both single and clustered tumour cells and that cell clusters can rapidly disseminate to the lungs and produce metastasis more efficiently than single cells in animal models[62,65,66]. Direct evidence has also been provided that clusters exhibit superior survival and colony-forming potential both in culture[67] and *in vivo*[67,68]. Furthermore, on the day of submission of this thesis, Ewald group published a study demonstrating that the intercellular adhesion protein E-cadherin is required and promotes metastasis in mouse and human models of invasive ductal carcinomas[354].

It has been recently reported the cancer cells leading collective invasion are characterised by cytokeratins K14 and K5[67,69]. These leading cells were found in all stages of disseminative spread but were rare in the primary tumour and in the macro metastasis[67,69], suggesting different epithelial molecular programs driving tumour growth and dissemination. In this model, disseminating breast cancer clusters retain cytokeratins together with desmosomal adhesion proteins, and more in general the epithelial cell adhesion machinery, and do not require EMT to accomplish metastasis[70].

Further questions arise from these observations, including at what stage and how tumour cells escape from the primary tumour, how is this related to chronic medical treatment and what their molecular properties are as they transit to the distant site.

Dissecting the relationship between chronic exposure to treatment and activation of the metastatic program in the clinical setting is challenging. Surprisingly, patients that present with synchronous metastasis at diagnosis still experience significant response to endocrine treatments. *In vitro* models well reflect this response, indeed almost all ERα positive breast cancer cell lines have been so far derived from untreated metastatic patients but display significant response to targeted treatment. In addition, recent reports have highlighted how breast cancer cells can disseminate very early during tumorigenesis[47,48]. These observations lead to the hypothesis that the metastatic program does not necessarily predate the drug-resistance program; however, previous reports indicate that drug resistance might influence the development of metastatic potential[53,55]. For example, a specific drug-induced epigenetic reprogramming is activated in cells acquiring resistance to AI but not in cells developing resistance to Tamoxifen[53]. Resistance to AI treatment is partly driven by epigenetic activation of the cholesterol biosynthesis pathway and chromatin accessibility data in AI treated cells show enrichment for the sterol response elements (SREs)[53].
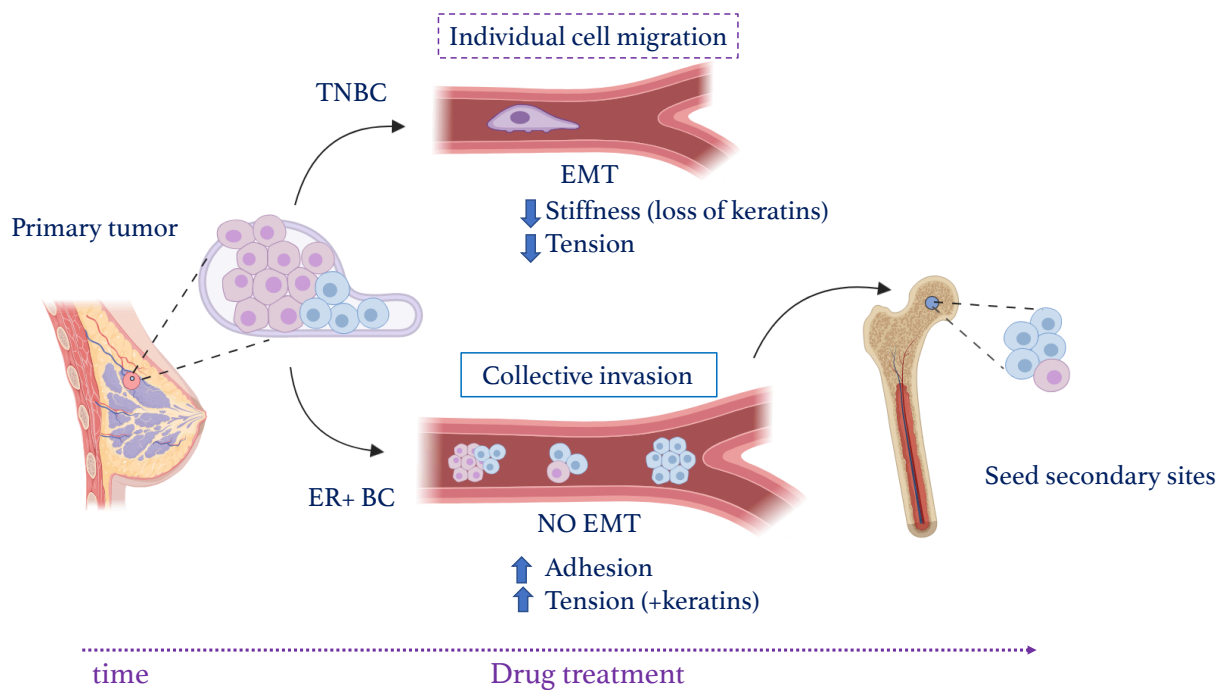
Figure 1.6: Mechanisms of metastatic invasion in breast cancer. Whereas TNBC spread through single-cell invasion typical of EMT, recent studies have shown a model for metastatic spread of ER breast tumour cells that is based on collective invasion.

1.2 SREBPs in physiology and disease

1.2.1 Sterol Regulatory Element Binding Proteins – molecular features

1.2.1.a Genes, isoforms and structure

The Sterol-regulatory Element Binding Proteins (SREBPs) are a family of transcription factors regulating lipid homeostasis by controlling the expression of downstream enzymes involved in endogenous cholesterol, fatty acid (FA), triacylglycerol (TCA) and phospholipid (PL) biosynthesis[71]. SREBPs were first described more than 25 years ago by Brown, Goldstein and colleagues as the master regulators of cholesterol and lipid biosynthesis[72, 73].

SREBPs are basic-helix-loop-helix-leucine zipper (bHLH-LZ) transcription factors (TF) synthetized as inactive precursors bound to the membrane of the endoplasmic reticulum (ER)[74] (Fig. 1.2.1a). The inactive form is constituted by three domains: i) a NH2-terminal domain containing the transactivation domain, ii) two hydrophobic transmembrane segments connected by a short loop and iii) a COOH-terminal regulatory domain[75] (Fig. 1.2.1a). There are three members in the SREBP family sharing 47% of homology. SREBP-1a and 1c are produced from the transcription of different promoters of the *SREBF1* gene[74] on chromosome 17p11.2, SREBP2 is the transcript of the *SREBF2* gene[76] located on chromosome 22q13 (Fig. 1.2.1a). SREBP1a is a more potent transcriptional activator than 1c because of its longer NH2-terminal transactivation domain that is able to strongly bind to cAMP-response element binding protein (CREB)-binding protein (CREBBP)[77]. SREBP1c is widely expressed in most tissues, whereas the isoform 1a is predominantly expressed in highly proliferative cells such as macrophage, spleen and intestine[78].

SREBP transcription factors (TF) are finely regulated to sense cellular energy states, playing a key role in cellular energy homeostasis[75].
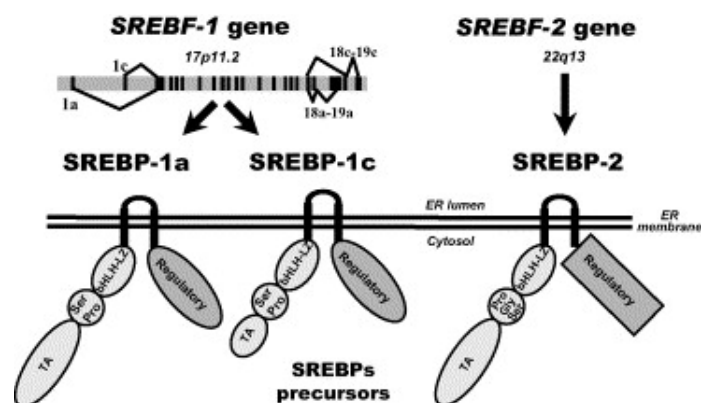
Figure 1.2.1a (previous page): SREBPs genes and structure. SREBPs family is composed of three members: SREBP-1a and 1c produced from a single gene (*SREBF-1)* on human 17p11.2 and SREBP-2 from a separate gene (*SREBF-2)* on human chromosome 22q13. SREBPs are bHLH-LZ transcription factors synthetized as inactive precursors bound to the ER membranes. Each SREBP precursor is organized into three domains: (a) an NH_2-terminal domain containing the transactivation domain and the bHLH-LZ region for DNA binding and dimerization; (b) two hydrophobic transmembrane spanning segments interrupted by a short loop that projects into the lumen of the ER; and (c) a COOH-terminal segment regulatory domain (Adapted from Eberle 2004 [71]).

1.2.1.b Regulation: proteolytic cleavage, transcription, post-translational modifications and transcriptional coactivators

From yeasts to humans SREBPs are highly conserved, therefore the expression of lipogenic genes is regulated according to species-specific requirements[75]. As such, SREBP is regulated by palmitate in Drosophila[79], by hypoxia in fission yeast[80] and by sterols in mammals[81].

Different isoforms play different roles in the physiological modulation of lipid synthesis[71]. SREBP1a strongly activates global lipid synthesis and growth, whereas SREBP1c primarily controls energy storage through nutritional regulation of FA and triglycerides. SREBP2 mediates cholesterol metabolism-related gene expression[82–84]. However, when overexpressed, the isoforms exhibit functional overlap.

Key events in the activation and regulation of SREBPs involve several steps of trafficking between cellular compartments such as cleavage, recycling and degradation. SREBPs normally reside in the ER in complex with SCAP (SREBP cleavage-activating protein) and INSIG (insulin-induced gene)[85–89] (Fig. 1.2.1b). In response to sterol depletion, SREBP-SCAP migrate to the Golgi and, through the sequential action of the Golgi-localised Site-1 and Site-2 Proteases (S1P, S2P respectively), the N-terminal domain is proteolytically released[75,90] (Fig. 1.2.1b). The cleaved SREBP then translocates into the nucleus where it binds to the promoter of several genes involved in cholesterol synthesis and uptake, thus restoring sterol homeostasis in a feedback regulation loop[73,81,91–93] (Fig. 1.2.1b). SREBPs are also self-regulated by a transcriptional positive feedback[94–96].

Cleavage and trafficking of SREBPs precursors end with the nuclear translocation of the mature form of SREBPs. It has been demonstrated that the mature forms of SREBPs are modified by phosphorylation[97–101], acetylation[102], sumoylation[103], and ubiquitination[97,99,104,105]. Not only mature but, also SREBPs precursor forms are subject to proteasome-dependent degradation via ubiquitylation. Heat shock protein (HSP) 90 regulates SREBP by binding to

and stabilizing SCAP-SREBP complex; inhibition of HSP90 leads to proteasome-dependent degradation of SCAP-SREBP protein[106]. Furthermore, after dissociation from the complex SCAP/SREBP, Insig1 is ubiquitinated and degraded in proteasomes. Ubiquitination is not necessary for release of SCAP/SREBP from Insig1, but it establishes a requirement for synthesis of new Insig1 for feedback inhibition. When the new Insig1 and cholesterol converge on SCAP, SCAP/SREBP binds to Insig1, preventing ubiquitination[107]. As a result, treating cells with proteasome inhibitors increases nuclear levels of SREBPs and target gene expression.

SREBPs further interact with various transcriptional co-activators (TFs) such as CBP and p300, which acetylate and stabilise SREBPs by preventing ubiquitination[102,105]. These modifications regulate the stability and/or transcriptional activity of the active transcription factors. Transcriptional coactivators and cooperating TFs provide yet another level of regulatory control of SREBP activity[75]. In human hepatocarcinoma cells, SREBP1 cooperates with its associated factors, nuclear factor Y (NFY) and simian-virus-40-protein-1 (SP1), to regulate the expression of a subset of target genes through direct interaction[96,108]. To date, a variety of TFs activated in response to extracellular stimuli has been reported to modulate SREBP transcriptional activity. For instance, the Liver X receptor (LXR) which is activated by oxysterols, regulates SREBP activity by direct binding[109–111] (Chen G 2004). Liver X receptor alpha (LXRα) is a nuclear hormone receptor highly expressed in the hepatic tissue that mediates the induction of SREBP1c by insulin[109].
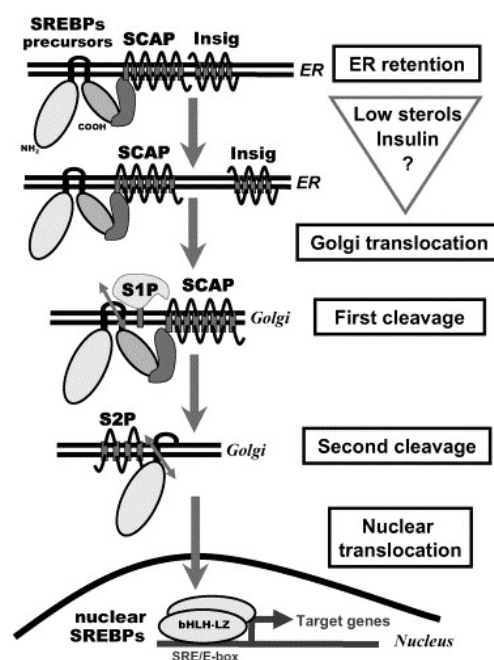
Figure 1.2.1b (previous page): SREBPs activation by proteolytic cleavage. SREBPs are tightly associated with the SCAP escorting protein, in turn interacting with the Insig proteins, which retain the SCAP/SREBP complex in the ER compartment. Upon appropriate conditions (low sterol concentrations, insulin or other stimuli), the interaction between Insig and SCAP decreases and allows the escorting by SCAP of SREBPs to the Golgi apparatus where the S1P cleaves SREBP at a luminal site. The S2P then cleaves releasing the NH2-terminal SREBPs domain (nSREBPs). This domain containing the bHLH-LZ region is then translocated to the nucleus where it will bind its target genes (Adapted from Eberle 2004 [71]).

1.2.2 Modulation of SREBPs activity and processing – physiology

1.2.2.a Nutritional regulation: Sterol-dependent activation

Physiologically, sterols control SREBP activation by modulating its ER-to-Golgi transport. SCAP, forms a complex with SREBP to exit the ER, binds COPII vesicles coat proteins (SEC23 and SEC24) and transports to the Golgi. When SCAP is depleted, SREBPs become unstable and easily broke down with a consequent marked decline in the expression of target genes[87] (Fig. 1.2.2a). When cholesterol levels increase, sterols directly bind SCAP preventing the complex to be embedded on COPII vesicles[88,89]. Excess ER cholesterol also promotes binding of SCAP to additional ER-retention membrane proteins, namely Insig[89] (Fig. 1.2.2a). Insulin-induced gene 1 protein (Insig-1) and Insig-2, control tissue- and signal-specific regulation of SREBP transport. They derive their name because *INSIG1* was initially identified as a gene highly induced by insulin[88] whereas INSIG2, primarily expressed in the liver, is downregulated by insulin. *INSIG1* is a direct SREBP target gene and it binds oxysterols such as 22-, 24-, 25-, 27-hydroxycholesterol[112]. Oxysterol derivatives of cholesterol accumulate under conditions of excess of cholesterol and bind to Insig1 to promote ER-retention of SCAP-SREBP[89,112–114] (Fig. 1.2.2a). Insig proteins accelerate the degradation of SREBP target gene 3-hydroxy-3-methylglu-taryl-CoA reductase (*HMGCR*), further regulating cholesterol homeostasis. SREBP activation both increases cholesterol and Insig1 providing a convergent, negative feedback regulation of the SREBP-SCAP complex transport and proteolytic activation[107].
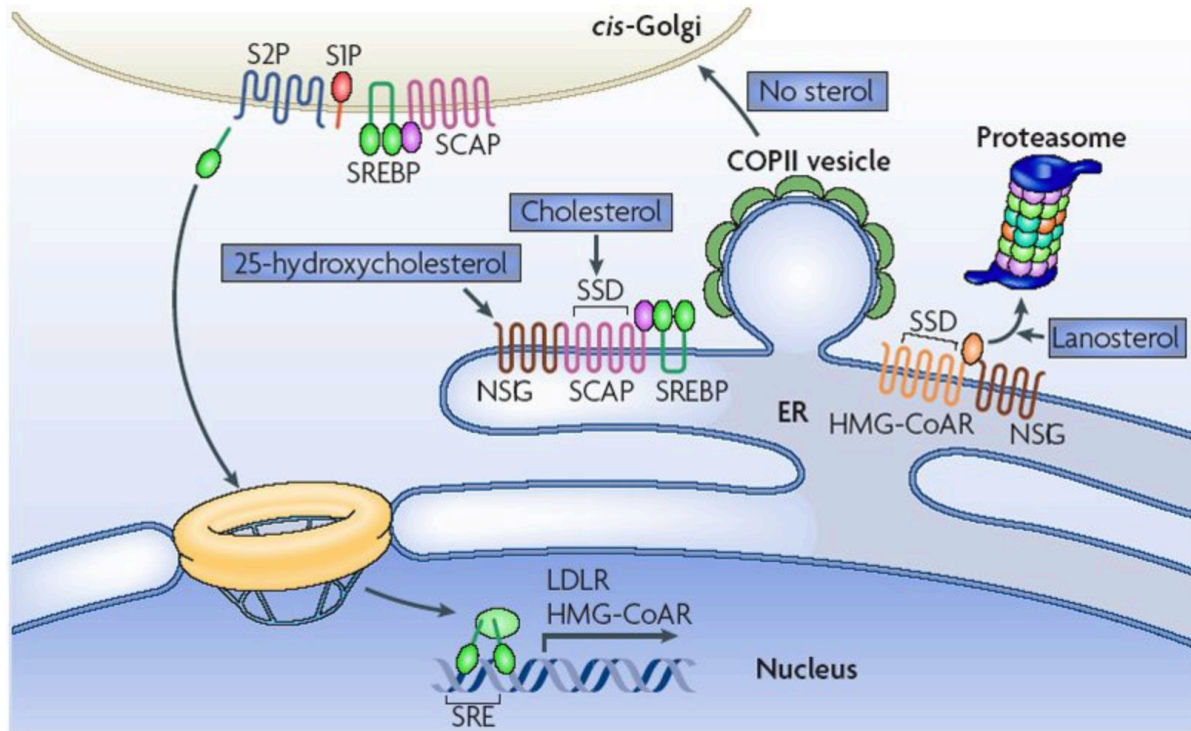
Figure 1.2.2a: SREBPs sterol-dependent activation. Under conditions of ample sterol in the ER, INSIG prevents entry of the SREBP–SCAP complex to COPII-coated vesicles. SREBP transported to the nucleus activates sterol-regulated genes (such as *HMGCR* and the low-density lipoprotein receptor (*LDLR*)). HMGR is also post-transcriptionally regulated by sterol, with INSIG binding of the protein leading to its proteasomal degradation. SSD, sterol-sensing domain; SRE, sterol regulatory element (Adapted from Ikonen 2008[91]).

1.2.2.b Insulin-dependent regulation: the PI3K-AKT-mTOR-SREBP pathway

In contrast to sterol regulation of SREBP2, the gene expression and activity of SREBP1 are dependent on the energy state[81] and partially attributed to mammalian target of rapamycin (mTOR)[115]. mTOR is a protein kinase and, as part of the mTOR complex 1 (mTORC1), senses local and systemic nutrients through insulin signalling. mTORC1 plays a key role in cell growth, survival, ageing and metabolism promoting protein synthesis, inhibiting autophagy and as a central regulator of lipid homeostasis[116–118]. In particular, functional importance of its signalling has been demonstrated in controlling mammalian lipid metabolism including lipid synthesis, oxidation, transport, storage and lipolysis, as well as adipocyte differentiation and function[119]. Overall, mTOR functions as a critical anabolic signal integrator between three major nutritional pathways: glucose, protein and lipid metabolism[116].

Lipogenesis is an anabolic process activated in nutrient-rich states and insulin is a major anabolic hormone. Insulin activates hepatic SREBP1c both transcriptionally and post-translationally leading to the upregulation of fatty acid synthesis[75]. The effects of insulin on SREBP1c expression are mediated by mTOR, as major downstream effector of phosphatidylinositol 3-OH-kinase/ proteinkinase B (PI3K/Akt)-dependent pathway[117,120] (Fig. 1.2.2b, c). Chronic activation of mTORC1 is known to cause insulin resistance. In conditions of chronic hyperinsulinemia, mTORC1-dependent hepatic SREBP-driven lipogenesis is continuously hyperactivated. SREBP1 also exhibits insulin-independent induction through hyperglycaemia, although the role of glucose in the regulation of SREBP1c is still controversial with both no effect on and induction of transcription been reported in hepatocytes[71] (Fig. 1.2.2b, c). It has been reported that PI3K/Akt activates SREBPs[121]. Nevertheless, the precise molecular mechanisms are still controversial and differ between SREBP isoforms; proposed mechanisms include increased trafficking and processing of SREBP, reduced degradation, and involvement of the downstream signalling hub, mTORC1[121]. mTOR signalling further branches to SREBP1 and lipogenesis via: i) Hepatic p70 S6 kinase (S6K), ii) Cyclic adenosine monophosphate-responsive element (CRE) binding protein-regulated transcription coactivator (CRTC) 2, iii) phosphatidate phosphatase lipin-1 (LPIN1) and iv) E4 promoter-binding protein 4 (E4BP4) (Fig. 1.2.2d).

i)   S6K is the major downstream effector of mTORC1. It has been shown that S6K regulates cleavage-dependent activation of SREBP1. TORC1–S6K1 interaction is crucial for SREBP activation and sustained lipogenesis and hepatosteatosis under conditions of insulin resistance[122]. Moreover, S6K1 phosphorylates LXRα impairing its ability to transactivate the expression of SREBP1[115]. To what extent the S6K1-LXRα connection stimulates SREBP1 expression when mTORC1 is activated is yet unknown.

ii)   CRTC2 is the master regulator of glucose metabolism. mTORC1-mediated phosphorylation of CRTC2 facilitates SREBP1 translocation from ER to Golgi. By releasing inhibitory SEC31, mTORC1 allows the formation of SEC23-SEC24 complex to maintain COPII vesicles function providing for the intracellular shuttling of SREB-SCAP complex[123].

iii)   Lipin-1 is a phosphatase required for triglycerides (TG) synthesis and a transcriptional coactivator essential for adipose tissue development[124]. mTORC1 activates SREBP through phosphorylation of Lipin-1. In hepatoblastoma cells,

SREBP1 directly activates LPIN1 transcription, suggesting a possible mechanism for feedback regulation of SREBP1 activity[95].

iv) E4BP4 is a transcription factor involved in clock and immune regulation; it has been proposed as a downstream regulator of mTOR signalling via its activation of the AKT-mTORC1-SREBP1c pathway in hepatocytes through stabilisation of SREBP1c[125].

1.2.2.c Starvation-dependent regulation: AMPK, cAMP-PKA

AMPK (AMP-activated protein kinase) is an evolutionarily conserved protein kinase, master regulator of cellular homeostasis. AMPK coordinates cell growth, autophagy and metabolism and it is required for embryonic growth and development[126]. The metabolic sensor AMPK is physiologically activated in response to a broad range of stresses such as glucose deprivation and hypoxia providing cells with the flexibility to adapt and survive metabolic stress[126]. Once activated, AMPK maintains energy balance by switching from anabolic to catabolic pathways to generate ATP. Thus, AMPK can restrain cell growth by:

i) inhibiting protein synthesis through direct phosphorylation of mTORC1 signalling pathway. Amino acid levels can regulate and activate SREBP expression through mTORC1 in the lysosome. Both protein synthesis and autophagy can be regulated in this way. A sensor of amino acids deficiency responds to nutrient deprivation by suppressing protein translation and decreasing levels of SREBP1 and lipogenic enzymes[125,126].

ii) blocking FA and cholesterol biosynthesis through direct phosphorylation of the enzymes HMGR and acetyl-CoA carboxylase 1 (ACC1) and inhibition of SREBP, required for new membrane formation in proliferating cells. Phosphorylation of SREBP1c by AMPK is necessary for inhibition of proteolytic processing and transcriptional activity of SREBP1c[126–128].

iii) inducing cell-cycle arrest and apoptosis through stabilization of p53, regulation of the cyclin-dependent kinase inhibitors p21, and of the Yes-associated protein (YAP) hippo signalling pathway, while promoting cell survival mechanisms during metabolic stress[126,128].

A further mechanism of starvation-dependent regulation is provided by the cAMP-PKA pathway. The protein kinase A (PKA) is a family of enzymes whose activity is dependent on cellular levels of cyclic AMP (cAMP), thus PKA is also known as cAMP-dependent protein

kinase. PKA links extracellular starvation signals, such as glucagon and adrenaline, to adaptive responses to energy depletion. PKA inhibits lipogenesis by: i) downregulating SREBP1c gene expression[127], ii) phosphorylating and disrupting the DNA-binding activity of SREBP1[129,130], iii) phosphorylating upstream LXR[131].

1.2.2.d Additional signals to SREBP: PUFA

In addition to sterol, polyunsaturated fatty acids (PUFA) control SREBP1 activity by altering its exit from ER. PUFA decrease the proteolytic processing of SREBP1 by stabilising Insig1[95,132–134] (Fig. 1.2.2 b, c). Ingestion of PUFA reduces hepatic SREBP1c activity decreasing lipogenesis and plasma TG. PUFA-dependent inhibition occurs at multiple levels: i) the primary mechanism is through suppression of proteolytic cleavage, ii) decreased transcription, iii) accelerated mRNA decay, iv) proteasomal degradation of nuclear SREBP1c[95,125,132–134]. It has been suggested that PUFA-mediated SREBP1 cleavage may be regulated in an ER-to-Golgi transport-independent manner. This indicates that SREBP cleavage could be activated in the absence of ER-to-Golgi transport by controlling the localisation of SP1 and SP2. Another study showed that, by blocking phosphatidylcoline (PC) synthesis, the proteolytic activation of SREBP1 is stimulated. This further adds to mechanisms of SREBP regulation that are independent of sterol-mediated trafficking [95,132–134].

Collectively, combined regulation of SREBP transcription, proteolytic activation, and nuclear activity allows SREBPs to integrate nutrient signals from multiple pathways to control metabolism.
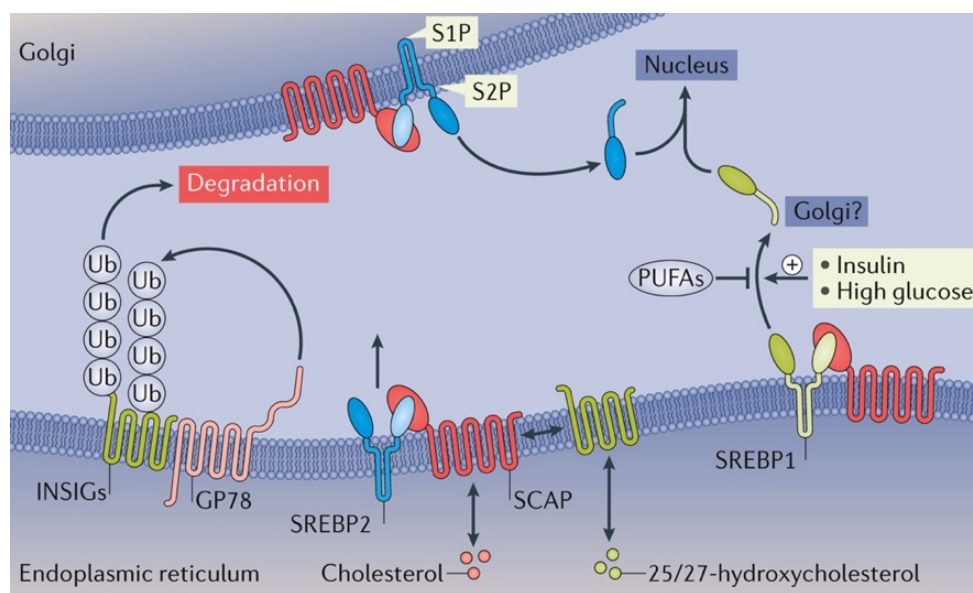
Figure 1.2.2b (previous page): SREBPs activation and degradation. In the presence of cholesterol and oxycholesterols (25-hydroxycholesterol and 27-hydroxycholesterol), the SREBP2–SCAP complex is retained in the ER together with INSIGs. In the absence of sterols, INSIGs become ubiquitylated (Ub) and are rapidly degraded. Proteolysis of SREBP1 is not strongly sterol-regulated, but rather is inhibited by PUFAs and induced by insulin or high-glucose conditions. SREBP1 activation remains incompletely understood. GP78, E3 ubiquitin-protein ligase (Adapted from Shimano, H. & Sato, R. 2017 [125]).
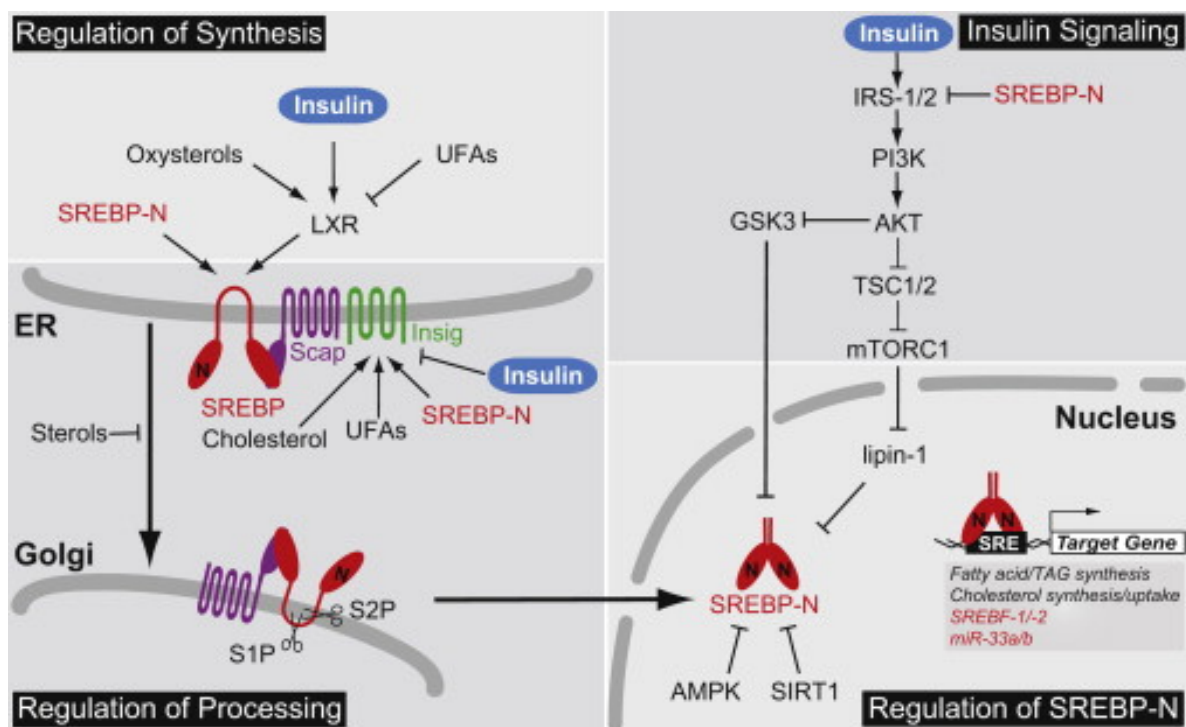
Figure 1.2.2c: Multivalent Regulation of SREBPs. Regulation of SREBPs occurs at the level of SREBP synthesis, proteolytic activation, transcriptional activity, and degradation. In addition, nuclear SREBPs are highly regulated by posttranslational modification, including phosphorylation, acetylation, and ubiquitinylation. Nuclear form of SREBP, SREBP-N; unsaturated fatty acids, UFAs; triacylglyceride, TAG (Adapted from Shao and Espenshade 2012 [95]).
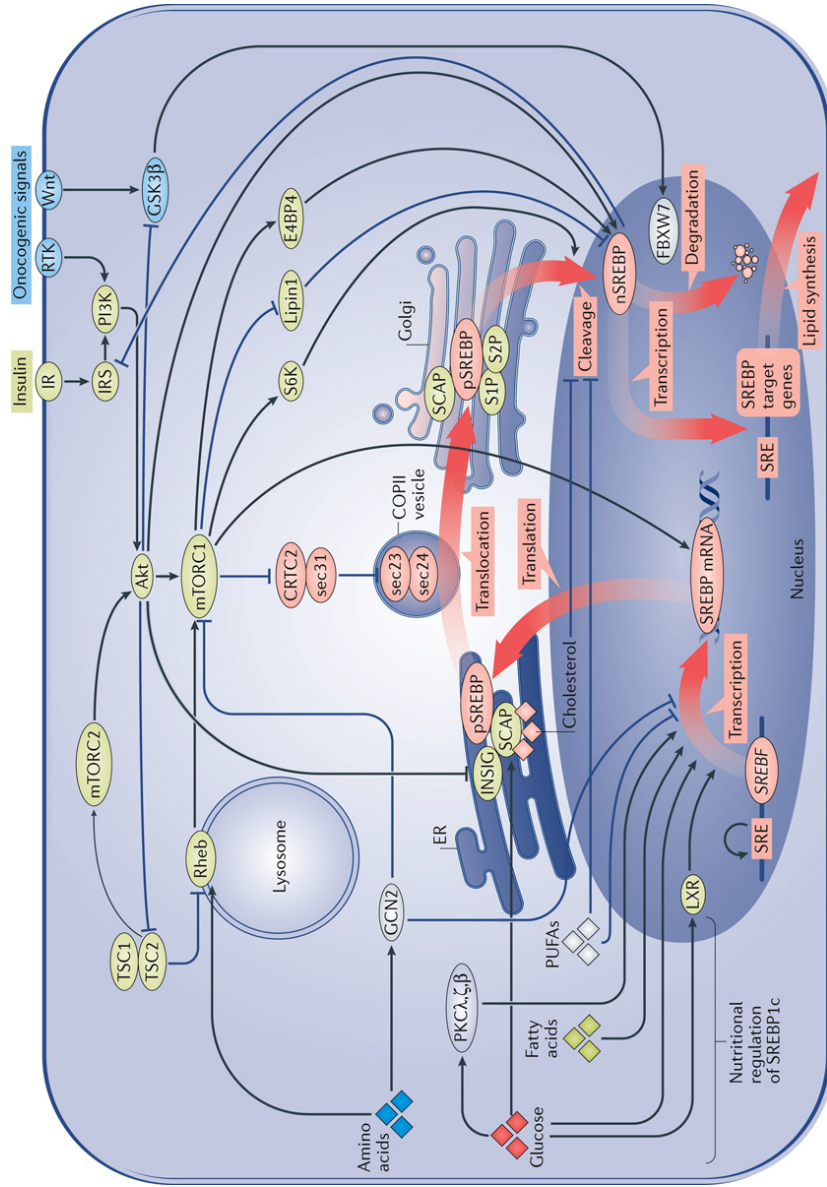
Figure 1.2.2d: Nutritional and growth signalling to SREBPs. Anabolic states activate SREBP-mediated lipogenesis via the PI3K–AKT pathway and further link to the mTOR pathway. mTOR signalling branches to SREBP1 and lipogenesis via ribosomal protein S6 kinase (S6K), phosphatidate phosphatase lipin-1, CREB-regulated transcription coactivator 2 (CRTC2) and E4 promoter-binding protein 4 (E4BP4). Amino acids and energy hub molecules also influence mTOR–SREBP pathways. Other nutrients, liver X receptor (LXR) and atypical protein kinase C (PKC) isoforms (λ, ζ and β) activate the SREBP1c promoter for nutritional regulation. Signals involved in energy depletion, such as cAMP–protein kinase A (PKA), 5′-AMP-activated protein kinase (AMPK) and sirtuins, generally inhibit SREBP1c expression. (Adapted from Shimano, H. & Sato, R. 2017[124]).

1.2.2.e Emerging roles for SREBP

Fine-tune regulation of lipid homeostasis is essential for cell viability. Hepatocytes are normally non-dividing and frequently challenged by fluctuations in lipid supply and demand. Thus, in the liver, lipids and insulin are the dominant signals to control SREBP-dependent lipogenesis. However, in non-hepatic cells, lipids level may not change dramatically, thereby minimising the sterol-dependent regulation of SREBP. In these settings, other signals may dominate, revealing new models of sterol-independent regulation of SREBP. Consistent with this, the application of ChIP-seq technology to the SREBP pathway has revealed that SREBPs control transcriptional programs extending beyond lipid synthesis[108]. Genome-wide analysis and other techniques and new tools became available in the past decades, facilitating the study of SREBPs roles, broadening their spectrum of action to non-hepatic tissues and demonstrating new wider functions in metabolism and beyond. The diverse mechanisms of SREBP regulation provide ample opportunity for context-specific regulation of SREBP. Given their critical roles, a great interest developed during the years leading to extensive studies of SREBP functions in the context of physiology and disease. Emerging roles played by SREBP have been described in apoptosis, inflammation, immune system, liver disease, diabetes and circadian rhythm.

Whether SREBPs promote or inhibit apoptosis is controversial and it seems to be both cell- and context dependent. For instance, SREBP has been found to be both protective and cause of death in pancreatic β cells[125]. Caspase 3, a major apoptosis effector, is one of the proteases activating SREBPs independently of sterol regulation suggesting a role for SREBP in the apoptotic cascade. However, SREBPs are also activated upon bacterial toxin challenge directly regulating the expression of anti-apoptotic genes to promote cell survival[95]. Indeed, an anti-apoptotic role in macrophages has been reported for SREBP1a during the proinflammatory phase of the immune response[125] (Fig. 1.2.3). The immune system uses phagocytosis as a major mechanism to remove pathogens. Phagocytosis promotes membrane biogenesis by activating processing of both SREBP1a and SREBP2[75] (Fig. 1.2.3). SREBP1a is highly expressed in cells of the immune system, such as macrophages and dendritic cells (Fig. 1.2.3). ChIP experiments in macrophages indicate that LPS (lipopolysaccharide stimulating lipogenesis) enhances activation of SREBP1a promoter to increase lipogenesis. In addition, SREBP1a activates a component of the inflammasome displaying a role in the inflammatory response (Fig. 1.2.3). Inflammatory factors such as tumour necrosis factor and hepatitis C virus infection induce SREBP1c, suggesting a mutual bidirectional interaction between SREBP1c activation and ER

stress and inflammation in the hepatic tissue. Besides, the presence of a feedback clock has been proposed since hepatic SREBP1 function is also influenced by circadian rhythm. Indeed, SREBPs have been involved in the regulation of metabolic circadian rhythm via interactions with clock genes encoded proteins[135]. Whereas nutrients regulate mainly SREBP1 subcellular trafficking, circadian clock also influences SREBP1 promoter activity through protein interactions and stability[125]. The link between SREBP and the circadian clock has been described in cancer as well. Lung cancer in mice can distally induce the endogenous circadian reorganisation in the liver, associated with the disruption of AKT-AMPK-SREBP signalling pathway[136].

1.2.3 Modulation of SREBPs activity and processing – pathology

1.2.3.a The PI3K-AKT-mTOR-SREBP pathway in cancer

The PI3K-AKT-mTORC1 pathway is a well-known pro-survival axis constitutively activated in cancer with prominent roles in neoplastic transformation, growth, drug resistance and metastasis[137]. In melanoma, The PI3K-AKT-mTORC1-SREBP axis controls cell growth independently of BRAF mutation and dependent on sterol regulation instead[125,138]. It has been reported that the oncogenic PI3K or K-Ras signalling converging on the activation of mTORC1 in breast epithelial cells, is sufficient to induce SREBP-driven *de novo* lipogenesis[139]. The same authors further showed that oncogenic stimulation of mTORC1 is associated with increased SREBP promoting aberrant growth and proliferation of cancer cells in primary human breast cancer samples[139].

PI3K-AKT-SREBP pathway controls *de novo* lipid biosynthesis through glucose and glutamine [137]. Rapidly proliferating tumour cells depend more on glucose and glutamine for extensive *de novo* lipogenesis because of the action of oncogenic growth signalling molecules. Some cancer cells preferentially use glutamine as the main precursor to synthesize FA by reprogramming glutamine metabolism (glutaminolysis). SREBP can directly induce glutamine-derived carbon flux into lipid precursors at the sacrifice of the normal tricarboxylic acid (TCA) cycle. As a result, cancer cells do not simply shift to an anabolic phenotype, but they rather actively reprogram their metabolism via SREBP, thus rendering this TF a core hallmark of cancer[125].

1.2.3.b Unfolded protein response and ER stress-dependent regulation

Interactions between SREBP and ER stress are bidirectional and context dependent. Physiologically, mTOR-SREBP signalling is central for unfolded protein response (UPR) and ER homeostasis, whereas ER stress contributes to mTOR-SREBP-induced hepatosteatosis in the liver[140]. In the energy abundant state, SREBP1 activation associated with protein synthesis precipitates ER stress. The adaptive UPR initially compensate regaining homeostasis. Nevertheless, chronic ER stress with a prolonged UPR leads to hyperactivation of SREBP1c that, in the liver, precipitates steatosis and inflammation[141]. These events eventually increase the risk of fibrosis and cancer. Additionally, SREBP has been linked to ER stress-related lipotoxicity contributing to metabolic diseases such as obesity, diabetes, dyslipidaemia, hepatosteatosis and atherosclerosis[140,141].

The potential link between the UPR and lipogenesis is supported by evidence showing that: i) S1P and S2P are shared proteases between SREBP and the ER-stress sensor ATF6[143,144], ii) the main ER stress regulator suppresses ER stress-regulated activation of SREBP1c in the liver[144], iii) AMPK inhibits both ER stress and SREBP1c[127], iv) PUFAs protect pancreatic β cells from ER stress-related damage through inhibition of SREBP1c possibly preventing the development of diabetes[145], v) silencing SREBPs induce UPR and ER stress with depletion of MUFAs (Monounsaturated Fatty acids) and PUFAs and consequent apoptosis due to radical oxygen species (ROS) accumulation[141], vi) cellular stress induces proteolytic activation of SREBP through depletion of Insig1[142].
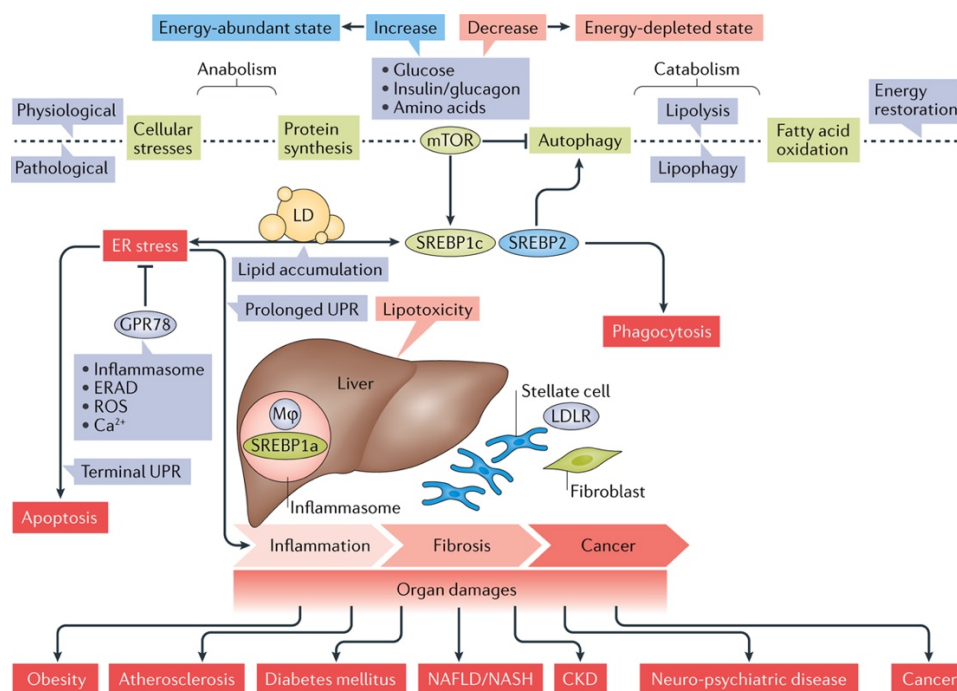
Figure 1.2.3 (Previous page)**:** Lipotoxicity mediated by SREBPs. In the energy abundant state, SREBP1 activation associated with protein synthesis precipitates ER stress. The adaptive UPR initially works to regain homeostasis, but chronic ER stress with a prolonged UPR causes further SREBP1c activation, which aggravates steatosis, cellular stress and inflammation. Finally, chronic inflammation increases the risk of fibrosis and cancer. Steatosis–inflammation–fibrosis mediated by lipotoxicity is a final common pathway to organ pathologies of immunometabolic disorders such as obesity, atherosclerosis, diabetes mellitus, nonalcoholic fatty liver disease (NAFLD), nonalcoholic steatohepatitis (NASH), chronic kidney disease (CKD) and neurological disorders. Severe cell stressors induce apoptosis through a terminal UPR. SREBP could be involved in regulating phagocytosis, autophagy and the innate immune responses of macrophages. (Adapted from Shimano, H. & Sato, R. 2017[125]).

1.2.3.c Transcriptional regulation by oncogenes and tumour suppressors: YAP/TAZ, MYC, p53, RB

In order to meet the high bioenergetic demands for cell growth, several oncogenic signalling molecules are involved in the activation of protein and lipid biosynthesis:

i) The YAP and TAZ proto-oncogenes are inhibited by the Hippo tumour-suppressor pathway. YAP/TAZ promote tissue proliferation, organ growth, cancer stem cell properties, metastatic potential and resistance to cancer therapy[146,147]. It has been shown that the SREBP/mevalonate pathway promotes YAP/TAZ nuclear localization and transcriptional activity[148]. Mechanistically, the geranylgeranyl pyrophosphate (GGPP), produced by the mevalonate cascade, activates YAP/TAZ by inhibiting their phosphorylation and promoting their nuclear accumulation. Thus, these findings indicate that mevalonate–YAP/TAZ axis is required for proliferation and self-renewal of breast cancer cells [148].

ii) c-Myc is an important proto-oncogene transcription factor regulating growth of both normal and cancer cells. In cancer c-Myc promotes tumour initiation, progression and survival. In prostate cancer, SREBP2 directly induces c-Myc activation to drive stemness and metastasis[149]. It has also been reported that SREBP1 promotes reprogramming by interacting with c-Myc in a translocation-dependent manner. SREBP1 interacts with c-Myc facilitating its binding to and promoting the expression of downstream pluripotent targets[150]. Gouw and colleagues recently demonstrated that MYC regulates lipogenesis to promote tumorigenesis through SREBP1[151]. Using human and mouse tumour-derived cell lines, tumour xenografts, and four conditional transgenic mouse models of MYC-induced tumours the authors showed that MYC induces SREBP1, activates FA synthesis and drives FA chain elongation from glucose and

glutamine[151]. Inhibition of FA synthesis blocked tumorigenesis and induced tumours regression in both xenograft and primary transgenic mouse models, revealing the vulnerability of MYC-induced tumours to the inhibition of lipogenesis. Upon MYC induction across different cancers, *in vivo* lipidomic changes were described as well[151]. These results further confirmed previous findings showing oncogenic levels of MYC to be linked to increased glutaminolysis resulting in glutamine addiction of MYC-transformed cells[152,153].

iii) More than 50% of human tumours are characterised by mutations of the TP53 gene. Previous studies have suggested that missense mutations confer tumour-promoting functions to p53. A possible mechanism has been previously proposed where the upregulation of the mevalonate pathway in breast tumours might be mediated by mutated p53 and SREBP and SCAP[154,155]. Nevertheless, a recent detailed analysis of p53 missense mutations in human leukemia by using genome editing, mouse models and clinical data showed that a dominant-negative effect drives selection of TP53 missense mutations in myeloid malignancies. The authors found no evidence that p53 missense mutations confer an oncogenic gain of function[156].

iv) Retinoblastoma protein (RB) is a tumour suppressor involved in the senescence pathway against malignant transformation. It has been shown that loss of RB activates SREBP eventually leading to DNA damage response and cellular senescence[125]. A recent study described subclonal RB1 loss in intermediate- to high-risk primary prostate cancer suggesting that this may be an early event in the development of mCRPC (metastatic androgen independent or castration-resistant prostate cancer) [213].

1.2.3.d Transcriptional regulation: dual role of AMPK as a central metabolic switch in cancer

Despite the well-known roles played by AMPK in physiological processes, its paradoxical context-dependent functions in regulating metabolic plasticity in cancer are still matter of discussion. Current opinions suggest that, whether early in tumorigenesis AMPK may first act as a tumour suppressor, in the advanced stages of the disease it may rather function as an oncogene contributing to therapy resistance and cancer recurrence[128]. Furthermore, decreased AMPK activation has been implicated in human metabolic disorders associated with increased cancer risk such as obesity and the metabolic syndrome[157].

The tumour suppressor role of AMPK has been reported to act through several mechanisms: i) inhibition of de novo FA synthesis inducing cell-cycle arrest (metabolic role), ii) induction of

mitotic spindle assembly/chromosome segregation abnormalities (non-metabolic role), iii) suppression of the oncogenic MEK–ERK signalling and consequent impairment of cell proliferation and cell-cycle progression via phosphorylation of the oncogene BRAF, iv) counteraction of the EMT, v) loss of AMPK activity contributing to tumorigenesis through hyperactivation of YAP, vi) inactivation of AMPK via ubiquitination and degradation leading to inhibition of autophagy and activation of mTORC1 signalling[126 158,159].

During tumour evolution stresses ranging from drug exposure to hypoxia, matrix detachment and starvation activate the AMPK pathway. At this advanced stage, AMPK is hypothesized to drive cancer progression by promoting metabolic plasticity, resistance to cellular stress and, therefore, cell survival. Several mechanisms by which the AMPK pathway supports this plasticity have been described including: i) induction of autophagy, ii) transcriptional changes induced by phosphorylation of the core histone H2B, iii) promotion of FA oxidation (FAO) to generate ATP and iv) increase of intracellular NADPH levels through the activation of FAO/inhibition of FA synthesis to neutralize cytotoxic ROS [126].

Recently, Khan and Frigo proposed a model of spatiotemporal regulation of AMPK complexes as one of the mechanisms responsible of this kinase's role in cancer[128]. The mechanistic explanation provided could assist in understanding how AMPK complexes regulate downstream metabolic processes to be either tumour suppressive or oncogenic [128], sheding light on the potential beneficial use of drugs targeting AMPK in combination with inhibitors of the lipogenic pathway in cancer.

1.2.3.e Transcriptional regulation by microRNAs

A further regulation of SREBPs is provided by microRNAs (miRNAs), small endogenous RNA molecules measuring 18-24 nucleotides in length and occurring in eukaryotes only. They do not code proteins but regulate post-transcriptional and translational gene expression. Among the lipid-related microRNAs (miRs), miR-33a has been described to be located in the intron of *SREBF2* and miR-33b in the human intron of *SREBF1*[160–162]. The miR-33 system regulates lipid homeostasis by modulating HDL biogenesis and cholesterol efflux. miR-33 also targets SREBP1c and has been shown to inhibit breast cancer metastasis[163,164]. Two miRs, miR-185 and 342, control lipogenesis and cholesterogenesis in prostate cancer cells by inhibiting SREBP1 and 2 expression and downregulating their target genes, including FASN and HMGCR. Both miRs inhibited tumorigenicity, cell growth, migration and invasion in PCa[162]. Their expression was found significantly decreased in PCa cells compared to non-cancerous

epithelial cells. It has further been shown that restoring miR-185 and 342 led to caspase-dependent apoptotic death in PCa cells[165]. In glioblastoma, EGFR/PI3K signalling upregulates SCAP/SREBP1 in turn activating miR-29 by directly binding to its promoter. A negative feedback loop has been shown for miR-29 that is able to suppress SCAP/SREBP1 and inhibit tumour growth[166].

Given their involvement in basic cellular functions and in the pathogenesis of many cancers, microRNAs are emerging as attractive candidates as prognostic biomarkers and therapeutic targets in cancer.

1.2.3.f SREBP as a novel therapeutic target in cancer

Given the so far discussed actions of SREBP1 in cancer, it is not surprising that there has been an increasing interest over the years in developing compounds able to disrupt its function at different levels. Some of the emergent therapeutic strategies are here discussed.

Fatostatin is an inhibitor of SREBPs that was originally developed to block insulin-induced adipogenesis[167]. This compound directly binds SCAP at a site distinct from the sterol-binding domain and hinders ER-to-Golgi transport of the complex SREBP-SCAP. Fatostatin: i) blocks hepatic lipid accumulation and body weight gain in obese mice; ii) inhibits cell growth by impeding intracellular shuttling in a SCAP-independent manner[168]; iii) has antiproliferative effects that are mediated via inhibition of mitotic microtubule spindle assembly[169]; iv) has shown to be a promising anticancer agent in breast[170] (reduce ERα and block cell invasion in AI resistant cells[53,170]), prostate (both AR-positive and metastatic AR-negative PCa) and pancreatic cancer[171–173] (Fig. 1.2.4b).

Betulin is a natural compound abundant in birch bark that inhibits the maturation of SREBPs by directly interacting with SCAP. Betulin improves hyperlipidaemia, insulin resistance and atherosclerotic plaques[174]. Betulin decreases hepatocarcinoma development and progression through reduction of SREBP-driven lipogenesis and attenuated inflammatory responses by down-regulation of tumour-promoting cytokines, including interleukin 6 (IL6), tumour necrosis factor alpha and IL1b[175].

Xanthohumol is a flavonoid found in hops and a novel SREBP inactivator that hampers COPII vesicle-mediated ER-to-Golgi transport[176]. This drug also affects the development of obesity, hepatic steatosis, and improves atherosclerotic plaque formation.

Nelfinavir and its analogues block S2P cleavage leading to suppression of proteolytic activity and accumulation of SREBP1 precursor and ATF6[177]. Nelfinavir is able to inhibit castration resistant PCa proliferation *in vitro*.

Sibilin is a natural compound isolated from the seeds of milk thistle plant (Silybum marianum) and widely consumed as a hepatoprotective agent. Through activation of AMPK, SREBP1 phosphorylation is increased in turn inhibiting SREBP1 nuclear translocation. In this way, Sibilin decreases nuclear protein levels of SREBP1 and their target genes in PCa cells leading to reduced lipid and cholesterol accumulation with consequent cell cycle arrest and inhibition of PCa cell proliferation. Sibilin also blocked androgen-induced lipid accumulation and prevented the development of androgen-independent LNCaP cell clones via targeting SREBP1[178].

Among the boron-containing small molecules, BF175 is a novel compound that can specifically block the binding of the Mediator complex to SREBP1a-TAD *in vitro*[179]. The recruitment of the Mediator complex to the SREBP transactivation domains (TADs) has been described to be required for the SREBP transcriptional activity. BF175 effect results in an inhibition of the SREBP transcriptional activity and a decrease of SREBP target gene expression in cultured hepatocytes. BF175 can improve lipid homeostasis in the mouse model of diet-induced obesity, decreasing hepatic and blood levels of lipids. These results suggest that blocking the interaction between SREBP-TADs and the Mediator complex by small molecules may represent a novel approach for treating diseases with aberrant lipid homeostasis.

Other agents targeting SREBP through cAMP-PKA or AMPK signalling pathways are: GLP1R, CB2R, GPR119 and FGF19[125]. Other AMPK activators [157]: i) indirect: Metformin, Thiazolidinediones, Resveratrol, ii) direct: AICAR (5-aminoimidazole-4-carboxamide ribo-nucleotide), PT-1, S396 (inhibits the transcriptional activity of SREBP[128]) and MT 63–78. MT 63–78 is a specific and potent direct AMPK activator able to inhibit PCa cell growth both in androgen sensitive and CRPC models, inducing mitotic arrest, and apoptosis[159].

There are several studies testing lipid lowering drugs anticancer potential but their protective role in the clinical setting is still greatly debated and highly controversial. *In vitro* studies have shown that statins, HMG-CoA reductase inhibitors, inhibit cancer cell growth[181] and antagonize breast cancer progression by interfering with ERα activation[53]. A nationwide population-based prospective cohort study[180], found an association between lipophilic statins, such as simvastatin, and a reduced risk of breast cancer recurrence. No association was found for hydrophilic statin users indicating that a possible beneficial effect of statins might be dependent on their composition. Prospective and population studies in PCa also showed an association between statins and reduced progression with a lower risk of developing metastatic or fatal prostate cancer[181–187]. Nevertheless, there are discordant data in the literature with several reports finding no such association[188–192]. Importantly, hypercholesterolemia and statin treatment influence serum cholesterol levels however, whether these changes affect intratumoral cholesterol is not clear[193,194]. As such, given the scientific plausibility that cholesterol promotes cancer progression[53,195–198], albeit in the face of equivocal epidemiologic data, it is important to understand the molecular mechanisms underlying metabolic reprogramming in the development of drug resistance.

1.2.4 The "lipogenic phenotype": a lesson from Prostate cancer

A century ago, Otto Warburg first described aerobic glycolysis, a metabolic reprogramming that tumour cells undergo by increasing the uptake of glucose and converting it to lactate, even in normal oxygen conditions[199]. Since then, a growing body of literature has recognised this phenomenon, termed the Warburg effect, as a hallmark of cancer with cell metabolism shifting from a catabolic to an anabolic state[60]. Metabolic adaptation in tumours extends beyond the Warburg effect, indeed more mechanisms of metabolic reprogramming have been described in different cancer types. Prostate Cancer (PCa), for example, does not always show the classic glycolytic switch presenting an aberrant increase of *de novo* lipogenesis from glucose and glutamine instead[200]. Since the early stages of tumoral transformation, *de novo* lipid biosynthesis correlates with tumour progression, poorer prognosis and shorter patient survival [200–202]. PCa cells exhibit higher cholesterol levels than juxtaposed normal cells[201]. It has also been suggested that the increased cellular cholesterol within mitochondrial membranes renders cells resistant to many chemotherapeutics[203]. This altered lipid metabolism, also called "lipogenic phenotype", is linked with PCa pathogenesis. Rewiring the lipogenic program consists in the dysregulation of biosynthetic, remodelling and catabolic processes.

1.2.4.a Crosstalk between lipogenesis and oncogenic signals promotes tumour growth and progression

Physiologically, normal tissues maintain low levels of lipogenic enzymes preferring the use of dietary lipids more than endogenous biosynthesis as energy supply. The high metabolic demand that cancer cells develop during uncontrolled proliferation made them adapt to the use of alternative metabolic pathways with increased synthesis of *de novo* fatty acids and cholesterol, independently from the circulating lipid levels[204-206].

Many key enzymes of the lipid biosynthesis, such as SCD1 (Stearoyl-CoA-Desaturase), ACLY (ATP Citrate Lyase), ACC (Acetyl-CoA Carboxylase) and FASN, are reported to be overexpressed in androgen sensitive and resistant PCa and showed to be required for PCa growth and survival[200]. Constitutively active SCD1 promotes cell survival, proliferation and transformation, increases tumorigenicity and invasiveness, and correlates with a higher Gleason grade[200]. SCD1 plays a role in tumour progression through the stimulation of ACC, inactivation of AMPK phosphorylation, activation of Akt pathway and increased ratio of MUFAs to SFAs (Saturated fatty acids) [201].

Since diet supplies most FAs, endogenous synthesis is minimal, thus FASN expression is either low or undetectable in most normal human tissue. However, with development of PCa, these transformed cells overexpress FASN, and this marked contrast in expression has led FASN to be proposed as a candidate oncogene [204-207, 216]. FASN overexpression occurs in the early stages of PCa, and is associated with tumour progression, bone metastasis, poor prognosis and reduced disease-free survival [204,206,216] (Fig. 1.2.4a). FASN is involved in the control of cell cycle progression, energy homeostasis (improved redox balance), resistance to oxidative stress and cell polarity. Interference with FASN activity causes arrest in G1 phase of the cell cycle or apoptosis[201,204]. FASN overexpression is induced by Phosphatase and Tensin homolog (*PTEN*) knockdown in PCa cells and correlates with the activation of Akt, suggesting a coordinated feedback between lipogenesis and oncogenic signals to promote tumour growth and progression[205]. About 70% of advanced and 42% of primary PCa have *PTEN* loss, which in turn results in the activation of PI3K pathway[200]. This pathway is responsible for the increase in cell survival, metastasis and castration-resistant growth in PCa. It is also responsible for the activation of SREBP and LDLR and is involved in, cholesteryl ester (CE). Studies on PCa bone metastasis revealed elevated levels of LDLR that is responsible for LDL uptake and for maintenance of intracellular cholesterol homeostasis [204]. PCa cells esterify cholesterol in lipid droplets to avoid cellular toxicity due to high intracellular cholesterol levels and maintain cholesterol levels independently of the free cholesterol concentration. In this way, cancer cells

can keep SREBP constantly active [204]. It has been recently suggested that a concomitant loss of Promyelocytic Leukemia (*PML*) in *PTEN*-null cap, found in 20% of mCRPC, promotes metastatic progression through reactivation of MAPK (Mitogen-Activated Protein Kinase) signalling and subsequent hyperactivation of an aberrant SREBP pro-metastatic lipogenic program[208]. Treatment with Fatostatin can block the lipidomic profile and the metastatic effects of these changes in the PML and PTEN double-null PCa [208].

As a further consequence of *PTEN* loss, chromosome 8q is amplified, including the *MYC* gene, in about 30% of prostate tumours [200]. *AKT* and *MYC* are the most prevalent driving oncogenes in all cancers and they both induce the expression of FASN in PCa[209]. Whereas AKT1 has been associated with accumulation of aerobic glycolysis metabolites, overexpression of MYC is linked with dysregulated lipid metabolism[303]. MYC is frequently amplified in the late stages of prostate cancer and it has been suggested that MYC-driven tumours may rely more heavily on lipid metabolism[209]. Although *de novo* synthesised fatty acids are typically derived from glucose, some cancer cells rely more on glutamine metabolism[200]. MYC function is essential for glutamine-dependent metabolism to maintain macromolecule synthesis and its overexpression induces glutamine conversion for final NADPH production.

The androgen receptor (AR), as key driver of PCa, also regulates MYC expression, glutamine transporters and glutamine uptake[201]. PCa cells convert glutamine to alpha-ketoglutarate and shuttle it into the mitochondria to replenish the TCA cycle intermediates for anaplerotic reactions; under hypoxic conditions glutamine is used, rather than glucose, as major lipid precursor[205]. Hypoxia can lead to oxidative stress that has been linked to PCa development and progression in several studies[205]. A hypoxic environment induces AKT/HIF1α-mediated activation of SREBP1, which increases transcription of FASN. SREBP1 can also induce transcription of NOX5 (NADPH (Nicotinamide Adenine Dinucleotide Phosphate) oxidase 5), a prominent producer of ROS and regulator of PCa cell growth, suggesting another feed forward mechanism including lipogenesis and prostate cancer growth[205].

FASN expression can be induced downstream of AKT1 via mTORC1-mediated SREBP1 activation[200]. The activity of Akt and mTORC1 is required for the nuclear accumulation of mature SREBP1, directly regulating its expression. SREBP1 function is also essential for Akt-dependent regulation of cell size, suggesting that Akt/mTORC1 signalling axis regulates protein and lipid synthesis to meet cancer cell heavy metabolic demands[200] (Fig. 1.2.4a). Furthermore, FASN activity is associated with palmitoylation of known oncogenes including *k-RAS* and *WNT1*, regulation of ER function to sustain membrane biogenesis and resistance to genotoxic insults [210,217] (Fig. 1.2.4a).

Further crosstalk between genomic changes and lipid metabolism have been reported in prostate cancer cells. In 5% of primary and 37% of advanced tumours, the tumour suppressor Rb is inactivated, enhancing N-Ras through induction of SREBP1 and 2[200]. The mevalonate pathway is significantly upregulated in p53 mutant cells, reported in about 3-20% of PCa cases[200]. Moreover, a recent study found an amplification and overexpression of the pyruvate dehydrogenase complex (PDC), responsible of converting pyruvate into acetyl-coA for entry into the TCA in mitochondria [211]. The important role played by acetyl-coA does not consist only in sustaining metabolic activity of mitochondria, but it is additionally supporting histone acetylation and enhancer activity in the nucleus. The principal effect of targeting the PDC complex is tumour suppression by abrogating lipid biosynthesis[211].
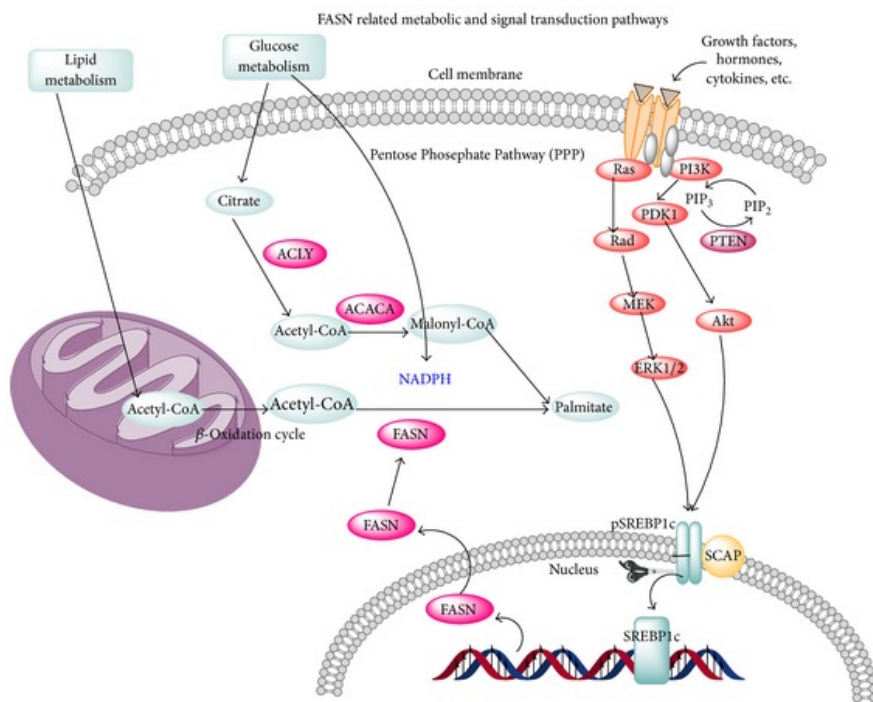
Figure 1.2.4a: Lipid pathways in cancer. De novo fatty acid synthesis is an important hallmark of cancer cells, differentiating it from normal cells. The overexpression of FASN allows for the *de novo* synthesis of essential lipids for the formation of cell membrane and for the production of extra energy via beta-oxidation and lipid modification of proteins. The binding of the growth factor and the growth factor receptor results in the activation of their downstream PI3k-Akt and Ras signal transduction pathway. The FASN expression is regulated by several growth factors, including steroid hormone, and steroid hormone receptors (such as ER, PR, and AR). (Adapted from Cheng 2014[212]).

1.2.4.b Androgens and Androgen Receptor crosstalk with SREBP1 in PCa progression

With the use of genome-wide expression profiling, evidence is mounting that most androgen gene networks are reactivated in CRPC progression[198,214]. Two contending hypotheses stand to explain these observations[214]: i) AR is aberrantly activated by signalling pathways or by coactivators in the absence of androgens, ii) androgen-regulated pathways within prostate cancer cells are activated by alternative sources of androgenic steroids.

1.2.4.b.i AR aberrant activation by SREBP1 signalling pathway

In prostate cancer, SREBP1 plays a crucial role in the activation of the lipogenic phenotype through an established crosstalk with androgens and androgen receptor[215]. Staining of human prostate tumours showed elevated levels of SREBP1 protein compared with normal prostate tissue[215]. SREBP1 induces PCa cell proliferation, migration and invasion *in vitro* and promotes PCa tumour growth and castration-resistant progression *in vivo*[198,215]. It has been further demonstrated that many of the cholesterol synthesis enzymes downstream of SREBPs (HMGCS, SQLE, squalene monooxygenase, lanosterol synthase, farnesyl diphosphate synthase) are induced during progression of the disease[198]. The clinical and animal data collectively indicate that SREBP1 expression and nuclear translocation play a critical role in the regulation of PCa development and progression to castration-resistance[215]. Blocking SREBP translocations with Fatostatin in PCa: i) suppressed cell proliferation and anchorage-independent colony formation in both androgen-responsive LNCaP and androgen-insensitive C4-2B PCa cells, ii) reduced *in vitro* invasion and migration in both cell lines causing G2/M cell cycle arrest, iii) induced apoptosis by increasing caspase-3/7 activity and cleavage of caspase-3 and PARP (Poly (ADP-Ribose) Polymerase), iv) significantly inhibited subcutaneous C4-2B tumour growth and markedly decreased serum PSA level compared to the control group *in vivo* animal results, v) decreased the expression of AR and its target gene PSA *in vitro* and *in vivo*[171,172,208] (Fig. 1.2.4b).

AR protein levels can be reduced by treatment with fatostatin and simvastatin, similarly AR expression and activity are reduced through inhibition of ACACA (Acetyl-coA Carboxylase Alpha), FASN or SCD-1[171,200]. Furthermore, concomitant overexpression of AR and FASN in prostate is sufficient to induce adenocarcinoma in mice[216].
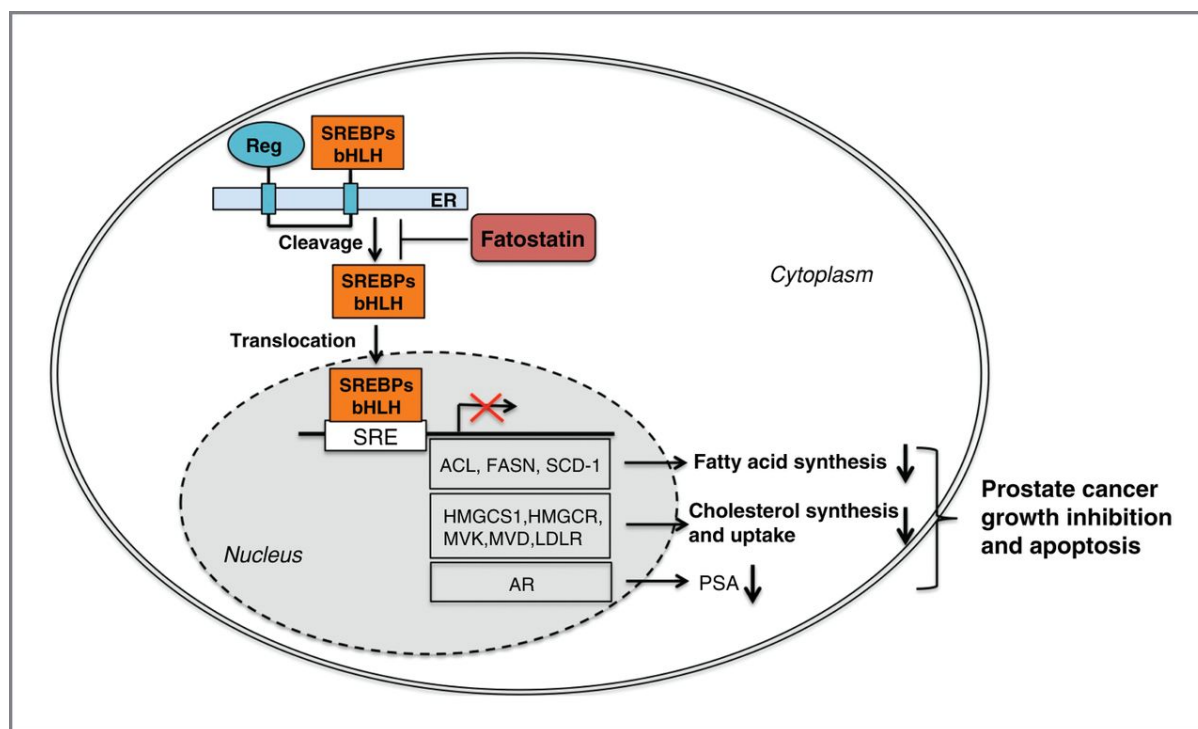
Figure 1.2.4b: Fatostatin action on SREBPs in cancer. By inhibiting the nuclear translocation and transcriptional activity of SREBPs, fatostatin decreases the expression of SREBP downstream target genes, suppresses expression of AR and its target gene PSA (Adapted from Xiangyan Li 2014[171]).

Activation of the androgen receptor by androgens increases expression of lipogenic enzymes in a SREBP1c-dependent manner[218]. A positive feedback loop promotes this signalling pathway since binding sites for SREBP1 are also found in the AR gene of prostate cancer cells[219]. AR is regulated by B2-microglobulin (B2-M) in a MAPK/SREBP1-dependent manner[220]. B2-M is a component of the housekeeping major histocompatibility complex class I molecule found on PCa cells. Inhibition of B2-M decreases the interaction between SREBP1 and its binding site in the AR promoter region, resulting in decreased AR expression and lipogenesis [220]. It has been shown that SREBP1 inhibition can down-regulate AR levels [171] and FASN inhibition causes ER stress response associated with AR pathway deregulation[221], further confirming the mutual regulation between AR signalling and FA metabolism. The emergence of AR splice variants (AR-Vs) such as AR-V7 has been described as one of the different mechanisms of resistance to AR-directed therapies[222]. AR-V7 lacks the C-terminal ligand-binding domain of full-length AR and functions as a constitutively active, ligand-independent transcription factor driving growth of mCRPC cells *in vitro* and *in vivo*[223]. Increased lipid biosynthesis (potentially reactivated by AR-V7) is associated with poor

outcome in multiple clinical cohorts, indicating the AR-V-mediated reactivation of lipid biosynthesis may drive the disease recurrence[224].

Although these studies indicate that modulation of lipid metabolism may impact AR signalling, it has been observed that overall cholesterol homeostasis is unaffected by changing androgen receptor activity in PCa cells[225]. This does not negate the relationship between androgens and cholesterol homeostasis but may rather suggest that other factors compensate for altered androgen receptor activity.

It has been suggested that androgens may activate the SREBP pathway by a mechanism that differs from the one observed under sterol depletion[218]. Indeed, androgens activate the SREBP pathway with minor effects on SREBP precursor levels and a major increase in the expression of SCAP[218,219,226]. SCAP plays a pivotal role in the lipogenic effects of androgens in tumour cells[227]. In this positive feedback loop, androgens stimulate the expression of SREBP1 through SCAP[218]. In turn, SREBP1 regulates the expression of the androgen receptor[219,227].

1.2.4.b.ii Activation of androgen-regulated pathways via de novo cholesterol biosynthesis

An additional hypothesis is that intra-tumoral androgens may derive from *de novo* synthesised cholesterol of the mevalonic acid pathway, as well as circulating cholesterol from the diet[200,205]. Indeed, increased cholesterol levels along with elevated expression of enzymes involved in steroidogenesis have been described in PCa bone metastasis[205,228]. These data suggest that CRPC progression may not be entirely independent of androgen-driven activity, but non-testicular sources of androgens could be capitalised on for AR activation[200,205]. Locke and colleagues showed that androgen levels within CRPC tumours are sufficient for AR activation, whereas corresponding serum androgens remain low in mice after castration[229]. This study suggested that androgens driving CRPC progression are synthesised *de novo* within the prostate cancer and increased by using a feed forward biosynthesis pathway as many of the enzymes are increased by androgens[229].

Androgen synthesis is often described as a classic steroidogenic pathway through dehydroepiandrosterone (DHEA) and testosterone. However, a "backdoor pathway" has been described as an alternative synthesis pathway that uses progesterone as the primary steroidal precursor of dihydrosterone, bypassing testosterone as an intermediate[214] (Fig. 1.2.4c). It has been reported that CRPC tumours producing relatively high concentrations of progesterone are capable of *de novo* synthesis of androgenic steroids[229]. These findings suggest that

progesterone may be involved in an adaptation mechanism whereby the cancer cell initially adjust to androgen deprivation by producing more progesterone to act as an androgen precursor[214,229]. Progesterone, like dihydrotestosterone, is known to induce cholesterol synthesis in prostate cancer cells[230]. Tumour progesterone levels are high after castration and enzymes necessary for progesterone synthesis from cholesterol (CYP11A1 (Cytochrome P450 family) and StAR (Steroidogenic Acute Regulatory Protein)) and metabolism (CYP17A1 and SRD5A1 (Steroid 5 Alpha-Reductase 1)) are increased in CRPC tumours[231]. Compared with untreated primary prostate tumours, castration-resistant metastasis showed significant increase in the expression of FASN (~10-fold change), HSD3B1, HSD3B2 (Hydroxy-Delta-5 Steroid Dehydrogenase, 3 Beta- and Steroid Delta-Isomerase 1 and 2), CYP17A1 (~17-fold change), AKR1C3 (Aldo-Keto reductase Family) and HSD17B3, key enzymes required for metabolism of progestins to adrenal androgens and their subsequent conversion to testosterone[231](Fig. 1.2.4c). CYP17A1 has also been demonstrated to have squalene epoxidase activity suggesting it may have a dual role in CRPC steroid metabolism[214].

Marked up-regulation of CYP19A1 (30-fold change), which mediates the aromatisation of testosterone to oestradiol, was also observed in metastases and is consistent with prior reports demonstrating up-regulated expression of aromatase in malignant versus benign prostate epitelium[232]. Furthermore, amplification of the CYP19A1 gene (CYP19A1[amp]) has been demonstrated to occur in 21.5% of AI-treated BCa patients[54] causing increased aromatase activity, oestrogen independent ERα binding to target genes and decreased sensitivity to AI treatment. These data indicate that AI treatment selects for acquired CYP19A1[amp] and promotes local autocrine oestrogen signalling in AI-resistant metastatic patients[54].

Transcripts encoding the full complement of enzymes comprising the steroidogenic pathway were detectable in the majority of primary and metastatic prostate tumours examined [231] and in six prostate cell lines and mice models [214]. This observation suggests that adaptive modulation of steroidogenic pathways to castrate environment may occur within the tumour before initiation of the metastatic cascade with therapy promoting adaptation and evolution of a more aggressive phenotype. While a role for *de novo* steroidogenesis *per se* in primary prostate tumours is less likely, these observations suggest that the selective pressure of androgen-deprivation therapy (ADT) may lead to upregulated expression of these enzymes and reconstitution of tumour androgen levels in CRPC. In summary, these studies propose that metastatic prostate cancers may adapt to low systemic testosterone levels by maintaining intratumoural androgens through the modulation of enzymes involved in intracrine steroidogenesis from *de novo* cholesterol biosynthesis.

Considering what is known about the pathobiology of cholesterol in prostate cancer, it is clear that these cancer cells have evolved mechanisms to bypass the tight homeostatic regulation of intracellular cholesterol representing a potential vulnerability for intervention. The lipogenic phenotype has been further documented in various cancer types other than PCa including breast, ovarian[361] and colorectal cancer[362,363]. These findings taken together with other published studies[53,54,196–199,233] paves the way for a deeper investigation into the role played by SREBP1 in development of BCa resistance[232].
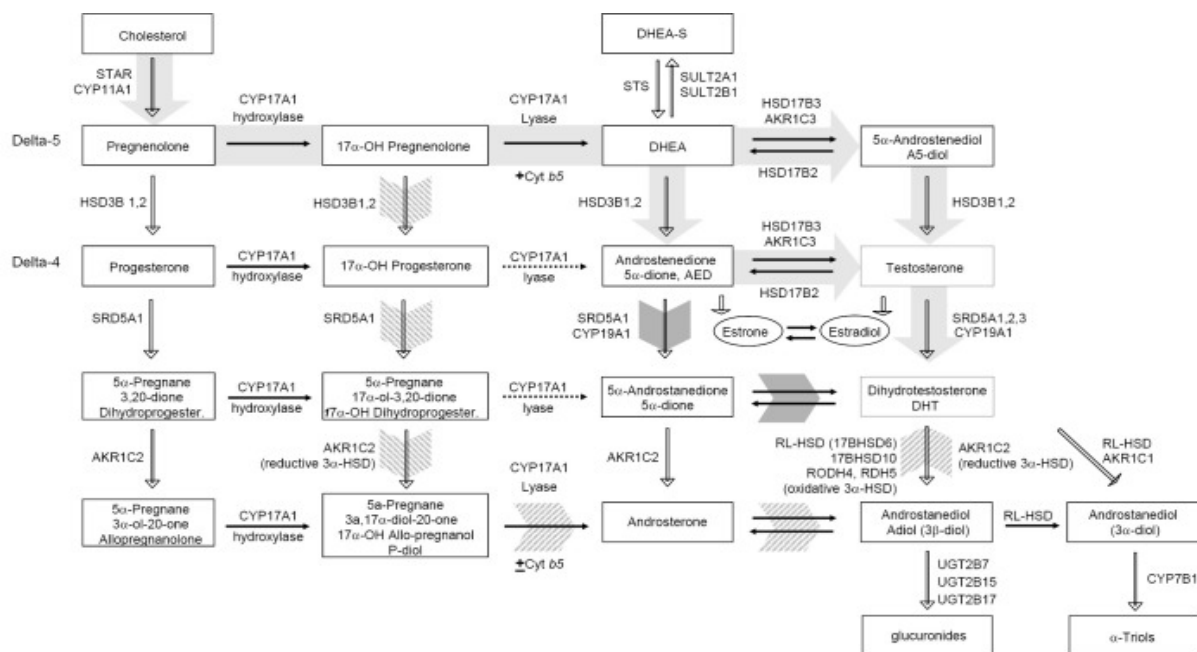
Figure 1.2.4c: Classical and non-classical pathways of androgen biosynthesis. Cholesterol is first converted to precursors pregnenolone and progesterone that are converted to the adrenal androgens DHEA and androstenedione (AED). DHEA (from intrinsic or circulating sources depending on the tissue) subsequently form testosterone (T) which is converted to DHT. In the backdoor pathway (hatched arrows) the progestin intermediates generate DHT. Alternatively, (dark grey arrows) AED can also be converted first to 5α-Androstanedione and then to DHT (Adapted from Mostaghel 2013[214]).

1.3 Cytoskeleton: intermediate filaments (IFs)

The cytoskeleton is the main structural framework within eukaryotic cells cytoplasm determining cell shape and facilitating a variety of cell functions. It consists of three-dimensional (3D) networks composed of three kind of cytoskeletal filaments ranging from microfilaments (~6-8 nm), intermediate filaments (~10 nm), to the largest microtubules (~25 nm)[233]. Intermediate filaments (IFs) are assembled from a diverse group of evolutionary conserved proteins and are specified in a tissue-, cell type-, and context-dependent fashion in the body[234]. There are ~70 genes encoding for IFs proteins in the human genome, of which 54 code for keratins[235]. IFs are classified into 6 major subtypes (Table 1.3.1) all sharing the property of self-assembly into filaments as hetero- or homopolymers.

Subgrouping	Proteins	Cell type specificity
Type I	Keratins	Soft complex epithelia (skin, oral mucosa, etc.)
Type II		Soft simple epithelia (liver, gut, kidney, etc.)
		Hard epithelia (hair, nail, oral papillae)
Type III	Vimentin, Desmin GFAP, Peripherin syncoilin	Various (fibroblasts, leukocytes, endothelium muscle, astrocytes, glia, peripheral nerves)
Type IV	NF-L, NF-M, NF-H a-internexin synemin, nestin	CNS & neurons CNS & neurons Muscle, neural stem cells
Type V	Lamins A, B & C	Nucleus
Orphan	Filensin, Phakinin	Lens

Table 1.3.1: General classification of intermediate filament proteins (Adapted from Chung 2013[234]).

1.3.1 Keratins in health:

1.3.1a Keratin filaments in epithelial cells: assembly, properties and functions

Keratin genes are subdivided into two types, 28 type I and 26 type II, clustering respectively on chromosome 17q21.1 and 12q13.13 (except type I keratin 18 located in the type II keratin gene domain)[237,238] (Fig. 1.3.1). This clustered arrangement of genes might suggest the existence of a hierarchical regulation during embryogenesis and adult-tissue homeostasis[239] (Fig. 1.3.1). Whereas type I keratins are generally smaller and acidic (40-56.5 kDa, pI 4.5-6.0), type II are larger and basic-neutral in charge[239] (50-70 kDa, pI 6.5-8.5). From a biophysical and biomechanical prospective, keratin filaments behave like relatively weak gels when dispersed in solution[240]. When cross-linked into networks, however, keratin filaments become

more similar to solid material and are able to maintain their viscoelastic properties while at the same time withstand broad regimens of deformation[240]. Keratin filaments can stretch three times their initial length before breaking[241,242]. Although not as dynamic as F-actin, keratins show a significantly greater ability than F-actin and microtubules to resist mechanical stress *in vitro*[243]. These properties have direct implications for their functional importance *in vivo* and make IFs unique among cytoskeletal proteins.

Keratins are generally considered very static proteins conferring mechanical strength to epithelial cells. They play an important role in cytoprotection from mechanical and non-mechanical stressors[245] forming cytoskeletal networks, greatly contributing at the cytoarchitecture and adhesion, cell size and growth, and protection from apoptosis[244–246]. Naturally occurring mutations disrupting filaments structure and networking result in cellular fragility, impaired responses to stress and several diseases affecting skin (such as epidermolysis bullosa simplex caused by mutations in K5 or K14), hair, cornea (corneal dystrophy associated with K3/K12) and liver[247,248]. A compendium of keratin mutations and consequent diseases has been compiled and is publicly available as The Intermediate Filament database[247] (www.interfil.org).

Keratins participate in the regulation of epithelial tissue growth in three interconnected ways: i) regulation of cell cycle progression, ii) regulation of protein synthesis and iii) activation and expression of immune and inflammatory mediators[239,249,250]. Fine-tuned epithelial functions such as translation, signalling, vesicle trafficking[246] and wound healing[250,251] have been attributed to keratins. Epithelia stress response[252], cell survival, migration and metabolism[253] are dynamic processes regulated by keratins.

Figure 1.3.1 (following page)**:** Phylogenic tree of human keratins. (A) Comparison of the primary structure of human keratins. Two major branches corresponding to the types I and II keratins that are further segregated into major subgroupings (colour coded). (B) Location and organization of genes encoding types I and II keratins in the human genome (Adapted from Coulombe 2013, Jacob 2018[234,238]).

A

K39 K34 K35 K37 K38 K23
K32 K33b K36 K20
K31
K33a
K40
K74
K72 K73 K71
K3
K77 K76 K75 K5
K4 K6a,b,c

Type II
keratins

Type I
keratins

K1
K2
K79
K78
K8
K7

K15
K13
K19
K17
K16
K14
K12
K24

K83
K81 K86
K85
K84 K82 K80

K10
K26
K18 K25 K28
K9 K27

0.1
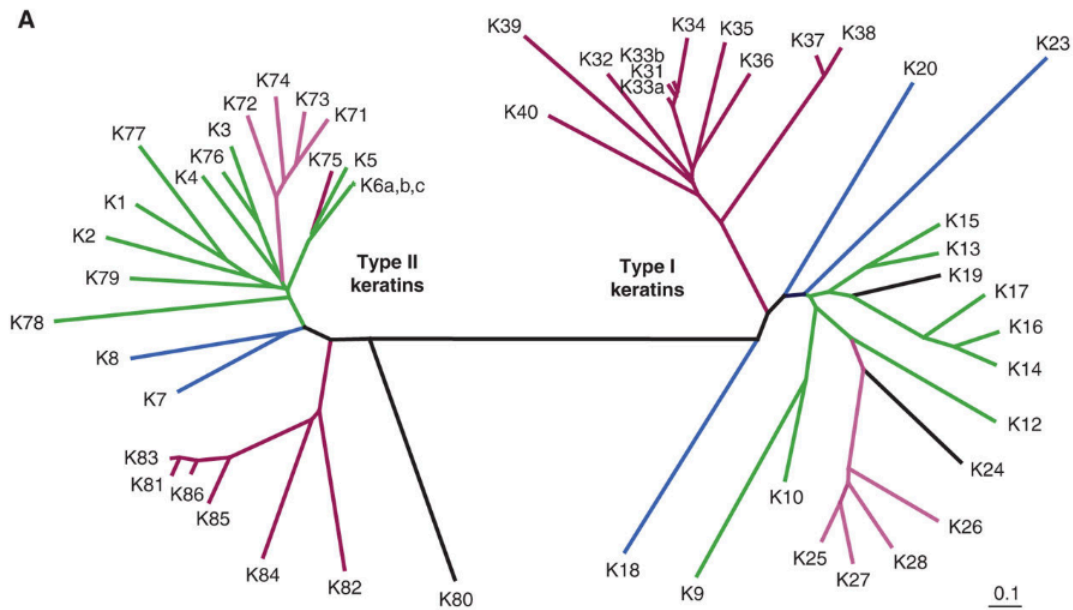
B

Type I keratin gene cluster - Human chromosome 17q21.2

centromere ... KRT222P KRT24 KRT223P KRT25 KRT26 KRT27 KRT28 KRT10 KRT12 KRT20 KRT23
KRT39 KRT40 KRT33A KRT33B KRT34 KRT31 KRT41P KRT37 KRT38 KRT221P KRT32 KRT35 KRT36
KRT13 KRT15 KRT19 KRT9 KRT14 KRT16 KRT17 KRT42P ... telomere

Type II keratin gene cluster - Human chromosome 12q13.13

centromere ... KRT80 KRT7 KRT121P KRT122P KRT81 KRT86 KRT23 KRT123P KRT85 KRT84 KRT82
KRT124P KRT75 KRT6B KRT6C KRT6A KRT5 KRT71 KRT74 KRT72 KRT73 KRT125P KRT2 KRT1
KRT77 KRT126P KRT127P KRT128P KRT76 KRT3 KRT4 KRT79 KRT78 KRT8 KRT18 ... telomere

1.3.1.b Keratin static and dynamic organisation of cell structure

Keratins are pivotal to cell structure providing strength and mechanical resilience and yet, they remain dynamic and flexible allowing for a rapid networking remodelling without a network disruption. Epithelial cytoskeleton continuously adjusts to dynamic cellular processes through a perpetual keratin turnover. Keratin filament system is not homogenous, but it is organised into temporally and spatially distinct subdomains. The peripheral region consists of young filaments that are continuously replenished by integration of precursors recruited form the soluble pool. The region closer to the nucleus contains older filaments that disassemble to support peripheral network renewal[254]. The network formed in this way is functionally and structurally organised in a highly dynamic manner. A continuous spatiotemporal cycle of filaments assembly and disassembly starts with small-sized precursors at the cell periphery, often in close proximity to lamellipodia focal adhesions. Then, precursors elongate and integrate into the network determining an inward movement toward the nucleus in a centripetal

direction[254] (epithelial, non-epithelial cells and cancer cells like MCF7). Keratin filaments bundle mature in a stable network surrounding the nucleus and anchoring to desmosomes and hemi-desmosomes or, alternatively, they disassemble and turn over rapidly diffusing through the cytoplasm and starting another cycle form the cell periphery[255]. As a crucial consequence of keratins cycling, Windoffer and colleagues suggested that the intrinsically nonpolar keratin network acquires spatial orientation with a centripetal organisation. Whereas the younger and more dynamic peripheral filaments might be able to react to structural and functional requirements imposed on the cytoskeleton, the older and more stable central filaments anchoring desmosomes and hemi-desmosomes, might be rather responsible for the cellular mechanical stability[254]. The same authors also proposed this as a mechanism of continuous probing of the immediate extracellular surroundings until new physical contact with either other cells or extracellular matrix can be established and stabilised through desmosomes and hemi-desmosomes[254]. The dynamics of the keratin network enhance the adaptability and the functions of moving cells especially because it doesn't require protein biosynthesis thus providing the cell with a variety of options to respond to environmental challenges within a small-time frame.

A perpetual filament turnover cycle supports the plethora of keratin functions through a multistep process that keeps the cytoskeleton in motion while maintaining an intact network[255]. Regulation of cycling is linked to post-translational modifications, particularly phosphorylation, and it is therefore targeted by signalling pathways.

1.3.1.c Keratin regulation: post-translational modifications and protein interactions

In response to stress, keratin expression is commonly altered to allow for structural reorganisation[252,256]. Keratin cytoskeleton disintegrates in mammary epithelial cells under metabolic stress such as combined glucose and oxygen deprivation mimicking the tumour microenvironment[256]. Keratin reorganisation depends on stress-duration and is a severity-dependent response to mechanical forces regulated by post-translational modifications. Among the several type of post-translational modifications (phosphorylation, O-linked glycosylation, ubiquitination, acetylation, SUMOylation and transamination[257]), phosphorylation is commonly considered the major regulator of keratin properties such as solubility, conformation and filament structure[258]. In general, phosphorylation levels are low in basal conditions and increase under a variety of cellular stresses including drug-induced apoptosis, heat stress, shear stress and metabolic stress[257,258]. Coupled with the keratin filament reorganisation, this

dynamic phosphorylation/dephosphorylation cycle provides a reversible shock-absorber-type mechanism for cells to cope with intracellular and extracellular stresses[258].

1.3.1.d Keratin regulation: endocrine control of keratins

Epithelial and hair keratins are regulated by several endocrine signals such as glucocorticoids, vitamins and hormones with the physiological process of keratinisation consisting of terminal differentiation of keratinocytes being deeply influenced by hormones[259].

Glucocorticoids (GCs) inhibit keratins expression, which are generally upregulated during inflammatory skin response and wound healing. GCs influence keratin expression through two main independent mechanisms: i) direct binding of the GC receptor (GR) leading to suppression of transcription and ii) indirect control of keratin genes expression by blocking the induction of AP-1. AP-1 is a transcription factor upregulating keratin genes transcription by binding to their regulatory regions. GR stimulation up-regulates several hair keratins and keratin-associated protein genes that are important for the disulphide bonds formation and consequent strong hair keratin structure[259].

Excess of Vitamin A, a precursor of Retinoic Acid (RA), inhibits keratinisation whereas hypovitaminosis A causes epidermal hyperkeratosis and keratinisation of epithelia like conjunctiva and cornea that are physiologically otherwise non-keratinised. RA regulates keratin synthesis by: i) controlling the expression of secondary regulators of epidermal keratinocyte differentiation and ii) by nuclear receptors direct action on gene regulatory sites. Through the latter mechanism, RA suppresses the expression of specific, disease-associated keratin genes. It has been shown that, in T47D breast cancer cells, inhibiting RA decreases cell growth and increases the expression of K8, K18, K19, which are markers of luminal differentiation[259].

Thyroid hormones (TH), in particular the active form triiodothyronine (T3), affect keratin expression inducing keratinocytes proliferation.

The vitamin D active metabolite 1,25(OH)2D (D3) can also regulate keratinocyte differentiation. Increased levels of calcium, regulated by vitamin D, inhibit keratinocyte proliferation and induce terminal differentiation. D3 was found to up-regulate K13 in a model for human colon cancer and the absence of its receptor led to the induction of undifferentiated basal cell carcinomas[259].

Major effects of androgens on hair growth are well-documented. It has also been speculated that androgens might regulate keratin expression in prostate epithelium, probably through

transcriptional regulation of HOXC13 (Homeobox C13). In humans, HOXC13 is known to be involved in controlling the expression of several hair keratin genes. Oestrogens, likewise, might contribute to the regulation of some keratins. It has been reported that the expression of some hair keratins such as K2, K14, K15, K17, K19, K37 and K75 is stimulated by 17β oestradiol. In MCF7, the up regulation of K19 contributes to the cytoskeletal and nuclear network reorganisation, potentially increasing the metastatic potential of BCa cells upon exposure to oestrogens[259].

1.3.2 Keratins in cancer: canonical and non-canonical functions for IFs

1.3.2.a Diagnostic and prognostic markers in epithelial tumours

Given their characteristic cell-type, tissue-type expression patterns in epithelial cells, and the availability of specific antibodies, keratins have been so far extensively used as immunohistochemical diagnostic and prognostic tumour markers (Tables 1.3.2, 1.3.3).

Cancer site and subtype	*Keratin expression*	
Biliary duct	K7, K8, K18–20	
Bladder, transitional cell	K5[a], K7, K8, K18, K19, K20[a]	
Breast	K5[a,b], K6[a,b], K7, K8, K14[a,b], K17[a,b], K18, K19	
Cervix	K4–8, K10, K13–19	
Colon	K7[a], K8, K18–20	
Kidney, clear cell	K8, K18, K19[a]	
Papillary	K7, K8, K18, K19	
Chromophobe	K7, K8, K18, K19[a]	
Liver	K7[a], K8, K18, K19[a], K20[a]	
Lung, adenocarcinoma	K7, K8, K18, K19	
Small cell	K8, K18, K19[a]	
Ovary, adenocarcinoma[c]	K7, K8, K18, K19	
Pancreas	K5[a], K7, K8, K18, K19, K20[a]	
Pleura (mesothelioma)	K5, K7[a], K8, K18, K19	
Prostate	K8, K18, K19	
Skin, squamous	K1, K4–6, K8[d], K10, K13–17, K18[d], K19[d]	
Merkel cell	K8, K18, K20	
Stomach	K7[a], K8, K18, K19, K20[a]	
Uterus	K5[a], K7, K8, K18, K19	

[a]Focal/heterogeneous staining in some, but not all, cases.
[b]Focal or extended staining in basal-like tumors.
[c]Non-mucinous.
[d]In poorly differentiated cases.

Table 1.3.2: Keratins as diagnostic markers in tumour pathology (Adapted from, Karantza 2011[256]).

Cancer site	Keratin expression pattern	Detection site	Prognosis
Biliary duct	High K19 fragment (CYFRA21-1)	Serum	Worse
Breast	K5/6, K17	Tumor	Worse
	K19 mRNA	CTCs[a]	Worse
	Reduced K18 mRNA	Tumor	Worse
	Ubiquitinated K8 and K18 fragments	Tumor	Worse
Colon	Reduced K8, K20	Tumor	Worse
	Persistent or higher K18 fragment (M30) after primary tumor resection	Serum	Worse
Kidney	K8, K18	CTCs[a]	Worse
	K7, K19	Tumor	Better
Liver	K10, K19	Tumor	Worse
Lung	High K18 fragment (M30)	Serum	Worse
Pancreas	K20	Tumor	Worse
	K20	Serum	Worse
Prostate	K8, K18, K19 before surgery	Bone marrow	Worse
Skin (melanoma)	K18	Tumor	Worse
Stomach	K20	Peritoneal fluid	Worse
Uterus	Loss of K5/K6	Tumor	Worse

[a]Circulating tumor cells.

Table 1.3.3: Keratins as prognostic markers in tumour pathology (Adapted from, Karantza 2011[256]).

Use of keratins as diagnostic markers in tumour pathology is by far their most common application in the field of cancer. In cases remaining unclear on the basis of clinical presentation and conventional histopathology, keratin typing is especially valuable for correct tumour identification and subsequent selection of the most appropriate treatment plan.

Most breast adenocarcinomas, including ductal and lobular subtypes, constitutively express K7, K8, K18 and K19. Whereas K8 staining is predominantly peripheral in ductal carcinoma, in lobular carcinoma its pattern is more ring-like and perinuclear[260]. In basal-like subtype, poorly differentiated carcinoma, K5/6, K14 and K17 are also expressed as they are characteristic of the basal cells of stratified epithelium[14].

Beyond their well-established role as diagnostic markers in cancer, keratins have also been recognised as prognostic indicators in a variety of epithelial tumours (Table 1.3.3). Detection of disseminated keratin-positive tumour cells in the bone marrow of prostate cancer patients before surgery is an independent risk factor for metastasis within 48 months[261]. Some authors reported higher levels of K16 to correlate with poorer survival among breast cancer patients with metastatic relapses whereas higher levels of K5, K6 and K17 have been linked to a worse prognosis in triple negative breast cancer[262–264].

1.3.2.b Functional role in tumorigenesis – cell shape and invasive behaviour

Whether keratins play a role in tumorigenesis is still an open question. Some studies have shown that loss of K8 results in colorectal hyperplasia and inflammation in mice. Others reported K8 overexpression resulting in: i) loss of acinar architecture, dysplasia and hyperplasia in the pancreas, ii) preneoplastic alterations in the skin of aging mice and iii) malignant progression of benign skin tumours[256].

As AKT/mTOR pathway is frequently abnormally activated in aggressive tumours, it has been proposed that keratin-dependent activation of Akt signalling may play a role during tumorigenesis. Mechanistically, AKT/mTOR pathway might interact with keratins in multiple ways. K17 can stimulate the mTOR pathway regulating in this way protein synthesis, in turn, AKT isoforms regulate intermediate filament expression in epithelial cancer cell lines. IFs accumulation results in atypical PKC (Protein Kinase C) signal indicating that Akt1 may depend on keratins[243,252,256].

Regarding the role of keratins in cancer cell invasion and metastasis, supporting evidence indicated that these IFs might also be important in migration by influencing cell shape[265]. This process is well characterised at a physiologic level. Indeed, in case of epithelial injury, K5/K16

filaments reorganise from a pan-cytoplasmic to a perinuclear pattern that leads to keratinocyte migration into the wound site[266].

When epithelial tumour cells are incubated with sphingosylphosphorylcoline (SPC), K8-K18 filament are rapidly reorganised switching from a pan-cytoplasmic pattern to a ring-like juxtanuclear distribution[267]. SPC is a bioactive lipid present in high density lipoproteins (HDL) and found at increased level in blood and in malignant ascites of patient with ovarian cancer. The consequent redistribution changes the cellular shape and its viscoelastic properties increasing cellular elasticity and enhancing cell migration. This reorganisation is accompanied by phosphorylation, requires metabolic energy, is specific for SPC and independent of F-actin or microtubules.

Increased cell invasiveness after SPC treatment was later confirmed by another group showing how SPC enhances cell's deformability and increases cell's migration speed on flat surfaces[268]. By using a microchannel 3D based approach, they also described a drastic increase in the migration speed independently from SPC. This study suggest that the dimensionality of the environment strongly affects the migration phenotype and that the spatial cytoskeletal keratin organisation correlates with the tumour cell's invasive potential[268]. Interestingly, they distinguished two characteristic patterns of motion: a smooth sliding motion characterised by an equidistant movement of the front and rear of the cell, and a stepwise push-and-pull behaviour characterised by a variation of cell length in an oscillatory manner[268]. They propose a two-component model to explain the enhanced invasive behaviour upon SPC treatment. First, the motor unit with the lamellipodium in the front and the acto-myosin assembly at the rear; second, the passive and voluminous cell body being pulled upon migration. When the cell invades the channel, the keratin network compresses the nuclear region deforming the cell body.

In line with a model of continuous probing of the immediate extracellular surroundings, a dynamic keratin network enhances the adaptability and the functions of moving cells. Indeed, keratins can contribute to invasive behaviour in tumour cells even through interactions with the extracellular environment[265].

1.3.2.c Functional role in tumorigenesis – mechano-transduction and migration

During metastatic dissemination, tumour cells migrate through the extracellular matrix (ECM), a 3D microenvironment in the connective tissue made of dense and complex molecules. This infiltrative process requires a high grade of cellular deformability, especially of the nuclear

region. Nuclear deformation reflects the pulling and pushing behaviours by cancer cells to move their nucleus across confining environments. Nucleus has limited deformability, calculated at about 10% or the original diameter[269]. It has been hypothesised that mechanical stress during tumour infiltration triggering nuclear deformation may eventually lead to nuclear envelope rupture[270]. Although nuclear rupture can be rapidly repaired, potential consequences such as genomic alterations, DNA damage, double-strand breaks, and chromosomal-copy number changes could eventually contribute to tumour progression by favouring genomic instability and chromosomal rearrangements.

Decreased matrix pore size or increased nuclear stiffness trigger the collagenolytic program through transmembrane matrix-metalloproteinases (MT-MMPs)[271]. MT-MMPs proteolytically enlarge pores in the matrix facilitating tumour cells migration through dense tissues. MT-MMPs are concentrated in the invadopodia, actin-rich membrane protrusions forming from the leading edge of the invading cell[272]. One of the questions raising form this model is about how mechanical constraints on the nucleus can trigger invadopodia formation. As migrating cancer cells have to constantly adapt to environmental changes in the ECM, these responses have to happen on a rapid timescale. This open the avenue to emerging mechanisms of regulation such as epigenetic regulation and mechano-transduction intended as the ability of the nucleus to detect and respond to external forces[273]. As cells sense different environmental cues during migration, they become polarised initiating a series of cytoskeletal changes and signalling events ultimately leading their front and rear to acquire different functional properties. Protrusions such as filopodia, lamellipodia and invadopodia are generated from the leading edge of the cell membrane promoting contact with the ECM and directional migration[274–277]. Cell attachment to the ECM is maintained through mechano-transduction involving IFs. In order to maintain the epithelial morphogenesis, when the cell is subject to tension, hemi-desmosomes can act as mechano-sensors triggering intracellular signalling pathways. During collective cell migration, local traction forces are integrated to intra- and inter-cellular tension by cell-cell junctions[276,277]. In order to explore and move into the surrounding environment cells form filopodia and lamellipodia[278]. Lamellipodia are the main organelle for cell locomotion. They are formed by actin network, interacting and attaching to the environment via different adhesion molecules including integrins and cadherins. Filopodia originate from the basis of lamellipodia and penetrate into the surrounding environment through extensions formed by tight bundled actin fibres. They are considered as the sensory organs of the cell able to transduce extracellular signals like nutrients and chemo attractants. Filopodia- and

lamellipodia-like structures have been correlated with the invasive capacity of tumour cells and cancer progression[278].

In order to invade a three-dimensional matrix, cells form invadopodia. These protrusions are constituted by an actin-rich core surrounded by integrins and integrin-associated proteins like vinculin, and paxillin. They are provided of strong ECM proteolytic activity via the expression of MT-MMPs. However, filopodia and lamellipodia are highly interactive and inter-convertible structures with filopodia that can transform in lamellipodia by initiating actin nucleation typical of lamellipodia formation. Invadopodia share with lamellipodia the branched actin network and with filopodia the actin bundling structure, probably representing a hybrid of lamellipodia and filopodia structures.

1.3.3 Keratin 80

Keratin 80 is a largely unknown keratin. Its existence was undiscovered until the completion of a reference sequence of the human genome by the Human Genome Sequencing Project (International Human Genome Sequencing Consortium 2001) that allowed the identification of some novel keratins including Kb20[236,237]. Following international consensus at the 2004 Gordon Conference on Intermediate Filaments in Oxford, Kb20 was designated as K80 by the Human Genome Nomenclature Committee (HGNC)[235]. However, it has been properly characterised for the first time only few years ago[279] (Fig. 1.3.2).

Keratin 80 gene (*KRT80*) is located at the centromeric end of the type II keratin gene domain on chromosome 12q13.13 (Fig. 1.3.2). It encodes a 452 amino acid protein with a molecular mass of 50 kDa[237]. Compared to other type II keratins, Keratin 80 possesses a number of highly unusual properties [279]. Evolutionary tree analysis showed that *KRT80* is positioned between the epithelial and the trichocytes' keratins [279] (Fig. 1.3.2). In evolutionary terms, *KRT80* is one of the oldest keratins, conserved down to fish. K80 protein is encoded by an ancient gene that acquired the capacity of alternative splicing only in the mammalian line, comprising mainly primates but also cow[279] (Fig. 1.3.3).

K80 is ubiquitously expressed in all type of epithelia: stratified keratinising/non-keratinising, hard-keratinising, non-stratified tissues, simple epithelia in addition to the highly complex hair follicle[279]. Although its pre-eminently epithelial distribution, K80 is structurally closer to hair than to epithelial keratins showing a higher sequence affinity with the former[236]. This is also reflected in the non-alpha-helical domains containing a relatively high number of cysteine and proline residues along with the complete absence of GGX or GGG repeats present in many

type-II epithelial but not hair keratins[237]. Considering that there are no other type I keratins exhibiting structural intermediate properties between hair and epithelial, K80 structural properties make it unique within the IFs.

Peculiarly, and best seen in differentiating cells of the stratified epithelia, K80 accumulated along the cell margins close to the desmosomal plaques[279] (Fig. 1.3.4). Instead of exhibiting the conventional distribution of cytoplasmic IF network, they are tightly interwoven with the cytoplasmic IF bundles[237]. K80 IFs distribution represents an ideal mean to uniformly stabilise the cell membrane and render it able to cope with the high cellular deformability[279]. During terminal differentiation, K80 staining abruptly changes from cell margins localisation to be strongly cytoplasmic in the last living cell without leading to alterations in cell shape or volume[279] (Fig. 1.3.4). Cells of non-stratified simple epithelia also show K80 staining in the apical cytoplasm. In simple epithelial cells, K80 is apparently involved in the reinforcement of the IF zone of the apical region probably contributing to the stabilisation and maintenance of the polar nature of the cells[279].

Whereas alternative mRNA splicing has never been reported for any keratin, their high diversity has evolved mainly from gene duplication[238]. Langbein firstly demonstrated an alternative keratin splicing identifying functional protein variants for K80[279]. They distinguished the large protein variant previously described by Hesse and Rogers as K80 from the smaller carboxyl termini-truncated variant that they designated as K80.1[279] (422 amino acids and a molecular mass of 47 kDa). K80.1 expression is restricted to only distinct sub compartments of the hair follicle and the filiform tongue papilla. Keratins self-assemble to form 10-nm intermediate filaments on its own or, more commonly, in partnership with other IF proteins[238]. The remarkably extensive expression range underlies an unexpected unprecedented *in vivo* promiscuity of K80 that forms IFs with at least 21 different type I keratin partners. Considering K80.1 restricted expression, it is quite surprising that even the splice variant can partner with up to 16 different type I keratins[279].

Little is known about the function of keratin 80 in cancer. A recent study described KRT80 as an independent prognostic factor for patients with colorectal carcinoma (CRC)[280]. KRT80 was found highly expressed in two of the three cell lines used (high in SW620 and Caco-2, lower in RKO cells) and manipulation of KRT80 expression influenced migration and invasion in 2D. Augmented expression of K80 in CRC cells change their morphology from round to polygonal and promoted the expression of phospho-AKT (Ser 473) but not of phospho-AKT (Thr 308) or total AKT. They observed increased K80 expression in cancer, according to the tumour grade, compared to normal mucosa[280].
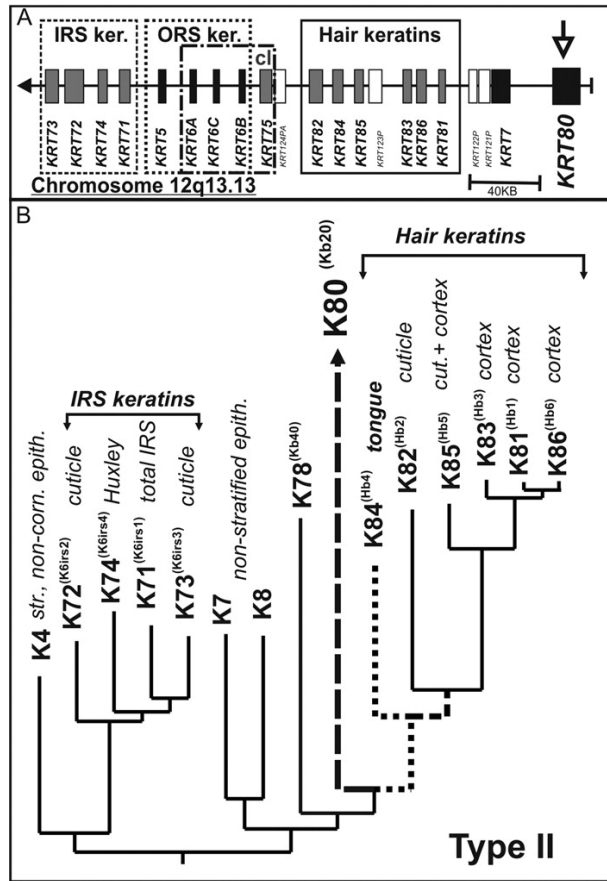
Figure 1.3.2: Analysis of the human K80 gene. A, localization of at the centromeric end of the type II keratin gene domain on chromosome 12q13.13. B, partial phylogenetic tree analysis of the α-helical rod domains of human type II keratins (Adapted from Langbein 2010[279]).
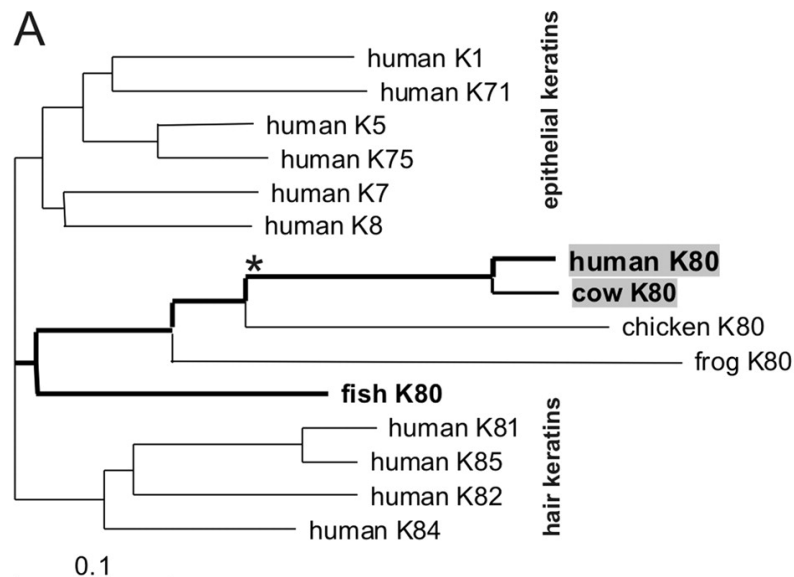
Figure 1.3.3: Evolution of the KRT80 gene. Phylogenetic analysis of K80 proteins of various species and representative human type II epithelial and hair keratins. The branch lengths in the tree are proportional to the number of substitutions per site (scale: 0.1 substitution per site) (Adapted from Langbein 2010[279]).
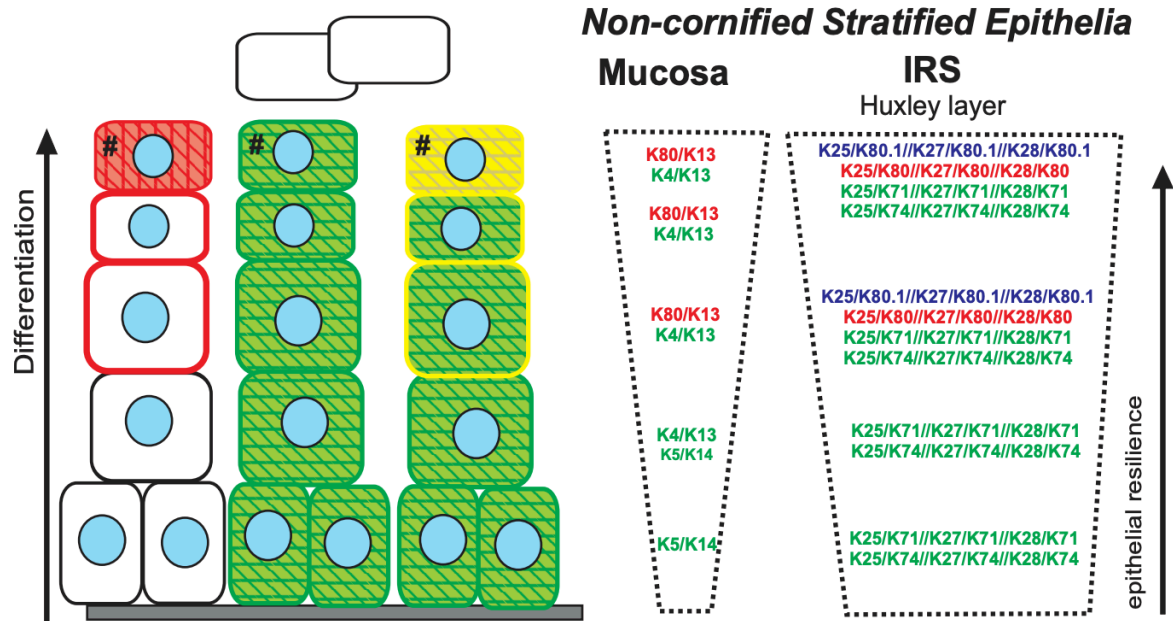
Figure 1.3.4: Expression scheme and cellular localization of K80 in epithelia. K80 is given in red, other keratins in green and the respective co-localizations in yellow. # last living cells. Pairing variants of the keratins of the various tissues are indicated at the right side. Increasing numbers of keratins parallel an increase of epithelial resilience (Adapted from Langbein 2010[279]).

1.4 Hypothesis and Aims

Increased *de novo* lipid synthesis is now a well-established hallmark of cancer as tumour cells reprogram their metabolism to provide energy and the essential building blocks required to maintain their aberrant survival and growth[59]. Our group has previously shown how breast cancer cells developing resistance to AI endogenously trigger transcription of the super pathway of cholesterol biosynthesis to promote sustained oestrogen independent ERα activation[53]. Considering what is known about the pathobiology of lipids in cancer, it is plausible to hypothesise that invading cells evolve mechanisms to bypass the tight homeostatic regulation of intracellular cholesterol adapting to their new environmental conditions (Fig. 2). With this idea in mind, we sought to investigate the role of SREBP1 as key regulator of *de novo* lipid biosynthesis in hormone-dependent cancers resistant to endocrine therapy (ET).

The main aims of the project were to:

1. Investigate SREBP1 regulation in hormone-dependent cancer cells
2. Examine genome-wide profiling of SREBP1 binding in breast and prostate cancer cells
3. Identify canonical and non-canonical targets of SREBP1

4. Investigate the role of Keratin 80 as a SREBP1 non-canonical target in breast cancer
5. Examine the effects of Keratin 80-driven cytoskeletal changes on tumour stiffness, migration and invasion
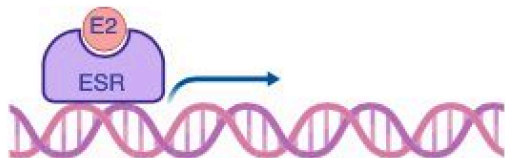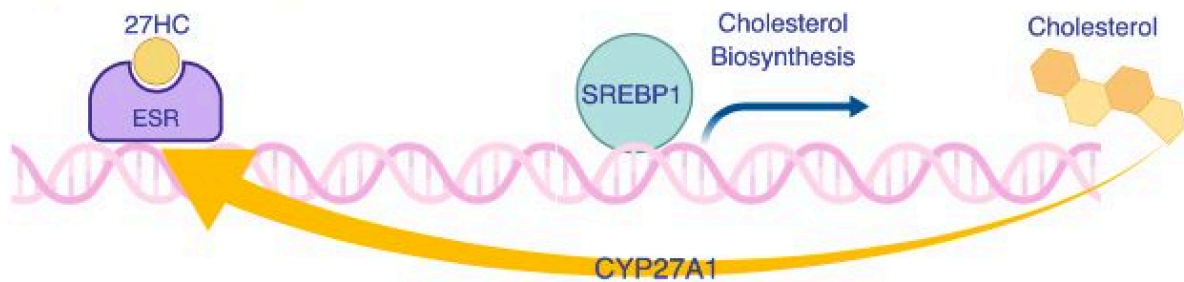
Figure 2: Schematic model of SREBP1 activation in AI-resistant ERα breast cancer. (A) Circulating oestrogens promote proliferation of ERα breast cancers. ERα activates gene transcription (arrow). (B) In post-menopausal women, inhibition of the conversion of androgens to oestrogens is sufficient to impair ERα binding and block ERα driven transcription. (C) Long-term exposure to aromatase inhibitors can lead to permanent epigenetic reprogramming especially near genes involved in cholesterol biosynthesis. In this scenario the chromatin structure near these genes become accessible to transcription factors such as SREBP1 leading to endogenous cholesterol biosynthesis. Ultimately cholesterol is transformed into ERα ligands (i.e. 27-hydroxil cholesterol) and becomes an alternative fuel for ERα mediated transcription. (oestradiol, E2; Oestrogen receptor α, ESR; Sterol regulatory binding element 1, SREBP1; 27 hydroxycholesterol, 27HC), (Modified from Perone 2016[281]).

"Ever tried. Ever failed. No matter.
Try again. Fail again. Fail better."
Samuel Beckett

Chapter 2: Materials and Methods

2.1 Materials
2.1.1 Reagents and materials

All chemicals were purchased from Sigma-Aldrich (Gillingham, UK) and VWR International (Lutterworth, UK) unless otherwise stated. Molecular biology reagents, such as DNA and protein markers, cDNA synthesis kits, Polymerase Chain Reaction (PCR) kits and the QubitTM dsDNA HS assay kit were obtained from Thermo Scientific (Life Technologies, Paisley, UK) and New England Biolabs (NEB; Hitchin, UK). Kits for DNA preparation were obtained from Invitrogen (Life Technologies, Paisley, UK) and kits for RNA preparation, plasmid DNA preparation and PCR purification were obtained from QIAGEN Ltd (Crawley, UK). Oligonucleotides were synthesised by Invitrogen. Library preparation kits were purchased from NEB or Illumina (Saffron Walden, UK). SYBR® Green real-time PCR reagents were supplied by Applied Biosystems (Life Technologies, Paisley, UK). Protease and phosphatase inhibitor cocktail tablets were obtained from Roche Diagnostics Ltd (Sussex, UK).

2.1.2 Sundries

1.5ml and 2ml microcentrifuge tubes and pipette tips were obtained from STARLAB Ltd (Milton Keynes, UK). 0.2ml Thermo-PCR tubes were obtained from Thermo Fisher Scientific (Life Technologies). Tissue culture and microbiology grade plastics and serological stripette pipettes were from Corning Incorporated and were obtained from Appleton Woods (Birmingham, UK). MicroAmp Fast 96-well reaction plate (0.1ml) was purchased from Applied Biosystems (Life Technologies).

2.1.3 Cell culture reagents

Cells were maintained at $37^O C$ in a humidified air atmosphere supplied with 5% CO_2 in RS Biotech Galaxy R+ CO_2 incubators (New Brunswick, Eppendorf, Cambridge, UK). Tissue culture work was carried out in BIOMAT 2 Microbiological Safety Cabinets (Contained Air Solutions Ltd, Manchester, UK). Dulbecco's Modified Eagle's Medium (DMEM), L-Glutamine-Penicillin-Streptomycin solution (200mM L-glutamine, 10,000 units/ml penicillin and 10 mg/ml streptomycin in 0.9% sodium chloride) (PSG), 0.02% ethylenediaminetetraacetic acid (EDTA) solution and 10X trypsin solution were supplied by

Sigma-Aldrich. Phenol red-free DMEM, Opti-MEM reduced serum medium were supplied by Gibco (Life Technologies). Fetal calf serum (FCS) and dextran-coated, charcoal-treated FCS (DC-FCS) were obtained from First Link Ltd (Birmingham, UK).

2.1.4 Chemicals

17β-oestradiol (E2), Fulvestrant and 4-Hydroxytamoxifen (OHT) were obtained from Sigma-Aldrich. Di-hydroxy-testosterone (DHT) was a kind gift of Damien Leach. Stock solutions were prepared by dissolving solid in ethanol at a concentration of 10mM and were then stored at -20°C.

2.1.5 General Stock Solutions

Solutions were made double distilled deionised water (ddH$_2$O), autoclaved and stored at room temperature unless otherwise stated. Phosphate buffered saline (PBS) (10x): 137 mM NaCL, 3mM KCl, 8mM Na$_2$HPO$_4$, 1.5mM KH$_2$PO$_4$, pH 7.4. Tris-EDTA (TE) buffer: 10mM Tris-HCl pH 8.0, 1mM Na-EDTA. Tween 20: 10% Tween 20 in PBS.

2.1.6 Microbiological reagents

Bacterial media were prepared using ddH$_2$O and autoclaved, before being stored at 4°C. Selection antibiotics were added to media freshly, prior to inoculation with bacteria. Luria-Bertani (LB)-broth and LB-broth agar capsules were obtained from MP Biomedical, LLC (Illkirch, France) (Table 2.1).

2.1.7 General Equipment

Large volume centrifugations were performed using a Sorvall RC6 Plus centrifuge with Fiberlite® F21-8 x50y and Fiberlite® F14-6 x 250y Fixed-Angle Rotors (Thermo Scientific). Bench top microcentrifuges, Sorvall Pico® (Sorvall, Leicester, UK) or Heraeus® Biofuge Pico® (Thermo Scientific) were used for volumes less than 2ml. Incubations were carried out in a water bath (Grant SUB Aqua Pro, Grant Instruments Ltd, Shepreth, UK), incubator oven (LEEC, Nottingham, UK), shaking incubator (New Brunswick Scientific Company Incorporated, Edison, USA) or block heater (Grant-bio, Wolf Laboratories, York, UK). Sonication was performed using a Bioruptor® sonication system (Diagenode, Liege, Belgium). PCR was carried out using an Applied Biosystems Veriti® 96 well thermal cycler (Life Technologies) and real-time quantitative PCR was carried out using an Applied Biosystems

7900HT Fast Real-Time PCR System (Life Technologies). Western blotting gel electrophoresis tanks and transfer apparatus were from Hoeffer (GE Healthcare Life Science, Buckinghamshire, UK) and Bio-Rad. DNA and RNA quantification were carried out using either spectrophotometric measurements from a Nanodrop® ND-1000 (Labtech International, UK) or using a Qubit® 2.0 fluorometer (Life Technologies). Samples were mixed using a Vortex-Genie 2 (Scientific Industries, London, UK). RNA and DNA quality assessment were carried out using an Agilent 2100 Bioanalyser (Agilent Technologies UK). RNA sequencing and ChIP-sequencing were performed using the Nextseq500 and HiSeq 4000 (Illumina). Absorbance readings were determined using Tecan Sunrise® absorbance microplate reader (Tecan UK Ltd, Reading, UK). Chemiluminescence imaging was performed using the Fusion Solo system (Vilber Lourmat, Collégien, France).

LB-broth	10g tryptone-B, 5g yeast extract-B, 10g NaCl in 1 litre ddH_2O
LB-plates	10g tryptone-B, 5g yeast extract-B, 10g NaCl, 15g agar-B in 1 litre ddH_2O, autoclaved, then cooled to 50^oC prior to addition of selection antibiotic. LB-agar was poured into Sterilin Petri dishes (Thermo Fisher Scientific), allowed to set and stored at 4^oC for up to one month.
S.O.C medium	2% Tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM $MgSO_4$, 20mM glucose. Purchased from Invitrogen.
Ampicillin	Ampicillin sodium salt (Sigma-Aldrich), dissolved in ddH_2O to 100mg/ml and sterilised by filtration through a 0.45μm filter, was used at a final concentration of 100μg/ml and was added to media freshly, prior to inoculation with bacteria.
Kanamycin	Kanamycin sulphate salt (Sigma-Aldrich), dissolved in ddH_2O to 50mg/ml and sterilised by filtration through a 0.45μm filter, was used at a final concentration of 50μg/ml and was added to media freshly, prior to inoculation with bacteria.

Table 2.1: Microbiological reagents.

2.2 Methods

2.2.1 Tissue culture

MCF7, T47D and ZR-75-1 breast adenocarcinoma cell lines were a kind gift of Prof. Philippa Darbre (Table 2.2), the prostate cancer cell line LNCaP and derived resistant clones were a kind gift of Prof. Jun Luo. All cells were routinely cultured in DMEM containing 10% FCS and PSG, also called full medium (FM) or red medium (RM). MCF7 were further supplemented with 17β-oestradiol (E2) ($1 \cdot 10^{-8}$ M final) (Table 2.3). LTED (Long Term Oestradiol Deprived) cells were derived from the parental cell lines upon one-year of oestradiol deprivation, mimicking aromatase inhibitor resistance. Long-term hormone deprived cells (LTED, LNCaP95 and ZR-75 Clone 11a) were cultured in phenol-red free DMEM supplemented with 10% DC-FCS and PSG, also called white medium (WM) or stripped medium (SM) (Table 2.3). MCF7 Fulvestrant resistant (MCF7F/MCF7 FulvR) and MCF7 Tamoxifen resistant (MCF7T/MCF7 TamR) cell lines was derived from MCF7 upon one-year treatment with Fulvestrant and Tamoxifen respectively (Fig. 2.1) and routinely cultured in FM. Short-term hormone depletion was achieved by culturing cells for up to 72 hours in WM. For E2 treatment, MCF7 cells were stimulated with E2 ($1 \cdot 10^{-8}$ M final) for 45 minutes following 72 hours of E2 deprivation. For DHT treatment, LNCaP cells were stimulated with DHT ($100 \cdot 10^{-9}$ M final) for 4 hours following 72 hours of DHT deprivation. Cell lines were routinely tested for Mycoplasma using MycoAlertTM Mycoplasma Detection Kit (Lonza, Cambridge, UK). Cells were incubated at 37°C in 5% CO_2, and maintained in a logarithmic phase of growth, in 75 cm^2 or 150cm^2 cell culture flasks (Corning) and passaged when cells reached 70-80% confluence. Medium was removed and cells were washed twice with PBS. Adherent cells were detached by addition of 1x trypsin solution (Sigma) in 0.02% EDTA solution (Sigma), pre-warmed to 37°C and incubated at 37°C for 3 minutes. After trypsinisation, 9ml of pre-warmed culture medium was added to neutralize the action of trypsin. Cells were serially passaged into tissue flasks by "splitting" at a ratio of 1:2 to 1:4. Cell lines in continuous culture were discarded after sub culturing for up to 25 passages. Cell lines were frozen and maintained for long-term storage in liquid nitrogen. Cultured cells growing in log-phase (70-80% confluency) were detached following the procedure used above for sub culturing, transferred to a 15ml falcon tube and pelleted by centrifugation at 1200rpm for 3 minutes. Cell pellets were resuspended in the appropriate volume of freezing medium (90% FCS, 10% DMSO). The cell suspension was aliquoted into 1.5 ml cryogenic vials in 1 ml volumes. Vials were

transferred to a Mr Frosty® cell freezing tub (Nalgene, Thermo Fisher Scientific), and left to cool in the-80 °C freezer for at least 24 hours, before being transferred on dry ice to the cell store in liquid nitrogen for long-term storage. For recovery of frozen cells, vials were rapidly warmed at 37°C for 3-5 minutes, and then pipetted drop-wise into a T75 flask containing 15 ml warmed medium.

Cell Line	Tissue	Cell type	ER status	PR status	HER2 status
MCF7	Mammary gland, breast; derived from metastatic pleural effusion	Epithelial	+	+	-
ZR75-1	Mammary gland, breast/duct; derived from metastatic fluid: ascites	Epithelial	+	+	-
T47D	Mammary gland, breast; derived from metastatic pleural effusion	Epithelial	+	+	-

Table 2.2: Breast cancer cell lines with corresponding hormone receptor status.

Cell line	Media	Serum	Antibiotics	Drug
MCF-7	DMEM	10% FBS	1% Glutamine/Pen/Strep	10^{-8} M 17β-oestradiol
MCF-7 Tam	DMEM	10% FBS	1% Glutamine/Pen/Strep	10^{-7} M 4-OH Tamoxifen
MCF-7 Fulv	DMEM	10% FBS	1% Glutamine/Pen/Strep	10^{-7} M Fulvestrant
LTED	DMEM (Phenol Red Free)	10% Double charcoal stripped FBS	1% Glutamine/Pen/Strep	
LTED Tam	DMEM (Phenol Red Free)	10% DCS FBS	1% Glutamine/Pen/Strep	10^{-7} M 4-OH Tamoxifen
LTED Fulv	DMEM (Phenol Red Free)	10% DCS FBS	1% Glutamine/Pen/Strep	10^{-7} M Fulvestrant

Table 2.3: Culture conditions.
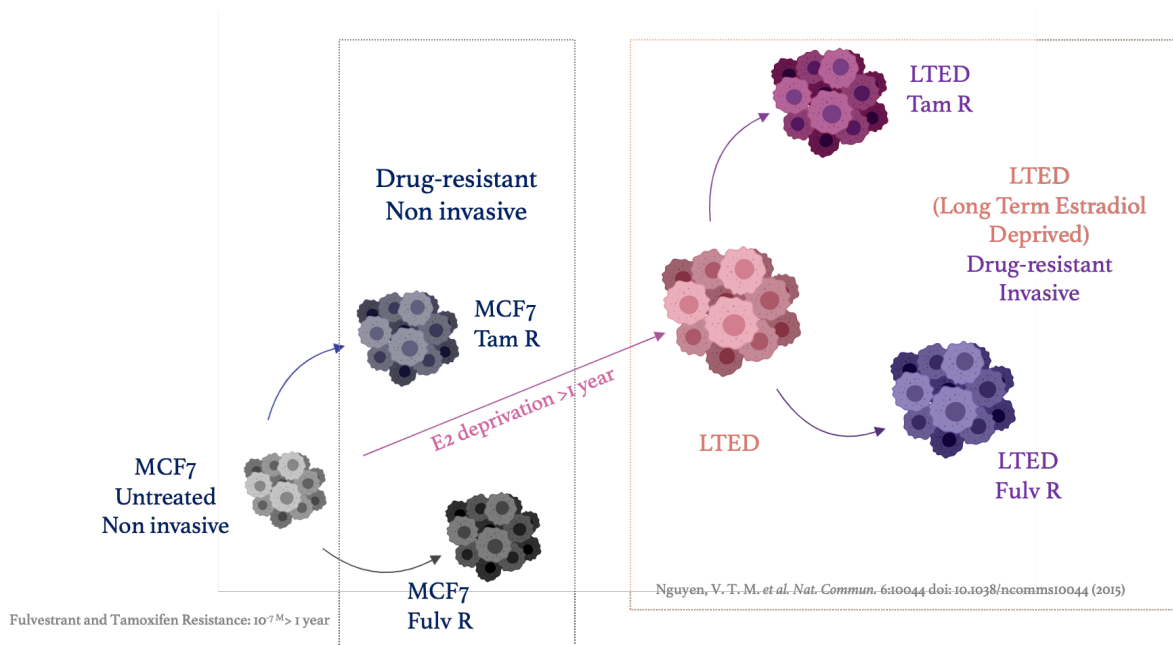
Figure 2.1: Breast cancer cell lines.

2.2.2 Generation of manipulated cell lines

For KRT80 overexpression, a full length KRT80 cDNA clone Myc-DKK-tagged was obtained from OriGene and transformed into DH5α competent cells (Invitrogen). CYP19A1 overexpressing cells were obtained by transfecting MCF7 cells with full length CYP19A1 (RC205890, OriGene Technologies) and selection using G418. Plasmidic DNA was isolated using Maxi-Prep Kit (QIAGEN) and transfected in MCF7 and LTED cells using X-tremeGENE 9 DNA Transfection Reagent (Roche) following manufacturer's instructions. Transfected cells, carrying Neomycin resistance, were selected with G418 (SIGMA), used at a final concentration of 1 mg/mL for MCF7 and 0.5 mg/mL for LTED. Knock-down of KRT80 was achieved by transfection of two different shRNA expression vectors and a scrambled negative control obtained from OriGene. Cells carrying the corresponding construct were selected with Puromycin (Sigma-Aldrich) at a final concentration of 1 ug/mL for MCF7 and 0.5 ug/mL for LTED cell line. NucLight Red Lentivirus (IncuCyte, 4627) was used to infect MCF7 and generate MCF7 mKate2. Stable and polyclonal cell populations were established after Zeocin selection (300 μg/ml). Small interfering RNA (siRNA) for CYP19A1 were obtained from ThermoFisher (siSilencer Select pre-validated s3875 and s3877) and Silencer negative control (Ambion; AM4611) were used for transfection experiments. The siRNAs were re-suspended in nuclease-free water at a concentration of 50μM, equivalent to 50pmol/μL. 1.5 x 10^5 cells were seeded, per well, using a 6-well plate (Costar; #3516). Cells were seeded in a 6-well plate; at a density of 3 x 10^3 cells. Following 24 hours, cells were then transfected with siRNA using Lipofectamine 3000 (Invitrogen; L3000015). Using the manufacturer protocol for a 6-well plate transfection, per well, 7.5μL of Lipofectamine 3000 was added to 125uL of Opti-MEM in one eppendorf whilst the 50μM stock siRNA (5 to 20nM final concentration per well) was added to 125uL of Opti-MEM. After five minutes, contents of both eppendorfs were mixed thoroughly and incubated for 15 minutes. The complex was then added to the wells containing phenol-free DMEM with 10% DCFCS and PSG. Compared to Lipofectamine 2000 (Invitrogen; #11668019) protocol the "siRNA-Lipofectamine 3000-Opti-MEM" can be added to culture media with PSG without a decrease in transfection efficiency. Following 24 hours, cells were then transfected as previously described above. Cells were harvested for protein/survival/gene expression analysis following at least 48 hours of transfection. The efficiency of the transfections was assessed either by RT-qPCR and/or western blotting.

2.2.3 Sulforhodamine B (SRB) growth assays

CYP19A1siRNA and full length CYP19A1 for overexpression were transfected at 5nM final concentration two days prior to SRB analysis (day 0). Twenty-four hours following seeding of 4×10^3 cells in 96-well plates, the medium was replaced with fresh medium supplemented with letrozole. After three days of culture in the presence of increasing amount of letrozole, cells were fixed using 100μl 40% w/v trichloroacetic acid (TCA) in ddH_2O, at $4^{\circ}C$ for a minimum of 1 hour. Cells were washed five times with distilled, deionised water and then stained with 0.4% w/v SRB (Sigma) in 1% v/v acetic acid in ddH_2O, for 1 hour at room temperature. Excess dye was removed following five washes in 1% acetic acid, and then the cells were left to dry at room temperature. Absorbance was determined using a Tecan SunriseTM microplate reader (Tecan), following the addition of 100μl of 10mM Tris-base to each well, and 10 minutes of incubation with shaking to resuspend the dye. Experiments were conducted using five technical replicates and three independent biological replicates. Average growth was plotted with error bars showing standard error of the mean (SEM).

2.2.4 3D Organoid assay

250,000 cells were resuspended in 1 mL of the corresponding medium and 20 μL drops were placed in the lid of a 10 cm dish (Corning). The lid was flipped over the dish containing 5 mL of medium in order to prevent evaporation. Hanging drops were incubated for 5 days at 37% C in a humidified atmosphere, during which formation of organoids was achieved. Before being included in 3D matrix for the invasion assay, the organoids were collected and labelled with 10 μM CellTracker® Green CMFDA (Thermo Fisher, Waltham, USA) dye by incubating them in serum free media for 45 minutes at 5% CO_2. Labelling solution was removed, and spheroids were washed in cell medium. To follow, spheroids were centrifuged at 300 rpm, immersed in 10 μL of phenol red free Matrigel® (BD Biosciences) and placed in a 24 well-plate (Corning) The appropriate media containing G418 or puromycin was subsequently added to the well. Brightfield images were acquired at days zero and day two using an EVOS microscope (Advanced Microscopy Group, Life Technologies). Images were analyzed using Fiji ImageJ software and fold-change area was calculated using the following formula: Area (fold-change) = Area Day 2/Area Day 0.

2.2.5 Immunofluorescence and confocal microscopy

Organoids were washed with PBS and fixed for 15 minutes with 4 % PFA/PBS. Fixation was stopped by rinsing with 100 mM Glycine/PBS. Cells were permeabilized with 0.5% Triton/PBS X-100 and unspecific binding was blocked with blocking solution (5% BSA, 0.2 % Triton X-100, 0.05% Tween in PBS) for 90 minutes. Organoids were then incubated with primary antibody (Rabbit Anti-KRT80 1:200, Sigma-Aldrich) for 2 hours, washed three times with washing buffer (0.2 % Triton X-100, 0.1 % BSA, 0.05% Tween in PBS), and incubated with secondary antibody (Goat Anti-Rabbit Alexa Fluor 555 1:200, Invitrogen) for 45 minutes. Organoids were washed with immunofluorescence buffer for 20 minutes and PBS for 10 minutes. Finally, organoids were mounted in Moviol (AppliChem) containing 5 μg/mL of DAPI (Lonza) and visualized using a Zeiss LSM-780 inverted confocal microscope.

2.2.6 Tissue specimens

Seventy-five human breast specimens and ten metastatic lymph nodes were selected from Histopathology Department at Charing Cross Hospital, with the previous approval of Imperial College Healthcare NHS Trust Tissue Bank. Thirty-two human primary and 32 matched metastatic breast samples were selected from European Institute of Oncology (IEO, Milan, Italy) Tissue Bank. Immunohistochemistry staining was scored using a quick score system by two independent investigators, one of them a consultant pathologist (SS). Score was calculated as follows: S=3 (strongly stained cells), S=2 (moderate staining), S=1 (poorly stained cells) and S=0 (absence of staining).

2.2.7 Statistical analysis

Data is presented as mean ± SD (standard deviation) in most figures. Whenever this is not the case, the figure legends state the exact details. Statistical tests were carried out using the GraphPad Prism 7.0 software (GraphPad Software, USA). For the comparisons of two experimental conditions a two-tailed student was performed and for multiple comparisons a two-way ANOVA was used. P-values less than 0.05 were deemed significant with asterisks indicating the level of significance: *$p<0.05$, **$p<0.01$ and ***$P<0.005$.

2.2.8 Survival analysis

Publicly available breast cancer datasets were identified in GEO (https://www.ncbi.nlm.nih.gov/geo/), EGA (https://www.ebi.ac.uk/ega/home), and TCGA

(https://cancergenome.nih.gov/). Only cohorts including at least 30 patients and with available follow-up data were included. Samples derived using different technological platforms (Affymetrix gene chips, Illumina gene chips, RNA-seq) were processed independently. For KRT80, the probe set 231849_at was used in the Affymetrix dataset, the probe ILMN_1705814 was used in the Illumina dataset and the gene 144501 was used in the RNA-seq dataset. Cox proportional hazards survival analysis was performed, Kaplan-Meier plots were derived to visualize survival differences. In the multivariate analysis, the RNA expression of ERα, HER2, and MKI67 were used as surrogate markers for ER and HER2 status, and for proliferation. In this, the probe sets 205225_at, 216836_s_at, and 212021_s_at were used for ERα, HER2, and MKI67, respectively. The survival analysis was performed for relapse-free survival (RFS), overall survival (OS), and post-progression survival (PPS). PPS was computed by extracting the RFS time from the OS time for patients having both RFS and OS data and having an event for RFS. Censoring data for PPS was derived from the OS event. The survival analysis was performed in the R statistical environment.

2.2.9 Designing and Cloning gRNAs

Guide RNAs required to direct the Cas9 (CRISPR-associated protein 9 nuclease) enzyme were designed and cloned into expression vectors. The web application for the Design and Optimization (CRISPR-DO) of guide sequences was used for CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) design and optimization. Designed sgRNA efficiency (0.88) and specificity (88.06) are indicated in Figure S1 A, B. In brief two gRNAs of 29 bp oligonucleotide sequences were designed complimentary to both the plus and minus strands of the SREBP1 binding motif we wished to target ending in the required Protospacer Adjacent Motif (PAM) sequence (Fig. 2.2 and S1 A, B). The PAM sequence is necessary for the Cas9 nuclease to target the DNA and denotes a sequence ending in NGG where N is any base pair followed by two guanines. The oligonucleotides (qRNAs) were then assembled into full double stranded DNA and cloned into a gRNA Cloning Vector which was a kind gift from George Church's lab (Addgene plasmid # 41824). Assembly involved annealing the gRNAs, extension and amplification through PCR, cloning into the vector using a Gibson Assembly Kit (New England Biolabs, Massachusetts, USA). The assembled vectors were transformed into chemically competent bacteria, grown overnight to allow amplification before harvesting using a GeneJET Plasmid miniprep kit (ThermoFisher Scientific). Correct assembly was validated using DNA Sanger Sequencing services provided by GENEWIZ Sanger DNA sequencing service (GENEWIZ, Takeley, UK).

SREBP1KO_1F

TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG*CCTCCAGAAGTACACGGCG*

SREBP1KO_1R

GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC*CGCCGTGTACTTCTGGAGG*C

TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG*CCTCCAGAAGTACACGGCG*
                                      C*GGAGGTCTTCATGTGCCGC*CAAAATCTCGATCTTTATCGTTCAATTTTATTCCGATCAG

SREBP1KO_2F

TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG*TGCCTCCAGAAGTACACGG*

SREBP1KO_2R

GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC*ACGGAGGTCTTCATGTGCC*

TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG*TGCCTCCAGAAGTACACGG*
                                        C*ACGGAGGTCTTCATGTGCC*CAAAATCTCGATCTTTATCGTTCAATTTTATTCCGATCAG

Figure 2.2: Two different CRISPR gRNA targeting SREBP1.

2.2.10 Gibson assembly cloning

2.2.10.a. Annealing, extension and end repair

CRISPR oligonucleotide pairs were annealed using the following reaction mix: 8μl 5x Phusion® HF buffer (NEB), 2μl forward oligonucleotide (100μM), 2μl reverse oligonucleotide (100μM) and ddH2O up to a total volume of 40μl. The reaction tubes were placed in boiling water and allowed to cool over 2 hours. Extension and end repair were performed by mixing 20μl of the annealed oligonucleotide with the following reaction mix: 4μl 5x Phusion® HF buffer, 0.8μl 10mM dNTP mix, 0.4μl Phusion® DNA polymerase and ddH2O to a total volume of 20μl (all from NEB). This reaction mix was then incubated at 72°C for 10 minutes, and briefly micro centrifuged.

2.2.10.b. PCR rescue of double stranded DNA template

Double stranded DNA templates were then PCR-rescued using the following PCR reaction mix: 5μl 5x Phusion® HF buffer (NEB), 1μl end repaired oligonucleotide, 5μl CRISPR primer mix (5μl 100μM PCR primer A and 5μl 100μM PCR primer B, together with ddH2O up to a total volume of 50μl), 1μl 10mM dNTP mix (NEB), 0.5μl Phusion® DNA polymerase (NEB) and ddH2O up to a total volume of 50μl. Cycling conditions are as follows: 98°C for 30

seconds, 25 cycles of 98°C for 15 seconds, 50°C for 15 seconds and 72°C for 15 seconds, with a final elongation step of 72°C for 2 minutes. Products were then checked using agarose gel electrophoresis. The remaining PCR product was purified using QIAquick PCR Purification Kit (QIAGEN) and diluted to a stock of 1ng/μl based on NanodropTM (Thermo Scientific) results.

2.2.10.c. Vector preparation

The vector was linearized, by restriction enzyme digestion with 1μl AFl II (20u/μl), 10μl 10x NE buffer 4, 1μl 100x BSA (100μg/ml), 5μg of vector DNA and ddH2O up to a total volume of 100μl. This reaction mix was incubated at 37°C overnight. Digestion was checked for completion using agarose gel electrophoresis, and then heat inactivated at 65°C for 20 minutes, followed by purification as above.

2.2.10.d. Gibson Assembly reaction

The Gibson Assembly reaction was set up as follows: 50ng of Afl II linearized vector DNA, 6.37ng of 100bp insert DNA, 10μl of Gibson Assembly® Master Mix (NEB) and ddH2O up to a volume of 20μl. The reaction mix was incubated in a thermocycler at 50°C for 60 minutes, and then stored on ice until transformation.

2.2.11 Transformation of bacterial competent cells with plasmid DNA

50μl of high efficiency chemically competent 10B E. Coli cells (NEB #C30191JH) were transferred to a chilled 1.5ml tube, and 2μl of Gibson assembled DNA product was added and mixed gently. Suspension was placed on ice for 30 minutes, and then heat shocked at 42°C for 30 seconds, before being placed on ice again for 2 minutes. 950μl of S.O.C. medium (Invitrogen) was added and contents were transferred to a 7ml plastic bijou tube. Tubes were incubated with shaking (250rpm) at 37°C for 60 minutes. 100μl of transformed cells were then spread onto warmed plates with appropriate antibiotics. Plates were then incubated overnight at 37°C.

2.2.12 Plasmid DNA purification

A single bacterial colony picked from a bacterial culture plate was transferred into 1.5ml pre-warmed LB broth containing the appropriate selection antibiotic. The culture was incubated for 8-16 hours at 37 °C with shaking at 200rpm. Plasmid DNA was prepared using QIAprep Spin Miniprep Kit (QIAGEN), according to manufacturer's instructions. For large-scale plasmid preparations for nucleofection, 500μl of the 1.5ml small culture was added to 200ml LB broth containing the appropriate selection antibiotic and incubated for 16 hours at 37 °C with shaking at 200 rpm. Plasmid DNA was prepared using QIAGEN Plasmid Maxi Kit (QIAGEN), according to manufacturer's instructions. Following elution in 400 μl of TE buffer, 400 μl of Phenol Chloroform isoamyl alcohol (25:24:1 saturated with 10nM Tris, pH8, 1nM EDTA) was added, followed by mixing and centrifugation at 13000 rpm for 5 minutes. The top layer (DNA) was transferred to a new 1.5 ml Eppendorf tube. 40 μl of 3M Na Acetate (pH5.2) and 1ml 100% ethanol was added, and the tube was inverted to mix, followed by incubation in -80 °C freezer for 15 minutes. The supernatant was then removed following a 10-minute centrifugation at 13000rpm, at 4 °C. The pellet was then washed twice gently in 700 μl 70% ethanol, and air dried for 10-15 minutes, before finally being re-suspended in 200 μl TE buffer. Plasmid DNA concentration and purity was determined by measuring absorbance at 260 and 280 nm with the NanoDrop® ND-1000 spectrophotometer.

2.2.13 CRISPR Transfection and Clone Isolation

CRISPR gRNA vectors were co-transfected with a pCas9-GFP plasmid which was a kind gift from Dr Simak Ali (Originally from Kiran Musunuru Addgene plasmid # 44719) into MCF7 cells using the Amaxa Cell Line Nucleofector Kit V (Lonza, Basel, Switzerland). 2 x106 MCF7 cells were suspended in the Nucleofector Solution and a 1:1 ratio solution (2 μg : 2 μg) of the gRNA and Cas9-GFP plasmids were added before transfer to a provided cuvette. The cuvette was placed in a 4D- Nucleofector System (Lonza, Basel, Switzerland) to electroporate the cells using program P-020. Cells were replated and left 6 hours before the media was replaced. Forty-eight hours after transfection cells were detached using 1 mM EDTA (Ethylenediaminetetraacetic acid), centrifuged and resuspended in FACS (phosphate-buffered saline (PBS) containing 1% FCS and 5 mM EDTA). Cell suspensions were then sorted at the Imperial College London MRC Flow Cytometry Facility using GFP expression. Cells were first sorted into 96 well plates using single cell sorting, after four plates were sorted the

remaining cells were sorted into a total population pool of clones.

2.2.14 CRISPR Validation Using Sanger Sequencing

CRISPR cells were lysed using a series of lysis buffers, first 10 minutes rotating at 4°C in Lysis Buffer 1 (LB1 - 50 mM HEPES-KOH, pH 7.5, 40 mM NaCL, 1mM EDTA, 10% glycerol, 0.5% NP-40, 0.15% Triton X-100), then a further 10 minutes rotating at 4°C in Lysis Buffer 2 (10 mM Tris- HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA and 0.5 mM EGTA) before sonication in Lysis Buffer 3 (10 mM TRIS- HCl, pH 8.0, 100mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate and 0.5% N-lauroylsarcosine). Cell lysates were sonicated (12 high cycles of 30 seconds on and 30 seconds off) to fracture the chromatin using a Biorupter Pico sonicator (Diagenode). DNA was then extracted using Phenol/chloroform extraction; Phenol/chloroform was added to samples at a 1:1 ratio vortexed and centrifuged, following this DNA is precipitated out using 5 M NaCl and 100% ethanol and finally eluted in water. DNA concentration was measured by a fluorometric quantitation assay using a Qubit 3.0 Fluorometer (ThermoFischer Scientific) using its standard protocol. The area of interest was subsequently amplified through a Polymerase Chain Reaction (PCR) using 100 ng of DNA and the Phusion High-Fidelity PCR Master Mix with High Fidelity buffer (New England Biolabs, Massachusetts, USA). Cloning primers (forward and reverse) were designed for SREBP1 (Table 2.4) to clone a 541 bp fragment spanning the area. PCR cycles were set; denaturation 98°C (30 seconds), 30 cycles of 98°C (5 seconds), 60°C (30 seconds), 72°C (30 seconds), final extension 72°C (5 minutes). PCR products were cleaned up using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) as per manufacturer's standard protocol. Purified PCR samples were sent for sequencing with sequencing primers (Table 2.5) using the GENEWIZ Sanger DNA sequencing service (GENEWIZ, Takeley, UK). The facility uses ABI 3730xl DNA analysers, with templates and sequencing primers being supplied by the user, following provided guidelines. DNA sequencing chromatograms generated via this service were viewed and analysed using SnapGene Viewer (Fig. S1B, GSL Biotech LLC, Chicago, USA).

PRIMER NAME CODE	SEQUENCE	TARGET
YPamp1 CRISPR	TGCCCCAAAGCTCAGAAGAG	CRISPR1/2 SREBP1 amplifying primers (expected size 541)
YPamp2 CRISPR	CAGCGTCTACCATAGCCCTG	CRISPR1/2 SREBP1 amplifying primers (expected size 541)

Table 2.4: Cloning CRISPR primers.

PRIMER NAME CODE	SEQUENCE	TARGET
YPseq1 CRISPR	ACAGCAGGCTCTGGAGGG	CRIPSR1/2 SREBP1 F
YPseq2 CRISPR	TTGTGAACTTGGGGCTCTGG	CRIPSR1/2 SREBP1 R

Table 2.5: Sequencing CRISPR primers.

2.2.15 Incucyte growth assays

Cells were seeded at a density of 100 cells per well, in 96-well plates, in FM. 24 hours later, plates were washed twice with PBS and the medium was replaced with fresh WM in order to start hormone starvation. Medium was changed every 3 days, with fresh medium. Images per well were taken every 6 hours for the first seven days and then every 24 hours for up to 47 days, using the IncuCyte ZOOM (Essen Bioscience, Welwyn Garden City, UK). Confluency (%) was calculated using the IncuCyte ZOOM software package (Essen Bioscience). Data were analysed and plotted using Prism6. Individual cells were counted longitudinally to verify absence/presence of proliferation.

2.2.16 Extraction of total RNA

For RNA preparation from cultured cells, culture medium was removed, and cells were washed twice with PBS. Cells were collected following scraping in RLT buffer (QIAGEN) containing 1% (v/v) β-mercaptoethanol. Cells were homogenised by centrifugation through a QIAShredder spin column (QIAGEN) at 13,000rpm for 2 minutes. RNA extraction was carried out using the RNeasy Mini Preparation Kit (QIAGEN) following manufacturer's instructions, with DNase treatment, using RNase-Free DNase (QIAGEN), prior to elution to remove genomic DNA. RNA prepared was eluted in 30μl of RNase-free water. RNA concentration and purity were determined using the NanoDrop® ND-1000 spectrophotemeter, to measure absorbance at 260 and 280nm.

2.2.17 Quality assessment of total RNA using bioanalyzer

The quality of RNA was assessed using an RNA 6000 Nano chip with the Agilent 2100 Biaoanalyzer (Agilent Technologies). The chip was prepared and primed with gel-dye mix according to manufacturer's instructions. RNA samples were diluted to be between 25-250 ng/μl. RNA samples and RNA 6000 Ladder (Agilent Technologies) were denatured with

heating at 70°C for 2 minutes. 5 μl of RNA 6000 Nano Marker (Agilent Technologies) were loaded into the wells of the RNA 6000 Nano chip, followed by 1 μl of denatured RNA sample or RNA 6000 Ladder. The chip was mixed by vortexing for 1 minute, before being loaded into the Agilent 2100 Bioanalyzer. Eukaryotic Total RNA Nano Series II assay was run on 2100 Expert Software (Agilent Technologies).

2.2.18 RNA-sequencing (RNA-seq)

Total RNA from each sample was quantified by Qubit® 2.0 fluorometer (Thermo Fisher Scientific) and quality checked by Agilent Bioanalyzer RNA 6000 Nano Chip. All samples have high quality RNA integrity number (RIN) with a RIN score > 7. One microgram of total RNA from each sample was used as starting material for paired-end RNA-seq library preparation using NEBNext rRNA Depletion Kit (NEB #E6310) and NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB #E7770) following the manufacturer's instructions. Libraries were sequenced on an Illumina Next Seq machine (#NextSeq500). Reads were processed using Kallisto and DEGS were called using Sleuth.

2.2.19 Complementary DNA (cDNA) preparation by reverse transcription

cDNA was prepared from total RNA using (thin-walled) PCR tubes and reagents from reverse transcriptase (RT) iScript cDNA synthesis kit (Bio-Rad; #1708891). Reactions were prepared from 1 μg of RNA, as shown in Table 2.5. The reaction mix was then incubated in a thermal cycler (Applied Biosystems; Veriti; # 4375786) using the following thermal profile settings: 1. Priming; 5 minutes at 25°C, 2. Reverse transcription; 20 minutes at 46°C, RT inactivation; 1 minute at 95°C and Optional step; 4°C holding temperature. cDNA samples are diluted 1 in 10 with nuclease-free water and subsequently, 2.5 μL of diluted cDNA was used for each qPCR reaction.

.Component	Volume per reaction (μL)
5X iScript Reaction Mix	4
IScript Reverse Transcriptase	1
Nuclease-free water	Variable
RNA template (100fg-1μg)	Variable
Total volume	**20**

Table 2.6: Reagents required per reaction to reverse transcribe 1μg of total RNA to cDNA using the iScript kit.

2.2.20 Single cell RNA-FISH

Cell were cultured, fixed and pretreated according to the protocol for the RNAscope® (Single-cell RNA Fluorescence in Situ Hybridization (sc-RNA FISH)) Multiplex Fluorescent Reagent Kit v2 Assay provided by Advanced Cell Diagnostics (ACD, #323100, Nunc Lab Tek II 2 Well Glass Slides, #154461K). The assay was run following the manufacturer's instructions, hybridization was performed overnight. PerkinElmer TSA Plus Fluorophores (fluorescein, NEL741001KT and Cyanin 3, NEL744001KT) were diluted at 1:1300 and assigned to the channels HRP-C1 and HRP-C2, respectively. Samples were imaged using a ×60 objective with a Ti Nikon microscope equipped with a spinning disk (CAIRN) and analysed in Image J.

Fluorophores	Part number
PerkinElmer TSA Plus Fluorescein System	NEL741001KT
PerkinElmer TSA Plus Cyanine 3 System	NEL744001KT

Table 2.7: RNA scope fluorophores.

Catalog Number	Product Name	Probe Type	Channel	Species Name
320861	RNAscope® 3-plex Positive Control Probe	RNAscope® Control Probes	2	Homo sapiens
310043	RNAscope® Negative Control Probe - DapB	RNAscope® Control Probes	1	Other*
300031	RNAscope® Target Probe C1	RNAscope® 20ZZ probe named Hs-KRT80 targeting 294-1445 of NM_182507.2	1	Homo sapiens
300031	RNAscope® Target Probe C2	RNAscope® RNAscope® Probe - Hs-SQLE 465071	2	Homo sapiens
300031	RNAscope® Target Probe C2	RNAscope® 20ZZ probe named Hs-SREBF1 targeting 958-2002 of NM_001005291.2	2	Homo sapiens
470561	RNAscope® Probe Hs-HMGCR	RNAscope® Target Probes	1	Homo sapiens

Table 2.8: RNAscope probes. *Bacillus subtilis dihydrodipicolinate reductase (dapB) gene.

2.2.21 Quantitative real-time PCR

Real time-qPCR (RT-qPCR) reactions were carried out in 10 uL volume containing 5 μL of Sybergreen mix (ABI; 4472918), 0.5 μL of primer (2.5 uM final concentration), 2.5 μL of genomic DNA and 2 μl of DNASE/RNASE–free water. A three-step cycle programme and a melting analysis were applied. The cycling steps were as follows: 10s at 95°C, 30s at 60°C and 30s at 72°C, repeated 40 times. For the case of assessing gene expression, the expression of the test gene was normalised to the *28S* housekeeping gene and relative to the appropriate control using the 2-ΔΔCT method. Primer sequences for used for gene expression profiling in Table 2.9.

TARGET	SEQUENCE- Forward Primer	SEQUENCE - Reverse Primer
28S	CGATCCATCATCCGCAATG	AGCCAAGCTCAGCGCAAC
KRT80	GAGGAACTGCGCAAAGTGAG	TGCGCTTGGAGATCTCATCC
HMGCR	AGTGACACTGACCATCTGCA	AGGATGGCTATGCATCGTGT
SREBP1	AGCACCTAGGGAAAGGCTTC	GAGGAGGCTTCTTTGCTGTG
SREBP2	ATCGCTCCTCCATCAATGAC	TTCCTCAGAACGCCAGACTT
MSMO1	TCACTTCTGTATCCAGCTGCC	CACAACCAAAGCATCTTGCCA
SQLE	TTGGAGAATTCCTGCAGCCG	GGGCATCAAGACCTTCCACT

Table 2.9: Primer sequences of targets used for RT-qPCR experiments.

2.2.22 Protein extraction, quantification and western blotting

Cells were harvested in 50μL ice-cold RIPA buffer (50 mM Tris- HCl at pH 8.0, with 150 mM sodium chloride, 1.0% Igepal CA-630 (NP-40), 0.5% sodium deoxycholate, and 0.1% sodium dodecyl sulphate) (Sigma; #R02780), supplemented with 1X protease (Roche; #11697498001) and 1X phosphatase (Sigma; #93482) inhibitor cocktail. The cell pellet and RIPA were mixed by pipetting up and down, incubated at 4°C for 30 minutes and vortexed every 5 minutes. Cell lysates were then centrifuged at 13,000 rpm for 30 minutes at 4°C. The supernatants were transferred to a new 1.5mL eppendorf tube and the pellets were discarded. Protein concentration was measured using BCA Assay Kit (Thermo Fisher) following manufacturer's instructions. In brief, bovine serum albumin (BSA) standards were generated (1 to 30μg) in a 50μL volume. 5μl protein lysates were diluted 10x in ddH2O to obtain a 50μL volume. BCA reagents A and B were mixed with a ratio of 50 to 1, and 1mL of mixed BCA reagent were added to each sample and standard. The samples were mixed and incubated at 37°C for 30 minutes. The samples were then transferred into a micro-cuvette and the optical density was

recorded at absorbance reading of 562 nm. The protein concentration of the samples was determined by comparing the measured protein values to the BSA standard curve. With regard to western blotting, 20µg of protein per sample, were mixed with 4X Bolt sample buffer (Life Technologies; #B0007), 10X Bolt sample reducing agent (Life Technologies; #B0009), ddH2O and heated at 95$^{\text{O}}$C prior to loading. Protein lysate were loaded into BOLT 4-12% Bis-Tris Plus Gel (Life Technologies; NW04120BOX). The pre-made gel was placed into a mini gel tank (Life Technologies; #A25977) containing 1X Bolt running buffer (Life Technologies). Electrophoresis was carried out at 90V for 35 minutes to allow proteins to adequately run through and also until the bromophenol blue dye reached the bottom of the gels. The gels were transferred into a Biotrace nitrocellulose membrane (VWR; #PN66485) using a TE-22 transfer unit (Hoefer GE Healthcare) at 100V for 90 minutes. The membrane was incubated in blocking buffer for 45 minutes at room temperature to reduce non-specific binding of primary antibody. The membrane was then incubated with the diluted primary antibodies (Anti-KRT80 for shRNA and IHC from HPA 077836 and 077918, Atlas Antibodies (1:200 dilution), Anti-SREBP1 H-160 sc-8984 Santa Cruz Biotechnology (1:200 dilution), (Guinea Pig Anti-KRT80 1:5,000; Mouse Anti-DKK 1:1,000, OriGene; Mouse Anti-β-Actin 1:10,000) in blocking buffer at 4$^{\text{O}}$C and allowed to shake overnight. After primary antibody incubation, the membrane was washed three times in PBST (5 minutes per wash on a rocking platform) and then incubated for 1 hour with the HRP-GAPDH (Abcam; #ab9482 (1:5000 dilution)) conjugated antibody (for the loading control membrane) which was diluted in 5% BSA/PBST and goat anti-rabbit IgG (H+L) Cross Absorbed secondary antibody, HRP 1:20000 dilution (ThermoFisher Scientific; #31462). The membranes (including the loading control membrane) were washed three times in PBST. Amersham ECL start Western Blotting Detection reagent (GE Healthcare Life Sciences; #RPN3243) was used for chemiluminescent imaging using the Fusion solo (Vilber; Germany) imager.

2.2.23 Preparation of cell lines for ChIP

Cell lines were seeded at a density of 4×10^6 cells in a 15cm^2 dish (Corning; #CLS430597). The cells were maintained for 48 hours to allow a 75-85% confluency to be reached. The cells were either treated or not depending on the experimental design prior to fixation. 10ml of 1% formaldehyde in "Solution A" (Table 2.2) was added to each dish and incubated at 37$^{\text{O}}$C for 10 minutes. Cold glycine (1M) was then added at a ratio of 1/10 the volume of the media in

the 15cm^2 dish and incubated at 4°C for 8 minutes. The media was then removed, and the cells were washed three times with cold autoclaved phosphate buffered saline (PBS). The cells were harvested using a cell scraper (Corning; #CLS3008) using 500μL of cold PBS supplemented with 1X EDTA-free protease inhibitor (PI) (Sigma; #S8830) to each dish. The cell suspension was collected in a 14ml and centrifuged at 3000rpm for 3 minutes (Eppendorf; #5810). The fixed cell pellets were either stored at -80°C, until required, or subsequently used for ChIP.

2.2.24 Chromatin Immunoprecipitation (ChIP)

For ChIP, cells were fixed with 1% formaldehyde for 10 min at 37°C and reaction was quenched with 0.1M glycine. The cells were subsequently washed twice with PBS after which they were lysed in lysis buffer (LB) 1, for 10 minutes, then for 5 minutes in LB 2 and subsequently eluted in LB 3 for sonication (all buffers are listed in Table 2.10). DNA was sheared using the Bioruptor® Pico sonication device (High, 10 cycles of 30'' on and 30'' off) (Diagenode). Sheared chromatin was cleared by centrifugation. Magnetic beads were precoated by adding 10 μg of antibody Rabbit-anti- SREBP1 (H-160): sc-8984 (Santa Cruz Biotechnology, Inc.); and 4 ug of Rabbit- anti-Histone H3 acetyl K27 antibody (abcam, ab4729); Rabbit- anti-Histone H3 (monomethyl K4) antibody (abcam, ab8895); Rabbit- anti-Histone H3 (dimethyl K4) antibody (abcam, ab7766) to 100 μl magnetic beads per ChIP (Dynabeads protein A, Life technologies) and incubated for 6 hours on a rotating platform at 4°C. Diluted sheared chromatin was added to the coated magnetic beads and incubated on a rotating platform at 4°C O/N. 10 μl of sheared chromatin taken as input and treated the same. The next day magnetic bead complexes were washed three times with RIPA buffer (50 mM HEPES pH 7.6, 1mM EDTA, 0.7% Na deoxycholate, 1% NP-40, 0.5 M LiCL) and two times with TE buffer (10mM Tris pH 8.0, 1 mM EDTA). DNA is O/N eluted from the beads in 100 μl de-crosslinking buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% SDS) at 65 °C. After overnight de-crosslinking, DNA was treated with 2.7 μl of 1mg/ml RibonucleaseA (RNaseA) for 30 min at 37°C and subsequently incubated with 1.3 μl of 20 mg/ml proteinase K (Invitrogen) for 1 hour at 55 °C. Then DNA extraction was performed using SPRI magnetic beads (Beckman Coulter, B23318). After elution in TE buffer, DNA was quantified using Qubit (ThermoFisher Scientific; Qubit 3.0 Fluorometer; #Q33216) high sensitivity assay (ThermoFisher Scientific; #33216). Quantitative polymerase chain reaction (qPCR) was then carried out (Applied Biosystems; #7900HT Real time PCR, #StePOnePlus, primers used are listed in Table 2.11). If sufficient enrichment is seen in the antibody treatment samples over

the 'input' samples and compared to internal negative controls, these undergo DNA size selection (which aims to retain DNA fragments between 200-300 base pairs, recognisable for the Illumina HiSeq Sequencer) and library preparation.

As further described in Results, the ChIP protocol has been optimised for SREBP1 antibodies modifying the number of cells used (from 5 million each cell line to 50 million), quantity of antibody used (from 4 μg to 10 μg) and Dynabeads protein A, Life technologies (100 μl instead of 20 μl) accordingly. Instead of phenol-chloroform extraction, DNA is extracted using SPRI beads (Beckman Coulter, B23318) (1:1) and following Cleanup Library prep for NEBNExt Ultra II DNA Library Prep Kit for Illumina (New England Biolabs) protocol. The rationale behind these changes was to get more 'pulled-down' DNA fragments in order to measure the enrichment with quantitative PCR and have enough starting material for a library prep.

2.2.25 Library preparation and ChIP-seq data analysis

Prior to sequencing, ChIP samples were library prepared using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, NEBNext Ultra II DNA library prep kit for Illumina, #E7770, NEBNext Multiplex Oligos for Illumina, #E7335L). Adaptor ligated DNA was size selected with SPRI (Solid Phase Reversible Immobilisation Beads) magnetic beads (Beckman Coulter, B23318) which aim to retain DNA fragments between 200-300 base pairs (bp), recognisable for the Illumina sequencer (#NextSeq500). After library preparation, we performed qPCR, high sensitivity DNA quantification and size selection measurement (Agilent Bioanalyzer 2100 system + High sensitivity DNA measurement assay; 5067-4626) before sending samples for sequencing. Raw sequencing files processed by the Illumina NextSeq500 sequencer were obtained in "FASTQ" format. **TF-ChIP-seq data analysis.** In order to remove low-quality bases and adapter sequences, FASTQ files were initially processed using *fastx_trimmer* from FASTX-Toolkit (-f 1 -l 50; http://hannonlab.cshl.edu/fastx_toolkit/) and Trim Galore (-q 20; https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), respectively. Trimmed reads were then aligned to the hg38 reference genome using bowtie2 v2.2.9[282]. The output of Bowtie is the "SAM" file extension format, for both input (control) and ChIP samples. To discard multi-mapping reads, SAMTools *view* v1.3.1 with -q 10 was used[283]. Considering only de-duplicated reads, unfiltered TF-bound locations were identified using Model-based analysis for ChIPSeq (MACS v1.4) for peak calling[284] with parameters --gsize=mm --nomodel --shiftsize=100, and sample-matched input DNA as control. MACS outputs result in "BED" file format and "WIG" files. Genome-wide coverage profiles were

generated by virtue of *genomeCoverageBed* from BEDTools v2.25.0[285] and normalized to RPM (Reads Per Million sequenced reads). Peaks called either on the mitochondrial or random chromosomes, as well as those overlapping the ENCODE blacklist[286], were excluded. Reproducible peaks across replicates were defined according to the following conditions: (1) an overlapping peak was independently called in both replicates; (2) an average enrichment *vs* input of at least 2-fold (the normalized reads density in the input was estimated using a window of 10 kb centered on the peak summit, separately for each replicate); (3) an RPKM of at least 2, separately for both replicates (Reads Per Kilobase per Million sequenced reads); (4) an average peak size greater than 0.5 kb but smaller than 2 kb. **Copy Number estimations from ChIP-seq input DNA.** Copy number was estimated using de-duplicated reads from input DNAs via CNVkit v 0.9.6[287]. cnvkit.py *batch* was run with the following parameters: -n -m wgs, followed by cnvkit.py *call* (-y -m threshold -t=-1.1, -0.4,0.3,0.7). **Differential binding analysis.** Three separate master lists were obtained for MCF7, T47D and LNCaP cells. Regions in these master lists were annotated to the TSS of the nearest RefSeq gene (table downloaded from the UCSC genome browser on Dec 13, 2018), using a custom script. Regions within 2.5 kb of a TSS were defined as TSS-proximal, those located further as TSS-distal. Samples were hierarchically clustered (using Ward's method) based on the normalized expression values (RPKM) across the regions included in each master lists. Pairwise distance between samples was measured as one minus the Spearman's correlation coefficient between each pair. For each cell type, reproducible peaks from any of the profiled conditions were stitched together using *mergeBed* from BEDTools. Peaks overlapping with chromosomal regions predicted to have a copy number equal or higher than 5 in any of the cell lines considered were excluded to avoid artefactual differential calls. *edgeR* (v3.26.3)[288] was then used for normalization and to estimate the differentially bound regions in each condition, using one of them as reference (MCF7 red media, T47D red media and LNCaP, respectively). After re-quantification of RPKM for each region, *estimateCommonDisp* and *estimateTagwiseDisp* were run separately, followed by Relative Log Expression (RLE) normalization. Differential binding was evaluated using the function *exactTest*. A region was defined as differentially bound if showing a *q*-value <= 0.05 (Benjamini-Hochberg correction). Publicly available nuclear receptors (ERα and AR) ChIP-seq data were overlapped with the SREBP1-bound regions identified in our MCF7, T47D and LNCaP cells' datasets by using Giggle[289]. **TF-binding sites enrichment analysis.** Pscan-ChIP[290] was used to identify over-represented, known TF-motifs in the sequence of a given set of genomic regions. The analysis were run on the webserver (http://159.149.160.88/pscan_chip_dev/) using *Promoters* as background, and *Jaspar 2018 NR*

as database of known binding sites. Over-represented motifs were then sorted by global *p*-value.

2.2.26 Cancer hotspot mutations

I performed targeted capture using NEB Cancer Hotspot panel modified to include ESR1 ligand binding domain (NEB E7000X). Sonicated Input material from ChIP-seq analysis (frozen tissues) was used as an input (minimum 50ng) as specified by the manufacturer. Sequencing was performed on a NextSeq Illumina machine by multiplexing 24 samples per lane in two lanes (Single End 75bp flow cell). Single-end 75-base pairs reads were aligned to the hg38 human reference genome using bwa1 version 0.7.15 (parameters: -q 0). Samtools (PMID: 19505943) version 1.3.1 was then used to obtain indexed bam files. Aligned reads from each captured sample were pre-processed using Picard (http://broadinstitute.github.io/picard) version 2.6.0, applying functions AddOrReplaceReadGroups (parameters: RGID=1 RGLB=lib1 RGPL=illumina RGPU=unit1 RGSM=1) and sortSam (parameters: SORT_ORDER=coordinate). GATK 2 version 3.6 was then used for variant identification. PCR duplicates were marked using the MarkDuplicates function from Picard (parameters REMOVE_DUPLICATES=False AS=True). Re-alignment around indels was performed using functions RealignerTargetCreator and IndelRealigner from GATK (known indels from the GATK bundle: Mills_and_1000G_gold_standard.indels.hg38.vcf). This step was followed by base quality score recalibration (GTAK BaseRecalibrator). Mutect2 (part of GATK v3.6) was finally run separately on each capture, without control samples. The identified variants were then annotated to known SNPs (1000G_phase1.snps.high_confidence.hg38.vcf in the GATK bundle) and to COSMIC 3 version 34 (hg38). Variants showing alternate allele frequency lower than 1% were excluded from further analyses. Those supported by evidence from both alleles and covered by ten or more reads were retained. Variants overlapping known SNPs were excluded. Among the remaining variants, only those previously reported in COSMIC were kept. As a final step, those protein-coding variants predicted as "Neutral" by FATHMM 4were filtered out. Reads were quality controlled with FastQC v0.11.5 and aligned to the human hg38 reference using bowtie v1.1.2 5 with default parameters. The generated sequence alignments were converted into binary files (BAM), then sorted and indexed using the SAMtools.

Reagents	Composition/Concentrations	Storage
Solution A + Formaldehyde	50 mM Hepes-KOH, 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 1% formaldehyde	Room temperature
Glycine	1M	4 °C
Lysis Buffer 1 (LB1)	50 mM Hepes-KOH, pH 7.5; 140 mM NaCl; 1mM EDTA; 10% Glycerol; 0.5% NP-40 or Igepal CA-630; 0.25% Triton X-100	4 °C
Lysis Buffer 2 (LB2)	10 mM Tris-HCL, pH8.0; 200 mM NaCl; 1 mM EDTA; 0.5 mM EGTA	4 °C
Lysis Buffer 3 (LB3)	10 mM Tris-HCl, pH 8; 100 mM NaCl; 1 mM EDTA; 0.5 mM EGTA; 0.1% Na-Deoxycholate; 0.5% N-lauroylsarcosine	4 °C
RIPA buffer	50 mM Hepes-KOH, pH 7.5; 500 mM LiCl; 1 mM EDTA; 1% NP-40 or Igepal CA-630; 0.7% Na-Deoxycholate	Room temperature

Table 2.10: Reagents used for ChIP.

TARGET	SEQUENCE- Forward Primer	SEQUENCE - Reverse Primer
SQLE (PROMOTER)	ACGCAACATCGCCCTAGC	CCCCTACCCCTGCTCCTG
KRT80 (PROMOTER)	TAGGTCTCCCTCCCAACAGG	TTCCTTTTGCCAACTCCCCA
KRT80 alt (PROMOTER)	GCAATGGAAGGCATGCTAGC	TTCCTGTTTGGCAGCCACA
ACLY (PROMOTER)	GCCTGCTGGGACTTGTAGTC	CTCAAGCGATCAGGCCACAA
ACACA (PROMOTER)	TGACTACGTCCGGGGTTACT	ATCGCCTCACATCGCTCC
FASN prox (PROMOTER)	CGTCTCTCTGGCTCCCTCTA	AAGCTGTCAGCCCATGTGG
FASN distal (PROMOTER)	GCAGCAGCAACCAATCGG	CCATCACCCTATCGCCTAGC
SCD1 (PROMOTER)	AGGAGAAACAGAGGGGAGGG	GGCTTCTGTAAACTCCGGCT
GPAM (PROMOTER)	GAAACACCGCAAGAGGAGGA	AAGCCCTGAAATGCACTCCA
RPL27 (PROMOTER)	TTGTCTACTGCTGGAGCTGC	CTTTGCCACGCTGATTGTGT
LINC00263 (PROMOTER)	GGTTACTGCCGGTCACAGAC	AATCAGACAAGCTTGGGGGC
ATP6AP1 (PROMOTER)	GAGACACGTACAGCCAACCA	CAGATCAGGTGACCGTTGCC
KRT80 (ENHANCER)	TGTCAGAAGCTGTTTGAGGCA	GAGATGCCACAGAAGCTGGT
SREBP1-c (PROMOTER)	GGGCTCGAGTTTCACCCC	CAGTAACCCCGGCAGACG
KRT80 2 (ENHANCER)	CACCACCCTGCTTTGTCAGA	CAGAAGCTGGTGTTCTGCCA
KRT80 3 (ENHANCER)	AGAAGCTGTTTGAGGCAGAGA	ATGCCACAGAAGCTGGTGTT
KRT80 4 (ENHANCER)	CTGCACCACCCTGCTTTGTC	CTCTCCCCACCTGGCAGAAG
KRT80 5 (ENHANCER)	CTGGTGCAGTGGGGTGAG	CCTCTCCCCACCTGGCAG
Hetero-chromatinic region ML5	TGGGTGGTGTCATCTGGTAA	GGATGGAATGGATCAGATGG
Hetero-chromatinic region; non-functioning ERE ML13	CTTCTTCCTTCCGGCTTTCT	AGCTGGGAGAGGACACACAC

Table 2.11: List of primers used for ChIP-qPCR.

"The more perturbing the stimulations,
the more compelling and vital
is the need to adapt"
Guido Morselli, Dissipatio H.G.

Chapter 3: Results I-SREBP1 regulation in hormone-dependent cancer cells

3.1 SREBP1-driven cholesterol biosynthesis activation in hormone-dependent cancer cells

It has been previously reported by our group that AI resistant cells upregulate lipid/cholesterol biosynthesis via global epigenetic reprogramming[53]. In order to gain more detailed insights into global changes in transcription factor (TF) occupancy, the accessible chromatin landscape was investigated in breast cancer cells using high-depth DHS-seq (DNase I hypersensitivity sequencing). DHS-seq revealed that chromatin regions increasing DNA accessibility in AI resistant breast cancer cells triggering *de novo* cholesterol biosynthesis were enriched for the SREBP1 binding site[53]. RNA-seq data suggest a role as a transcriptional regulator specifically for SREBP1 rather than SREBP2, that is not found expressed in BCa cells [53]. By using reverse transcription RT-qPCR, I validated these results confirming higher levels of SREBP1 mRNA in LTED cells, but no increase of SREBP2 (Fig. 3.1 A, B). LTED (Long-term oestradiol deprived from MCF7 and T47D cells) are models used to study the development of resistance to AI in breast cancer without showing any reduction in ERα number or function[291].

Higher mRNA levels of SREBP1 were accompanied by increase in the expression of downstream key enzymes in LTED cells compared to MCF7 (Fig. 3.1 B), further supporting the role of SREBP1 in CB regulation. The whole super pathway of cholesterol biosynthesis consists of 24 genes coding for enzymes necessary to build cholesterol molecules from acetyl-derived carbons. Our group previously showed that, at a transcriptional level, the CB super pathway, with particular regard to the rate-limiting enzymes, is upregulated in MCF7-LTED cells (22/24 genes) as well as in T47D-LTED cells (21/24 genes) compared to their respective parental cells [53]. This activation may be specific for long-term hormone deprived cell lines since the invasive drug-resistant cell lines MCF7T (MCF7 Tamoxifen-resistant) and MCF7F (MCF7 Fulvestrant-resistant) upregulate only 2/24 and 3/24 genes respectively, of which none transcribing for rate limiting enzymes[53]. Furthermore, the non-tumorigenic MCF10A cells have much lower mRNA levels for all CB genes compared with non-invasive T47D or MCF suggesting that lipogenesis may be activated during AI-resistance development[53].

Using single cell RNA-seq (scRNA-seq), a recently published study from our lab demonstrated that cholesterol biosynthesis pathway is profoundly reprogrammed by ET-resistant cells (LTED)[292]. To extend the characterization of the altered SREBP1 transcriptional activity in invasive resistant cells versus the parental ones, the RNAscope imaging approach was

exploited. Single-cell RNA Fluorescence in Situ Hybridization (sc-RNA FISH or RNAscope) is an *ISH*-based technique which, in combination with confocal microscopy, can provide a detailed insight about the single RNA molecules spatial distribution at a single cell level. In addition to SREBP1 I also included in the analysis the two rate-limiting enzymes HMGCR and SQLE, and KRT80 (Keratin 80, see chapter 4) as direct target of SREBP1. A first attempt was performed seeding cells at a density of 100,000 cells/slide, but it resulted in an extensive cell overlapping, hampering a clear analysis of the localization of the target mRNA (images not shown). Therefore, after optimisation of the technique, I used a final concentration of 15,000 cells/slide in order to obtain a single molecule-resolution image. Preliminary analysis based on images acquired at a single cell level displays no difference in SREBP1 mRNA expression between MCF7 and LTED cells (Fig. 3.1 C, D). In contrast, a consistent difference can be detected for the other targets SQLE, HMGCR and KRT80 that are highly expressed in LTED (Fig. 3.1 C, D).

At a protein level, immunoblotting showed similar quantity of SREBP1 in both MCF7 and MCF7-LTED (Fig. 1 G), T47D and T47D-LTED (Fig. 3.1 E) whereas increased amount of SQLE was observed in LTED cells (Fig. 3.1 F). Although clear differences were observed when downstream key enzymes were studied, LTED and MCF7 showed similar amount of SREBP1 suggesting that the abundance of SREBP1 transcript *per se* is not likely to be the driver of the increased *de novo* lipogenic activity.

SREBP1 activates cholesterol genes by binding their promoter after nuclear translocation in response to cellular cues[82]. Fatostatins are drugs designed to block SREBP1 nuclear translocation[167]. It has been previously shown by our group that Fatostatin treatment induced a significant reduction for several key genes in the CB pathway including HMGCR and SQLE specifically in LTED cells[53]. Furthermore, treating LTED cells with Fatostatin was sufficient to reduce ERα and block cell invasion[53].

Previous evidence indicate that ET-driven epigenetic reprogramming of breast cancer metabolism leads to the activation of *de novo* CB pathway. In addition, the effects of blocking SREBP1 trafficking to the nucleus on ERα and cell invasion further supports the possibility that SREBP1-driven metabolic chenges may play an important role in AI-resistant breast cancer cells. However, whether SREBP1 activation, rather than the amount of SREBP1 transcript *per se,* may be responsible for the increased *de novo* lipogenic activity is further investigated in the next experiments I performed.

A   SREBP1 RT-qPCR

B

C
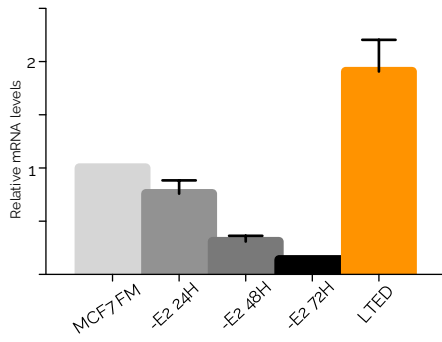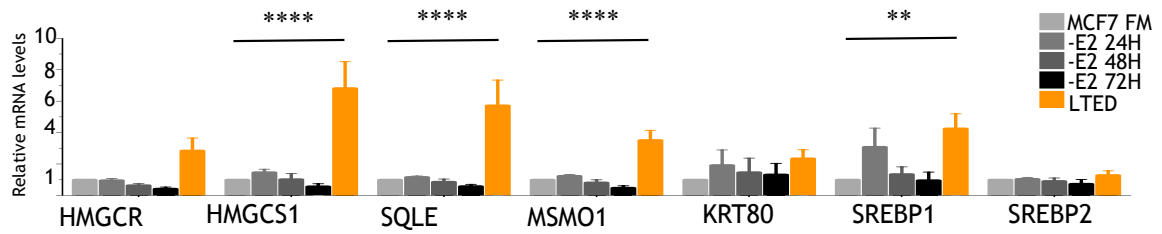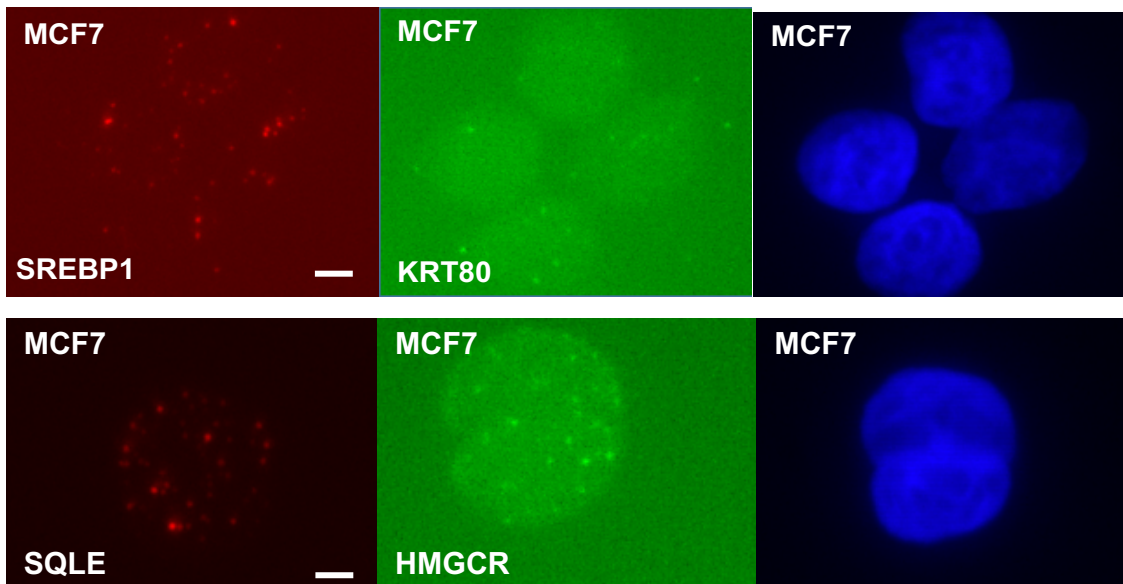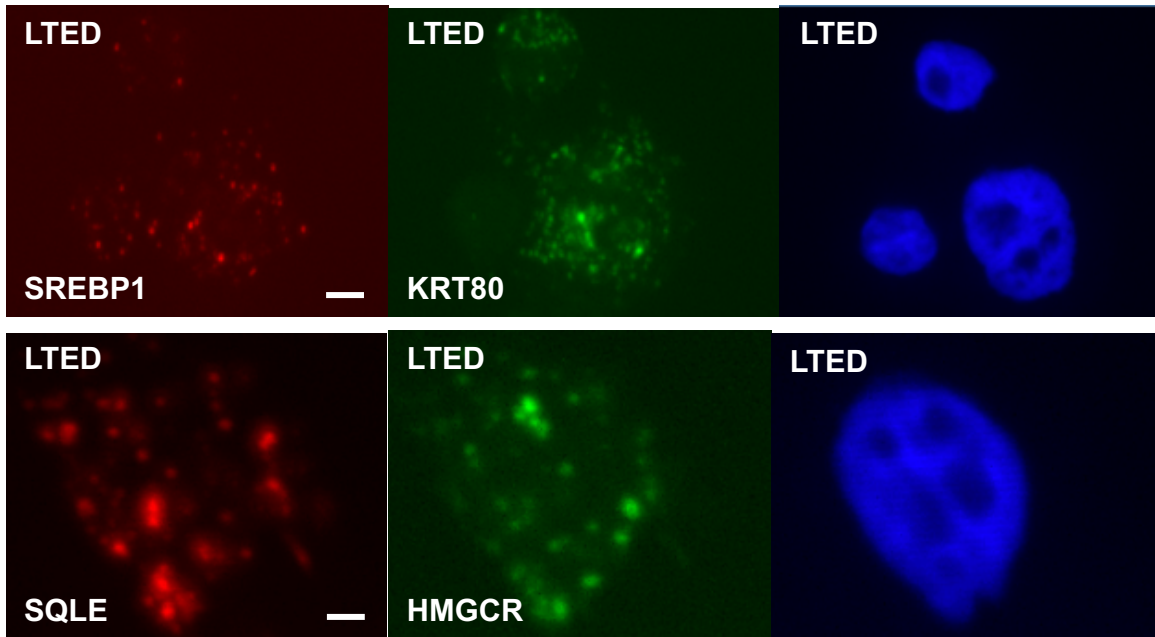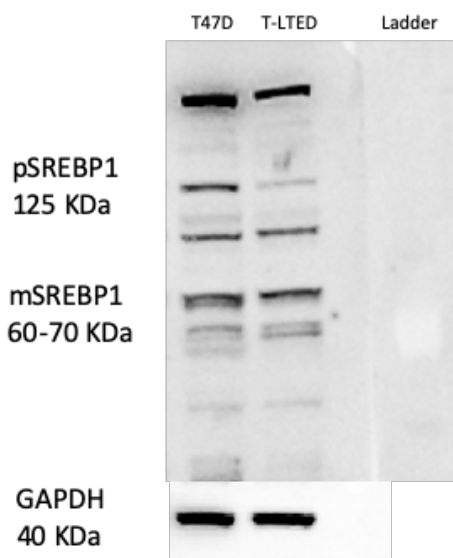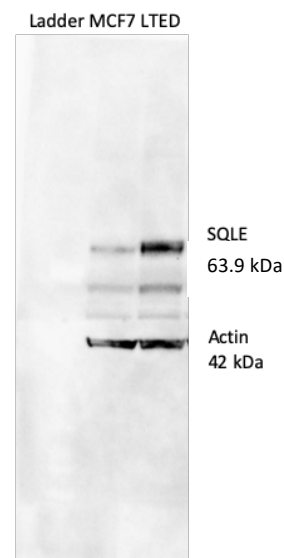
Figure 3.1: SREBP1-driven CB in breast cancer cells. A) RT-qPCR data in MCF7 and LTED showing relative mRNA levels of SREBP1, B) SREBP2 and key enzymes regulated by the transcription factor. The RT-qPCR are representative of six-time course biological replicates. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, not statistically significant if not otherwise specified. Two-way ANOVA and Dunnett's multiple comparison test. MCF7 FM: MCF7 Full Media. –E2 24 H: MCF7 starved for 24 hours. –E2 48 H: MCF7 starved for 48 hours. –E2 72 H: MCF7 starved for 72 hours. C-D) Single cell RNAscope images acquired using confocal microscopy in C) MCF7 and D) LTED. SREBP1 and SQLE in red, KRT80 and HMGCR in green, nuclei in blue. Scale bars, 10 μm. E-G) Immunoblotting for E) SREBP1 in T47D and T-LTED cells, F) SQLE in MCF7 and LTED cells. pSREBP1: premature form of SREBP1, mSREBP1: mature form of SREBP1.

3.2 SREBP1 levels change during ERα BCa progression in clinical samples

To test whether breast cancer progression after therapy could be related to SREBP1 *in vivo*, Immunohistochemistry (IHC) for SREBP1 has been performed by our collaborators at the European Institute of Oncology in Milan, Italy (IEO). Two groups of patients with primary disease were included in the clinical dataset and treatment consisted of either Tamoxifen (Tam) for pre-menopausal women or Aromatase Inhibitors (AI) for post-menopausal patients. Matched tumour biopsies were derived from the same patients before ET (primary group) and at relapse after treatment (mets group). Interestingly, staining score for SREBP1 was significantly higher in metastatic samples of patients treated either with Tam (*p= 0.0032*) or AI (*p= 0.0043*) (Fig. 3.2 A and B respectively) compared to the corresponding matched primaries. When we compared tissues from primary tumours of the two groups of patients before ET, SREBP1 was found significantly increased in the post-menopausal group that was going to receive AI treatment (*p= 0.031*) (Fig. 3.2 C). These data suggest a possible correlation with the development of an AI-driven resistance program (untreated primary vs long-term oestrogen deprived metastasis, Fig. 3.2 A, B) and also possibly with the patients' age hence with the lower circulating oestrogen levels (pre- vs post-menopausal, Fig. 3.2 C). Taken together, these data warrant further investigation about the role of SREBP1 in hormone depleted conditions *in vivo*.
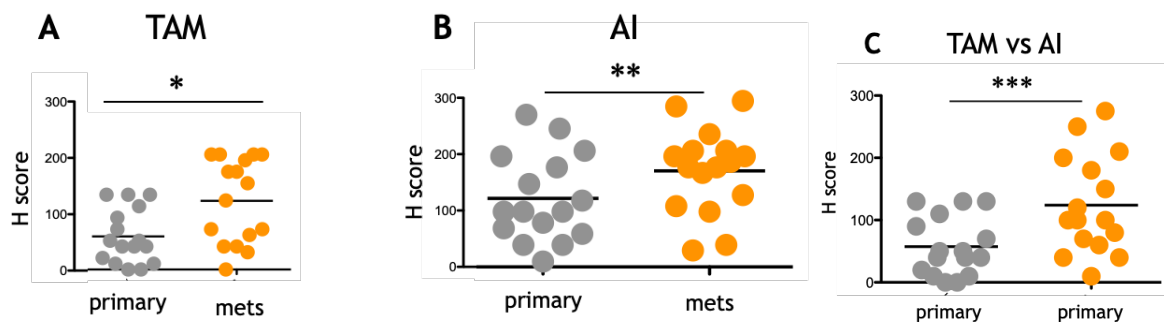
Figure 3.2: SREBP1 is increased in metastatic AI-treated breast tumours. IHC for SREBP1 performed on tumour biopsies from primary and metastatic patients treated with either Tamoxifen (Tam) or Aromatase Inhibitors (AI). Primary samples (grey), metastatic samples (mets, orange). (A) Primary tumours Tam treated vs secondary tumours Tam treated *p= 0.0032, Student T-Test. (B) Primary tumours AI treated vs secondary tumours AI treated ** p= 0.0043, Student T-Test. (C) Primary tumours AI treated vs Primary tumours Tam treated *** p= 0.031, two-way ANOVA. (IHC performed by Giancarlo Pruneri at IEO in Milan, Italy, figure by Luca Magnani).

3.3 Long-term hormone deprivation upregulates SREBP1-dependent cholesterol biosynthesis

To dissect the kinetic recruitment of SREBP1 during hormone starvation and investigate whether SREBP1 is sensitive to acute deprivation or oestradiol stimulation, I performed time course studies. To this end, MCF7 and T47D cells were incubated in stripped media (T0) without any additional hormone stimulation and harvested after 6, 12 and 24 hours. After mRNA extraction, I performed RT-qPCR that showed no significant increase of SREBP1 relative mRNA levels for the described times compared to T0 (Fig. 3.3 A). To further evaluate acute hormone starvation, I prolonged the deprivation to up to 72 hours. To this end, cells were seeded and incubated in either full or stripped media without any additional hormone stimulation and harvested at different time points (24, 48 and 72 hours). RT-qPCR data show relative mRNA levels of SREBP1 and its target genes increased in LTED but not in the MCF7 in either full media or starved conditions (Fig. 3.1 A). The same experimental conditions were used for T47D and T47D-LTED confirming the activation of a SREBP1-dependent lipogenic program after long-term oestrogen deprivation (Fig 3.3 B).

Several studies have previously highlighted the importance of the lipogenic pathway in prostate cancer progression (see Introduction, Paragraph 1.2.4). For this reason, time course studies were performed in LNCaP and in the correspondent long term androgens deprived LNCaP95 cell lines using the same experimental conditions (Fig. 3.3 C). During the first 24 hours of androgen deprivation, SREBP1 and key enzyme mRNA levels are elevated, and they further increase in LNCaP95 (SREBP1 10 to 17-fold, Fig. 3.3 C). Although SREBP2 levels do not raise during the first 24 hours of androgen deprivation, they increase when acute starvation is extended to up to 72 hours, and a further increase is found in the long-term androgen deprived LNCaP95 (Fig. 3.3 C).

It has been previously reported that, in prostate cancer, androgen-induced increase in the expression of SCAP (responsible for ER-to-Golgi SREBP1 transport) levels causes the maturation of SREBP precursor proteins and lead to the coordinated activation of transcription of SREBP target gene[218]. To evaluate whether hormone stimulation may exert the same effect in breast cancer with oestradiol activating SREBP1 through SCAP, I performed additional time course studies. To this end, MCF7 cells were cultured in stripped media for up to 72 hours and then stimulated by adding 17β-oestradiol (E2). After the indicated periods of time (4, 8 and 24 hours) cells were harvested, RNA was isolated and RT-qPCR analysis performed using SCAP,

HMGCR, HMGCS1, SQLE and SREBP1 primers (Fig. 3.3 D). A trend is observed, although not significant, for SREBP1, SCAP and CB enzyme mRNA levels in MCF7. In contrast to the action exerted by androgens in PCa, these data suggest that oestradiol stimulation does not activate CB in BCa cells.

Because the results so far indicate that activation of SREBP1 is also observed in long-term androgen deprived conditions, I next tested for SCAP mRNA levels in LNCaP cells. At a transcriptional level, SCAP expression was found increased in both acute and long-term starvation conditions in LNCaP (2- and 3.5-fold change respectively) although not statistically significant (Fig. 3.3 E).

In summary, RT-qPCR time course studies showed that short-term oestradiol deprivation is not sufficient to induce SREBP1-dependent CB activation that is instead consistently upregulated in resistant long-term hormone deprived cell lines. We can therefore hypothesise that if SREBP1 target activation is enhanced by hormone removal, this does not occur at an early stage during treatment, but it may rather be a cellular program that is selected/elicited after a long time from starting the endocrine treatment. This hypothesis is also in agreement with scRNA-seq analysis in ETR BCa cells[292]. One possibility is that other mechanisms of regulation such as glucose-dependent signalling or likely involving the nuclear receptors (investigated and reported below), rather than the hormones themselves, may be involved in the sterol biosynthesis activation.
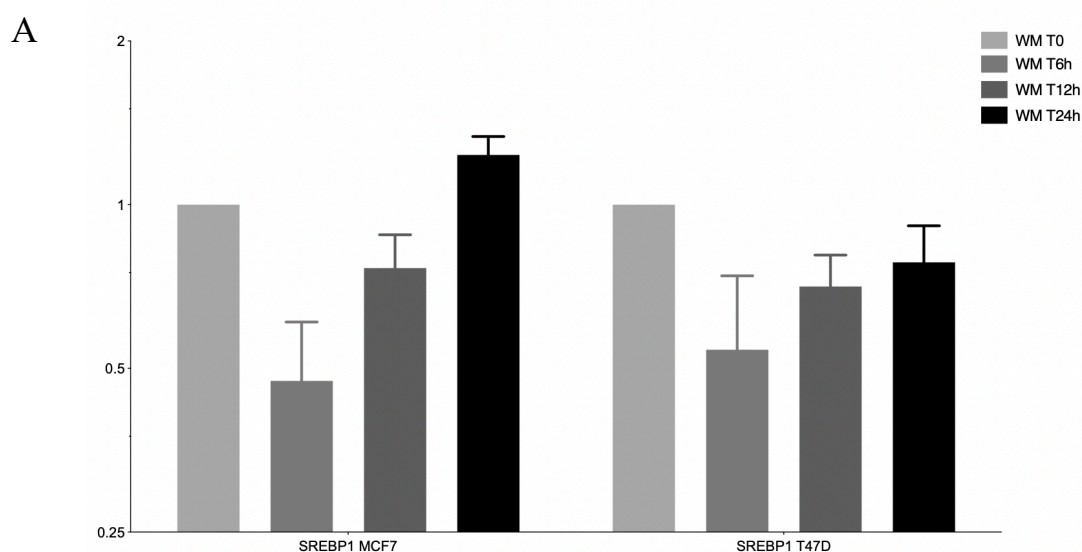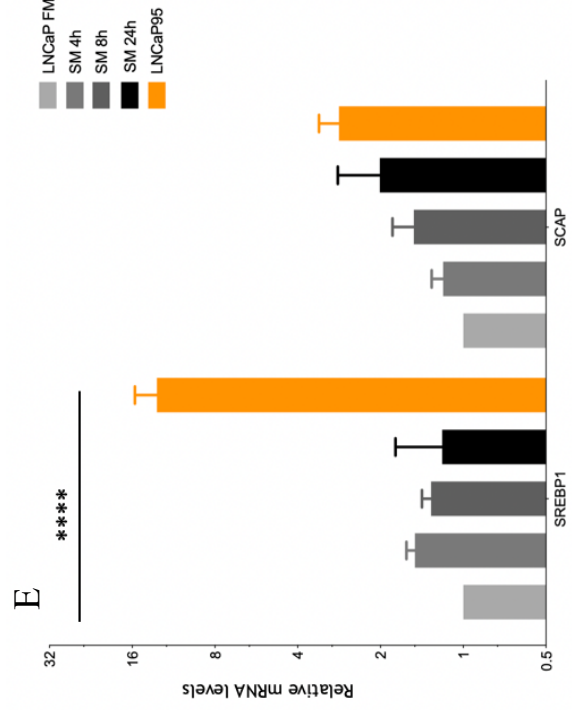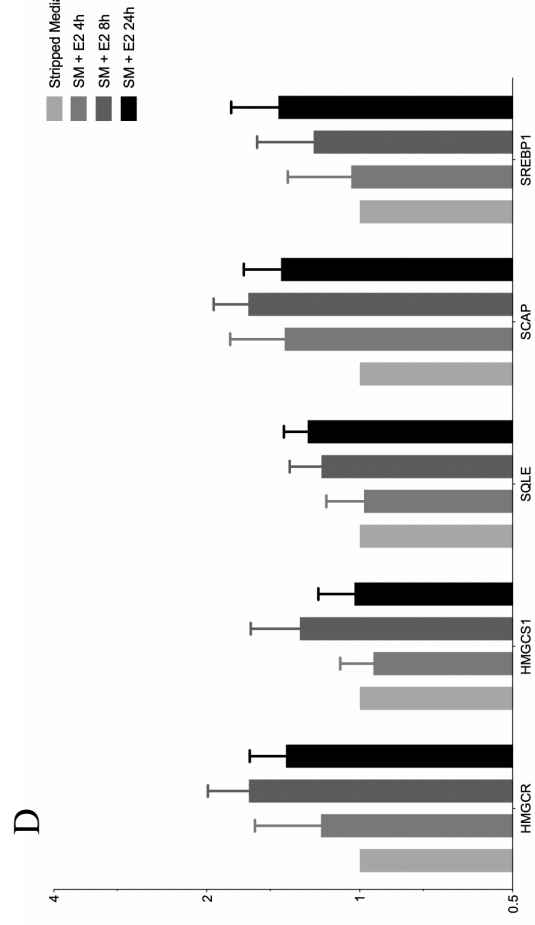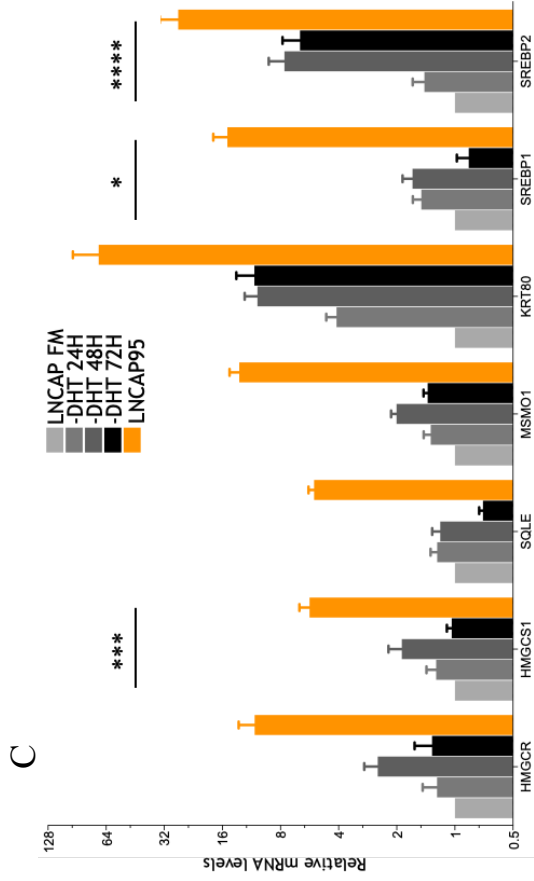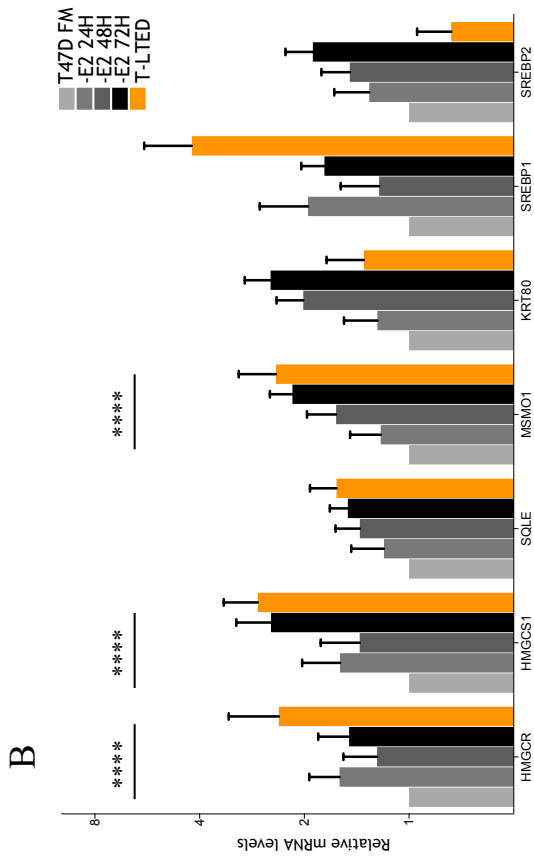
Figure 3.3 (previous pages): RT-qPCR time course studies. A) SREBP1 mRNA levels in MCF7 and T47D cells cultured in starved conditions (WM, white media) for 0, 6, 12 and 24 hours (T0, T6h, T12h, T24h respectively). B-C) RT-qPCR data in B) T47D (grey), T-LTED (orange) and C) LNCaP (grey), LNCaP95 (orange) showing mRNA levels of SREBP1, SREBP2 and key enzymes. Cells were cultured in either full media (FM), or starved conditions for 24, 48 and 72 hours. D) RT-qPCR data in MCF7 cells cultured in stripped media and then stimulated by adding E2 for 4, 8 and 24 hours (SM + E2 4h, SM + E2 8h and SM + E2 24h respectively). E) RT-qPCR data in LNCaP cultured in full media (FM, grey), or stripped media for 4, 8 and 24 hours (SM 4h, SM 8h and SM 24h respectively) and LNCaP95 (orange). The RT-qPCR are representative of six-time course biological replicates. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$, not statistically significant if not otherwise specified. Two-way ANOVA and Dunnett's multiple comparison test. T47D FM: T47D Full Media. –E2 24H: T47D starved for 24 hours. –E2 48H: T47D starved for 48 hours. –E2 72H: T47D starved for 72 hours. LNCaP FM: LNCaP Full Media. –DHT 24H: LNCaP starved for 24 hours. – DHT 48H: LNCaP starved for 48 hours. – DHT 72H: LNCaP starved for 72 hours.

3.4 Switch of metabolic dependency in resistant breast cancer cells

At a physiologic level, SREBP signalling is activated by oxysterols, glucose and the insulin pathway (see 1.2.2). In some cancers, like glioblastoma, it has been reported that glucose dependent SCAP glycosylation is required for SREBP activation[293]. However, it is also known that SREBP1-dependent *de novo* lipid biosynthesis can be activated by increased levels of glutamine[294]. Therefore, we next asked the question whether changes in metabolic dependencies might determine the different SREBP1 activation in starved cells compared to the parental ones.

In order to test whether glucose is driving SREBP1 binding, I performed ChIP-qPCR in different culture conditions. MCF7 cells were cultured in phenol-red free DMEM (4.5 g/L D-Glucose) and supplemented with 10% fetal calve serum (FCS) and E2 (MCF7 FCSE2). At the same time, MCF7 cultured in full medium were used as a control (MCF7 RM: DMEM 1 g/L Glucose with 10% FCS and E2). Then I performed SREBP1 ChIP-qPCR in cells cultured in both conditions (biological duplicates). ChIP-qPCR primers MSM01 and OLMALINC were used as positive, and ML13 as negative controls to check immunoprecipitated DNA versus Input and compared against negative controls. No amplification was detected for either SREBP1 or SQLE primers and no significant changes were observed for MSM01 and OLMALINC levels in MCF7 FCSE2 compared to MCF7 RM, suggesting that, in this particular model, the activation of SREBP1 is not dependent on glucose supply (Fig. 3.4 A).

In order to further test metabolic dependencies of breast cancer cells, metabolic profiling of

MCF7 and LTED cells was performed by our collaborators at the SYSBIO Centre of Systems Biology in Milan, Italy (unpublished data from Rohit Bharat).

Acute starvation resulted in divergence of metabolic propensity from glucose towards glutamine (data not shown). This metabolic switch from a highly glycolytic phenotype towards a glucose independent phenotype was also confirmed in long-term hormone deprivation (LTED versus MCF7) indicating a nutrient dependence shift upon resistance development. This was observed by the decreased glucose consumption and a concomitant increase in glutamine uptake in long-term hormone starved cells (Fig. 3.4 B, C). Furthermore, measuring the ratio of lactate secretion over glucose consumption showed an increased lactate production in LTEDs. One mole of glucose (6 Carbon molecule) produces two moles of Lactate (3 Carbon molecule), therefore a ratio of 2 is indicative of a glycolytic phenotype. Analysis in MCF7 cells showed a ratio of 2 indicating that all/most of the glucose was fermented towards lactate production and not much for mitochondrial respiration. An increased ratio (>2) suggested that the produced lactate is not only coming via glucose but also through other sources like glutamine or via uptake of pyruvate and conversion to lactate (Fig. 3.4 D).

Comparison of total cellular Acetyl-CoA levels were performed in all cell lines in normal growth condition as well as in nutrient perturbed conditions (low Glucose and low Glutamine) (Fig. 3.4 E). Acetyl-CoA serves as the fundamental molecule for key metabolic pathways like lipid synthesis, cholesterol synthesis, acetylation and TCA cycle. Higher Acetyl-CoA levels were found in LTED when compared to other cells. In these cells glutamine serves as the predominant source for Acetyl-CoA, as perturbation using low levels of glutamine resulted in significant decrease in Acetyl-CoA levels in LTED. Not much difference was observed when cells were subjected to low glucose levels further suggesting that their metabolism is not dependent on glucose (Fig. 3.4 E).

Mass spectrometric quantitation revealed not only increased cholesterol levels in LTEDs, but also increased palmitate levels suggesting an increase in fatty acid synthesis in LTED compared to non-resistant cells (Fig. 3.4 F, G). ^{13}C Isotope resolved fluxomics showed an increased mole percent enrichment of labelled isotopes in palmitate obtained via [U-^{13}C$_6$]-Glutamine, proving the increased utilization of glutamine towards lipid synthesis via reductive carboxylation in LTED cells (Fig. 3.4 H).

Cholesterol and fatty acids serve as key constituents of cellular membranes including that of mitochondria. Furthermore, oxidation of glutamine is a phenomenon occurring in

mitochondria. In order to meet the energy demands, high dependence on glutamine would require cells to have well-functioning mitochondria. For these reasons, mitochondrial parameters like respiration and morphology were further investigated. Confocal imaging showed that LTED cells were characterised by an increase in mitochondrial number and network and also by increased mitochondrial respiration mediated via glutamine (Fig. 3.4 I, L).

Taken together, these results suggest that the activation of the SREBP1 pathway in our model may not be dependent on glucose supply. Indeed, we observed a switch of metabolic dependency in drug-resistant ERα breast cancer cells from a glucose- to a glutamine-dependent phenotype and this was found to be associated with increased *de novo* cholesterol and fatty acid synthesis. Whether and how SREBP1 activation may be directly driven by a cellular switch of metabolic dependency needs to be worked out mechanistically.

D

E

Acetyl-CoA

F

cholesterol

Original Conc.

Normalized Conc.

G

Palmitate

- MCF-7
- MCF-T
- MCF-F
- LTED

H

MPE PALMITATE
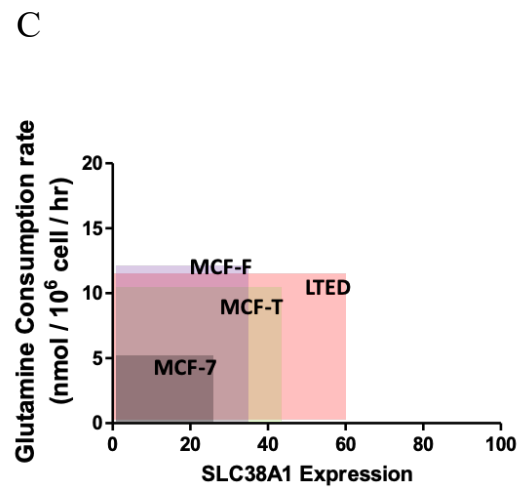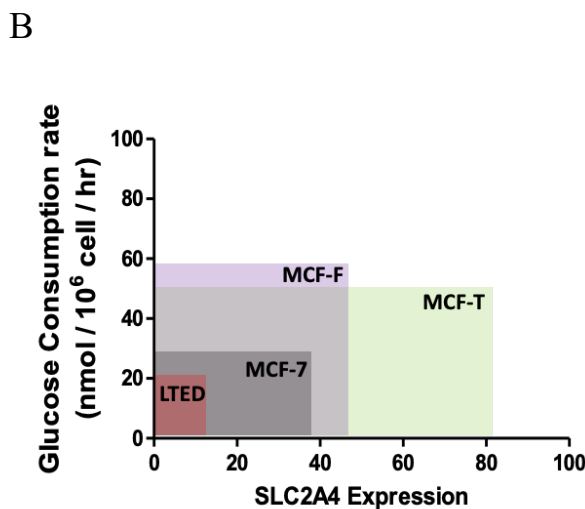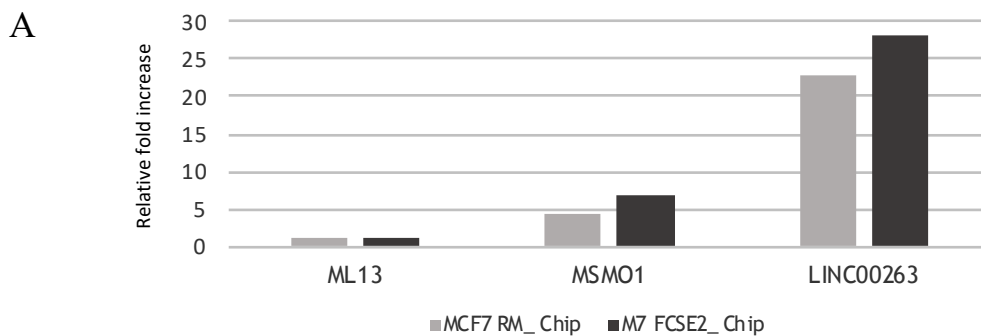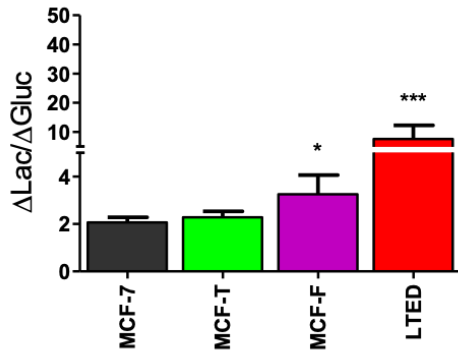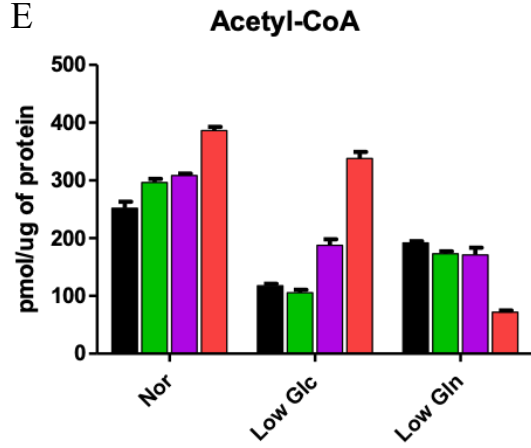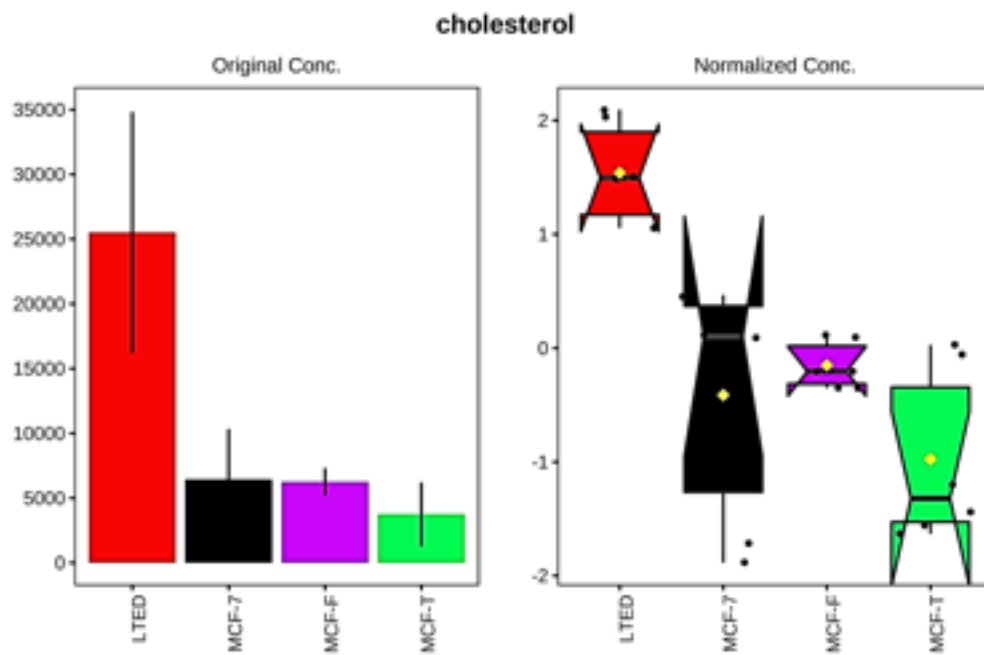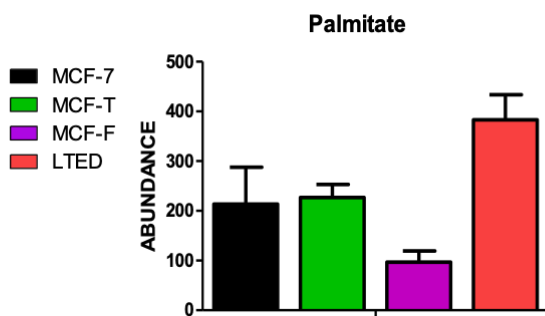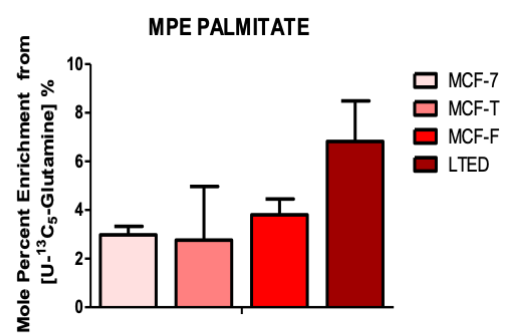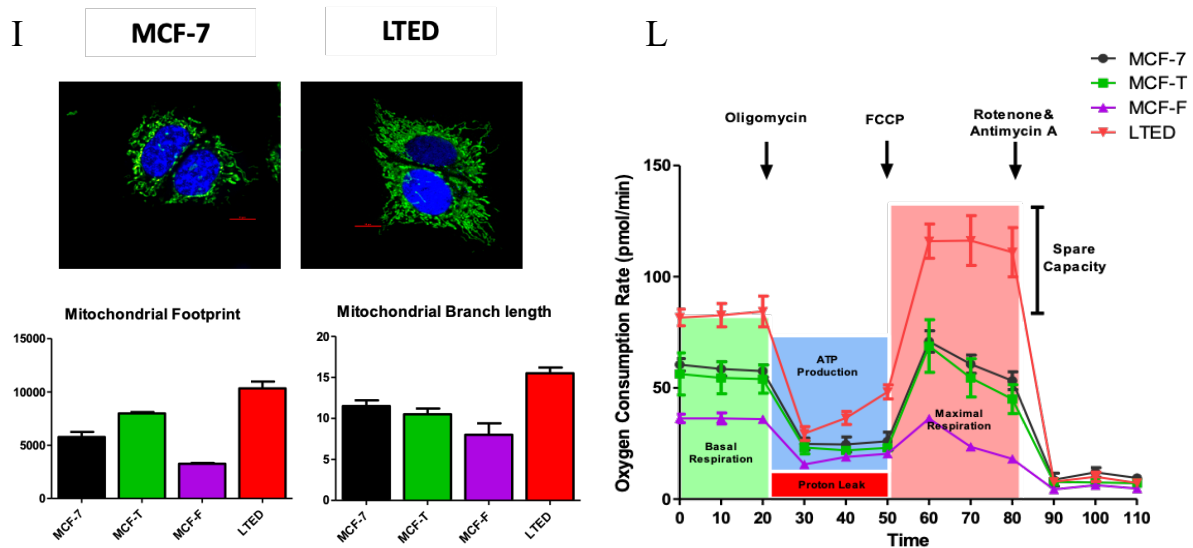
- MCF-7
- MCF-T
- MCF-F
- LTED

Figure 3.4: Metabolic profiling of breast cancer cells. Previous pages: A) SREBP1 ChIP-qPCR in MCF7 cells cultured in full medium (MCF7 FM) and in phenol-red free DMEM, supplemented with 10% fetal calve serum (FCS) and E2 (MCF7 FCSE2). Y axis: Relative fold increase compared to ML13, negative control. B) Comparison of Glucose consumption (y-axis) measured using YSI 2950 Biochemistry Analyser and mRNA expression levels of Glucose transporter SLC2A4/GLUT4 (x-axis, data obtained from RNA-seq data[53]) in MCF7, MCF7-T, MCF7-F and LTED. C) Comparison of Glutamine consumption (y-axis) measured using YSI 2950 Biochemistry Analyser and mRNA expression levels of Glutamine transporter SLC38A1/SNAT1 (x-axis, data obtained from RNA-seq data[53]) in MCF7, MCF7-T, MCF7-F and LTED. D) Ratio of Lactate secretion over Glucose consumption measured using YSI 2950 Biochemistry analyser. E) Comparison of total cellular Acetyl-CoA levels measured using PicoProbe™Acetyl-CoA Fluorometric Assay Kit (BioVision, Milpitas, CA) in normal growth condition as well as nutrient perturbed conditions (low Glucose and low Glutamine). F) Relative quantification of cellular cholesterol levels in ET resistant and sensitive cells using Gas chromatography–mass spectrometry (GC-MS). G) Quantification of cellular Palmitate level in ET resistant and sensitive cells using GC-MS. H) Mole percent enrichment of palmitate using 13-C labelled glutamine. I) Confocal micrographs of mitochondria stained with MitoTracker Green (Green) and nuclei stained with Hoechst 33342 (Blue) in MCF-7 and LTED along with computational quantification of mitochondrial footprint and mitochondrial network using MiNA. L) Differences in mitochondrial respiration depicted by oxygen consumption rates (OCR) under basal condition and upon subsequent treatment with drugs altering the mitochondrial activity namely Oligomycin (1µM), Uncoupler FCCP (0.5µM) or electron transport inhibitor Rotenon &Antimycin A (2µM). The Measurements were taken in normal conditions to measure Basal Respiration, after addition of Oligomycin to measure respiration used for ATP production, after addition of FCCP to measure maximal respiratory capacity and after rotenone/Antimycin (that shuts down the mitochondria) to measure non mitochondrial respiration. Data are normalized to protein content of the cells. MCF7-T: MCF7 Tamoxifen-resistant, MCF7-F: MCF7 Fulvestrant-resistant.

117

3.5 Suppression of SREBP1 expression blocks cell proliferation in starved conditions

To investigate the effects of SREBP1 on cell proliferation and survival, especially in hormone-starved conditions (mimicking AI treatment), I have generated three individual SREBP1 *null* clones by using clustered regularly interspaced short palindromic repeats / CRISPR-associated protein 9 (CRISPR/Cas9) technology. Different clones were isolated after FACS sorting, validated with Sanger sequencing and tested separately (Fig. S1 A, B and 2.2.9-2.2.14). Figures S1 D and E respectively show immunoblotting analysis suggesting clone 1a and 1c were indeed SREBP1 *null* whilst the mature form of the protein was still present in clone 1b along with sequencing chromatograms indicating the exact sites of the cuts performed in the isolated CRISPR clones.

During the first week of growth assay, cells proliferation was monitored every 6 hours: parental cells survived and formed discrete colonies, albeit their proliferation was only marginal as expected in starved conditions; after an initial stall, MCF7 CRISPR clone 1a decreased their growth whereas MCF7 CRISPR clone 1c cells were not able to form any colonies and proliferate from time 0 (Fig. 3.5 A, B). An exception was observed for the MCF7 CRISPR clone 1b cells that grew very fast since the beginning. Continuous monitoring for up to 39 days indicates total absence of proliferation for MCF7 CRISPR clone 1a and 1c in contrast to a constant growth of the clone 1b. To rule out the possibility of a technical bias, I performed a biological replicate of the growth assay for all the CRISPR clones and parental MCF7 cells (Fig. 3.5 C, D). Analysis confirmed previous results for all cell CRISPR edited lines. Indeed, after 28 days clone 1b cells reached confluency and were therefore split (blue line drop) and kept in Incucyte for continuous imaging monitoring up to 47 days (Fig. 3.5 C, D). More data need to be collected to correctly interpret the clone 1b abnormal proliferation. Further characterisation of all silenced cell lines is also needed in order to investigate the phenotypical changes due to SREBP1 deletion. However, both clone 1a and clone 1c behave in a similar manner blocking cell proliferation during steroid starvation, further strengthening the idea that SREBP1 may play an important role in cancer cells response to stress induced by hormone deprivation conditions. Further experiments are needed in order to investigate the effects of SREBP1 suppression on cell cycle, apoptosis and Ki-67 in cancer cells.

By using a vast array of techniques, SREBP1-driven cellular metabolic changes have been investigated *in vitro* and *in vivo* suggesting a possible role in the development of drug-resistance. Evidence so far indicates that, during hormone deprivation, SREBP1 transcript and

protein levels do not change suggesting that cholesterol biosynthesis may be upregulated by modulating SREBP1 signalling. Therefore, in order to investigate SREBP1 recruitment to the chromatin, I next performed Chromatin Immunoprecipitation coupled with high-throughput sequencing (ChIP-Seq) using the antibody anti-SREBP1 in different cell lines.

A

B
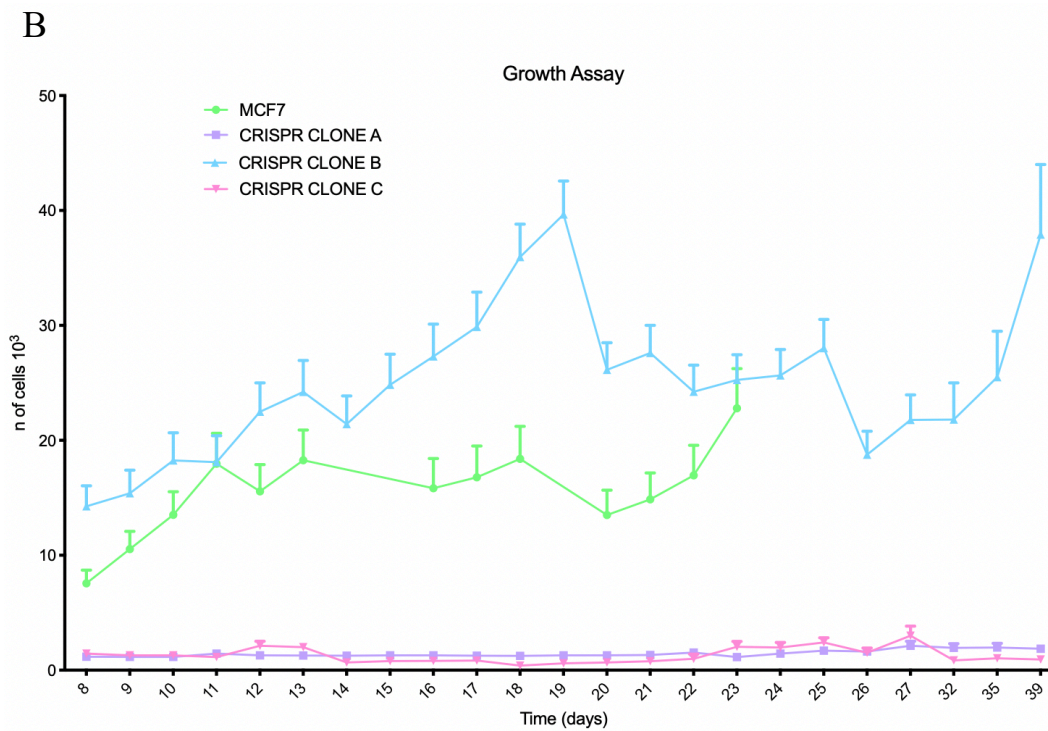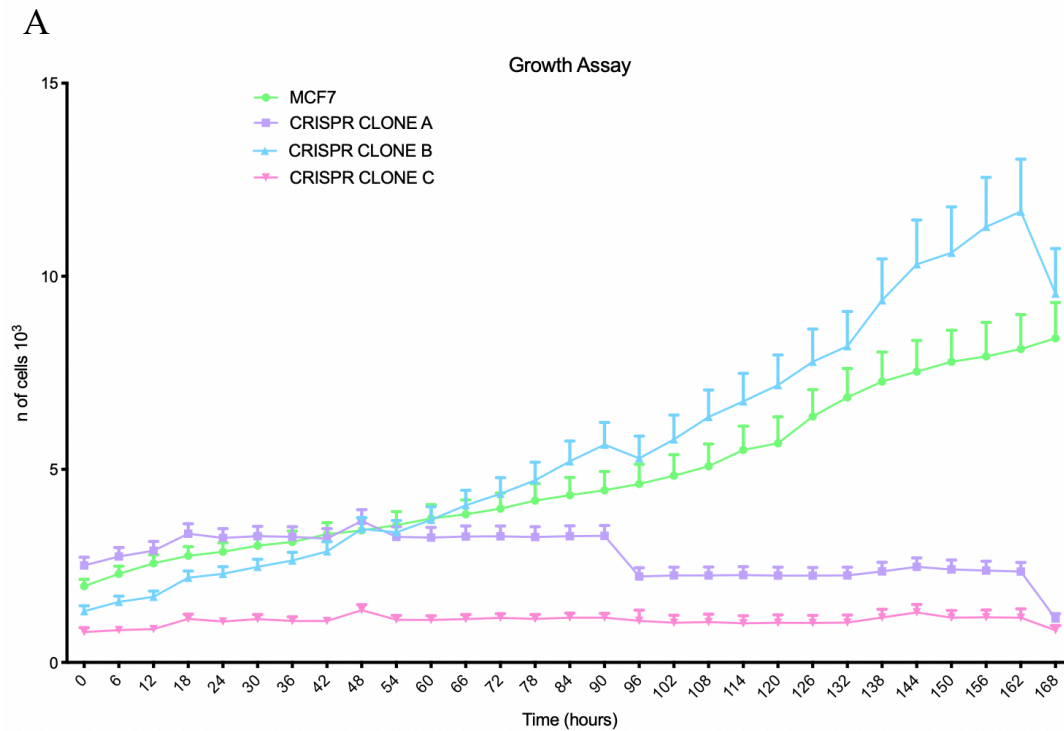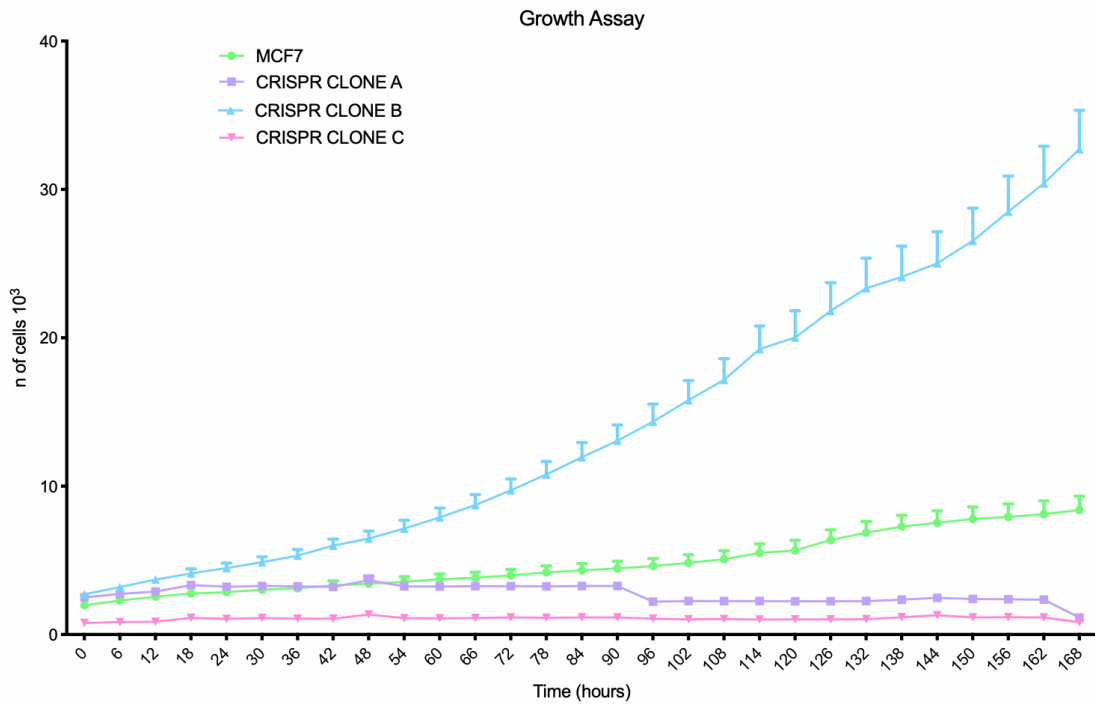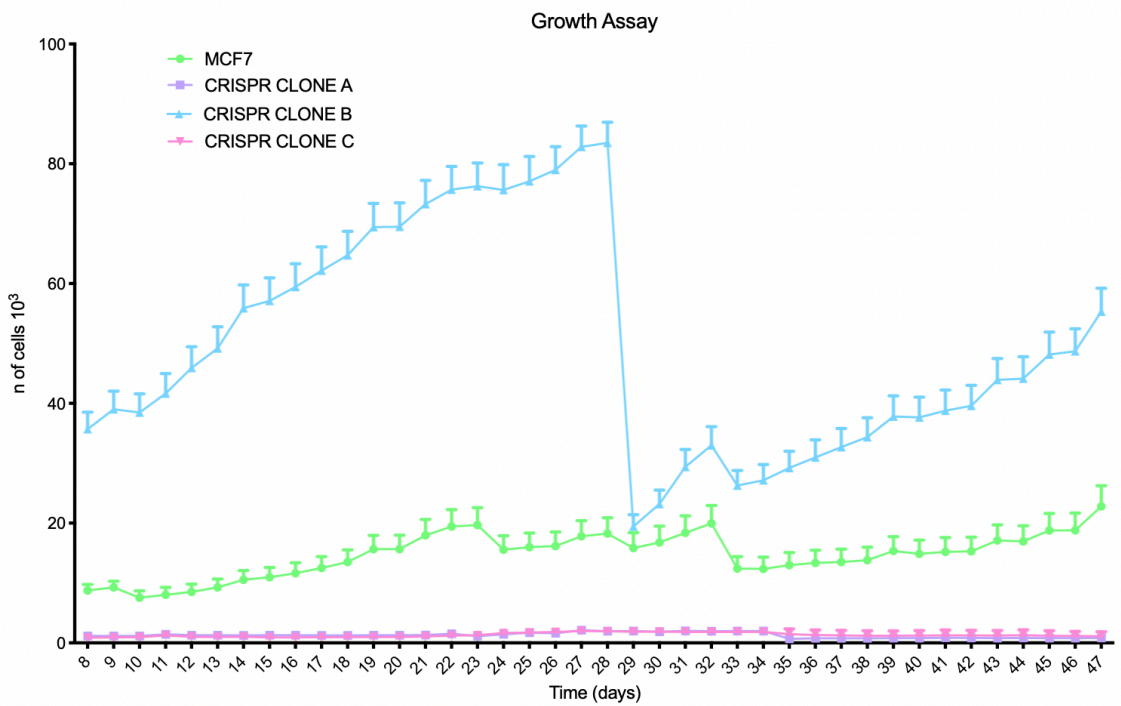
Figure 3.5: Cell proliferation of SREBP1 silenced cells in starvation conditions. A-D) Parental MCF7 and SREBP1 CRISPR clones cell growth assay performed by continuous imaging monitoring with Incucyte: images acquired every 6 hours for the first week, A); every 24 hours from day 8 to 39, B); C-D) biological replicate with images acquired every 6 hours for the first week C), and every 24 hours from day 8 to 47, D).

3.6 Chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) against SREBP1

3.6.1 Overview of ChIP-seq

In order to understand transcriptional regulation, genome-wide mapping of protein-DNA interactions is essential. Since the crosstalk between chromatin and transcription is a dynamic process, mapping of binding sites for TFs is vital for deciphering the gene regulatory networks that underpin and control several biological processes[295]. Hence ChIP has become a very important tool to investigate the mechanisms of transcriptional regulation and involves assaying protein-DNA binding *in vitro* and *in vivo*[296]. Laboratory techniques mapping TF occupancy genome-wide by using ChIP were developed almost twenty years ago[296,297]. ChIP assays involved targeting a protein contacting chromatin (e.g. a TF or a cofactor) by immunoprecipitation from cross-linked cells, to be compared to their total DNA[298]. The enriched genomic sites were then identified by DNA hybridisation to microarray (ChIP-on-chip)[296,297] which then progressed to ChIP coupled with next generation sequencing (NGS), commonly known as ChIP-Seq[299,300]. ChIP-Seq enables sequencing of millions of small DNA fragments in a single run, which allows for large-scale experiments to be carried out[295].

3.6.2 Optimisation of ChIP against SREBP1

The ChIP protocol has been detailed by *Schmidt et al.*[301] and can also be found in Chapter 2 (2.2.24-2.2.26). The procedure has been further improved and thoroughly used in our lab during the past years[302]. I have used the same protocol to perform ChIP using antibodies against histone post-translational modifications (or marks; H3K27ac, H3K4me1, H3K4me2) in MCF7 and LTED cell lines[329] (see 4.2). Despite previous extensive validation, when performing SREBP1 ChIP very low quantity of DNA was pulled down for ChIP samples and qPCR showed no enrichment when comparing DNA from ChIP to Input samples. Therefore, I further optimised several steps of the protocol in order to pull enough DNA to allow a proper analysis of the results through ChIP-qPCR and following library preparation for sequencing.

Schematically, the following steps were optimised:

1. A concentrated version of the SREBP1 (H-160) sc-8984 antibody, the SREBP1 (H-160) X Santa-Cruz was used.

2. In order to increase the pulled chromatin sample:

   - number of cells per condition was increased from 1 x 10^7 to 5 x 10^7 cells,
   - increased concentration of the antibody from 4 ug up to 10 ug,
   - the relative absorbance of antibody on beads accordingly adjusted increasing Dynabeads concentration from 20 uL to 100 uL.

3. Sonication has been previously optimised for 1 x 10^7 cells. Therefore, in order to do not affect sonication efficacy, a total of 5 dishes each containing 1 x 10^7 cells per condition were processed as they were different samples (i.e. 5 dishes corresponding to 5 different microcentrifuge tubes).

4. DNA samples were then eluted by doubling the concentration of buffers used:

   - Lysing Buffer LB3 from 800 ul to 1600 ul
   - 10% triton accordingly increased from 90 uL to 170 uL (1600uL LB3+100uL beads= 1700 uL, 10% triton= 170uL).

5. Day 2 was performed decreasing the number of washes with RIPA buffer from 6 to 3 to avoid losing DNA material.

6. In order to extract the processed DNA material more efficiently, equal volumes of SPRI beads were added to each sample. Solid-phase reversible immobilization (SPRI) paramagnetic bead technology is used to selectively bind DNA fragments obtained after sonication to paramagnetic beads. Isolated DNA samples from Input and ChIP was then quantified and processed for downstream applications (ChIP qPCR, library preparation and sequencing).

3.7 Genome-wide profiling of SREBP1 binding in hormone-dependent cancer cells

3.7.1 SREBP1 chromatin immunoprecipitation in breast and prostate cancer cell lines

The initial analysis was aimed at gaining further understanding about the role of SREBP1 in the ERα-positive BCa regulatory network. Firstly, I conducted SREBP1ChIP-seq on MCF7, T47D, ZR-75-1 and derived LTED cell lines. ZR-75-1 and T47D cell lines, like the MCF7, are luminal breast cancer cells derived from metastasis of an invasive ductal carcinoma. Unlike MCF7, T47D are mutated for TP53[304]. Secondly, to test for the dynamic recruitment of

SREBP1, I added two more experimental conditions: WM, WM+E2 (Fig. 3.6A). I used the same experimental design to expand the cistrome in prostate cancer cell lines by using LNCaP and the derived long-term androgen deprived LNCaP95. Lastly, sequencing data were further analysed for transcription binding (TF) sites enrichment analysis and identification of differential binding sites.

The MCF7 and T47D SREBP1 ChIP (all experimental conditions) were performed in biological triplicates, with two out of three replicates processed for sequencing. The third replicate also passed ChIP-qPCR validation but was not taken to the library stage. The LNCaP SREBP1 ChIP (all experimental conditions) were performed and sequenced in biological duplicates. Since no replicates were performed for SREBP1 ChIP in ZR-75-1 (all experimental conditions) these samples were excluded from downstream analysis.

The bioinformatic ChIP-seq analysis pipeline is summarized in Fig. 3.6, 3.7. It included: quality and adapter trimming, alignment of the raw reads to the GRCh38 Human genome[305] using Bowtie2 (unique reads), peaks called using MACS v1.4 (de-duplicated reads) and filtering against ENCODE blacklists (see 2.2.26). Traditional analyses of ChIP-seq data involve identifying peaks of high read density in the genome. These peaks represent putative binding sites for the target protein. Peak calling refers to a computational method whereby regions in the showing enrichment of sequenced reads given some expectation, that usually accounts for local biases inferred from the input DNA. MACS had become a *de facto* standard to identify enriched genomic loci in binding profiles of transcription factors.

Further statistical analyses were performed to take into account information from biological replicates for each cell line and all the experimental conditions. First of all, peaks were filtered using stringent parameters (Fig. 3.6 D). Reproducible peaks across replicates were defined according to the following conditions: (1) an overlapping peak was independently called in both replicates; (2) an average enrichment *vs* input of at least 2-fold in both replicates; (3) an RPKM (Reads Per Kilobase per Million sequenced reads) of at least 2, separately for each replicate; (4) an average peak size greater than 0.5 kb but smaller than 2 kb.

The total number of reproducible peaks between replicates varies depending on the cell line used. After filtering and considering all experimental conditions, three master lists were generated according to the cell line (MCF7, T47D, LNCaP) (Fig. 3.7 A-D). The first observation is that the number of peaks called in any condition is very similar between different

breast cancer cell lines with MCF7 showing 229 peaks and T47D 183 peaks (Fig. 3.7 A-D). When comparing to ZR-75-1, the 217 peaks called showed 94.8% concordance with MCF7. Interestingly, there is a significant difference in the number of peaks called between breast and prostate cancer cell lines with LNCaP cells reaching a total of 1,189 SREBP1 bound regions (Fig. 3.7 A-D).

A transcription factor-binding sites enrichment analysis was performed using Pscan-ChIP in order to identify over-represented, known TF-motifs in the sequence of a given set of genomic regions. For every cell line, the most significantly over-represented motifs were those for SREBF1 and SREBF2 (Fig. 3.8 and Table 3.1 with global p-values). This result can be considered a further, positive quality control confirming that the generated profiles represent *bona fide* binding sites for SREBP1, consistently across all the cell lines profiled.

In order to investigate the genomic distribution of SREBP1-bound sites relative to known genes, a TSS annotation analysis was performed. Regions in the three master lists were annotated to the TSS of the nearest RefSeq gene. Regions within 2.5 kb of a TSS were defined as TSS-proximal, those located further as TSS-distal. This analysis highlighted a different distribution of regions contacted in BCa and PCa, with BCa showing more binding to promoter regions than PCa cell lines (Fig 3.9 A; TSS 76% and 62.3% in MCF7 and T47D respectively versus 28.1% in LNCaP). We reasoned that this difference could be artefactual, due to spurious, TSS-distal peaks in PCa that we were unable to filter out (about 1,000 more peaks in PCa compared to BCa). Box plots in Fig. 3.9 B, C, D showing enrichment vs input, are relative to the reproducible peaks in samples where at least 10 peaks were called. In line with this, we found a statistically significant difference in the enrichment between proximal and distal regions for LNCaP cells (*p*-values: 0.022 for RM, 7.9e-0.5 for LNCaP 95; Wilcoxon rank-sum test; Fig. 3.9B) whereas there no difference was found in MCF7 (*p*-values: 0.46, 0.93 for MCF7-LTED; Wilcoxon rank-sum test; Fig. 3.9C) and T47D (*p*-values: 0.86 for RM, 0.97 for T47D- LTED; Wilcoxon rank-sum test; Fig. 3.9D).

In summary, SREBP1 ChIP-seq peaks were filtered using stringent parameters and only reproducible peaks across replicates were considered for downstream analysis. TF motif analysis confirmed that the generated profiles represent *bona fide* binding sites for SREBP1 consistently across all the cell lines profiled. TSS annotation analysis highlighted more binding to promoter regions in BCa cells. The number of peaks called in any condition was found to be very similar across the three breast cancer cell lines. Nevertheless, a profound difference was found in the number of peaks called between breast and prostate cancer cell lines.

A
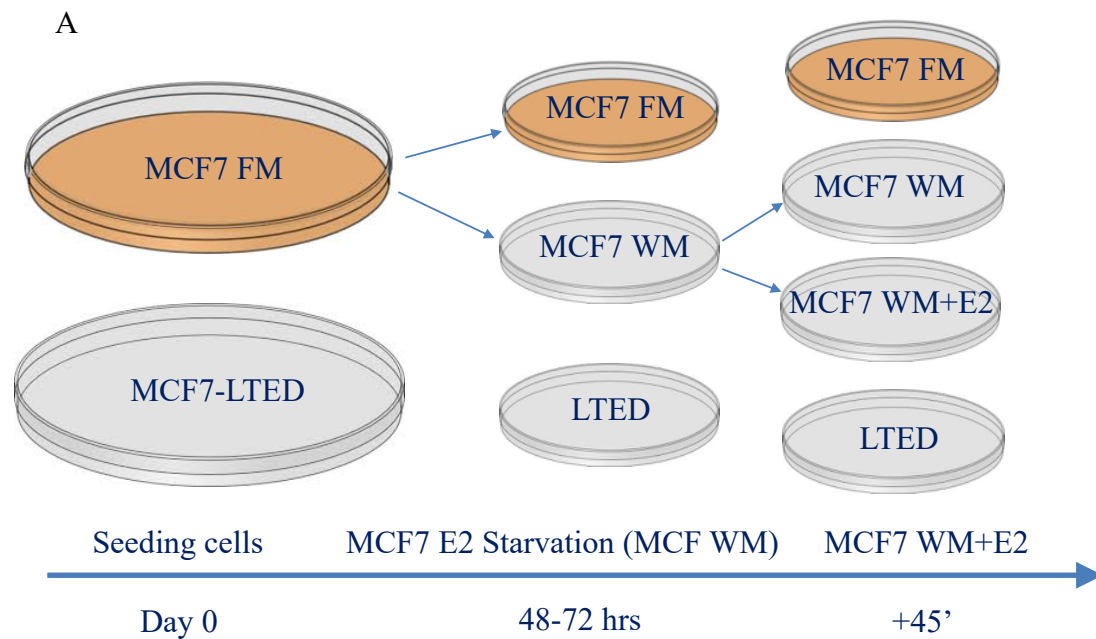
Cell line	Media	Serum	Antibiotics	Drug
MCF-7	DMEM (FM)	10% FBS	1% Glutamine/Pen/Strep	10^{-8} M 17-Beta Estradiol
LTED	DMEM (Phenol Red Free) (WM)	10% Double charcoal stripped FBS	1% Glutamine/Pen/Strep	

B

C
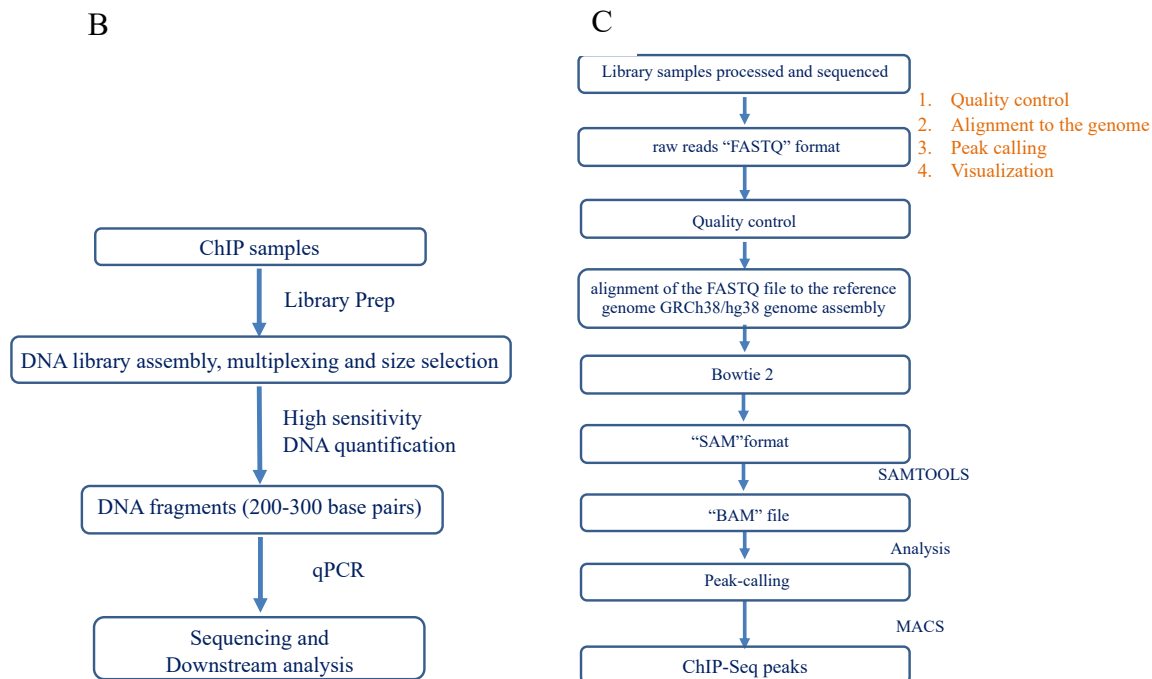
D
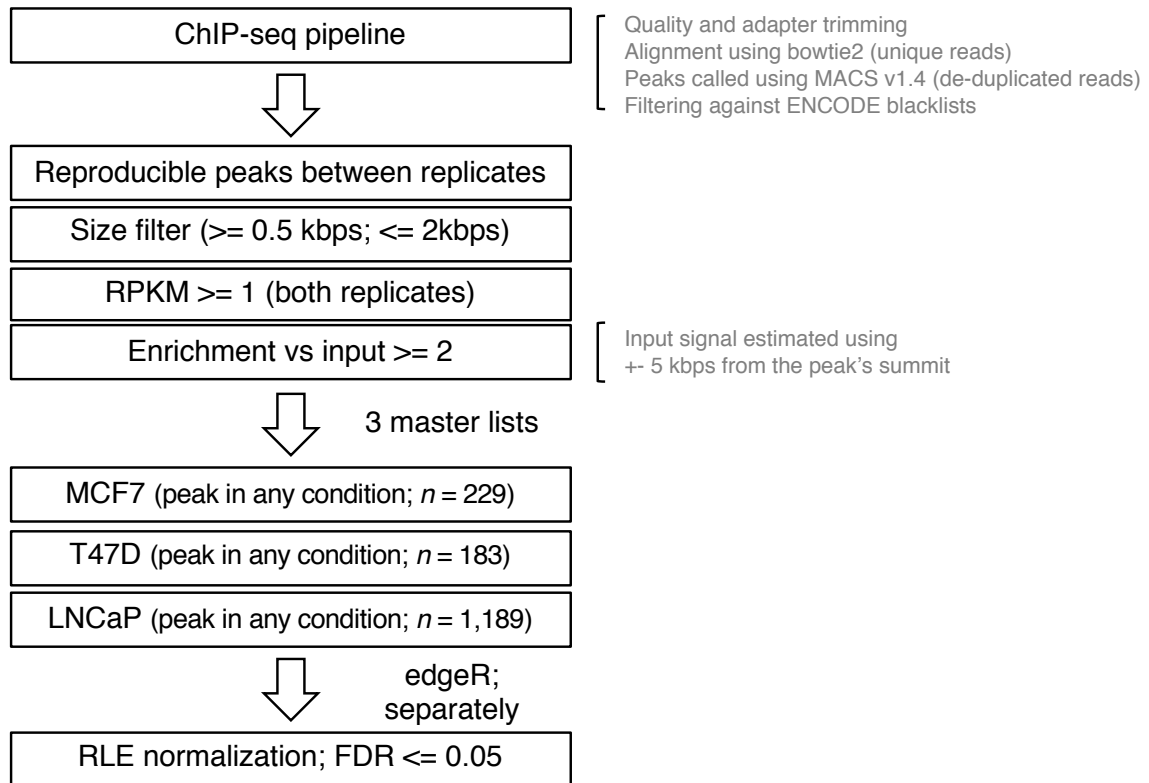
Figure 3.6 (Previous page): ChIP-seq workflow and analysis pipeline. A) ChIP samples preparation. B) ChIP-seq workflow. C-D) Computational ChIP-seq analysis pipeline. FASTQ files were initially processed using fastx_trimmer from FASTX-Toolkit and Trim Galore. Trimmed reads were then aligned to the hg38 reference genome using bowtie2. To discard multi-mapping reads, SAMTools was used. Unfiltered TF-bound locations were identified using MACS and sample-matched input DNA as control. Genome-wide coverage profiles were generated using BEDTools and subsequently normalized to RPM (Reads Per Million sequenced reads). Peaks identified either on the mitochondrial or random chromosomes, as well as those overlapping the blacklist provided by the ENCODE consortium, were excluded. Only peaks found as reproducible in biological replicates and fulfilling stringent criteria on absolute signal and relative enrichment were considered for further analysis. Kbps: Kilobases. RPKM: Reads Per Kilobase per Million sequenced reads, RLE: Relative Log Expression. FDR: false discovery rate. (Analysis performed by Iros Barozzi, and figure D created by Iros Barozzi).
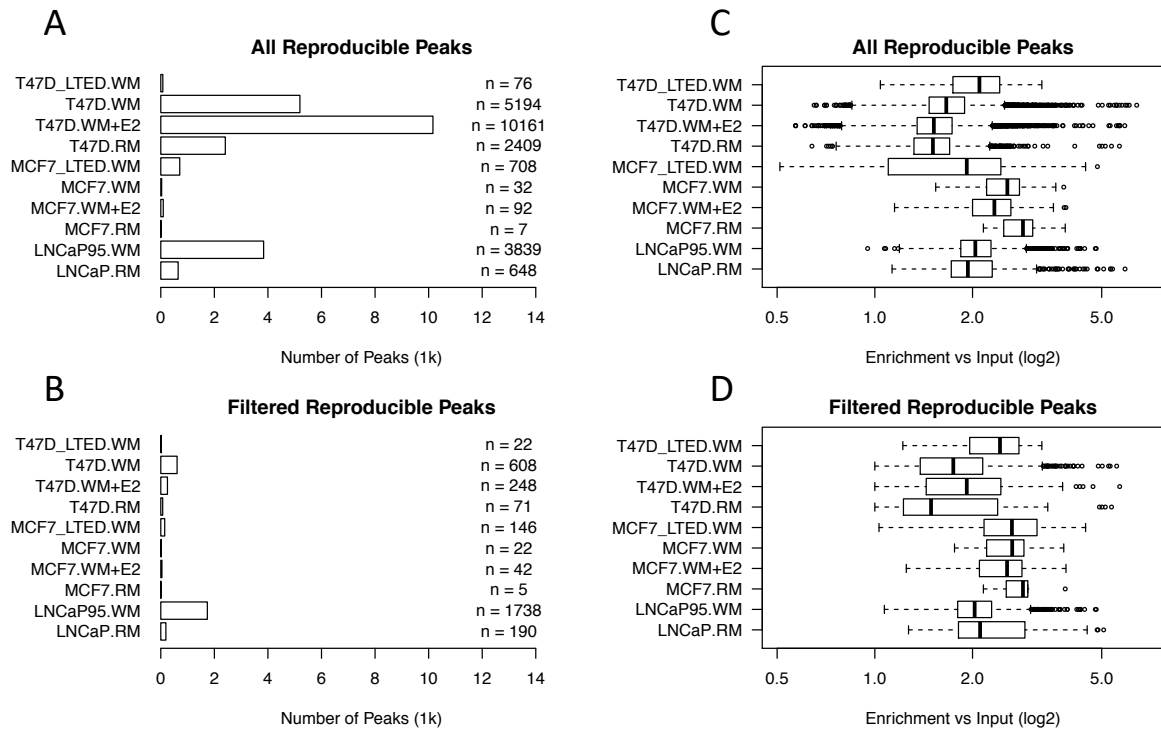
Figure 3.7: SREBP1 reproducible peaks. Bar plots showing the total number of reproducible peaks in the SREBP1 ChIP-seq performed in breast and prostate cancer cell lines. Panels on the left (A, B) show the total number of peaks called before and after stringent filtering. Box plots on the right (C and D) show the distributions of enrichment values for the peaks before and after filtering. (Analysis performed by and figure created by Iros Barozzi).

SREBP1 motif   SREBP1 motif (var.2)   SREBP2 motif   SREBP2 motif (var.2)
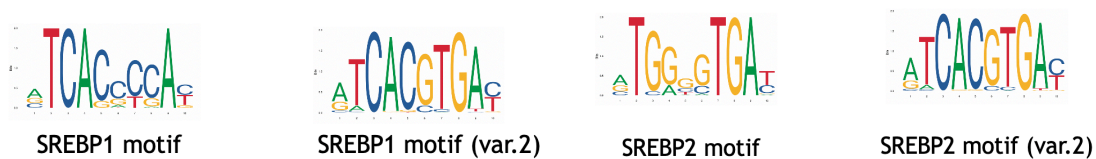
	ID	MCF7	T47D	LNCAP
SREBF1	MA0595.1	1.35E-20	3.70E-09	6.99E-37
SREBF2	MA0596.1	1.10E-18	1.11E-07	2.33E-45
SREBF1 (var.2)	MA0829.1	2.17E-13	0.0065849	9.12E-18
SREBF2 (var.2)	MA0828.1	4.41E-14	0.00221574	1.20E-15

Figure 3.8 and Table 3.1: Transcription factor-binding sites enrichment analysis. For every cell line, the most significantly over-represented motifs identified by using Pscan-ChIP were those for SREBF1 and SREBF2 (var.2: variant 2). Global p-values for each cell line are shown in the table. (Analysis performed by Iros Barozzi).
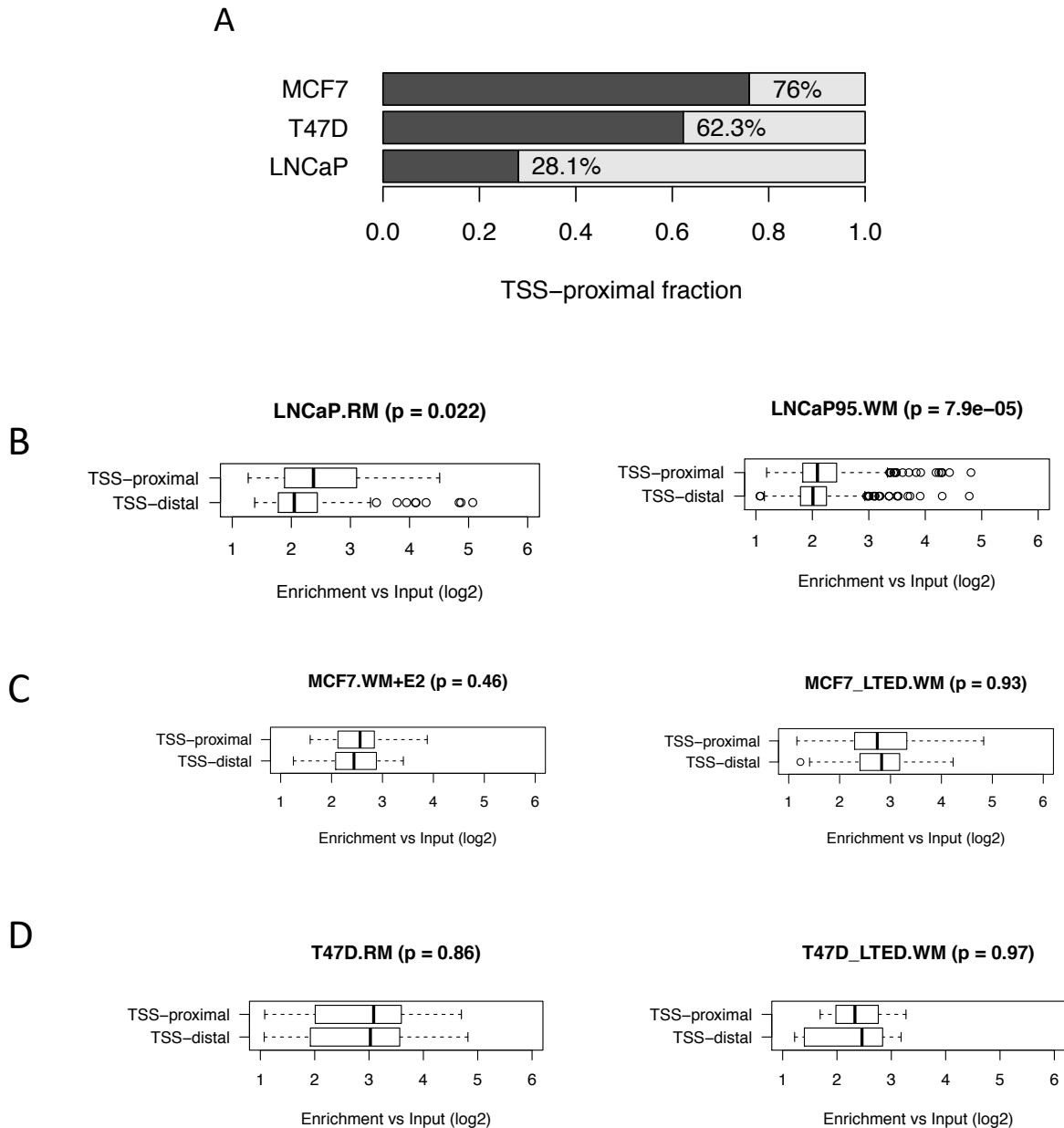
Figure 3.9: Genomic annotation of the SREBP1-bound regions. Regions in the three master lists (one per cell line) were annotated to the TSS of the nearest RefSeq gene. TSS-proximal: regions within 2.5 kb of a TSS, TSS-distal: regions located further than 2.5kb. A) Peaks identified in BCa cell lines are skewed to promoter regions compared to PCa cell lines (76% and 62.3% in MCF7 and T47D respectively versus 28.1% in LNCaP). Box plots in B), C), D) show the enrichment vs input of proximal and distal regions for peaks identified in LNCaP, MCF7 and T47D cells, respectively. P-values computed using the Wilcoxon rank-sum test. (Analysis performed by and figure created by Iros Barozzi).

3.7.2 SREBP1 binding profiles distinguish cancer cells based on the tissue of origin

Results from hierarchical clustering are commonly shown as a tree diagram called dendrogram. The clustering procedure starts with each object in a separate cluster. At each step, the two clusters that are most similar are joined into a single new cluster, until all clusters are merged. The vertical axis of the dendrogram represents the distance (or dissimilarity) between clusters, the horizontal axis represents the objects and clusters. Clusters at one level join with clusters in the next level up, using a degree of similarity. Each joining (fusion) of two clusters is represented on the graph by the splitting of a line into two vertical lines. The vertical position of the split corresponds to a specific distance (dissimilarity) between the two clusters. The hierarchical relationship and overall clustering of the SREBP1 profiles using either MCF7 or LNCaP peaks as a reference are consistent (Fig. 3.10 A, B respectively). Focusing on the structure of the dendrogram, Fig. 3.10 shows how the replicates of the long-term hormone deprived cell lines are most similar in both BCa and PCa. Overall, samples from the same cell line (T47D, LNCaP and MCF7 all conditions) tend to cluster together and apart from those of other cell lines.

The dendrogram in Fig. 3.10 B shows two main clusters (PCa versus BCa cell lines), with the BCa cells splitting in two further branches, one for MCF7 and the other one for T47D cells. In all of the branches, the long-term hormone deprived cells are hierarchically positioned further from the samples from other conditions (Fig. 3.10 A, B).

More generally, both dendrograms support high consistency between biological replicates. This analysis also confirm what previously showed at a transcriptional level, that acute starvation is not sufficient to activate SREBP1 and that stimulating cells with oestradiol after acute hormone deprivation is not sufficient to appreciate a change in the metabolic program since the WM and WM+E2 conditions cluster very closely to the RM for all cell lines (Fig. 3.10).

Fig 3.10 are informative of how SREBP1 binding differs between cancer cell lines: even though the number of genome-wide bound regions is very small (around 200 in MCF7), this is enough to hierarchically organise and divide MCF7 from T47D, and breast from prostate cancer (3 distinct branches of cell lines). Despite a 6-fold increase between peaks called in LNCaP and MCF7 (most of which of low-quality; see previous paragraph), when clustering is done using LNCaP regions (Fig. 3.10 B) the final clustering is very similar to that obtained using the regions derived from peaks in MCF7 cells. The results of the clustering are indeed robust and consistent independently of the chosen set of regions.

Taken together, these analysis suggest that: 1. Biological replicates are consistent for each cell line; 2. There is no clear difference between parental and short-term starvation (RM, WM, WM+E2); 3. Differences are found between parental and long-term starved derived cell lines in MCF7; 4. The largest distance in the hierarchical organisation is found between different cell lines, especially in breast versus prostate cancer, suggesting that SREBP1 binding might be cancer specific, at least to some extent.

Although clustering showed a global trend in the difference between cell lines, it is inconclusive in terms of peaks that are specifically and robustly enriched in one cell line *vs* the others. For this reason, a more formal differential analysis was performed.

Figure 3.10 (following page): Hierarchical clustering of the SREBP1. Dendograms show profiles for breast and prostate cancer cells, using either peaks identified in MCF7 in A) or LNCaP in B) as a reference. Samples were hierarchically clustered based on the normalized expression values (RPKM) across the regions included in each master list. Pairwise distance between samples was measured as one minus the Spearman's rank correlation coefficient between each pair of samples. (Analysis performed by and figure created by Iros Barozzi).

A

Height

1.5
1.0
0.5
0.0

MCF7_LTED.WM.rep2
MCF7_LTED.WM.rep3
MCF7.RM.rep2
MCF7.WM.rep2
MCF7.RM.rep1
MCF7.WM.E2.rep2
MCF7.WM.E2.rep1
MCF7.WM.rep1
T47D_LTED.WM.rep1
T47D_LTED.WM.rep2
T47D.RM.rep2
T47D.WM.E2.rep2
T47D.RM.rep3
T47D.WM.rep1
T47D.WM.E2.rep1
T47D.WM.rep2
LNCaP.RM.rep1
LNCaP.RM.rep2
LNCaP95.WM.rep1
LNCaP95.WM.rep2

B

Height

1.4
1.2
1.0
0.8
0.6
0.4
0.2

LNCaP95.WM.rep1
LNCaP95.WM.rep2
LNCaP.RM.rep1
LNCaP.RM.rep2
T47D.RM.rep3
T47D.WM.rep1
T47D.WM.E2.rep1
T47D.WM.E2.rep2
T47D.RM.rep2
T47D.WM.rep2
T47D_LTED.WM.rep1
T47D_LTED.WM.rep2
MCF7.RM.rep2
MCF7.WM.rep2
MCF7.RM.rep1
MCF7.WM.E2.rep2
MCF7_LTED.WM.rep2
MCF7_LTED.WM.rep3
MCF7.WM.E2.rep1
MCF7.WM.rep1

3.7.3 Differential binding analysis identifies SREBP1 non-canonical target genes in cancer cells

ChIP-seq is widely used to identify binding sites for a target protein in the genome. An important application is to identify changes in local protein binding between different treatment conditions, i.e. to detect differential binding (DB). DB leverages biological replicates and it is able to identify quantitative changes in the binding profile between experimental conditions. Therefore, a differential binding analysis was performed providing for the *de novo* detection of differentially SREBP1-bound genomic regions.

Before differential binding calling, the lists of regions on which to perform the analysis have been further filtered based on copy number alteration (CNA). The aim is to avoid that a region is called as differentially bound only because of biases in the specific region, due to a different genomic copy number among the cell lines. Copy number was estimated using de-duplicated reads from input DNA from all cell lines. When considering the largest set of SREBP1 binding sites (LNCaP), the signals in the MCF7-LTED vs that in LNCaP show a strong, positive correlation with estimated copy number (Fig. 3.11; SCC: spearman's correlation coefficient 0.75). In line with this, when performing the analysis without accounting for CNAs a substantial number of peaks called in PCa cells turns out to be even more bound by SREBP1 in MCF7. In order to avoid artefactual differential calls, peaks overlapping with chromosomal regions predicted to have a copy number equal or higher than 5 in any of the cell lines considered were excluded.

After this filter and normalization, differentially bound regions were estimated in each condition using one of them as a reference (Table 3.2, MCF7 RM, T47D RM and LNCaP RM). SREBP1 differentially binds 5 regions in MCF7-LTED compared to MCF7 RM (up-regulated, Table 3.2). No differential binding was found between parental and long-term starved cells in any of the other cell lines. This is in part due to the very stringent criteria applied for the statistical analysis, as wells as to the low number of replicates (n = 2) and a mid to high variability between replicates. Indeed, looking at the ChIP-seq tracks a clear modification in SREBP1 binding is observed in regions that do not show any statistically significant difference in our analysis (Fig. 3.12, Fig. 4.1c). Furthermore, by applying GREAT[306] to SREBP1 ChIP-seq data set, significant biological pathways were obtained highlighting how genes involved in processes such cholesterol, lipid and fatty acids biosynthesis and metabolism are enriched for the binding of SREBP1 in their vicinity (Fig. 3.12).

Using MCF7 RM as a reference, we found that SREBP1 differentially binds 15 and 10 regions in LNCaP RM and in LNCaP95 respectively (Table 3.2). If we use PCa cells as a reference, both BCa cell lines showed upregulated regions (27 in MCF7-LTED, 51 in T47D RM and 5 in T47D-LTED respectively, Table 3.2).

In summary, after further filtering and normalizing for CNAs, differential binding calling was performed in each condition. The differentially bound regions found to be specific for each cell line are listed in tables 3.2 to 3.8 and further discussed in Discussion and Future work section.
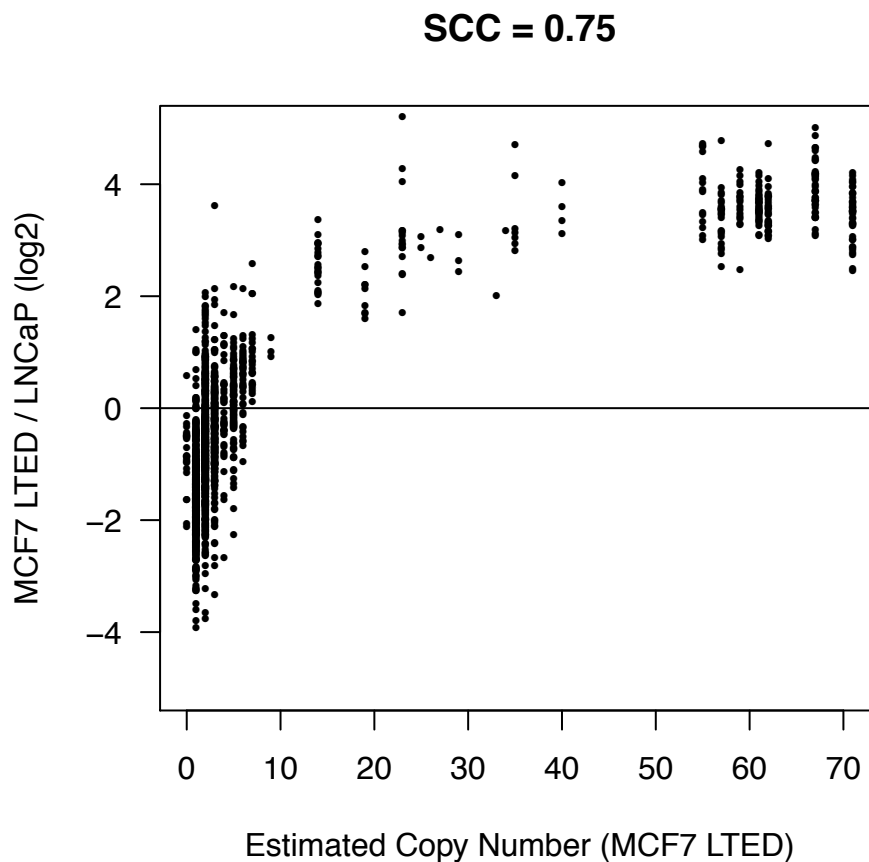
Figure 3.11: Estimation of copy number alterations. CNA in different cell lines confound the identification of cell line specific SREBP1-bound sites. When considering the largest set of SREBP1 binding sites (LNCaP), the signals in the MCF7-LTED vs that in LNCaP show a strong, positive correlation (0.75) with the local copy number estimated from input DNA (SCC: Spearman's rank correlation coefficient, Analysis performed by and figure created by Iros Barozzi).
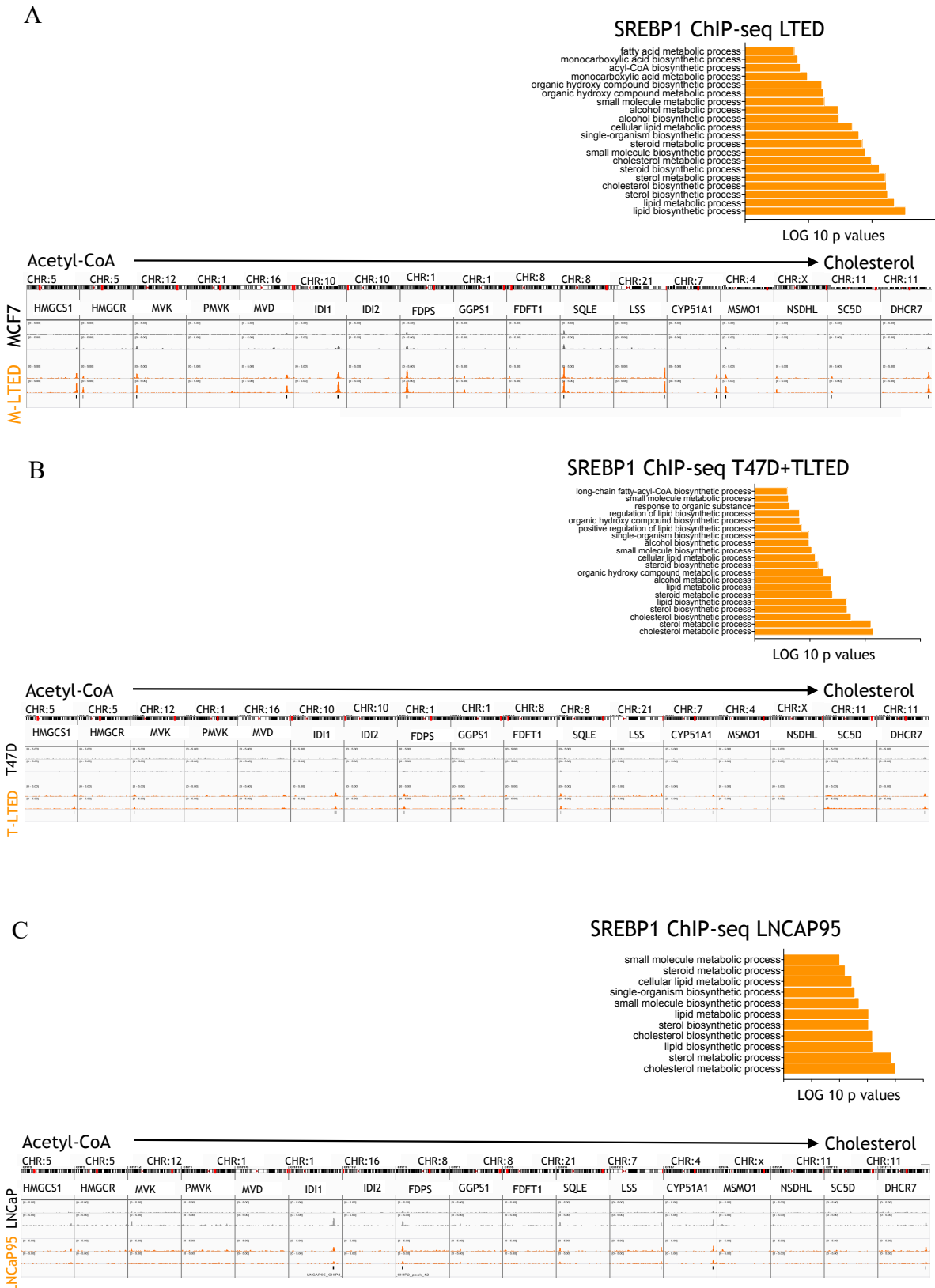
Figure 3.12: SREBP1 ChIP-seq peaks and biological pathways in BCa and PCa cells. A) MCF7 (grey) and MCF7-LTED (orange), B) T47D (grey), T-LTED (orange), C) LNCaP (grey), LNCaP95 (orange). Regions shown are genes involved in the CB starting from Acetyl-CoA. (Analysis performed by and figure created by Luca Magnani).

vs_MCF-RM

vs_MCF-RM	LNCaP.RM	LNCaP95.WM	MCF7_LTED.WM	MCF7.WM	MCF7.WM.E2	T47D_LTED.WM	T47D.RM	T47D.WM	T47D.WM.E2
unchanged	212	214	224	229	229	226	211	215	224
up	15	10	5	0	0	1	9	8	3
down	2	5	0	0	0	2	9	6	2

vs_T47D-RM

vs_T47D-RM	LNCaP.RM	LNCaP95.WM	MCF7_LTED.WM	MCF7.RM	MCF7.WM	MCF7.WM.E2	T47D_LTED.WM	T47D.WM	T47D.WM.E2
unchanged	177	177	172	175	179	177	182	183	183
up	1	0	1	4	1	1	1	0	0
down	5	6	10	4	3	5	0	0	0

vs_LNCaP-RM

vs_LNCaP-RM	LNCaP95.WM	MCF7_LTED.WM	MCF7.RM	MCF7.WM	MCF7.WM.E2	T47D_LTED.WM	T47D.RM	T47D.WM	T47D.WM.E2
unchanged	1168	1135	1182	1185	1178	1171	1117	1168	1161
up	0	27	0	0	2	5	51	3	2
down	21	27	7	4	9	13	21	18	26

Table 3.2: Differential binding analysis. *edgeR* was used for normalization and to estimate the differentially bound regions in each condition, using one of them as reference (MCF7 red media, T47D red media and LNCaP, respectively). Differential binding was evaluated using the function *exactTest*. A region was defined as differentially bound if showing a *q*-value ≤ 0.05 (Benjamini-Hochberg correction), (Analysis performed by and figure created by Iros Barozzi).

3.7.4 Evidence of a crosstalk between SREBP1 and nuclear receptors

An ever-increasing body of literature suggests that the cholesterol-driven signalling in cancer might be centered on the potential crosstalk between nuclear receptors and cholesterol-specific TFs[53,200]. Nuclear receptors can be activated other than by sexual hormones (oestrogens, androgens) also by compounds derived from cholesterol biosynthesis, such as 25 and 27 hydroxycholesterols (25HC, 27HC) in breast cancer cells[197,307,308]. Moreover, it has been shown that 27HC can increase metastatic invasion in mice xenografted with MCF7 cells[309] and statins treatment is able to reduce ERα binding to DNA and abrogate cell invasion[53]. An extensive genome-wide ERα binding has also been previously confirmed by ChIP-seq analysis in LTED cells, despite oestrogen-deprived conditions[53]. It is well known that, androgen receptor (AR) activation is regulated by SREBP1 in prostate cancer (see 1.2.4). Therefore, to answer the question whether SREBP1 binding is dependent on the hormone receptor, I performed SREBP1 ChIP-seq also in MCF7F cultured either in full or stripped media. As previously discussed, Fulvestrant is a drug that through ER immobilization leads to the increased ER turnover (see 1.1.3). By integrating RNA-seq and H3K27ac ChIP-seq in ETR (ET-resistant) cells, it has been showed that fulvestrant resistant cells acquire an ERα-negative status, while the rest of the panel remains ERα positive[53]. We reasoned that by culturing MCF7F with or without oestradiol we can investigate the role of hormones in regulating directly SREBP1 binding in breast cancer cells. ChIP-seq analysis after applying the same stringent filtering criteria used for peak calling in other cell lines (Fig. 3.7) return no binding in MCF7F cultured in either FM or WM. Whereas absence of hormone alone cannot achieve SREBP1 activation, these results further suggest that SREBP1 action may require the presence of the nuclear receptor.

Next, publicly available nuclear receptors (ERα and AR) ChIP-seq data were overlapped with the SREBP1-bound regions identified in our MCF7, T47D and LNCaP cells' datasets. Using a combined score (accounting for both statistical significance and effect size), a significant overlap is observed between either ERα (Fig. 3.13 A) or ERα-related transcription factors (Fig. 3.13 C) with SREBP1 peaks in breast cancer indicating that SREBP1 binds mostly the same regions that are bound by ERα. The same analysis was then performed using AR ChIP-seq data. We found that there is a significant enrichment between AR and SREBP1 peaks in LNCaP cells whereas the score in breast cancer was comparatively low (Fig. 3.13 B, MCF7 and T47D used as a control) suggesting a possible crosstalk between AR and SREBP1 in prostate cancer on chromatin.

Considering previously published observations from our group and taking together the analysis here presented, it is tempting to speculate that there is a link between SREBP1 activation and ERα signalling in breast cancer as SREBP1 has been previously proved to regulate AR in PCa. These data suggest, for the first time, a possible crosstalk between the master regulator of the lipid biosynthesis and the hormone receptor ERα. Further experiments are needed to fully address the potential interaction of NRs and SREBP1 and functionally characterize the mechanisms of activation/regulation in cancer.
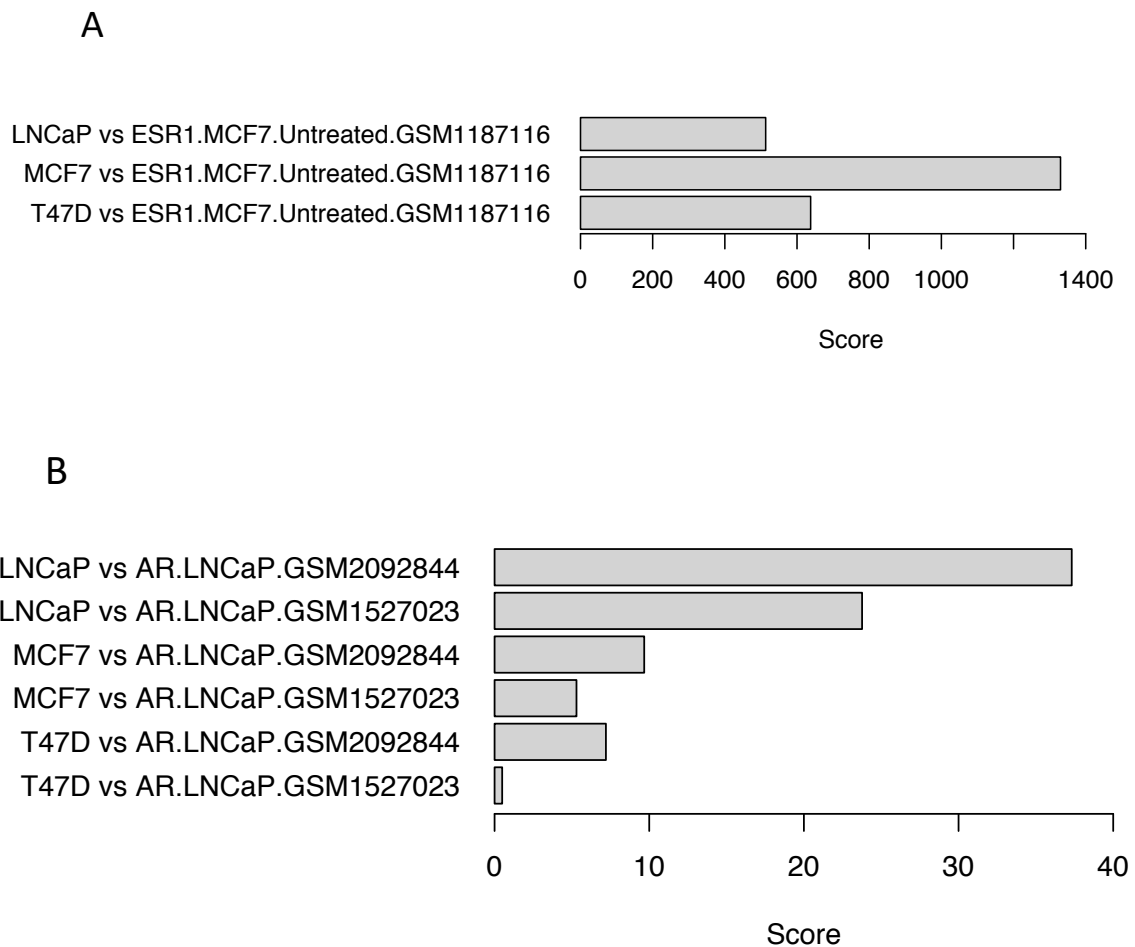
Figure 3.13: Nuclear receptors ChIP-seq dataset analysis. Bar plots highlighting the overlap between the SREBP1 cistrome in different cell lines and publicly available ChIP-seq profiles for nuclear receptors (ERα and AR; A and B) as well as with ERα-related and general transcription factors (profiles in MCF7; C). A) ERα ChIP-seq data overlap with SREBP1 ChIP-seq data in MCF7 and T47D cells, with LNCaP used as a control. B) AR ChIP-seq data overlap with SREBP1 ChIP-seq data in prostate cancer cells, with MCF7 and T47D used as a control. Following page: C) ERα-related transcription factors ChIP-seq data overlap with SREBP1 ChIP-seq data in MCF7 (grey) and T47D (black) cells. Scores are a combination of effect size and statistical significance, as calculated using GIGGLE. (Analysis performed by and figure created by Iros Barozzi).

grey = MCF7; black = T47D

Score

C

138
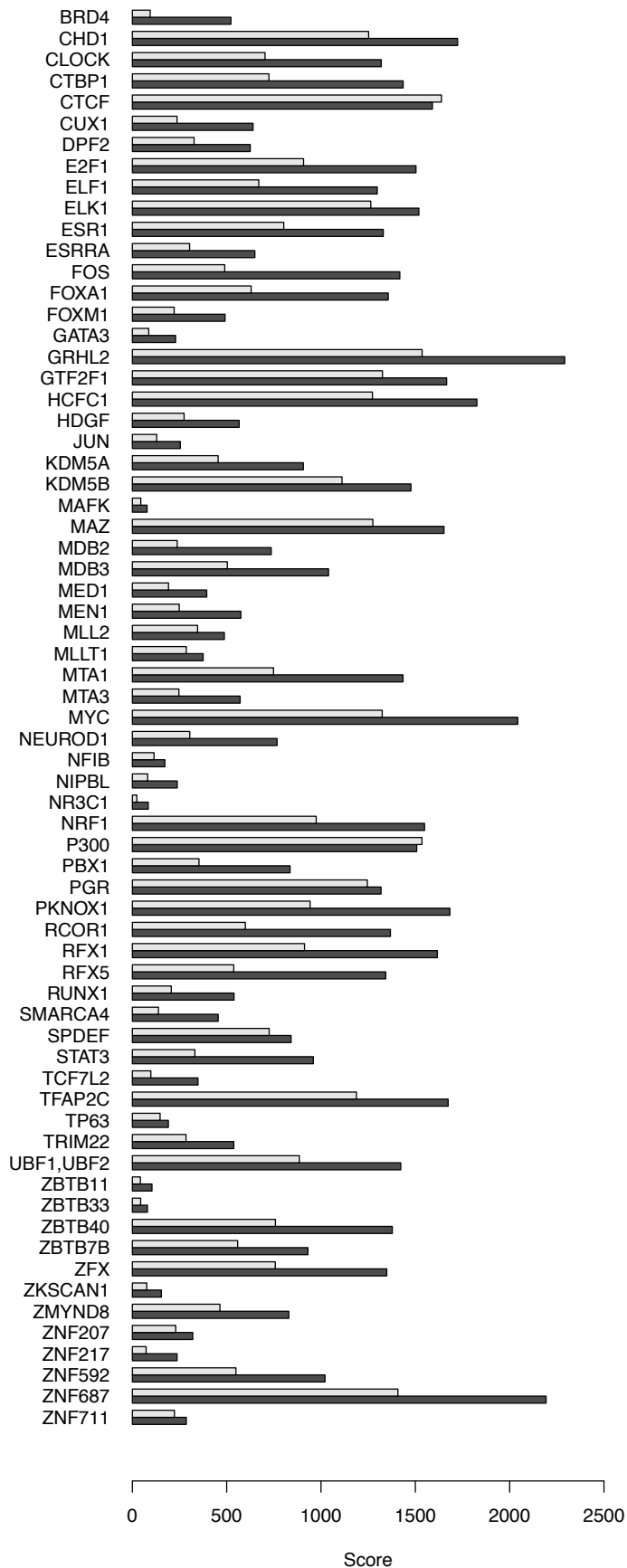
3. 8 Discussion and Future work

3.8.1 SREBP1 activation in hormone-insensitive cancer cell lines

In this chapter, I investigated the role of SREBP1 in the development of ET-resistant cancer phenotype. Cancer cells becoming hormone independent activate a lipogenic phenotype that is mainly driven by SREBP1. IHC analysis showed increased SREBP1 staining in metastatic lesions of AI-treated patients compared to the respective matched untreated primary. Although measurements from primary and metastatic samples are still limited, the observed SREBP1 staining increase during BCa progression suggests that SREBP1 might be playing a role upon AI-driven resistance development *in vivo* as well. To test whether SREBP1 differentially binds in BCa, I performed ChIP on matched primary and AI-treated metastatic samples. Although high cellularity lesions (>80%) were used, a first attempt did not seem to have worked. Similar to *in vitro* experiments, further optimization of the protocol is needed in order to successfully perform this technique on tumour samples. Possible limitations could be ascribed to the antibody efficiency that is much different compared to the antibody anti H3K27ac, previously successfully used in tumour samples in our lab, and to the substrate (SREBP1 in chromatin) that is much lower than histone marks. Alternatively, SREBP1 ChIP could be performed on cells derived from either pleural effusion or ascites of metastatic patients or on tumour samples of patient-derived xenograph (PDX) models. To further support our *in vivo* data, future larger, multiple independent datasets are required. These could include IHC for SREBP1 staining (differentiating nuclear versus cytoplasmic) in primary versus matched samples from AI treated metastatic patients, possibly dividing cases per tumour grade and Ki-67.

Our data suggest that there is an activation of the SREBP1-driven CB pathway upon resistance development. Thus, we asked the question whether this metabolic reprogramming happens during the early or late phases of ET. Time course studies showed that only long-term hormone deprived cancer cell lines are characterised by the activation of SREBP1-driven CB pathway. One of the possibilities about the reasons why there is no activation of the CB in the short-term starvation is that SREBP1 could be modulated through signalling and need other stimuli or co-factors to start the transcription. Alternatively, SREBP1 could be activated in a specific phenotypic subpopulation. The hypothesis that ET could select for a subgroup of treatment-naïve cancer cells is supported by a recent study published from our group dissecting at a single cell resolution, the phenotypic heterogeneity and plasticity during the early and late phases of ET[292]. They identify a rare sub-population of pre-adapted cells (PA) which undergoes further transcriptomic reprogramming (upregulation of CB and ER signalling) and copy number

changes to acquire full resistance. PA cells displayed features of mixed epithelial and mesenchymal traits, along with up-regulation of p53 pathway, cell polarity (apical junction components), apoptosis, and hypoxia[292]. Although analysis in PA cells didn't display anything associated to CB, cholesterol homeostasis-related genes were found[292]. The PA signature is also found significantly enriched in circulating tumour cells (CTCs) and even higher levels in clusters of CTCs providing a further link between drug-induced adaptation and metastatic invasion. Therefore, a multi-step model for ET resistance development involving both non-genetic and genetic contributions is proposed[292]. This model would partly explain the delayed relapse common to ET treated patients by hypothesising that a group of cells (PA-like) may be selected by ET entering a quiescent state for up to more than ten years. In the same way we can speculate that CB may be one of the factors contributing to exit from quiescence and progress of disease.

Although the factors contributing to the progression of breast and prostate cancer remain incompletely understood, hormones have long been recognized to play a central role in this process. For many decades, the hormone dependency of these types of cancer has been exploited therapeutically by hormone ablation strategies. Standard therapy for non-organ confined prostate cancer, for example, aims to block the production or action of androgens. Although initially successful, in a still high percentage of cases these forms of therapy fail, and a hormone depletion independent (i.e. androgen depletion independent (ADI) in PCa) disease emerges, for which currently no cure is available. Despite low circulating levels of functional hormones, hormone independent cancers still rely on functional nuclear receptors (AR and ERα) that are critical for the proliferation and survival of drug-resistant tumour cells. The crosstalk between androgens and the activation of a lipogenic phenotype through SREBP has been widely investigated during the last decades (see 1.2.4). It has been reported that LNCaP cells respond to androgen stimulation, independently of the sterol content, by an induction of SREBP1 and SREBP2 expression (1.3 and 2.1-fold) and the upregulation of target enzymes expression[219]. In our time course studies, LNCaP cells responded to both acute and long-term androgen starvation by a higher induction of SREBPs and downstream enzymes. At a first sight, our data may appear in contrast with many published studies. Considering that *in vivo* experiments have shown androgen-mediated changes in lipogenic genes expression in non-cancerous androgen-responsive cells as well[219], one possibility is that androgen stimulation of SREBP-mediated lipid pathway could represent a more general mechanism of physiological regulation. Androgen-induced maturation of SREBP precursors is also

conserved in other androgen-sensitive cell systems[219]. However, LNCaP xenograft model as well as clinical specimens of prostate cancer demonstrated an up regulation of SREBPs and their downstream effector genes during progression to androgen independence[198]. In the clinical setting various grade of disease, especially androgen independent tumours, show up regulation of FAS. As the androgen independent phenotype emerges, enzymes critical for lipogenesis and cholesterol synthesis are activated and likely contribute significantly to cell survival of prostate cancer[198]. SREBP1 confers growth advantages in both hormone-naïve and castration resistant xenograft mouse models[215]. Interestingly, it has been demonstrated that SREBP1 expression is increased in advanced forms of human prostate cancer[198]: whereas it was found expressed only in 20% of normal/benign prostate tissues, expression of SREBP1 protein increased with higher Gleason grades of disease (50% in grade 3, 71% in grade 5). Nuclear SREBP1 was detected prevalently in grade 4 and 5 prostate cancers suggesting that expression of SREBP1 is closely linked with the development of aggressive pathologic features in PCa. Taken together our data reconcile with the hitherto published results indicating that SREBP1 acts as a key player in the development of a hormone independent, drug resistant cancer phenotype.

Whereas bulk mRNA analysis indicated increased levels of SREBP1 in LTED, preliminary data from RNAscope imaging showed an equal amount of transcript found in either parental or LTED cells. With the current evidence we cannot draw any conclusion about the transcriptional levels of SREBP1. Firstly, RT-qPCR and RNAscope are very different techniques analysing the bulk mRNA levels in the whole population the former, and the single mRNA molecules at a single cell level the latter. Secondly, RNAscope imaging data here presented were just preliminary, based on few single cells and not representative of the whole cell population. Thirdly, there is the possibility that RNA probes might target specific splice variants whereas SREBP1 primers are designed to recognise them all. Last consideration to take into account about this discrepancy is that SREBP1 may be activated in a subpopulation of reprogrammed cells. With current evidence we are not able to distinguish whether a subgroup of activated cells may be responsible for the increased amount of transcript in LTED cells. Despite these differences, downstream CB enzymes were up regulated in LTED at both transcriptional and protein level suggesting that the amount of SREBP1 transcript *per se* is not likely to be the driver of the increased *de novo* lipid activity. However, whether and how SREBP1 activation, rather than the amount of SREBP1 transcript *per se,* may be responsible for the increased *de novo* lipogenic activity needs further investigation.

Future work aiming at extending RNA scope analysis to other BCa and PCa cell lines using a higher number of single cells, is needed in order to obtain representative results about the upregulation of SREBP1 pathway. Immunoblotting and mRNA fractionation studies in several cell lines would also be informative of possible changes in the SREBP1 activation dynamics. Furthermore, it would be important to address the post-translational modifications involved in processing this TF by studying phosphorylation and ubiquitin-proteasome-dependent degradation in stressful conditions such as hormone deprivation. Besides, regulation of SREBP1 trafficking through SCAP and INSIG may play an important role in processing the mature functional transcription factor.

An alternative pathway to the classic sterol-induced one of SREBP1 activation in LNCaP cells suggest that SCAP, not SREBP, is the primary target of hormone action[218]. AR directly stimulates SCAP transcription (the SCAP Androgen Responsive Element (ARE) binds the AR DNA Binding Domain (DBD)) favouring translocation of SREBP precursors to the Golgi[218,219]. Pharmacologically or genetically targeting SCAP can inhibit SREBP activity and significantly attenuate tumour growth *in vitro* characterised by: G1 cell-cycle-arrest, decreased basal oxygen consumption, increased ROS and a profound defect in mitochondrial respiratory capacity[227]. Although our data do not provide proof for a functional relationship between the observed increase in both SREBP1 and SCAP in androgen starvation conditions, they certainly point toward such a relationship. Based on our current results, we can only speculate that the starvation-dependent activation of SREBP1 pathway may be mediated through SCAP in LNCaP cells. Further studies are needed to characterise the role of SCAP in relation to steroid deprivation to shed light on these data.

As investigation into the regulation of the SREBP pathway goes on, a picture of ever-increasing complexity emerges. Indeed, SREBP is also regulated through glucose/glutamine-dependent reactions. Cheng and colleagues showed that SREBP activation requires glucose in glioblastoma (GBM) cells and that SCAP levels decrease considerably in low glucose conditions[293]. Glucose controls SCAP because of its requirement for SCAP glycosylation. The same authors show that inhibitors of N-linked, but not O-linked, glycosylation decreases SCAP levels and prevent SREBP activation[293]. Overexpression of SCAP increased tumour weight relative to control and decreased survival in their model[293]. The requirement of glucose for SCAP N-glycosylation permits glucose supply to control SREBP-dependent lipogenesis in GBM. To what extent physiological changes in glucose are able to influence SCAP

glycosylation and function is not known. Blood glucose concentrations are tightly controlled, and such low glucose may only occur in poorly vascularised tumours. Besides, cancer cells frequently increase glutamine uptake, and glutamine can also serve as a precursor of N-acetylglucosamine. Whether glutamine can also supply signal to SCAP and SREBP-driven lipogenesis needs to be elucidated. In contrast, a different study showed a connection between nutrient sensor O-GlcNAcylation, SREBP1 and lipid synthesis in breast cancer[310]. O-GlcNAcylation is required for lipid biosynthesis through the regulation of SREBP1 in an AMPK-dependent manner. In addition, they found that O-GlcNAcylation also regulates lipid metabolism, and SREBP in adult lipogenic tissue such as lactating mammary gland suggesting that cancer cells have co-opted this normal regulation of *de novo* lipid synthesis [310].

Our results point towards a glutamine-dependent rather than glucose-dependent activation of SREBP1 pathway in LTED. This metabolic switch from a highly glycolytic phenotype towards a glucose independent phenotype in long-term hormone deprivation indicates a nutrient dependence shift upon resistance development. Metabolic analysis performed on SREBP1 *null* cancer cells in different nutrient conditions would further help to identify SREBP1-dependent metabolic changes. Future studies may also shed light on the role of glutamine in the activation of SCAP-dependent SREBP1-driven lipogenesis in ETR cancer cells.

The importance of the role played by SREBP1 in cancer cell survival and growth withstanding hormone depletion was revealed by suppressing SREBP1 expression in cells cultured in absence of oestradiol. Despite SREBP1 silencing being quite effective in abrogating cell growth, there was one cell line, CRISPR clone 1b, that showed an aberrant behaviour. Future work performing RNA-seq of both silenced and overexpressing SREBP1 cell lines is warranted to transcriptionally characterise the effects of SREBP1 manipulation in cancer. Furthermore, performing CRISPR/Cas9 in other BCa (LTED, T47D and T47D- LTED) and in PCa cell lines would help to further elucidate the role of SREBP1 in ETR cancer cell survival. Besides, auxiliary experiments should clarify why SREBP1 appears to be needed to survive the acute stress imposed by endocrine treatment while cholesterol upregulation seems to occur only at later stages of resistance acquisition. This apparent disentanglement between SREBP1 activity and the SREBP1-driven canonical pathway suggest that this TF might mediate other stress-induced non canonical pathways in cancer cells.

Additional experiments using stable expression of a Tamoxifen-inducible SREBP1 construct or transient transfection of SREBP1-HaloTag[311] in CRISPR/Cas9 KO cancer cells could be

informative of the mechanisms of activation of SREBP1 signalling. The HaloTag is a modified bacterial enzyme designed to covalently bind to a synthetic ligand of choice and fuse to a protein of interest[311]. This bacterial enzyme is designed to facilitate visualization of the subcellular localization, immobilization, or capture of the binding partners of a protein of interest. Since bacterial dehalogenases are relatively small and the reactions used are foreign to mammalian cells, there is no interference by endogenous mammalian metabolic reactions. Once the HaloTag fusion protein is expressed, there is a wide range of potential areas of experimentation including enzymatic assays, cellular imaging, protein arrays and determination of sub-cellular localization. Imaging tools such as Incucyte, inverted confocal microscopy and super-resolution microscopy could be used to study protein localization, trafficking and turnover of SREBP1 in different experimental conditions: before and after stimulation with E2, before and after treatment with several compounds (such as Fatostatin). Translational future work would be provided by screening a full battery of compounds coupled with metabolic and transcriptional profiling of parental and LTED cell lines under treatment (see 1.2.3.f).

3.8.2 SREBP1 cistrome in hormone-dependent cancer cell lines

Optimisation of the SREBP1 ChIP protocol allowed for: 1. Testing for the antibody able to efficiently pull-down chromatin samples. 2. Reaching the correct balance between cell number, antibody and relative Dynabeads concentration used. 3. Obtaining samples with a DNA concentration sufficient enough to look for enrichment between Input and ChIP samples and to then proceed with sequencing and downstream analysis.

Genome-wide profiling of SREBP1 binding in hormone-dependent cancer cells showed that biological replicates were consistent for each cell line. Furthermore, the most significantly over-represented TF-motifs in the filtered peaks of every cell line were those previously identified for SREBF1 and SREBF2, confirming that the generated profiles represent *bona fide* binding sites for SREBP1. The number of peaks called in any condition was very similar between different breast cancer cell lines. However, a significant difference was found between breast and prostate cancer even though limited to a small set of potential target genes.

All ChIP-seq performed in different cell lines were then analysed together through hierarchical clustering. This extended out previous observations at a transcriptional level to the chromatin

level, i.e. no difference in SREBP1 binding between acute starvation conditions and parental cells. The largest difference is found between different cancer cell lines leading to the hypothesis that SREBP1 binding may be cancer specific.

Because SREBP1 may be regulated by alternative pathways other than the classic sterol-related one, the relationship between this TF and the NRs was investigated at a chromatin level. It is well known that SREBP1 regulates AR promoter activity and expression and cell viability in prostate cancer cells[220]. Bicalutamide, an AR antagonist, abolished the stimulatory effects of androgens and no androgens-induced lipogenic response is observed in AR-negative PCa cell lines consistent with the requirement of the AR to mediate the activation of the SREBP pathway in androgen-induced lipogenesis [219]. In turn SREBP1 promotes human prostate tumour growth in mouse xenograph models and supports the development of a castration-resistant progression phenotype through the induction of AR, FASN and ROS. In breast cancer, not much is known about the interplay between ERα and SREBP1[215]. Our group has formerly shown that endogenous CB activity promotes ERα binding in LTED cells also contributing to their invasive phenotype[53]. By blocking SREBP1 trafficking to the nucleus in LTED cells, ERα signal is impaired; blocking endogenous CB using Lovastatin (HMGCR inhibitor) or Terbinafine (SQLE inhibitor) induced a significant reduction in ERα recruitment at 27HC unique sites in LTED cells, further suggesting that, by replacing oestrogen action, CB contributes to cell-autonomous ERα recruitment in LTED cells[53] (working model in Fig. 2). We found no binding of SREBP1 in MCF7F cultured in either FM or WM further suggesting that SREBP1 action may require the presence of the ERα. Moreover, when combining publicly available datasets from AR, ESR and ESR-related genes ChIP-seq with our SREBP1 ChIP-seq we observed: i) a significant overlap between both ERα and ER-related transcription factors and SREBP1 peaks in BCa and ii) a significant co-occurrence of AR and SREBP1 binding in LNCaP cells. These data indicate that SREBP1 tend to co-bind at regions occupied by ERα in BCa or AR in PCa, confirming the link between AR and SREBP1 in prostate cancer and suggesting a possible crosstalk between ERα and SREBP1 in breast cancer at the chromatin level.

Finally, differential binding analysis identified canonical targets (genes involved in the lipid metabolism), but also many SREBP1 non-canonical target genes in both BCa and PCa cell lines. By comparing different groups, here I discuss some of the regions of interest that could be considered for further studies and future work.

- SREBP1 target genes in LTED (Table 3.3, MCF7 RM vs LTED):

1 ALOX15B (Arachidonate 15-Lipoxygenase Type B) gene encodes a member of the lipoxygenase family that catalyses the peroxidation of free and esterified polyunsaturated fatty acids generating a spectrum of bioactive lipid mediators[315]. Lipoxygenases in humans are expressed in a tissue-specific fashion with the isoenzyme 15-LOX-2 (abbreviated 15B) expressed in skin, cornea, prostate, lung, and oesophagus. ALOX15B plays an important role in cholesterol homeostasis in human macrophages, it has also been implicated in the regulation of cytokine secretion by macrophages therefore playing a role in the immune response[312,313]. The link between SREBP and ALOX has been formerly described: silencing the 15-LOX isoforms impaired SREBP-2 signalling by inhibiting its processing into a mature transcription factor. It also reduced SREBP-2 binding to sterol regulatory elements and subsequent target gene expression, decreasing cellular cholesterol and the cholesterol intermediates as well as oxysterols activated macrophages.

Opposing actions are reported in the literature for both 15-LOX-1 and 15- LOX-2 in carcinogenesis[317]. This dispute extends to both solid tumours and haematological malignancies[317].

ALOX15B gene is located in a cluster of related genes on the short arm of chromosome 17. A recent study reported that heterozygous deletion of 17p13.1 region produces a greater effect on lymphoma and leukaemia development than *Trp53* deletion. The effect of 17p13.1 region loss on tumorigenesis involves co-deleted genes such as *Alox15b*, the suppression of which cooperates with *Trp53* loss to produce more aggressive disease. Their results imply that the selective advantage produced by human chromosome 17p deletion reflects the combined impact of *TP53* loss and the reduced dosage of linked tumour suppressor genes[316].

Diseases associated with ALOX15B include prostate cancer, in which it is thought to regulate progression through cell cycle and cell proliferation[315]. The tumour-suppressor function of 15-LOX-2 in normal prostate epithelial cells may be explained by the induction of replicative senescence. Indeed, 15-LOX-2 is overexpressed in age-dependent prostatic hyperplasia, but cell senescence may hinder progression to malignant transformation[315]. There is one report showing an upregulated expression of 15-LOX-2, but not 15-LOX-1, in macrophages obtained from patients with newly diagnosed renal tumour suggesting that 15-LOX-2 supports immune evasion through tumour associated macrophages[315]. In another study, the expression of 15-LOX-2 was significantly up-regulated in lung adenocarcinoma tissue samples (NSCLC) compared with adjacent normal ones whereas silencing 15-LOX-2 inhibited the proliferation

and migration of A549 cells *in vitro*[317]. Expression of 15-lipoxygenase-2 mRNA is also found strongly augmented in primary ovarian carcinoma and ovarian carcinoma metastases (20-fold) when compared with normal ovarian tissue[318]. Other reports attributed pro-tumorigenic actions to 15-LOX metabolites in BCa cells. A prospective study examined 120 human BCa tumour biopsies for the expression of various LOX types and found that the levels of 15-LOX-1 were reduced[315]. Other studies reported increased protein expression of 15-LOX-2 but not of 15-LOX-1 in a BCa cell line and localization of 15-LOX to sentinel lymph nodes and matching tumour stage in 13 tumour samples examined by tissue microarray[315].

2 G6P (Glucose-6-Phosphate Dehydrogenase) is a cytosolic enzyme encoded by a housekeeping X-linked gene whose main function is to produce NADPH, a key electron donor in the defence against oxidizing agents and in reductive biosynthetic reactions[319]. G6PD catalyses the rate-limiting step of the oxidative pentose-phosphate pathway. The main function of this enzyme is to provide reducing power (NADPH) and pentose phosphates for fatty acid and nucleic acid synthesis[319]. G6PD is remarkable for its genetic diversity. Many variants of G6PD, mostly produced from missense mutations, have been described with wide ranging levels of enzyme activity and associated clinical symptoms[319]. G6PD deficiency may cause neonatal jaundice, acute haemolysis, or severe chronic non-spherocytic haemolytic anaemia. An important paralog of this gene is H6PD. Upon entering a cell, glucose is phosphorylated by hexokinase (HK) to glucose-6-phosphate (G6P). Cancer cells seem to preferentially express HK2, which is thought to support anabolic metabolism, in particular, it has been reported that prostate cancer cells strongly upregulate HK2 and lipid synthesis in response to androgens[200]. G6P, produced by HK2, may take one of two fates: metabolism by the pentose phosphate pathway (PPP) or glycolysis. Cancer cells can shunt G6P through the PPP to produce NADPH, which is a critical reductant for fatty acid synthase. PPP intermediates can also return to glycolysis, producing pyruvate, which may be used later for fatty acid synthesis[200].

3 DIP2C gene encodes a member of the disco-interacting protein homolog 2 family. The protein shares strong similarity with a Drosophila protein which interacts with the transcription factor disco and is expressed in the nervous system[320].

4 HIF1A (Hypoxia-Inducible Factor 1 Alpha Subunit) is the master regulator of cellular and systemic homeostatic response to hypoxia by activating transcription of genes involved in energy metabolism, angiogenesis and apoptosis[321]. HIF-1 thus plays an essential role in

embryonic vascularization, pathophysiology of ischemic disease and tumour angiogenesis, growth and invasion[321]. Hypoxia exerts a selective pressure for malignant progression leading to the survival of subpopulations of cells able to adapt to conditions of poor nutrition and a hostile microenvironment[321]. HIF1α reduces apoptosis and is frequently overexpressed in tumour cells. Activation of HIF1α, initiate autophagy and aerobic glycolysis providing cells surrounding the cancer cells with the energy necessary to promote their growth. HIF1α is also associated with resistance to conventional chemo- and radiation therapy[321].

Interestingly, HIF has been reported to control fatty acid metabolism contributing to clear cell renal cell carcinoma (ccRCC) tumourigenesis[321]. HIF directly represses the rate-limiting component of mitochondrial fatty acid transport, carnitine palitoyltransferase 1A (CPT1A), therefore reducing FA transport in the mitochondria and increasing lipid deposition in ccRCC[322]. Hypoxia-induced-lipid storage has also been demonstrated to serve as a protective barrier against oxidative stress-induced toxicity in breast and glioma cell lines due to a HIF1α-dependent increase of FA uptake via FA binding proteins FABP3 and FABP7[323]. Furthermore, Lewis and colleagues investigated the effect of hypoxia and serum deprivation on SREBP activity and the expression of lipid metabolism genes in human glioblastoma multiforme (GBM) cancer cells[324]. They found that SREBP transcriptional activity was induced by serum depletion both in normoxic and hypoxic cells and that activation of SREBP was required to maintain the expression of FA and cholesterol metabolism genes under hypoxic conditions[324]. Additionally, they showed that hypoxia-induced expression of SCD, FABP3 and FABP7 was strongly dependent on SREBP function. Inhibition of SREBP blocked lipid biosynthesis and impaired cell survival in a three-dimensional spheroid hypoxic cancer cells model. Finally, poor survival in glioblastoma patients was found associated with a SREBP-defined gene signature[324].

5 LOC101928911 is an uncharacterised RNA Gene affiliated with the non-coding RNA (ncRNA) class[320].

- SREBP1 target genes in PCa (Tables 3.4 and 3.5, BCa (MCF7 RM and T47D RM) vs LNCaP):

Regions associated with metabolism:

FADS2 (Fatty Acid Desaturase 2) is a member of the fatty acid desaturase (FADS) gene family. Acts as a fatty acyl-coenzyme A (CoA) desaturase and regulates unsaturation of fatty acids. It is involved in biosynthesis of highly unsaturated fatty acids (HUFA) from the essential polyunsaturated fatty acids (PUFA) linoleic acid (LA) and alpha-linolenic acid (ALA) precursors. Catalyses the first and rate limiting step in this pathway which is the desaturation of LA and ALA[327]. Gene Ontology annotations related to this gene include iron ion binding and oxidoreductase activity.

ACACA (Acetyl-CoA Carboxylase Alpha) catalyses the rate-limiting reaction in the biogenesis of long-chain fatty acids. ACC-alpha is highly enriched in lipogenic tissues[320]. Among its related pathways are Development, Leptin signalling via PI3K-dependent pathway and Fatty acid metabolism.

CYP51A1 (Cytochrome P450 Family 51 Subfamily A Member 1): products of the CYP51 reaction are vital intermediates in pathways leading to the formation of cholesterol in humans[320]. As a member of this family, lanosterol 14α-demethylase is responsible for an essential step in the biosynthesis of sterols. Antifungal, such as azoles, are non-competitive inhibitors of this enzyme.

Further target genes involved in metabolic pathways: **ALOX15B** (previously described)**, SREBF2**, **SMAD3** (SMAD Family Member 3), **PDP2** (Pyruvate Dehydrogenase Phosphatase Catalytic Subunit 2), **MTHFR** (Methylenetetrahydrofolate Reductase) and **H6PD** (Hexose-6-Phosphate Dehydrogenase/Glucose 1-Dehydrogenase) paralog of G6PD.

Regions involved in the inflammatory response and protection from oxidative stress and senescence:

TP73 (Tumour Protein P73) is a member of the p53 family of transcription factors involved in cellular responses to stress and development. It maps to a region on chromosome 1p36 that is frequently deleted in neuroblastoma and other tumours and thought to contain multiple tumour suppressor genes[320]. Diseases associated with TP73 include Small Cell Cancer of the lung. P73 may be a tumour suppressor protein, it participates in the apoptotic response to DNA damage.

Isoforms containing the transactivation domain are pro-apoptotic, isoforms lacking the domain are anti-apoptotic and block the function of p53 and transactivating p73 isoforms[320].

MGST1 (Microsomal Glutathione S-Transferase 1) catalyses the reduction of lipid hydroperoxides. This protein is localized in the endoplasmic reticulum and mitochondrial membrane where it is thought to protect these membranes from oxidative stress[320]. Through its glutathione S-transferase and peroxidase activities, MGST1 is involved in cellular defence against toxic, carcinogenic, and pharmacologically active electrophilic compounds. Diseases associated with MGST1 include Ewing Sarcoma. Among its related pathways are Cytochrome P450 and Innate Immune System.

JUND/mir3188 (Jun D Proto-Oncogene/microRNA 3188) The protein encoded by this intron less gene is a member of the JUN family, and a functional component of the AP1 transcription factor complex[320]. This protein has been proposed to protect cells from p53-dependent senescence and apoptosis. T-Cell Leukemia has been associated with JUND. Among its related pathways are Toll-Like receptor Signalling Pathways and ERK Signalling.

HIF1A (previously described)**, AGTRAP** (Angiotensin II Receptor Associated Protein) is involved in the Oncogenic MAPK signalling, **PANX2** (Pannexin 2) is involved in Electric Transmission Across Gap Junctions, **BHLHE40** (Basic Helix-Loop-Helix Family Member E40). **DNAJB12** (DnaJ Heat Shock Protein Family (Hsp40) Member B12) and **ERP29** (Endoplasmic Reticulum Protein 29) are involved in Protein processing in endoplasmic reticulum. **PIK3C2A** (Phosphatidylinositol-4-Phosphate 3-Kinase Catalytic Subunit Type 2 Alpha) is involved in Vesicle-mediated transport. **SKAP2** (Src Kinase Associated Phosphoprotein 2) is involved in cell junction organization. **DIP2C** has been previously described. **MYL10** (Myosin Light Chain 10) involved in the Regulation of actin cytoskeleton.

- SREBP1 target genes in LNCaP95 (see table 3.6, upregulated in LNCaP RM vs LNCaP95):

CBX4 (Chromobox 4) is involved in C-MYB transcription factor network. Gene Ontology annotations related to this gene include chromatin binding and ligase activity. It is involved in the sumoylation of a p53/TP53 transcriptional coactivator, hence indirectly regulates p53/TP53

transcriptional activation resulting in p21/CDKN1A expression[325]. It has also been reported to be part of a Polycomb group (PcG) multiprotein PRC1-like complex, a complex class that via chromatin remodelling and modification of histones maintains the transcriptionally repressive state of many genes throughout development[325].

SEC63 (SEC63 Homolog, Protein Translocation Regulator). The Sec61 complex is the central component of the protein translocation apparatus of the ER membrane. The protein encoded by this gene and SEC62 protein are found to be associated with ribosome-free SEC61 complex. It is speculated that Sec61-Sec62-Sec63 may perform post-translational protein translocation into the ER[326]. This complex might also perform the backward transport of ER proteins that are subject to the ubiquitin-proteasome-dependent degradation pathway. Among its related pathways are UPR and Protein processing in ER[326].

TP73 and MYL10 have been previously described**.**

- SREBP1 target genes in BCa (Tables 3.7 and 3.8, PCa vs BCa):

Regions associated with metabolism:

ELOVL5 (ELOVL Family Member 5, Elongation of Long Chain Fatty Acids) is highly expressed in the adrenal gland and testis and is involved in the elongation of long-chain polyunsaturated fatty acids[320]. Among its related pathways are Fatty acid metabolism and alpha-linolenic acid (ALA) metabolism[320]. ELOVL5 catalyses the first and rate-limiting reaction of the long-chain fatty acids elongation cycle. It can also participate in the production of monounsaturated and of polyunsaturated very long chain fatty acids (VLCFAs) of different chain lengths that function as precursors of membrane lipids and lipid mediators.
ELOVL5 together with ELOVL7, another member of the FA elongase family, have been previously described to play an important role in mCRPC. Indeed, ELOVL7 has been found overexpressed in high grade PCa and regulated by the androgen pathway through SREBP1[200]. Its inhibition significantly affects *de novo* androgen synthesis [200]. It has been reported that the AR cistrome is largely retained in the CRPC stage, in particular the AR-activated lipid synthesis genes including ELOVL5/7[224]. Furthermore, silencing the expression of ELOVL7, leads to the regression of CRPC xenograft tumors[224].

LPIN1 (lipin 1, see 1.2.2.b)**, LRP8** (LDL Receptor Related Protein 8), **DHCR7** (7-Dehydrocholesterol Reductase), **HMGCS1** (3-Hydroxy-3-Methylglutaryl-CoA Synthase 1).

SCD (Stearoyl-CoA Desaturase, see 1.2.4)**, ACOT1** (Acyl-CoA Thioesterase 1)**, ACY3** (Aminoacylase 3)**, MMAB** (Metabolism of Cobalamin Associated B).

PLCH2 (Phospholipase C Eta 2) and **MTMR14** (Myotubularin Related Protein 14) are involved in the Inositol phosphate metabolism. **SEH1L** (SEH1 Like Nucleoporin) is involved in Interferon gamma signalling. **PTPRJ** (Protein Tyrosine Phosphatase Receptor Type J) is involved in innate and immune system. **ALCAM** (Activated Leukocyte Cell Adhesion Molecule) is involved in Embryonic and Induced Pluripotent Stem Cell Differentiation Pathways and Lineage-specific Markers. **MALAT1** (Metastasis Associated Lung Adenocarcinoma Transcript 1). **KRT19** (Keratin 19) is a type I cytokeratin widely used as a breast cancer biomarker (see Introduction, paragraph 1.3.2).

FGFBP1 (Fibroblast Growth Factor Binding Protein 1) encodes a secreted fibroblast growth factor carrier protein. FGFBP1 plays a critical role in cell proliferation, differentiation and migration by binding to fibroblast growth factors and potentiating their biological effects on target cells[328]. It has also been involved in tumor growth as an angiogenic switch molecule, and its expression has been associated with several types of cancer including pancreatic and colorectal adenocarcinoma[328].

LOC105374167 is an uncharacterised RNA Gene affiliated with the ncRNA class. Our group has described it as the SREBP1 binding site core enhancer 1 of Keratin 80. Keratin 80 is a largely unknown type II keratin, characterised for the first time by Langbein 2010[279]. We aimed to further investigate the functions of this particular keratin in breast cancer since previous findings from our group already revealed that type II keratins are one of the most hyperactivated when comparing parental untreated non-invasive BCa cells with the invasive resistant derived ones. Results are presented in the next chapter.

Following pages:

Table 3.3: Differential binding analysis. MCF7 RM vs LTED

Table 3.4: Differential binding analysis. MCF7 RM vs LNCaP

Table 3.5: Differential binding analysis. T47D vs LNCaP

Table 3.6: Differential binding analysis. LNCaP RM vs LNCaP95

Table 3.7: Differential binding analysis. LNCaP RM vs all BCa cell lines

Table 3.8: Differential binding analysis. LNCaP vs LTED

MCF7 RM vs LTED

chrom	Region		Closest Gene			Related Pathways	FDR
	start	end	Gene	nearestGene_distance	Description		
chr17	8057840	8055859	ALOX15B	17810	Arachidonate 15-Lipoxygenase Type B	Arachidonic acid metabolism	0.00019548
chrX	154547861	154546333	G6PD	-79	Glucose-6-Phosphate Dehydrogenase	Pentose phosphate pathway	0.02387298
chr10	437830	436012	DIP2C	-62579	Disco Interacting Protein 2 Homolog C	Unknown	0.044542926
chr14	61696655	61694432	HIF1A	143	Hypoxia Inducible Factor 1 Subunit Alpha	Adipogenesis	0.044542926
chr6	87773894	87772519	LOC101928911	-11653	Uncharacterised	Unknown	0.044542926

Table 3.3: Differential binding analysis. MCF7 RM vs LTED

153

MCF7 RM vs LNCaP						
	Region		Closest Gene			
chrom	start	end	Gene	nearestGene_distance	Description	Related Pathways
chr17	8055859	8057840	ALOX15B	17810	Arachidonate 15-Lipoxygenase Type B	Arachidonic acid metabolism
chr6	53358463	53360327	ELOVL5	-10216	ELOVL Fatty Acid Elongase 5	Fatty Acid Metabolism
chr17	37358117	37360327	ACACA	-87	Acetyl-CoA Carboxylase Alpha	Fatty Acid Metabolism
chr11	61813900	61816124	MIR1908	228	MicroRNA 1908/FADS1/FADS2	Regulation of lipid metabolism by Peroxisome proliferator-activated receptor alpha (PPARalpha).
chr5	129241457	129242519	LOC1079863	-18526	Uncharacterised	Unknown
chr10	72330832	72333361	DNAJB12	22904	DnaJ Heat Shock Protein Family (Hsp40) Member B12	Protein processing in endoplasmic reticulum
chr14	61694432	61696655	HIF1A	143	Hypoxia Inducible Factor 1 Subunit Alpha	Adipogenesis
chr19	18281233	18282604	MIR3188	-158	MicroRNA 3188/JUND JunD Proto-Oncogene, AP-1 Transcription Factor Subunit	IL-1 Family Signaling Pathways
chr22	41832221	41833894	SREBF2	-21	Sterol Regulatory Element-Binding Protein 2	Sterol Regulatory Element-Binding Proteins (SREBP) signalling
chr1	11805063	11806811	MTHFR	57	Methylenetetrahydrofolate Reductase	Metabolism of water-soluble vitamins and cofactors
chr12	16346958	16348209	MGST1	-59	Microsomal Glutathione S-Transferase 1	Cytochrome P450 and Metabolism
chr12	112012102	112014905	ERP29	156	Endoplasmic Reticulum Protein 29	Protein processing in endoplasmic reticulum
chr7	92134029	92135398	CYP51A1	32	Cytochrome P450 Family 51 Subfamily A Member 1	Sterol Regulatory Element-Binding Proteins (SREBP) signalling
chr16	66880091	66881361	PDP2	220	Pyruvate Dehyrogenase Phosphatase Catalytic Subunit 2	Pyruvate metabolism and Citric Acid (TCA) cycle
chr7	26864274	26865477	SKAP2	-132	Src Kinase Associated Phosphoprotein 2	Cell junction organization
chr10	436012	437830	DIP2C	-62579	Disco Interacting Protein 2 Homolog C	Unknown
chr3	4977326	4980278	BHLHE40	-609	Basic Helix-Loop-Helix Family Member E40	HIF-1-alpha transcription factor network
chr4	110193414	110194568	LOC1079863	-2100	Uncharacterised	Unknown
chr11	17207514	17208733	PIK3C2A	-127	Phosphatidylinositol-4-Phosphate 3-Kinase Catalytic Subunit Type 2 Alpha	Vesicle-mediated transport
chr8	29520969	29522217	LOC1053793	-7650	Uncharacterised	Unknown
chr22	30212241	30213334	LOC1053729	-5653	Uncharacterised	Unknown
chr15	67078935	67080194	SMAD3	13708	SMAD Family Member 3	Factors and pathways affecting insulin-like growth factor (IGF1)-Akt signaling and Adipogenesis

Table 3.4: Differential binding analysis. MCF7 RM vs LNCaP

T47D vs LNCaP

	Region		Closest Gene			
	start	end	nearestGene	nearestGene_distance	Description	Related Pathways
chr10	100345834	100347988	SCD	-103	Stearoyl-CoA Desaturase	Fatty acid metabolism
chr14	73536313	73537796	ACOT1	-169	Acyl-CoA Thioesterase 1	Fatty Acyl-CoA Biosynthesis
chr17	26603715	26604281	LOC1053717	485243	Uncharacterized	Unknown
chr16	10651821	10653906	TEKT5	28555	Tektin 5	Unknown
chr17	41532592	41533924	KRT19	-4869	Keratin 19	Tumoural biomarker
chr11	67649716	67650699	ACY3	452	Aminoacylase 3	Cytochrome P450 and Metabolism
chr12	109572702	109574523	MMAB	-59	Metabolism Of Cobalamin Associated B	Metabolism of water-soluble vitamins and cofactors
chr6	53358463	53360327	ELOVL5	-10216	ELOVL Fatty Acid Elongase 5	Fatty Acid Metabolism

Table 3.5: Differential binding analysis. T47D vs LNCaP

LNCap RM vs LNCaP95							
Region			Closest Gene				
chrom	start	end	nearestGene	nearestGene_distance	Description	Related Pathways	FDR
chr1	3701595	3702284	TP73	3894	Tumor Protein P73	TP53 Regulates Transcription of Cell Death Genes	0.0016026
chr1	223727847	223728754	LOC1053730	12147			0.0016026
chr7	156099406	156100101	LOC1053756	46236			0.00034318
chr1	53374964	53375668	LOC1079849	-6953			0.00053106
chr7	101679761	101680812	MYL10	-50990	Myosin Light Chain 10	Regulation of actin cytoskeleton	0.00060214
chr6	107964568	107965363	SEC63	-6687	SEC63 Homolog, Protein Translocation Regulator	Unfolded Protein Response (UPR) and Protein processing in endoplasmic reticulum	0.00060214
chr1	12674619	12675330	AADACL4	10385	Arylacetamide Deacetylase Like 4	Unknown	0.00344803
chr9	6680867	6682176	LOC1053759	-22832			0.00344803
chr7	89807413	89808126	LOC1079868	-246364			0.00350401
chr17	79853878	79854894	CBX4	-14972	Chromobox 4	C-MYB transcription factor network	0.00495718
chr10	132168221	132169289	DPYSL4	-16227			0.00816566
chr1	55912980	55913635	LOC1053787	31651			0.01004677
chr7	45087804	45088750	NACAD	617	NAC Alpha Domain Containing	Unknown	0.01004677
chr11	20939995	20940975	LOC1053765	-13184			0.01573239
chr5	123114946	123115875	PRDM6	16917	PR/SET Domain 6	Lysine degradation	0.01677155
chr16	1187938	1188712	CACNA1H	-7246	Calcium Voltage-Gated Channel Subunit Alpha1 H	TCR Signaling	0.01677155
chr1	212118809	212119654	MIR3122	41619			0.01848105
chr17	1720913	1721906	WDR81	-3130	WD Repeat Domain 81	Unknown	0.01848105
chr12	112252829	112253648	MIR6861	-89917			0.02524863
chr13	17337215	17337868	FAM230C	-857755			0.03365352
chr15	100448972	100449720	CERS3-AS1	76408	CERS3 Antisense RNA 1	Unknown	0.03458412

Table 3.6: Differential binding analysis. LNCaP RM vs LNCaP95

LNCap RM vs All Bca cells						
Region			**Closest Gene**			
chrom	start	end	nearestGene	nearestGene_distance	Description	Related Pathways
chr11	61827457	61829123	FADS2	258	Fatty Acid Desaturase 2	Fatty acid metabolism
chr1	3701595	3702284	TP73	3894	Tumor Protein P73	TP53 Regulates Transcription of Cell Death Genes
chr1	11737049	11738118	AGTRAP	1477	Angiotensin II Receptor Associated Protein	Oncogenic MAPK signaling
chr4	15905365	15906860	FGFBP1	32628	Fibroblast Growth Factor Binding Protein 1	Signaling by FGFR2
chr7	101679761	101680812	MYL10	-50990	Myosin Light Chain 10	Regulation of actin cytoskeleton
chr3	154226864	154227961	LOC1053741	24107	Keratin 80	

Table 3.7: Differential binding analysis. LNCaP RM vs all BCa cell lines

	Region		Closest Gene			Related Pathways	FDR
chrom	start	end	nearestGene	nearestGene_distance	Description		
chr5	173329185	173330903	STC2	-541	Stanniocalcin 2	Unknown	0.00015199
chr11	61827457	61829123	FADS2	258	Fatty Acid Desaturase 2	Fatty acid metabolism	0.00017494
chr18	12947361	12948637	SEH1L	6	SEH1 Like Nucleoporin	Interferon gamma signaling	0.00017494
chr3	105367982	105369547	ALCAM	2052	Activated Leukocyte Cell Adhesion Molecule	Embryonic and Induced Pluripotent Stem Cell Differentiation Pathways and Lineage-specific Markers.	0.00017494
chr1	3701595	3702284	TP73	3894	Tumor Protein P73	TP53 Regulates Transcription of Cell Death Genes	0.00022756
chr12	52147622	52149431	LOC1053697	-29487	Uncharacterised	Unknown	0.00022756
chr1	11737049	11738118	AGTRAP	1477	Angiotensin II Receptor Associated Protein	Oncogenic MAPK signaling	0.00041705
chr2	11743368	11744923	LPIN1	-2450	LPIN 1	Adipogenesis	0.00041705
chr10	100345834	100347988	SCD	-103	Stearoyl-CoA Desaturase	Fatty acid metabolism	0.00041705
chr4	15905365	15906860	FGFBP1	32628	Fibroblast Growth Factor Binding Protein 1	Signaling by FGFR2	0.00058958
chr21	46228039	46230232	MCM3AP-AS	-95	MCM3AP Antisense RNA 1	Unknown	0.00096191
chr1	2484188	2485123	PLCH2	8341	Phospholipase C Eta 2	Inositol phosphate metabolism	0.00139291
chr7	101679761	101680812	MYL10	-50990	Myosin Light Chain 10	Regulation of actin cytoskeleton	0.00442968
chr1	9244211	9245840	H6PD	-211	Hexose-6-Phosphate Dehydrogenase/Glucose 1-Dehydrogenase	Pentose phosphate pathway	0.00623502
chr1	1577188	1578006	SSU72	-2715			0.00635899
chr3	9648791	9649650	MTMR14	-212	Myotubularin Related Protein 14	superpathway of inositol phosphate compounds	0.00635899
chr1	53325845	53327068	LRP8	943	LDL Receptor Related Protein 8	Metabolism of fat-soluble vitamins	0.00635899
chr11	71447495	71449624	DHCR7	-128	7-Dehydrocholesterol Reductase	cholesterol biosynthesis II (via 24,25-dihydrolanosterol)	0.00635899
chr12	65168894	65170170	LEMD3	-38			0.00635899
chr1	212118809	212119654	MIR3122	41619			0.00665677
chr5	43312491	43314562	HMGCS1	-14	3-Hydroxy-3-Methylglutaryl-CoA Synthase 1	Sterol Regulatory Element-Binding Proteins (SREBP) signalling	0.00689166
chr12	56629600	56630983	BAZ2A	28			0.00899808
chr17	1720913	1721906	WDR81	-3130			0.00931565
chr5	16666010	16667142	MYO10	48044			0.00993521
chr11	48011639	48012837	PTPRJ	-879	Protein Tyrosine Phosphatase Receptor Type J	Innate Immune System	0.00993521
chr3	154226864	154227961	LOC1053741	24107			0.00993521
chr22	20702507	20703218	TMEM191A	1749			0.00993521
chr1	6153596	6155568	LOC1053766	22820			0.00993521
chr5	147782449	147783120	JAKMIP2	-28			0.00993521
chr1	234974603	234976286	LOC1079853	5398			0.00993521
chr7	158594726	158595766	MIR5707	3631			0.01052923
chr17	37518461	37519262	DUSP14	28442			0.01270934
chr7	158391535	158392402	LOC1079868	31620			0.01441632
chr22	50175513	50176665	PANX2	5359	Pannexin 2	Electric Transmission Across Gap Junctions	0.01503166
chr1	235073063	235074502	LOC1079857	2585			0.01520162
chr16	80540340	80541571	DYNLRB2	4			0.02150264
chr12	85089819	85090517	LRRIQ1	53788			0.02802241
chr12	110677057	110678661	HVCN1	11298			0.02802241
chr1	2887187	2888215	TTC34	-97964			0.02802241
chr1	3761760	3762798	CCDC27	9568			0.02802241
chr18	74269055	74269784	CYB5A	22561			0.02802241
chr1	3686415	3687735	TP73	-3596			0.02802241
chr5	123114946	123115875	PRDM6	16917			0.02892957
chr5	12674619	12675330	AADACL4	10385			0.02921459
chr17	6534764	6535570	PITPNM3	21330			0.03020306
chr12	1444551355	1444552468	LOC1053795	-320			0.0309654
chr1	46132310	46133775	PIK3R3	-6			0.03441521
chr11	65506957	65509008	MALAT1	10304	Metastasis Associated Lung Adenocarcinoma Transcript 1		0.03493047
chr6	143113375	143114065	AIG1	52860			0.03645847
chr16	2857874	2859310	PRSS22	38			0.03645847
chr1	3312279	3313452	LOC1053786	2295			0.03772402
chr1	223727847	223728754	LOC1053730	12147			0.03983467
chr12	62519390	62520261	MON2	20907			0.04636581

Table 3.8: Differential binding analysis. LNCaP vs LTED

3. 9 Conclusions

In this chapter, I investigated the role of SREBP1 in hormone-dependent cancer developing endocrine therapy resistance through hormone independence.

ET-driven epigenetic reprogramming of breast cancer metabolism leads to the activation of *de novo* CB pathway. SREBP1-driven cellular metabolic changes were investigated by using a vast array of techniques both *in vitro* and *in vivo*. RT-qPCR time course studies show that short-term oestradiol deprivation is not sufficient to induce SREBP1-dependent CB activation that is instead consistently upregulated in resistant long-term hormone deprived cell lines. We hypothesised that if SREBP1 target activation is enhanced by hormone removal, this does not occur at an early stage during treatment, but it may rather be a cellular program that is selected/elicited after a long time from starting the endocrine treatment. Metabolic analysis shows the activation of the SREBP1 pathway may not be dependent on glucose supply, at least in our cell line model. Indeed, we observe a switch of metabolic dependency in drug-resistant ERα breast cancer cells from a glucose- to a glutamine-dependent phenotype and this is found to be associated with increased *de novo* cholesterol and fatty acid synthesis. Whether and how SREBP1 activation may be directly driven by a cellular switch of metabolic dependency needs to be worked out mechanistically. In order to investigate SREBP1 recruitment to the chromatin, I performed SREBP1 ChIP-Seq in breast and prostate cancer cell lines. Optimization of ChIP protocol allowed for an efficient pull-down of SREBP1-bound chromatin in order to perform downstream analyses. We found that SREBP1 binding profiles distinguish cancer cells based on the tissue of origin (BCa vs PCa cells). Furthermore, a difference in SREBP1 binding is found between parental and long-term starved derived cell lines in MCF7. Our data confirm a significant co-occurrence between AR and SREBP1 binding sites in PCa and also suggest, for the first time, a possible crosstalk between SREBP1 and the ERα in BCa on chromatin. Lastly, differential binding analysis identifies novel SREBP1 non-canonical target genes in cancer cells. Future work aiming at further understanding the role played by SREBP1 in resistance development overall points towards four main directions:

1. exploring SREBP1 non-canonical targets
2. characterizing SREBP1 signalling transduction
3. profiling cellular metabolic switches
4. testing drugs for translational applications

"Resistance is futile"
The Borg Collective

Chapter 4: Results II - SREBP1 drives cell-autonomous cytoskeletal changes by Keratin 80 remodeling during ERα breast cancer progression

4.1 SREBP1 regulates KRT80 by binding to the core enhancer

Fine mapping with DHS-seq revealed that KRT80 is already accessible in MCF7. However, when looking at potential transcription factor footprints, digital mapping of chromatin accessible loci suggested different occupancy rates (Fig. 4.1A, B). In particular, the appearance of a SREBP1 footprint within the core enhancer E1 was found to be specific for LTEDs (Fig. 4.1A, B). Meta-analysis of available SREBP1 ChIP-seq data reveal that SREBP1 binds the *KRT80* core-E1 locus in lung adenocarcinoma A549 models but not in Hep2G liver cancer cells, leukemia K562 cells nor glioblastoma GB12878 cells. In agreement, only A549 transcribed significant amount of KRT80 while the E1-*KRT80* activity in Hep2G and all blood cancer cells lines is null[328].

To directly test if SREBP1 binds *KRT80* enhancer, I performed SREBP1 ChIP-seq in both untreated MCF7 and resistant LTED cells. As discussed in the previous chapter, SREBP1 was found at the promoter of genes involved in the cholesterol biosynthesis (Fig. 4.1C), especially in cholesterol driven LTED cells. In the LTEDs, a strong SREBP1 recruitment was found at the KRT80 core E1 enhancer, suggesting a role for SREBP1 in regulating KRT80 expression in oestradiol-deprived breast cancer cells (Fig. 4.1C). To confirm these results, SREBP1 was mapped in other independent BCa cell lines (T47D and T47D-LTED) and observed SREBP1 binding to the same core E1 loci (Fig. 4.1D). These data lead to the hypothesis that cholesterol biosynthesis and KRT80 activation are co-regulated by SREPB1 in hormone-deprived cell lines. This co-regulation appears to be conserved in ERα-positive BCa patients as shown by the strong correlation between the RNA of cholesterol biosynthesis genes and KRT80 transcripts (Fig. 4.1E). To test if KRT80 activation is mediated by SREBP1 regulation, I stably silenced SREBP1 in LTED cells using different short-hairpin RNA, namely SREBP1 sh1, sh2 and sh control (shctrl). mRNA levels of KRT80 and SREBP1 showed that in LTED SREBP1-knockdown cells there is a significant reduction of KRT80 compared to the control ones (Fig. 4.1F). Western blotting further confirmed at a protein level that constitutive depletion of SREBP1 leads to complete loss of K80 protein informing that K80 expression may be dependent upon SREBP1 binding to its enhancer (Fig. 4.1G). Intriguingly, while the majority of the DNA sequence within the core E1 enhancer is extremely conserved throughout

evolution, the footprint containing the SREBP1 motif is not under significant evolutionary constraint[328], leading to the speculation that the link between cholesterol biosynthesis and KRT80 activation might have developed relatively recently.

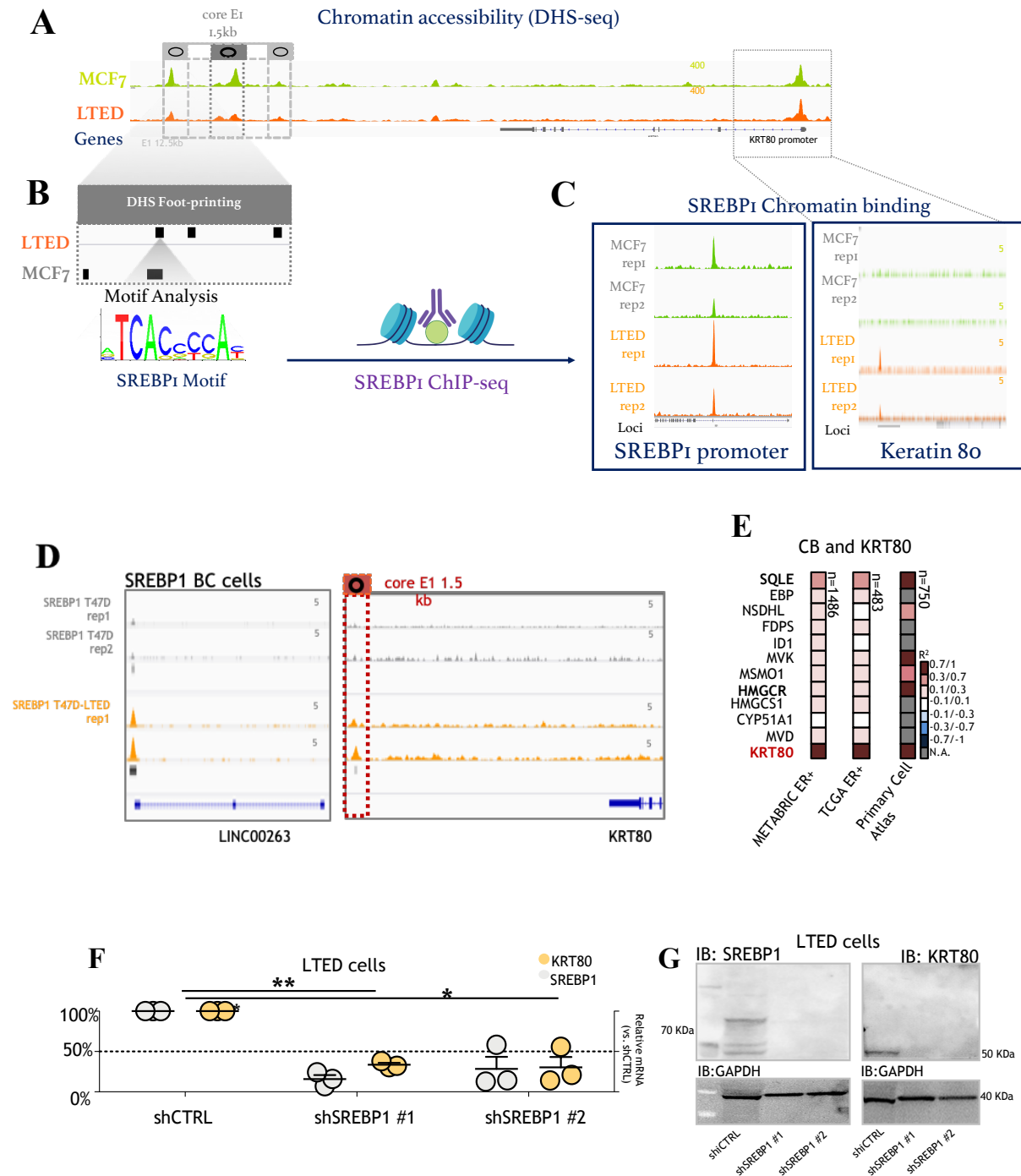**Figure 4.1:** De novo SREBP1 binding at *KRT80* enhancer drives KRT80. A) Open chromatin profiling via DHS-seq in MCF7 and LTED cells near the *KRT80* locus. B) Digital Foot-printing analysis shows differential occupancy status within the E1 *KRT80* enhancer. C) ChIP-seq analysis for SREBP1 at the E1 core enhancers in MCF7 and LTED cell lines. D) ChIP-seq signal at LINC00263 (also named OLMALINC) and *KRT80* E1

enhancer is shown for T47D and T47D-LTED cells. E) Metanalysis of KRT80 transcriptional correlation with SREBP1 canonical targets in ER breast cancer samples (performed by L.M.). F) Stable shSREBP1 silencing in LTED cells using two independent shRNA. Individual biological replicates are shown. Lines represent means and SD. Asterisk represent significant difference at p < 0.05. G) Stable shSREBP1 LTED cells were assessed for SREBP1 and K80 protein levels (Modified from Perone 2019, figures originally made by Luca Magnani).

4.2 Long-term hormone deprivation reprograms the Keratin Type II locus

AI treatment is standard of care for breast cancer, yet BCa cells frequently display drug-resistance and stronger metastatic potential at relapse, suggesting that chronic exposure to endocrine treatment (ET) might contribute in shaping the invasive potential[53]. The mechanism/s, order of events and molecular players mediating these phenomena are not well understood but it is likely to involve cytoskeletal re-arrangements as they are essential for cancer invasion and metastasis[256, 265, 268, 272, 275]. One possibility is that ET might indirectly promote invasive behaviors by selecting for interrelated phenotypes during tumour evolution[53,329,331]. Alternatively, AI treatment may directly contribute to the activation of invasive transcriptional programs. Chronic exposure to ET leads to coordinated activation and repression of regulatory regions such as enhancers and promoters[331]. This is shown by global changes in the localization of epigenetic marks histone 3 lysine 27 acetylation (H3K27ac) and histone 3 lysine 4 mono- and di-methylation (H3K4me1-2)[53,329,331].

Previous evidence suggest that the human genome is organized in three-dimensional compartments, known as topological associated domains (TADs), capable of restricting enhancer-promoter interactions[330]. We have previously observed coordinated epigenetic changes at the TAD level between untreated and chronically treated cells[53]. One of the most significantly hyper-acetylated (H3K27ac, top 5%) TAD in invasive resistant BCa cell lines compared to untreated were characterized by the presence of type II keratins[53]. These changes are not likely to originate from TAD re-organization, as TADs structures are generally conserved throughout differentiation or cancer transformation. Our data led to the hypothesis that epigenetic reprogramming might contribute to significant switches in the keratin organization in AI resistant cells[53]. To test this, we analyzed the transcriptional changes for all genes mapping to the type II Keratins TAD using RNA-seq (Fig. 4.2 A). These data highlight that only few keratins are differentially expressed in AI-resistant LTED cells when compared to parental, untreated MCF7 cells (Fig. 4.2 A). In particular, further targeted validation confirms that KRT80 is the only transcript consistently activated in invasive, resistant LTED

models and this is matched by an increase in protein levels in both MCF7 and T47D models (Fig. 4.2 B). Thus, in order to confirm the significant difference of H3K27ac within the *KRT80* enhancer E1 between MCF7 and LTED, I conducted ChIP-qPCR using antibodies anti-histone marks H3K27ac, H3K4me2 and H3K4me1 (Fig. 4.2 D). The results supported the hypothesis of an increased activation of *KRT80* enhancer specifically in LTED cells (Fig. 4.2 D).

Using single-cell RNA-seq, KRT80 transcripts were measured before and after acute oestradiol starvation (48-hours). The proportion of KRT80 positive cells significantly increased after hormone deprivation suggesting that it might be driven by transcriptional activation rather than selection of KRT80 positive clones (Fig. 4.2 E). To test this hypothesis, I performed continuous imaging proliferation assay by live-tracking cells to monitor their growth during the initial 48 hours of hormone starvation. Analysis indicated that there is no significant change in cell number, and more importantly, there is no significant cell death suggesting that cells stall within this time frame (Fig. 4.2 F). Therefore, over 90% of the cells measured in the single-cell RNA-seq experiment should be conserved before and after treatment. These results also reinforce the possibility that KRT80 is transcriptionally upregulated at the single cell level rather than increased through a selection process. Next, to validate scRNA-seq analysis, I conducted single-cell RNA Fluorescence in Situ Hybridization (sc-RNA FISH) for KRT80 and SREBP1 in MCF7 and LTED cells and found increased KRT80 transcript levels in LTEDs (Fig. 4.2 G). Then, I examined KRT80 expression in a second line of ERα breast cancer cells (T47D) and in their AI-resistant derivatives (T47D-LTED) and again found a significant increase in KRT80 expression in drug-resistant cells (Fig 4.2 H). Considering that Keratin 80 is widely expressed in epithelial cells, we expanded our analysis to several other breast cancer cell lines found elevated KRT80 transcripts independently from ERα status[329]. Interestingly, levels are significantly lower in breast myoepithelial cells with no oncogenic potential (MCF10A) and KRT80 transcripts are not upregulated in non-invasive Tamoxifen resistant MCF7-T cells (Fig. 4.2 H). Immunohistochemistry (IHC) analysis of two independent clinical datasets of matched primary and metastatic biopsies from endocrine treated patients confirmed that K80 positive cells significantly increase after endocrine therapy, particularly after pure AI treatment while showing a trend in Tamoxifen-treated patients (Fig. 4.2 I).

Collectively, these data support the previous observation that chronic AI treatment drive invasive behavior in hormone independent breast cancer cells by inducing upregulation of KRT80.

A

Hi-C interaction map (Chr:12)

Non-invasive/untreated
Non-invasive/resistant
Invasive/resistant

GM12878 10kb resolution

Type II Keratins TAD

H3K27ac RPKM-TAD

Global
Type II Keratins

MCF7
MCF7T
MCF7F
LTED
LTEDT
LTEDF

50950000   54180000

Untreated
Non invasive

Endocrine
therapy-resistant

Endocrine
therapy-resistant
Invasive

MCF7
MCF7T
MCF7F
LTED
LTEDT
LTEDF

Log2 (FPKM)

NR4A1 C12orf44 **KRT80** LOC283402 LOC283404 KRT86 KRT81 KRT7 KRT83 KRT85 KRT84 KRT82 KRT75 KRT6B KRT6C KRT6A KRT5 KRT71 KRT74 KRT72 KRT73 KRT2 KRT1 KRT77

B

Relative mRNA Fold Change (MCF7)

MCF7
MCF7T
MCF7F
LTED
LTEDT
LTEDF

NR4A1   ATG101   **KRT80**   KRT81   KRT7   KRT86

C

IB:KRT80

MCF7 MCF7T MCF7F LTED LTEDT LTEDF   MCF10A T47D T47D-LTED

IB:β-Actin

D

ChIP-qPCR
E1 Enhancer

○ H3K27a
□ H3K4me1
△ H3K4me2

Fold Enrichment *vs.* Internal Negative control

***

MCF7 LTED   MCF7 LTED   MCF7 LTED

E

Single-Cell RNA-seq

Cell Distribution (~5000 cells)

MCF7

MCF7 -E2 (48hrs.)

n>2000 p<1.7e-65

10.8% KRT80 +

39.9% KRT80 +

Normalized Expression (Tags per Cell)

F

Live Cell Imaging Count

■ +E2
■ -E2

Cell Number (10³)

Time (hours)

G

Single-Cell RNA-FISH

MCF7

LTED

3 representative field of views

SREBP1 probes
KRT80 probes

H

RT-qPCR: KRT80

Relative mRNA Fold Change (MCF10A)

**
*
**
*

MCF10A T47D T47D-LTED

I

IHC KRT80
Matched Biopsies (pre and post treatment)

Endocrine therapies
p=0.014

Quick Score

Dataset 1- ICL: Primary vs. Metastatic matched biopsies

AI only
p=0.045

H Score

Dataset 2-IEO: Primary vs. Metastatic matched biopsies

TAM only
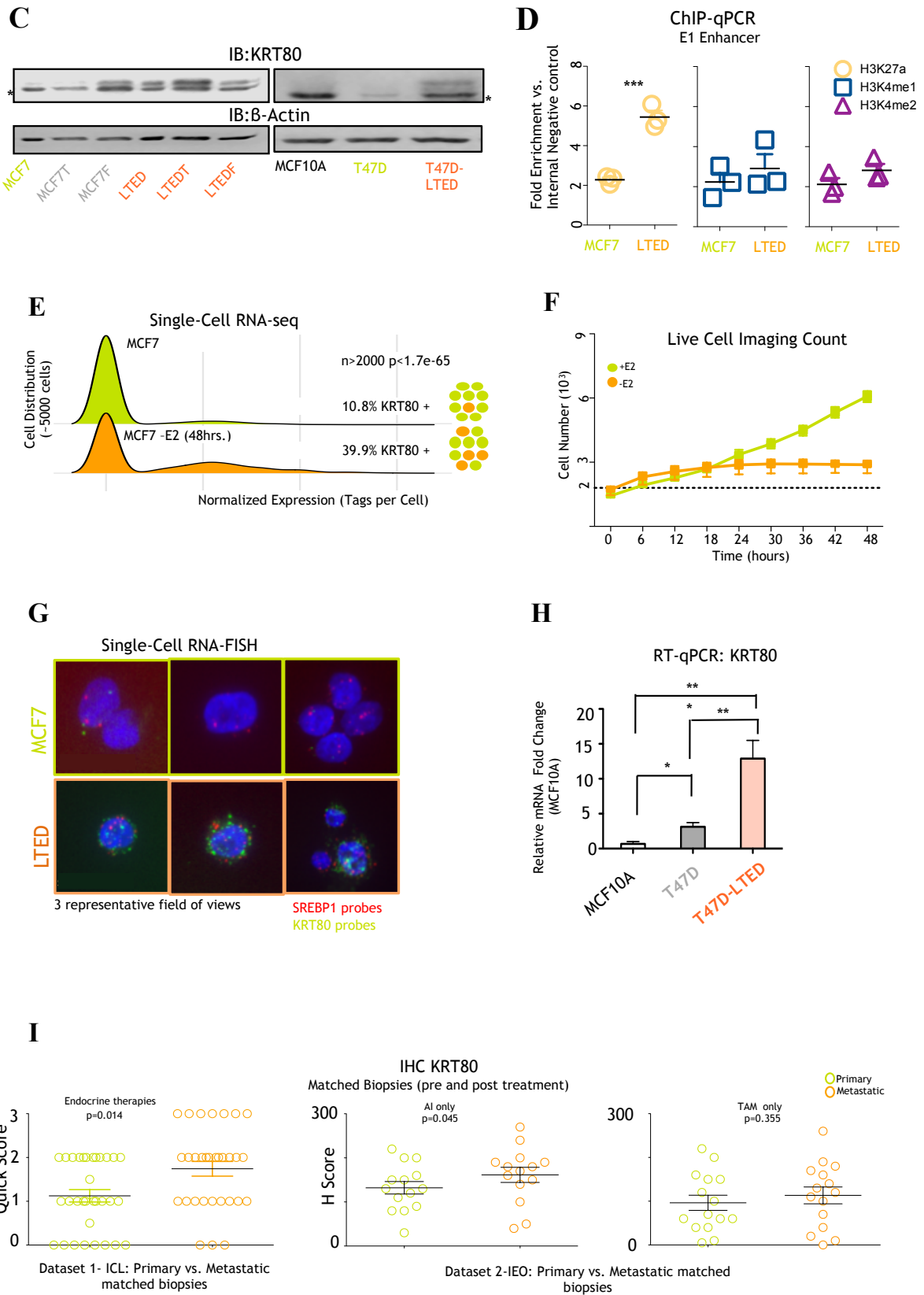p=0.355

○ Primary
○ Metastatic

Figure 4.2: AI treatment induces KRT80 expression via epigenetic reprogramming. Previous page: A) Hi-C 3D interaction in GM12878 cells. Bars represent the normalized median change in H3K27ac within the Type II-

Keratin TAD compared to the overall change in H3K27ac between parental MCF7 cells (green) and drug-resistant non-invasive (grey) and drug-resistant invasive (orange) counterparts. The bottom heatmap shows the normalized expression of RNA-seq data for protein coding genes within the Type II-Keratin in all BCa cell lines. B) Targeted RT-qPCR for genes that were found differentially regulated in RNA-seq analysis. Bars and lines represent mean and SD of three independent measurements. C) KRT80 protein levels in MCF7 and additional independent models of invasive drug-resistant breast cancer cell lines (experiments in A-C perfomed by A.R.M.). The asterisk represents an unspecific band. D) Targeted ChIP-qPCR for the E1 enhancer locus using H3K4me1, H3K4me2 and H3K27ac antibodies. Individual biological replicates, mean and SD are shown. Asterisks represent significance at the $p < 0.001$ level. E) Population level single-cell RNA-seq data for KRT80 expression are shown. KRT80 was identified in 10.8% of MCF7 cultured in estrogen rich media and in 39.9% of MCF7 deprived of estrogen for 48 h. The distribution of the two set of data was compared using a Fisher exact test (analysis performed by Iros Barozzi). Experiments were run comparing cells within 48 h in absence of major cell division/apoptosis (by Sung Pil Hong). F) Live- imaging cell counts of mate-labeled MCF7 cells grown in presence or absence of estrogen for 48 h. Dotted line represents an ideal stalling dynamic in cell number during the time of the assay. Mean and SD of three independent counts are shown. G) Representative single-molecule, single cell RNA-FISH for SREBP1 (red) and KRT80 (green) in MCF7 and LTED cells. H) KRT80 transcripts were also quantified in non-tumorigenic breast MCF10A cell lines and in drug sensitive and drug-resistant clones of T47D breast cancer cell lines. Bars and error bars represent the average and the SD of 3 biological replicates. Asterisks represent significance levels at *, **, *** P<0.05, 0.01 and 0.001 respectively (experiment perfomed by A.R.M.). I) Matched clinical specimens from breast cancer patients show an increase in KRT80 positive cells following mono-treatment (Dataset 1 performed by C.I., S.S. at ICL: Imperial College London, UK) or sequential treatment with aromatase inhibitors (Dataset 2 performed by G. P. at IEO: Istituto Europeo di Oncologia, Milan, Italy). Similar results were not significant in Tamoxifen-only treated patients. MCF7T, MCF7F, LTEDT, and LTEDF: MCF7 and LTED cells resistant to Tamoxifen or to Fulvestrant respectively; RPKM: reads per kilobase million; FPKM: Fragments Per Kilobase of transcript per Million mapped reads (Modified from Perone 2019, figures originally made by Luca Magnani).

4.3 Keratin 80 levels and localization dynamically change during ERα BCa progression in clinical samples

To further explain the potential for KRT80 in breast cancer progression and further explore the findings from the previous section, we next characterized K80 expression patterns in both benign and cancerous patients' tissues via IHC. In contrast with other broadly characterized human keratins in breast epithelia[14,67,333] (K5, K14, K8, K18 or K19), K80 presents a polarized pattern towards the lumen within healthy ducts and lobules (Fig. 4.3). These data were substantially reproduced using two additional antibodies (Appendix II, Fig. S2 A). No staining was evident in the myoepithelial, stromal or adipocytes compartments (Fig. 4.3) in agreement with E1 activity prediction and mRNA levels in various breast compartments[329]. On the other

hand, tissues with E1 inactivity, such as ovaries, show diffuse negativity with only a very small clone of epithelial cells showing evidence of K80 positivity[329]. Additionally, ten benign breast lesions were examined including fibroadenomas, fibrocystic changes, tubular adenomas or intraductal papilloma. Interestingly, K80 exhibited the same apical granular and polarized pattern in all benign lesions (Fig. 4.3). However, in high cellularity lesions, a more disorganized pattern was observed with K80 distribution dramatically changing in advanced (grade 3) invasive carcinomas and metastatic BCa (Fig. 4.3). These data strongly suggest that K80 expression and pattern might change in concert with the augmented invasive potential of luminal breast cancer cells during cancer progression.

We reasoned that, if KRT80 underlies increased invasive potential, patients characterized by high K80 expression at diagnosis should have poorer prognosis than K80 low counterparts. Meta-analysis of all publicly available patients' derived datasets containing KRT80 RNA measurements (Affymetrix-based[335], Illumina-based METABRIC[334] and RNA-seq TCGA[336]) consistently show that high KRT80 mRNA levels are significantly associated with all poor prognostic endpoints (Overall Survival, Recurrence- and Metastasis-free survival and Post-Progression Survival, univariate and multivariate analysis, Fig. S2 B-I). This association is independent from ERα, Ki67, HER2 or gene expression biomarkers (Fig. S2 D-E). We next focused on ERα patients (METABRIC and TCGA, Fig. S2 B, G-I) and found again that patients with high KRT80 expression had significant worse prognosis. Interestingly, KRT80 expression appears to strongly stratify patients that have early relapse and shorter survival being also the only type II keratin expressed in breast cancer that is significantly associated to all clinical endpoints (Fig. S2 C). We then looked within the TCGA dataset for which patients we could extrapolate the adjuvant treatment and found that KRT80 expression can retrospectively stratify patients prior to AI but not to Tamoxifen or chemotherapy (Fig. S2 I) suggesting that high KRT80 might indicate the existence of cells with *de novo* SREBP1 activity in which AI treatment might be ineffective.

Analysis of *in vivo* data provides clinical evidence suggesting that KRT80 upregulation during breast cancer progression might contribute to the development of invasive potential.

Figure 4.3: (Following page) K80 dynamics in treated progressing breast cancer patients. Immunocytochemistry (IHC) analyses show changes in KRT80 protein distribution. KRT80 was imaged using IHC in a series of human samples collected at Charing Cross Hospital Imperial College NHS Trust (ICL London, UK, performed by C.I., C.D., S.S.). Tissues were collected to cover a large spectrum of benign and malignant lesions including metastatic samples from breast cancer patients. Yellow arrows in the bottom panels highlight cells with KRT80 expanded cytoplasmic staining (Modified from Perone 2019, figures originally made by Luca Magnani).

Progression

Invasive

Metastatic lymph node

IDC Grade 3

IDC Grade 1

In Situ

LCIS

High Grade DCIS

Intermediate DCIS

Benign

Fibroadenoma / 2

Fibroadenoma

Apocrine Cyst

Normal Breast

Lobule

Duct

Dataset 3 ICL ICL: Diagnostic Material from a panel of 80 patients diagnosed with breast-related pathologies
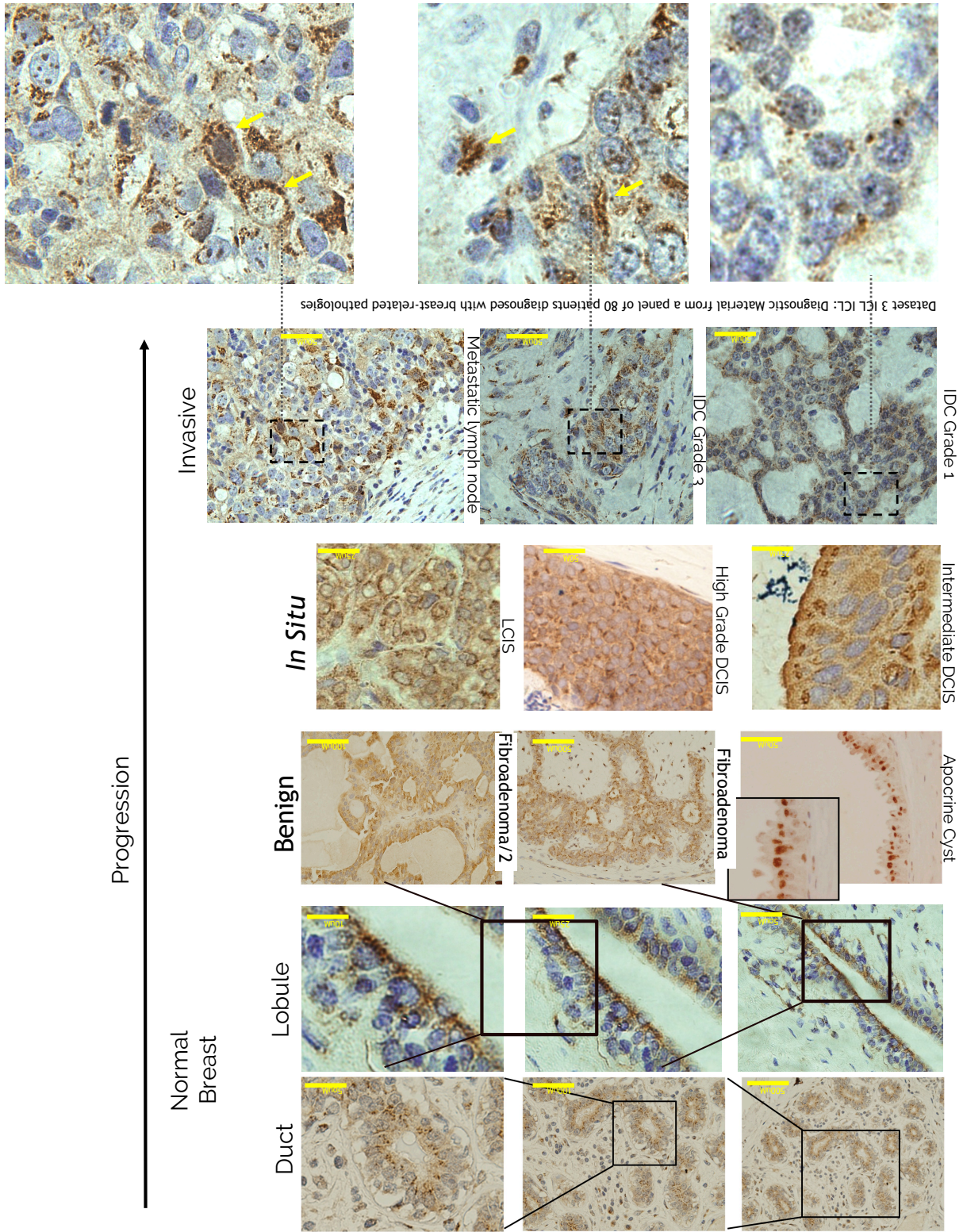
4.4 Keratin 80 identifies invasive cells within 3D structures

To test if KRT80 is essential for the invasive phenotype, its expression was modulated in non-invasive and invasive ERα breast cancer cells and confirmed KRT80 knockdown or overexpression measuring mRNA and protein levels (Fig. S3 A, B). Stable depletion of KRT80 in invasive LTED cells lead to decreased migration in 2D and invasion in 3D (Fig. 4.4 B, D). Interestingly, spheroids showed defect in cell adhesion after KRT80 depletion, with single cells detaching from the spheroid mass (Fig. 4.4 B). Conversely, stable over-expression of KRT80 significantly increased the invasive potential in otherwise non-invasive MCF7 cells, even in the absence of chronic oestradiol deprivation (Fig. 4.4 C, D). Our results show that KRT80 over-expression leads to a significant increase in spheroid size in MCF7, while KRT80 depletion greatly reduces spheroid size in LTED cells (Fig. 4.4 B-D). Immunostaining showed that K80 positive cells clustered at the invasive front in LTED spheroids (Fig. 4.4 E), a pattern reminiscent of the leading cells characterizing epithelial tumours during collective invasion[67,69]. To confirm that invasion was driven by active motion rather than proliferation at the border of the spheroids, the invasion assays was repeated using proliferation sensitive CMFDA live labelling (Fig. 3S C). At the concentration used in these experiments, fluorescence is lost in just 2-3 cell divisions. The images in figure 3S C-E depict retention of fluorescent dye in the invading cells, suggesting that these cells actively move into the Matrigel interface rather than undergoing cell division at the border. Labelled cells maintained their invasive properties while KRT80 suppression still blocked invasion (Fig. 3S C-E). These data are supported by live imaging of spheroid invasion previously performed in the same cell lines[53]. Confocal immunofluorescence imaging was used to detect the expression and distribution of K80 in untreated and treated MCF7 compared to LTED cells. Whilst K80 was minimally expressed or undetectable in MCF7s, its signal was strongly enriched in the cytoplasm of all LTED cells after treatment (Fig. S4 A). When a similar analysis was performed using MCF7 and LTED in 3D spheroid cultures, it displayed patterns with K80 enriched in cells facing the extra-cellular matrix (Fig. S4 B, C). This distribution was similar to K14 "leading cells" in basal breast carcinomas[67, 69]. Additionally, confocal microscopy shows the presence of K80 in *bona fide* invasive ultra-structures (white arrows), suggesting that K80-rich intermediate filaments might participate in active cellular movement (Fig. S4 B, C). Interestingly, K80 was also strongly expressed in the cytoplasm of circulating tumor cancer cells obtained from pleural effusion material of three patients with AI-resistant metastatic breast cancer (Fig. S4 C). In summary, highly expressed K80 in the leading cells of drug-

resistant cancer cell spheroids might play a central role in promoting active epithelial migration and invasion trough the extracellular matrix.
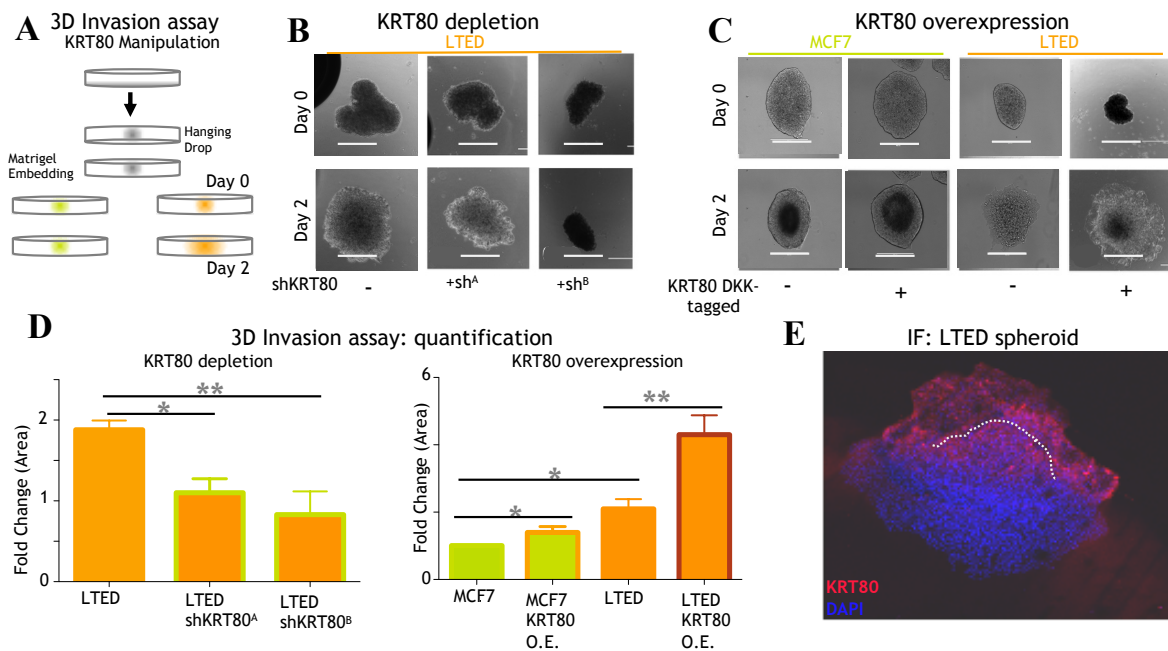

Figure 4.4: KRT80 directly promotes cell invasion. A) Design of the 3D invasion assay. Organoids were derived from treatment naive (green; MCF7) or invasive AI resistant (orange; LTED) breast cancer cells. KRT80 expression was manipulated via ectopic overexpression or sh-mediated stable depletion. Organoids were embedded in Matrigel and monitored for 48 h. B-C) Representative brightfield images of KRT80-manipulated organoids. Panels show results obtained in KRT80 depleted cells D), and in KRT80 over-expressing cells. D) (DKK-tagged KRT80). Small inset number represent normalized fold area changes of each represented experiment. Bars scale = 400 μm. E) Quantification of the area fold change in organoids KRT80 knock down or overexpressing KRT80 LTED cells in 3D invasion assay normalized to MCF7 (*p < 0.05, **p < 0.01, Student t test; n = 3 biological triplicates in which at least 4 organoids were measured). Data is presented as mean ± SD. F) Confocal microscopy of matrigel embedded invasive AI resistant LTED organoids (Experiments performed by M.F., modified from Perone 2019, figures originally made by Luca Magnani).

4.5 KRT80 directly promotes cytoskeletal re-arrangements at invasive ultra-structures

Following our initial observations, we aimed to characterize these KRT80-driven cytoskeletal changes in higher details. LTED cells can effectively invade matrigel extra-cellular matrix via the emission of filopodia and lamellipodia at the leading edge and K80 was specifically observed in cells at the forefront of invasion in 3D spheroids (Fig. 4.5 C, D). Confocal microscopy analyses informed that LTED and MCF7-KRT80 cells presented an intricate

network of K80 filaments that significantly overlap actin fibers. This K80 network was prominent in lamellipodium-like structures in leading cells (Fig. 4.5 B, C). Conversely, in KRT80[low] cells (i.e. MCF7 and LTED-sh[a]), K80 staining was punctuated and border cells presented strong cortical actin and no prominent lamellipodia (Fig. 4.5 A, D). Confocal images and relative quantitative analysis indicated that K80 expression was associated with a significant increase of F-actin at lamellipodial structures, with smaller compensating changes at the cell cortex and cytosol depending on the system (MCF7 or LTED) (Fig. S5). Together, these results suggest that the generation of a network of K80 positive filaments do not affect actin polymerization but rather reorganize the actin cytoskeleton to promote lamellipodia formation; in line with these results, cells expressing K80 presented a higher proportion of lamellipodia when compared to their K80 low counterparts (Fig. S4 C). Focal adhesion growth and maturation are tightly coupled with the forward movement of the lamellipodia[337], are associated to cell stiffness/cellular tension[275,337] and are particularly relevant in the generation of forces required for migration and invasion in complex settings. Collectively, these data show that KRT80 can directly contribute to cell invasive phenotype by reprogramming the cytoskeleton of breast cancer cells forming lamellipodia and promoting migratory phenotype.
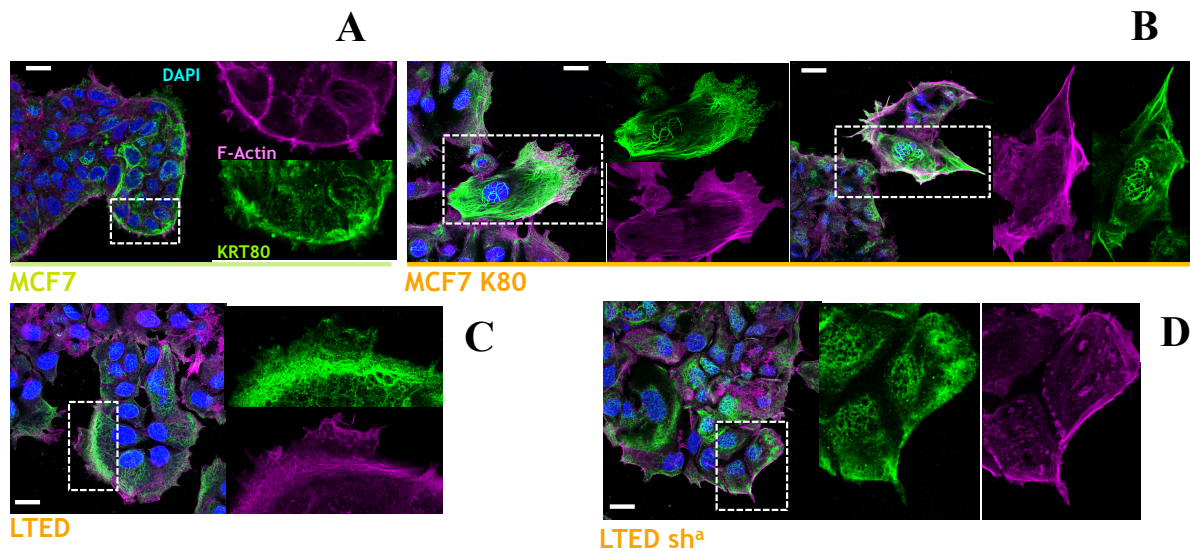
Figure 4.5: KRT80 induces invasion-associated cytoskeletal changes. Representative confocal microscopy images showing F-actin (magenta), K80 (green) and DAPI (blue) staining of A) MCF7, B) MCF7-K80, C) LTED and D) LTED-sha cells. Scale bars represent 25 μm. Zoom-up magnifications of squared areas showing F-actin (magenta), K80 (green) and DAPI (blue) staining in cells located at the border of clusters. Single channel images for F- actin and K80 are also shown. Scale bars, 10 μm (experiments, analysis and figures originally made by A.F. and F.C. at ICR London, modified from Perone 2019).

4.6 KRT80-driven mechano-transduction is independent of YAP/TAZ signaling pathway

Cell migration initiates by traction generation through reciprocal actomyosin tension and focal adhesion reinforcement, but continued motility requires adaptive cytoskeletal remodeling and adhesion release[338]. The transcriptional coactivators YAP and TAZ modulate cell mechanics by limiting cytoskeletal and focal adhesion maturation to enable persistent cell motility and 3D vasculogenesis. YAP/TAZ also establishes a transcriptional feedback axis necessary to maintain a responsive cytoskeletal equilibrium and persistent migration[339]. Emerging evidence indicates that deregulation of YAP and TAZ mediators of the Hippo pathway signalling may be a major mechanism of intrinsic and acquired resistance to various targeted and chemotherapies promoting tissue proliferation and organ growth[147]. In response to various therapies, numerous upstream signals could impinge on components of the Hippo pathway to activate YAP/TAZ. Furthermore, it was reported that YAP/TAZ activity is controlled by the SREBP/mevalonate pathway[148] by demonstrating that statin treatment inhibits YAP/TAZ nuclear localization and transcriptional responses. The authors of the study found that mevalonate–YAP/TAZ axis is required for proliferation and self-renewal of TNBC cells. In this study, YAP/TAZ activation is promoted by increased levels of mevalonic acid produced by p53-induced SREBP transcriptional activity[148].

Studies based largely on 2D culture described YAP as a universal mechano-transducer. Conversely, recent reports suggest that mechano-transduction occurs independently of YAP in breast cancer patient samples and mechanically tunable 3D cultures[340]. This work highlights the context-dependent role of YAP in mechano-transduction suggesting that YAP does not mediate mechano-transduction in breast cancer[340].

Considering these recent evidences, we wanted to investigate the role of YAP/TAZ pathway in our model system. Indeed, we reasoned that potential cytoskeletal changes induced by KRT80 manipulation might also affect migration and invasion by impinging on YAP/TAZ signalling activation (Lamar 2012, Dupont 2011). Despite evidence of elevated YAP/TAZ signalling being detected in invasive LTED cells (Fig. S6), direct KRT80 manipulation did not further alter YAP/TAZ target genes (no evidence that Keratin 80 affects YAP/TAZ or that is a YAP/TAZ target). These data suggest that SREBP1 may influence YAP/TAZ activity and mechano-transduction as previously described [148] with YAP/TAZ activation primarily driven by *de novo* cholesterol biosynthesis present in LTED cells.

In summary, we observed higher YAP/TAZ activation in LTED, possibly via SREBP1-melanovate pathway. This might represent an additional mechanism of drug resistance driven by SREBP1 and independent of KRT80.

4.7 KRT80 cytoskeletal changes promote increased tumour stiffness

Despite being expressed in a wide range of epithelial tissues, K80 presents an unusual structure that is closer to hair keratins more than to other epithelial keratins[279]. Indeed, K80 structural features implies a different pattern of disulfide bonding in its secondary structure. This pattern differentiates hair and epithelial keratins, and results in a tougher and more durable structure of hair keratins compared to epithelial ones[244]. We thus hypothesized that epigenetic reprogramming might lead to changes in the mechanical properties of invasive BCa cells via KRT80 upregulation. To test this, changes in the viscoelastic properties of KRT80 over-expressing or knockdown cells were measured using magnetic tweezers[341]. A significant increase in cellular stiffness (inversely correlated to cell compliance/deformability) was observed at the single cell level after KRT80 over-expression in MCF7 and LTED cells (Fig. 4.6 A). Conversely, KRT80 depletion in LTED cells resulted in a significant loss of cellular stiffness (Fig. 4.6 A). It was previously suggested that cells acquiring invasive behavior increase their cell stiffness in response to changes in the extra-cellular matrix[342,343]. Our data suggest that changes in cell stiffness can also be cell-autonomous and directly imparted by epigenetic reprogramming in response to chronic treatment and that these changes might lead to increased invasive potential. Finally, we investigated if differences in tumour stiffness between patients could be explained at least in part by different K80 content in breast cancer cells. To test this hypothesis, twenty women with suspected BCa were prospectively recruited and radiological exam using Shear Wave ultrasound[338,344] was used to measure tissue stiffness in the normal stroma, peri-tumoral stroma and suspected lesions (Fig. 4.6 B). Elastography was performed prior to biopsies collection and all cases were further stained using two K80 antibodies. Eighteen out of twenty patients were later diagnosed with either Invasive Ductal or Invasive Lobular Carcinomas. By comparing tumour measurements with nearby normal stroma, our data showed that invasive carcinomas are stiffer than surrounding stroma (Fig. 4.6 B). Interestingly, measurements taken at the tumour-stroma interface, where the leading edge of tumour growth occurs, showed that these regions are in itself stiffer than the tumour, possibly reflecting a more substantial deposition of K80 positive cells. Therefore, we investigated if the increased tumour stiffness correlates with increased K80 content. To do so, we focused on

patients with high cellularity cancers, to avoid potential confounding factors such as adipocyte and immune infiltrate. Linear regression analysis showed that the percentage of K80 positive cells within a lesion significantly correlated with intra-tumor stiffness (Fig. 4.6 D), suggesting that K80 contribute to the physical property of cancer cells in clinical samples. Interestingly, meta-analysis of tumor and matched nearby tissue from TCGA show increased KRT80 mRNA in the tumor biopsies (Fig. S7). We also observed that strong K80 positivity characterized most breast cancer cells invading the nearby micro-environment in nearly all biopsies (Fig. 4.3). These *in vitro* and *in vivo* data further support the hypothesis that breast cancer cells characterised by high K80 content are mechanically stiffer.
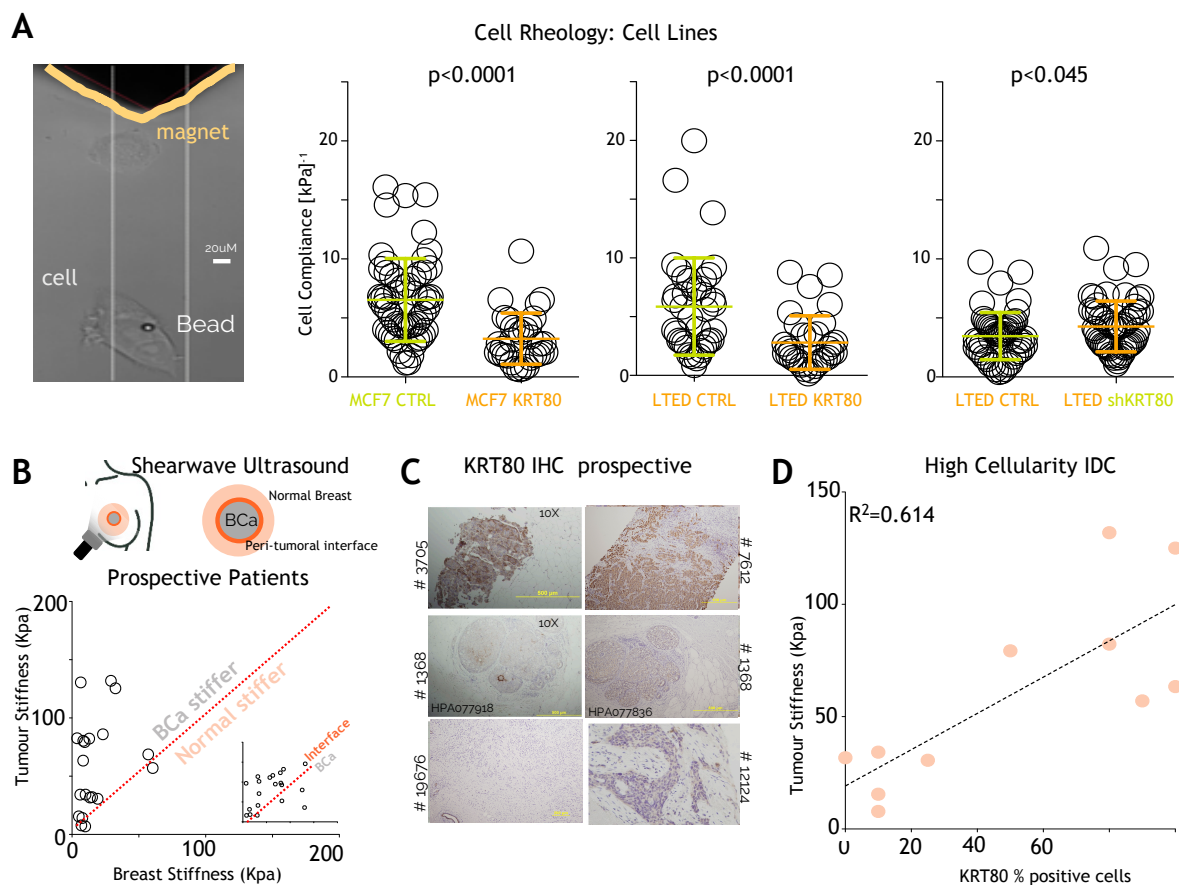
Figure 4.6: K80 levels are associated with changes in cell stiffness. A) Magnetic tweezers (yellow) were used to measure the biomechanical properties of individual cells with or without KRT80 manipulation. Changes in cell compliance (deformation) were measured in KRT80 over-expressing cells for both parental MCF7 and AI-resistant LTED cells (left and middle graphs). Changes in cell compliance (deformation) were measured after stable KRT80 depletion in LTED cells (right graph). Significance was calculated with a student t- test and reported. B) Shearwave Ultrasound measurements in prospectively recruited patients. Measures were collected at three independent location for each patient (see diagram). Plots show matched tissue stiffness for cancer vs.

normal (large panel) and peri-tumoral interface vs. cancer (small inset). C) K80 cells in diagnostic material from prospective patients assessed with ultrasound were counted using IHC. D) Plots show matched tissue stiffness against the percentage of K80-positive cells for each individual patient. Simple linear regression was applied to calculate the correlation coefficient between these two values (experiments in A performed by A.C. and A.L. at ICL, Department of Bioengineering; measurements and analysis in B, C, D performed by S.S., C.D., figures originally made by L.M., modified from Perone 2019)

4.8 KRT80-driven cytoskeletal changes promote migration and invasion

To test if KRT80 manipulation drives ancillary phenotypes synergistic to cytoskeletal changes, I performed RNA-seq in cells transfected with KRT80 but where SREBP1 is not yet activated (non-invasive MCF7 cells, Fig. 4.7A-B). Ectopic KRT80 expression led to clear transcriptional differences dominated by the reprogramming of a small set of genes (Fig. 4.7 A, B). Pathway analyses of upregulated genes suggest involvement in cytoskeletal rearrangements (Fig. 4.7 C). Amongst this small set of genes, particularly striking is the strong KRT80-dependent induction of cortactin (*CTTN*), a factor directly linked to actin rearrangements, lamellipodia formation and cancer cell invasion[345,346], that was confirmed by immunofluorescence (Fig. 4.7 D). In addition, a significant upregulation of Septin 9 (*SEPT9*) was also detected. *SEPT9* is a member of the septin family directly linked to actin fiber formation, focal adhesion maturation and motility[347,348] (Fig. 4.7 B). Genes upregulated in response to KRT80 activation have prognostic value, even when other classical clinical features are considered (Fig. 4.7 E). In strong agreement with KRT80 prognostic features, our data suggest that these genes may be involved in early metastatic invasion. RNA-seq analysis also revealed several genes negatively regulated by KRT80 induction playing central roles in cancer biology including: i) negative regulators of migration (PCDH10, CADM1), ii) tumor suppressors such as CDKN1A (p21) and PDCD2, iii) genes involved in DNA repair (RAD50), iv) chromatin remodelers as SMARCE1 and CHD4 and v) tumor specific antigens (CD276). These findings suggest a direct link between cytoskeletal reprogramming and several other oncogenic phenotypes (Fig. 4.7 B). Together, these results further support that KRT80 manipulation is sufficient to activate genes driving evident cytoskeletal rearrangements that ultimately induce invasive behaviors in BCa and poorer prognosis. We cannot speculate at the moment if this is driven by a cytoskeleton-transcriptional feedback or if it is mediated by some specific transcriptional factors or some more complex signalling mechanism between cytoskeleton dynamics and the control of cell cycle. Aberrant cytoskeletal architecture characterizes tumour cells and it is associated with

cell migration and invasion[345]; yet the mechanisms underlying cytoskeletal reorganization in tumour cells are not well understood. Here, we have uncovered a novel and causal link between endocrine therapy resistance, intra-tumoral stiffness and augmented invasive potential in luminal BCa (as schematized in the working model in Fig. 4.7 F).
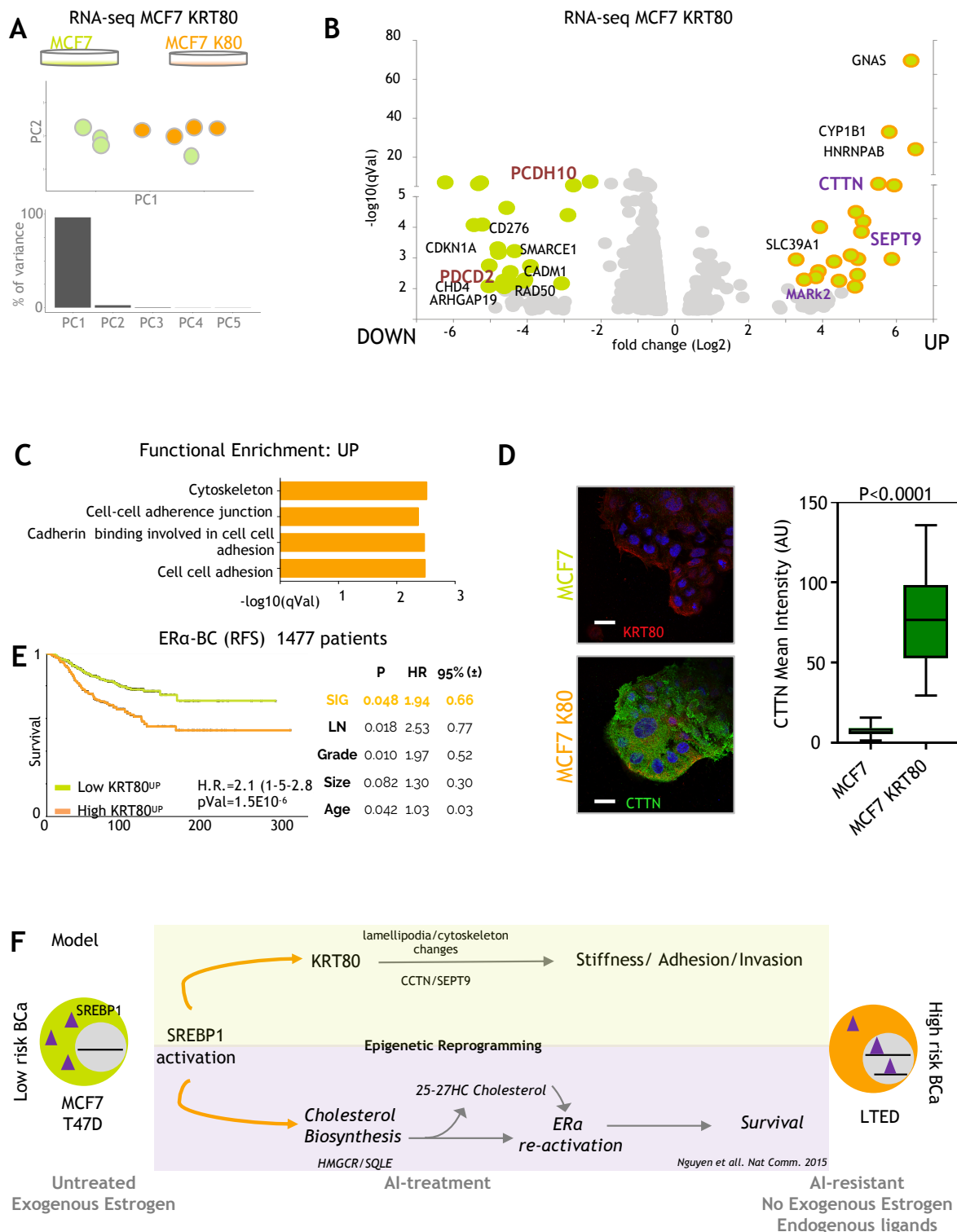
Figure 4.7 (previous page): KRT80-changes induce transcriptional changes of cytoskeletal genes. A) Principal Component Analysis (PCA) of RNA-seq profiled MCF7 breast cancer cells or MCF7 cells with ectopic expression of KRT80. B) Volcano plots of over-expressed or under-expressed genes in MCF7 cells following KRT80 ectopic expression. C) Functional enrichment for upregulated genes following KRT80 ectopic expression. B-C) Analysis performed by Giacomo Corleone. D) Representative confocal microscopy images showing F-actin (magenta), cortactin (CTTN, green) and DAPI (blue) staining of MCF7-control and MCF7-KRT80 cells (performed by A.F., F.C. at ICR London). Scale bars represent 25 μm. Graph shows mean fluorescence intensity of cortactin in MCF7-control and MCF7-KRT80 cells (n = 40, MCF7; n = 4, MCF7-KRT80 individual cells). E) Kaplan-Meier plot of ERα-positive breast cancer patients dichotomized to average high or low expression for genes upregulated in response to KRT80 over-expression (Panel B, analysis performed by L.M.). Multivariate statistics are shown on the right inside table. F) Current model: long-term AI treatment promotes constitutive activation of SREBP1 leading to pro-survival re-activation of estrogen receptor, and global cytoskeletal re-arrangements. Cytoskeletal re-organization leads to direct biomechanical changes and promotes pro-invasive behavior (figures originally made by L.M., modified from Perone 2019).

4.9 Discussion and Future Work

In this chapter, we investigated the mechanisms linking treatment-induced epigenetic reprogramming to cytoskeletal changes directly promoting invasive potential in breast cancer cells. Dissecting the relationship between drug-resistance and tumour metastatic program development in the clinical setting is rather problematic due to the fact that, in most cases, relapsing patients have been chronically treated with adjuvant ET. Starting from breast cancer cell lines and their invasive and resistant counterparts, we investigated how chronic exposure to standard of care hormone-deprivation treatment can trigger cytoskeletal rearrangements through SREBP1 regulation, thus promoting the acquisition of more invasive behavior.

Epigenetic reprogramming leads directly to cytoskeletal changes including the increased development of invasive cytoplasmic structures in cells that have acquired aggressive potential. In AI resistant cells, epigenetic reprogramming contributes to significant switches in the keratin organization[53] , in particular RNA-seq data highlight that only few keratins are differentially expressed in LTED cells when compared to parental, untreated MCF7 cells (Fig. 4.2A). Given that RNA-seq also identified upregulation of NR4F1, KRT86 and KRT81 in some resistant/invasive cells, it does raise the question that are the effects of KRT80 on increasing stiffness, lamellipodia formation and migration unique to KRT80 or partially shared with these other keratins. We noticed that NR4A1 is not a Keratin, but a nuclear receptor and is

upregulated only in MCF7 Fulvestrant resistant but not in the invasive AI cell lines. KRT86 and 81 are the only other Keratins upregulated in AI resistant lines, but they have no prognostic significance in breast cancer.

KRT80 is an unusual type II keratin, highly conserved during evolution and possessing a number of highly peculiar properties. Whether other keratins evolved their high diversity from gene duplication rather than alternative mRNA splicing, K80 is the only one to have a splice protein variant K80.1[279]. However, the role and function of K80.1, generally and in cancer, is as yet unknown. The regulation of KRT80 expression either through partnering with other keratins like type I or post-translational modifications such as phosphorylation, is yet unexplored. On the other hand, KRT80 is emerging as a prognostic marker in other cancers (Fig. S8 A) suggesting it might have a more conserved role in driving invasive phenotypes. Therefore, we pinpoint these changes to the increased expression of the previously uncharacterized keratin, K80, as a parallel event to the activation of SREBP1-driven *de novo* cholesterol biosynthesis in AI resistant cells (Fig. 4.7 F).

Upon long-term AI treatment, SREBP1 mediates the activation of pro-survival pathways[53] by promoting the cell-autonomous production of endogenous ERα ligands. In addition, SREBP1 is also recruited at the *KRT80* enhancer, a non-canonical SREBP1 target, leading to KRT80 transcription in drug-treated cells. Despite our data showing that SREBP1 is essential for KRT80 activation, it alone might not be sufficient for activation. How SREBP1 is capable to sense AI-mediated stress needs to be worked out mechanistically, but overall these data support SREBP1 as a potential target to antagonize BCa progression. We also describe an unexpected role for intermediate filaments in promoting cancer cell invasion by showing for the first time that KRT80 promotes actin cytoskeleton rearrangements. We show that KRT80 accumulation contributes to cell-autonomous cytoskeletal changes driving increased cell stiffness at invadopodia. Notably, these data were confirmed in ERα-positive patients through prospective measurement of tumour stiffness coupled with K80 immunostaining.

The effect of increasing stiffness in metastatic invasion is highly debated. Changes in extra-cellular matrix stiffness can force focal adhesions, growth factor signalling and breast malignancy[349,350]. However, a significant body of clinical literature has linked increased breast tumour stiffness to poorer prognosis [69,338] and lymph node positivity [69,338,344] independently of changes in extracellular matrix stiffness. We reasoned that a model in which KRT80 upregulation in BCa cells leads to increased stiffness and augmented collective invasion might

partly reconcile all these observations. Several mechanisms of invasion have been identified, including some that involve switches between epithelial to mesenchymal phenotypes to allow for single-cell invasion[275] through loss of keratins leading to a decreased stiffness. Nonetheless, BCa metastatic lesions are characterized by defined epithelial features[351] and recent reports suggest breast cancer cells do not need to undergo EMT to form lung metastasis[70]. Solid tumours can use a myriad of multicellular invasion programs collectively termed "collective invasion"[274]. Recent studies have shown that epithelial cancer can invade through collective invasion mediated by leading cells acquiring well defined keratin patterns[67]. This mode of invasion for breast cancer cells is also supported by recent genetic evidence showing that keratins, such as KRT14, can play critical roles in collective invasion[67] and multi-clonal metastatic seeding[67,69], two processes driving BCa progression[67]. Indeed, invading cells releasing from the spheroid clusters maintain cell-cell contacts, in agreement with KRT80 potential role in cell-cell adhesion[279] as KRT80 is also found in desmosomes (Fig. S8 B). This metastatic route is also strongly reminiscent of tumor budding in colon cancer[352]. In line with this model, we found KRT80 ranking consistently in the top 1% of genes with increased expression in colon cancer compared to normal colon, despite no evidence of gene amplification[329] (4th in the TCGA colorectal dataset, 20-fold increase, $p<7*10^{-30}$).

Confocal microscopy analysis showing an increase in focal adhesion formation in cells expressing KRT80 may seem in contrast with the increased migration hereby described. Focal adhesions (FAs) are the points of contact of a cell with the substrate, and therefore more or bigger FAs are intuitively associated with a stronger adhesion that may difficult migration. However, FAs are critical for the generation of the tensional forces required for cytoskeletal rearrangements that elicit the cell shape changes and generation of structures required for migration. In particular, the forward movement of the lamellipodium requires focal adhesion growth and maturation at the leading edge to propel the cell forward[337]. Simultaneously, focal adhesion disassembly occurs mainly at the rear where the cell needs to detach to move the cell body forward. We rationalize that the increased focal adhesion at the leading-edge results from the reorganization of F-actin and it is a direct consequence of the generation of lamellipodia-like structures. Thus, the novel network of KRT80 positive filaments is promoting cell motility and invasive behaviors by reinforcing the organization of the lamellar actin fiber network and the stability of focal adhesions.

A recent study identified SREBP1 to be regulated by acto-myosin contractility and mechanical forces imposed by the ECM[353]. The authors reported SREBP1 control by mechanical cues to

be dependent on geranylgeranyl pyrophosphate, another key bio-product of the mevalonate pathway, and impacting on stem cell fate in mouse and on fat storage in Drosophila. Drosophila has only a single homologue of SREBP which activity is not dependent on cholesterol[79,355]; in this model inhibition of actin polymerisation impacts on SREBP activation[353]. They show that activation of AMPK by ECM stiffening and geranylgeranylated RhoA-dependent acto-myosin contraction inhibits SREBP1 activation in TNBC and MCF10A cells.

Another recently published report describes how extracellular physical cues may affect the mechanical properties of the Golgi apparatus impacting on lipid metabolism[356]. The authors showed that conditions of reduced actomyosin contractility lead to accumulation of SCAP/SREBP to the Golgi and activation of SREBP, in turn driving lipid synthesis[356]. This occurs independently of YAP/TAZ, mTOR and AMPK, and in parallel to feedback control by sterols identifying a general mechanism centred on SREBP that links the physical cell microenvironment to a key metabolic pathway[356]. Although it might seem in contrast with our results, some experimental conditions about this study are to be taken into account when discussing their general conclusions[356]: i) primary cell line models are used for the experiments (RPE1, 3T3L1, WI38, hPSCs, HEK293, human fibroblast-like fetal lung cells, human retina, human embryonic kidney, mouse adipocytes respectively) and two TNBC cell lines (non-metastatic MCF10ATk1 and MDA231); ii) ECM mechanical cues regulate lipid synthesis through SREBP2; iii) *in vivo* analysis are conducted in soft normal skin and matched stiff keloid tissue of 7 patient samples. In summary the authors suggest that, mostly in primary cell line models, the extracellular microenvironment forces impact on the intracellular lipid metabolic pathway via SREBP2[356].

We observed higher YAP/TAZ activation in LTED, possibly via SREBP1-melanovate pathway. This might represent an additional mechanism of drug resistance driven by SREBP1 and independent of KRT80. Nevertheless, further analyses in substrates with different stiffness are needed to assess whether KRT80 may participate in YAP activation in soft or physiologically relevant substrates rather than in hard plastic/glass culture conditions as the ones used. Characterization of YAP transcriptional activity by using reporters in soft conditions or in KRT80 silenced LTED cells could be helpful to further rule out whether KRT80 is involved in the regulation of this pathway.

Moreover, future work is needed to address the crosstalk between the tumour microenvironment and cellular motility and stiffness. A human triculture 3D *in vitro* tumour microenvironment system (TMES) has recently been engineered recapitulating the transcriptomic changes occurring *in vivo* in both patients and xenograft models[357]. In the view

of testing novel compounds impinging on SREBP1 signalling, this model could also provide a unique platform to evaluate therapies in cell lines, patient tumour material and with applicability for patient avatars.

At a clinical level, an interesting point to address with future work would be characterising the differences between invasive and non-invasive phenotypes at a transcriptional level. To this end, a cohort of patients including low, middle and high-grade ERα breast cancer, lobular breast cancer and TNBC could be recruited. All cases could be stained for IHC using K80, E-cadherin, vimentin, other cytokeratins such as K5, high and low molecular weight K14-K8, cytokeratin AE1/AE3 and Cam 5.2. Subsequently, further characterisation by using digital spatial profiling with NanoString nCounter gene expression system technology[358] could be performed to capture and count individual mRNA transcripts directly measuring mRNA expression levels *in vivo*. The study would aim to: i) characterise the phenotypic heterogeneity within individual patients, that might not to be genetically driven and ii) study the dynamics of cancer cell invasive behaviour development.

To further explore the link between intra-tumoral stiffness and AI resistance, future work comparing patients responding or not to therapy is needed. Larger longitudinal clinical studies measuring stiffness and KRT80 activation in endocrine neo-adjuvant-treated patients should be linked to long-term monitoring for distal relapse. Moreover, to investigate if there is a correlation between KRT80 and SREBP1 in clinical samples, especially after endocrine therapy, IHC should be performed to measure intranuclear SREBP1 levels in pre and post-AI treated patients' samples. This measurement could be then further correlated to K80 levels.

Our results show that changes in KRT80 expression and localization in breast cancer track invasive potential in clinical samples where KRT80 also promotes stiffer cancer lesions. As KRT80 expression has been found to be increased in lung, colorectal, prostate and cervical cancers as well, future work aiming at expanding our current results to other types of cancer would be pivotal to understand the role of KRT80 as a biomarker in the development of drug resistance program in cancer.

4.10 Conclusions

The findings described in this chapter have two major implications. Firstly, we show that invasive cancer cells are not necessarily "softer" as previously reported. As also other studies described a correlation between stiffness with invasive potential in clinical samples[338,344], it might be possible that increased stiffness is required to efficiently invade stiffer extra-cellular compartments especially using collective invasion mechanisms where cell-to-cell contact is essential[67-69,354]. Although, at diagnosis, breast cancer cells develop in one of the "softest" microenvironments, they also have a well-known organotropism for bone tissue[359] suggesting that invading cells might benefit from a stiffer cell structure. Additionally, the activated stroma of invasive breast cancers significantly increases in stiffness concomitantly with BCa progression. Our data then support a model in which endocrine therapy resistance promotes changes in the cellular rheology underlying the escape from the primary site. AI resistant cells then develop into resilient cancer cells capable of matching their cytoplasm fluid mechanics to stiffer matrixes (i.e. cancer-associated stroma[360]) and large bones. Secondly, our data suggest that mechanical changes in breast cancer cells are not uniquely driven by functional changes in the extra-cellular matrix but can also be determined by epigenetic changes in response to specific endocrine therapies. This is extremely important because it points at SREBP1 inhibition as a potential therapeutic point of entry to reverse cell stiffness. Our data warrant for further mechanistical studies and clinical investigation aimed at modifying the mechanical properties of metastatic BCa to prevent or delay invasion.

"The rule to fight on the disputed territory is that
who adapts to the situation, wins
and who limits himself to fight, loses."
Sun Tzu, The Art of War

Chapter 5: Results III- Collaborative published work

5.1 "Going off the grid: ERα breast cancer beyond estradiol." Y Perone and L Magnani

This review briefly summarized the previously published findings from our lab from which my PhD research project would be based on. It discussed how epigenetic reprogramming is connected to endogenous cholesterol biosynthesis in cancer cells in the process of becoming resistant to endocrine therapy. We particularly focused on how cholesterol metabolites, including 27HC, can act as ERα ligands in these cells. From the reviewed data, we then proposed a working model in which cholesterol biosynthesis promotes autocrine, pro-invasive signalling via activation of a series of closely related transcription factors. We then examined how our proposed model is influenced by aromatase inhibitors. Finally, we review how the field is considering the development of anti-cholesterol therapeutics and companion biomarkers to stratify and treat ERα breast cancer patients (Fig. 5.1).
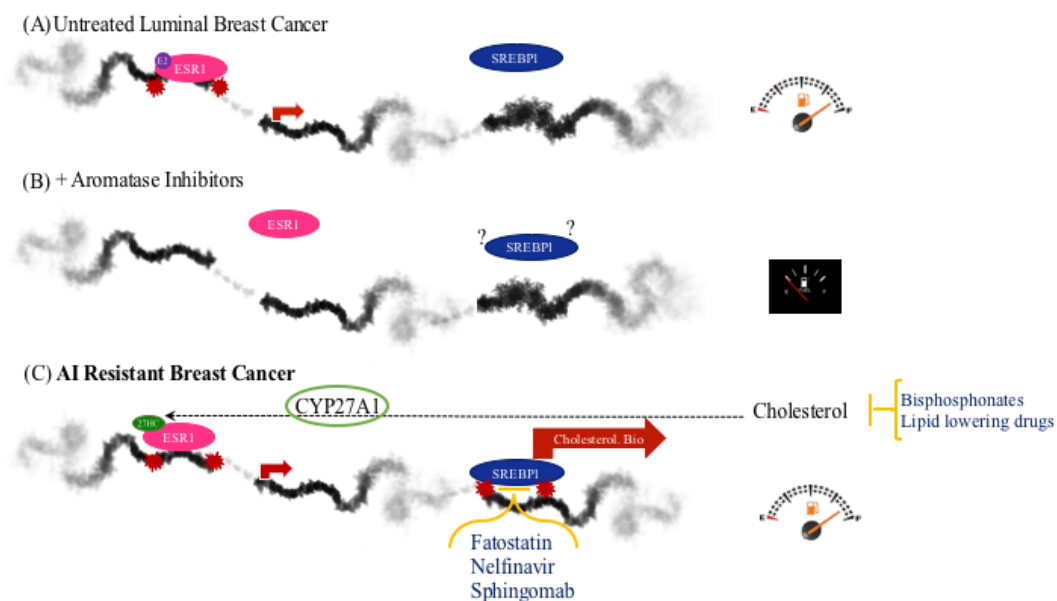
Figure 5.1: Schematic model of AI-resistant ERα breast cancer. Epigenetically hyperactivated cholesterol biosynthesis becomes an alternative fuel for ERα mediated transcription. Potential therapeutic molecules targeting the cholesterol machinery are included in brackets. DNA is represented in black and grey, red arrow indicates gene transcription. The key molecules involved in this working model are shown in the schematic (oestradiol, E2; Estrogen receptor 1, ESR1; Sterol regulatory binding element 1, SREBP1; 27 hydroxycholesterol, 27HC), (modified from Perone 2016, figure originally made by Luca Magnani).

5.2 "Acquired *CYP19A1* amplification is an early specific mechanism of aromatase inhibitor resistance in ERα metastatic breast cancer" L Magnani, G Frige`, RM Gadaleta, G Corleone, S Fabris, H Kempe, PJ Verschure, I Barozzi, V Vircillo, S Hong, **Y Perone**, M Saini, A Trumpp, G Viale, A Neri, S Ali, MA Colleoni, G Pruneri and S Minucci.

More than 20% of luminal BCa patients treated with adjuvant ET, relapse within 10 years and eventually progress to incurable metastatic disease. In this study we demonstrate that the therapy used has a profound influence on the genetic landscape of relapsed diseases (Fig. 5.2 A). Tumour evolution is shaped by many factors, potentially involving external selective pressures induced by therapies. We found that 21.5% of AI-treated, relapsed patients acquires CYP19A1 (encoding aromatase) amplification (CYP19A1^{amp}). Relapsed patients also developed numerous mutations targeting key breast cancer-associated genes, including ESR1 and CYP19A1. Notably, CYP19A1^{amp} cells also emerged *in vitro*, but only in AI-resistant models. CYP19A1 amplification caused increased aromatase activity and oestrogen independent ERα binding to target genes. Thus, CYP19A1^{amp} cells showed decreased sensitivity to AI treatment. These data indicate that AI treatment selects for acquired CYP19A1^{amp} and promotes local autocrine oestrogen signalling in AI-resistant metastatic patients.
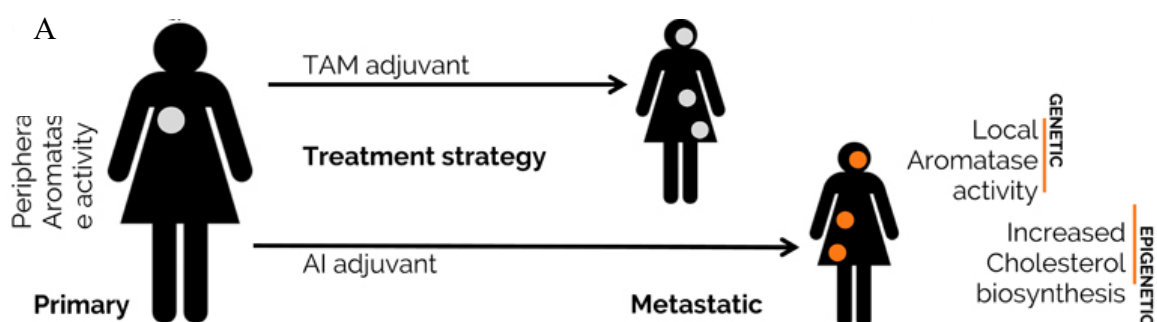
To test the role of *CYP19A1^{amp}* on AI sensitivity, I depleted CYP19A1 mRNA in the LTED cells using two independent siRNA (siCYP19A1-1, siCYP19A1-2) and overexpressed it in parental MCF7 cells using a CYP19A1 open reading frame construct. Then, I validated CYP19A1 silencing and overexpression by RT-qPCR (Fig. 5.2 B, C).
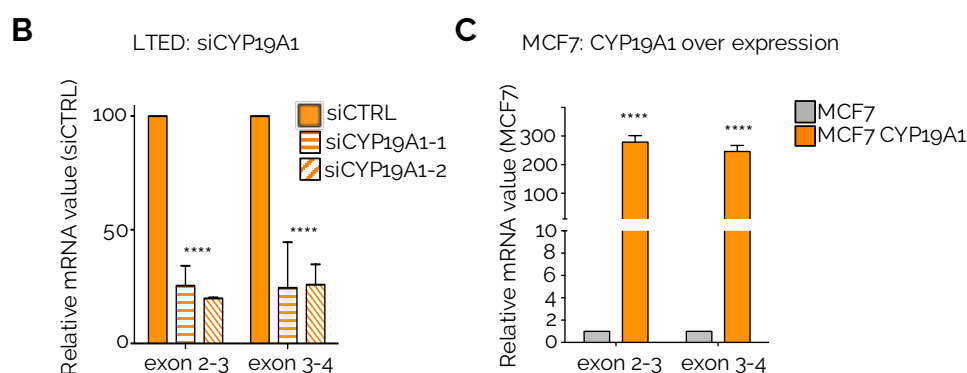
Figure 5.2: Manipulation of CYP19A1 in breast cancer cells. A) Previous page: Working hypothesis for therapy-specific breast ... and epigenetic changes collaborate to increase tumour fitness by creating an ... metastatic sites in patients treated with AI therapy. B) CYP19A1 mRNA was ... siRNA as detected via RT-qPCR. Data is presented as mRNA expression l... NA against CYP19A1 relative to cells transfected with control siRNA using two alternative intron-spanning primers. Data represents mean and SEM from 3 independent experiments. Asterisks represent significant difference after one-way ANOVA with Dunnett's post-test. C) CYP19A1 was overexpressed using a CYP19A1 ORF. Bars represent mRNA levels measured with qRT-PCR and compared to a control plasmid using two alternative intron-spanning primers. Data represents mean and SEM from 3 independent experiments. Asterisks represent significant difference after a Student T-Test (modified from Magnani 2017, figure originally made by Luca Magnani).

To investigate cell proliferation and cytotoxicity, sulforhodamine B assay (SRB) was conducted to test the proliferative/cytotoxic effects of Letrozole in both overexpressing MCF7 and silenced LTED cells (Fig. 5.3). In LTED cells, siCYP19A1 significantly increased the sensitivity of these cell to AI treatment (Fig. 5.3 A). Letrozole, however, did not affect MCF7 grown in oestrogen-supplemented conditions (Fig. 5.3 B). These results suggest that: i) CYP19A1 over-expression did not confer any growth advantage to MCF7 cells grown in the presence of oestradiol (Fig. 5.3 B), ii) CYP19A1 over-expression was sufficient to relieve cell growth arrest in MCF7 cultured in absence of oestrogens (Fig. 5.3 B), iii) interestingly, this effect could not be antagonized by letrozole (Fig. 5.3 B). Collectively, these data indicate that CYP19A1 amplification might induce reduced sensitivity to AI treatment.
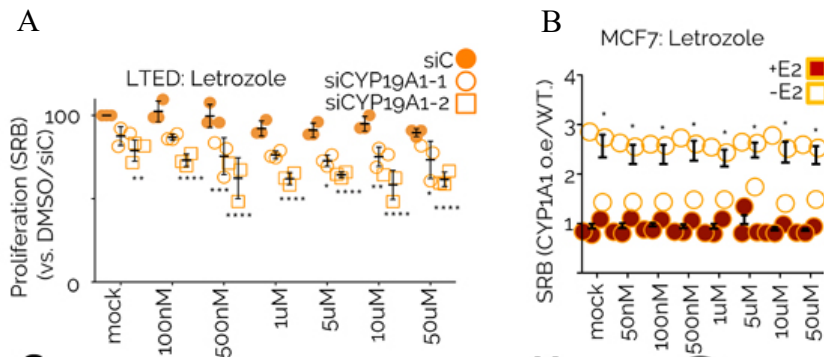
Figure 5.3: CYP19A1 amplification in BCa cells endogenously activates ERα and develop tolerance to AI. A) LTED cells treated with siRNA against CYP19A1 (siCYP19A1-1 and siCYP19A1-2) have increased sensitivity to AI. SRB, sulforhodamine B assay; siCTRL, control siRNA targeting luciferase. B) CYP19A1-overexpressing cells have a growth advantage over wild-type in the absence of oestradiol. Relative increase in growth rate is shown as the ratio of the growth of CYP19A1-overexpressing cells to CYP19A1 wild-type cells under letrozole challenge. Data are mean ± s.e.m. from 3 independent experiments. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, two-way ANOVA and Bonferroni's post-test. (modified from Magnani 2017, figure originally made by Luca Magnani).

CYP19A1 copy number alteration (CNA) was exclusively identified in LTED cells and AI-treated patients *in vivo* indicating that CYP19A1 CNA is acquired during treatment. Similarly, CYP19A1 has also been previously reported to be upregulated in malignant versus benign prostate epithelium, and in metastatic PCa compared to localised disease[232] (see 1.2.4). Mechanistically, AI targets CYP19A1, the aromatase enzyme mediating the conversion of testosterone to oestradiol, with the aim to remove circulating oestrogens. Testosterone, along with other sex hormones, are produced from cholesterol. *De novo* cholesterogenesis is upregulated in LTED cells via epigenetic reprogramming to promote autonomous ERα activation. Indeed, LTED cells have a ten-fold increase in ERα protein expression[53] despite no evidence of ESR1 amplification or ER activating mutations in these cells[302].

In summary, our study indicate that hormone deprivation promotes acquired CYP19A1 amplification that, in turn, triggers ERα activity by converting testosterone possibly obtained via epigenetically driven *de novo* cholesterol biosynthesis.

5.3 "Enhancer mapping uncovers phenotypic heterogeneity and evolution in patients with luminal breast cancer" DK Patten, G Corleone, B Győrffy, **Y Perone**, N Slaven, I Barozzi, E Erdős, A Saiakhova, K Goddard, A Vingiani, S Shousha, LS Pongor, DJ Hadjiminas, G Schiavon, P Barry, C Palmieri, RC Coombes, P Scacheri, G Pruneri and L Magnani.
Nature Medicine 2018 doi: 10.1038/s41591-018-0091-x

The degree of intra- and inter-patient phenotypic heterogeneity and its role in tumour evolution is still poorly understood. Phenotypic changes can be transmitted via transcriptional programs maintained through the epigenetic modulation of regulatory regions such as promoters and enhancers (Fig. 5.4). In this study, we identified key regulatory elements commonly shared across of ERα breast cancer patients by investigating the regulatory landscape of primary and metastatic clinical samples. We found that shared regions contain a unique set of regulatory information including the motif for transcription factor Ying Yang 1 (YY1). We identified YY1 as a critical determinant of ERα transcriptional activity promoting tumour growth in most luminal patients. YY1 also contributed to the expression of genes mediating resistance to endocrine treatment. By tracking the clonality of SLC9A3R1-positive cells, a *bona fide* YY1-ERα-regulated gene, we showed that endocrine therapies select for phenotypic clones under-represented at diagnosis.

Collectively, our data showed that epigenetic mechanisms significantly contribute to phenotypic heterogeneity and evolution in systemically treated breast cancer patients.
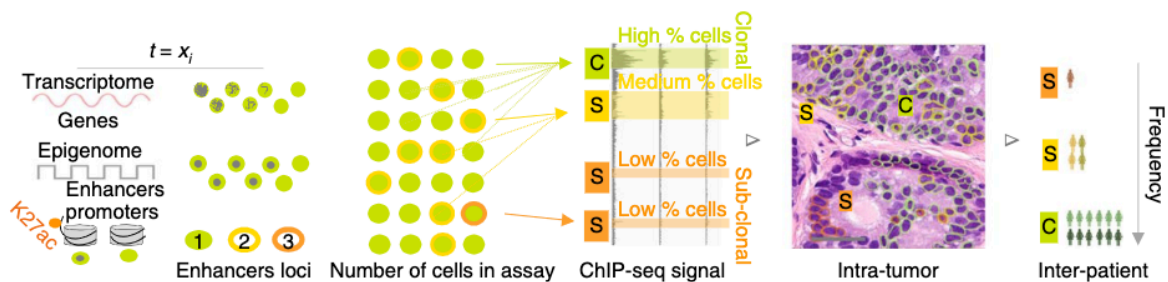
Figure 5.4: Assessment of intra- and inter-tumour epigenetic heterogeneity. Main hypothesis of the study. RNA is ultimately an analogue signal in which each individual cell, at any given time, can contribute a stochastic amount of RNA, while transcriptional data from bulk tissue represent an average over a million cells. For chromatin data, at any given time ($t = X_i$), each cell can only contribute a deterministic value to the bulk signal, generally from two alleles. Therefore, the relative strength of ChIP–seq data is dependent on the number of cells carrying an epigenetic signal at discrete loci. C and S represent strong and medium/weak signal, respectively. Clonal regulatory regions are commonly shared by BCa patients, whereas weak enhancers are more patient specific (Modified from Patten 2018, figure originally made by Luca Magnani).

Genetic alterations may affect the DNA accessibility and ultimately be a source of patient-specific bias. Theoretically, mutations of genes previously associated with development of drug-resistance could explain the metastatic progression without the need of involving epigenetic aberrations. Thus, in order to investigate the potential synergy between epigenetic changes and genetic drivers, I conducted an updated Cancer Hotspot panel including all the major mutations commonly found in cancer and further customised for the ESR1 mutation. I performed targeted capture sequencing by using input DNA material from a total of 55 primary and metastatic samples constituting patients' clinical dataset and cell lines MCF7 and LTED (Fig. 5.5). Downstream analysis was performed, and we found that PIK3CA mutation affected 13/55 patients, while ESR1 was found only in 4 out of 13 metastatic patients (Fig. 5.5). However, there were no consistent coding mutations that could justify on their own the development of resistance in the metastatic samples. Undoubtedly, future studies may further exploit this clinical dataset to identify any evidence of genetic/epigenetic interplay in primary and metastatic ERα breast cancer patients.
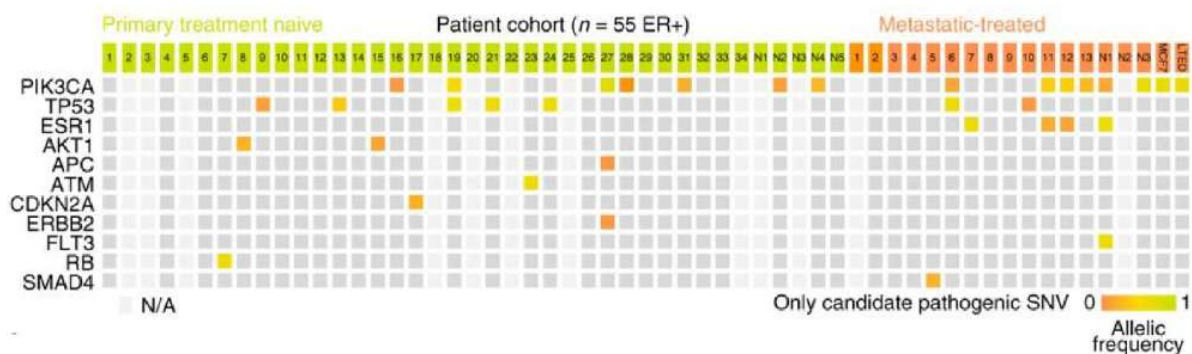
Figure 5.5: Targeted sequencing for common genes affecting BCa. Mutational analysis for common cancer driver genes in the patients' cohort selected for the study. Fifty-five ERα-positive BCa samples with H3K27ac ChIP–seq were profiled to build a comprehensive compendium of clinically relevant active regulatory regions (primary n = 39, metastatic n=16, MCF7 and LTED cells). Colours show the allele frequency of identified mutations. Colour scheme: from green=100% allele frequency to orange=0% allele frequency (Analysis performed by Iros Barozzi, figure produced by Luca Magnani, modified from Nature Medicine 2018).

Conclusions

The work presented here aimed to provide a further understanding of the mechanisms of activation and the role of SREBP1 in hormone-dependent cancer cells. We found that SREBP1-driven lipogenesis is activated in hormone depleted resistant breast and prostate cancer cells. Kinetically, short-term hormone deprivation is not sufficient to induce SREBP1-dependent CB activation that is instead consistently upregulated in resistant long-term hormone deprived cell lines. Preliminary *in vivo* data indicate that the number of SREBP1-positive cells is increased in metastasis of AI-treated breast cancer patients compared to the matched pre-treatment primary. These results further support the hypothesis that SREBP1 may be important in driving *de novo* cholesterol biosynthesis in endocrine therapy resistant breast cancer cells. Lipidic profiling suggests that the activation of the SREBP1 pathway may be glutamine-dependent. Indeed, a switch of metabolic dependency upon resistance development from a glucose- to a glutamine-dependent phenotype is associated with increased *de novo* cholesterol and fatty acid synthesis. However, our data indicate that transcriptional activation of the SREBP1 gene *per se* is not likely to be the driver of the increased *de novo* lipid activity, suggesting that cholesterol biosynthesis may be upregulated by modulating SREBP1 signalling. Therefore, in order to investigate SREBP1 recruitment to the chromatin, I performed optimization of ChIP protocol that allowed for a genome-wide profiling of SREBP1 binding in breast and prostate cancer cells. Downstream analysis showed a difference in SREBP1 cistrome between parental and long-term starved derived cell lines in MCF7. Furthermore, SREBP1 binding profiles distinguished cancer cells based on the tissue of origin (breast versus prostate cancer cells) suggesting that, at least to some extent, SREBP1 binding might be driving cancer specific phenotypes. Our data also confirmed a significant co-occurrence between AR and SREBP1 binding sites in PCa and suggested a possible crosstalk between SREBP1 and the ERα in BCa on chromatin. Finally, differential binding analysis identified novel SREBP1 non-canonical target genes in cancer cells. Keratin 80 was then further investigated as one of the SREBP1 non-canonical targets in relationship to potential changes in invasive properties.

In LTED cells, epigenetic reprogramming contributes to significant switches in the keratin organization directly leading to cytoskeletal changes including the increased development of invasive cytoplasmic structures in cells that have acquired aggressive potential. We pinpoint these changes to the increased expression of Keratin 80, as a parallel event to the activation of

SREBP1-driven *de novo* cholesterol biosynthesis in LTED cells. We showed that KRT80 accumulation contributes to cell-autonomous cytoskeletal changes driving increased cell stiffness and directly contributing to actin polymerization at invadopodia. Notably, these data were confirmed in ERα-positive patients through prospective measurement of tumour stiffness coupled with K80 immunostaining.

In conclusion, our data support a model in which endocrine therapy resistance promotes constitutive chromatin activation of SREBP1 leading to pro-survival re-activation of oestrogen receptor[53], and global cytoskeletal re-arrangements[329] (working model in Fig. 4.7.F). Cytoskeletal re-organization leads to direct biomechanical changes and promotes pro-invasive behaviour. Finally, our results suggest that mechanical changes in breast cancer cells are not uniquely driven by functional changes in the extra-cellular matrix but can also be determined by intrinsic epigenetic changes in response to specific endocrine therapies. This is extremely important because it points at SREBP1 inhibition as a potential therapeutic point of entry to reverse cell stiffness. Our data warrant for further mechanistical studies and clinical investigation aimed at modifying the mechanical properties of metastatic BCa to prevent or delay invasion.

Bibliography

1. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca Cancer J Clin* **68,** 394–424 (2018).

2. PAFFENBARGER, R. S., KAMPERT, J. & CHANG, H.-G. CHARACTERISTICS THAT PREDICT RISK OF BREAST CANCER BEFORE AND AFTER THE MENOPAUSE. *Am J Epidemiol* **112,** 258–268 (1980).

3. Epstein, F. H., Clemons, M. & Goss, P. Estrogen and the Risk of Breast Cancer. *New Engl J Medicine* **344,** 276–285 (2001).

4. Endogenous Hormones and Breast Cancer Collaborative Group, *et al.* Sex hormones and risk of breast cancer in premenopausal women: a collaborative reanalysis of individual participant data from seven prospective studies. *Lancet Oncol* **14,** 1009–1019 (2013).

5. Collaborative Group on Hormonal Factors in Breast Cancer. Type and timing of menopausal hormone therapy and breast cancer risk: individual participant meta-analysis of the worldwide epidemiological evidence. *Lancet* (2019). doi:10.1016/s0140-6736(19)31709-x

6. Yager, J. D. & Davidson, N. E. Estrogen carcinogenesis in breast cancer. *The New England journal of medicine* **354,** 270–82 (2006).

7. James, R. E. *et al.* Postmenopausal Serum Sex Steroids and Risk of Hormone Receptor–Positive and -Negative Breast Cancer: a Nested Case–Control Study. *Cancer Prev Res* **4,** 1626–1635 (2011).

8. Huh, S. *et al.* The Proliferative Activity of Mammary Epithelial Cells in Normal Tissue Predicts Breast Cancer Risk in Premenopausal Women. *Cancer Res* **76,** 1926–1934 (2016).

9. Katz, T. A. Potential Mechanisms underlying the Protective Effect of Pregnancy against Breast Cancer: A Focus on the IGF Pathway. *Frontiers Oncol* **6,** 228 (2016).

10. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406,** 747 (2000).

11. Wallden, B. *et al.* Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. *Bmc Med Genomics* **8,** 54 (2015).

12. Cheang, M. C. *et al.* Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer. *Jnci J National Cancer Inst* **101,** 736–750 (2009).

13. Osborne, K. C., Yochmowitz, M. G., Knight, W. A. & McGuire, W. L. The value of estrogen and progesterone receptors in the treatment of breast cancer. *Cancer* **46,** 2884–2888 (1980).

14. Sørlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc National Acad Sci* **98,** 10869–10874 (2001).

15. Sørlie, T. Introducing Molecular Subtyping of Breast Cancer Into the Clinic? *J Clin Oncol* **27,** 1153–1154 (2009).

16. Duffy, S. W. *et al.* Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *J Med Screen* **17,** 25–30 (2010).

17. Roy, S. & Vadlamudi, R. K. Role of Estrogen Receptor Signaling in Breast Cancer Metastasis. *Int J Breast Cancer* **2012,** 654698 (2012).

18. Ali, S. & Coombes, C. R. Endocrine-responsive breast cancer and strategies for combating resistance. *Nat Rev Cancer* **2,** nrc721 (2002).

19. Sommer, S. & Fuqua, S. Estrogen receptor and breast cancer. *Semin Cancer Biol* **11,** 339–352 (2001).

20. Osborne, K. C. & Schiff, R. Mechanisms of Endocrine Resistance in Breast Cancer. *Annu Rev Med* **62,** 233–

247 (2011).

21. Johnston, S. & Dowsett, M. Aromatase inhibitors for breast cancer: lessons from the laboratory. *Nat Rev Cancer* **3,** nrc1211 (2003).

22. Ma, C. X., Reinert, T., Chmielewska, I. & Ellis, M. J. Mechanisms of aromatase inhibitor resistance. *Nat Rev Cancer* **15,** 261–275 (2015).

23. Gruber, C. J., Tschugguel, W., Schneeberger, C. & Huber, J. C. Production and Actions of Estrogens. *New Engl J Medicine* **346,** 340–352 (2002).

24. Auchus, M. & Auchus, R. J. Human Steroid Biosynthesis for the Oncologist. *J Invest Med* **60,** 495 (2012).

25. Frenkel, B. *et al.* Regulation of adult bone turnover by sex steroids. *J Cell Physiol* **224,** 305–310 (2010).

26. Mendelsohn, M. E. Protective effects of estrogen on the cardiovascular system. *Am J Cardiol* **89,** 12–17 (2002).

27. Viña, J. & Lloret, A. Why Women Have More Alzheimer's Disease Than Men: Gender and Mitochondrial Toxicity of Amyloid-β Peptide. *J Alzheimer's Dis* **20,** S527–S533 (2010).

28. Buzdar, A. & Howell, A. Advances in aromatase inhibition: clinical efficacy and tolerability in the treatment of breast cancer. *Clin Cancer Res Official J Am Assoc Cancer Res* **7,** 2620–35 (2001).

29. Miller, W. L. & Auchus, R. J. The Molecular Biology, Biochemistry, and Physiology of Human Steroidogenesis and Its Disorders. *Endocr Rev* **32,** 81–151 (2011).

30. Cui, J., Shen, Y. & Li, R. Estrogen synthesis and signaling pathways during aging: from periphery to brain. *Trends Mol Med* **19,** 197–209 (2013).

31. Taylor, S. E., Martin-Hirsch, P. L. & Martin, F. L. Oestrogen receptor splice variants in the pathogenesis of disease. *Cancer Lett* **288,** 133–148 (2010).

32. Speirs, V., Skliris, G., Burdall, S. & Carder, P. Distinct expression patterns of ER and ER in normal human mammary gland. *J Clin Pathol* **55,** 371–374 (2002).

33. Couse, J. F. & Korach, K. S. Estrogen Receptor Null Mice: What Have We Learned and Where Will They Lead Us? *Endocr Rev* **20,** 358–417 (1999).

34. Fisher, C. R., Graves, K. H., Parlow, A. F. & Simpson, E. R. Characterization of mice deficient in aromatase (ArKO) because of targeted disruption of the cyp19 gene. *Proc National Acad Sci* **95,** 6965–6970 (1998).

35. Krege, J. H. *et al.* Generation and reproductive phenotypes of mice lacking estrogen receptor β. *Proc National Acad Sci* **95,** 15677–15682 (1998).

36. Carroll, J. S. *et al.* Genome-wide analysis of estrogen receptor binding sites. *Nat Genet* **38,** ng1901 (2006).

37. Zhao, C., Putnik, M., Gustafsson, J.-Å. & Dahlman-Wright, K. Microarray analysis of altered gene expression in ERβ-overexpressing HEK293 cells. *Endocrine* **36,** 224–232 (2009).

38. Leygue, E., Dotzlaw, H., Watson, P. & Murphy, L. Altered estrogen receptor alpha and beta messenger RNA expression during human breast tumorigenesis. *Cancer Res* **58,** 3197–201 (1998).

39. Treeck, O., Lattrich, C., Springwald, A. & Ortmann, O. Estrogen receptor beta exerts growth-inhibitory effects on human mammary epithelial cells. *Breast Cancer Res Tr* **120,** 557–565 (2010).

40. Parker, J. S. *et al.* Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *J Clin Oncol* **27,** 1160–1167 (2009).

41. Ali, S., Buluwela, L. & Coombes, C. R. Antiestrogens and Their Therapeutic Applications in Breast Cancer

and Other Diseases. *Annu Rev Med* **62,** 217–232 (2011).

42. Johns, N. & Dixon, J. M. Should patients with early breast cancer still be offered the choice of breast conserving surgery or mastectomy? *European J Surg Oncol Ejso* **42,** 1636–1641 (2016).

43. Castaneda, S. A. & Strasser, J. Updates in the Treatment of Breast Cancer with Radiotherapy. *Surg Oncol Clin N Am* **26,** 371–382 (2017).

44. (EBCTCG), E. *et al.* Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10 801 women in 17 randomised trials. *Lancet* **378,** 1707–1716 (2011).

45. Beatson, G. T. On the Treatment of Inoperable Cases of Carcinoma of the Mamma: Suggestions for a New Method of Treatment, with Illustrative Cases. *Transactions Medico-chirurgical Soc Edinb* **15,** 153–179 (1896).

46. Palmieri, C., Patten, D. K., Januszewski, A., Zucchini, G. & Howell, S. J. Breast cancer: Current and future endocrine therapies. *Mol Cell Endocrinol* **382,** 695–723 (2014).

47. Harper, K. L. *et al.* Mechanism of early dissemination and metastasis in Her2+ mammary cancer. *Nature* **540,** 588 (2016).

48. Hosseini, H. *et al.* Early dissemination seeds metastasis in breast cancer. *Nature* **540,** 552 (2016).

49. Davies, C. *et al.* Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet* **381,** 805–816 (2013).

50. Ferreira, A. R. *et al.* Effectiveness of adjuvant ovarian function suppression in premenopausal women with early breast cancer: a multicenter cohort study. *Clin Breast Cancer* (2019). doi:10.1016/j.clbc.2019.06.003

51. (EBCTCG), E. *et al.* Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. *The Lancet* **386,** 1341–1352 (2015).

52. Pan, H. *et al.* 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. *The New England Journal of Medicine* **377,** 1836–1846 (2017).

53. Nguyen, V. T. *et al.* Differential epigenetic reprogramming in response to specific endocrine therapies promotes cholesterol biosynthesis and cellular invasion. *Nature Communications* **6,** 10044 (2015).

54. Magnani, L. *et al.* Acquired CYP19A1 amplification is an early specific mechanism of aromatase inhibitor resistance in ERα metastatic breast cancer. *Nature Genetics* **49,** 444 (2017).

55. Varešlija, D. *et al.* Adaptation to AI Therapy in Breast Cancer Can Induce Dynamic Alterations in ER Activity Resulting in Estrogen-Independent Metastatic Tumors. *Am Assoc Cancer Res* **22,** 2765–2777 (2016).

56. (EBCTCG), E. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* **365,** 1687–1717 (2005).

57. (EBCTCG), E. *et al.* Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials. *Lancet Oncol* **19,** 27–39 (2018).

58. Guan, J. *et al.* Therapeutic Ligands Antagonize Estrogen Receptor Function by Impairing Its Mobility. *Cell* **178,** 949-963.e18 (2019).

59. Hanahan, D. & Weinberg, R. A. The Hallmarks of Cancer. *Cell* **100,** 57–70 (2000).

60. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144,** 646–674 (2011).

61. Massagué, J. & Obenauf, A. C. Metastatic colonization by circulating tumour cells. *Nature* **529,** 298–306 (2016).

62. Talmadge, J. E. & Fidler, I. J. AACR Centennial Series: The Biology of Cancer Metastasis: Historical Perspective. *Cancer Res* **70,** 5649–5669 (2010).

63. Friedl, P. & Gilmour, D. Collective cell migration in morphogenesis, regeneration and cancer. *Nat Rev Mol Cell Bio* **10,** nrm2720 (2009).

64. Ye, X. & Weinberg, R. A. Epithelial–Mesenchymal Plasticity: A Central Regulator of Cancer Progression. *Trends Cell Biol* **25,** 675–686 (2015).

65. ZEIDMAN, I. & BUSS, J. Transpulmonary passage of tumor cell emboli. *Cancer Res* **12,** 731–3 (1952).

66. Liotta, L., Saidel, M. & Kleinerman, J. The significance of hematogenous tumor cell clumps in the metastatic process. *Cancer Res* **36,** 889–94 (1976).

67. Cheung, K. J. *et al.* Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proceedings of the National Academy of Sciences* **113,** E854–E863 (2016).

68. Aceto, N. *et al.* Circulating Tumor Cell Clusters Are Oligoclonal Precursors of Breast Cancer Metastasis. *Cell* **158,** 1110–1122 (2014).

69. Cheung, K. J., Gabrielson, E., Werb, Z. & Ewald, A. J. Collective Invasion in Breast Cancer Requires a Conserved Basal Epithelial Program. *Cell* **155,** 1639–1651 (2013).

70. Fischer, K. R. *et al.* Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* **527,** 472 (2015).

71. Eberlé, D., Hegarty, B., Bossard, P., Ferré, P. & Foufelle, F. SREBP transcription factors: master regulators of lipid homeostasis. *Biochimie* **86,** 839–848 (2004).

72. Wang, X. Nuclear protein that binds sterol regulatory element of low density lipoprotein receptor promoter. II. Purification and characterization. (1993). Available at: http://www.jbc.org/content/268/19/14497.long.

73. Yokoyama, C. *et al.* SREBP-1, a basic-helix-loop-helix-leucine zipper protein that controls transcription of the low density lipoprotein receptor gene. *Cell* **75,** 187–197 (1993).

74. Hua, X., Wu, J., Goldstein, J. L., Brown, M. S. & Hobbs, H. H. Structure of the human gene encoding sterol regulatory element binding protein-1 (SREBF1) and localization of SREBF1 and SREBF2 to chromosomes 17p11.2 and 22q13. *Genomics* **25,** 667–673 (1995).

75. Osborne, T. F. & Espenshade, P. J. Evolutionary conservation and adaptation in the mechanism that regulates SREBP action: what a long, strange tRIP it's been. *Genes & Development* **23,** 2578–2591 (2009).

76. Miserez, A. R., Cao, G., Probst, L. C. & Hobbs, H. H. Structure of the Human Gene Encoding Sterol Regulatory Element Binding Protein 2 (SREBF2). *Genomics* **40,** 31–40 (1997).

77. Toth, J. I., Datta, S., Athanikar, J. N., Freedman, L. P. & Osborne, T. F. Selective Coactivator Interactions in Gene Activation by SREBP-1a and -1c. *Molecular and Cellular Biology* **24,** 8288–8300 (2004).

78. Shimomura, I., Shimano, H., Horton, J., Goldstein, J. & Brown. Differential expression of exons 1a and 1c in mRNAs for sterol regulatory element binding protein-1 in human and mouse organs and cultured cells. *Journal of Clinical Investigation* **99,** 838–845 (1997).

79. Seegmiller, A. C. *et al.* The SREBP Pathway in Drosophila Regulation by Palmitate, Not Sterols. *Dev Cell* **2,** 229–238 (2002).

80. Hughes, A. L., Todd, B. L. & Espenshade, P. J. SREBP Pathway Responds to Sterols and Functions as an Oxygen Sensor in Fission Yeast. *Cell* **120,** 831–842 (2005).

81. Brown, M. S. & Goldstein, J. L. The SREBP Pathway: Regulation of Cholesterol Metabolism by Proteolysis

of a Membrane-Bound Transcription Factor. *Cell* **89,** 331–340 (1997).

82. Horton, J. D., Goldstein, J. L. & Brown, M. S. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *Journal of Clinical Investigation* **109,** 1125–1131 (2002).

83. Shimano, H. *et al.* Elevated levels of SREBP-2 and cholesterol synthesis in livers of mice homozygous for a targeted disruption of the SREBP-1 gene. *Journal of Clinical Investigation* **100,** 2115–2124 (1997).

84. Horton, J. Sterol regulatory element-binding proteins: transcriptional activators of lipid synthesis. *Biochemical Society Transactions* **30,** 1091–1095 (2002).

85. Nohturfft, A., Brown, M. S. & Goldstein, J. L. Topology of SREBP Cleavage-activating Protein, a Polytopic Membrane Protein with a Sterol-sensing Domain. *Journal of Biological Chemistry* **273,** 17243–17250 (1998).

86. Sakai, J. *et al.* Identification of Complexes between the COOH-terminal Domains of Sterol Regulatory Element-binding Proteins (SREBPs) and SREBP Cleavage-Activating Protein. *Journal of Biological Chemistry* **272,** 20213–20221 (1997).

87. Brown, A. J., Sun, L., Feramisco, J. D., Brown, M. S. & Goldstein, J. L. Cholesterol Addition to ER Membranes Alters Conformation of SCAP, the SREBP Escort Protein that Regulates Cholesterol Metabolism. *Molecular Cell* **10,** 237–245 (2002).

88. Yang, T. *et al.* Crucial Step in Cholesterol Homeostasis Sterols Promote Binding of SCAP to INSIG-1, a Membrane Protein that Facilitates Retention of SREBPs in ER. *Cell* **110,** 489–500 (2002).

89. Adams, C. M. *et al.* Cholesterol and 25-Hydroxycholesterol Inhibit Activation of SREBPs by Different Mechanisms, Both Involving SCAP and Insigs. *Journal of Biological Chemistry* **279,** 52772–52780 (2004).

90. DeBose-Boyd, R. A. *et al.* Transport-Dependent Proteolysis of SREBP Relocation of Site-1 Protease from Golgi to ER Obviates the Need for SREBP Transport to Golgi. *Cell* **99,** 703–712 (1999).

91. Ikonen, E. Cellular cholesterol trafficking and compartmentalization. *Nature Reviews Molecular Cell Biology* **9,** 125–138 (2008).

92. Bennett, M. K., Lopez, J. M., Sanchez, H. B. & Osborne, T. F. Sterol Regulation of Fatty Acid Synthase Promoter COORDINATE FEEDBACK REGULATION OF TWO MAJOR LIPID PATHWAYS. *Journal of Biological Chemistry* **270,** 25578–25583 (1995).

93. Osborne, T. F. & LaMorte, V. J. Molecular Aspects in Feedback Regulation of Gene Expression by Cholesterol in Mammalian Cells. *Methods* **16,** 42–48 (1998).

94. Wang, X., Sato, R., Brown, M. S., Hua, X. & Goldstein, J. L. SREBP-1, a membrane-bound transcription factor released by sterol-regulated proteolysis. *Cell* **77,** 53–62 (1994).

95. Shao, W. & Espenshade, P. J. Expanding Roles for SREBP in Metabolism. *Cell Metabolism* **16,** 414–419 (2012).

96. Inoue, J., Sato, R. & Maeda, M. Multiple DNA elements for sterol regulatory element-binding protein and NF-Y are responsible for sterol-regulated transcription of the genes for human 3-hydroxy-3-methylglutaryl coenzyme A synthase and squalene synthase. *J Biochem* **123,** 1191–8 (1998).

97. Sundqvist, A. *et al.* Control of lipid metabolism by phosphorylation-dependent degradation of the SREBP family of transcription factors by SCFFbw7. *Cell Metabolism* **1,** 379–391 (2005).

98. Bengoechea-Alonso, M. T., Punga, T. & Ericsson, J. Hyperphosphorylation regulates the activity of SREBP1 during mitosis. *P Natl Acad Sci Usa* **102,** 11681–11686 (2005).

99. Punga, T., Bengoechea-Alonso, M. T. & Ericsson, J. Phosphorylation and Ubiquitination of the Transcription Factor Sterol Regulatory Element-binding Protein-1 in Response to DNA Binding. *Journal of Biological Chemistry* **281,** 25278–25286 (2006).

100. Bengoechea-Alonso, M. T. & Ericsson, J. A Phosphorylation Cascade Controls the Degradation of Active SREBP1. *J Biol Chem* **284,** 5885–5895 (2009).

101. Bengoechea-Alonso, M. & Ericsson, J. The phosphorylation-dependent regulation of nuclear SREBP1 during mitosis links lipid metabolism and cell growth. *Cell Cycle* **15,** 2753–2765 (2016).

102. Giandomenico, V., Simonsson, M., Grönroos, E. & Ericsson, J. Coactivator-Dependent Acetylation Stabilizes Members of the SREBP Family of Transcription Factors. *Molecular and Cellular Biology* **23,** 2587–2599 (2003).

103. Hirano, Y., Murata, S., Tanaka, K., Shimizu, M. & Sato, R. Sterol Regulatory Element-binding Proteins Are Negatively Regulated through SUMO-1 Modification Independent of the Ubiquitin/26 S Proteasome Pathway. *J Biol Chem* **278,** 16809–16819 (2003).

104. Hirano, Y., Yoshida, M., Shimizu, M. & Sato, R. Direct Demonstration of Rapid Degradation of Nuclear Sterol Regulatory Element-binding Proteins by the Ubiquitin-Proteasome Pathway. *Journal of Biological Chemistry* **276,** 36431–36437 (2001).

105. Sundqvist, A. & Ericsson, J. Transcription-dependent degradation controls the stability of the SREBP family of transcription factors. *Proceedings of the National Academy of Sciences* **100,** 13833–13838 (2003).

106. Kuan, Y.-C. *et al.* Heat Shock Protein 90 Modulates Lipid Homeostasis by Regulating the Stability and Function of Sterol Regulatory Element-binding Protein (SREBP) and SREBP Cleavage-activating Protein. *J Biol Chem* **292,** 3016–3028 (2017).

107. Gong, Y. *et al.* Sterol-regulated ubiquitination and degradation of Insig-1 creates a convergent mechanism for feedback control of cholesterol synthesis and uptake. *Cell Metabolism* **3,** 15–24 (2006).

108. Reed, B. D., Charos, A. E., Szekely, A. M., Weissman, S. M. & Snyder, M. Genome-Wide Occupancy of SREBP1 and Its Partners NFY and SP1 Reveals Novel Functional Roles and Combinatorial Regulation of Distinct Classes of Genes. *Plos Genet* **4,** e1000133 (2008).

109. Chen, G., Liang, G., Ou, J., Goldstein, J. L. & Brown, M. S. Central role for liver X receptor in insulin-mediated activation of Srebp-1c transcription and stimulation of fatty acid synthesis in liver. *Proceedings of the National Academy of Sciences of the United States of America* **101,** 11245–11250 (2004).

110. Repa, J. J. *et al.* Regulation of mouse sterol regulatory element-binding protein-1c gene (SREBP-1c) by oxysterol receptors, LXRα and LXRβ. *Genes & Development* **14,** 2819–2830 (2000).

111. Ou, J. *et al.* Unsaturated fatty acids inhibit transcription of the sterol regulatory element-binding protein-1c (SREBP-1c) gene by antagonizing ligand-dependent activation of the LXR. *Proceedings of the National Academy of Sciences* **98,** 6027–6032 (2001).

112. Radhakrishnan, A., Ikeda, Y., Kwon, H., Brown, M. S. & Goldstein, J. L. Sterol-regulated transport of SREBPs from endoplasmic reticulum to Golgi: Oxysterols block transport by binding to Insig. *Proceedings of the National Academy of Sciences* **104,** 6511–6518 (2007).

113. Engelking, L. J. *et al.* Severe facial clefting in Insig-deficient mouse embryos caused by sterol accumulation and reversed by lovastatin. *Journal of Clinical Investigation* **116,** 2356–2365 (2006).

114. Sun, L.-P., Seemann, J., Goldstein, J. L. & Brown, M. S. Sterol-regulated transport of SREBPs from endoplasmic reticulum to Golgi: Insig renders sorting signal in Scap inaccessible to COPII proteins. *Proceedings of the National Academy of Sciences* **104,** 6519–6526 (2007).

115. Bakan, I. & Laplante, M. Connecting mTORC1 signaling to SREBP-1 activation. *Current Opinion in Lipidology* **23,** 226 (2012).

116. Portsmann, T., Santos, C. R., Lewis, C., Griffiths, B. & Schulze, A. A new player in the orchestra of cell growth: SREBP activity is regulated by mTORC1 and contributes to the regulation of cell and organ size.

Biochemical Society Transactions **37,** 278–283 (2009).

117. Porstmann, T. *et al.* SREBP Activity Is Regulated by mTORC1 and Contributes to Akt-Dependent Cell Growth. *Cell Metabolism* **8,** 224–236 (2008).

118. Lewis, C. A., Griffiths, B., Santos, C. R., Pende, M. & Schulze, A. Regulation of the SREBP transcription factors by mTORC1. *Biochemical Society Transactions* **39,** 495–499 (2011).

119. Ricoult, S. J. & Manning, B. D. The multifaceted role of mTORC1 in the control of lipid metabolism. *EMBO reports* **14,** 242–251 (2013).

120. Porstmann, T. *et al.* PKB/Akt induces transcription of enzymes involved in cholesterol and fatty acid biosynthesis via activation of SREBP. *Oncogene* **24,** 6465 (2005).

121. Krycer, J. R., Sharpe, L. J., Luu, W. & Brown, A. J. The Akt–SREBP nexus: cell signaling meets lipid metabolism. *Trends in Endocrinology & Metabolism* **21,** 268–276 (2010).

122. Owen, J. L. *et al.* Insulin stimulation of SREBP-1c processing in transgenic rat hepatocytes requires p70 S6-kinase. *Proc National Acad Sci* **109,** 16184–16189 (2012).

123. Han, J. *et al.* The CREB coactivator CRTC2 controls hepatic lipid metabolism by regulating SREBP1. *Nature* **524,** 243 (2015).

124. Peterson, T. R. *et al.* mTOR Complex 1 Regulates Lipin 1 Localization to Control the SREBP Pathway. *Cell* **146,** 408–420 (2011).

125. Shimano, H. & Sato, R. SREBP-regulated lipid metabolism: convergent physiology — divergent pathophysiology. *Nature Reviews Endocrinology* **13,** 710–730 (2017).

126. Zadra, G., Batista, J. L. & Loda, M. Dissecting the Dual Role of AMPK in Cancer: From Experimental to Human Studies. *Mol Cancer Res* **13,** 1059–1072 (2015).

127. Li, Y. *et al.* AMPK Phosphorylates and Inhibits SREBP Activity to Attenuate Hepatic Steatosis and Atherosclerosis in Diet-Induced Insulin-Resistant Mice. *Cell Metabolism* **13,** 376–388 (2011).

128. Khan, A. S. & Frigo, D. E. A spatiotemporal hypothesis for the regulation, role, and targeting of AMPK in prostate cancer. *Nat Rev Urol* **14,** 164–180 (2017).

129. Dong, Q. *et al.* Phosphorylation of sterol regulatory element binding protein-1a by protein kinase A (PKA) regulates transcriptional activity. *Biochem Bioph Res Co* **449,** 449–454 (2014).

130. Lu, M. & Shyy, J. Y. Sterol regulatory element-binding protein 1 is negatively modulated by PKA phosphorylation. *Am J Physiol-cell Ph* **290,** C1477–C1486 (2006).

131. Yamamoto, T. *et al.* Protein Kinase A Suppresses Sterol Regulatory Element-binding Protein-1C Expression via Phosphorylation of Liver X Receptor in the Liver. *J Biol Chem* **282,** 11687–11695 (2007).

132. Nakakuki, M. *et al.* A novel processing system of sterol regulatory element-binding protein-1c regulated by polyunsaturated fatty acid. *Journal of biochemistry* **155,** 301–13 (2014).

133. Takeuchi, Y. *et al.* Polyunsaturated Fatty Acids Selectively Suppress Sterol Regulatory Element-binding Protein-1 through Proteolytic Processing and Autoloop Regulatory Circuit. *J Biol Chem* **285,** 11681–11691 (2010).

134. Yahagi, N. *et al.* A Crucial Role of Sterol Regulatory Element-binding Protein-1 in the Regulation of Lipogenic Gene Expression by Polyunsaturated Fatty Acids. *Journal of Biological Chemistry* **274,** 35840–35844 (1999).

135. Gilardi, F. *et al.* Genome-Wide Analysis of SREBP1 Activity around the Clock Reveals Its Combined Dependency on Nutrient and Circadian Signals. *Plos Genet* **10,** e1004155 (2014).

136. Masri, S. *et al.* Lung Adenocarcinoma Distally Rewires Hepatic Circadian Homeostasis. *Cell* **165,** 896–909 (2016).

137. Guo, D., Bell, E., Mischel, P. & Chakravarti, A. Targeting SREBP-1-driven lipid metabolism to treat cancer. *Current pharmaceutical design* **20,** 2619–26 (2014).

138. Yamauchi, Y., Furukawa, K., Hamamura, K. & Furukawa, K. Positive Feedback Loop Between PI3K-Akt-mTORC1 Signaling and the Lipogenic Pathway Boosts Akt Signaling: Induction of the Lipogenic Pathway by a Melanoma Antigen. *Cancer Research* **71,** 4989–4997 (2011).

139. Ricoult, S., Yecies, J., Ben-Sahra, I. & Manning, B. Oncogenic PI3K and K-Ras stimulate de novo lipid synthesis through mTORC1 and SREBP. *Oncogene* **35,** 1250–1260 (2016).

140. Sanchez-Alvarez, M. *et al.* Signaling Networks Converge on TORC1-SREBP Activity to Promote Endoplasmic Reticulum Homeostasis. *Plos One* **9,** e101164 (2014).

141. Griffiths, B. *et al.* Sterol regulatory element binding protein-dependent regulation of lipid synthesis supports cell survival and tumor growth. *Cancer & Metabolism* **1,** 1–21 (2012).

142. Lee, J. & Ye, J. Proteolytic Activation of Sterol Regulatory Element-binding Protein Induced by Cellular Stress through Depletion of Insig-1. *J Biol Chem* **279,** 45257–45265 (2004).

143. Ye, J. *et al.* ER Stress Induces Cleavage of Membrane-Bound ATF6 by the Same Proteases that Process SREBPs. *Molecular Cell* **6,** 1355–1364 (2000).

144. Kim, J. *et al.* ER Stress Drives Lipogenesis and Steatohepatitis via Caspase-2 Activation of S1P. *Cell* (2018). doi:10.1016/j.cell.2018.08.020

145. Lee, J., Zhang, X., Feramisco, J. D., Gong, Y. & Ye, J. Unsaturated Fatty Acids Inhibit Proteasomal Degradation of Insig-1 at a Postubiquitination Step. *Journal of Biological Chemistry* **283,** 33772–33783 (2008).

146. Lee, C. *et al.* Tumor metastasis to lymph nodes requires YAP-dependent metabolic adaptation. *Science* **363,** 644–649 (2019).

147. Nguyen, C. & Yi, C. YAP/TAZ Signaling and Resistance to Cancer Therapy. *Trends in Cancer* (2019). doi:10.1016/j.trecan.2019.02.010

148. Sorrentino, G. *et al.* Metabolic control of YAP and TAZ by the mevalonate pathway. *Nature Cell Biology* **16,** 357–366 (2014).

149. Li, X. *et al.* SREBP-2 promotes stem cell-like properties and metastasis by transcriptional activation of c-Myc in prostate cancer. *Oncotarget* **7,** 12869–12884 (2016).

150. Wu, Y. *et al.* Srebp-1 Interacts with c-Myc to Enhance Somatic Cell Reprogramming. *Stem Cells* **34,** 83–92 (2016).

151. Gouw, A. M. *et al.* The MYC Oncogene Cooperates with Sterol-Regulated Element-Binding Protein to Regulate Lipogenesis Essential for Neoplastic Growth. *Cell Metab* (2019). doi:10.1016/j.cmet.2019.07.012

152. Schulze, A. & Harris, A. L. How cancer metabolism is tuned for proliferation and vulnerable to disruption. *Nature* **491,** 364–373 (2012).

153. Jones, N. P. & Schulze, A. Targeting cancer metabolism – aiming at a tumour's sweet-spot. *Drug Discovery Today* **17,** 232–241 (2012).

154. Freed-Pastor, W. A. *et al.* Mutant p53 Disrupts Mammary Tissue Architecture via the Mevalonate Pathway. *Cell* **148,** 244–258 (2012).

155. Bathaie, Z. S., Ashrafi, M., Azizian, M. & Tamanoi, F. Mevalonate Pathway and Human Cancers. *Current*

molecular pharmacology **09,** 1–1 (2016).

156. Boettcher, S. *et al.* A dominant-negative effect drives selection of TP53 missense mutations in myeloid malignancies. *Science* **365,** 599–604 (2019).

157. Flavin, R., Zadra, G. & Loda, M. Metabolic alterations and targeted therapies in prostate cancer. *The Journal of Pathology* **223,** 284–295 (2011).

158. Scaglia, N., Tyekucheva, S., Zadra, G., Photopoulos, C. & Loda, M. De novo fatty acid synthesis at the mitotic exit is required to complete cellular division. *Cell cycle (Georgetown, Tex.)* **13,** 859–68 (2014).

159. Zadra, G. *et al.* A novel direct activator of AMPK inhibits prostate cancer growth by blocking lipogenesis. *Embo Mol Med* **6,** 519–538 (2014).

160. Rottiers, V. & Näär, A. M. MicroRNAs in metabolism and metabolic disorders. *Nature Reviews Molecular Cell Biology* **13,** 239 (2012).

161. Aryal, B., Singh, A. K., Rotllan, N., Price, N. & Fernández-Hernando, C. MicroRNAs and lipid metabolism. *Current Opinion in Lipidology* **28,** 273–280 (2017).

162. Walker, A. K. & Näär, A. M. SREBPs: regulators of cholesterol/lipids as therapeutic targets in metabolic disorders, cancers and viral diseases. *Clinical Lipidology* **7,** 27–36 (2012).

163. Lin, Y. *et al.* MicroRNA-33b Inhibits Breast Cancer Metastasis by Targeting HMGA2, SALL4 and Twist1. *Sci Rep-uk* **5,** srep09995 (2015).

164. Manning, B. D., Kelsey, I., Torrence, M., Zbinden, M. & Byles, V. mTORC1 suppresses PIM3 expression via miR-33 encoded by the SREBP loci. *Scientific Reports* **7,** 1 (2017).

165. Li, X. *et al.* MicroRNA-185 and 342 Inhibit Tumorigenicity and Induce Apoptosis through Blockade of the SREBP Metabolic Pathway in Prostate Cancer Cells. *Plos One* **8,** e70987 (2013).

166. Ru, P. *et al.* Feedback Loop Regulation of SCAP/SREBP-1 by miR-29 Modulates EGFR Signaling-Driven Glioblastoma Growth. *Cell Reports* **16,** 1527–1535 (2016).

167. Kamisuki, S. *et al.* A Small Molecule That Blocks Fat Synthesis By Inhibiting the Activation of SREBP. *Chemistry & Biology* **16,** 882–892 (2009).

168. (邵威) W., Machamer, C. E. & Espenshade, P. J. Fatostatin blocks ER exit of SCAP but inhibits cell growth in a SCAP-independent manner. *Journal of Lipid Research* **57,** 1564–1573 (2016).

169. Gholkar, A. A. *et al.* Fatostatin Inhibits Cancer Cell Proliferation by Affecting Mitotic Microtubule Spindle Assembly and Cell Division. *Journal of Biological Chemistry* **291,** 17001–17008 (2016).

170. Brovkovych, V. *et al.* Fatostatin induces pro- and anti-apoptotic lipid accumulation in breast cancer. *Oncogenesis* **7,** 66 (2018).

171. Li, X., Chen, Y.-T., Hu, P. & Huang, W.-C. Fatostatin Displays High Antitumor Activity in Prostate Cancer by Blocking SREBP-Regulated Metabolic Pathways and Androgen Receptor Signaling. *Molecular Cancer Therapeutics* **13,** 855–866 (2014).

172. Li, X., Wu, J., Chung, L. & Huang, W.-C. Anti-cancer efficacy of SREBP inhibitor, alone or in combination with docetaxel, in prostate cancer harboring p53 mutations. *Oncotarget* **6,** 41018–41032 (2015).

173. Siqingaowa, Sekar, S., Gopalakrishnan, V. & Taghibiglou, C. Sterol regulatory element-binding protein 1 inhibitors decrease pancreatic cancer cell viability and proliferation. *Biochemical and Biophysical Research Communications* **488,** 136–140 (2017).

174. Tang, J.-J. *et al.* Inhibition of SREBP by a Small Molecule, Betulin, Improves Hyperlipidemia and Insulin Resistance and Reduces Atherosclerotic Plaques. *Cell Metabolism* **13,** 44–56 (2011).

175. Li, N. *et al.* Inhibition of the sterol regulatory element-binding protein pathway suppresses hepatocellular carcinoma by repressing inflammation in mice. *Hepatology* **65,** 1936–1947 (2017).

176. Miyata, S., Inoue, J., Shimizu, M. & Sato, R. Xanthohumol Improves Diet-induced Obesity and Fatty Liver by Suppressing Sterol Regulatory Element-binding Protein (SREBP) Activation. *Journal of Biological Chemistry* **290,** 20565–20579 (2015).

177. Guan, M. *et al.* Nelfinavir Induces Liposarcoma Apoptosis through Inhibition of Regulated Intramembrane Proteolysis of SREBP-1 and ATF6. *Clinical Cancer Research* **17,** 1796–1806 (2011).

178. Nambiar, D. K., Deep, G., Singh, R. P., Agarwal, C. & Agarwal, R. Silibinin inhibits aberrant lipid metabolism, proliferation and emergence of androgen-independence in prostate cancer cells via primarily targeting the sterol response element binding protein 1. *Oncotarget* **5,** 10017–10033 (2014).

179. Zhao, X. *et al.* Inhibition of SREBP Transcriptional Activity by a Boron-Containing Compound Improves Lipid Homeostasis in Diet-Induced Obesity. *Diabetes* **63,** 2464–2473 (2014).

180. Mo, H. & Elson, C. E. Studies of the isoprenoid-mediated inhibition of mevalonate synthesis applied to cancer chemotherapy and chemoprevention. *Experimental biology and medicine (Maywood, N.J.)* **229,** 567–85 (2004).

181. Ahern, T. P. *et al.* Statin Prescriptions and Breast Cancer Recurrence Risk: A Danish Nationwide Prospective Cohort Study. *JNCI: Journal of the National Cancer Institute* **103,** 1461–1468 (2011).

182. Platz, E. A., Clinton, S. K. & Giovannucci, E. Association between plasma cholesterol and prostate cancer in the PSA era. *International Journal of Cancer* **123,** 1693–1698 (2008).

183. Platz, E. A. *et al.* Men with Low Serum Cholesterol Have a Lower Risk of High-Grade Prostate Cancer in the Placebo Arm of the Prostate Cancer Prevention Trial. *Cancer Epidemiology Prev Biomarkers* **18,** 2807–2813 (2009).

184. Hamilton, R. J. *et al.* Statin medication use and the risk of biochemical recurrence after radical prostatectomy. *Cancer* **116,** 3389–3398 (2010).

185. Nielsen, S. F., Nordestgaard, B. G. & Bojesen, S. E. Statin Use and Reduced Cancer-Related Mortality. *New Engl J Medicine* **367,** 1792–1802 (2012).

186. Geybels, M. S. *et al.* Statin Use in Relation to Prostate Cancer Outcomes in a Population-based Patient Cohort Study. *Prostate* **73,** 1214–1222 (2013).

187. Mucci, L. A. & Stampfer, M. J. Mounting Evidence for Prediagnostic Use of Statins in Reducing Risk of Lethal Prostate Cancer. *J Clin Oncol* **32,** 1–2 (2013).

188. Yu, O. *et al.* Use of Statins and the Risk of Death in Patients With Prostate Cancer. *J Clin Oncol* **32,** 5–11 (2013).

189. Agalliu, I., Salinas, C. A., Hansten, P. D., Ostrander, E. A. & Stanford, J. L. Statin Use and Risk of Prostate Cancer: Results from a Population-based Epidemiologic Study. *Am J Epidemiol* **168,** 250–260 (2008).

190. Coogan, P. F., Kelly, J., Strom, B. L. & Rosenberg, L. Statin and NSAID use and prostate cancer risk. *Pharmacoepidem Dr S* **19,** 752–755 (2010).

191. Haukka, J. *et al.* Incidence of cancer and statin usage—Record linkage study. *Int J Cancer* **126,** 279–284 (2010).

192. Chang, C., Ho, S., Chiu, H. & Yang, C. Statins increase the risk of prostate cancer: A population-based case–control study. *Prostate* **71,** 1818–1824 (2011).

193. YuPeng, L. *et al.* Cholesterol Levels in Blood and the Risk of Prostate Cancer: A Meta-analysis of 14

Prospective Studies. *Cancer Epidemiology and Prevention Biomarkers* **24,** 1086–1093 (2015).

194. Gazzerro, P. *et al.* Pharmacological Actions of Statins: A Critical Appraisal in the Management of Cancer. *Pharmacological Reviews* **64,** 102–146 (2012).

195. Simigdala, N. *et al.* Cholesterol biosynthesis pathway as a novel mechanism of resistance to estrogen deprivation in estrogen receptor-positive breast cancer. *Breast Cancer Research* **18,** 58 (2016).

196. Nelson, E. R., Chang, C. & nnell, D. P. Cholesterol and breast cancer pathophysiology. *Trends in Endocrinology & Metabolism* **25,** 649–655 (2014).

197. Nelson, E. R. *et al.* 27-Hydroxycholesterol Links Hypercholesterolemia and Breast Cancer Pathophysiology. *Science* **342,** 1094–1098 (2013).

198. Ettinger, S. L. *et al.* Dysregulation of Sterol Response Element-Binding Proteins and Downstream Effectors in Prostate Cancer during Progression to Androgen Independence. *Cancer Research* **64,** 2212–2221 (2004).

199. WARBURG, O. On respiratory impairment in cancer cells. *Sci New York N Y* **124,** 269–70 (1956).

200. Zadra, G., Photopoulos, C. & Loda, M. The fat side of prostate cancer. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1831,** 1518–1532 (2013).

201. Giunchi, F., Fiorentino, M. & Loda, M. The Metabolic Landscape of Prostate Cancer. *European Urology Oncology* **2,** 28–36 (2018).

202. Leon, C. G. *et al.* Alterations in cholesterol regulation contribute to the production of intratumoral androgens during progression to castration-resistant prostate cancer in a mouse xenograft model. *The Prostate* **70,** 390–400 (2010).

203. Rysman, E. *et al.* De novo Lipogenesis Protects Cancer Cells from Free Radicals and Chemotherapeutics by Promoting Membrane Lipid Saturation. *Cancer Res* **70,** 8117–8126 (2010).

204. Baron, A., Migita, T., Tang, D. & Loda, M. Fatty acid synthase: A metabolic oncogene in prostate cancer? *J Cell Biochem* **91,** 47–53 (2004).

205. Suburu, J. & Chen, Y. Q. Lipids and prostate cancer. *Prostaglandins & Other Lipid Mediators* **98,** 1–10 (2012).

206. Swinnen, J. V. *et al.* Selective activation of the fatty acid synthesis pathway in human prostate cancer. *Int J Cancer* **88,** 176–179 (2000).

207. Swinnen, J. V. *et al.* Overexpression of fatty acid synthase is an early and common event in the development of prostate cancer. *International Journal of Cancer* **98,** 19–22 (2002).

208. Chen, M. *et al.* An aberrant SREBP-dependent lipogenic program promotes metastatic prostate cancer. *Nature Genetics* 1–13 (2018). doi:10.1038/s41588-017-0027-2

209. Priolo, C. & Loda, M. Untargeted metabolomics for profiling oncogene-specific metabolic signatures of prostate cancer. *Molecular & Cellular Oncology* **2,** e1001197 (2015).

210. Little, J. L., Wheeler, F. B., Fels, D. R., Koumenis, C. & Kridel, S. J. Inhibition of Fatty Acid Synthase Induces Endoplasmic Reticulum Stress in Tumor Cells. *Cancer Res* **67,** 1262–1269 (2007).

211. Chen, J. *et al.* Compartmentalized activities of the pyruvate dehydrogenase complex sustain lipogenesis in prostate cancer. *Nature Genetics* 1–10 (2018). doi:10.1038/s41588-017-0026-3

212. Cheng, C., Wang, Z. & Chen, J. Targeting FASN in Breast Cancer and the Discovery of Promising Inhibitors from Natural Products Derived from Traditional Chinese Medicine. *Evid-based Compl Alt* **2014,** 232946 (2014).

213. Sowalsky A.G., *et al.* Neoadjuvant-Intensive Androgen Deprivation Therapy Selects for Prostate Tumor Foci with Diverse Subclonal Oncogenic Alterations. *Tumor Biology and Immunology* 10.1158/0008-5472 (2018).

214. Mostaghel, E. A. Steroid hormone synthetic pathways in prostate cancer. *Transl Androl Urology* **2,** 212–227 (2013).

215. Huang, W.-C., Li, X., Liu, J., Lin, J. & Chung, L. Activation of Androgen Receptor, Lipogenesis, and Oxidative Stress Converged by SREBP-1 Is Responsible for Regulating Growth and Progression of Prostate Cancer Cells. *Molecular Cancer Research* **10,** 133–142 (2012).

216. Migita, T. *et al.* Fatty Acid Synthase: A Metabolic Enzyme and Candidate Oncogene in Prostate Cancer. *JNCI: Journal of the National Cancer Institute* **101,** 519–532 (2009).

217. Fiorentino, M. *et al.* Overexpression of fatty acid synthase is associated with palmitoylation of Wnt1 and cytoplasmic stabilization of β-catenin in prostate cancer. *Lab Invest* **88,** labinvest200897 (2008).

218. Heemers, H. *et al.* Androgens stimulate lipogenic gene expression in prostate cancer cells by activation of the sterol regulatory element-binding protein cleavage activating protein/sterol regulatory element-binding protein pathway. *Mol Endocrinol Baltim Md* **15,** 1817–28 (2001).

219. Heemers, H. V., Verhoeven, G. & Swinnen, J. V. Androgen Activation of the Sterol Regulatory Element-Binding Protein Pathway: Current Insights. *Molecular Endocrinology* **20,** 2265–2277 (2006).

220. Huang, W.-C., Zhau, H. E. & Chung, L. W. Androgen Receptor Survival Signaling Is Blocked by Anti-β2-microglobulin Monoclonal Antibody via a MAPK/Lipogenic Pathway in Human Prostate Cancer Cells. *J Biol Chem* **285,** 7947–7956 (2010).

221. Zadra, G. *et al.* Inhibition of de novo lipogenesis targets androgen receptor signaling in castration-resistant prostate cancer. *Proceedings of the National Academy of Sciences* 201808834 (2018). doi:10.1073/pnas.1808834116

222. Dehm, S. M., Schmidt, L. J., Heemers, H. V., Vessella, R. L. & Tindall, D. J. Splicing of a Novel Androgen Receptor Exon Generates a Constitutively Active Androgen Receptor that Mediates Prostate Cancer Therapy Resistance. *Cancer Res* **68,** 5469–5477 (2008).

223. Chan, S., Li, Y. & Dehm, S. M. Androgen Receptor Splice Variants Activate Androgen Receptor Target Genes and Support Aberrant Prostate Cancer Cell Growth Independent of Canonical Androgen Receptor Nuclear Localization Signal. *J Biol Chem* **287,** 19736–19749 (2012).

224. Han, W. *et al.* Reactivation of androgen receptor-regulated lipid biosynthesis drives the progression of castration-resistant prostate cancer. *Oncogene* **37,** 710 (2017).

225. Krycer, J. & Brown, A. Does Changing Androgen Receptor Status during Prostate Cancer Development Impact upon Cholesterol Homeostasis? *Plos One* **8,** e54007 (2013).

226. Yang, T., Goldstein, J. L. & Brown, M. S. Overexpression of Membrane Domain of SCAP Prevents Sterols from Inhibiting SCAP/SREBP Exit from Endoplasmic Reticulum. *Journal of Biological Chemistry* **275,** 29881–29886 (2000).

227. Williams, K. J. *et al.* An Essential Requirement for the SCAP/SREBP Signaling Axis to Protect Cancer Cells from Lipotoxicity. *Cancer Research* **73,** 2850–2862 (2013).

228. Thysell, E. *et al.* Metabolomic Characterization of Human Prostate Cancer Bone Metastases Reveals Increased Levels of Cholesterol. *Plos One* **5,** e14175 (2010).

229. Locke, J. A. *et al.* Androgen Levels Increase by Intratumoral De novo Steroidogenesis during Progression of Castration-Resistant Prostate Cancer. *Cancer Res* **68,** 6407–6415 (2008).

230. Swinnen, J. V., Ulrix, W., Heyns, W. & Verhoeven, G. Coordinate regulation of lipogenic gene expression by androgens: Evidence for a cascade mechanism involving sterol regulatory element binding proteins. *Proc National Acad Sci* **94,** 12975–12980 (1997).

231. Montgomery, B. R. *et al.* Maintenance of Intratumoral Androgens in Metastatic Prostate Cancer: A Mechanism for Castration-Resistant Tumor Growth. *Cancer Research* **68,** 4447–4454 (2008).

232. Ellem, S. J., Schmitt, J. F., Pedersen, J. S., Frydenberg, M. & Risbridger, G. P. Local aromatase expression in human prostate is altered in malignancy. *J Clin Endocrinol Metabolism* **89,** 2434–41 (2004).

233. Herrmann, H. & Aebi, U. Intermediate Filaments: Structure and Assembly. *Csh Perspect Biol* **8,** a018242 (2016).

234. Chung, B.-M., Rotty, J. D. & Coulombe, P. A. Networking galore: intermediate filaments and cell migration. *Curr Opin Cell Biol* **25,** 600–612 (2013).

235. Schweizer, J. *et al.* New consensus nomenclature for mammalian keratins. *J Cell Biology* **174,** 169–174 (2006).

236. Rogers, M. A. *et al.* Characterization of New Members of the Human Type II Keratin Gene Family and a General Evaluation of the Keratin Gene Domain on Chromosome 12q13.13. *Journal of Investigative Dermatology* **124,** 536–544 (2005).

237. Hesse, M., Zimek, A., Weber, K. & Magin, T. M. Comprehensive analysis of keratin gene clusters in humans and rodents. *Eur J Cell Biol* **83,** 19–26 (2004).

238. Jacob, J. T., Coulombe, P. A., Kwan, R. & Omary, B. M. Types I and II Keratin Intermediate Filaments. *Cold Spring Harbor Perspectives in Biology* **10,** a018275 (2018).

239. Ma, L., Yamada, S., Wirtz, D. & Coulombe, P. A. A 'hot-spot' mutation alters the mechanical properties of keratin filament networks. *Nat Cell Biol* **3,** 503–506 (2001).

240. Wagner, O. I. *et al.* Softness, strength and self-repair in intermediate filament networks. *Experimental Cell Research* **313,** 2228–2235 (2007).

241. Sivaramakrishnan, S., DeGiulio, J. V., Lorand, L., Goldman, R. D. & Ridge, K. M. Micromechanical properties of keratin intermediate filament networks. *Proceedings of the National Academy of Sciences* **105,** 889–894 (2008).

242. Janmey, P. A. Mechanical properties of cytoskeletal polymers. *Curr Opin Cell Biol* **3,** 4–11 (1991).

243. Kim, S., Wong, P. & Coulombe, P. A. A keratin cytoskeletal protein regulates protein synthesis and epithelial cell growth. *Nature* **441,** 362 (2006).

244. Coulombe, P. A. & Omary, M. B. 'Hard' and 'soft' principles defining the structure, function and regulation of keratin intermediate filaments. *Curr Opin Cell Biol* **14,** 110–122 (2002).

245. Kim, S. & Coulombe, P. A. Intermediate filament scaffolds fulfill mechanical, organizational, and signaling functions in the cytoplasm. *Gene Dev* **21,** 1581–1597 (2007).

246. Szeverenyi, I. *et al.* The Human Intermediate Filament Database: comprehensive information on a gene family involved in many human diseases. *Hum Mutat* **29,** 351–360 (2008).

247. Omary, B. M., Coulombe, P. A. & McLean, W. H. Intermediate Filament Proteins and Their Associated Diseases. *New Engl J Medicine* **351,** 2087–2100 (2004).

248. Magin, T. M., Vijayaraj, P. & Leube, R. E. Structural and regulatory functions of keratins. *Exp Cell Res* **313,** 2021–2032 (2007).

249. Kellner, J. C. & Coulombe, P. A. Keratins and protein synthesis: the plot thickens. *J Cell Biology* **187,**

157–159 (2009).

250. DePianto, D. & Coulombe, P. A. Intermediate filaments and tissue repair. *Exp Cell Res* **301,** 68–76 (2004).

251. Cheng, F. & Eriksson, J. E. Intermediate Filaments and the Regulation of Cell Motility during Regeneration and Wound Healing. *Csh Perspect Biol* **9,** a022046 (2017).

252. Toivola, D. M., Strnad, P., Habtezion, A. & Omary, M. B. Intermediate filaments take the heat as stress proteins. *Trends Cell Biol* **20,** 79–91 (2010).

253. Vijayaraj, P. *et al.* Keratins regulate protein biosynthesis through localization of GLUT1 and -3 upstream of AMP kinase and Raptor. *J Cell Biology* **187,** 175–184 (2009).

254. Windoffer, R., Wöll, S., Strnad, P. & Leube, R. E. Identification of Novel Principles of Keratin Filament Network Turnover in Living Cells. *Mol Biol Cell* **15,** 2436–2448 (2004).

255. Windoffer, R., Beil, M., Magin, T. M. & Leube, R. E. Cytoskeleton in motion: the dynamics of keratin intermediate filaments in epithelia. *J Cell Biology* **194,** 669–678 (2011).

256. Karantza, V. Keratins in health and cancer: more than mere epithelial cell markers. *Oncogene* **30,** 127 (2011).

257. Snider, N. T. & Omary, B. M. Post-translational modifications of intermediate filament proteins: mechanisms and functions. *Nat Rev Mol Cell Bio* **15,** nrm3753 (2014).

258. Ku, N.-O. & Omary, B. M. A disease- and phosphorylation-related nonmechanical function for keratin 8. *J Cell Biology* **174,** 115–125 (2006).

259. Ramot, Y., Paus, R., Tiede, S. & Zlotogorski, A. Endocrine controls of keratin expression. *Bioessays* **31,** 389–399 (2009).

260. Lehr, H.-A., Folpe, A., Yaziji, H., Kommoss, F. & Gown, A. M. Cytokeratin 8 Immunostaining Pattern and E-CadherinExpression Distinguish Lobular From Ductal BreastCarcinoma. *Am J Clin Pathol* **114,** 190–196 (2000).

261. Weckermann, D. *et al.* Perioperative Activation of Disseminated Tumor Cells in Bone Marrow of Patients With Prostate Cancer. *J Clin Oncol* **27,** 1549–1556 (2009).

262. van de Rijn, M. *et al.* Expression of Cytokeratins 17 and 5 Identifies a Group of Breast Carcinomas with Poor Clinical Outcome. *The American Journal of Pathology* **161,** 1991–1996 (2002).

263. de Rudland, S. *et al.* Statistical Association of Basal Cell Keratins with Metastasis-Inducing Proteins in a Prognostically Unfavorable Group of Sporadic Breast Cancers. *The American Journal of Pathology* **179,** 1061–1072 (2011).

264. Joosse, S. A. *et al.* Changes in Keratin Expression during Metastatic Progression of Breast Cancer: Impact on the Detection of Circulating Tumor Cells. *Clinical Cancer Research* **18,** 993–1003 (2012).

265. Hendrix, M. J., Seftor, E. A., Chu, Y.-W., Trevor, K. T. & Seftor, R. E. Role of intermediate filaments in migration, invasion and metastasis. *Cancer Metast Rev* **15,** 507–525 (1996).

266. Paladini, R., Takahashi, K., Bravo, N. & Coulombe, P. Onset of re-epithelialization after skin injury correlates with a reorganization of keratin filaments in wound edge keratinocytes: defining a potential role for keratin 16. *J Cell Biology* **132,** 381–397 (1996).

267. Beil, M. *et al.* Sphingosylphosphorylcholine regulates keratin network architecture and visco-elastic properties of human cancer cells. *Nature Cell Biology* **5,** ncb1037 (2003).

268. Rolli, C. G., Seufferlein, T., Kemkemer, R. & Spatz, J. P. Impact of Tumor Cell Cytoskeleton Organization on Invasiveness and Migration: A Microchannel-Based Approach. *PLoS ONE* **5,** e8726 (2010).

269. Wolf, K. *et al.* Physical limits of cell migration: Control by ECM space and nuclear deformation and tuning by proteolysis and traction force. *J Cell Biology* **201,** 1069–1084 (2013).

270. Ferrari, R., Infante, E. & Chavrier, P. Nucleus–Invadopodia Duo During Cancer Invasion. *Trends in Cell Biology* (2018). doi:10.1016/j.tcb.2018.11.006

271. de Belo, V., Parente, J., Tanus-Santos, J. & Castro, M. M. Matrix metalloproteinase (MMP)-2 decreases calponin-1 levels and contributes to arterial remodeling in early hypertension. *Biochem Pharmacol* **118,** 50–58 (2016).

272. Eddy, R. J., Weidmann, M. D., Sharma, V. P. & Condeelis, J. S. Tumor Cell Invadopodia: Invasive Protrusions that Orchestrate Metastasis. *Trends Cell Biol* **27,** 595–607 (2017).

273. Aureille, J., Belaadi, N. & Guilluy, C. Mechanotransduction via the nuclear envelope: a distant reflection of the cell surface. *Curr Opin Cell Biol* **44,** 59–67 (2017).

274. Friedl, P. & Alexander, S. Cancer Invasion and the Microenvironment: Plasticity and Reciprocity. *Cell* **147,** 992–1009 (2011).

275. Seltmann, K., Fritsch, A. W., Käs, J. A. & Magin, T. M. Keratins significantly contribute to cell stiffness and impact invasive behavior. *Proceedings of the National Academy of Sciences* **110,** 18507–18512 (2013).

276. Cheung, K. J. & Ewald, A. J. Illuminating breast cancer invasion: diverse roles for cell–cell interactions. *Current Opinion in Cell Biology* **30,** 99–111 (2014).

277. Aceto N., Toner M., Maheswaran S., Haber D.A. En Route to Metastasis: Circulating Tumor Cell Clusters and Epithelial-to-Mesenchymal. Transition*Trends in Cancer*. 10.1016 (2015)

278. Yilmaz, M. & Christofori, G. EMT, the cytoskeleton, and cancer cell invasion. *Cancer and Metastasis Reviews* **28,** 15–33 (2009).

279. Langbein, L., Eckhart, L., Rogers, M. A., Praetzel-Wunder, S. & Schweizer, J. Against the Rules: Human Keratin K80 TWO FUNCTIONAL ALTERNATIVE SPLICE VARIANTS, K80 AND K80.1, WITH SPECIAL CELLULAR LOCALIZATION IN A WIDE RANGE OF EPITHELIA. *Journal of Biological Chemistry* **285,** 36909–36921 (2010).

280. Li, C. *et al.* Keratin 80 promotes migration and invasion of colorectal carcinoma by interacting with PRKDC via activating the AKT pathway. *Cell Death & Disease* **9,** 1009 (2018).

281. Perone, Y. & Magnani, L. Going off the grid: ERα breast cancer beyond estradiol. *Journal of Molecular Endocrinology* **57,** F1–F5 (2016).

282. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9,** 357 (2012).

283. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).

284. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol* **9,** R137 (2008).

285. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).

286. Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci Rep-uk* **9,** 9354 (2019).

287. Talevich, E., Shain, H. A., Botton, T. & Bastian, B. C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *Plos Comput Biol* **12,** e1004873 (2016).

288. 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).

289. Layer, R. M. *et al.* GIGGLE: a search engine for large-scale integrated genome analysis. *Nat Methods* **15,** 123 (2018).

290. Zambelli, F., Pesole, G. & Pavesi, G. PscanChIP: finding over-represented transcription factor-binding site motifs and their correlations in sequences from ChIP-Seq experiments. *Nucleic Acids Res* **41,** W535–W543 (2013).

291. Darbre, P. D. & Daly, R. J. Transition of human breast cancer cells from an oestrogen responsive to unresponsive state. *J Steroid Biochem Mol Biology* **37,** 753–763 (1990).

292. Hong, S. *et al.* Single-cell transcriptomics reveals multi-step adaptations to endocrine therapy. *Nat Commun* **10,** 1–14 (2019).

293. Cheng, C. *et al.* Glucose-Mediated N-glycosylation of SCAP Is Essential for SREBP-1 Activation and Tumor Growth. *Cancer Cell* **28,** 569–581 (2015).

294. Inoue, J. *et al.* Glutamine stimulates the gene expression and processing of sterol regulatory element binding proteins, thereby increasing the expression of their target genes. *FEBS Journal* **278,** 2739–2750 (2011).

295. Park, P. J. ChIP–seq: advantages and challenges of a maturing technology. *Nat Rev Genet* **10,** 669–680 (2009).

296. Ren, B. *et al.* Genome-Wide Location and Function of DNA Binding Proteins. *Science* **290,** 2306–2309 (2000).

297. Horak, C. E. & Snyder, M. ChIP-chip: A genomic approach for identifying transcription factor binding sites. *Methods Enzymol* **350,** 469–483 (2002).

298. Landt, S. G. *et al.* ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22,** 1813–1831 (2012).

299. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-Wide Mapping of in Vivo Protein-DNA Interactions. *Science* **316,** 1497–1502 (2007).

300. Robertson, G. *et al.* Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* **4,** nmeth1068 (2007).

301. Schmidt, D. *et al.* ChIP-seq: Using high-throughput sequencing to discover protein–DNA interactions. *Methods* **48,** 240–248 (2009).

302. Patten, D. K. *et al.* Enhancer mapping uncovers phenotypic heterogeneity and evolution in patients with luminal breast cancer. *Nat Med* 1 (2018). doi:10.1038/s41591-018-0091-x

303. Priolo, C. *et al.* AKT1 and MYC induce distinctive metabolic fingerprints in human prostate cancer. Cancer Res (2014). doi: 10.1158/0008-5472.CAN-14-1490.

304. Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10,** 515–527 (2006).

305. Tyner, C. *et al.* The UCSC Genome Browser database: 2017 update. *Nucleic Acids Res* **45,** D626–D634 (2017).

306. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28,** 495 (2010).

307. Wu, Q. *et al.* 27-Hydroxycholesterol Promotes Cell-Autonomous, ER-Positive Breast Cancer Growth. *Cell Reports* **5,** 637–45 (2013).

308. Umetani, M. *et al.* 27-Hydroxycholesterol is an endogenous SERM that inhibits the cardiovascular effects

of estrogen. *Nature Medicine* **13,** 1185–1192 (2007).

309. Phimister, E. G., Warner, M. & Gustafsson, J.-A. On Estrogen, Cholesterol Metabolism, and Breast Cancer. *New Engl J Medicine* **370,** 572–573 (2014).

310. Sodi, V. *et al.* Nutrient sensor O-GlcNAc transferase controls cancer lipid metabolism via SREBP-1 regulation. *Oncogene* **37,** 924 (2017).

311. Los, G. V. *et al.* HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *Acs Chem Biol* **3,** 373–382 (2008).

312. Krieg, P., Marks, F. & Fürstenberger, G. A Gene Cluster Encoding Human Epidermis-type Lipoxygenases at Chromosome 17p13.1: Cloning, Physical Mapping, and Expression. *Genomics* **73,** 323–330 (2001).

313. Weigert, A., Strack, E., Snodgrass, R. G. & Brüne, B. mPGES-1 and ALOX5/-15 in tumor-associated macrophages. *Cancer Metast Rev* **37,** 317–334 (2018).

314. Snodgrass, R. G. *et al.* A Novel Function for 15-Lipoxygenases in Cholesterol Homeostasis and CCL17 Production in Human Macrophages. *Front Immunol* **9,** 1906 (2018).

315. Klil-Drori, A. J. & Ariel, A. 15-Lipoxygenases in cancer: A double-edged sword? *Prostag Oth Lipid M* **106,** 16–22 (2013).

316. Liu, Y. *et al.* Deletions linked to TP53 loss drive cancer through p53-independent mechanisms. *Nature* **531,** 471 (2016).

317. Yang, L. *et al.* 15-lipoxygenase-2/15(S)-hydroxyeicosatetraenoic acid regulates cell proliferation and metastasis via the STAT3 pathway in lung adenocarcinoma. *Prostag Oth Lipid M* (2018). doi:10.1016/j.prostaglandins.2018.07.003

318. Roffeis, J., Hornung, D., Kuhn, H. & Walther, M. 15-Lipoxygenase-2 is differentially expressed in normal and neoplastic ovary. *Eur J Cancer Prev* **16,** 568–575 (2007).

319. van SCHAFTINGEN, E. & GERIN, I. The glucose-6-phosphatase system. *Biochem J* **362,** 513–532 (2002).

320. O'Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **44,** D733–D745 (2016).

321. Tameemi, W., Dale, T. P., Al-Jumaily, R. M. & Forsyth, N. R. Hypoxia-Modified Cancer Cell Metabolism. *Frontiers Cell Dev Biology* **7,** 4 (2019).

322. Du, W. *et al.* HIF drives lipid deposition and cancer in ccRCC via repression of fatty acid metabolism. *Nat Commun* **8,** 1769 (2017).

323. Bensaad, K. *et al.* Fatty Acid Uptake and Lipid Storage Induced by HIF-1α Contribute to Cell Growth and Survival after Hypoxia-Reoxygenation. *Cell Reports* **9,** 349–365 (2014).

324. Lewis, C. *et al.* SREBP maintains lipid biosynthesis and viability of cancer cells under lipid- and oxygen-deprived conditions and defines a gene signature associated with poor survival in glioblastoma multiforme. *Oncogene* **34,** 5128–5140 (2015).

325. Vandamme, J., Völkel, P., Rosnoblet, C., Faou, P. & Angrand, P.-O. Interaction Proteomics Analysis of Polycomb Proteins Defines Distinct PRC1 Complexes in Mammalian Cells. *Mol Cell Proteomics* **10,** M110.002642 (2011).

326. Lang, S. *et al.* Different effects of Sec61α, Sec62 and Sec63 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. *J Cell Sci* **125,** 1958–1969 (2012).

327. Ge, L., Gordon, J. S., Hsuan, C., Stenn, K. & Prouty, S. M. Identification of the Δ-6 Desaturase of Human

Sebaceous Glands: Expression and Enzyme Activity. *J Invest Dermatol* **120,** 707–714 (2003).

328. Abuharbeid, S., Czubayko, F. & Aigner, A. The fibroblast growth factor-binding protein FGF-BP. *Int J Biochem Cell Biology* **38,** 1463–1468 (2006).

329. Perone, Y. *et al.* SREBP1 drives Keratin-80-dependent cytoskeletal changes and invasive behavior in endocrine-resistant ERα breast cancer. *Nat Commun* **10,** 2115 (2019).

330. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485,** 376 (2012).

331. Magnani, L. *et al.* Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer. *Proc National Acad Sci* **110,** E1490–E1499 (2013).

332. Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518,** 331 (2015).

333. Gusterson, B. A., Ross, D. T., Heath, V. J. & Stein, T. Basal cytokeratins and their relationship to the cellular origin and functional classification of breast cancer. *Breast Cancer Res* **7,** 143 (2005).

334. Győrffy, B., Surowiak, P., Budczies, J. & Lánczky, A. Online Survival Analysis Software to Assess the Prognostic Value of Biomarkers Using Transcriptomic Data in Non-Small-Cell Lung Cancer. *Plos One* **8,** e82241 (2013).

335. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486,** 346 (2012).

336. Network, T. Comprehensive molecular portraits of human breast tumours. *Nature* **490,** 61 (2012).

337. Parsons, T. J., Horwitz, A. & Schwartz, M. A. Cell adhesion: integrating cytoskeletal dynamics and cellular tension. *Nat Rev Mol Cell Bio* **11,** 633 (2010).

338. Evans, A. *et al.* Invasive Breast Cancer: Relationship between Shear-wave Elastographic Findings and Histologic Prognostic Factors. *Radiology* **263,** 673–677 (2012).

339. Mason, D. E. *et al.* YAP and TAZ limit cytoskeletal and focal adhesion maturation to enable persistent cell motility. *J Cell Biol* **218,** jcb.201806065 (2019).

340. Lee, J. Y. *et al.* YAP-independent mechanotransduction drives breast cancer progression. *Nat Commun* **10,** 1848 (2019).

341. Bausch, A. R., Möller, W. & Sackmann, E. Measurement of Local Viscoelasticity and Forces in Living Cells by Magnetic Tweezers. *Biophys J* **76,** 573–579 (1999).

342. Tozluoğlu, M. *et al.* Matrix geometry determines optimal cancer cell migration strategy and modulates response to interventions. *Nat Cell Biol* **15,** ncb2775 (2013).

343. Charras, G. & Sahai, E. Physical influences of the extracellular environment on cell migration. *Nat Rev Mol Cell Bio* **15,** nrm3897 (2014).

344. Youk, J., Gweon, H. & Son, E. Shear-wave elastography in breast ultrasonography: the state of the art. *Ultrasonography* **36,** 300–309 (2017).

345. Yamaguchi, H. & Condeelis, J. Regulation of the actin cytoskeleton in cancer cell migration and invasion. *Biochimica Et Biophysica Acta Bba - Mol Cell Res* **1773,** 642–652 (2007).

346. Schnoor, M., Stradal, T. E. & Rottner, K. Cortactin: Cell Functions of A Multifaceted Actin-Binding Protein. *Trends Cell Biol* **28,** 79–98 (2018).

347. Dolat, L. *et al.* Septins promote stress fiber–mediated maturation of focal adhesions and renal epithelial

motility. *J Cell Biology* **207,** 225–235 (2014).

348. Stanbery, L. *et al.* High SEPT9_v1 Expression Is Associated with Poor Clinical Outcomes in Head and Neck Squamous Cell Carcinoma. *Transl Oncol* **3,** 239–245 (2010).

349. Acerbi, I. *et al.* Human breast cancer invasion and aggression correlates with ECM stiffening and immune cell infiltration. *Integr Biol* **7,** 1120–1134 (2015).

350. Levental, K. R. *et al.* Matrix Crosslinking Forces Tumor Progression by Enhancing Integrin Signaling. *Cell* **139,** 891–906 (2009).

351. Weigelt, B., Peterse, J. L. & Veer, L. J. Breast cancer metastasis: markers and models. *Nat Rev Cancer* **5,** nrc1670 (2005).

352. Grigore, A., Jolly, M., Jia, D., Farach-Carson, M. C. & Levine, H. Tumor Budding: The Name is EMT. Partial EMT. *J Clin Medicine* **5,** 51 (2016).

353. Bertolio, R. *et al.* Sterol regulatory element binding protein 1 couples mechanical cues and lipid metabolism. *Nature Communications* **10,** 1326 (2019).

354. Padmanaban V, Krol I, Suhail Y, Szczerba B. M., Aceto N, Bader J. S., Ewald, A. J E-cadherin is required for metastasis in multiple models of breast cancer *Nature* (4 September 2019).

355. Dobrosotskaya, I., Seegmiller, A., Brown, M., Goldstein, J. & Rawson, R. Regulation of SREBP Processing and Membrane Lipid Production by Phospholipids in Drosophila. *Science* **296,** 879–883 (2002).

356. Romani, P. *et al.* Extracellular matrix mechanical cues regulate lipid metabolism through Lipin-1 and SREBP. *Nature Cell Biology* **21,** 338–347 (2019).

357. Gioeli, D. *et al.* Development of a multicellular pancreatic tumor microenvironment system using patient-derived tumor cells. *Lab on a chip* **19,** 1193–1204 (2019).

358. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol* **26,** nbt1385 (2008).

359. Savci-Heijink, D. C., Halfwerk, H., Koster, J. & van de Vijver, M. J. A novel gene expression signature for bone metastasis in breast carcinomas. *Breast Cancer Res Tr* **156,** 249–259 (2016).

360. Calvo, F. *et al.* Cdc42EP3/BORG2 and Septin Network Enables Mechano-transduction and the Emergence of Cancer-Associated Fibroblasts. *Cell Reports* **13,** 2699–2714 (2015).

361. Gansler, T. S. *et al.* Increased expression of fatty acid synthase (OA-519) in ovarian neoplasms predicts shorter survival. Human Pathology (1997).

362. Ogino S. *et al.* Fatty acid synthase overexpression in colorectal cancer is associated with microsatellite instability, independent of CpG island methylator phenotype. Human Pathology (2007).

363. Ogino S. *et al.* Cohort Study of Fatty Acid Synthase Expression and Patient Survival in Colon Cancer. Journal of Clinical Oncology (2008)
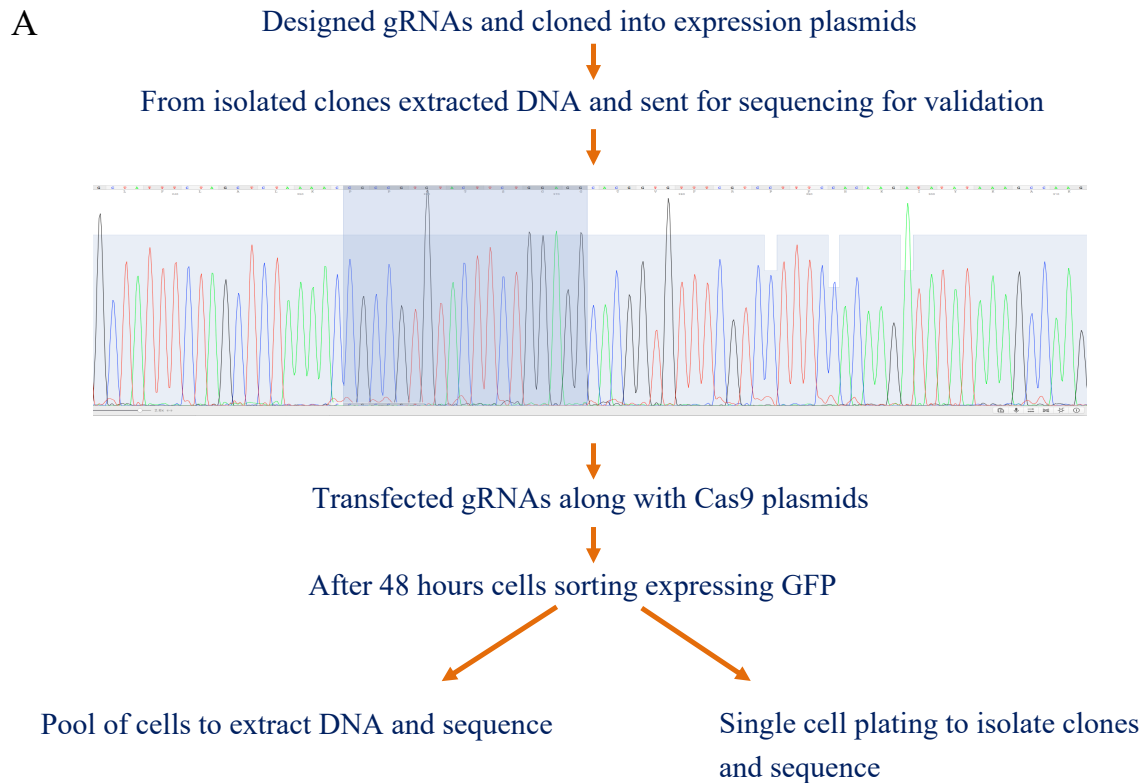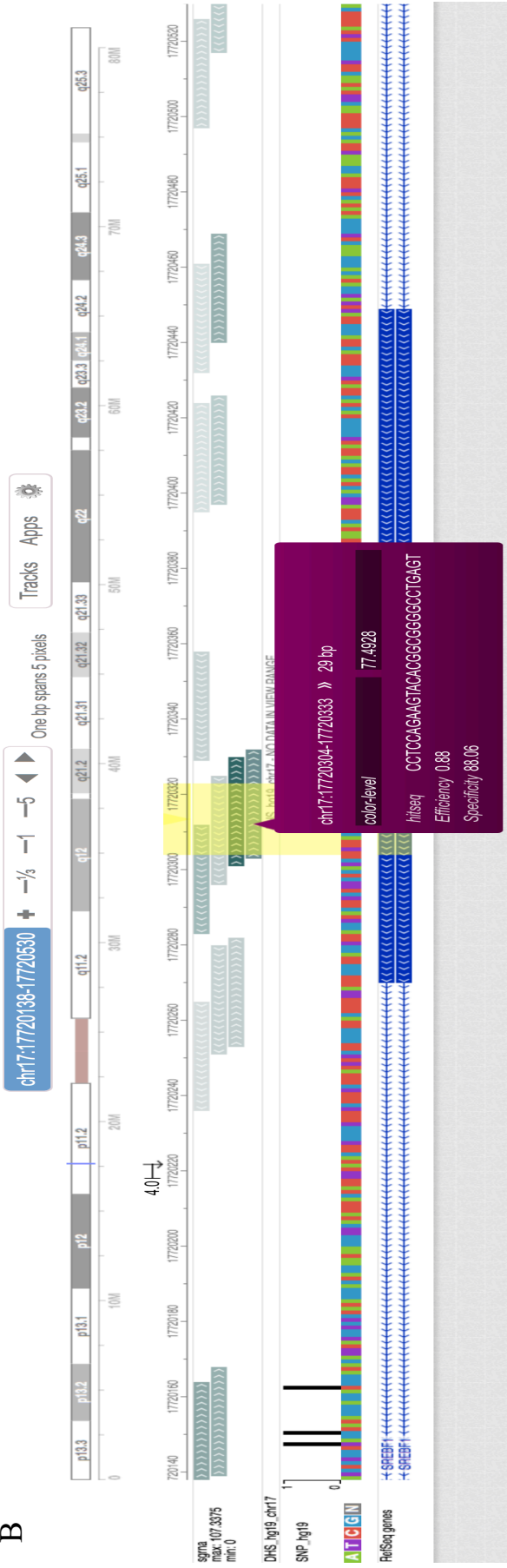
Appendix I: Supplementary Figures – Chapter 3

A   Designed gRNAs and cloned into expression plasmids

From isolated clones extracted DNA and sent for sequencing for validation

Transfected gRNAs along with Cas9 plasmids

After 48 hours cells sorting expressing GFP

Pool of cells to extract DNA and sequence

Single cell plating to isolate clones and sequence

Figure S1: A) Schematic of the CRISPR strategy used for generation of SREBP1 MCF7 using CRISPR/Cas9 technology. Following pages B-D): Validation of exact mutations in CRISPR clones and Karyotyping. B) The web application for the Design and Optimization (CRISPR-DO) of guide sequences was used for CRISPR design and optimization. Designed sgRNA efficiency (0.88) and specificity (88.06) are indicated. D) SREBP1 Immunoblotting analysis in MCF7 untransfected (UT) cells and CRISPR clones 1A, 1B, 1C. D) Sequencing chromatograms of genomic DNA prepared from the MCF7 SREBP1 CRISPR clones, annotated with mutations and CRISPR targeting site positions. E) Graphic of sequencing results (black bars) for Clone B, compared against the SREBP1 binding site (green bar). In different sequencing results distinct deletions have been identified. Due to the indication of multiple copies of the SREBP1 loci karyotyping analysis was performed identifying three copies of chromosome 17 in some cells C).
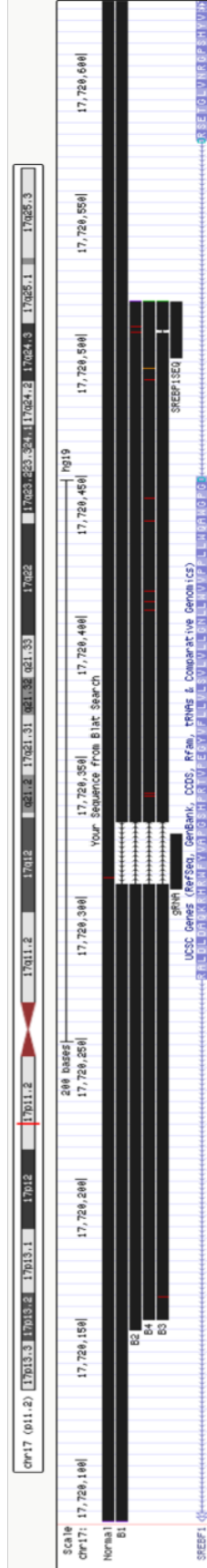
B

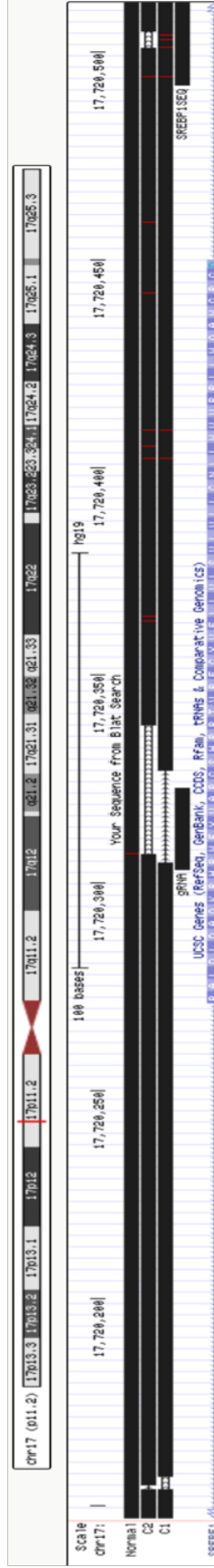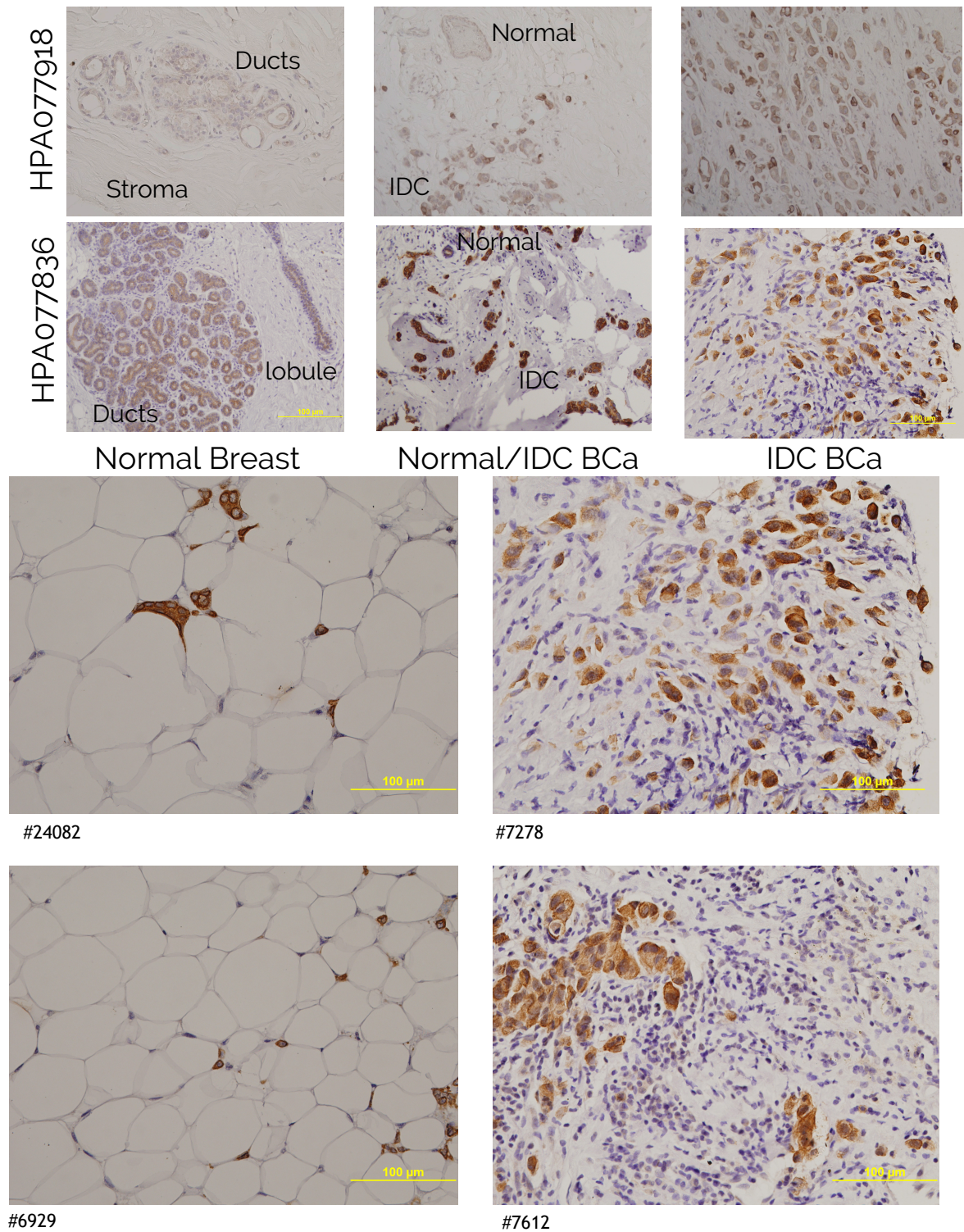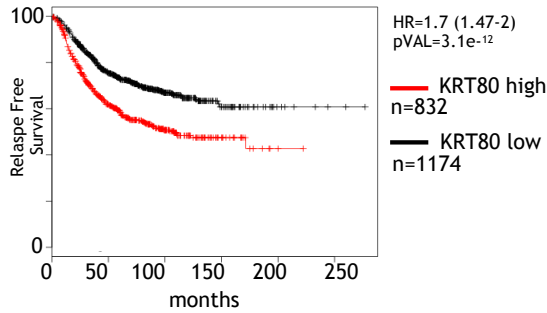chr17:17720138-17720530

Tracks   Apps

+ −⅓ −1 −5 ▲ ▼   One bp spans 5 pixels

4.0

sgrna
max: 107.3375
min: 0

DHS_hg19_chr17

SNP_hg19

A T C G N

RefSeq genes

+ SREBF1
+ SREBF1

chr17:17720304-17720333 ≫ 29 bp

color-level   CCTCCAGAAGTACACGGCGGGGCCTGAGT

hitseq   77.4928

Efficiency 0.88

Specificity 88.06

S_hg19_chr17 NO DATA IN VIEW RANGE

p13.3 p13.2 p13.1 p12 p11.2 q11.2 q12 q21.2 q21.31 q21.32 q21.33 q22 q23.1 q23.2 q23.3 q24.1 q24.2 q24.3 q25.1 q25.3

720140 17720160 17720180 17720200 17720220 17720240 17720260 17720280 17720300 17720320 17720340 17720360 17720380 17720400 17720420 17720440 17720460 17720480 17720500 17720520

0 10M 20M 30M 40M 50M 60M 70M 80M

C

Karyotype Analysis

17   17   17

D

CRISPR 1A   CRISPR 1B   CRISPR 1C   MCF7 UT

SREBP1

pSREBP1 125 KDa

mSREBP1
60-70 KDa

GAPDH
40 KDa

213

E

Clone 1A

Clone 1B

Clone 1C

Appendix II: Supplementary Figures - Chapter 4

A

Normal Breast   Normal/IDC BCa   IDC BCa

#24082

#7278

#6929

#7612

B

METABRIC ER+

HR=1.7
pVAL=8e⁻⁸

Overall Survival — high / low — n=1486
Time (Years)

Untreated ER+ — p=0.0029, p=0.038, n=300
ET treated ER+ — p=1.9*10⁻⁷, p=3.3*10⁻⁶, n=1013
Untreated ERa negative — p=0.01, p=0.014, n=433
H.R

5 years / 25 years

METABRIC ER+
Endocrine treated only

Follow Up — 5 yrs / 25 Yrs.
p=1.9*10⁻⁷, p=3.3*10⁻⁶, n=1013
Odds Ratio
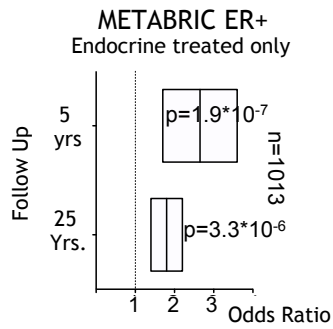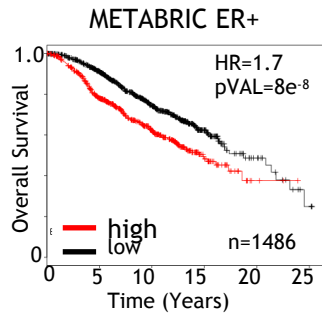
C

METABRIC ER+ 5 yrs.
Breast Cancer expressed KRTs

KRT86
KRT81
KRT80 — pVAL=8e⁻⁸
KRT19
KRT17
KRT7
KRT4
Odds Ratio

D

AFFY all cases

HR=1.7 (1.47-2)
pVAL=3.1e⁻¹²
KRT80 high n=832
KRT80 low n=1174
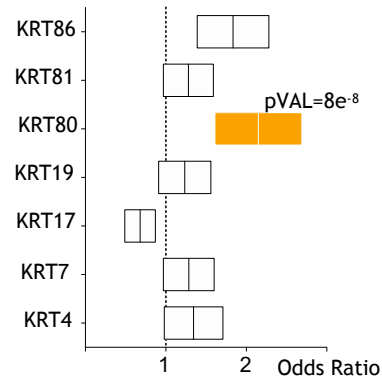
Relaspe Free Survival
months

E

AFFY all cases

Multivariate	P val	H.R	CI
Ki67	<0.05	1.19	(1-1.42)
ERα	<0.0001	0.65	(0.54-0.77)
HER2	<0.03	1.21	(1.01-1.45)
Gene expression	<0.0001	1.58	(1.35-1.85)

F

AFFY all cases

Post-Progression Survival
median survival=40.8
median survival=19.97
months
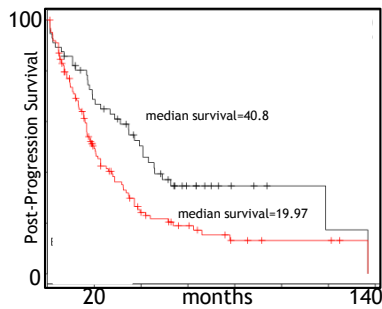
KRT80 high n=124
KRT80 low n=56
HR=1.7 (1.18-2.59)
pVAL=0.005

AFFY all cases

Distal Metastasis Free Survival
months
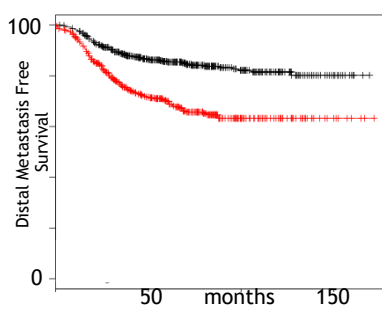
KRT80 high n=524
KRT80 low n=425
HR=2.31 (1.75-3.05)
pVAL=1.1e⁻⁹

AFFY all cases

OS
median survival=136.8
median survival=54.6
months
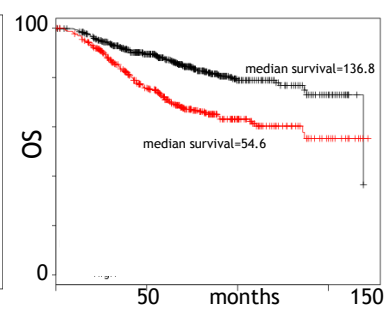
KRT80 high n=124
KRT80 low n=56
HR=2.13 (1.62-2.79)
pVAL=2.3e⁻⁸

216
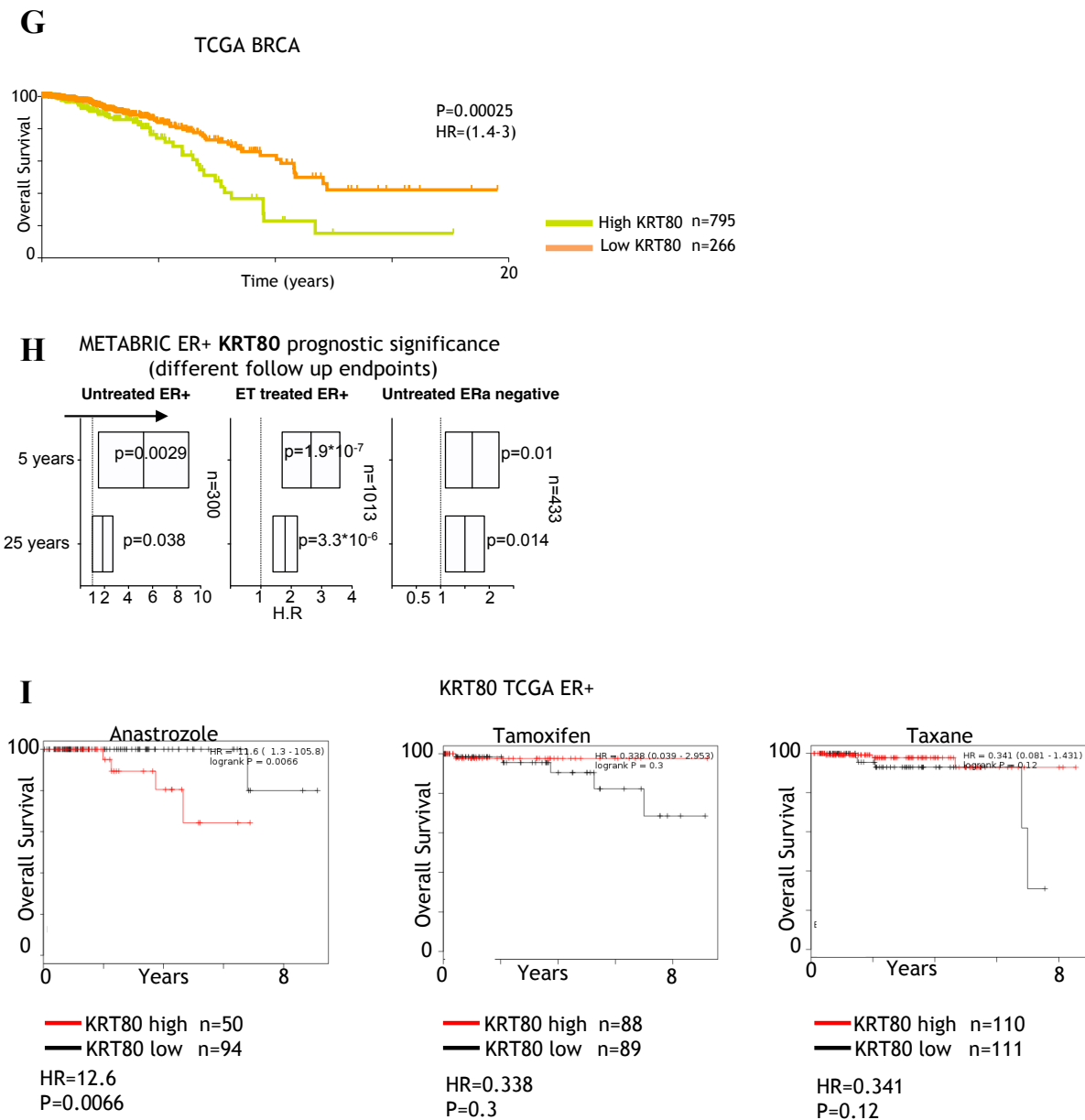
Figure S2: A) KRT80 expression in diagnostic material has prognostic significance. Analysis were performed on the METABRIC RNA-seq splitting patient in high and low KRT80 expression. Two distinct follow-ups for an additional sub-cohort is also shown (endocrine-treated patients). Floating bars show minimum-maximum and average hazard ratios. B) Representative images of IHC using two independent antibodies for KRT80 in breast tissues and breast cancer samples (by C.D.). Antibody are labelled accordingly and were obtained from the Protein Atlas Initiative. Bottom four panels show actual samples from the prospective trial in which Shearwave Elastography was conducted. C) Other transcribed type-II Keratins in breast cancer samples from METABRIC are not associated with prognostic significance. Floating bars show minimum-maximum and average hazard ratios. D) KRT80 expression in diagnostic material has prognostic significance. Analysis were performed on a meta-collection of batch-normalized microarray from GEO browser. Patients were split into high and low KRT80 expression. E) Multivariate correction analysis for panel E. F) KRT80 expression in diagnostic material has prognostic significance in the context of post-progression survival (PPS), distal-metastasis free survival (DMFS) and overall survival (OS). Number of patients in each arm is indicated. Hazard ratios with confidence intervals

and univariate p-values are also shown. G) Kaplan-Meier analysis using KRT80 independent cohort of BC patients (TCGA) profiled with RNA-seq. H) KRT80 investigated in three sub cohorts derived from the METABRIC cohorts and were tes power in the short-term (<5 years) or long-term (<25years). Hazard ratios are plott overall survival). I) KRT80 prognostic association with overall survi post-surgical adjuvant treatment. C-J) (Modified from Perone 2019, made by Luca Magnani).
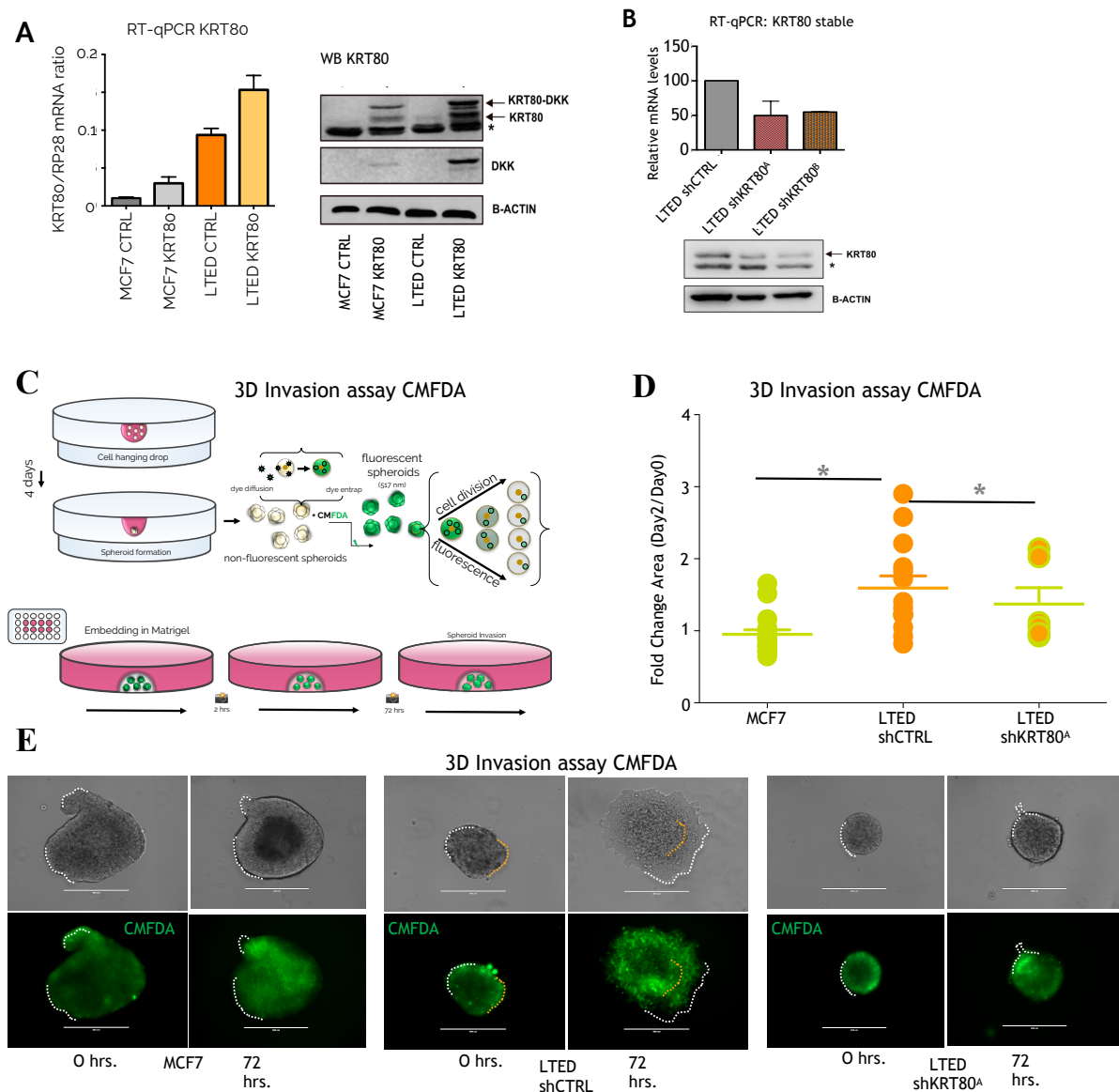
Figure S3: A) RT-qPCR measurements and Western Blot analysis of KRT80 mRNA levels in breast cancer cells transfected with KRT80-DKK constructs. Both KRT80 and DKK tagged are shown. B) RT-qPCR and Western Blot analysis of KRT80 mRNA and protein in cells transfected with stable shKRT80 constructs. C) Replication dependent labeling of breast cancer spheroids. Cells were labeled with CMFDA that is converted to its membrane-impermeant fluorescent form by cytosolic esterase to entrap the dye. Active replication can dilute the dye until

218

disappearance within 2–3 cell cycles. D) Quantification of the area fold change in organoids treated with CMFDA. Lines represent mean and SD. Asterisks represent significance level $p < 0.05$ after Student t test. E) Representative images of CMFDA tagged spheroids. Invasive borders are highlighted by dotted white lines. Representative original borders are highlighted by yellow dotted lines. Bars scale = 400 μm. (Experiments performed with the help of A.R.M. and P.M., modified from Perone 2019, figures originally made by Luca Magnani).

A — KRT80: IF 2D

MCF7 KRT80 DAPI   MCF7T   MCF7TF   LTED   LTEDT   LTEDF

B — KRT80: IF 3D

MCF7F   LTED   LTEDT   LTEDF

KRT80 DAPI

KRT80: IF 3D

KRT80   MCF7F   LTED

KRT80 DAPI   LTEDT   LTEDF

C — IF: PLEURAL EFFUSION

Pt. A   Pt. B   Pt. C   KRT80   DAPI
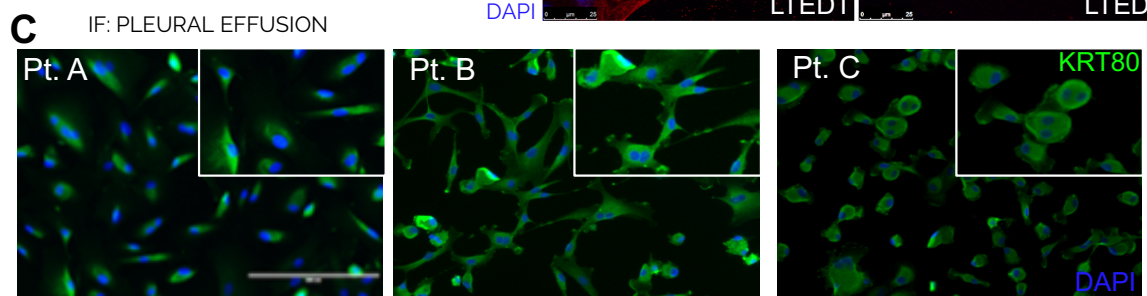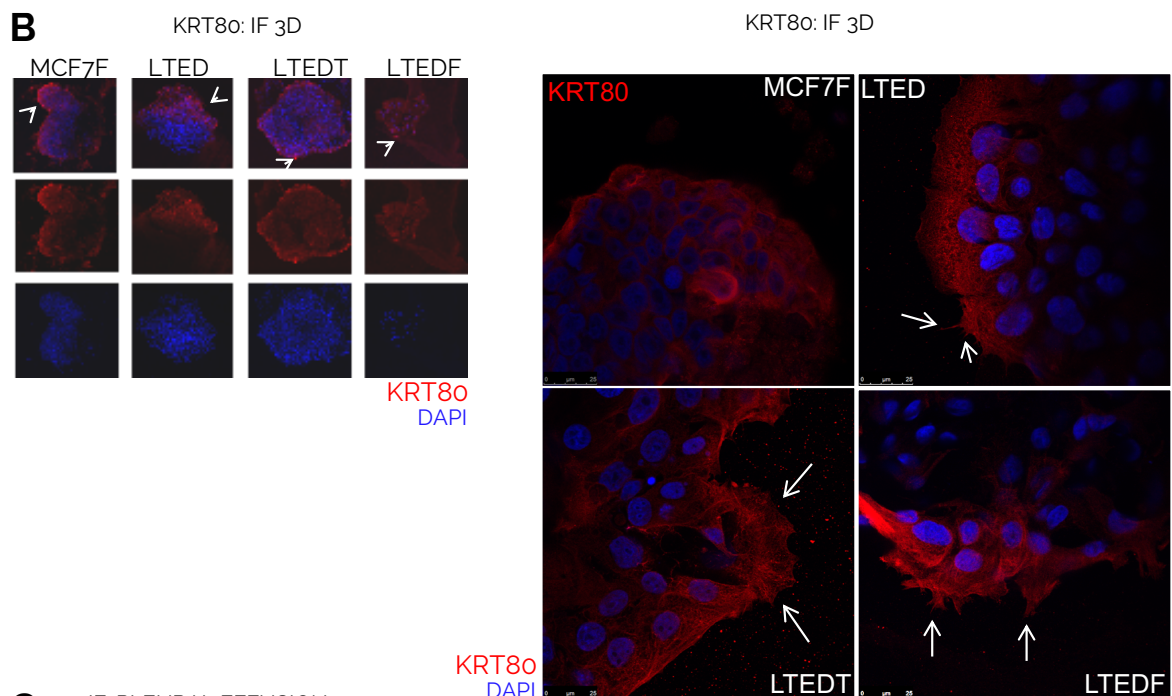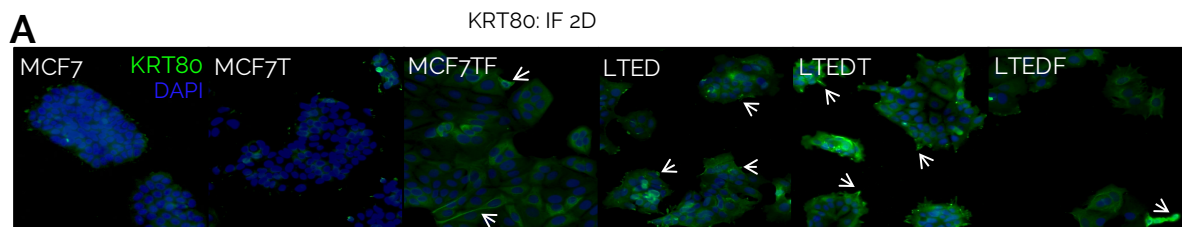
patient	type	histology	treatment	secondary
PT.A	Pleural fluid	ER+, HER-2 negative	2007-Gem/Taxol/EC chemotherapy as part of Neo-tAnGo trial, August 2008-Tamoxifen switched to anastrozole completed 2013	2014 - Liver metastasis (Liver mets receptors are ER 8/8 PgR 4/8 HER2-)
PT.B	Pleural Fluid	ER+, PR+ HER2 negative	2012-Letrozole + Zometa; Sep 2013 Irosustat; March 2014 Everolimus and Exemestane	Bone and lung mets
PT.C	Pleural Fluid	ER 4/8, PgR 0/8, HER2	2013-FECT, radiotherapy, Letrozole	

Figure S4 (previous page): KRT80 is preferentially distributed at the margin of 3D cultures. A) IF analysis of KRT80 in MCF7 cells and drug-resistant derivatives. B) IF imaging of KRT80 in breast cancer cell lines grown in 3D spheroids. Blown up of spheroids margins are shown in the right panels C) IF analysis of KRT80 in cancer cells isolated from fresh pleural effusion from breast cancer patients (experiments, analysis and figures originally made by A.F. and F.C. at ICR London, modified from Perone 2019).
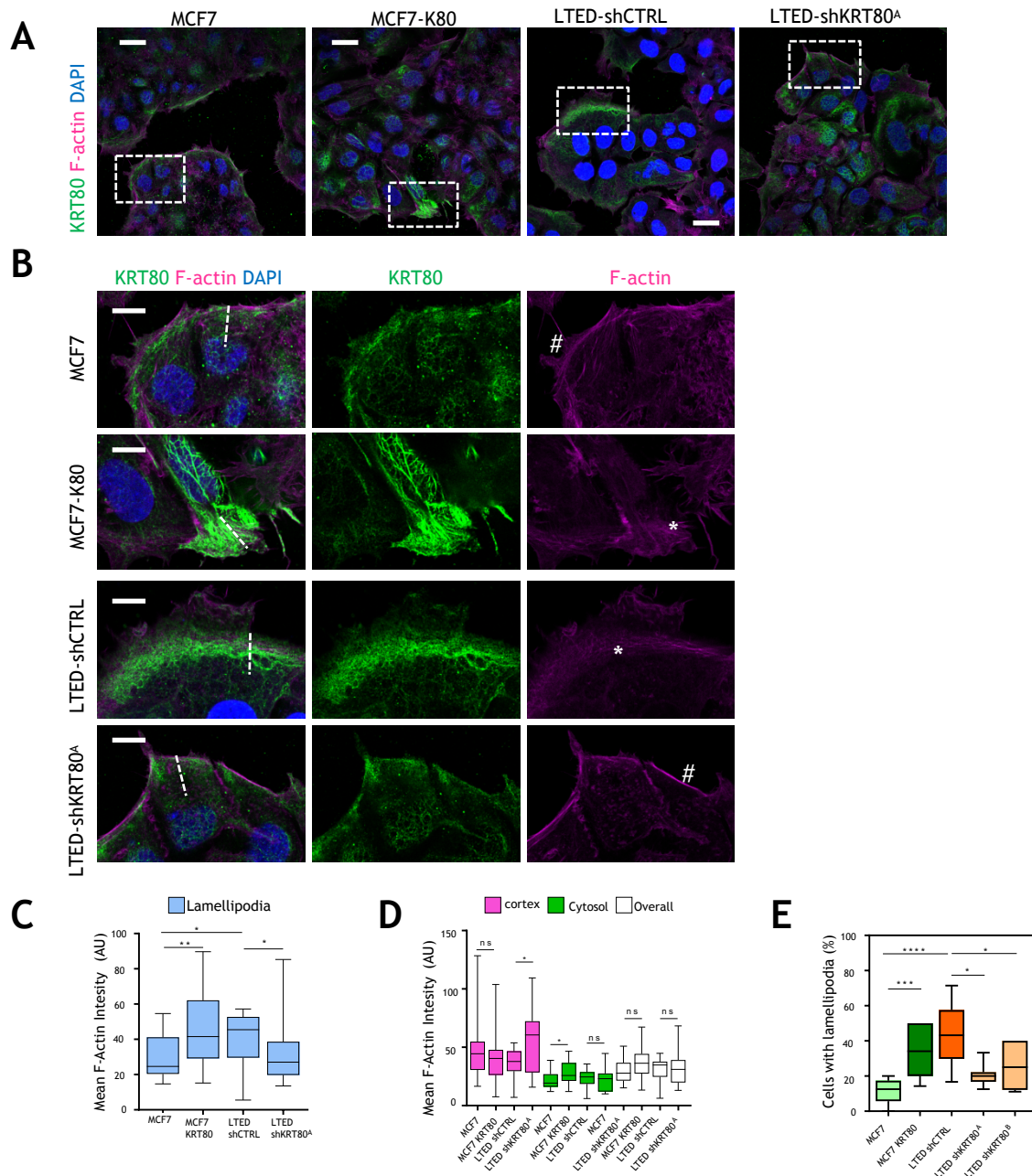
Figure S5: KRT80 induces invasion-associated cytoskeletal changes. A) Representative confocal microscopy images showing F-actin (magenta), KRT80 (green) and DAPI (blue) staining of MCF7-control, MCF7-KRT80, LTED-control and LTED-sha cells. Scale bars represent 25 μm. B) Zoom-up magnifications of areas indicated in A), showing F-actin (magenta), KRT80 (green) and DAPI (blue) staining in cells located at the border of clusters. Single channel images for F- actin and KRT80 are also shown. Scale bars, 10 μm. Asterisks indicate lamellipodia-

like structures in MCF7-KRT80 and LTED cells, and hashtags indicate cortical actin areas in MCF7 and LTED-sha cells. Graphs on the right show line scan analysis for F-actin and KRT80 fluorescence across the leading edges of cells, as indicated in the broken line in the merged images. C), D) Graphs show quantification of F-actin fluorescence intensity at lamellipodial regions (C) and at cell cortex, cytosol and overall (i.e., whole cell) (D) in MCF7-control, MCF7-KRT80, LTED-control and LTED-sha cells (n = 19, MCF7; n = 20, MCF7- KRT80; n = 14, LTED; n = 16, LTED-sha individual cells). E) Graph shows quantification of percentage of cells with clear lamellipodia and membrane ruffles in MCF7-control, MCF7-KRT80, LTED-control, LTED-sha and LTED-shb cells (n = 8, MCF7; n = 12, MCF7-KRT80; n = 12, LTED; n = 7, LTED-sha; n = 6, LTED-shb fields of view), (experiments, analysis and figure originally made by A.F. and F.C. at ICR London, modified from Perone 2019).
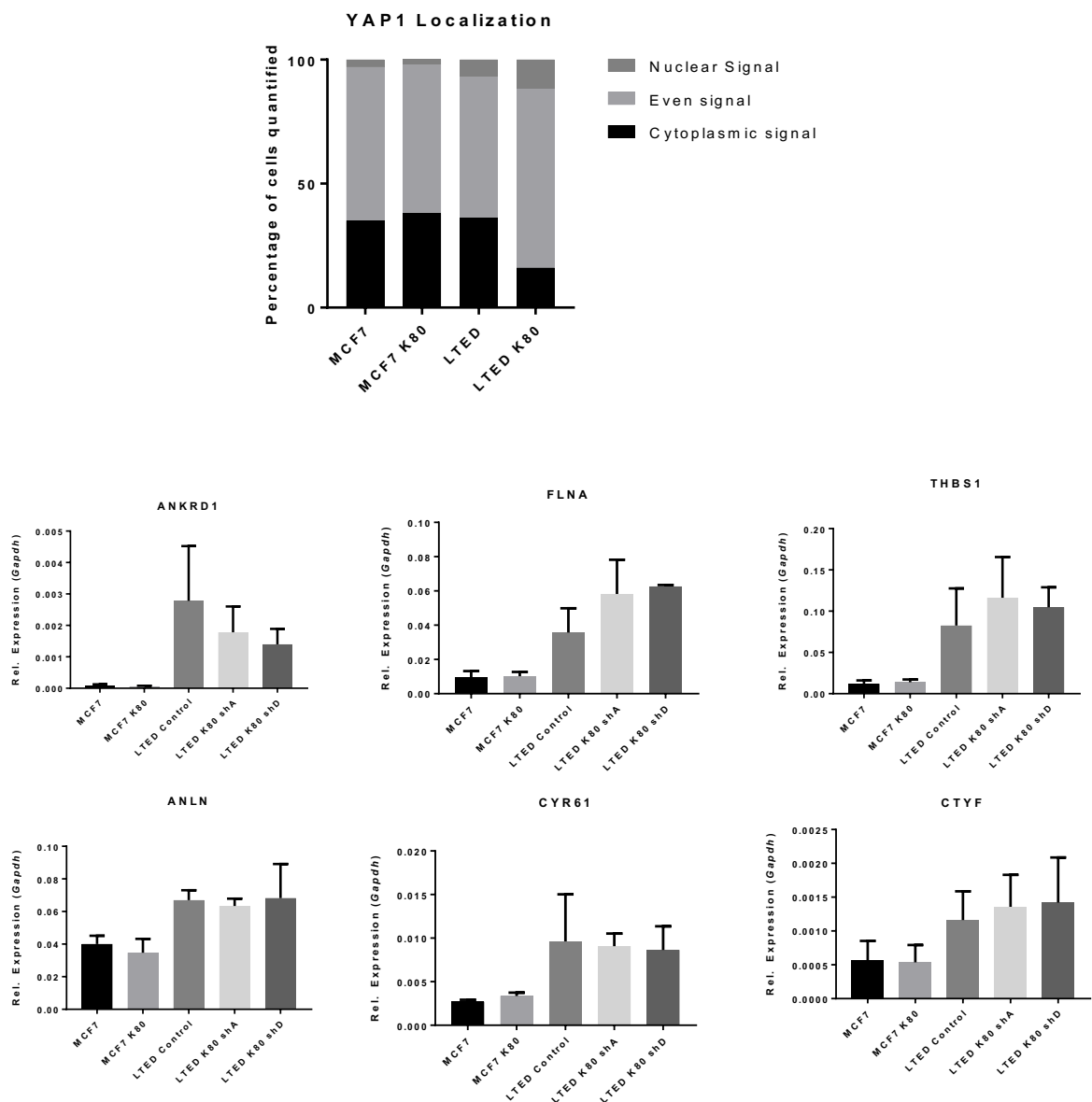

Figure S6: Quantification of immunofluorescence images taken by confocal microscopy of nuclear versus cytoplasmic YAP. Cells (of similar confluency) were stained and scored manually depending on their YAP intensity (i.e. higher staining in the nucleus, higher in the cytoplasm or equal intensity in the cytoplasm). RT-
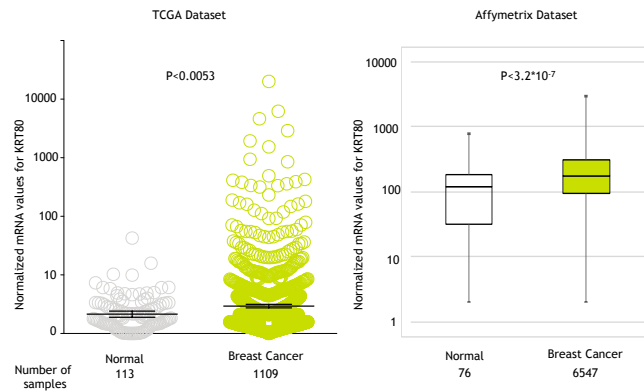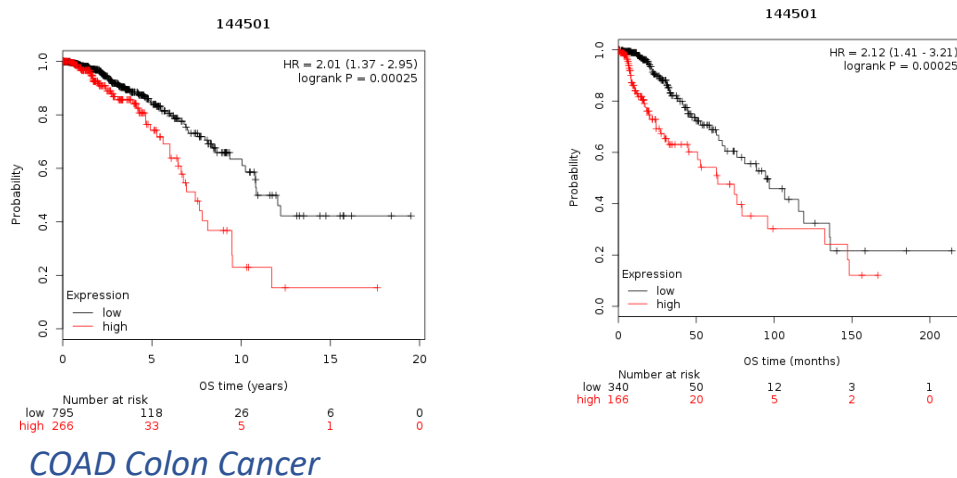
Figure S7: KRT80 mRNA expression in normal breast. Meta-analysis of TCGA data and Affymetrix GEO data comparing expression levels for KRT80 in normal samples and cancer samples (Analysis and figure made by Luca Magnani, modified from Perone 2019).
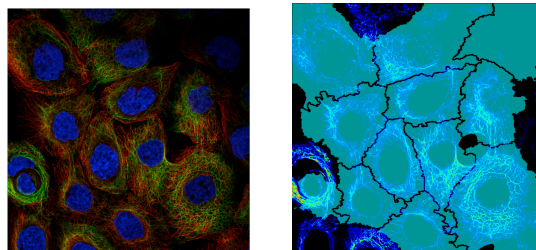
A

COAD Colon Cancer

B

Figure S8: A) Kaplan-Meier analysis using KRT80 expression levels in colon cancer. Overall survival (OS) in months (right panel) and years (left _____ in _____ A-431epithelial cells, green: KRT80, red: Microtubules, blue: DAPI, mod. Mod_____ ein _____ riginally made by L.M.

Appendix III: Grants, Prizes and Awards

Prizes and Awards	Awarding Body	Conference	Year
Outstanding Poster Award	Gordon Research Conference	Gordon Research Conference on Hormone-Dependent Cancers (USA)	2019
Meeting bursary	Gordon Research Seminar	Gordon Research Seminar on Hormone-Dependent Cancers (USA)	2019
FEBS Letters 2018 Cover Contest Award	The Federation of European Biochemical Societies (FEBS)	N/A	2019
Biochemical Society General Travel Grant	The Biochemical Society	ENDO 2019 Conference New Orleans (USA)	2019
Outstanding Abstract Award	Endocrine Society	ENDO 2019 Conference New Orleans (USA)	2019
Early Career Forum Travel Award	Endocrine Society	ENDO 2019 Conference New Orleans (USA)	2019
EACR25 Meeting Bursary	European Association of Cancer Research (EACR)	EACR25 Amsterdam	2018
British Association of Cancer Research/Cancer Research UK BACR/CRUK Student Travel Award	European Association of Cancer Research (EACR)	EACR25 Amsterdam	2018
Best Poster Award	Fusion Conferences	Nuclear Receptors Conference, Mexico	2018
Student Grant	Fusion Conferences	Nuclear Receptors Conference, Mexico	2018
Society for Endocrinology BES Registration Grant	Society for Endocrinology (SfE BES)	SfE BES 2017 Conference, Harrogate, UK	2017
Third Prize in Research as Art Competition	Imperial College of London	N/A	2017

223

Qaplà!