Université de Montréal

Mechanisms Controlling Luminal Identity of Breast Tumours

par
Houssam Ismail

Programmes de Biologie Moléculaire
Institut de recherche en immunologie et en cancérologie (IRIC)
Faculté de Médecine

Thèse présentée à la Faculté des Études Supérieures
en vue de l'obtention du grade de Doctorat
en Biologie Moléculaire
option Biologie des Systèmes

Décembre, 2017

Résumé

Le cancer du sein est une maladie hétérogène qui comprend trois types principaux de tumeurs identifiables par analyse de profils d'expression génique : les tumeurs luminales, HER2+ et basal-like. Ces sous-types pourraient résulter d'un blocage à différents stades de la différenciation épithéliale mammaire. Cependant, les mécanismes qui déterminent ces sous-types restent peu explorés. Les tumeurs luminales, qui représentent deux tiers des tumeurs mammaires, sur-expriment le récepteur des oestrogènes alpha (ERα), un facteur de transcription (FT) ligand-dépendant. De plus, elles expriment des niveaux élevés de deux autres FT luminaux, FOXA1 et GATA3, impliqués dans la régulation de l'expression et de l'activité de ERα. En clinique, l'action prolifératives du récepteur peut être bloquée par des anti-oestrogènes, mais malheureusement la maladie progresse dans ~50% des cas dû à une résistance au traitement, soulignant le besoin de thérapies additionnelles. Les inhibiteurs d'histones désacétylases (HDACi) sont une classe d'agents anti-cancer qui inhibent la proliferation des tumeurs hématologiques et solides. Dans les cellules tumorales mammaires, ils suppriment l'expression de ERα et sa signalisation et induisent la différentiation ou l'apoptose. Ces propriétés ont conduit à l'utilisation combinée d'anti-oestrogènes et HDACi pour le traitement des tumeurs du sein ER+.

Dans ce travail, nous avons effectué une analyse de corrélation dans de grands jeux de transcriptomes de tumeurs mammaires et caractérisé un groupe de gènes luminaux hautement corrélés contenant six FT du lignage luminal : *ESR1* (le gène codant pour ERα), *FOXA1*, *GATA3*, *SPDEF*, *XBP1* et *AR*. Nous démontrons via une déplétion par siRNA que FOXA1 régule l'expression de ces gènes de manière prépondérante, et ce malgré l'enrichissement en sites de liaisons de ChIP-Seq d'autres FT luminaux. Ce groupe peut être subdivisé en deux sous-groupes, l'un centré sur *ESR1* et *GATA3* et l'autre sur *FOXA1*, suggérant l'existence de tumeurs exprimant ces sous-groupes de manière différentielle. En effet, l'expression de *FOXA1*, *ESR1* et *GATA3* permet de ségréger les tumeurs luminales (expression élevée) et basal-like (expression faible). De plus, nous avons observé que les tumeurs de type molecular apocrine, décrites comme dépendantes de la signalisation par le récepteur des androgènes AR et fréquemment HER2+, présentent un profil *FOXA1*^{high} *ESR1*^{low} *GATA3*^{low}. L'expression ectopique de ERα dans les cellules tumorales SK-BR-3, de type molecular apocrine, (1) induit une ouverture de la chromatine à des éléments cibles liés faiblement par FOXA1 et (2) récapitule une partie de la réponse oestrogénique des cellules

luminales. Ces résultats complémentent des études précédentes décrivant FOXA1 comme un factor pionnier pour la liaison de ERα à ses éléments de réponse. L'absence d'expression de *ESR1* dans les cellules SK-BR-3, en dépit de hauts niveaux d'expression de son facteur régulateur amont FOXA1, corrèle avec une liaison faible de FOXA1 et des marques de chromatine actives H3K9/K14 sur les régions régulatrices amont de *ESR1*.

Dans la seconde partie de notre étude, nous montrons de plus que les HDACi reprogramment la transcription des cellules de cancer du sein, entraînant une suppression des FT du lignage luminal FOXA1 and GATA3, et également de plusieurs protéines de remodelage de la chromatine, et des traits de différenciation lactogénique. La suppression de l'expression des FT luminaux par des mécanismes transcriptionnels et post-transcriptionnels implique l'acétylation de substrats protéiques restant à identifier.

De manière globale, l'identification de réseaux transcriptionnels qui sous-tendent les sous-types de tumeurs mammaires et d'interrupteurs moléculaires contrôlant leur expression/fonction aidera à déterminer si leur modulation permet de reprogrammer les cellules de cancer du sein vers un phénotype moins agressif.

Mots-clés : cancer du sein, ERα, FOXA1, GATA3, FT luminaux, HDAC, inhibiteur d'HDAC, reprogrammation transcriptionnelle, acétylation.

Abstract

Breast cancer is a heterogeneous disease comprising three major molecular subtypes that can be identified by gene expression profiling: luminal, HER2-positive, and basal-like. These subtypes are thought to arise from a block at different stages of breast epithelial cell differentiation. However, the mechanisms underlying subtype specification remain largely unexplored. Luminal tumours, which account for two-thirds of breast tumours, overexpress estrogen receptor alpha (ERα), a ligand-inducible transcription factor (TF). These tumours additionally express high levels of FOXA1 and GATA3, two luminal-lineage TFs, that have been reported to regulate both expression and activity of the receptor. In the clinic, the proliferative actions of the receptor can be blocked with antiestrogens, but unfortunately ~50% of patients will relapse due to resistance, highlighting the need for additional therapies. Histone deacetylase inhibitors (HDACis) are a class of anticancer agents with anti-proliferative activity in both solid and haematological tumours. In breast cancer, these drugs were shown to suppress ERα expression/signalling, and induce differentiation and apoptosis. These properties have led to the combined use of HDACis and antiestrogens in clinical trials for treatment of ERα-positive breast tumours.

In the present work, we have utilized gene correlation analysis in large breast tumour transcriptome datasets and characterized a cluster of highly correlated luminal genes containing six luminal-lineage TFs: *ESR1* (gene encoding ERα), *FOXA1*, *GATA3*, *SPDEF*, *XBP1* and *AR*. We show using siRNA-mediated depletion that FOXA1 dominantly positively regulates expression of the cluster despite its enrichment in ChIP-Seq binding sites of other luminal TFs. Interestingly, this cluster partitions into two subclusters, one centered around *FOXA1*, and the other, around *ESR1* and *GATA3*, suggesting the presence of tumours that differentially express either subcluster. Indeed, expression of *ESR1*, *FOXA1* and *GATA3* can readily segregate luminal tumour samples (high expression) from basal-like ones (low expression). Furthermore, we found that molecular apocrine tumours, previously described as frequently HER2-positive, and androgen receptor signalling-dependent, are $FOXA1^{high}$ $ESR1^{low}$ $GATA3^{low}$. Ectopic expression of ERα in molecular apocrine SK-BR-3 cells could (1) recapitulate part of the estrogen response observed in luminal cells and (2) also result in chromatin opening and transcriptional regulation of target genes weakly pre-bound by FOXA1. This is complementary to previous studies describing FOXA1 as a factor that pioneers ERα binding at its target genes. Notably, lack of ERα expression in SK-BR-3 cells,

despite high expression of its upstream regulatory factor FOXA1, coincided with weak FOXA1 binding and reduced recruitment of active H3K9/K14 marks at upstream *ESR1* regulatory regions.

In the second part of our study, we further show that HDACis transcriptionally reprogram breast cancer cells resulting in downregulation of luminal-lineage TFs FOXA1 and GATA3, as well as many chromatin remodelling proteins, while inducing expression of lactogenic differentiation genes. Shutdown of luminal TFs occurred through both transcriptional and post-transcriptional mechanisms that involved modulated acetylation of protein substrates to be identified.

Altogether, identification of transcriptional networks that demarcate breast tumour subtypes and molecular switches controlling their function/expression will help determine whether their perturbation can rewire breast cancer cells towards a less aggressive phenotype.

Keywords: breast cancer, ERα, FOXA1, GATA3, luminal TFs, HDAC, HDAC inhibitors, transcriptional reprogramming, acetylation.

Table of Contents

List of Tables

List of Figures

Chapter One.

Chapter Two.

Chapter Three.

Chapter Four. Discussion and Perspectives

Abbreviations

AF - activation function
AI - aromatase inhibitor
alpha 1-AT - alpha 1-antitrypsin
AML - acute myeloid leukaemia
AP1 - activator protein 1
APL - acute promyelocytic leukaemia
AR - androgen receptor
ARE - androgen response element
ATAC-Seq - Assay for Transposase-Accessible Chromatin using sequencing
bZIP - basic leucine zipper
CBP - CREB-binding protein
CHIP - C-terminal HSC70-interacting protein
ChIP - chromatin immunoprecipitation
CNA - copy number alteration
CREB - cAMP response element-binding protein
Csnb - casein beta
CtBP1 - c-terminal binding protein 1
CTCL - cutaneous T-cell lymphoma
DBD - DNA-binding domain
Dbh - dopamine β-hydroxylase
DHT - dihydrotestosterone
DNMT1 - DNA methyltransferase 1
E2 - 17β-estradiol
EGF - epidermal growth factor
EGFR2/HER2 - epidermal growth factor receptor 2
EMT - epithelial-mesenchymal transition
ER - endoplasmic reticulum
ERE - estrogen response element
ERSE - endoplasmic reticulum stress response element
ERα - estrogen receptor alpha
FACS - fluorescence-activated cell sorting
FAIRE - formaldehyde assisted isolation of regulatory elements
FOG1 - friend of GATA-1
FOX - forkhead box
FOXA1 - forkhead box A1
FOXRE - FOXA1 response element
GATA3 - GATA binding protein 3
GCDFP15 - gross cystic disease fluid protein 15

Gcg - proglucagon
GH - growth hormone
GR - glucocorticoid receptor
HAT - histone acetyltransferase
HDAC - histone deacetylase
HDACi - Histone deacetylase inhibitor
HIV - human immunodeficiency virus
HNF3A - hepatocyte nuclear factor 3A
HRT - hormone replacement therapy
HSF1- heat shock factor 1
IGF - insulin-like growth factor
IGFBP5 - insulin-like growth factor binding protein 5
IHC - immunohistochemistry
IRE1 - inositol-requiring transmembrane kinase/endonuclease
KLK3 - prostate-specific antigen
KO - knockout
LBD - ligand-binding domain
LCoR - ligand-dependent corepressor
LIF - leukemia inhibitory factor
MAPK - mitogen-activated protein kinase
MaSC - mammary stem cell
MeCP2 - methyl cytosine binding protein 2
MEN1 - multiple endocrine neoplasia type 1
METABRIC - Molecular Taxonomy of Breast Cancer International Consortium
MNase - micrococcal nuclease
MRC - mammary repopulating cells
NF-κB - nuclear factor-κB
NR0B2 - nuclear receptor subfamily 0 group B member 2
PCAF - p300/CBP-associated factor
PI3K - phosphoinositide 3-kinase
PR- progesterone receptor
PSA - prostate-specific antigen
PTCL - peripheral T-cell lymphoma
Rb - retinoblastoma
REA - repressor of estrogen receptor activity
RIME - rapid immunoprecipitation mass spectrometry of endogenous proteins
ROS - reactive oxygen species
RUNX1 - runt-related transcription factor 1
SAHA - suberoylanilidehydroxamic acid

SAM - sterile alpha motif
SERD - selective estrogen receptor downregulators
SERM - selective estrogen receptor modulators
SIRT - sirtuin
SP1 - specificity protein 1
SPDEF - Sterile alpha motif (SAM) pointed domain containing ETS transcription factor
STAT - signal transducer and activator of transcription
TBP2 - thioredoxin-binding protein 2
TCGA - the cancer genome atlas
TEB - terminal end bud
TF - transcription factor
TFBS - transcription factor binding site
Th - tyrosine hydroxylase
TMPRSS2 - transmembrane protease serine 2
TTR - transthyretin
UPR - unfolded protein response
VDR - vitamin D receptor
Wap - whey acidic protein
WT - wild type
XBP1 - X-box protein 1

Acknowledgments

Sylvie, thank you for granting me the opportunity to pursue my graduate studies in your lab back in 2009. This has had a profound impact on me. Moving to Montreal and conducting research with you allowed me to grow both personally and scientifically, and for that, I am forever grateful. You have truly taught me how to be a scientist and an investigator. And to me, that is the greatest honour of all.

I would also like to thank all members of the lab, past and present. Working with you has been an absolute pleasure. In particular, I want to thank Martine Bail who patiently trained me at the bench when I first joined the lab. Our many intelligent and stimulating discussions constantly motivated me. Thanks for all your help, Martine. I could not have done this without you.

I wish to thank members of my PhD graduate committee, Dr. Lea Harrington and Dr. Michael Witcher, for their precious feedback and time.

I want to graciously thank all the agencies who provided funding for my studies: the Canadian Institutes of Health Research, Quebec Breast Cancer Foundation, and Canadian Breast Cancer Research Alliance.

To all my friends in Montreal and abroad, thank you so much for all your love and support. Alex, my Californian ray of sunshine, thank you for your unconditional support. I'm very grateful for your friendship and for your wisdom, which has helped keep me sane during difficult times. Justyna, thank you for making my PhD and life in Montreal a joyful experience. And thank you for the helpful discussions and comments on this thesis. You are and will always be my muse.

Hiba, thank you for being my home away from home.

io, my grey feline bestfriend and pawoneer factor expert, thank you profusely for your kind assistance in the writing of this thesis.

Lastly, thank you to Mona and Fouad, for all your support throughout the years and for all the laughs we have shared. I could not have done this without you.

Finally, mom and dad, thank you for always believing in my potential. Dad, you are my soul and my hero. I love you forever and I hope I made you proud.

Gaily bedight,
A gallant knight,
In sunshine and in shadow,
Had journeyed long,
Singing a song,
In search of Eldorado.

But he grew old—
This knight so bold—
And o'er his heart a shadow
Fell as he found
No spot of ground
That looked like Eldorado.

And, as his strength
Failed him at length,
He met a pilgrim shadow—
"Shadow," said he,
"Where can it be—
This land of Eldorado?"

"Over the Mountains
Of the Moon,
Down the Valley of the Shadow,
Ride, boldly ride,"
The shade replied,—
"If you seek for Eldorado!"

— Edgar Allan Poe, *Eldorado*, 1849

Chapter One.

Introduction

1. The mammary gland

The mammary gland is a remarkable tissue in mammals in that it has the unique capacity to produce milk to sustain the newborn. It is composed of two main tissue compartments: the epithelium and the mammary fat pad or stroma. Two main cell lineages make up the mammary epithelium: luminal cells, which line a lumen, and basal cells (mainly contractile myoepithelial cells), which surround luminal cells and are positioned adjacent to the basement membrane (1). Both of these cell types form ductal structures that invade the stromal compartment and terminate in milk-producing alveolar units during pregnancy and lactation. The stromal compartment is primarily composed of adipocytes, in addition to fibroblasts, the extracellular matrix (ECM) and cells of the haematopoietic system. The epithelial tissue compartment is highly dynamic as it is subject to tightly coordinated hormonal regulation during different stages of mammary differentiation.

1.1 The life and death of the mammary gland

Most of our knowledge of mammary gland development comes from mouse studies, because of the feasibility of gland manipulation. Despite some differences in hormonal regulation and the architecture of the human mammary gland, the mouse mammary gland has been used for decades as a model system for understanding the biology of the human mammary gland. Development of the mouse mammary gland begins prenatally in the embryo with the majority of its morphogenesis occurring postnatally during puberty and reproduction. At birth, the mouse mammary gland is seen as a small rudimentary ductal tree connected to the nipple. As the body develops, this structure grows in an allometric fashion, until the onset of puberty when highly proliferative structures, termed terminal end buds (TEB), begin to form at the tip of every duct (2). These bulbous structures are made up of two main cell compartments: (i) the highly proliferative outer ‘cap cell’ layer, which is thought to contain mammary stem cells (MaSCs) that can differentiate into myoepithelial cells that would envelope the duct as it elongates, and (ii) a multi-cellular layer of ‘body cells’, which is thought to comprise luminal and alveolar progenitors that differentiate into mature luminal cells (2, 3). In addition, apoptosis in the ‘body cell’ layer of the TEB is thought to assist in the creation of a lumen (4). Our current understanding of the contribution of hormonal cues to development of the mammary gland owes largely to elegant endocrine-ablation studies performed in the 1950s (5, 6). In these studies, the minimal hormonal requirements for mammary gland development were identified through surgical removal of ovaries, which produce estrogens and

progesterone, pituitary gland, which synthesizes prolactin and growth hormone (GH), and adrenal gland, which produces cortisol, coupled with subsequent hormone replacement. Indeed, ductal elongation and side-branching into the fat pad ensues following exposure to circulating hormones 17β-estradiol (E2) and progesterone, which elicit their proliferative effects through binding estrogen receptor alpha (ERα) and progesterone receptor (PR), respectively (7–9). Following puberty, and at each menstrual cycle, expansive proliferation takes place in the mammary gland followed by tissue remodelling and cell death should fertilization/implantation not take place. However, in the case of pregnancy, exposure to prolactin, estrogen and progesterone causes gross expansion of the epithelial compartment to fully occupy the fat pad **(Figure 1)**. During pregnancy, secretory alveolar units begin to form. These are small sacs composed of an inner epithelial cell layer that lines the lumen encapsulated by a myoepithelial cell outer layer which contracts to push milk through ducts. Upon binding to prolactin, the prolactin receptor activates a number of signalling pathways, primarily the JAK2/STAT5 cascade, in addition to mitogen-activated protein kinase (MAPK) and phosphoinositide (PI) 3-kinase (PI3K) (10). Collectively, these signals drive terminal lactogenic differentiation of epithelial cells within alveoli and generate a lactation-competent gland. This stage of terminal differentiation is also known as secretory differentiation and occurs in two phases: secretory differentiation I (Initiation) and II (Activation). The Initiation phase begins around mid-pregnancy where alveolar cells become primed for milk production. However, only limited secretion of milk components is detected at this stage both in mouse and human (11). At parturition, a drastic drop in progesterone levels is thought to trigger the Activation phase, which lasts 3 weeks postpartum (10, 12). This phase is marked by transcriptional activation of milk protein genes, such as whey acidic protein (*Wap*) and casein beta (*Csnb*) (10), lipid biosynthesizing enzymes (13), accumulation of lipid droplets (13), Golgi apparatus maturation and mitochondrial expansion (14, 15). Unlike mouse, in humans, this phase takes place shortly after parturition (11). Lack of suckling at weaning marks the beginning of involution, the final stage of mammary gland morphogenesis, whereby alveolar cells undergo massive cell death and the gland regresses back to its pre-pregnancy stage. Involution happens in two phases. The first phase is reversible (first 48-72 h) in that milk production can restart if suckling is resumed (10). In addition, it marks the onset of p53-dependent apoptosis (16). The second phase (8-10 days) is irreversible, and is accompanied by a second wave of apoptosis that is p53-independent and mediated through increased expression of serine proteases and matrix metalloproteinases (10, 17). Notably, adipocyte

cell differentiation takes place during this phase to replace the dying epithelial tissue and replenish the fat pad (17, 18). Although the exact trigger of involution remains unclear, signalling by members of the signal transducer and activator of transcription (STAT) protein family has been shown to play a role. Particularly, STAT3 propagates pro-apoptotic signals that antagonize pro-survival signals of STAT5 (19). Following its activation by Leukemia Inhibitory Factor (LIF), STAT3 induces expression of two inhibitory subunits of PI3K, namely p50α and p55α, which result in reduced activation of the serine-threonine protein kinase AKT1 and its pro-survival signals (20, 21). Another mechanism by which STAT3 can regulate involution is through induction of insulin-like growth factor (IGF) binding protein 5 (IGFBP5) expression, which inhibits pro-proliferative IGF1 signalling (22).

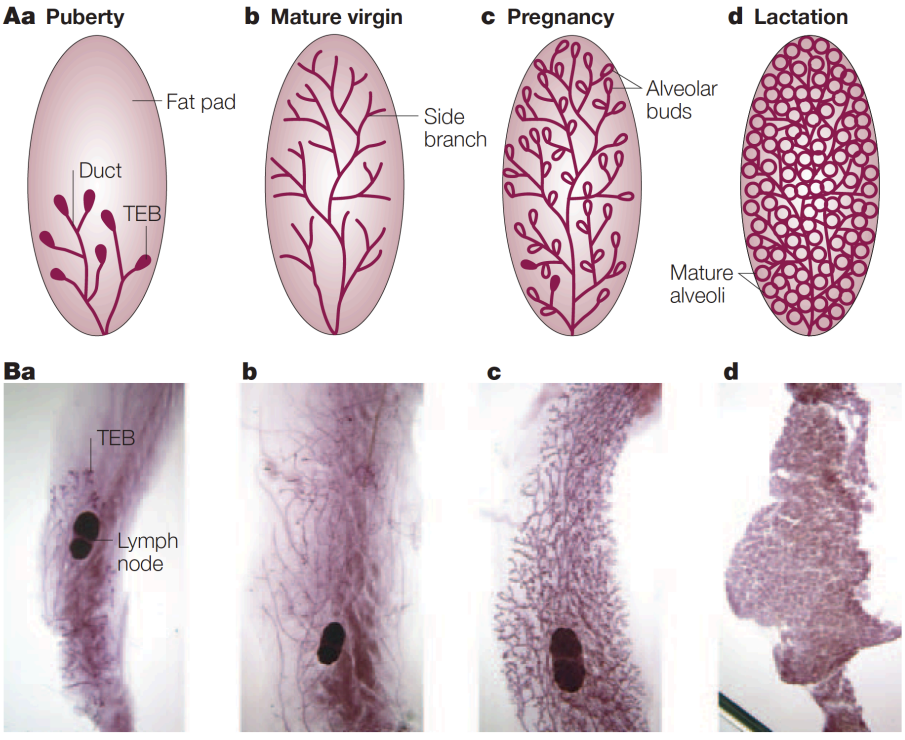


Figure 1. Mouse mammary gland development during puberty, pregnancy and lactation. Schematic (Aa–d) and wholemount (Ba–d) representation of the different stages of mammary gland development. A rudimentary ductal tree is observed at birth which grows gradually until puberty when estrogens and progesterone drive ductal side branching and invasion of the fat pad. The bulbous terminal end buds are also shown. During pregnancy, massive epithelial cell expansion takes place and the fat pad is fully invaded by alveolar structures following exposure to prolactin and progesterone resulting in a lactation-competent gland. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology **6**, 715-725, copyright (2005). (Reference (1))

1.2 Epithelial Hierarchy in the Mammary Gland

Transplantation studies have been instrumental in shaping our understanding of the epithelial hierarchy in the mammary gland. In 1959, DeOme et al. demonstrated for the first time that transplantation of normal mammary epithelial tissue of a donor mouse into a de-epithelialized mammary fat pad of a host mouse can fully reconstitute the mammary ductal tree (23). This was one of the first lines of evidence to suggest existence of mammary repopulating cells (MRCs), thought to be enriched for MaSCs. Indeed, Hoshino et al. and others subsequently corroborated these findings where transplantation of any portion of the mouse mammary gland into cleared fat pads could regenerate a ductal tree that was lactation-competent (24–26). The advent of fluorescence-activated cell sorting (FACS) technology allowed the characterization of cell surface markers of MRCs. Following dissociation of the mouse mammary gland into its different cell populations and exclusion of endothelial ($CD31^{+}$) and haematopoietic ($CD45^{+}$) cells, fat pad transplantation of a single cell that was $CD29^{high}CD24^{+}CD31^{-}CD45^{-}$ had the capacity to generate a fully functional bi-lineage mammary gland (27, 28). These findings suggested that a multipotent $CD49f^{high}$ $CD29^{high}CD24^{+}$ MaSC can fully reconstitute mature luminal, alveolar and myoepithelial cells of the mammary gland. However, recent lineage-tracing experiments revealed that the role of multipotent MaSCs may be restricted to embryogenesis and luminal ($CD29^{lo}CD24^{+}$) and myoepithelial ($CD29^{high}CD24^{+}$) unipotent stem cells are the main contributor to maintenance and expansion of the epithelial compartment in the adult mouse mammary gland (29). Remarkably, the myoepithelial unipotent stem cell-enriched population could regenerate a full mammary gland when transplanted into cleared fat pads, indicating that under conditions of transplantation assays, it either acquires ‘multipotent stem cell’ behaviour or dormant multipotent stem cells become reactivated to combat injury and restore homeostasis (29, 30). Despite the progress that has been made in understanding the hierarchy in the mammary gland, several questions remain unanswered. What does the differentiation path between mammary stem cells and mature committed cells entail? Are there several luminal and myoepithelial intermediary progenitors? What is the degree of plasticity of these cell populations? Alveolar progenitors that differentiate to mature alveolar cells during pregnancy have also been described (31). These cells were identified as a subset of the luminal lineage that are ERα-negative $Sca1^{-}CD49b^{+}$ and express higher levels of milk proteins in the virgin state in comparison to ERα-positive $Sca1^{+}CD49b^{+}$ luminal progenitors (31).

2. Luminal transcription factors

Transcription factors (TFs) are remarkable proteins because they can cause the selective expression of subsets of genes that ultimately dictate cell identity. One important example of how TFs can reprogram cell fate is the conversion of fibroblast cells into induced pluripotent stem cells following exogenous introduction of a specific cocktail of factors (32). As mentioned in the previous section, E2 can regulate development of the mammary gland through ERα, a ligand-dependent TF. Forkhead box A1 (FOXA1) and GATA binding protein 3 (GATA3), two other luminal lineage-restricted TFs, gained considerable attention over the recent years following gene expression profiling studies of breast cancer cells and purified normal mammary epithelial cell populations. Genetic knockouts (KO) of these TFs in mice revealed their important roles in development of the normal mammary gland.

2.1 The estrogen receptor

Estrogens play key roles in the development and function of normal sexual and reproductive functions in both males and females. In addition, estrogens are known to have protective biological effects in musculoskeletal, immune, bone as well as central nervous systems (33). They are also protective against cardiovascular disease, where they can attenuate ventricular hypertrophy and prevent ischemic arrhythmias (34). In humans, ERα and ERβ, both of which belong to the nuclear receptor superfamily of ligand-inducible TFs, mediate the effects of estrogens. They are encoded by separate genes on different chromosomes: *ESR1* on chromosome 6 encodes ERα and *ESR2* on chromosome 14 encodes ERβ. These two receptors share evolutionarily-conserved structural and functional domains: a ligand-binding domain (LBD) that is composed of 12 α-helices (H1-12), a hinge region that contains a nuclear localization signal, and a highly conserved DNA-binding domain (DBD) flanked by two transcriptional activation function (AF) domains. AF1 is constitutively active whereas AF2 activation is dependent on ligand binding (35).

The studies described in this thesis are focused exclusively on the functions of ERα. Liganded ERα can mediate the effects of E2 through either genomic or non-genomic (rapid) actions as summarized in **Figure 2.** The genomic actions of ERα can either be ligand-dependent or independent and result in direct or indirect (via tethering) engagement of E2 target genes.

2.1.1 ERα signalling

Liganded ERα binds as a dimer to estrogen response elements (EREs) in the promoters or enhancers of its target genes. The ERE consensus sequence has been described as a 15bp palindromic inverted repeat that contains a 3bp spacer: 5′-(A/G)GGTCAnnnTGACC(T/C)-3′ (36). Upon binding E2, ERα undergoes a conformational change that repositions helix 12 over the ligand-binding pocket. This results in the formation of an interaction surface that exposes several amino acid residues (Asp538, Leu539, Glu542 and Met543) that mediate recruitment of transcriptional coactivators (37). This is essential for the receptor to gain full transcriptional activity. The p160 family of homologous steroid receptor coactivators comprising SRC-1, SRC-2 (GRIP1) and SRC-3 (AIB1 or ACTR) are amongst the most characterized coactivators of ERα (38). Not only can these factors engage in direct or indirect recruitment of other coactivators, but they also harbour intrinsic, albeit weak, histone acetyltransferase activity (HAT) that can assist in chromatin opening and allow subsequent recruitment of basal transcriptional machinery (39, 40). Structurally, these coactivators contain highly conserved LXXLL motifs that can interact with the hydrophobic groove of the LBD of ERα that becomes revealed following E2-induced rearrangement of H12 (41, 42). Interaction of SRC proteins with ERα also provides a platform for recruitment of additional cofactors that harbour histone modifying and chromatin remodelling activities. These include the histone acetyltransferases p300, CREB-binding protein (CBP) and p300/CBP-associated factor (PCAF) (40, 43, 44), protein arginine methyltransferase CARM1 (45–47), as well as SWI/SNF complexes (48, 49). *TFF1* (formerly pS2), an E2-target gene (50), has been historically used as a model gene to study transcriptional activation by ERα. It has been demonstrated through chromatin immunoprecipitation (ChIP) and re-ChIP, that ERα and its cofactors can cycle the *TFF1* gene promoter on and off in 45-minute cycles (51). Interestingly, at the end of each cycle, ERα clearance is associated with a decrease in binding of p300, CBP and PCAF and increased recruitment of histone deacetylase 1 (HDAC1) and 7 before a second round of transcription begins and chromatin is opened again (51). Consistent with the findings that ERα can recruit HDAC activities to an induced gene in the presence of E2, the receptor can also repress E2-dependent transcription through recruiting corepressors including ligand-dependent corepressor (LCoR) and c-terminal binding protein 1 (CtBP1), which could in turn recruit HDAC3 and 6 (52). An interaction between the LBD of ERα and repressor of estrogen receptor activity (REA) was also observed, where REA inhibited ERα's transcriptional activity by competing with SRC-1

binding (53) or by recruiting HDAC1 (54). On the other hand, ERα can engage its target genes indirectly via protein-protein interactions. This allows the receptor to regulate genes that don't harbour consensus EREs. Indeed, ERα can interact with activator protein 1 (AP1) (55), specificity protein 1 (SP1) (56), nuclear factor-κB (NF-κB) (57), cAMP response element-binding protein (CREB) (58), p53 (59), and recently discovered, runt-related transcription factor 1 (RUNX1) (60), to regulate expression of its target genes. The genomic mechanisms described above could also take place in a ligand-independent manner. For instance, phosphorylation of unliganded ERα can stimulate its transcriptional activity. Several serine residues in the AF1 domain (S104/106/118) of ERα are targets of phosphorylation of MAPK following activation of growth factor signalling pathways such as IGF-1 and epidermal growth factor (EGF) (61–63). Another example of activation of unliganded ERα is through its interaction with cyclin D1, which can act as a bridge to mediate the receptor's interaction with coactivator proteins SRC1 and 3 (64, 65). This interaction cannot take place in a direct manner in the absence of E2-induced conformational changes of ERα, as mentioned above. Lastly, a small fraction of plasma membrane-bound ERα can exert rapid, nongenomic actions by interacting with signalling molecules such as G-proteins (66), as well as c-Src (67) which results in activation of the MAPK and PI3K/AKT protein kinase signalling pathways in the cytoplasm. Even though ERα does not contain an intrinsic trans-membrane domain, it is thought to associate with the plasma membrane through palmitoylation of Cys447 and interaction with caveolin 1 (68). Ultimately, these nongenomic actions of ERα also result in genomic changes following activation of downstream TFs by activated kinase pathways.

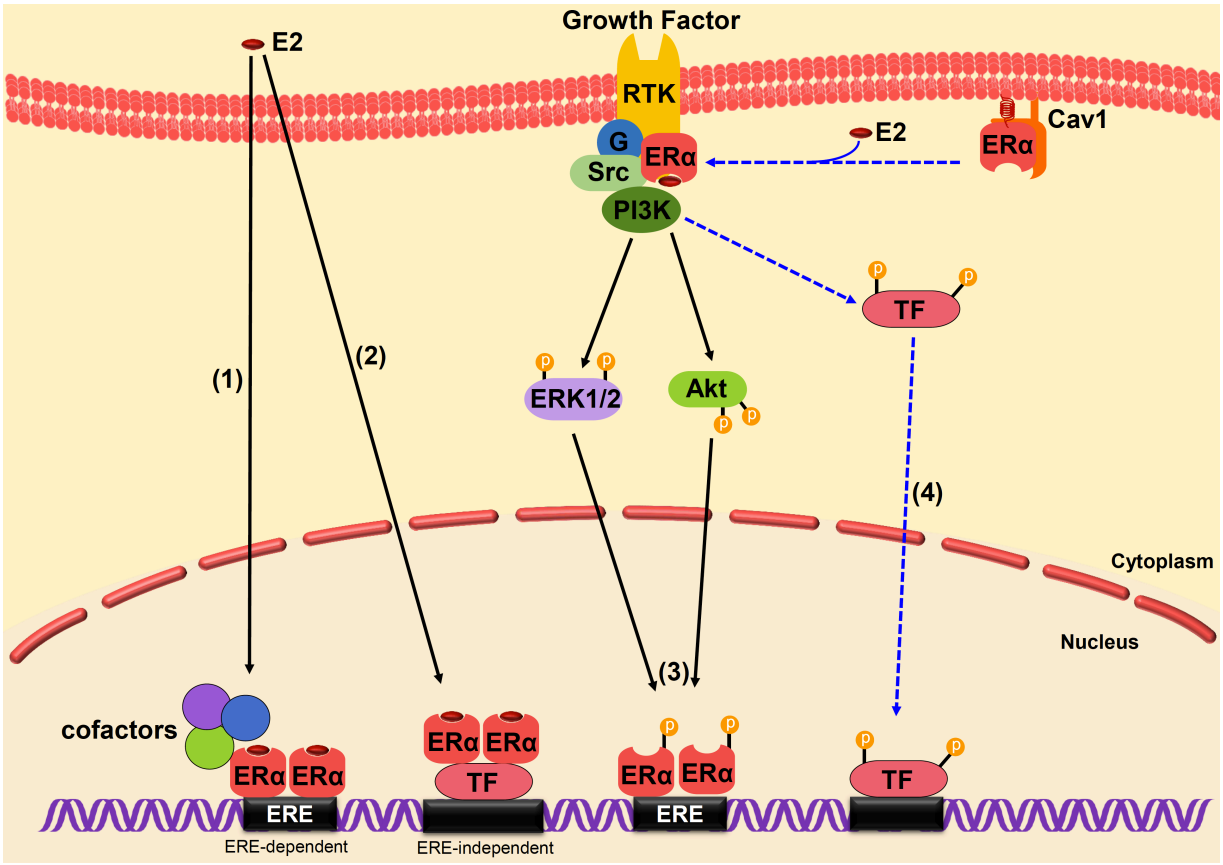


Figure 2. Genomic and nongenomic actions of ERα.
(1) Upon binding to its ligand, E2, ERα dimers bind to EREs with the help cofactor proteins or (2) activates expression of its target genes by tethering to other TFs. Additionally, unliganded ERα can be phosphorylated following activation of growth factor signalling cascades, which activates the receptor (3). Finally, ERα can activate protein kinase signalling cascades in the cytoplasm resulting in activation of downstream TFs through increased phosphorylation (4). This is known as nongenomic ERα signalling.

2.1.2 Lessons from ERα KO mice

In the mammary gland, ERα is expressed in a subset of luminal epithelial cells surrounding the ductal lumen, but not detectable in myoepithelial cells **(Figure 3)**. ERα can also be detected in the stromal layer (69). The first ERα knockout mouse was generated in 1993 by Lubahn et al. by insertion of the neomycin gene into exon 2 of the mouse *Esr1* gene using classical genetic recombination in embryonic stem cells. Deletion of ERα resulted in infertility in both sexes. In females, this was associated with hypoplastic uteri that lose responsiveness to estrogen stimulation. In addition, these mice are not capable of ovulation. Males exhibited abnormal spermatogenesis and reduced sperm counts, and decreased aggression and mating behaviour (70). Importantly, while the mammary gland of ERα KO mice develops normally in the prenatal and prepubertal stages, it remains a simple rudimentary ductal tree and does not undergo any additional morphogenesis at puberty (71). As mentioned earlier, PR plays important roles in regulating this process. Expectedly, it was found that mammary glands of ERα KO mice had reduced *PGR* mRNA expression (the gene encoding PR) consistent with a positive regulatory role of ERα in regulating PR expression (71). Thus, the lack of mammary gland development in these mice is likely due to the combined loss of ERα expression and the subsequent decrease in PR expression. The finding that deletion of ERα does not affect prenatal development of the mammary gland is consistent with low ERα expression in the MaSC population (72). However, it remains unclear what impact deletion of ERα has on the ERα-positive $Sca1^{+}CD49b^{+}$ luminal progenitor population previously described (31). Notably, in comparison, ERβ KO mice were fertile although females exhibited infrequent pregnancies and smaller litter size (73). In addition, these mice harboured normal ovarian, uterine and testicular development. Interestingly, mammary glands of these mice develop normally and can fully differentiate to the stage of lactation. This data indicated that, of the two receptors, ERα is the main player in regulating development of the mammary gland.

Additional studies have assessed the individual contribution of stromal as well as epithelial ERα in mammary gland development. Cunha et al. were the first group to address this question (74). They generated four different combinations of recombined neonatal epithelial and stromal tissue from ERα KO and wild type (WT) mice: (i) WT epithelium and WT stroma, (ii) WT epithelium and ERα KO stroma, (iii) ERα KO epithelium and WT stroma, and (iv) ERα KO epithelium and ERα KO stroma. Following transplantation under the kidney capsule of nude athymic mice, no ductal growth could be detected in (ii) whereas ductal

growth and branching morphogenesis was observed in (iii). This study indicated that stromal, but not epithelial ERα, was required for mammary gland development (74). However, a subsequent study by Mueller et al. contrasted these findings where they found that both stromal and epithelial ERα is required for mammary growth postpuberty (75). Importantly, it was later discovered that the ERα KO mouse generated by Lubahn et al. and used in these subsequent studies expresses an alternatively spliced form of ERα that retains the DBD and LBD/AF2 domains of the WT receptor and partially retains transcriptional transactivation potential (76, 77). This prompted the generation of a new ERα KO mouse (78), where it was found that deletion of ERα in the mouse mammary gland also resulted in a rudimentary ductal tree that could not develop past the pubertal stage (79). In addition, this study found that the epithelial ERα, but not the stromal, is needed for development of the mammary gland. Lastly, Feng et al. employed conditional deletion of ERα in the mammary gland at different stages of its development and confirmed its importance not only for ductal elongation but also for the process of alveologenesis and lactation (80). Furthermore, consistent with the important roles of progesterone mentioned above, a similar phenotype was observed in PR KO mice, where ductal side-branching and alveologenesis were also impaired (81).

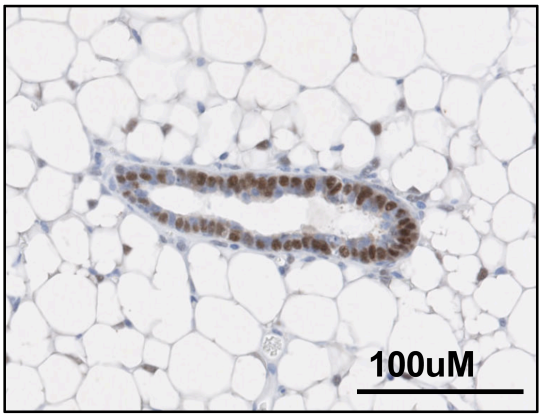


Figure 3. ERα is expressed in luminal mammary epithelial cells.
A representative image of nuclear ERα immunohistochemical staining in a virgin mouse mammary gland. Brown nuclei mark ERα-positive luminal cells lining a duct.

2.2 FOXA1

The forkhead box (FOX) family of TFs comprises 50 members in humans that are unique in their 'winged-helix' DBD and their capacity to spatially and temporally regulate gene expression during development (82). These TFs can be divided into 19 subgroups (FOXA to FOXS) based on homology of their DBDs. The family acquired its name following the discovery of its founding member in 1989, where mutations in the *fkh* gene in Drosophila Melanogaster embryos were found to cause homeotic transformation of the foregut and

hindgut which become replaced by ectopic head structures (83). That same year, FOXA1 or hepatocyte nuclear factor 3A (HNF3A), the first member of the FOXA subgroup, was described in a seminal study by Costa et al. as a regulator of the liver-specific genes Transthyretin (*TTR*) and alpha 1-antitrypsin (alpha 1-AT; *SERPINA1*) (84). Structurally, the highly-conserved forkhead DBD, which lies at the centre of the protein, consists of three α-helices (α1, 2, 3), three β-sheets (β1, 2, 3) and two wing (loop) domains (W1 and 2) arranged in α1-β1-α2-α3-β2-W1-β3-W2 order. This imparts a butterfly-like structure to FOXA1, where helix α3 primarily binds the DNA major groove and W2 binds the minor groove (85). The consensus FOXA1 response element (FOXRE) has been described as 5'-A(A/T)TRTT(G/T)RYTY-3' (86). In addition to its forkhead domain, FOXA1 also contains N- and C-terminal transactivation domains as well as conserved nuclear localization signal sequences. Because the structure of FOXA1 resembles that of linker histones H1 and H5, both of which also contain winged-helix domains, FOXA1 is able to make contacts with the sides of nucleosome cores (85). It was previously reported that four basic amino acids (K40 and 52 and R42 and 94) present in H1 are indispensable for its ability to compact nucleosomes (87). Unlike linker histone binding, FOXA1 binding does not result in chromatin condensation because this TF lacks any basic amino acids at its nucleosome binding interface (88). Indeed, addition of these residues at structurally similar sites in FOXA1 resulted in chromatin compaction as evidenced by protection of DNA from micrococcal nuclease (MNase) digestion (88). The current paradigm of transcriptional regulation follows that chromatin remodelling proteins can facilitate chromatin opening to allow TF binding and recruitment of the basal transcriptional machinery. The finding that FOXA1 opened compacted chromatin at the albumin gene enhancer *in vitro* even in the absence of chromatin remodelling enzymes lead to coining of the term 'pioneer factors' (89). Pioneer factors are defined as factors that have the capacity to bind heterochromatin and facilitate its opening either intrinsically or via recruitment of additional factors. The hypothesis regarding the mechanism of action of FOXA pioneer factors follows that their concomitant binding to both DNA and core histones abrogates inter-nucleosomal interactions mediated by histones H3 and H4 (89) and additionally stabilizes the position of underlying nucleosomes (90). This in turn may allow FOXA1 to modulate binding of additional TFs through phasing their local binding sites on and off of nucleosomes. The role of FOXA1 as a pioneer factor and regulator of ERα signalling in breast cancer will be further discussed in section **3.2.1 and Chapter 2.**

While FOXA1 has a positive regulatory role in the expression of liver-specific genes, livers of FOXA1 KO mice develop normally. It was determined that this phenotype results from FOXA2 compensation, which when deleted alone has no effect on liver development, but when deleted together with FOXA1 abrogates fetal liver development (91). FOXA1 can also regulate development of a number of tissues including pancreas, kidney, brain and lung. FOXA1 KO mice suffer from weakness, hypoglycemia and dehydration and die postnatally. This was in part due to reduced secretion of glucagon into the circulation because of disrupted expression of the proglucagon gene (*Gcg*) which is normally controlled by FOXA1 in pancreatic islet α-cells (92). It remains unclear why these mice suffer severe dehydration. Although kidney defects were reported, these could not readily explain the dehydration phenotype. Conditional FOXA1 knockouts in the kidney tissue will ultimately be crucial to further interrogate this phenotype. Brain defects in FOXA1 KO mice involved impaired maturation of midbrain dopaminergic neurons as a result of loss of expression of Nurr1, engrailed 1 and neurogenin 2, important factors for their differentiation process (93). Lung development defects have also been reported and these included delayed alveolarization associated with decreased expression of epithelial differentiation markers (94).

Up until recently, we were unable to examine whether FOXA1 plays any role in development of the mammary gland because, as mentioned earlier, mice die shortly after birth. Bernardo et al. devised an elegant study where they employed the use of orthotopic transplantation and renal capsule grafting to be able to monitor growth of mammary glands of FOXA1 KO mice postnatally until puberty (95). Similar to the ERα KO mouse phenotype, a rudimentary mammary ductal tree developed normally in the prenatal stage of FOXA1 KO mice, however it failed to grow beyond that and remained an anlage, even after several weeks of exposure to pregnancy hormones. Notably, it was found that, similar to ERα ductal localization, FOXA1 is also expressed in luminal cells but not myoepithelial cells in the WT gland **(Figure 4)**. Interestingly, FOXA1 mRNA levels increased as differentiation progressed from a stem cell to a luminal progenitor to a mature luminal cell in the WT gland. This raised the possibility that FOXA1 KO mice might have defects in luminal cell specification. However, using immunohistochemical staining (IHC), the authors found no major difference in expression/localization of luminal (Cytokeratin 8) or myoepithelial (α-Smooth muscle actin) cell markers between WT and KO mice. Importantly, while transplanted FOXA1 KO mammary epithelium was not capable of expansive proliferation when recipient mice became pregnant, it was capable of generating alveoli that stained positively for milk proteins and lipid droplets.

Additionally, glands with deletion of a single FOXA1 allele (heterozygotes) had increased alveoli formation in comparison to WT glands following exposure to pregnancy hormones. These data indicated that (i) FOXA1 is necessary for ductal outgrowth and invasion of the mammary anlage at puberty, and (ii) FOXA1 deletion enhances the process of alveologenesis during pregnancy. Several questions regarding FOXA1 biology in mammary gland development remain unresolved. Is FOXA1 important for luminal cell differentiation? What is the proportion of mature vs. progenitor/stem cells in FOXA1 KO mice? Does FOXA1 deletion result in accumulation of a less differentiated progenitor or stem cell population and/or depletion of the more differentiated mature luminal population? Using FACS analysis rather than non-quantitative IHC staining will help address these questions. Lastly, does FOXA1 truly inhibit alveologenesis? And does it alter the proportion of the alveolar progenitor cell population? To more directly answer these questions, mice that overexpress FOXA1 at different stages of mammary gland development will need to be generated.

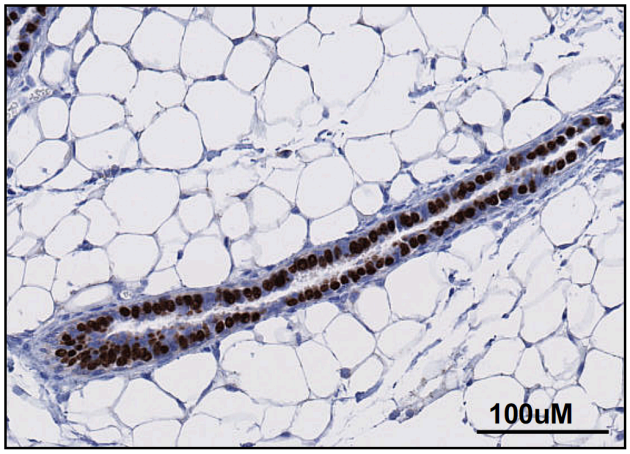


Figure 4. FOXA1 is expressed in luminal mammary epithelial cells.

2.3 GATA3

In humans, the GATA family of TFs comprises six members (GATA1-6), found on different chromosomes. Historically the GATA family has been divided into two subfamilies (GATA1-3 and GATA4-6), based on expression in hematopoietic and cardiac tissue, respectively. However, recent studies suggest that their functions extend to other types of tissues, including endothelium, breast and prostate. GATA proteins comprise two N-terminal transactivation domains and a dual zinc finger DBD that is highly conserved (>70%) amongst all GATA proteins and recognizes 5'-(A/T)GATA(A/G)-3' sequences (96). Each of the two zinc

fingers is followed by highly conserved and distinct basic regions (basic region 1 and 2) that have less defined functions.

The GATA consensus site is highly abundant in the human genome with about seven million GATA motifs present (97). Interestingly, not all GATA sites are bound by GATA proteins. In fact, ChIP-Seq studies have shown that GATA1 and 2 occupy only about 0.1 to 1% of these sites in erythroblast cells (97, 98). In addition, 90% of bound sites were found away from promoters suggesting that long-range interactions are necessary to regulate transcription. Indeed, GATA1 can induce higher order chromatin loops by interacting with its cofactor Friend of GATA-1 (FOG1) (99) or the chromatin remodelling factor BRG1 at the β-globin locus (100). While the mechanisms by which these factors choose their sites remain unclear, it has been reported that histone modifications (H3K9ac, H3K4me1 and me2) can demarcate GATA-bound sites. In addition, FOG1 (99, 101) and acetylation of GATA1 (102) have been reported to increase DNA accessibility at GATA sites. A potential pioneer factor role of this family of proteins was originally suggested by Bossard et al. where similar to FOXA1, GATA1-4 proteins were also found present at the albumin enhancer in endoderm cells that had not yet undergone liver specification (103). Recently, Takaku et al. addressed the pioneer factor role of GATA3 by comparing DNA accessibility before and after expression of this factor in a GATA3-negative breast cancer cell line (104). GATA3 expression resulted in opening of a subset of sites that coincided with concomitant BRG1 binding and enhanced gene expression. Notably, the same study found that GATA3-bound sites harboured high nucleosome density in the absence of GATA3 expression. Additionally, GATA3 was able to bind reconstituted nucleosomes *in vitro*, suggesting that GATA3 can bind closed chromatin.

The *GATA3* null mutation is embryonic lethal (11 days post coitum) due to extreme internal bleeding and neural as well as hematopoietic abnormalities with a block in T-cell differentiation (105). In addition, *GATA3 -/-* embryos exhibited noradrenaline deficiency in the sympathetic nervous system due to loss of expression of tyrosine hydroxylase (*Th*) and dopamine β-hydroxylase (*Dbh*) (106). It was found that administration of catechol intermediates to pregnant dams delayed embryonic lethality (12.5-16.5 days post coitum) thereby indicating that noradrenaline deficiency was a major cause of *GATA3 -/-* embryonic lethality (106). No defects were reported in *GATA3 -/+* mice (105), although recent evidence suggests that they suffer from sensorineural deafness (107). In the WT mouse mammary gland, GATA3 is expressed in luminal epithelial cells following a similar pattern of expression to that of ERα and FOXA1 **(Figure 5)**. Furthermore, GATA3 was identified as the most highly

expressed TF in the mammary epithelium. Conditional deletion of *GATA3* in the mammary gland using the *MMTV-cre* system results in severe defects in mammary gland development, including impaired ductal branching and elongation and failure to generate TEBs at puberty (108, 109). *Wap* promoter-Cre-mediated deletion of *GATA3* during pregnancy revealed impaired alveologenesis and lactation similar to the phenotype observed in ERα knockout mice. In addition, *GATA3* loss caused an accumulation of the $CD61^{+}CD29^{lo}CD24^{+}$ luminal progenitor population and a concomitant reduction in the $CD61^{-}$ mature luminal cell population, both of which were conversely depleted and expanded, respectively, as the mammary gland progressed to the pregnancy stage in WT mice (109). This indicates that loss of GATA3 causes a block in luminal cell differentiation. Importantly, ectopic expression of *GATA3* into the MaSC population lead to increased expression of milk protein genes, *Wap* and *Csnb* indicating that GATA3 is essential in driving differentiation along the alveolar lineage.

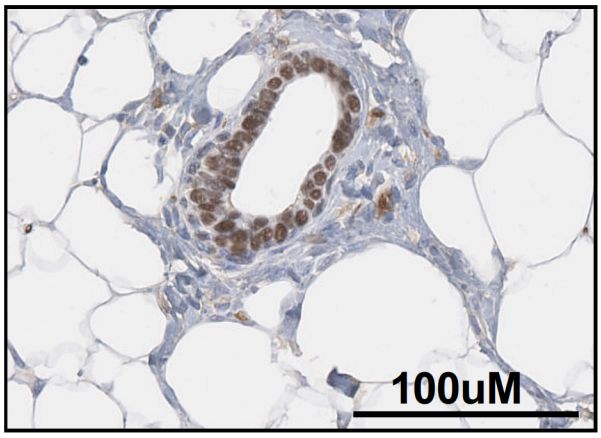


Figure 5. GATA3 is expressed in luminal mammary epithelial cells.

2.4 ETS family members: ELF5 and SPDEF

The human ETS family of TFs comprises 28 members that have a broad range of functions in development, cell fate specification and differentiation (110). These TFs share a highly conserved ETS winged helix-turn-helix domain DBD that recognizes the sequence 5'-GGA(A/T)-3'. While previous studies have reported a role of ELF5 in controlling the process of alveologenesis (111–113), it remains unclear what role other ETS family members have in mammary gland development. *ELF5* null mammary glands undergo normal ductal morphogenesis at puberty, however they are unable to produce alveoli during pregnancy and this results in lactational insufficiency (111). Conversely, forced expression of ELF5 at the

pubertal stage of mammary gland development activated alveolar differentiation, increased expression of *Wap* and *Csnb,* and lead to precocious milk production (111). Impaired alveologenesis in *ELF5 -/-* mice was associated with reduced expression and phosphorylation of STAT5, which was identified as direct ELF5 transcriptional target (112). In addition, *ELF5* null mammary glands exhibited increased expression of Notch1 and Notch2 leading to hyperactivation of Notch signalling and subsequent accumulation of MaSC population (113).

Sterile alpha motif (SAM) pointed domain containing ETS transcription factor (*SPDEF*), another ETS TF, was originally described as highly expressed in prostate tissue where it acts as a cofactor of androgen receptor (AR) to regulate expression of prostate-specific antigen (*KLK3*), a widely used diagnostic marker for prostate cancer (114). SPDEF KO mice are viable at birth and show defects in differentiation of a number of tissues, including lung, stomach and intestine (115–117). Regrettably, our knowledge of the function of this TF in the context of the mammary gland is limited to one published report. Buchwalter et al. found that *SPDEF* expression, like *FOXA1*, was highest in mature luminal cells and lowest in the MaSC population (118). Consistent with this observation, the authors show that knockdown of SPDEF in human luminal progenitor cells resulted in repression of luminal-lineage marker genes. On the other hand, forced expression in a non-transformed mammary epithelial cell line resulted in induction of luminal-lineage markers and repression of myoepithelial markers. Collectively, this data suggests that SPDEF regulates mammary luminal cell differentiation, although the mechanism by which this happens remains unclear.

2.5 XBP1

X-box protein 1 (XBP1), another luminal-lineage TF, has been recently described as an essential factor in mammary gland development. XBP1 is a ubiquitously expressed basic leucine zipper (bZIP) TF that was originally discovered as a regulator of MHC class II genes in B cells (119). It is most notably linked to the unfolded protein response (UPR). Upon accumulation of unfolded proteins in the endoplasmic reticulum (ER), the *trans*-acting TRF6 binds to the *cis*-acting ER stress response element (ERSE) in the *XBP1* promoter increasing its transcription (120). XBP1 mRNA subsequently undergoes splicing of 26 nucleotides through an ER transmembrane endoribonuclease and kinase called inositol-requiring transmembrane kinase/endonuclease 1 (IRE1). This results in a shorter transcript called XBP1(S) that encodes for a potent transactivator that translocates to the nucleus and activates transcription of UPR genes (121). Studies from *XBP1*-deficient mice reveal its

crucial roles in the development of secretory cells, hepatocyte and B-lymphocyte terminal differentiation (122–124). Targeted deletion of *XBP1* in the mammary gland hinders ductal branching at the virgin state, but does not alter the gland's capacity to produce alveoli. However, the subsequent process of lactation is severely impaired due to reduced lobuloalveolar development and reduced proliferation of the epithelial compartment because of increased ER stress (125, 126).

2.6 The androgen receptor (AR)

Like ERα, AR is a member of the nuclear receptor superfamily of ligand-inducible TFs and is essential for male sex differentiation (127). Upon binding testosterone ligand (produced primarily in the testes) or its derivative dihydrotestosterone (DHT) (produced intracellularly), heat shock proteins are displaced and AR dimerizes and undergoes a conformational change similar to that of ERα, where helix 12 rearrangement results in coactivator binding. AR binds to androgen response elements (ARE; GGTACAnnnTGTTCT) at enhancers and promoters of target genes, such as prostate-specific antigen (PSA, *KLK3*) and transmembrane protease serine 2 (*TMPRSS2*) (128, 129). In the prostate, AR is necessary for the growth and survival of normal cells. Genetic aberrations such as fusions between the promoter of *TMPRSS2* and the coding region of ETS TFs occur in a significant proportion of prostate tumours, which results in ETS factor overexpression (130). In addition, a number of mutations that target AR can increase risk of development of prostate cancer (131). Thus, androgen suppression therapy is administered be it surgical (castration) or chemical (treatment with gonadotropin-releasing hormone analogues). AR can also engage in non-genomic actions, interacting with c-src and the p85α subunit of PI3K to activate MAPK and AKT signalling cascades and enhance cell proliferation (132). In the human mammary gland, AR mainly localizes to nuclei of luminal and basal cells with some IHC staining detected in nuclei of stromal cells (133). In addition, AR -/- mice show defects in ductal morphogenesis during puberty (134). In addition, these mice produce alveoli and can lactate albeit much less than WT animals. Interestingly, it was found that MAPK signalling is inactive in AR -/- mammary glands, resulting in reduced expression of cyclin D1 and cell proliferation. Importantly, expression of ERα target genes was also reduced in AR -/- mammary glands, suggesting crosstalk between these two signalling pathways.

3. Breast cancer

Breast cancer remains the most frequently diagnosed cancer in women in Canada, followed by lung and colorectal cancer. An estimated 26,300 new breast cancer cases are expected in 2017, accounting for 25% of all new cancer cases in women (135). On average, one in eight Canadian women is expected to develop breast cancer within her lifetime (135). A number of risk factors contribute to development of breast cancer, including personal and family history, sex, age, *BRCA* tumour suppressor gene mutation status, density of breast tissue, reproductive history, and exposure to hormones. The implementation of provincial mammographic screening programs in 1988 resulted in an increase in breast cancer incidence in early 1990s, due to increased capacity for detection. The incidence dropped again in 2002, concurrent with a decline in hormone replacement therapy (HRT) use amongst post-menopausal women (135–137). Of note, breast cancer is very rare in males, accounting for less than 1% of new cases expected in Canada in 2017 (135).

While breast cancer is the second leading cause of cancer-related deaths amongst women after lung cancer, its mortality rate has been steadily declining owing to increased screening measures that allow for early detection of the disease, as well as major research advances in breast cancer treatment. Breast cancer diagnosis, prognosis and treatment depend on a number of factors including tumour type, size and histological grade, hormone receptor status, age, proliferative index (Ki67 positive) and whether metastasis/lymph node infiltration has occurred. These will dictate whether a patient will receive local therapy (surgery with or without adjuvant radiation therapy) with or without systemic therapy (chemotherapy, hormonal therapy and targeted therapy). Chemotherapy can be administered either in the neo-adjuvant or adjuvant setting, to shrink a large tumour prior to surgery or to prevent relapse due to residual cancer cells, respectively. Hormonal and targeted therapies, which will be discussed further below, depend on hormone receptor (ERα and PR) and epidermal growth factor receptor 2 (EGFR2 or HER2) protein expression.

3.1 Breast cancer is not a single disease

In the clinic, breast cancer has traditionally been classified into different histologic types based on histopathology (ductal or lobular *in situ* carcinoma vs. invasive carcinoma), pathological grade (grade 1: well differentiated, grade 2: moderately differentiated and grade 3: poorly differentiated), axillary lymph node status, and expression of protein markers (ERα, PR, and HER2) (138). The accuracy of the current methods for predicting the course of

breast cancer development has been questioned, as it can sometimes lead to ineffective treatment regimens.

The advent of gene expression profiling technologies over the last two decades revolutionized the way we perceive breast cancer. Two landmark studies in the early 2000s by Perou et al. and Sorlie et al. examined gene expression profiles of normal and tumourigenic breast samples using cDNA microarrays and unsupervised hierarchical clustering. This approach allowed the identification of several molecular subtypes of breast cancer (luminal A, luminal B, basal-like, HER2-enriched and normal-like), each characterized by its own signature and a different clinical outcome (139, 140), which have been confirmed by several other groups (141–143). In the clinic, these subtypes parallel those identified by IHC. Luminal tumours are ERα/PR+ HER2-/+, HER2-enriched tumours are ERα/PR- HER2+, and basal-like tumours are ERα/PR/HER2-. Normal-like breast tumours are also ERα/PR/HER2- and are characterized by expression of markers of adipose tissue and basal myoepithelial cells (140). Of note, the existence of the normal-like tumour subtype has been questioned due to the possibility of contamination with surrounding normal tissue. Two additional subtypes recently emerged from gene expression profiling studies by other groups: claudin-low and molecular apocrine (144, 145). Notably, these different molecular subtypes are thought to have arisen from blocks in cell differentiation together with genetic defects, as discussed below. Aligning the transcriptomes of particular molecular subtypes with those of their hypothesized cell of origin results in a tentative model **(Figure 6).**

Together, results from these studies engendered several commercially available prognostic signature tests, including the RT-qPCR-based OncotypeDx® and microarray-based Mammaprint®. The former estimates the risk of recurrence in ERα-positive early breast cancer, whereas the latter calculates a high-risk or low-risk recurrence score in lymph node negative breast cancers (<5cm) in patients younger than 61 years of age (138). However, the implementation of genomic tests into everyday use in the clinic to replace currently used histopathological methods is unlikely, due to high costs and time consumption. Instead, these tests will likely complement existing clinical methods for additional confirmation.

To fully capture the molecular basis underlying the observed breast tumour heterogeneity, correlate it with clinical features and tailor appropriate therapies, efforts were aimed at constructing large tumour datasets that integrate information at genomic, transcriptional, translational and epigenetic levels. The goals of these studies were to either

refine the previously reported molecular subtypes of breast cancer by identifying additional features (the METABRIC dataset) or to study the genomic complexity of the currently existing ones (the TCGA dataset). The Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) dataset comprised 2000 breast tumour samples, which were subjected to microarray-based gene expression profiling coupled with copy number alteration (CNA) analysis (146). This led to the identification of two novel breast cancer subgroups. One was ERα-positive and characterized by high mortality rate and amplifications in the *cis*-acting 11q13/14 region that contains driver genes such as *CCND1* (cyclin D1). The second subgroup comprised both ERα-positive and ERα-negative breast tumours that had good prognosis and were devoid of CNAs, but exhibited increased lymphocyte infiltration and a strong immune/inflammation gene signature. On the other hand, the Cancer Genome Atlas (TCGA) dataset comprised a total of 817 breast tumour samples that were studied using different platforms: RNA-Seq, whole-exome DNA sequencing, miRNA-Seq, SNP arrays, DNA methylations arrays, as well as reverse phase protein arrays (147, 148). Results from these studies revealed a plethora of genetic and epigenetic aberrations that could recapitulate the tumour subtypes previously identified by gene expression profiling.

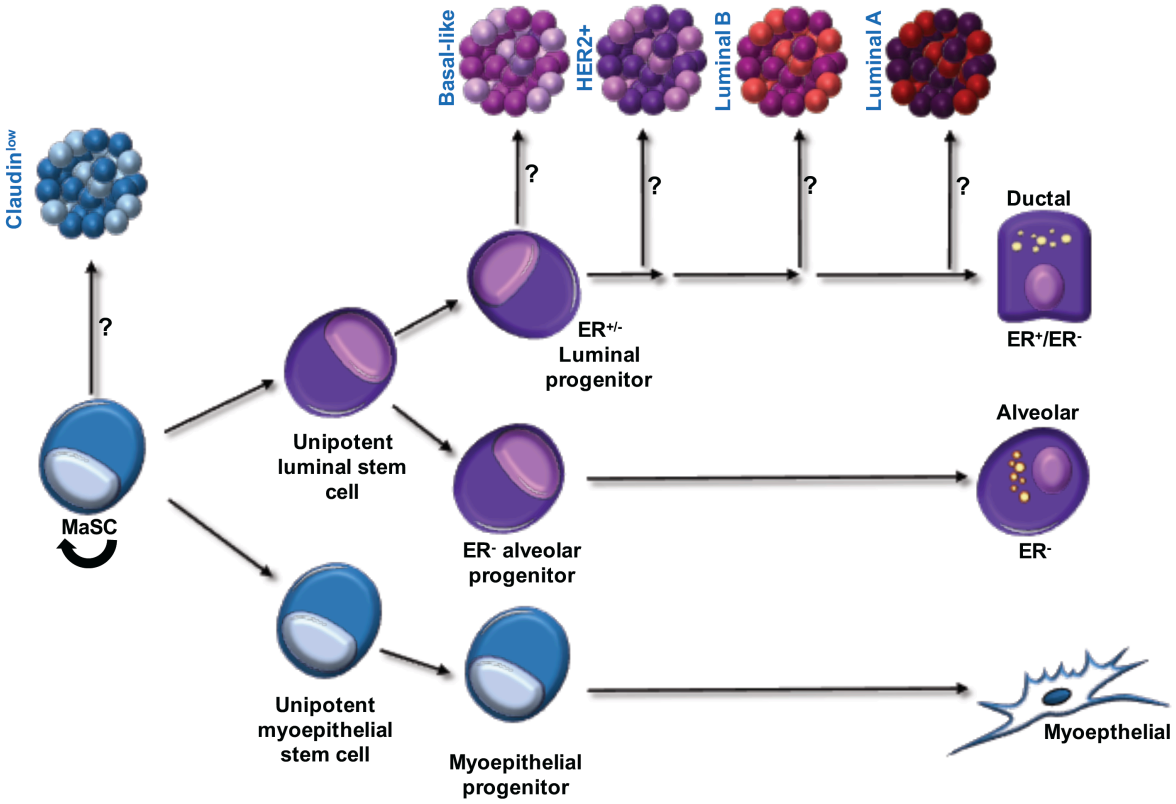


Figure 6. A tentative model outlining the different molecular subtypes of breast cancer and their hypothesized cell of origin. Adapted from Visvader et al. (149)

3.1.1 Luminal tumours

Luminal tumours, which represent about two-thirds of breast tumours, express ERα, and are dependent on estrogens for growth. Importantly, these tumours were dubbed 'luminal' because their gene signature comprises high expression of several luminal-lineage genes in addition to *ESR1*, including *FOXA1* and *GATA3* as well as *TFF3, SPDEF* and *XBP1 (140)*. Importantly, while the normal cell(s) of origin of these tumours remain(s) unclear, comparative gene expression profiling studies reveal that the transcriptome signature of these tumours closely aligns with that of normal luminal mammary epithelial cells (149). Luminal tumours can be further subdivided into two subgroups: luminal A and B. Luminal A tumours (ERα+ PR+ HER2-) are generally associated with smaller tumour size, lower grade and a better prognosis owing to high expression of ERα, which can be targeted by competitive inhibitors called antiestrogens. On the other hand, luminal B tumours (ERα+ PR+ HER2-/+) express lower levels of ERα and luminal-specific genes, have a higher proliferative index, and a worse clinical outcome in comparison to luminal A tumours (140). A third subgroup of luminal tumours termed luminal C has also been described (150). Similar to Luminal B tumours, Luminal C tumours are highly proliferative, although they position at the boundary between ERα+ and ERα- tumours. In addition, 40% of these tumours overexpress HER2 (150).

3.1.1.1 Hormonal therapy

The original work by Thomas Beatson in 1878 indicated that ovarian ablation (oophorectomy) abolished lactation in rabbits. Addressing the Edinburgh Medico-Chirurgical Society in 1896, Beatson pointed out that, "*This fact seemed to me of great interest, for it pointed to one organ holding control over the secretion of another and separate organ*." This finding lead Beatson to the discovery that surgical removal of ovaries was in fact beneficial in an advanced breast cancer patient (151). It wasn't until 60 years later that Jensen discovered ERα, which was subsequently identified as the target of ovarian estrogens (152). As mentioned earlier, exposure to estrogens is an important risk factor in breast cancer. In fact, early age at menarche, late age at menopause as well as use of hormone contraceptives and HRT have been linked to increased risk of developing breast cancer (153). Today, hormonal therapy for ERα-positive breast tumours comprises antiestrogens and aromatase inhibitors (AIs). Antiestrogens act as competitive inhibitors of ERα to block its transcriptional activity, and subsequently block E2-mediated tumour growth. These drugs can be divided into two subclasses: selective estrogen receptor modulators (SERMs) and selective estrogen receptor downregulators (SERDs). SERMs act by preventing recruitment of coactivator complexes

through inducing a conformational change in ERα that instead leads to recruitment of corepressors (37). Tamoxifen is an example of a SERM that is now considered the gold standard of care for antiestrogen treatment in ERα-positive breast tumours. While this drug acts as an ERα antagonist in the breast, it can have agonistic effects in the endometrium and the bone. Even though this aids in prevention of osteoporosis, it also increases the risk of developing endometrial cancer. Regrettably, 50% of patients do not respond to tamoxifen treatment due to intrinsic (de novo) resistance (154). In addition, 40% of patients who initially respond to tamoxifen eventually develop resistance (acquired) and relapse (155). The majority of these patients still express ERα and respond to SERDs and AIs in second-line treatment. Mechanisms of tamoxifen resistance (reviewed in (156)) include defects in its metabolism due to single nucleotide polymorphisms of cytochrome P450 2D6 (CYP2D6), loss of ERα expression, increased phosphorylation of the receptor due to activation of MAPK and PKA signalling, altered expression of ERα coregulator proteins, increased HER2 signalling, and presence of gain-of-function mutations in the LBD of ERα which result in ligand-independent activation of the receptor (157). ICI 182,780 (Fulvestrant), a SERD that binds ERα with much higher affinity than tamoxifen, induces sumoylation of the receptor, which correlates with its increased degradation and inhibition of its transcriptional activity (Traboulsi T. and Mader S., *manuscript under review* and (158)). Lastly, AIs act by inhibiting the enzyme aromatase, which catalyses the conversion of androstenedione to estrone and testosterone to estradiol. AIs are primarily administered in post-menopausal women in combination with tamoxifen (159).

3.1.1.2 Genetic defects in luminal breast tumours

The seminal TCGA network study published in 2012 revealed a multitude of mutations affecting luminal tumours (147). The alpha catalytic subunit of PI3K (*PIK3CA*) was found as the most frequently mutated gene in luminal tumours affecting 45% and 29% of luminal A and B tumours, respectively. Additionally, mutations targeting *MAP3K1* and *MAP2K4* have also been observed in luminal A tumours where they are more prevalent than in luminal B tumours. *TP53* mutations are also present in luminal tumours with a higher rate in luminal B (29%) than A (12%) tumours and less frequently than in HER2 and basal-like tumours.

As mentioned earlier, loss of ERα expression can lead to endocrine resistance, however the vast majority of distant metastases that are resistant to hormonal treatment still express an active ERα. The updated TCGA dataset of primary breast tumours (817 samples)

reveals that *ESR1* mutations are present in only 0.6% of the cases (148, 160, 161), however recent studies have reported a higher frequency of *ESR1* mutations in metastatic ERα-positive breast cancers (162–166), that affect the LBD of ERα (39 samples, 21%) (reviewed in (157)). These mutations, the majority of which affect tyrosine 537 and aspartate 538, confer ligand-independent activation of the receptor and increased proliferation. On the other hand, FOXA1 mutations were also rare affecting only 3.7% of all breast cancers in the TCGA dataset. Most of these mutations were missense mutations, and although many of them clustered immediately downstream of the forkhead domain, their impact on FOXA1 protein function remains unclear. This is important as previous studies have shown a role for FOXA1 in mediating the response to tamoxifen (167, 168). Recently, FOXA1 gene amplification/overexpression was shown to contribute to tamoxifen insensitivity of endocrine-resistant breast cancer cells in a mechanism that involved interleukin-8 induction (169). In addition, a G>A mutational hotspot was discovered in the promoter region of *FOXA1* in ERα-positive breast cancers at position -81 relative to the TSS, and it results in increased binding of E2F1 leading to FOXA1 overexpression (170).

Interestingly, *GATA3* is the third most mutated gene in luminal breast cancers with a prevalence of approximately 14% and 15% in luminal A and B tumours, respectively, in the TCGA dataset. The *GATA3* gene consists of 6 exons located at 10p14 and encodes for a 48kDa protein that is composed of two N-terminal transactivation domains and two zinc-finger DNA-binding motifs. Each of the two zinc fingers is followed by highly conserved and distinct basic regions (basic region 1 and 2) that have less defined functions. The reported mutations, which cluster in (i) basic region 1 (intron 4: 2 base-pair deletion hotspot), (ii) the second zinc finger motif and its subsequent basic region 2 and (iii) the C-terminal region (147, 171) **(Figure 7),** are predicted to result in a loss of function. The first zinc finger is targeted by a missense mutation converting methionine 293 to a lysine or an arginine. Two residues in basic region 1 are targeted: R305Q (missense mutation) and X308 (splice site mutation: residue 308 marks beginning of exon 5). Interestingly, mutating arginine at position 305 to alanine results in a dominant negative GATA3 protein as shown in reporter assays driven by the T-cell receptor alpha and beta enhancers (172). Mutations targeting the second zinc finger are expected to abrogate binding to DNA. For instance, the hydrophobic Leucine 327, is mutated to a histidine and has been suggested to play a role in maintaining the stability of the DNA/protein complex as it makes direct hydrogen bonds in the major groove (173). Also, R330, which makes interactions with the sugar phosphate backbone of DNA (173) and is

essential for the nuclear localization of the protein (174), is mutated to Glutamic acid via a frameshift mutation. In addition, one study recently reported increased protein stability and reduced DNA-binding of a truncated form of GATA3 that arises from a heterozygous frameshift mutation (D336fs) in the second zinc finger in MCF-7 cells (175). Of note, this mutation has not been reported in primary tumour samples. Deletion studies indicate that the C-terminal region, encompassing basic region 2, is required for the protein's transactivation capacity in *in vitro* CAT assays (172). Additionally, the study of the crystal structure of the C-terminal zinc finger of mouse GATA3 has identified several residues that are important for stabilizing the protein's binding to DNA. Several residues within basic region 2 are subject to mutation in luminal tumours. This region contains part of the highly conserved YXKXHXXXRP motif (aa 344-354; YYKLHNINRPL in humans and mice), where mutation of any amino acid has been reported to disrupt binding to DNA and abrogate GATA3-controlled Th2 cytokine production and cell differentiation (176). Several residues within this motif are also found mutated: Y344, K346, I350, X350_splice, N351, L354. Importantly, the NRPL residues within this motif can act as a homodimerization interface, with their mutation being reported to cause a decrease in GATA1 self-association (177). Arginine 364, which plays a key role in binding the minor groove of DNA, is mutated in breast cancer (R364S) and its disruption has been reported to prevent GATA3 binding to its consensus WGATAR sequence (173). From the clinical standpoint, GATA3 mutations have been associated with improved overall survival (178) and response to AI therapy (179), although their actual impact on GATA3 function remains unclear due the scarcity of functional assays. Thus, additional studies are needed to examine what contribution GATA3 mutants may have in luminal breast tumourigenesis.

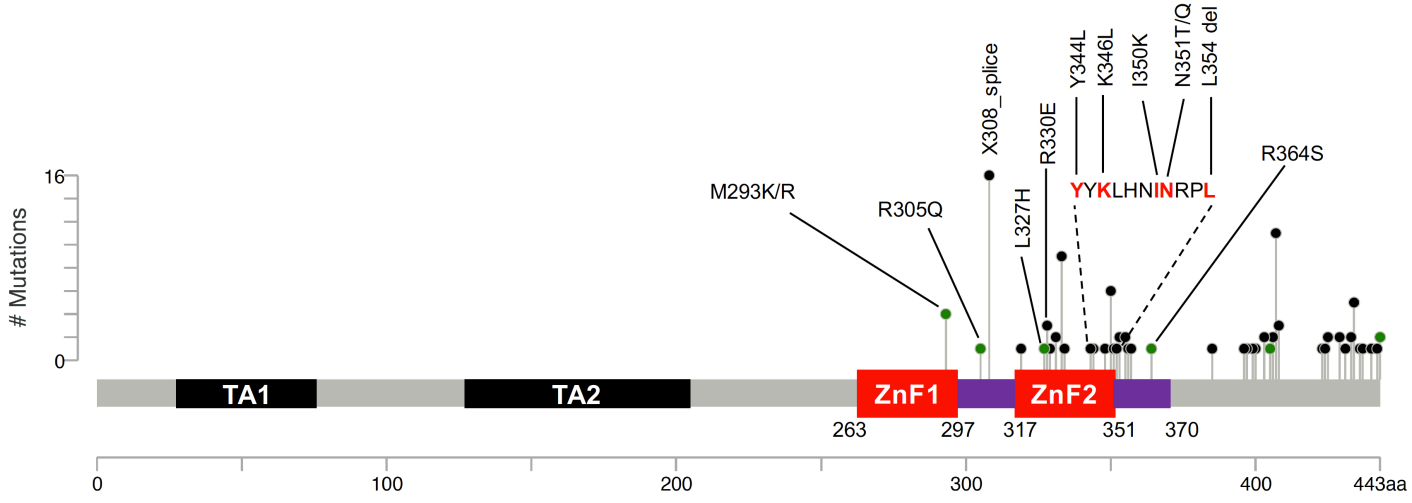


Figure 7. Lollipop representation of GATA3 mutations in breast tumours from the TCGA dataset.

3.1.2 HER2-enriched tumours

HER2-enriched tumours represent about 15-20% of all breast tumours (180). These tumours are characterized by HER2 overexpression most often due to amplification of the HER2 amplicon (80% of the cases) located at 17q12 (147). HER2 amplification correlates with worse prognosis and shorter disease-free and overall survival time (181), due to high tumour grade, increased cell proliferation, lymph node positivity, and large tumour size (182). HER2 encodes a receptor tyrosine kinase which dimerizes with other EGF receptors, leading to auto and/or transphosphorylation of tyrosine residues in EGFR proteins and constitutive downstream activation of MAPK and PI3K signalling. HER2-enriched tumours have a high frequency of *TP53* (72%) and *PIK3CA* (39%) mutation and *MDM2* (30%) and *CCND1* (38%) gene amplification (147).

3.1.2.1 HER2-targetted therapy

HER2-enriched tumours are treated with Trastuzumab (Herceptin®), a humanized monoclonal antibody, in combination with adjuvant chemotherapy. Trastuzumab recognizes the extracellular domain IV of HER2 (183) and blocks its activation by causing receptor internalization and increased degradation through the c-Cbl ubiquitin ligase (184), activation of natural killer cells and antibody-dependent cell-mediated cytotoxicity (185), and inhibition of RAS/MAPK and PI3K/AKT activation by interfering with HER2/HER3 ligand-independent dimerization (186).

3.1.3 Basal-like tumours

Basal-like tumours represent the most aggressive and least differentiated subtype of breast cancer with low overall survival rates. Tumours of this subtype have high incidence in *BRCA1* (DNA repair gene) germline mutation carriers (141, 187, 188) and premenopausal women of African-American ancestry (189). Furthermore, they are distinguished by their high tumour grade, often lack of expression of ERα, PR or HER2, and high expression of basal (myoepithelial) markers such as cytokeratin 5, 6, 14, 17 as well as EGFR. Of note, while the majority of basal-like breast cancers are clinically defined as ERα- PR- HER2- or triple-negative based on the traditional IHC classification, some of the tumours in this subtype do

not conform to the aforementioned IHC definition (reviewed in (190)). Despite the small number of patients in their study (82 patients), Rouzier et al. reported that about 14% of basal-like breast cancers are HER2+ and a smaller proportion (5%) is ERα-positive (191). Similarly, not all triple-negative tumours harbour the basal-like molecular signature. Amongst all breast cancer subtypes, basal-like tumours have the highest frequency of *TP53* mutations (80%), most of which are nonsense and frame shift mutations (147). In addition, they suffer from loss of retinoblastoma (Rb) in 30% of the cases (147). Together with the high frequency of *BRCA1* and *2* somatic mutations (20%) (147), these genetic defects culminate in genomic instabilities and increased proliferation (192, 193).

Recent studies indicate that basal-like tumours with *BRCA1* mutations may arise from a luminal progenitor cell population. Lim et al. analyzed mammary epithelial cell populations in preneoplastic mammary tissue from patients heterozygous for a *BRCA1* mutation (194). They discovered that luminal progenitor cells were expanded in *BRCA1* mutation carriers and that these cells exhibited enhanced clonogenic activity *in vitro*. Importantly, comparative gene expression profiling revealed that the gene signature of luminal progenitor cells was most correlated with signatures of preneoplastic *BRCA1* mutation carriers and basal-like tumours. These findings were corroborated in breast cancer mouse models where luminal progenitor cells heterozygous for a *BRCA1* mutation showed increased basal differentiation features prior to formation of basal-like tumours (188) without evidence of implication of basal progenitor cells in this process (195). Regrettably, basal-like tumours do not benefit from targeted therapy and are treated with chemotherapy. Additional studies are needed to fully understand the molecular mechanisms that underlie tumourigenesis in this subtype.

3.1.4 Claudin-low tumours

Claudin-low tumours were identified in 2007 as a new subtype of breast cancer that is marked by low expression of tight junction-associated proteins claudin 3, 4, 7 and occludin (144). In addition, these tumours are poorly differentiated, express low levels of E-cadherin and are characterized by a signature that has epithelial-mesenchymal transition (EMT) features (high expression of TWIST1 and 2, SNAI1 and 2, ZEB1 and 2) and closely aligns with the signature of MaSCs (196, 197). Clinically, most of these tumours are triple-negative and express low levels of luminal genes and, surprisingly, proliferation genes, suggesting that unlike basal-like tumours, claudin-low tumours are slow-cycling (144). Claudin-low tumours are treated with chemotherapy but show poor survival outcomes owing to their

mesenchymal/stem cell-like features that render these tumours highly intrinsically resistant (198).

3.1.5 Molecular apocrine tumours

In 2005, molecular apocrine tumours emerged as a new subtype of breast cancer representing about 8-14% of all cases (145). While tumours of this subtype harbour some apocrine features present in androgen-dependent apocrine scent glands, they are not classified as apocrine carcinomas as they don't meet the strict histopathological criteria. In addition to being ERα-negative, these tumours are often HER2+ and AR+, and harbour increased AR signalling, although there is a discord between mRNA and protein levels of AR in a subset of these tumours. Lehmann-Che et al. found that 43% of molecular apocrine tumours were AR- by IHC despite having high mRNA expression (199). Overexpression of gross cystic disease fluid protein 15 (GCDFP15), the product of the AR target gene *PIP* (200), combined with HER2-positivity and ERα-negativity was more accurate in identifying tumours of this subtype (199).

Gene expression profiling of molecular apocrine tumours indicates that they share some features of luminal breast tumours, such as expression of luminal lineage genes as well as ERα targets despite absence of the receptor (201). It was hypothesized that increased AR signalling may recapitulate the luminal features of these tumours. Interestingly, FOXA1 is overexpressed in 90% of molecular apocrine breast tumours (199) (as determined by IHC), and has been shown to play an important role in controlling AR signalling in the 'molecular apocrine' AR-positive breast cancer cell line MDA-MB-453 (202), similar to its role in prostate cancer cells (203). FOXA1 binding significantly matched that of AR across the genome and siRNA-mediated knockdown of FOXA1 in these cells resulted in reduced global AR binding without a change in expression of the receptor, and modulated expression of molecular apocrine genes (202). siRNA-mediated knockdown of AR, on the other hand, resulted in moderate reduction of mRNA expression of a number of luminal tumour markers such as *SPDEF* and *XBP1*. Furthermore, reduced proliferation of FOXA1-depleted MDA-MB-453 cells correlated with a growth inhibitory effect of the anti-androgen bicalutamide in these cells (202). The authors' conclusion that FOXA1 regulates the molecular apocrine phenotype through regulating AR signalling is not strongly supported, particularly as the authors do not address the impact of AR knockdown globally on expression of 'molecular apocrine' signature genes. Additional studies and cell line models are needed to fully understand transcriptional

regulation in AR+ and AR- molecular apocrine tumours and whether FOXA1 can control the phenotype of these cells independently of AR.

3.2 Luminal transcription factors in breast tumours

3.2.1 The ERα-FOXA1-GATA3 TF network: A tale of three musketeers

Large genomic and transcriptomic studies over the past two decades have revealed the great complexity of transcriptional regulation in breast cancer. As mentioned earlier, analysis of large transcriptomic data using gene correlations has allowed the identification of several subtypes of breast cancer, each marked by its own signature. These studies together with ChIP-Seq experiments characterizing ERα binding across the genome have been monumental in increasing our understanding of how the receptor works in luminal breast cancer cells. In luminal breast tumours, expression of FOXA1 and GATA3 is highly correlated with that of ERα. This prompted several studies to ask the question whether a cross-regulatory network exists between these factors and what impact this might have on the differentiation state of luminal tumours. Findings of these studies are summarized below.

3.2.1.1 ERα, FOXA1 and GATA3 exist in a transcriptional cross-regulatory network

Remarkably, the strong positive correlation between ERα, FOXA1 and GATA3 expression in luminal breast tumours is associated with favourable prognosis (204–207). This is in agreement with data from preclinical models where both FOXA1 and GATA3 regulate the differentiation state of luminal tumours. Conversely, these three TFs are not expressed in basal-like tumours, which are poorly differentiated. In the MMTV-PyMT breast cancer (luminal-like) mouse model, deletion of *GATA3* accelerated progression from adenoma to carcinoma and increased frequency of tumour-initiating cells whereas its reintroduction inhibited tumour dissemination potential and induced differentiation (208, 209). Furthermore, ectopic expression of *GATA3* in MDA-MB-231 basal-like cells inhibited metastatic potential and promoted their differentiation in mouse xenografts in a mechanism that was dependent on induction of miR-29b (210). On the other hand, FOXA1 and ERα expression was lost during mammary tumour progression in the MMTV-PyMT mouse model (211). siRNA-mediated knockdown of FOXA1 in luminal ERα-positive human breast cancer cell lines, MCF-7 and T-47D, reduced expression of luminal markers and enhanced expression of basal-like markers, which resulted in a less-differentiated phenotype that was accompanied by

increased cell migration and invasion (212). A recent study corroborated these findings where it was found that exogenous expression of TWIST1, a TF that regulates the process of EMT, in MCF-7 cells, directly inhibited FOXA1 expression, increasing invasion, migration and metastasis (213). These studies thus reveal that expression of ERα, FOXA1 and GATA3 in luminal breast tumours is associated with a well-differentiated phenotype.

Mouse knockout studies reveal the presence of a complex cross-regulatory transcriptional network composed of ERα, FOXA1 and GATA3, consistent with the degree of similitude in mammary gland phenotypes of their respective knockouts. As mentioned above, all three TFs are co-expressed in normal luminal mammary epithelial cells. This raises the question whether these three TFs exist in a cross-regulatory network in mammary epithelial cells. While it was found that GATA3 can bind a site 0.8kb upstream of the first exon of *FOXA1* in mouse primary mammary epithelial cells (108), GATA3 KO in the mammary gland had no impact on FOXA1 expression (95), but did reduce ERα expression in luminal ductal cells (109). ERα KO mice exhibited no change in expression of FOXA1 or GATA3, indicating ERα is downstream of the other two TFs (95). *FOXA1* deletion in the mammary gland resulted in loss of ERα expression and its downstream target, *PR*, without a significant change in *GATA3* mRNA expression (95). These results place both FOXA1 and GATA3 upstream of ERα in the normal mammary epithelium. They also indicate that GATA3 by itself is not sufficient for ERα expression, as FOXA1 null glands still lost expression of the receptor despite maintained expression of GATA3.

Similar results have been reported in breast cancer cells, although regrettably, original data from these studies are sometimes incorrectly interpreted by authors of the work and also by subsequent studies that cite this work. Eeckhoute et al. first reported that siRNA-mediated knockdown of GATA3 in MCF-7 and T-47D cells reduced ERα mRNA and protein levels, which correlated with GATA3 and p300 binding at two enhancer sites upstream of promoters E and F of *ESR1* (214) along with increased recruitment of the H3K9 di/tri-methyl demethylase KDM4B, and addition and removal of H3K9/K18Ac and H3K9me2/3 marks, respectively (215). Notably, the Eeckhoute et al. study additionally suggested that ERα regulates expression of GATA3 based on (i) E2-induced recruitment of ERα to an enhancer site downstream of *GATA3* and (ii) induction of *GATA3* mRNA following E2 treatment, albeit very weakly (1.5 folds). Although mere TF binding may not necessarily merit transcriptional regulation, several subsequent reports have cited the Eeckhoute et al. study incorrectly indicating that ERα positively regulates expression of GATA3. The role of FOXA1 in

regulating ERα expression was examined by two separate groups. Hurtado et al. found that siRNA-mediated knockdown (48 h) of FOXA1 in MCF-7 cells cultured in hormone-depleted medium does not alter ERα expression (167). However, two subsequent reports contradict this finding. Caizzi et al. observed a reduction in ERα protein levels following siRNA-mediated knockdown (48 h) of FOXA1 in MCF-7 cells cultured in hormone-depleted medium (216). On the other hand, Bernardo et al. reported that FOXA1 is recruited to a site at the proximal promoter of *ESR1* and its knockdown (36 h) in MCF-7 cells in hormone-replete medium results in reduction of ERα expression without a change in expression of GATA3 (95). Of note, the presence of steroidal hormones and phenol-red (which acts as a weak estrogen (217)) in culture media can interfere with estrogen-regulation experiments. To study estrogenic regulation, it is imperative to use a culture medium that is phenol-red free and deplete serum of steroidal hormones via charcoal-dextran treatment. Under hormone-replete conditions, estrogens can enhance ERα protein turnover (218). Thus, different cell culture conditions may explain the different findings reported by these studies.

Contradictory studies have reported that FOXA1 expression is both induced (219) and repressed (220–222) by E2 treatment. E2 treatment represses expression of *ESR1* and *GATA3* (222). These results indicate that both FOXA1 and GATA3 are upstream of ERα in breast cancer cells and additionally seem to act in parallel to regulate the receptor's expression in normal mammary epithelial cells. Repressed expression by E2 treatment may indicate negative feedback through ERα, but this remains to be confirmed through siRNA-mediated knockdown of the receptor in the presence of E2. Additional studies are also needed to examine whether there is crosstalk between FOXA1 and GATA3 in breast cancer cells.

Lastly, the ERα-FOXA1-GATA3 network can influence expression of two other luminal-lineage TFs: SPDEF and XBP1. *XBP1* is an E2-target gene in breast cancer (223) and knockdown of FOXA1 can result in its downregulation (212). SPDEF expression is activated by FOXA1 and ERα (118, 212) and repressed by GATA3 (118). Notably, SPDEF knockdown has no impact on mRNA levels of *ESR1, FOXA1,* and *GATA3* (118).

3.2.1.2 FOXA1 and GATA3 regulate ERα binding to its target genes

The complexity of this network is further increased by the fact that these three TFs can influence each other's function. GATA3 was originally found to bind to several GATA sites flanking an ARE in the upstream enhancer of the AR target gene *KLK3* in LNCaP prostate

cancer cells (224). Mutation of these GATA sites reduced androgen-mediated induction of a luciferase reporter under the control of the *KLK3* promoter, a finding that indicated potential cooperativity between GATA3 and AR in regulating its target genes (224). The first evidence that GATA3 may be implicated in regulation of E2 target genes was the finding that GATA6 binds several GATA sites in the promoter of *TFF1*, and enhances transcription of a luciferase reporter under the control of the *TFF1* promoter in gastric and intestinal cell lines (225). On the other hand, FOXA1 but not ERα microinjection into *Xenopus Laevis* oocytes lead to chromatin opening at the estrogen-responsive vitellogenin B1 gene promoter, suggesting that this might facilitate binding of ERα (226). In addition, mutation of a FOXA1 binding site immediately downstream of an ERE in the CYP3A4 proximal promoter reduced E2-mediated transactivation of this gene's promoter in a reporter assay (227). FOXA1 was also shown to bind a FOXRE immediately upstream of the *TFF1* TATA box and transactivate a *TFF1* reporter plasmid (228). Interestingly, the FOXRE and GATA elements in the *TFF1* proximal promoter were within a 350 bp window downstream of an imperfect ERE (229) that is necessary for E2-mediated induction of *TFF1* (230). Together, these studies based on single genes lead to the hypothesis that FOXA1 and GATA3 may cooperate with ERα to regulate its downstream targets. In 2005, using ChIP-on-chip two separate groups showed that ERα-binding sites in MCF-7 cells are enriched in FOXREs and that knockdown of FOXA1 reduces E2-mediated recruitment of ERα to its target genes, which correlated with their reduced expression (219, 231). This result was further corroborated by another study where *CCND1* gene regulation by E2 was abrogated following knockdown of FOXA1 in a mechanism involving decreased binding of ERα to an upstream enhancer (232). However, these studies did not report on whether FOXA1 knockdown altered ERα protein levels, which could explain decreased recruitment of ERα and subsequent transcriptional activation. Importantly though, the majority of the ERα mapped sites reported by Carroll et al. were located at distal enhancers, suggesting that ERα might primarily engage in long-range chromatin interactions for transcription of its target genes (231). Indeed, our lab previously showed through 3C experiments that *GREB1*, an E2 target gene, is regulated by ERα binding at three EREs within 20kb from the TSS through chromatin loops (233). In addition, through the ChIP-PET technique, the majority of ERα-bound *cis*-regulatory elements were found to interact with promoters of downstream E2 target genes through chromatin loops (234). Because the previous studies were limited to either studying ERα binding on proximal promoter regions in a genome-wide manner (219) or on specific chromosomes (231), genome-wide ERα ChIP-

on-chip experiments were performed without a limit on the binding site mapping window and confirmed enrichment of FOXREs as well as EREs and identified additional TF motifs including those of Oct, AP1, and C/EBP enriched at ERα binding sites (235). GATA motifs were also found enriched at sites where ERα is bound (236). Consistent with the TF binding site (TFBS) predictions, FOXA1 ChIP-on-chip experiments in MCF-7 cells revealed that 50-60% of ERα binding sites overlap with a FOXA1 binding site (237). Enrichment of E2 target genes near those overlapping sites (within 20kb) was also observed. This degree of overlap was also observed for AR in LNCap cells, indicating that FOXA1 may be implicated in regulating the function of other hormone receptors as well. Binding of FOXA1 occurred predominantly at distal enhancers in a cell-type specific manner and correlated with the presence of active enhancer marks H3K4me1 and 2 and DNA hypomethylation (237, 238). Importantly, reduction of the aforementioned histone marks by overexpression of the histone demethylase KDM1 resulted in decreased FOXA1 recruitment, whereas FOXA1 depletion had no impact (237). Although these results indicate that FOXA1 may be recruited to sites already marked by H3K4me1 and 2, it remains to be determined whether this holds true for all FOXA1 binding sites across the genome, as the KDM1 overexpression experiment was performed on only a select number of FOXA1 sites. One important question arises from these findings: does FOXA1 binding correlate with chromatin opening? Using Formaldehyde Assisted Isolation of Regulatory Elements (FAIRE), which can identify 'open chromatin' regions (239), Eeckhoute et al. found that a significant proportion of FOXA1 binding sites had surprisingly low FAIRE signals (240). Interestingly, while these sites did not correlate with gene expression in a given cell line, a subset of them were found active in another cell type indicating that FOXA1 may require additional signals to elicit chromatin opening and transcriptional regulation. Indeed, 22% of low FAIRE FOXA1-bound sites in MCF-7 cells had a high FAIRE signal while being bound by FOXA1 in LNCaP cells (240). These sites were also more likely to bind AR in LNCAP cells. Conversely, 17% of low FAIRE FOXA1-bound sites in LNCaP cells had a high FAIRE signal while being bound by FOXA1 in MCF-7 cells and those sites were more likely to also be bound by ERα (240). These data are consistent with the paradigm that differences in TF activity/expression dictate cell-type specific features.

Subsequent ChIP-Seq experiments have revealed that knockdown of FOXA1 reduced signal intensity of 90% of ERα binding events by at least 50% and abrogated regulation of 95% of E2 target genes, without a change in protein levels of the receptor (167). A proportion of ERα binding sites also overlapped with binding sites of the coactivators SRC-1 (12%),

SRC-2 (41%) and SRC-3 (37%) that were induced in an E2-dependent manner (241). Unlike the SRC proteins, p300 and CBP were present at ERα binding sites prior to E2 stimulation. However, the proportion of ERα binding sites that overlapped with P300 (51%) and CBP (45%) binding was markedly increased by E2 treatment (241). GATA3 was also shown to bind in concert with ERα. A set of the most highly active ERα enhancers defined by H3K27Ac marks and increased eRNA transcription were also found bound by GATA3 in an E2-dependent manner (242, 243). However, to address the importance of global GATA3 binding at ERα-bound sites, Theodorou et al. depleted GATA3 using siRNAs and showed that this can cause redistribution of about one third of all ERα binding events, half of which exhibited increased or decreased binding intensity (244). Interestingly, sites that exhibited stronger ERα signal intensity following GATA3 depletion, also exhibited stronger FOXA1 binding intensity and a gain in p300 recruitment and active histone marks. Conversely, sites on which GATA3 knockdown caused weaker ERα binding, exhibited reduced FOXA1, P300 and active histone mark recruitment. These data indicated that not only can GATA3 influence ERα cofactor binding, but also that of FOXA1 to a certain extent, although the mechanism by which this happens remains unclear (244).

Together, these data present a model whereby both FOXA1 and GATA3 are thought to act as 'pioneer factors' that bind and open chromatin in a first step and allow for the binding of ERα and other cofactors like the SRC proteins, CBP and p300, resulting in transcriptional regulation of downstream targets. These sites bound by a multitude of TFs have been recently referred to as MegaTrans enhancers, i.e super-enhancers bound by 1-2 MDa TF complexes acting in *trans* (242).

Whereas Hurtado et al. showed that knockdown of ERα or ICI 182,780 treatment had no impact on FOXA1 binding on a select number of sites and in a genome-wide manner, respectively (167), recent studies contradict this finding. Kong et al. showed that 30% of ERα-bound sites induced following E2 treatment were also bound by FOXA1 and GATA3 (245). Importantly, E2 treatment revealed new ERα and FOXA1 co-bound sites that were solely bound by ERα in the vehicle control, suggesting that liganded ERα may contribute to recruitment of FOXA1 at these sites. This was in agreement with another report that identified enrichment of FOXREs at sites bound by unliganded ERα and found that shRNA-mediated knockdown of ERα reduced FOXA1 binding on a number of these sites (216). These findings suggest that the current 'FOXA1 pioneers ERα recruitment' model is subject to change and that both ERα and GATA3 may act as 'pioneer factors' for FOXA1 in breast cancer. Additional

studies are needed to identify the mechanism(s) by which ERα, FOXA1 and GATA3 cooperate to regulate transcription in luminal tumours.

3.3 Other upstream regulators of ERα expression

3.3.1 FOXC1, FOXO3A and FOXM1

Several other members of the FOX family of proteins have been implicated in regulation of ERα expression. High FOXC1 expression, which demarcates basal-like tumours, is associated with poor prognosis. In addition, overexpression of this TF in MDA-MB-231 and MCF-7 breast cancer cells causes increased proliferation, migration, and invasion (246). A recent study has shown that FOXC1 overexpression in luminal MCF-7, T-47D and ZR-75-1 cells reduces expression and the transcriptional activity of ERα in a mechanism that involves reduced binding of GATA3 and KDM4B at promoter C and an enhancer element just upstream of promoter D in the *ESR1* gene (247). Another TF, FOXO3A, has been shown to positively regulate expression of ERα in ERα-positive MCF-7 cells and NF639 breast cancer cells by binding to two forkhead elements in promoter B of the *ESR1* gene (248). Interestingly, activation of the PI3K/Akt pathway increases phosphorylation of FOXO3A and inhibits its activity by increasing its association with the cytoplasmic scaffold protein 14-3-3ζ which prevents it from travelling to the nucleus (249). Consistent with this finding, transfection of a constitutively active Akt reduced ERα expression (248). In addition, overexpression of HER2 or EGFR or a constitutively active RAF kinase in MCF-7 cells resulted in downregulation of ERα, which could be reversed by concomitant inhibition of MEK or co-transfection of dominant-negative ERK1/2 (250). Conversely, inhibition of the PI3K/Akt pathway by chemical agents reduced FOXO3A phosphorylation, increased its nuclear localization and ERα expression (248). A subsequent study showed that in fact FOXM1 expression, but not that of FOXO3A or its phosphorylated form, positively correlated with expression of ERα in breast cancer cell lines. Through overexpression and knockdown experiments, Madureira et al. showed that FOXM1 positively regulates ERα expression by binding several forkhead elements in the promoter of *ESR1* (251). Interestingly, FOXM1 is a transcriptional target of ERα (252) and plays an important role in controlling expression of G2/M genes of the cell cycle (253). Additionally, 30% of ERα binding sites overlapped with a FOXM1 binding site, consistent with the finding that both of these TFs can physically interact (254).

3.3.2 p53

Regulation of ERα expression and function by p53 has also been described. Overexpression of WT p53 in MCF-7 cells was shown to increase mRNA and protein levels of ERα, whereas conversely, its knockdown reduced expression of the receptor (255). This positive regulation by p53 was not dependent on its DNA-binding capacity but is rather mediated through protein-protein interactions (255). Consistent with this finding, p53 activation by ionizing radiation or doxorubicin treatment increased expression of ERα in a mechanism that involved formation of a p53 complex at the promoter of *ESR1* composed of CARM1, CBP, c-Jun, and Sp1 (256). Furthermore, induction of p53 expression by ionizing radiation or treatment with an agent that inhibits p53's interaction with MDM2 also stabilized ERα protein (257). The positive correlation in expression of both factors was also observed *in vivo* where tumours arising in MMTV-Wnt1 transgenic mice that were heterozygous or homozygous for p53 deletion expressed lower levels of ERα than their WT counterparts (258). On the other hand, p53 expression is induced by E2 treatment in MCF-7 cells in a mechanism that requires the presence of CTF-1 and NFκB-binding motifs in the *TP53* promoter as revealed by deletion mutant analysis (259). Another mechanism by which ERα can induce p53 expression is through induction of c-myc, an E2-target gene, which can bind a c-myc/Max response element in the *TP53* promoter (260). The *TP53* promoter harbours no consensus EREs, however two distal half-EREs have been shown to be important for regulation of p53 by ERα (261). Indeed, knockdown of ERα reduces p53 expression and its downstream targets, p21 and MDM2 (261). Together, these studies indicate that both ERα and p53 exist in a positive cross-regulatory loop. As mentioned earlier, *TP53* is frequently mutated in breast cancer, primarily ERα-negative basal-like tumours. Thus, based on the findings outlined above, it is plausible to speculate that *TP53* mutations in basal-like tumours may contribute to the absence of ERα expression (and potentially other luminal TFs) in these tumours. Additional studies will be needed to address this hypothesis.

3.3.3 NFκB

There is evidence for the interplay of NFκB and ERα (reviewed in (262)). In breast cancer cell lines and tumour samples, there is a negative correlation between activation of NFκB and ERα expression (263–268). Indeed, the NFκB subunit, RelB, has been shown to inhibit ERα expression through induction of Blimp1, a repressor zinc finger protein that binds a GAAA core sequence in promoter B of *ESR1* (269). Additionally, Belguise et al. showed that PKCθ

can activate c-Rel and Akt, which results in increased FOXO3A phosphorylation and reduction of ERα expression, thus providing evidence of indirect interactions between NFκB and the receptor (267). Another mechanism by which NFκB activity may inhibit ERα expression is through potential recruitment of EZH2 to the *ESR1* gene promoter. This is because (i) EZH2 knockdown has been shown to increase ERα expression in MCF-7 cells (270) and (ii) TNFα-mediated activation of NFκB results in inhibition of Notch1 gene expression via increased recruitment of EZH2 and DNMT3B in muscle cells (271). On the other hand, ERα can also inhibit NFκB activity. For instance, introduction of ERα into ERα-negative MDA-MB-231 breast cancer cells inhibited TNFα and TPA-mediated activation of an NF-kB CAT reporter (263). Furthermore, inhibition of ERα using the full antiestrogen ICI 182,780 resulted in enhanced c-Rel DNA binding (267).

3.3.4 BRCA1

BRCA mutations are associated with ERα-negative basal-like tumours. Although little is known about the interactions between BRCA proteins and ERα, it was reported that BRCA1 can engage the *ESR1* promoter together with Oct-1, and both of these factors act as positive regulators of the receptor's expression (272). Indeed, siRNA-mediated knockdown of WT BRCA1 in T-47D and MCF-7 cells resulted in decreased ERα protein levels, whereas overexpression of WT BRCA1 in *BRCA1*-mutant triple-negative HCC1937 breast cancer cells resulted in upregulation of ERα expression (272). In contrast with this finding, another study has reported a negative role of BRCA1, through its ubiquitin ligase activity, in regulating ERα's transcriptional activity by inducing mono-ubiquitination and competing to reduce p300-mediated acetylation (273). Acetylation of ERα positively regulates the receptor's transcriptional activity (274). Indeed, BRCA1 knockdown increased and overexpression decreased ERα acetylation in MCF-7 cells with no impact on protein levels. It remains unclear how various BRCA mutations impact ERα transcriptional activity, but one could speculate that the exact impact would depend on the type of mutation and the effect the mutation has on BRCA protein function and ubiquitin ligase activity. As ERα is known to autoregulate its own expression (275), changes in its activity may determine whether it is expressed or not in BRCA mutant tumours. ERα-positive breast cancers in *BRCA1* mutation carriers have been described, although they are considered 'incidental' and not a direct result of a dysfunctional BRCA protein (276).

3.3.5 Menin

Menin is a nuclear protein that is targeted by inactivating mutations in multiple endocrine neoplasia type 1 (MEN1), which encompasses a group of neoplasias affecting the endocrine system (parathyroid, pituitary, and pancreas). A role of this factor in regulating estrogen signalling has recently been revealed. Menin acts as a coactivator of ERα and is recruited to the *TFF1* promoter in an E2-dependent manner (277). Additionally, as part of the MLL1/MLL2 H3K4 methyltransferase complexes (278, 279), menin contributes to increased H3K4me3 marks at the *TFF1* promoter (277). Although siRNA-mediated knockdown of menin had no impact on ERα expression in this study (277), a more recent study showed that its inhibition either by shRNA or a chemical agent downregulates ERα expression in several ERα-positive breast cancer cell lines (280). Interestingly, menin was found to be associated with enhancers co-bound by ERα, FOXA1 and GATA3 and its depletion resulted in a drastic decrease in FOXA1 binding signal at those sites, suggesting that menin acts upstream of FOXA1. The decrease in ERα expression following menin depletion was not due to a change in FOXA1 expression (280). However, menin has been shown to act as an adaptor factor allowing physical interaction between GATA3 and c-myb to positively regulate expression of *GATA3* in primary human Th2 cells (281). Thus, menin may modulate ERα expression through modulating expression of GATA3, although this requires further investigation. Female patients suffering from MEN1 are at a higher risk of developing breast cancer, suggesting that menin acts as a tumour suppressor in these patients (282). *MEN1* gene mutations are not common in ERα-positive sporadic breast cancers (147) where menin seems to act as an oncogene based on its capacity to drive ERα signalling. Further investigation is needed to determine whether targeting menin to abrogate ERα expression could serve as a useful therapy in ERα-positive breast cancer.

4. Histone deacetylase inhibitors (HDACis) as anticancer agents for breast cancer treatment

4.1 HDACs

Eighteen HDACs have been identified in humans. They are subdivided into four classes based on their homology to yeast HDACs and their subcellular localization (283). Class I HDACs (HDAC1, 2, 3, and 8) are detected in the nucleus and share homology with the yeast RPD3 protein. Class II HDACs (HDAC4, 5, 6, 7, 9 and 10) are found both in the nucleus and cytoplasm and share homology with the yeast Hda1 protein. Class III comprises members of the sirtuin family (SIRT1, 2, 3, 4, 5, 6, and 7), which are homologous to the yeast Sir2 protein. Class IV has just one member, HDAC11, which does not share a sufficient level of conserved residues in its catalytic core domain with class I and II HDACs to be placed in either class (283, 284). One key difference between the sirtuin class of HDACs and classes I, II and IV is that sirtuins require NAD+ for their enzymatic activity, whereas the other three classes are Zn^{2+}-dependent.

In 1964, Allfrey et al. discovered that acetylation can lower histone capacity to inhibit RNA synthesis (285). It was subsequently determined that a positive correlation exists between histone acetylation and transcriptional activation (286). HDAC and HAT activities act in concert to maintain a homeostatic balance between deacetylation and acetylation activities, respectively. On the one hand, histone acetylation can disrupt the electrostatic interactions between the negatively-charged phosphodiester DNA backbone and the positively-charged lysine residues in amino-terminal histone tails, leading to a more relaxed chromatin structure that is permissive to TF binding and subsequent transcriptional regulation (287). On the other hand, histone deacetylation tightens histone-DNA interactions and results in a compact chromatin state that is not amenable to TF binding **(Figure 8)**. The current paradigm follows that in addition to HATs acting as 'writers' and HDACs as 'erasers', bromodomain-containing proteins can act as 'readers' that recognize acetylated peptides. Indeed, bromodomain-containing proteins can modulate transcription through various mechanisms (reviewed in (288)). For instance, they can be part of chromatin remodelling complexes as is the case with SWI/SNF complex core subunits SMARCA2/SMARCA4, which promote chromatin opening or ACF-WSTF complex core subunits BAZ1A/BAZ2A, which promote chromatin compaction. Bromodomain-containing proteins can also mediate histone acetylation, as is the case for the HATs P/CAF, GCN5L2, p300 and CBP or histone

methylation, as seen with ASH1L and MLL. Other bromodomain-containing proteins, such as those belonging to the BET family, can serve as scaffolds that recruit components of the transcriptional machinery or additional TFs. These effector proteins represent a second mode of action by which histone acetylation can regulate gene expression.

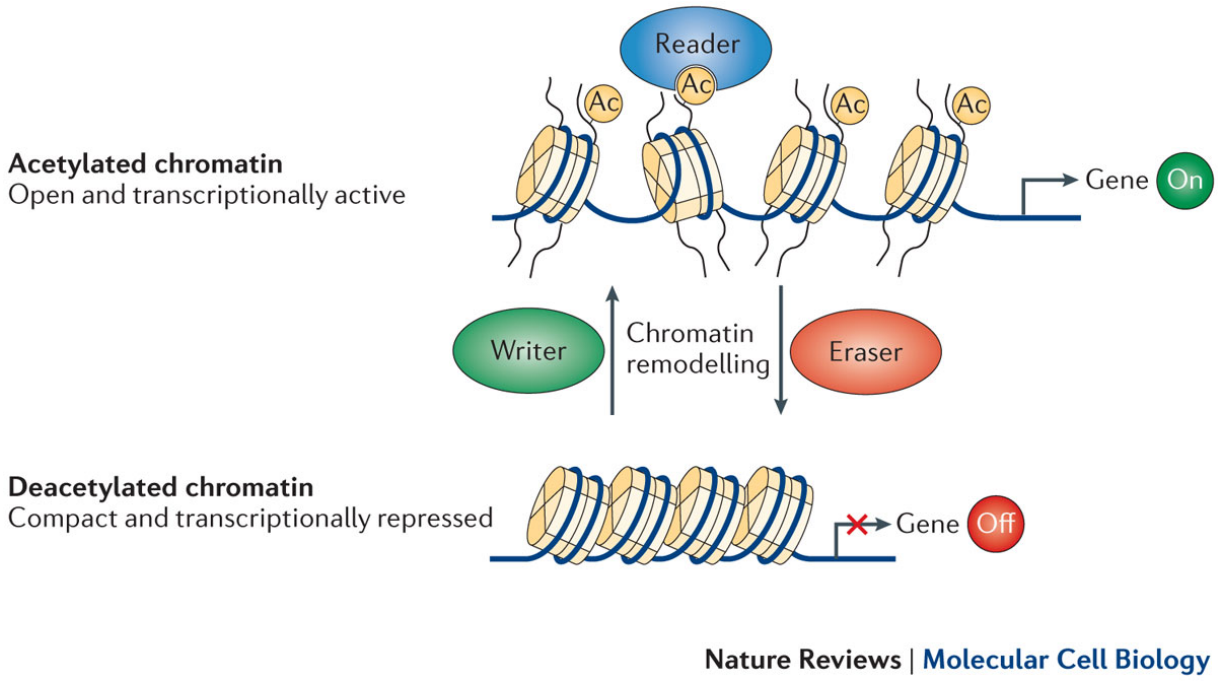


Figure 8. Impact of histone acetylation on gene expression. HATs (writers) catalyze addition of acetyl marks to side chains of lysine residues on N-terminal tails of all four core histone proteins. This results in a more relaxed chromatin state permissive to transcriptional activation. HDACs (erasers) catalyze removal of these marks resulting in a compact chromatin state and transcriptional repression. Bromodomain-containing proteins (readers) can include HATs, methyltransferases, TFs, chromatin remodeling factors, coactivator proteins, all of which can act as effector proteins that propagate the acetylation signal. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology **16,** 258-264, copyright (2015). (Reference (289))

Contrary to what their name implies, HDACs can deacetylate a variety of non-histone substrates, including TFs, chaperones and signal transducer proteins, and this acetylation can drastically impact their function. Acetylation of TFs can enhance their transcriptional activity/binding to DNA (ERα, GATA3, p53, RUNX3, E2F1, SREBP-1, AR) (274, 290–295) or repress it (BCL6, IRF7, FOXA1) (296–298). Acetylation can also increase protein stability (RUNX3, SREBP-1, SMAD7, c-myc, p53) (292, 294, 299–301) or turnover via the proteasome pathway (HIF1α, cyclin A) (302, 303). In some cases, protein-protein interactions can be enhanced (STAT3, Ubc12) (304, 305) or disrupted (HSP90, Ku70, NFκB) (306–308) by acetylation. Given the pleiotropic effects that HDACs have on proteins with diverse

biological functions, it is not surprising that their knockout in mice causes severe defects in proliferation and differentiation. Notably, most of these mice suffer from cardiac diseases, including hypertrophy and fibrosis (309).

4.1.1 HDACs as targets for cancer therapy

Cancer has been traditionally thought of as a disease arising from genetic aberrations in tumour suppressor genes and oncogenes, however there is a growing body of evidence that epigenetic defects can also contribute to the process of tumourigenesis. A role of HDAC proteins in cancer cells was first described in the mid-1970s by the laboratories of Leder and Riggs, who showed that the short-chain fatty acid N-butyrate reverses viral transformation of Friend erythroleukemia cells through increased histone acetylation (310), promoting their differentiation to normoblast-like cells that synthesize haemoglobin (311). It was later determined that N-butyrate is in fact an HDAC inhibitor (HDACi) (312, 313). It was subsequently revealed that HDACs play a role in leukaemogenesis as they are aberrantly recruited together with corepressors to promoters of differentiation genes by oncogenic fusion proteins such as PML/PLZF-RARα, and AML-ETO in acute promyelocytic leukaemia (APL) and acute myeloid leukaemia (AML), respectively (283). In addition to their aberrant recruitment to specific promoters, HDAC expression is also deregulated in several types of solid tumours. For instance, HDAC1 is overexpressed in breast (314), prostate (315), and gastric (316) cancers. HDAC2 is overexpressed in colon (317), gastric (318) and cervical (319) cancers. HDAC6 is also overexpressed in breast cancer (320) and cutaneous and peripheral T-cell lymphoma (CTCL and PTCL) (321).

4.2 HDACis: molecular mechanisms of action

Several HDACis are presently being assessed in clinical trials for treatment of haematological malignancies and solid tumours, where they show antiproliferative properties (322). These inhibitors are either naturally existing, or chemically synthesized, have different specificities and can be divided into four classes based on their chemical structure: short-chain fatty acids, hydroxamates, cyclic tetrapeptides and benzamides **(Figure 9)**. Short-chain fatty acids butyrate and valproic acid are weak HDACis acting at millimolar concentrations. Trichostatin A (TSA) is a hydroxamic acid derivative and was originally discovered as a naturally occurring antibiotic in *streptomyces hygroscopicus* due to its antifungal properties against Trychophyton

(323). It was subsequently shown that it induces differentiation of Friend murine erythroleukemia (324) and has HDACi activity (325). Another hydroxamate, vorinostat (suberoylanilidehydroxamic acid or SAHA), gained FDA approval in 2006 for treatment of CTCL. Romidepsin, a cyclic tetrapeptide and Panobinostat, a hydroxamate, are two other HDACis that have also been approved by the FDA for treatment of CTCL and multiple myeloma, respectively (322). These three drugs are currently being tested for treatment of patients infected with human immunodeficiency virus (HIV) to reactivate latent HIV in residual infected cells following antiretroviral therapy as a strategy to resensitize them for targeting (326). Both TSA and SAHA are pan-specific, targeting classes I, II and IV HDACs (283). A recent study suggests that both TSA and SAHA are more selective towards HDAC1, 2, 3 and 6, as determined by *in vitro* acetylation assays (327). The benzamide class of HDACis include Entinostat (MS-275), which is more selective towards HDAC1 and HDAC3. Entinostat is currently in a phase III clinical trial in combination with exemestane, an aromatase inhibitor, for treatment of recurrent hormone receptor-positive breast cancer.

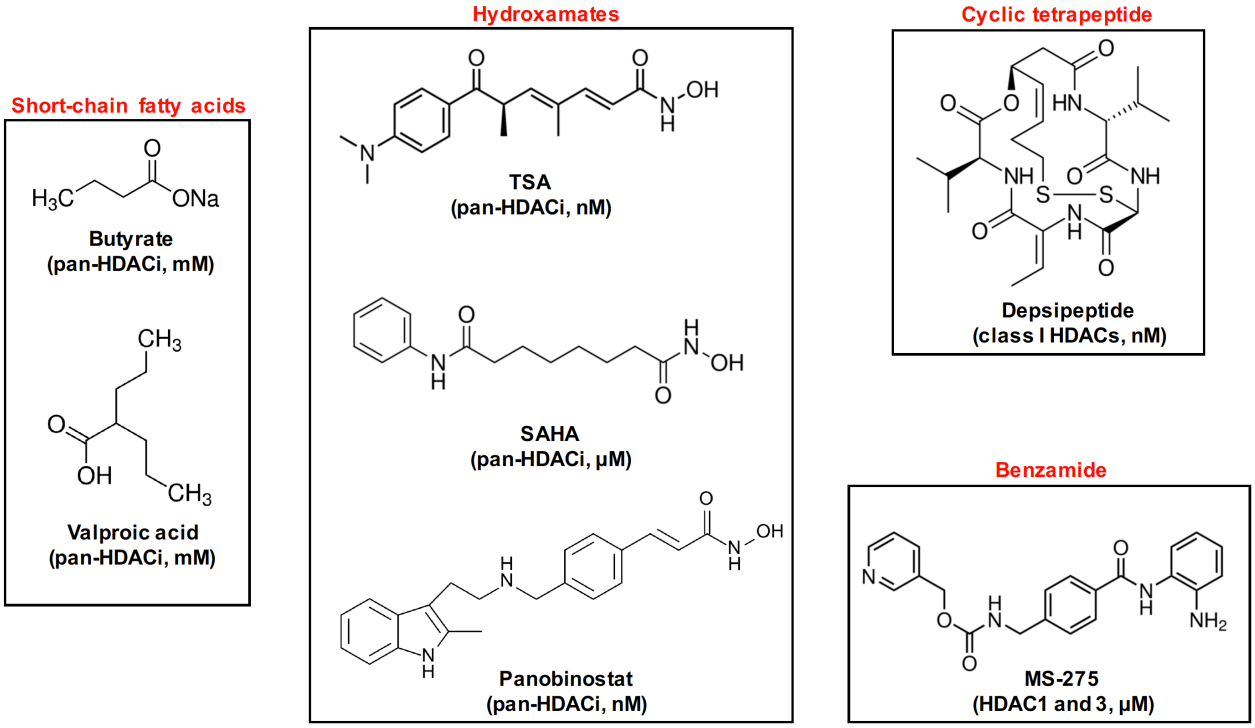


Figure 9. Structure of the four main classes of HDACis, their specificity and their potency (reviewed in (283, 309)).

While HDACi treatment results in global histone hyperacetylation, approximately 50% of HDACi target genes are repressed, owing to the capacity of these drugs to target non-

histone proteins. HDACis have antitumourigenic actions both *in vitro* and *in vivo,* as summarized in **Figure 10** (reviewed in (283)). Consistent with the initial discovery that these drugs act as differentiation agents in haematological malignancies, HDACis can also induce cell differentiation in several solid cancers including endometrial (328), ovarian (329), bone (330), and breast (331) cancers. Additionally, HDACis induce cell cycle arrest at the G1 phase via p53-independent induction of p21 and downregulation of cyclin D1, resulting in Rb hypophosphorylation (332). HDACis have also been shown to repress expression of DNA synthesis proteins thymidylate synthetase and CTP synthase, which correlates with a G1/S cell cycle arrest (333). These inhibitors can also induce expression of extrinsic death receptor (DR5, TRAIL) (334) and/or intrinsic (Bax, Bim, Bak) (335) apoptotic pathway genes, leading to apoptosis. Accumulation of reactive oxygen species (ROS) has also been reported in HDACi-treated cells (336). This was associated with induction of thioredoxin-binding protein 2 (TBP2), which negatively regulates the ROS scavenger thioredoxin (337). HDACis can also have immunomodulatory effects by inducing expression of co-stimulatory/adhesion molecules such as CD80, CD86 and ICAM1 (338), or MHC class I-related proteins MIC-A and MIC-B (339), which increases immunogenicity of tumours cells. Furthermore, HDACis have been shown to suppress expression of pro-inflammatory cytokines which reduces acute graft-versus-host disease (340).

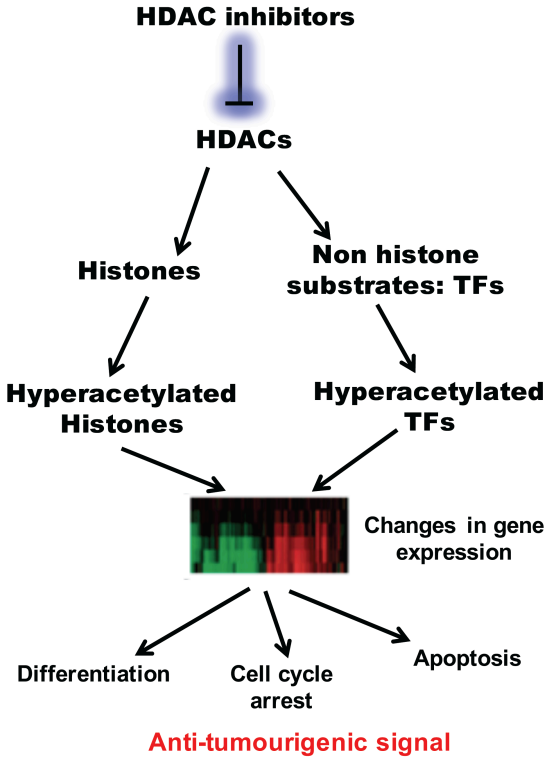


Figure 10. Molecular mechanisms of action of HDACis.

Inhibition of HDACs shifts the balance in favour of HATs, which generally results in global hyperacetylation of histone and non-histone substrates, including TFs. This results in global changes in gene expression profiling with 50% of genes up- or down-regulated. The consequence of HDACi treatment is an anti-tumourigenic signal marked by induction of differentiation, apoptosis, and/or cell cycle arrest.

4.2.1 HDACis in breast cancer

HDACis have been shown to have strong anti-proliferative properties in breast cancer. For instance, treatment of MCF-7 cells with SAHA results in G1/S and G2/M cell cycle phase arrest (331) along with induction of p21 (341) and reduced expression of cyclin D1 (342) and cyclin B (343). Interestingly, ERα-positive cell lines were more sensitive to HDACi-mediated p21 induction than ERα-negative cell lines (341). In fact, re-expression of ERα in these cells increased their sensitivity to growth inhibition by HDACis. Furthermore, HDACis can induce differentiation traits in breast cancer cells. For instance, Davis et al. showed that the HDACi sodium butyrate can induce formation of lipid droplets, a hallmark of lactating cells, in both normal and tumourigenic breast cells irrespective of ERα status (344). This was in agreement with studies by Munster et al. (331) and Travaglini et al. (343) who showed induction of milk fat globule proteins by SAHA and lipid droplets by valproic acid treatment, respectively. The antiproliferative actions of HDACis have also been reported *in vivo*. SAHA was shown to inhibit brain metastases in a preclinical mouse model of triple-negative breast cancer (345). Additionally, Hirokawa et al. showed that romidepsin treatment inhibited growth of both

estrogen-responsive and tamoxifen-resistant MCF-7 cell xenografts by inhibiting the kinase activity of PAK1, a serine/threonine-protein kinase (346). TSA was also shown to inhibit growth of MCF-7 cell xenografts (347) and ERα-negative MDA-MB-435 cells in combination with the DNA methyltransferase 1 (DNMT1) inhibitor 5-aza-2'-deoxycytidine (348).

4.2.1.1 HDACis repress ERα expression

Repression of ERα expression by HDACi treatment has been widely described by our lab and others and takes place at the transcriptional and post-transcriptional levels. We have previously reported that treatment of ERα-positive endometrial Ishikawa and MCF-7 breast cancer cells with TSA inhibited transcription from *ESR1* promoters A, B, C, and F, leading to a strong decrease in both mRNA and protein levels of the receptor (349). Counterintuitively, hypoacetylation of histones H3 and H4 was observed at these promoters, suggesting loss of activities that recruit HAT factors to the *ESR1* promoter and/or induction of activities that deacetylate histones (349). Transcriptional repression of ERα was not dependent on *de novo* protein synthesis as co-treatment with cycloheximide did not alter regulation by HDACis (342, 349). Reid et al. found that valproic acid treatment induced recruitment of methyl cytosine binding protein 2 (MeCP2) to the *ESR1* promoter, which might indicate that increased DNA methylation could be a potential mechanism of transcriptional repression of the receptor (350). Results observed with chemical inhibition of HDACs were corroborated by siRNA-mediated knockdown of HDAC1, 2 or 6, which also resulted in downregulation of ERα expression (351). Interestingly, inhibition of the class III sirtuin enzymes by sirtinol (350) or genetic deletion of *SIRT1* in mice (352) also inhibited expression of ERα. A reduction in mRNA stability of the receptor has also been shown to contribute to inhibition of its expression. Pryzbylkowski et al. showed that combined treatment with TSA and 5-aza-2'-deoxycytidine altered the subcellular localization of the RNA binding protein HuR, resulting in reduced ERα mRNA stability (353). Furthermore, transcriptional repression by TSA required the 3'UTR region of ERα, suggesting the possibility of regulation by miRNAs. Indeed, several miRNAs have been linked to regulation of ERα expression (reviewed in (354)), although regulation of their expression by HDACis remains unclear. Lastly, HDACis have been shown to increase ERα protein turnover in a mechanism involving HSP90 hyperacetylation and disruption of its function (306), causing dissociation from ERα and resulting in degradation of the receptor by the ubiquitin ligase C-terminal HSC70-interacting protein (CHIP) (355, 356).

4.2.1.2 HDACis and antiestrogens as a combined treatment for breast cancer

The *ESR1* gene promoter is hypermethylated in ERα-negative breast cancer cells, resulting in constitutive silencing of the gene (357, 358). There are several reports in the literature indicating that HDACis can relieve epigenetic silencing of *ESR1* in ERα-negative breast cancer cells, which in turn can sensitize them to antiestrogen treatment. Together with the strong antiproliferative properties of HDACis and their ability to inhibit expression of ERα in ERα-positive breast cancers, these observations serve as a rationale for the use of these inhibitors in combination with antiestrogens for a potentially more effective treatment. Indeed, co-treatment of MDA-MB-231 cells with 5-aza-2′-deoxycytidine and TSA lead to re-expression of ERα mRNA and protein. This sensitized cells to tamoxifen treatment and resulted in increased cell death associated with caspase 3 activation and subsequent PARP cleavage (359). In a subsequent study, the same group showed that another HDACi (LBH589) was sufficient to restore ERα expression in MDA-MB-231 cells with no need for DNMT inhibitors (360). LBH589 treatment abrogated the formation of a repressor complex composed of DNMT1, HDAC1, and the H3K9 methyltransferase SUV39H1 at the ESR1 gene promoter and reversed tamoxifen resistance (360). In contrast to these findings, Jang et al. reported that TSA treatment of ERα-negative breast cancer cell lines MDA-MB-231 and Hs578T restored their tamoxifen sensitivity through increased expression and transcriptional activity of ERβ but not ERα (361). It was found that abrogation of ERβ expression by siRNAs desensitized cells to tamoxifen. Another mechanism by which HDACis can restore tamoxifen sensitivity in ERα-negative breast cancer cells is through abrogated expression of EGFR, preventing activation of EGF-mediated signalling and downstream activation of PAK1, p38MAPK and AKT, all of which are involved in tamoxifen resistance (362). HDACis have also been reported to potentiate the anti-proliferative actions of antiestrogens in a variety of ERα-positive breast cancer cell lines and block agonist activity of tamoxifen in uterine endometrial cells (363).

These findings resulted in a phase II clinical trial that has assessed the clinical benefit from combining tamoxifen treatment with SAHA for treatment of ERα-positive metastatic breast tumours. The objective response rate was 19% (8/43 patients) and stable disease ≥ 24 months was observed in 9/43 patients (364). These results were encouraging as the expected response rate in this population of patients is 7.4% for fulvestrant (365). Several other phase I, II and III clinical trials are currently assessing combinations of entinostat/exemestane, SAHA/tamoxifen or entinostat/azacitdine for the treatment of metastatic ERα-positive and triple-negative breast cancers.

Chapter Two.

FOXA1 as a Master Regulator in Molecular Apocrine and Luminal Breast Cancer Subtypes

RATIONALE

Breast cancer is a heterogeneous disease with several tumour subtypes that can be identified by hierarchical clustering of tumours based on their gene expression profiles: Luminal, HER2-positive, molecular apocrine, basal-like and claudin-low. These subtypes are thought to arise from a block at different stages of breast epithelial cell differentiation. However, the mechanisms underlying subtype specification remain unclear. We have recently developed a tool called 'MiSTIC', which integrates data from several platforms including RNA-Seq, ChIP-Seq, as well as clinical features allowing for identification of clusters of co-expressed genes and generating testable hypotheses (366). By applying MiSTIC in the TCGA dataset of breast tumours, we identified a highly correlated cluster of genes in luminal tumours, which contribute to differentiation of breast tumour subtypes. We found that this cluster contains 45 genes including six luminal-specific transcription factors (ERα, FOXA1, GATA3, SPDEF, XBP1 and AR) that have documented roles in mammary gland development and luminal breast epithelial cell differentiation. Importantly, basal-like tumours that lack ERα, also lack expression of FOXA1 and GATA3. In addition, as mentioned earlier, expression of FOXA1 and GATA3 is associated with good prognosis and well-differentiated tumour status. Both TFs have been reported to act as 'pioneer factors' that positively regulate ERα expression and function. Thus, in this study, we are asking two questions: What is the TF hierarchy in luminal cells? And how is ERα expression/function regulated? This is important as unravelling mechanisms of regulation of ERα in luminal tumours can help understand how its expression is lost in ERα-negative tumours. This will ultimately lead to the design of therapies that aim to restore the receptor's expression to induce tumour differentiation.

FOXA1 AS A MASTER REGULATOR IN MOLECULAR APOCRINE AND LUMINAL BREAST CANCER SUBTYPES

Houssam Ismail[1], Salwa Haidar[1], Tatiana Traboulsi[1], David Laperrière[1], Eric Audemard[1], Sebastien Lemieux[1,2], and Sylvie Mader[1,3,*]

[1] Institute for Research in Immunology and Cancer (IRIC), Université de Montréal, Montréal, QC H3C 3J7, Canada.
[2] Computer Science and Operation Research, Université de Montréal, Montréal, QC H3C 3J7, Canada.
[3] Department of Biochemistry, Université de Montréal, Montréal, QC H3C 3J7, Canada, and Centre de Recherche du Centre Hospitalier Universitaire de l'Université de Montréal, Montréal, QC H2X 0A9, Canada.

* To whom correspondence should be addressed.

Tel: +1-514-343-7611; Fax: +1-514-343-6843; Email: Sylvie.Mader@umontreal.ca

Author contributions

H.I. and S.M. designed the study, analysed data and wrote the manuscript. H.I. designed all experiments, performed experiments and analysed all data. S.H. generated CRISPR-cas9 sg*FOXA1* plasmids, assisted in sample preparation and validation experiments for western analysis in cell lines in Figure 4 and Supplementary Figure 12. T.T. performed FAIRE experiments and generated the MCF-7 transcriptome dataset. D.L. generated the TFBS pipeline. E.A. processed raw transcriptome and ChIP-Seq data. S.L. provided assistance with the MiSTIC tool.

ABSTRACT

Transcriptome profiling in breast tumours has revealed different subtypes, but the underlying causes of subtype determination still require further investigation. Using correlation analysis in breast tumour datasets, we identified a cluster of highly-correlated luminal genes including *ESR1*, *FOXA1* and *GATA3*. Importantly, expression patterns of these TFs recapitulate stratification of luminal (*FOXA1*^{high} *ESR1*/*GATA3*^{high}), molecular apocrine (*FOXA1*^{high} *ESR1*/*GATA3*^{low}) and basal-like (*FOXA1*^{low} *ESR1*/*GATA3*^{low}) tumours. Downregulation of FOXA1 expression in luminal cells supported a role as a positive modulator of other luminal TFs, including *ESR1*. In molecular apocrine SK-BR-3 cells, which are ERα-negative and GATA3^{low}, FOXA1 maintained binding to response elements present in several luminal cluster genes, although its binding to *ESR1* regulatory regions was reduced and this correlated with the presence of repressive chromatin marks. Ectopic ERα expression in these cells resulted in E2-induced binding of the receptor to estrogen response elements (EREs) as well as pre-bound FOXA1 elements, with chromatin opening observed at some EREs with low FOXA1 binding. In addition, these cells exhibited activation of an E2-transcriptional response reminiscent of that observed in luminal cells, albeit several cell cycle genes were already active in an ERα/E2 independent manner. Collectively, these findings support transcriptional cooperativity between FOXA1 and ERα for expression of luminal genes, and suggest that differential expression of these master regulators drives breast tumour subtype specification.

INTRODUCTION

While breast cancer is the most frequently diagnosed cancer in women worldwide, it is a heterogeneous disease, with several “intrinsic” subtypes that have been identified through non-supervised clustering analyses based on gene expression profiling studies. These include the luminal, HER2-positive, basal-like, claudin-low and molecular apocrine subtypes, which can be identified by various gene classifiers albeit with only partial concordance in tumour partitioning between subtypes (139, 140, 144, 145, 367). Luminal tumours express ERα (encoded by the *ESR1* gene) and are candidates for hormonal therapies, such as antiestrogens or aromatase inhibitors that aim to block activity of the receptor (368). HER2-positive tumours harbour an amplification in the HER2-encoding gene, *ERBB2*, and can be targeted with inactivating antibodies or small molecule inhibitors of this receptor (369). Basal-like, molecular apocrine and claudin-low tumours are ERα-negative and are currently treated mainly with chemotherapy in the absence of common tumourigenesis driver events, although novel therapeutic regimens are currently being tested in subsets of patients based on the presence of specific genetic defects (370). Tentative models associating breast cancer subtypes with their hypothesized normal mammary epithelial cell counterpart have been proposed (367, 371, 372), postulating that tumours arise at different stages of epithelial cell differentiation, with ulterior steps being blocked by oncogenic defects present in the tumour (372). However, the determinants of tumour subtype specification remain largely unexplored, and could reflect both the impact of genetic defects on the expression of affected genes (e.g. *ERBB2* amplification) or alterations in gene expression networks due to deregulated expression of transcription factors (TFs).

Reprogramming of cell fate through modulation of expression of TFs has been widely described, perhaps most notably in conversion of fibroblasts to induced pluripotent stem cells (373) and in epithelial to mesenchymal transition in tumour cells (374). In the specific case of breast tumours, expression of ERα in two thirds of breast tumours specifies a luminal phenotype. However, genetic alterations also influence gene expression patterns, complicating the interpretation of non-supervised tumour subtype classification. For instance, HER2-positive tumours can be either ERα-positive or ERα-negative (375), the latter group overlapping with the molecular apocrine subtype that also includes HER2-negative tumours (199). It thus remains unclear to what extent transcriptional networks explain breast tumour subtypes and which transcription factors act as master regulators of tumour cell fate.

Addressing these questions may lead to the development of therapeutic strategies that rewire cancer cells through induced differentiation.

Transcription factors can have broad influence on cell differentiation by acting either as master transcription factors or pioneer factors. Master transcription factors are expressed early during differentiation processes and control the expression of downstream transcription factors. The term 'pioneer factor' emerged over the past decade to describe TFs that can bind closed chromatin, and elicit cell-specific enhancer competence, ultimately influencing the process of cell differentiation (376). ERα has long been recognized as a master regulator of the luminal subtype. In addition, FOXA1 and GATA3 have been described as pioneer TFs that facilitate 17β-estradiol (E2) signalling and control mammary gland development (95, 377–379). FOXA1 was found at approximately 50% of ERα-binding events in MCF-7 breast cancer cells via genome-wide ChIP-chip or ChIP-Seq analysis (237, 377). Furthermore, its depletion was found to result in deregulation of expression of roughly 95% of E2-target genes as well as E2-induced proliferation without an impact on ERα expression (377). On the other hand, a subsequent study reported a positive regulatory role of FOXA1 in ERα expression in MCF-7 cells (95). In the mouse mammary gland, FOXA1 is predominantly expressed in luminal progenitors and committed mature luminal cells and its ablation results in impaired ductal invasion and terminal end bud formation, and reduced ERα expression in luminal epithelial cells (95). GATA3 depletion was also reported to result in redistribution of nearly one-third of ERα-binding events and in an altered E2-transcriptional response (378). In addition, its knockdown has been shown to reduce expression of the receptor (380). GATA3 ablation in mouse models causes a block in differentiation of mammary stem cells, resulting in expansion of a $CD61^+$ luminal progenitor cell population (379). Other mammary luminal-lineage TFs include SAM pointed domain containing ETS (SPDEF), and X-box binding protein 1 (XBP1). SPDEF plays an important role in maintaining the luminal phenotype by regulating expression of luminal-lineage genes and is positively regulated by FOXA1 (118). *XBP1* is an ERα target gene in MCF-7 cells (223) and its deletion in mice leads to failure in mammary ductal branching and lactogenic differentiation (125).

In this study, we aimed to analyse the hierarchy of TFs that play a role in the specification of the luminal phenotype and examine how their expression impacts the hierarchical clustering of breast tumours. Using large breast tumour transcriptome datasets, we observed that FOXA1 expression pattern is broader than that of ERα and GATA3. In addition, we characterized its role as a master regulator of ERα-negative molecular apocrine

breast cancer cells, and in transcriptional cooperation with ERα in luminal breast cancer cells. Together, our results suggest that *FOXA1* and *ESR1*/*GATA3* expression orchestrates transcriptional programs implicated in specifying three main breast tumour subtypes: luminal, molecular apocrine and basal-like. We propose that these TFs, together with highly correlated target genes that reflect their activity, can form the basis of an informed classification of breast tumour subtypes reflecting the differential activity of the correspoding transcriptional networks.

MATERIAL AND METHODS

Cell culture conditions

All cell lines used in this study were purchased from ATCC. MCF-7 breast cancer cells were maintained in α-Minimal Eagle's Medium (α-MEM; Wisent) supplemented with 10% fetal bovine serum (FBS; Sigma), 1% penicillin/streptomycin (Wisent) and 1% L-glutamine (Wisent). T-47D and ZR-75-1 breast cancer cell lines were maintained in RPMI 1640 (Wisent) supplemented with 10% FBS, 1% penicillin/streptomycin, 1% L-glutamine, 10mM HEPES (Sigma), and 1% sodium pyruvate (Wisent). SK-BR-3 cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM; Wisent) supplemented with 10% fetal bovine serum, and 1% penicillin/streptomycin. For estrogen regulation experiments, cells were cultured in DMEM without phenol red supplemented with 2% L-glutamine, 1% penicillin/streptomycin, and 10% charcoal-dextran treated FBS (FBST) for 72 h. All cell lines were kept in a humidified 37°C incubator with 5% CO_2.

SiRNA Transfections

Cells were cultured in hormone-depleted media and allowed to adhere overnight. Two different siRNAs were used per gene. Cells were transfected with 40nM of ON-TARGETplus siRNA targeting *ESR1, FOXA1, GATA3, SPDEF, XBP1,* and *AR* using SilentFect reagent (BioRad) for a total duration of 72h. ON-TARGETplus Non-Targeting siRNAs were used as negative controls. Cells were stimulated with 25nM E2 (Sigma) for the last 24h of transfection. Cells were then washed twice with ice-cold phosphate buffered saline (PBS) and harvested for subsequent protein and RNA extractions. All siRNAs were purchased from GE Dharmacon and their sequences are provided in **Supplementary Table 1.**

RNA Extractions, Reverse Transcription and Real-Time Quantitative PCR

Total RNA was extracted using the TRIzol reagent according to the manufacturer's protocol (Invitrogen). 2µg of RNA were reverse transcribed to cDNA following specifications of the RevertAid H first minus strand cDNA synthesis kit (Thermo Fisher Scientific) using oligo(dT)18 primers. cDNA was subsequently diluted one in ten and expression levels of target genes were assessed by real-time quantitative PCR (qPCR) using the Universal ProbeLibrary system (Roche). *YWHAZ, RPLP0* and *TBP* were used as reference genes based on their stable expression levels as determined by RT-qPCR following different perturbations. Relative mRNA levels were determined using the $-\Delta\Delta C_T$ method (381). First,

ΔC_T is calculated by determining the difference between the C_T values of the target gene and the reference genes. $\Delta\Delta C_T$ is then calculated as the difference between ΔC_T of each condition (average of two siRNA) and that of the negative control (siControl). qPCR primer sequences and probe numbers are listed in **Supplementary Table 2.**

Transcriptome Sequencing and Analysis

Gene expression was computed with Kallisto (382) on the reference genome GRCh38 with the annotation of ensembl v85 (with cDNA and RNA). Kallisto was running with default parameters (100 bootstraps). Differentially expressed gene analysis was done with Sleuth [http://biorxiv.org/content/early/2016/06/10/058164]. Sleuth is an R package designed to work with Kallisto's output, which implements statistical algorithms (Wald test in our studies which generates a p-value and a q-value (FDR-adjusted p-value)) for differential analysis that leverage the bootstrap estimates. The Log2FC is calculated from the mean TPM value of each group of biological replicates. A Log2FC is calculated for each kallisto bootstrap and the reported value is the median of all of them.

MiSTIC

MiSTIC has been previously described (366). The Cancer Genome Atlas (TCGA) breast cancer transcriptome dataset including 754 tumours has been used for correlation analysis. For enrichement analyses, a p-value (Fisher's exact test), q-value (FDR-adjusted p-value using Benjamini–Hochberg) and an odds ratio are calculated.

Protein Extraction and Western Blot Analysis

Whole cell extracts were prepared using a total lysis buffer (50mM Tris-HCl pH 7.4, 5.0mM EDTA, 150mM NaCl, 0.5% Triton, 1.0% NP40. 2% SDS and freshly-added protease inhibitors (PMSF at 10mM; leupeptin, pepstatin, and aprotinin at 1 µg/ml). Extracts were homogenized by sonication and quantified using Lowry assay (BioRad). Equal amounts of proteins (20-60µg) were electrophoresed on an 8% SDS-polyacrylamide gel. Proteins were transferred onto polyvinylidene difluoride membranes (Millipore). Membranes were blotted with a rabbit anti-ERα (1 in 3000; Clone 60C from Millipore), rabbit anti-FOXA1 (1:3000; Abcam, ab23738), mouse anti-GATA3 (1 in 3000; BioLegend, Clone 16E10A23), rabbit anti-SPDEF (1:1000; Santa Cruz, sc-67022), rabbit anti-XBP1s (1:1000; Biolegend, 619501), mouse monoclonal anti-β-actin (1 in 10000; Sigma), rabbit anti-AR (1:1000; Santa Cruz, N-20), rabbit

anti-HER-2 (1:2000; Millipore Sigma, 06-562). HRP-conjugated anti-mouse and rabbit IgG were used as secondary antibodies (Cedarlane). Immunodetection was performed using enhanced chemiluminescence (PerkinElmer Life and Analytical Sciences) as per manufacturer's instructions.

Chromatin Immunoprecipitation (ChIP)

Cells were crosslinked by incubation with 1% formaldehyde for 10min at room temperature. Crosslinking was stopped by administration of 0.15M glycine for 5min. Cells were subsequently washed twice with ice-cold PBS and harvested and lysed for 10min on ice in cell lysis buffer (5mM Pipes pH8, 85mM KCl, and 0.5% NP-40) supplemented with a fresh mixture of protease and phosphatase inhibitors. The lysates were then centrifuged at 1000g for 5min at 4°C and nuclear pellets were kept and lysed for 10min on ice in nuclear lysis buffer (50mM Tris pH8.1, 10mM EDTA pH8, and 1% SDS) supplemented with a fresh mixture of protease and phosphatase inhibitors. To shear DNA, lysates were subjected to 4 rounds of sonications (8min each) using Bioruptor (Diagenode) set at 'medium' power with a 30sec interval between pulses. IPs were performed on 50µg of DNA diluted in 1/20 in dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7mM Tris pH8.1, 167mM NaCl) to attain a final SDS concentration of 0.05% and incubated with 4µg of antibodies overnight on a rotor at 4°C. A 1:1 mix of Dynabeads A and G (Invitrogen 10002D and 10004D) was then added for 2h on a rotor at 4°C to capture antibody-protein-DNA complexes. Beads were then washed twice with dialysis buffer (2mM EDTA, 50mM Tris pH8.1, and 0.2% Sarkosyl) and 4 times with wash buffer (0.5M LiCl, 1% NP40, 1% sodium deoxycholate, and 33.2mM Tris pH8.1) on a rotor at RT with each wash lasting 15min. Input DNA was precipitated overnight by addition of two volumes of 95% ethanol. Reverse-crosslinking and DNA elution was performed by heating beads for 30min at 65°C in elution buffer (50mM NaHCO3, 1% SDS). The eluates were subsequently incubated with Proteinase K (ThermoScientific, EO0491) for protein digestion and then DNA fragments were purified on EZ-10 columns (BioBasic BS664). RT-qPCR was subsequently performed using the Universal Probe Library (Roche) kit and the ViiA 7 Real-Time PCR System (ThermoFisher Scientific). ChIP results are shown as percentage of input (% input) or fold change (% input antibody/%input of IgG) as indicated. ChIP primer sequences and their respective probe numbers are provided in **Supplementary Table 3.**

FAIRE

Cells were crosslinked for 10min through incubation with 1% formaldehyde. Crosslinking was stopped by addition of 0.125M glycine for 5min. Cells were then washed twice with ice-cold PB, harvested and lysed in FAIRE lysis buffer (1% SDS, 10mM EDTA pH 8, and 50mM Tris pH 8.1) supplemented with a fresh mixture of protease and phosphatase inhibitors. DNA was sheared through one round of sonication using Bioruptor (Diagenode) set at high power with a 30sec interval between pulses for 15min. Lysates were centrifuged for 15min at 14,000 rpm at 4°C and supernatants were kept for three subsequent phenol:chloroform:isoamyl alcohol 25:24:1 (Sigma P3803) extractions using Phase Lock gel tubes (5 Prime, #2302810) keeping the aqueous phase ('open chromatin') each time. FAIRE and input DNA was precipitated with 2 volumes of 95% ethanol, 1/6.3 volume of 2M sodium acetate pH 5.2 and 20µg/mL of Glycogen (Fermentas #R0551) at -80°C for 48h. Samples were then incubated with RNAse A (BioShop #RNA675.100) and Proteinase K (ThermoScientific, EO0491) and reverse-crosslinked overnight at 65°C. EZ-10 columns (BioBasic BS664) were used to purify DNA fragments. Fragment abundance was subsequently assessed by RT-qPCR was using the Universal Probe Library (Roche) kit and the ViiA 7 Real-Time PCR System (ThermoFisher Scientific). FAIRE results are shown as percentage of input (% input). Primers and probe numbers used in FAIRE experiments were the same ones used in ChIP assays (see **Supplementary Table 3).**

CRISPR-Cas9-mediated gene recombination

CRISPR/Cas9 mediated FOXA1 knockdown experiments in T-47D and ZR-75-1 cells were performed using the lentiCRISPR v2 plasmid (Plasmid #52961-addgene). Briefly, five single-guide RNA (sgRNA) targeting *FOXA1* and 2 scrambled sgRNA controls were designed as described previously (383) (**Supplementary Table 4**). Each sgRNA was cloned into the lentiCRISPR plasmid by standard cloning methods. Lentivirus were packaged and produced in HEK-293T cells by co-transfecting 6µg of lentiCRISPR v2 plasmid, 1.5µg of pMDLg/pRRE, 1.5µg of pRev and 3µg of pMD2.G using lipofectamine 2000 (Invitrogen). Media was changed 18h after transfection. Viral supernatants were collected 32h to 72h after transfection and filtered through 0.45µm filters. T-47D and ZR-75-1 cells were plated and grown to 75% confluence and infected. 72 h after infection, cells were selected with puromycin (1 ug/ml) for 7-10 days. Protein extractions were performed on puromycin-selected cell populations.

Transient transfections and generation of stable SK-BR-3:ER cells.

An electroporation strategy was used for transient transfections in SK-BR-3 cells. Five million cells were counted and resuspended in 200µl of culture media. A 40µg DNA mix was prepared containing 6µg of plasmid DNA (PSG5-*ESR1* or PSG5 (Empty vector)), 34 µg of salmon sperm DNA (Invitrogen) and 0.2M NaCl. The cell and DNA mixture was then placed in a cuvette and pulsed using a Gene Pulser (Bio-Rad) at 250V and a capacitance of 975µF. Electroporated cells were then washed with fresh medium and allowed to adhere overnight. The following day, cells were treated with E2 (25nM) for 24 h and RNA and proteins were extracted for subsequent assays.

Retroviruses were produced by transient cotransfection of HEK293 cells with VSVG envelop plasmid, gag-pol packaging plasmid, and pBabe-HA-ESR1. The envelop and packaging plasmids were kind gifts from the Guy Sauvageau lab at IRIC. Filtered retroviral supernatants containing polybrene (8 µg /ml) were added to SK-BR-3 cells. Infected cells were selected with puromycin at a final concentration of 2 µg/ml. SK-BR-3 cells were infected in parallel with retroviruses expressing the empty vector (pBabe) as a negative control.

RESULTS

ESR1 expression strongly correlates with that of several luminal TFs in independent breast tumour transcriptome datasets

Using MiSTIC, a tool for correlation and enrichment analysis in large transcriptome datasets (366), we identified a cluster of 44 genes whose mRNA expression was highly correlated with that of *ESR1* in the Cancer Genome Atlas (TCGA) breast tumour RNA-Seq transcriptomes dataset (754 profiles) (**Figure 1)**. This cluster was also largely present in the METABRIC dataset (997 breast tumour expression microarrays), albeit of smaller size (27 genes) (**Supplementary Figure 1A)** (146). Correlation analysis performed on an in-house cohort of 150 breast tumours (Quebec-Alberta dataset) with associated RNA-Seq profiles also identified a similar cluster (40 genes, **Supplementary Figure 1B**). 27 genes were common to at least two out of three datasets, supporting the robustness of this association across platforms, sample size and transcriptome analysis methods (**Supplementary Figure 1C**). The TCGA cluster was found to be enriched in various gene signatures for normal luminal cells or luminal tumours (142, 371, 384–386) **(Supplementary Table 5).** In addition, expression of genes in this cluster was highest in luminal (A and/or B) and lowest in basal-like tumours using the PAM50 classifier **(Supplementary Figure 2)**, although overlap with the PAM50 classifier was reduced to *ESR1*, *FOXA1* and *MLPH.*

The high correlation in expression between genes in the luminal cluster in the TCGA dataset is not due to copy number variations affecting all genes in the cluster, since these are located on different chromosomes **(Supplementary Table 6)**. However, we note that several pairs of genes colocalize to the same chromosomal locus, often adjacent to each other (*GATA3-AS1* and *GATA3* (3kb); *ABCC8* and *KCNJ11* (7kb); *ARSG* and *SLC16A6* (7.8kb); *ANXA9* and *FAM63A*: 14kb; *ARMT1* and *RMND1* (47kb); *AGR2* and *AGR3* (67kb); *TFF1* and *TFF3* (50.6kb); *CCDC170* and *ESR1* (163kb)), suggesting either co-amplification/deletion, observed in less than 2% tumours for most gene pairs, or common transcriptional regulation via shared regulatory regions. Notably, the luminal cluster contains several luminal-lineage TFs in addition to *ESR1*: *FOXA1, GATA3, SPDEF, XBP1* (*AR* is present in the TCGA dataset but is absent in the top 50 genes most correlated with *ESR1* in the METABRIC and Quebec-Alberta datasets), whose expression has been reported to be strongly correlated in other breast tumour datasets (118, 139, 140, 144). We thus probed the flanking regions of cluster genes for the presence of ChIP regions associated with each TF in published ChIP-chip or ChIP-Seq datasets in MCF-7 cells (245, 387–390) and in LNCaP prostate cancer cells for AR

(391). Indeed, genes within the cluster were enriched for binding regions of ERα, FOXA1, GATA3, SPDEF and AR within 5 kb of their transcriptional start sites with statistical significance (**Figure 1 and Supplementary Figure 3**), with 41 out of 45 genes containing binding regions for at least three of these TFs. Of note, we did not observe enrichment of cluster genes with XBP1-bound regions (38), possibly reflecting roles of XBP1 not related to subtype specification, although the presence of bound regions at larger distances cannot be excluded. It is noteworthy that the XBP1 ChIP-Seq dataset was generated using a fusion protein between XBP1 and GFP, which may affect DNA binding patterns (390). Interestingly, we found that the cluster is also enriched in p300/CBP-binding regions within 25 kb from the TSS (40 out of 45 genes for both factors; *P*-value= 7.71 x 10^{-08}; Odds ratio= 7.98) (241). p300/CBP have been reported to act as transcriptional co-activators that are recruited to DNA in concert with ERα, FOXA1 and/or GATA3 to elicit a functional E2-response (245). These observations suggest that luminal TFs within the cluster may regulate expression of cluster genes, causing their high correlation in several independent datasets.

FOXA1 is a master regulator of luminal cluster genes

To explore the potential role of luminal TFs in regulating expression of highly-correlated genes, we designed two different siRNAs against each TF and monitored impact of their depletion on expression of cluster genes by RT-qPCR in MCF-7 cells in the presence of E2. Expression of most cluster genes was regulated by at least one and often several cluster TFs (**Supplementary Figure 4).** FOXA1 and ERα depletion had the widest influence on transcription of cluster genes, followed by that of GATA3, while SPDEF, XBP1 and AR had a minor impact on cluster gene expression (**Supplementary Figure 4)**. Notably, several genes within the cluster were induced by E2 treatment (*AGR*2, *AGR3, CA12, FSIP1, GPR77, SLC7A8, TFF1, TFF3* and *XBP1),* an effect that could be abrogated by knockdown of *ESR1*. Importantly, knockdown of *FOXA1* reduced E2-mediated induction of expression of several of these genes (*AGR*2, *AGR3*, *FSIP1*, *GPR77*, *TFF1, TFF3* and *XBP1*). In addition, *GATA3* knockdown had a similar impact on expression of some (*AGR*2, *AGR3, CA12, FSIP1,* and *SLC7A8)* but not all E2-induced targets within the cluster. An exception is *RABEP1*, a cluster gene that is bound by ERα, GATA3, SPDEF and AR, but is not affected by knockdown of these TFs **(Supplementary Figure 3 and 4)**. It remains possible that transcriptional regulation could only be observed following more complete knockdown or simultaneous knockdown of several TFs.

Even though several cluster genes contained flanking regulatory sequences co-occupied by several TFs, we did not observe an equal influence of targeting individual TFs. For instance, *CCDC170* is bound by ERα, FOXA1 and SPDEF (within 5 kb from the TSS), yet only knockdown of *FOXA1* affected its expression (**Supplementary Figure 3 and 4)**. On the other hand, *ANXA9* and *FBP1* contained FOXA1 and ERα peaks, but were upregulated by *ESR1* knockdown and unaffected by *FOXA1* knockdown. In addition, while *TFF1/TFF3* and *AGR2/AGR3* were downregulated by knockdown of both factors, *CA12* was upregulated by *FOXA1* knockdown but downregulated by *ESR1* knockdown. Thus, the presence of ChIP peaks is not sufficient to predict the existence/strength or the direction of regulation by the corresponding TF.

Interestingly, several genes were regulated at the transcriptional level despite lack of TF binding within 25 kb of their TSS, as in the case of FOXA1 positively regulating expression of *DEGS2* and *PRR15*. This might reflect the presence of long-range chromatin interactions with enhancers at distances farther than 25 kb, or indirect regulation. The tight correlation between expression of TFs present within this cluster indeed suggests that they might cross-regulate each other and form a transcriptional network regulating expression of cluster genes. Accordingly, knockdown of *FOXA1* resulted in significant downregulation of mRNA expression of *ESR1, GATA3, SPDEF*, *XBP1* and *AR* (**Figure 2A)**. Downregulation of *ESR1* expression by siRNA-mediated knockdown of *FOXA1* was verified at the protein level in MCF-7 cells **(Figure 2A and B)**. Decreased ERα protein levels were also observed in T-47D and ZR-75-1 cells either transfected with siRNA against *FOXA1* (**Figure 2C and D**) or infected with CRISPR-Cas9 sgRNA **(Supplementary Figure 5**). The general positive-regulatory role of FOXA1 in expression of luminal TFs (see also below for additional evidence of the role of FOXA1 as an upstream regulator of XBP1 and SPDEF) may explain indirect effects of FOXA1 downregulation on expression of cluster genes. For instance, *PRR15*, which is repressed by FOXA1 knockdown in the absence of a FOXA1-associated region, is bound by SPDEF, and si*SPDEF* results in its downregulation.

While siRNA-mediated depletion of *GATA3* in MCF-7 cells did not result in detectable changes in *ESR1* expression (**Figure 2A and B**), a reduction in ERα protein levels comparable to that obtained with si*FOXA1* was observed in T-47D and ZR-75-1 cells (**Figure 2C and D**). This is in agreement with the previous report by Eeckhoute et al. where GATA3 was shown to bind two enhancer sites upstream of the *ESR1* TSS and positively regulate expression of the receptor in T-47D cells (380). We note that unlike T-47D and ZR-75-1 cells,

which harbour WT *GATA3*, MCF-7 cells contain a heterozygous frameshift mutation in *GATA3*, which results in a truncated protein that has reduced binding to DNA but also increased stability (392, 393). siRNA-mediated knockdown of *ESR1* or *GATA3* caused a significant increase in *FOXA1* mRNA expression in MCF-7 cells, albeit without a significant change in its protein levels in all three cell lines at the examined time point (**Figure 2A and B**). In addition, regulation of *GATA3* by E2 was nonsignificant at the mRNA and protein levels in MCF-7 cells, suggesting a dominant role of FOXA1 in transcriptional regulation of luminal cluster genes, either directly via association with their regulatory sequences or through positive regulation of expression of ERα and/or of other luminal TFs.

Molecular apocrine tumours are *FOXA1*^{high}*ESR1*^{low}*GATA3*^{low}

MiSTIC generates a radial icicle representation of gene correlations resulting in peaks whose width and length is proportional to cluster size (gene number) and the correlation coefficient at which the cluster forms, respectively (366). Interestingly, examination of the luminal correlation cluster/peak in the TCGA dataset shown in Figure 3A reveals that it partitions into two sub-clusters at a higher correlation coefficient threshold. The *ESR1* sub-cluster also contains *GATA3*, and the *FOXA1* sub-cluster contains *SPDEF* and *XBP1*. Strikingly, stratifying breast tumours according to expression of *ESR1* and *FOXA1* identified three main populations of breast tumours: *FOXA1*^{high}*ESR1*^{high}, *FOXA1*^{low}*ESR1*^{low}, and *FOXA1*^{high}*ESR1*^{low} (**Figure 3B**). In addition, similar tumour partitioning was observed when comparing expression of *ESR1* to expression of any of the *FOXA1* sub-cluster genes (**Figure 3B and supplementary Figure 6A**), or of *FOXA1* with any of the *ESR1* sub-cluster genes (**Supplementary Figure 7**), reflecting the existence of three main breast tumour subpopulations defined by differential expression of the two sub-cluster metagenes, although tumour clustering was strongest with *FOXA1* and *ESR1*. Notably, separation of breast tumours in three subgroups based on expression of *ESR1* and *FOXA1* was also observed in the METABRIC and Quebec-Alberta datasets (Supplementary **Figure 8A and 8B, respectively**), reflecting the similar organization of the "luminal" correlation cluster in these datasets.

Using the CIT breast tumour classifier, which distinguishes luminal A/B/C, basal-like, normal-like and molecular apocrine breast tumours (394), *FOXA1*^{high}*ESR1*^{high} tumours were identified as luminal A/B/C (blue), *FOXA1*^{low}*ESR1*^{low} tumours as basal-like (black) while *FOXA1*^{high}*ESR1*^{low} tumours were enriched in molecular apocrine (red, either HER2-positive or

HER2-negative) tumours (**Figure 3B**). When compared to luminal tumours, molecular apocrine tumours expressed lower levels of the *ESR1* sub-cluster genes, but comparable levels of the *FOXA1* sub-cluster genes. Of interest, *AR,* which has previously been proposed as a marker and potential therapeutic target in molecular apocrine tumours (395, 396), is highly correlated with *FOXA1*, although with a correlation coefficient below the cut-off for sub-cluster partitioning (**Figure 3B).** On the other hand, when comparing expression levels of *ESR1* to those of *ESR1* sub-cluster genes, *CCDC170, C5AR2*, *AGR3* or *GATA3,* tumours partitioned mostly in two groups **(Figure 3C)**. Molecular apocrine tumours expressed low levels of any pair of *ESR1* sub-cluster genes, and segregated together with basal-like tumours (**Figure 3C and Supplementary Figure 6B)**. Therefore, molecular apocrine tumours are *FOXA1* metagene^{high} and *ESR1* metagene^{low}.

Lastly, the almost total absence of tumour samples that were *ESR1*^{high} and *FOXA1* sub-cluster^{low} (**Figure 3B and Supplementary Figure 6A,** top left quadrant) or *ESR1* sub-cluster^{high} and *FOXA1*^{low} **(Supplementary Figure 7,** lower right quadrant**)** indicates that *ESR1* sub-cluster genes are not expressed in the absence of FOXA1. This observation is consistent with our finding that FOXA1 is an upstream regulator of expression of *ESR1* and of several *ESR1* sub-cluster genes (**Figure 2 and Supplementary Figure 4)**.

FOXA1 directly regulates expression of a "luminal" sub-cluster in an ERα-independent manner in molecular apocrine SK-BR-3 cells

Breast cancer cell lines representative of the molecular apocrine subtype can be identified through use of the CIT breast cancer subtype classifier (394) on cell line transcriptomes and include the SK-BR-3 cell line (397). Indeed, SK-BR-3 cells express FOXA1 but are GATA3- and ERα-negative at the protein level, whereas luminal MCF-7, T-47D and ZR-75-1 cells express all three factors, and triple-negative MDA-MB-231, MDA-MB-436 and MDA-MB-468 cells express none of them (**Figure 4A)**. Additionally, SK-BR-3 cells are HER2-positive but express low levels of AR (**Figure 4A)**. Consistent with what was observed in tumours, SK-BR-3 cells exhibited low expression levels of *ESR1* sub-cluster genes except *CA12*, and high expression levels of most of the *FOXA1* sub-cluster genes (*SPDEF*, *XBP1*, *MLPH*, *PRR15* and *FOXA1***)** in existing RNA-Seq datasets (398) (**Figure 4B**); in addition, expression of *TFF1* and *TFF3* was low but detectable by RT-qPCR (data not shown)**.** MCF-7 and T-47D cells expressed high levels of genes in both sub-clusters, whereas MDA-MB-231 cells presented low expression levels of both sets of genes (**Figure 4B)**.

Next, we investigated whether FOXA1, XBP1, and/or SPDEF regulate expression of the *FOXA1* sub-cluster genes in SK-BR-3 cells, in spite of very low expression of ERα or GATA3. SiRNA-mediated depletion of *FOXA1* resulted in a decrease in mRNA expression of almost all *FOXA1* sub-cluster genes (*XBP1, SPDEF, TFF1, TFF3, AGR2, PRR15*; **Figure 4C**); downregulation of *SPDEF* and *XBP1* was also observed at the protein level (**Figure 4D**). Notably, while knockdown of *SPDEF* or *XBP1* also impacted expression of some cluster genes, including each other, it did not affect expression of *FOXA1* at the RNA or protein levels (**Figure 4C and D**), similar to our findings in MCF-7 cells **(Figure 2A)**. To investigate the mechanism by which FOXA1 regulates expression of this transcriptional network, we assessed enhancer/promoter occupancy by FOXA1 in sub-cluster genes in SK-BR-3 cells, based on a published FOXA1 ChIP-exonuclease dataset in MCF-7 cells (399). In addition, the presence of FOXA1 response elements (FOXREs) at FOXA1-bound sites was determined by TF binding site analysis (HNF3A_01.M01261, score cutoff 65%). Interestingly, FOXA1 binding coincided with one or several FOXREs at the promoter/enhancer regions of *TFF1*, *TFF3*, *SPDEF* and *XBP1* in MCF-7 cells (see UCSC genome browser views in **Supplementary Figure 9).** We confirmed FOXA1 association to *TFF1*, *TFF3*, and *SPDEF* promoter regions in SK-BR-3 cells using ChIP-qPCR **(Figure 4E**). FOXA1 binding to the *XBP1* enhancer site mapped in MCF-7 cells 13kb upstream of the *XBP1* TSS was also verified (**Figure 4E**). Together, these results suggest that in SK-BR-3 cells, FOXA1 regulates expression of genes most correlated with itself in the "luminal" cluster, in part directly via binding to FOXREs, and in part through SPDEF and XBP1. This highlights the role of FOXA1 as a master regulator of luminal/molecular apocrine genes, and together with its target genes, as a marker of both of these subtypes.

Expression of *ESR1* in molecular apocrine SK-BR-3 cells restores part of the E2-dependent gene regulation network in luminal cells

Although ectopic expression of *ESR1*, *FOXA1* and *GATA3* in the triple-negative (claudin-low) breast cancer cell line MDA-MB-231 has been reported to restore E2-dependent growth of these cells, reanalysis of the corresponding microarray expression data reveals that almost all E2-induced changes in gene expression in transfected cells are below a 1.4 fold change cut-off, suggesting a very modest activation of the E2-response. In addition, expression of luminal markers was only slightly increased (245). Given that SK-BR-3 molecular apocrine cells already express FOXA1, but not ERα or GATA3, we investigated whether ERα

transfection is sufficient to restore E2-dependent transcription. We transiently transfected SK-BR-3 cells with an expression vector for ERα (SK-BR-3:ER) or the parental vector (negative control) **(Supplementary Figure 10)** in the presence or absence of E2 for 24 h and sequenced the resulting transcriptomes for comparison with luminal MCF-7 cells. We identified 387 differentially-regulated genes by E2 upon ERα expression in SK-BR-3 cells (FC cut-off of 1.4 vs vehicle treatment; $q<0.05$). Strikingly, the majority of these genes were upregulated by ERα expression (377 genes) whereas only 10 genes were downregulated (**Figure 5A)**. In comparison, the repertoire of genes differentially regulated by E2 (16 h) was larger in MCF-7 cells, with a similar number of upregulated (1842) and downregulated genes (1483), compared to vehicle control ($FC>1.4$ or $FC<-1.4$ and $q<0.05$) as shown in **Figure 5A**. The vast difference in the number of E2-repressed genes between SK-BR-3:ER cells and MCF-7 cells may be due to the lack of expression of factors involved in E2 signalling, such as GATA3, or alternatively, activation of E2-independent growth regulatory pathways, such as *ERBB2* overexpression in SK-BR-3 cells. We cannot however exclude the possibility that stable ERα expression in these cells may produce a broader E2 response. Importantly, though, 95 genes upregulated by E2 in SK-BR-3:ER cells (constituting 25% of upregulated genes) were also induced by E2 treatment in MCF-7 cells (**Figure 5A**). This represents a much higher overlap than that observed in MDA-MB-231 cells upon ectopic co-expression of ERα, FOXA1 and GATA3, where only 6 genes (out of 14 with $FC\geq1.4$) were regulated by E2 in a similar manner to MCF-7 cells (245). Moreover, reanalysis of microarray gene expression profiles of MDA-MB-231 cells stably expressing ERα reported by Quintin et al. also revealed a small degree of overlap (~7%, 31 out of 433 genes) in the E2 response (upregulated genes) vs MCF-7 cells (400). Amongst the commonly upregulated genes in our dataset were well-characterized ERα targets in MCF-7 cells, such as *GREB1* (233, 401, 402), as well as genes in the ERα sub-cluster (*CA12, AGR3, C5AR2*), but also in the FOXA1 sub-cluster (*TFF1/TFF3*) (230, 378, 402), as validated by RT-qPCR (**Figure 5B**). Of note, none of these E2 target genes were significantly regulated in MDA-MB-231 cells expressing ERα, FOXA1 and GATA3 (245). However, some luminal cluster genes induced (*AFF3, CT62, XBP1, AGR2*) or repressed (*APBB2, HID1, C9orf152, GPR160*) in MCF-7 cells were not regulated in SK-BR-3 cells (**Supplementary Figure 11**). Conversely, luminal cluster genes *ANXA9 and FAM63A* were induced in SK-BR-3:ER cells, albeit these genes were not regulated by E2 (16 h) in MCF-7 cells, respectively.

Enrichment analysis for predicted TF binding sites (TFBS) in the flanking regions within 10 kb of TSS of genes commonly upregulated in both cell lines indicated a strong enrichment in EREs (2.4- to 4.9-fold using cut-offs between 70% and 85%) **(Figure 5C)**; accordingly, 66 of the 95 genes (69.5%) harboured ERα ChIP regions within 25 kb of their TSS in MCF-7 cells, the majority of which (43 genes) contained a predicted ERE (identified at 65% cut-off). In addition, enrichment in Ets1 (V_ETS1_B matrix M00339; up to about 4-fold), or NFKB (V_NFKB_Q6 matrix M00194; up to 2.5 fold) binding sites were also observed in commonly regulated genes **(Figure 5C)**, suggesting a role of these factors either in direct or indirect regulation of gene expression by ERα in both cell lines.

The 282 genes that were only significantly upregulated by E2 in SK-BR-3:ER cells included genes either upregulated without statistical significance (11 genes; FC>1.4 and q>0.06) or downregulated by E2 treatment in MCF-7 cells (21 genes; FC<-1.4 and q<0.05). In addition, 12 genes were not expressed in MCF-7 cells. Enrichment analysis in the remaining set of SK-BR-3 unique genes (n=238 out of 282) indicated a smaller enrichment in EREs (2.2- vs 4.9-fold at 85% cut-off). Interestingly, about 35% of these genes (84 out of 238) harboured ERα ChIP regions within 25 kb of their TSS in MCF-7 cells, almost half of which (41 genes) overlapped with an ERE (detected at 65% cut-off); thus in spite of their lack of regulation in MCF-7 cells after 16 h of E2 treatment, ERα binding to these genes suggests that the potential for their regulation by E2 in luminal cells might be preserved, possibly under different experimental conditions. While we did not find a detectable enrichment for Ets and NFKB motifs in the SK-BR-3 specific gene set, an enrichment in SMAD (V_SMAD_Q6_01 matrix M00974; up to about 3.8-fold) motifs was observed (**Figure 5C)**. Interestingly, SMAD2, SMAD3 and SMAD4 have been shown to physically interact with a constitutively active form of ERα, and SMAD3 enhanced transcriptional activation by ERα in luciferase assays (403). In addition, HER2 overexpression in MCF10A cells combined with TGF-β stimulation enhances binding of SMAD proteins to TGF-β target genes, suggesting cooperativity between TGF-β and HER2 signalling (404). These data indicate that unique ERα targets in HER2-overexpressing SK-BR-3 cells may be activated by E2 via tethering of the receptor to SMAD proteins.

ERα can open chromatin at EREs and binds to FOXA1-occupied response elements to regulate gene expression in molecular apocrine SK-BR-3 cells

To further examine the mechanisms of gene regulation by ERα in SK-BR-3:ER cells, we examined its recruitment at EREs or FOXREs in regulatory regions bound by either factor in MCF-7 cells **(Figure 6)**. The *GREB1* gene contains several regulatory regions associated with both TFs in MCF-7 cells, including the P3 region, which coincides with a FOXRE and the P4 enhancer, which contains two EREs. Association of FOXA1 was strong at the P3 site, but much lower at the P4 upstream enhancer in SK-BR-3 cells in the absence of ERα. Hormone-induced binding of ERα was detected at both elements in the transiently transfected cells, with a hormone-dependent increase of FOXA1 association at the P3 site, and a hormone-independent increase in FOXA1 association at the P4 site. Analysis of chromatin accessibility by formaldehyde assisted isolation of regulatory elements (FAIRE) (405) at these sites in SK-BR-3 cells stably expressing ERα **(Supplementary Figure 12)** indicated that accessibility of the P3 site, already strongly bound by FOXA1, was not regulated significantly by either ERα expression or E2 treatment **(Figure 6)**. On the other hand, chromatin accessibility was lower at the P4 upstream *GREB1* enhancer and was significantly increased by E2 treatment in ERα-expressing cells, but not in control cells. Similarly, ERα was recruited in an E2-dependent manner to EREs in regulatory regions of the *CA12* and *C5AR2* genes (both in the *ESR1* sub-cluster). The *CA12* element was strongly associated with FOXA1 in the absence of ERα despite the absence of predicted FOXREs, and was not further opened by E2 treatment in the presence of ERα. Conversely, the *C5AR2* ERE, which was weakly bound by FOXA1, was made more accessible by E2 treatment in the presence of exogenous ERα **(Figure 6).** Lastly, we also monitored recruitment of ERα at FOXREs found in regulatory regions of the two FOXA1 sub-cluster genes, *TFF1* and *XBP1* **(Supplementary Figure 13)**. ERα was recruited to the FOXA1-bound element present in the *TFF1* promoter, where an ERE is present in the vicinity of a FOXRE. ERα was also recruited to the FOXA1-bound *XBP1* enhancer, which contains a predicted FOXRE but no predicted ERE. These results are compatible with the positive regulation of *XBP1* and *TFF1* by FOXA1 independently of ERα (**Figure 4**), and with the regulation of *TFF1* by E2 in the presence of exogenous ERα **(Figure 5B)**. However, recruitment of ERα at the FOXRE of *XBP1* was not sufficient to confer significant regulation by E2 in SK-BR-3 cells (**Supplementary Figure 11)**, at least under our experimental conditions.

Overall, there was an excellent match between regions bound by ERα or FOXA1 in MCF-7 and SK-BR-3 cells containing EREs or FOXREs. In addition, co-occupancy of FOXA1 and ERα observed in MCF-7 cells, or lack thereof, was also verified in SK-BR-3:ER cells. Comparison of FOXA1 binding in the presence or absence of transfected ERα indicated that (i) FOXA1 may recruit ERα at FOXREs pre-bound by FOXA1 (*GREB1* and *XBP1* FOXA1 sites) and (ii) E2-dependent binding of ERα on elements not strongly pre-associated with FOXA1 (*C5AR2*, *GREB1* P4 site) can result in further chromatin opening without markedly affecting FOXA1 recruitment. Together, these results suggest that ectopic ERα expression in SK-BR-3 cells elicits ERα binding patterns and an E2-response that resemble those observed in luminal MCF-7 cells. The transcriptional response of ERα sub-cluster genes as well as some FOXA1 sub-cluster genes appears driven by ERα recruitment to high affinity EREs, which is sufficient for chromatin opening, but may also result from ERE-independent recruitment of ERα at already accessible, FOXA1-bound response elements.

ESR1 expression and FOXA1 recruitment to its upstream regulatory regions are epigenetically suppressed in SK-BR-3 cells

As we have observed that FOXA1 positively regulates *ESR1* expression in ERα-positive luminal cells, the lack of expression of the receptor in SK-BR-3 cells, which express FOXA1, remains unaccounted for. Copy number analysis of the *ESR1* locus did not reveal the presence of focal deletions in SK-BR-3 cells (cBioPortal, CCLE dataset). We carried out ChIP experiments to monitor FOXA1 recruitment to *ESR1* regulatory regions, and observed that FOXA1 binding to several enhancers mapped in MCF-7 cells (399) is suppressed in SK-BR-3 cells, while recruitment of FOXA1 at the *SPDEF* promoter occurs at similar levels in both cell lines (**Figure 7**). In addition, ERα is recruited at *ESR1* enhancer sites in MCF-7 cells containing an ERE (site 59) or a FOXRE (sites 5, 31) (**Supplementary Figure 14A and 14B**), consistent with the receptor's capacity to positively regulate its own expression (275). Lack of association of FOXA1 and ERα in upstream ERα regulatory regions in SK-BR-3 cells correlated with decreased recruitment of active chromatin marks H3K9/K14Ac at these elements (**Figure 7**). Finally, increased association with H3K27me3 was observed at promoter A of *ESR1* in SK-BR-3 as well as in triple-negative MDA-MB-231 cells (**Supplementary Figure 14C**). Together, these results indicate that lack of *ESR1* expression in SK-BR-3 cells coincides with epigenetic inactivation and incapacity of FOXA1 to access regulatory sites and open chromatin. Thus, while FOXA1 acts as a master regulator in both

molecular apocrine and luminal cells, its regulation of *ESR1* in luminal cell models is either lost in molecular apocrine cells or gained *de novo* in luminal tumour cells, even though its expression level is comparable in both tumour subtypes at the mRNA level, and between MCF-7 and SK-BR-3 cells at both the mRNA and protein levels **(Figure 4A and 4B)**.

Discussion

ESR1, *FOXA1* and *GATA3* are three TF-encoding genes whose expression is highly correlated in luminal breast tumours (139), however their respective contribution to gene regulation and breast tumour subtype specification still needs clarification. In the present study, we have characterized transcriptional regulation of a luminal cluster of genes whose expression is highly correlated across breast tumour samples. Despite that this cluster of genes was enriched in ERα, FOXA1 and GATA3 binding regions, we found that siRNA-based suppression of each TF favours a hierarchical rather than combinatorial regulation of gene expression, with FOXA1 playing a role as direct regulator of a significant fraction of these genes. Interestingly, regulation by these TFs was observed for genes whose expression was most correlated with them, particularly in the *FOXA1* and *ESR1* sub-clusters, although cross-regulations between the two sub-clusters were also observed. These results suggested that gene expression correlations generated by MiSTIC reflect specific transcriptional regulation. We do not exclude the possibility that some regulations may be missed due to the partial nature of gene suppression under our experimental conditions. GATA3 was found to regulate *ESR1* expression in T-47D and ZR-75-1 cells, consistent with the previous report by Eeckhoute et al. (380). In addition, like previously reported in MCF-7 cells (95), we also observed positive regulation of *ESR1* expression by FOXA1, consistent with the presence of bound FOXREs in *ESR1* regulatory regions. These results provide a molecular mechanism for the tight correlation that exists between these TFs in breast tumour samples. Notably, regulation of *ESR1* expression by FOXA1 was not observed by Hurtado et al. in MCF-7 cells (377), but this may result from shorter time frames of treatment with siRNAs and/or less complete suppression of *FOXA1* expression.

Molecular apocrine tumours were originally described as ERα-negative, and often HER2-positive and AR-positive with increased AR signalling (395, 396, 406). However, 43% of these tumours are AR-negative, as determined by immunohistochemical analysis, suggesting that additional factors may drive the phenotype of these tumours (199). Gene expression profiling of a set of 99 breast tumours has previously led to the identification of a subset of tumours characterized by expression of luminal markers *FOXA1, SPDEF, XBP1* and *AR* but lack of expression of *ESR1* (201). A significant proportion of these tumours were found to harbour strong apocrine morphological features, reminiscent of the molecular apocrine subtype. Robinson et al. subsequently hypothesized that AR signalling can recapitulate the luminal features of these tumours in the absence of ERα (396). By using

MDA-MB-453 cells as a molecular apocrine cell line model, FOXA1 was shown to be important for the majority of AR binding events across the genome (396). Moreover, *FOXA1* depletion modulated expression of molecular apocrine genes without affecting expression of AR (396). In this study, we have utilized SK-BR-3 cells as a cell line model representative of AR^{low} molecular apocrine tumours, as opposed to AR^{high} MDA-MB-453 cells (407). SiRNA-mediated inhibition of *FOXA1* and ChIP assays reveal that FOXA1 directly regulates luminal cluster genes, particularly those most correlated with its expression, via binding to cognate response elements, providing a key to their expression irrespective of ERα and AR levels. Expression of genes highly correlated with FOXA1 defines similar sub-populations of tumours in association with *ESR1* expression, albeit sometimes with lower resolution than FOXA1 itself. We note that despite the significant degree of overlap between AR and FOXA1 ChIP regions in MDA-MB-453 cells (396), expression of a 100-gene signature of molecular apocrine tumours (406) was largely unregulated by the AR agonist, dihydrotestosterone (DHT), in a transcriptome dataset of MDA-MB-453 cells (408). These results would indicate that transcriptional regulation by AR is not the main driver of the molecular apocrine phenotype as previously thought. In fact, our work indicates that FOXA1 represents a better marker in terms of segregating different subgroups of tumours based on its expression levels. Indeed, basal-like and molecular apocrine tumours present more of a continuum in terms of *AR* expression levels compared to *FOXA1*. Thus, FOXA1 appears to contribute to this phenotype by direct and independent transcriptional regulation, which complements its previously documented role as a mediator of estrogenic and androgenic signalling (409).

The partial degree of overlap in the E2 response as well as our observation that a subset of E2 target genes are uniquely regulated by ERα in SK-BR-3 vs MCF-7 cells prompted us to examine whether this is a result of expression of different sets of TFs. Analysis of TFs that are differentially expressed between both cell lines revealed reduced mRNA expression of ERα coactivators SRC-1, SRC-2 and SRC-3 in SK-BR-3 cells compared to MCF-7 cells (data not shown). It remains to be determined whether overexpression of these factors in SK-BR-3 cells can generate a larger degree of overlap in the E2 response with MCF-7 cells. We also note that *FOS*, *FOSB* and *JUN* expression was significantly higher in SK-BR-3 cells in comparison to MCF-7 cells. This was consistent with our finding that AP-1 motifs (V_AP1_C matrix M00199; up 1o 1.8 fold) are enriched in E2 target genes unique to SK-BR-3 cells, but not in commonly regulated ones **(Supplementary Figure 15)**. While estrogens both induce and repress gene expression in MCF-7 cells, mostly positive

regulation was observed in SK-BR-3 cells. Transcriptional repression was very weak and did not pass statistical significance cutoffs for selection of significant gene regulation. The reasons for this are unclear, but it is possible that the differential expression of some transcription factors between the two cell lines, such as GATA3, may explain this bias. Of note, several of the FOXA1 target genes are also regulated by estrogens in ERα-positive breast cancer cells, including *TFF1/3* and *XBP1*. This may be mediated via cooperative transcriptional activation, as in the *TFF1* gene that harbours chromatin regions with both FOXREs and EREs located in close proximity to each other. In addition, it is possible that the observed co-recruitment of ERα to FOXREs observed in SK-BR-3:ER cells results from interaction between the two factors, either directly or mediated via common cofactors or the transcriptional machinery. Indeed, binding of FOXA1 was not decreased upon ERα transfection, suggesting that these sites are still occupied by FOXA1 in the presence of ERα, rather than the receptor replacing FOXA1 at these elements. However, ERα was recruited to FOXA1 elements in the *XBP1* enhancer even though regulation of *XBP1* by E2 was not observed in SK-BR-3:ER cells under our experimental conditions. Notably, SRC-1, 2 and 3 have been shown to bind the *XBP1* enhancer in concert with ERα following E2 treatment in MCF-7 cells (241). Thus, their reduced expression in SK-BR-3 cells may explain lack of induction of this gene by E2. Ultimate identification of ERα-interacting partners as well as DNA-binding patterns in SK-BR-3:ER cells will help shed light on differential accessibility of EREs, and/or cooperativity with different factors/cofactors in the two cell lines. It also remains possible that some ERα-FOXA1 interactions are not transcriptionally-productive, but rather may reflect random contacts between TFs co-expressed in the same cell lines. Use of transcriptionally inactive ERα mutants or induction of binding of WT ERα by antiestrogens rather than estrogens in SK-BR-3 cells will be needed to clarify this question. Furthermore, ChIP-Seq against ERα performed in SK-BR-3 cells should further clarify the degree of overlap of bound chromatin with MCF-7 cells and the enrichment in TFBS at bound chromatin regions.

Ectopic expression of ERα in SK-BR-3 cells led to increased chromatin opening at some of the EREs surveyed. Constitutive and strong association of FOXA1 with these elements correlated with more open chromatin and little effect of ERα binding on chromatin accessibility in FAIRE experiments. This is consistent with the previously suggested role of FOXA1 as a pioneer factor enabling ERα binding to DNA. However, a direct role of estrogens in opening chromatin at EREs with weak association with FOXA1 and lower accessibility in

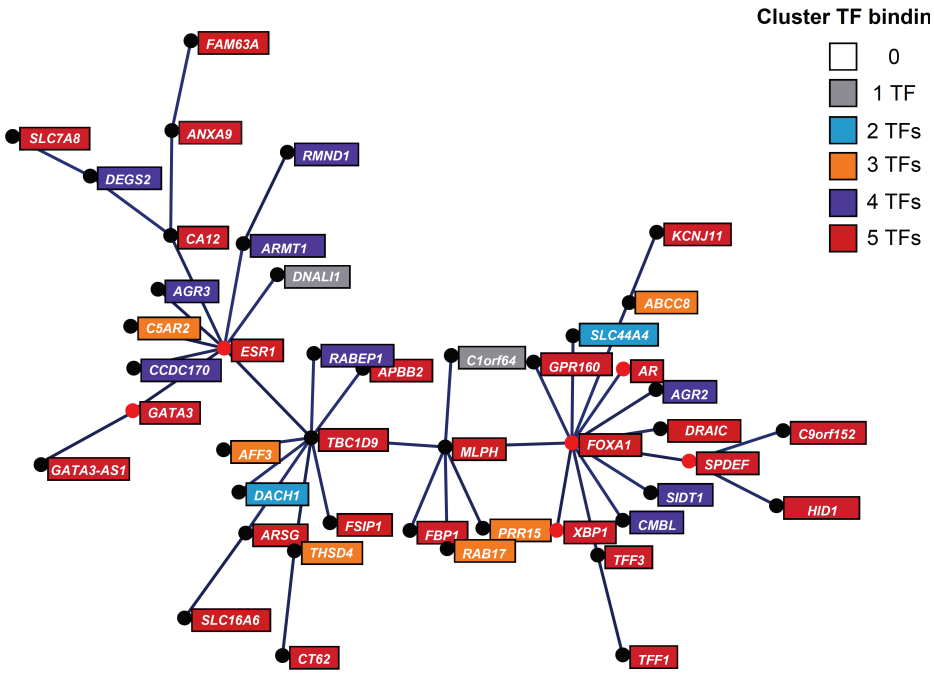
FAIRE experiments was also observed. It remains to be determined whether these effects are dependent on long-range interactions with FOXA1 bound at other regulatory sites or whether ERα can open chromatin in a FOXA1-independent manner on certain elements. We have attempted to examine this directly by depleting FOXA1 in SK-BR-3 cells stably expressing ERα, however FOXA1 knockdown resulted in reduced expression of the ectopically-expressed receptor (data not shown).

We have previously characterized a mitosis/ cell cycle cluster in breast cancer comprising 157 proliferation genes that are positively regulated by E2 treatment (366). Our SK-BR-3:ER RNA-Seq dataset indicated that E2 stimulation in these cells did not result in regulation of expression of this cluster. Therefore, we wondered whether these genes were already expressed in an ERα-independent manner in SK-BR-3 cells. Differential gene expression analysis between untreated SK-BR-3 (transfected with empty vector) and MCF-7 cells revealed that 79 out of 157 genes were expressed in SK-BR-3 cells to a level similar to that in E2-stimulated MCF-7 cells (**Supplementary Figure 16**). It is plausible to speculate that the *ERBB2* gene amplification plays a role in driving this cluster in SK-BR-3 cells. It will thus be of interest to stably express ERα and test whether this confers E2-dependent growth upon suppression of HER2 signalling using Herceptin.

While we have observed that FOXA1 acts as a positive regulator of *ESR1* expression in luminal breast cancer cell lines, its expression in SK-BR-3 cells was not accompanied with ERα-positivity. This suggests the existence of other positive upstream regulators of ERα expression absent in SK-BR-3 cells or gain of repressive mechanisms in these cells. GATA3 levels are low in SK-BR-3 cells, however transient expression of *GATA3* did not result in increased *ESR1* expression (data not shown). Another important regulator may be p53, which has been shown to upregulate *ESR1* expression in MCF-7 cells (255–257). Interestingly, MCF-7 cells harbor a WT version of *TP53* whereas SK-BR-3 express a mutant form. Loss of p53 function in SK-BR-3 cells preventing ERα expression is plausible as *TP53* loss-of-function mutations have been shown to be far more prevalent in ERα-negative basal-like tumours (410). Another possible explanation for the lack of ERα expression in these cells is absence of expression of the receptor in the corresponding cell of origin. On the other hand, *FOXA1* expression could be positively gained in both *ESR1* expressing and silent backgrounds. A recent study has reported the presence of a G>A mutational hotspot in the promoter region of *FOXA1* in ERα-positive breast cancers at position -81 relative to the TSS, which results in increased binding of E2F1 leading to FOXA1 overexpression (170). In

addition, *FOXA1* gene amplifications, albeit infrequent (about 2% of tumours), have been reported (169, 410). On the other hand, FOXA1 expression may be lost in basal-like tumours due to *BRCA1* mutations, which were reported to correlate with *FOXA1* promoter methylation and repression of its expression (411). Knockdown of *BRCA1* in MCF-7 cells can cause increased recruitment of H3K27me3 repressive marks to the *FOXA1* gene promoter and repression of its expression due to increased activation of EZH2, which is normally inhibited by direct interaction with BRCA1 (411). The positive regulatory role that BRCA1 plays in regulating *FOXA1* expression was also described by a another group (412). Future experiments will be necessary to examine whether overexpression of *BRCA1* mutants can abolish expression of *FOXA1* in luminal and molecular apocrine breast cancer cells.

In conclusion, we have shown that mRNA expression of *ESR1* and *FOXA1* sub-clusters largely recapitulates the luminal, molecular apocrine and basal-like phenotypes **(Figure 8)**, and that the role of FOXA1 in subtype specification reflects direct gene regulation mediated through binding cognate sites and/or tethering interactions, and indirect regulation via relay by TFs whose expression it modulates. These results point to a role of FOXA1 as a master regulator of gene expression in the molecular apocrine and luminal breast tumour subtypes.



P-value	Q-value	Odds	Name	Type	PMID
2.51E-07	2.76E-06	13.13	ESR1 MCF-7 cells E2 [25kb TSS]	ChIP-seq	20219941
8.02E-04	3.21E-03	3.18	FOXA1 MCF-7 cells E2 [25kb TSS]	ChIP-seq	21572391
6.19E-03	1.21E-02	2.57	GATA3 MCF-7 cells E2 [25kb TSS]	ChIP-seq	21878914
2.93E-03	8.15E-03	3.1	SPDEF MCF-7 cells [5kb TSS]	ChIP-seq	24043118
3.45E-02	3.45E-02	2.32	AR LNCaP cells DHT [25kb TSS]	ChIP-seq	21602788

Figure 1. The luminal cluster comprises six transcription factors whose expression is highly correlated in breast cancer. Shown is a minimal spanning tree of the 'luminal' cluster composed of 45 genes in the TCGA dataset. Genes encoding transcription factors (*ESR1, FOXA1, GATA3, SPDEF, XBP1* and *AR*) are indicated with a red node. Enrichment for TF binding within 5kb (SPDEF) or 25kb (ERα, FOXA1, GATA3, and AR) from the TSS is illustrated as a colour code. A p-value, q-value and an odds ratio are calculated for statistical significance.

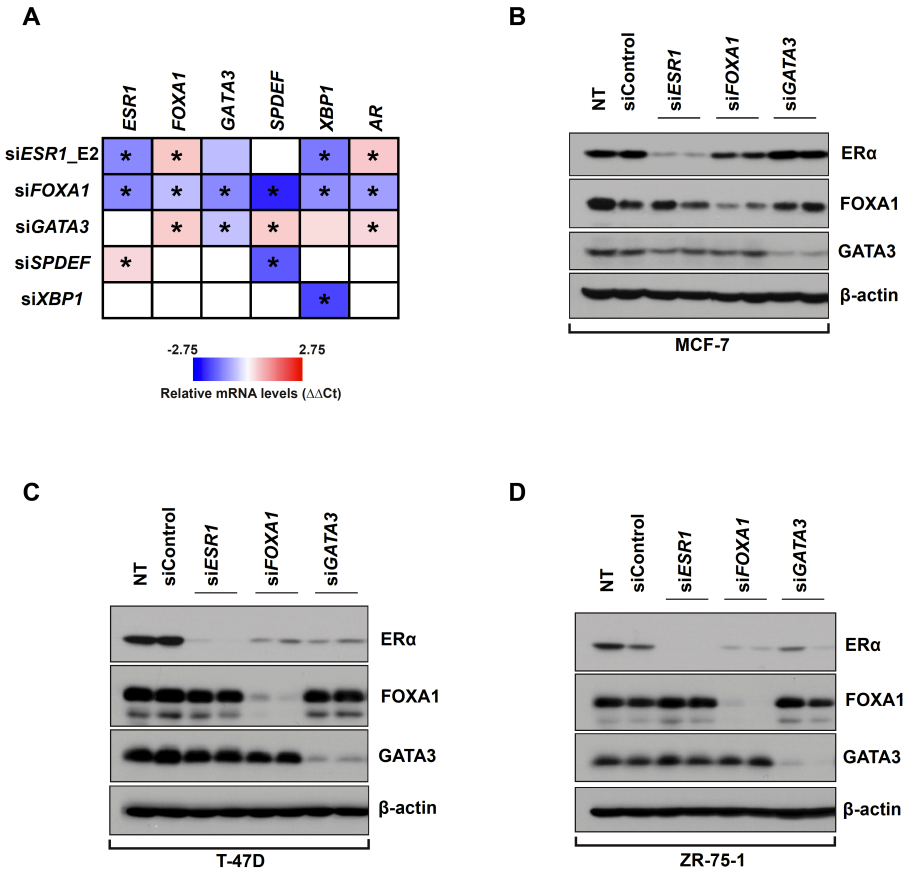


Figure 2. FOXA1 positively regulates expression of ERα in several luminal breast cancer cell lines.

(A) MCF-7 cells cultured in hormone-depleted medium were transfected with two different siRNAs targeting *ESR1, FOXA1, GATA3, SPDEF, XBP1* for 72 h. Cells transfected with si*ESR1* were treated with E2 (25nM) for the last 24 h of transfection. The mRNA level of each TF was measured using RT-qPCR and presented as a heatmap. Upregulation is represented in red and downregulation is represented in blue. White indicates absence of regulation (ratio of siRNA to sicontrol is lower than 1.3 or greater than -1.3). Asterisks indicate statistically significant regulations (N=3, $p<0.05$, student's paired t-test). Protein lysates from MCF-7 (B), T-47D (C), and ZR-75-1 (D) cells transfected with siRNAs in hormone-depleted medium were analysed by western blotting using antibodies against ERα, FOXA1, GATA3 (N=2). β-actin serves as a loading control.

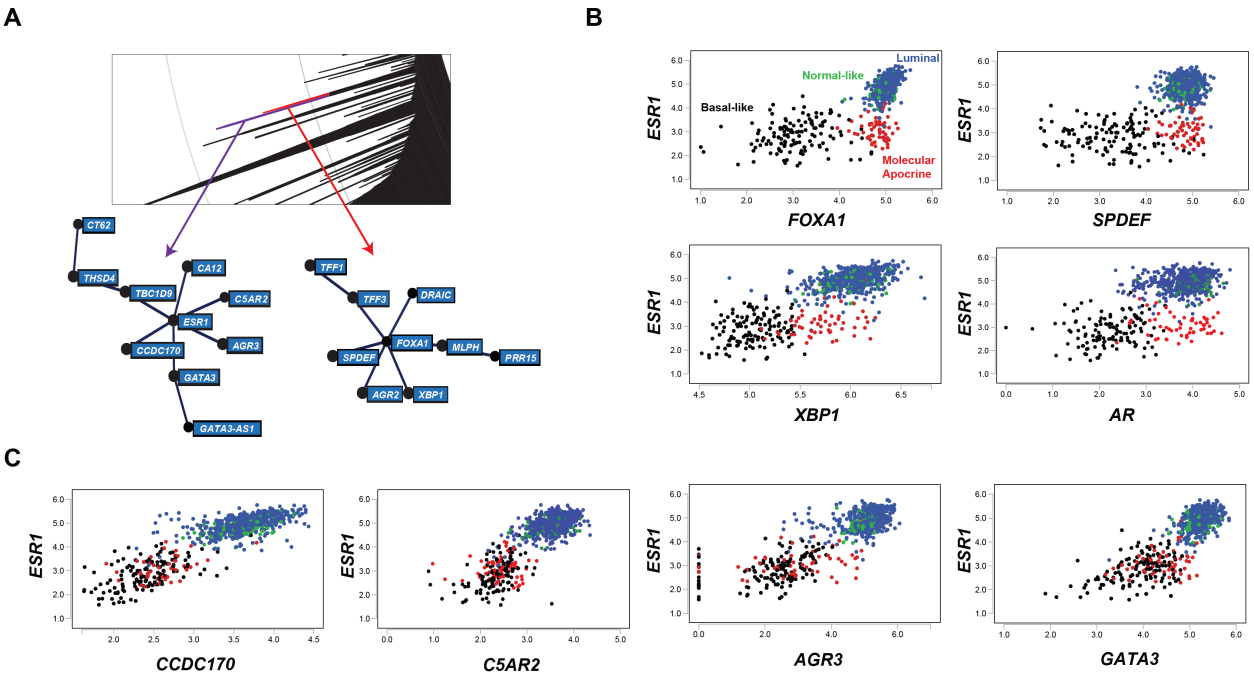


Figure 3. Molecular apocrine tumours are *FOXA1*^{high}*ESR1*^{low}*GATA3*^{low}.
(A) The luminal peak is partitioned into two sub-peaks one containing *FOXA1*, and the other, *ESR1*. (B) Breast tumours from the TCGA dataset were classified into four subpopulations using the CIT classifier: Normal-like (green), Luminal A, B, and C (blue), basal-like (black), and molecular apocrine (red). Tumours were then sorted according to expression levels of *ESR1* vs *FOXA1* sub-cluster genes (*FOXA1, SPDEF*, and *XBP1;* see Supplementary Figure 6A for additional *FOXA1* sub-cluster genes). *AR* expression was added as a known marker of molecular apocrine tumours. (C) TCGA dataset tumours were segregated according to expression levels of *ESR1* vs *ESR1* sub-cluster genes *CCDC170, C5AR2, AGR3*, and *GATA3* (see Supplementary Figure 6B for additional *ESR1* sub-cluster genes).

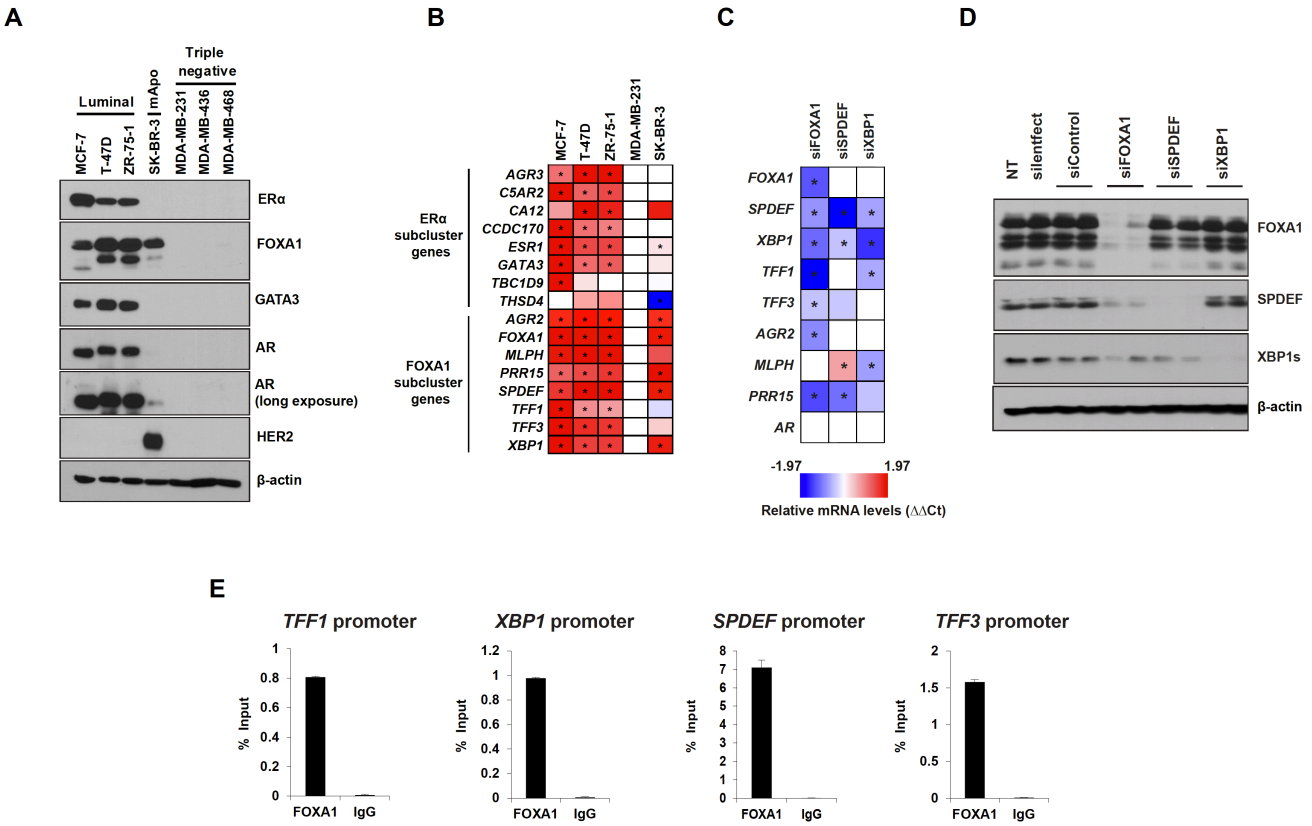


Figure 4. FOXA1 regulates expression of a 'luminal' sub-cluster in an ERα-independent manner in molecular apocrine SK-BR-3 cells.

(A) Protein levels of ERα, FOXA1, GATA3, AR and HER2 were assessed by western blots in luminal (MCF-7, T-47D, and ZR-75-1), molecular apocrine (SK-BR-3), and triple-negative (MDA-MB-231, MDA-MB-436, MDA-MB-468) cells. (B) Relative mRNA expression levels of *FOXA1* and *ESR1* sub-cluster genes in luminal (MCF-7, T-47D and ZR-75-1) and molecular apocrine (SK-BR-3) cells compared to triple-negative (MDA-MB-231) cells are represented as a heatmap. Each gene is given its own heatmap colour code to measure expression across all cell lines. Asterisks indicate statistically significant regulations ($q<0.05$) (C) SK-BR-3 cells cultured in hormone-depleted medium were transfected with two siRNAs targeting *FOXA1*, *SPDEF*, or *XBP1* and expression of genes from the *FOXA1* sub-cluster was measured by RT-qPCR. Asterisks indicate statistically significant regulations ($p<0.05$, N=3, student's paired t-test). (D) protein expression levels of FOXA1, SPDEF, and XBP1s are shown following knockdown of the indicated TFs (N=2). (E) FOXA1 binding to *FOXA1* sub-cluster genes: *TFF1, XBP1, SPDEF*, and *TFF3* in SK-BR-3 cells was examined by Chromatin IP (ChIP) and shown as percentage of input (N=3, one representative experiment is shown). IgG IPs were used as a negative control.

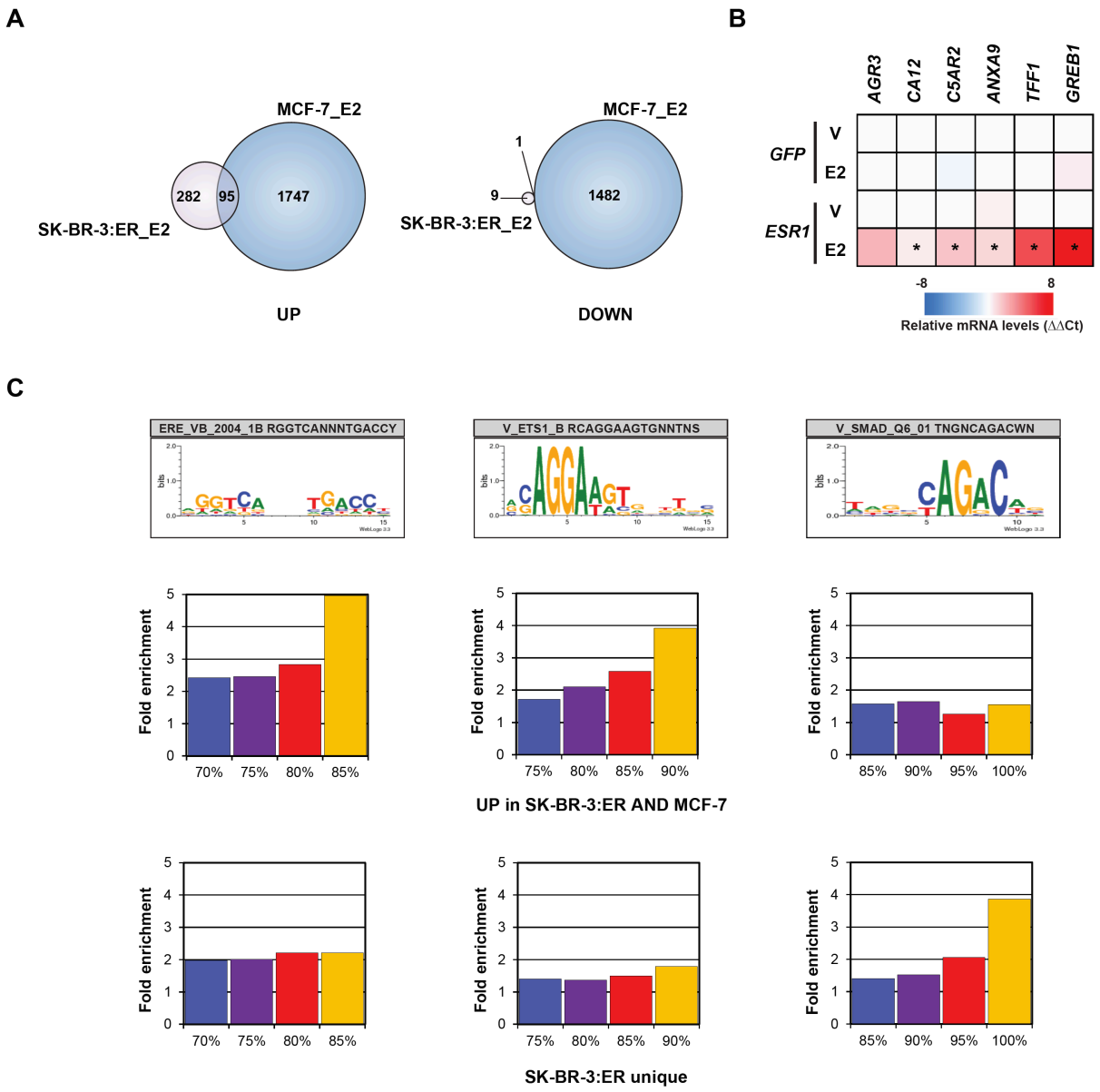


Figure 5. Ectopic ERα expression in SK-BR-3 cells restores part of the E2 transcriptional response observed in luminal MCF-7 cells. (A) Transcriptome profiling (RNA-Seq) was performed in MCF-7 cells treated with E2 (25nM) (16 h, N=3) and SK-BR-3 cells transiently expressing ERα (SK-BR-3:ER) and treated with E2 (25nM) (24 h, N=3). Venn diagrams showing the overlap of upregulated/downregulated genes (FC>1.4 or <-1.4 and q<0.05; compared to vehicle control) in MCF-7 and SK-BR-3:ER cells. (B) Relative mRNA levels of E2-target genes in SK-BR-3:ER cells treated with E2 (25nM) were assessed by RT-qPCR and represented as a heatmap. Asterisks indicate statistically significant regulations (p<0.05, N=3, student's paired t-test). (C) TF binding site enrichment analysis is shown at different matrix cutoffs in commonly upregulated genes between MCF-7 and SK-BR-3:ER cells or genes uniquely upregulated in SK-BR-3:ER cells.

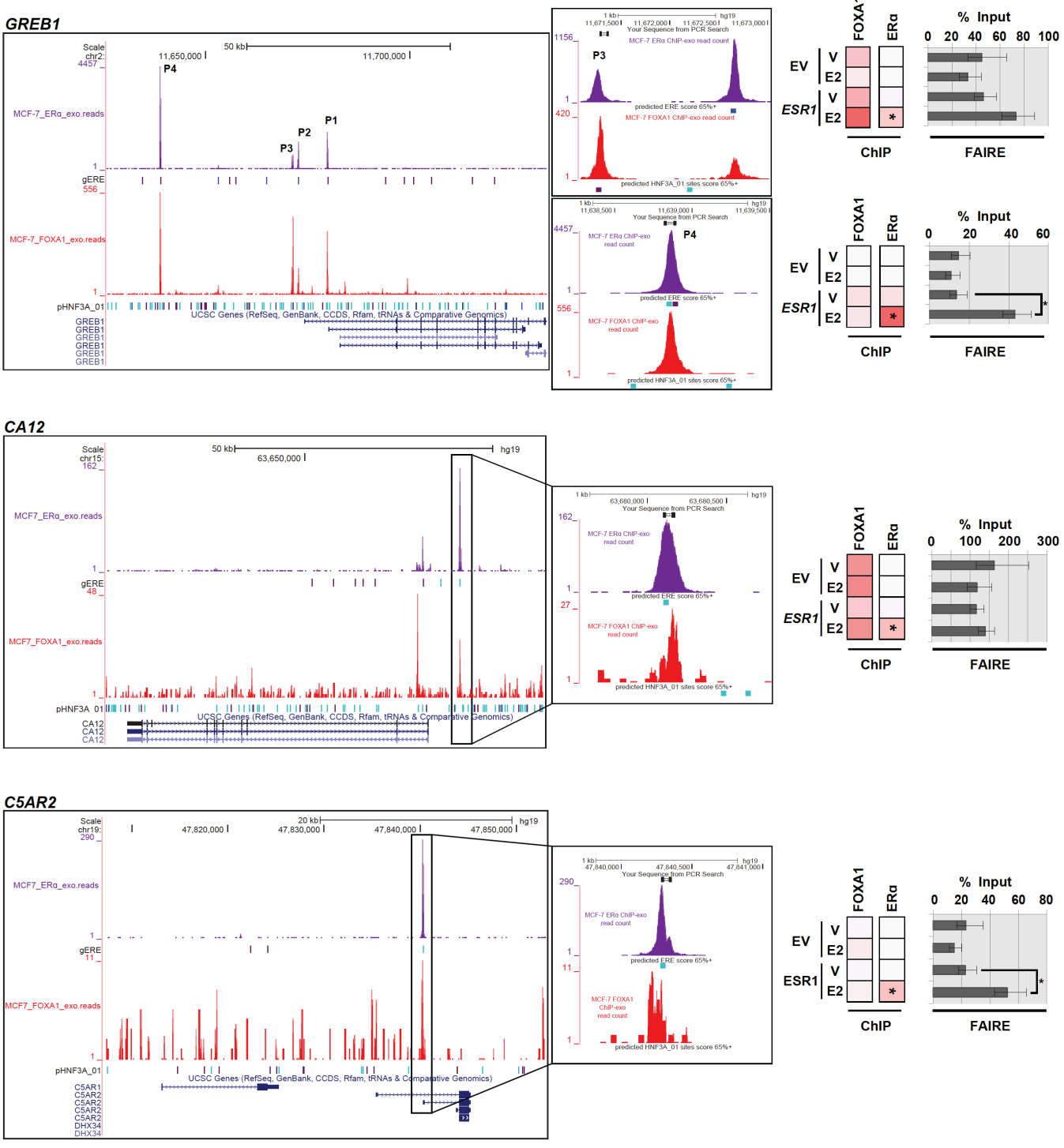

GREB1
Scale
chr2:
50 kb
11,650,000
11,700,000
P4
P1
P2
P3
MCF-7_ERα_exo.reads
gERE
MCF-7_FOXA1_exo.reads
pHNF3A_01
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
GREB1
Your Sequence from PCR Search
MCF-7 ERα ChIP-exo read count
predicted ERE score 65%+
MCF-7 FOXA1 ChIP-exo read count
predicted HNF3A_01 sites score 65%+
FOXA1
ERα
EV
ESR1
V
E2
ChIP
% Input
FAIRE
CA12
chr15:
63,650,000
hg19
CA12
63,680,000
63,680,500
C5AR2
chr19:
20 kb
47,820,000
47,830,000
47,840,000
47,850,000
47,840,500
ESR1

Figure 6. Interplay between ERα and FOXA1 binding in SK-BR-3:ER cells.
UCSC Genome Browser views of ERα and FOXA1 consensus sites as well as occupancy over promoter/enhancer regions of *GREB1*, *CA12*, and *C5AR2* in MCF-7 cells are shown; peaks probed are shown in close-ups. FOXA1 and ERα ChIP-qPCR was performed in SK-BR-3 cells transiently transfected with an ERα expression vector (*ESR1*) or with an empty vector (EV) and treated with vehicle (V) or E2 (25nM) for 24 h. Binding across sites shown is represented in a heatmap as the mean of fold enrichment vs IgG from three independent experiments. FAIRE was performed on the same sites probed in ChIP in SK-BR-3 cells stably expressing HA-ERα and treated with E2 (25nM) for 24 h. FAIRE enrichment is shown as the mean of percentage of input values from three independent experiments, and are represented as bar-graph. Error bars represent mean ± SEM. Asterisks denote statistical significance for differential intensities of peaks in *ESR1*_E2 vs *ESR1*_V ($p<0.05$ in a Student's paired t-test).

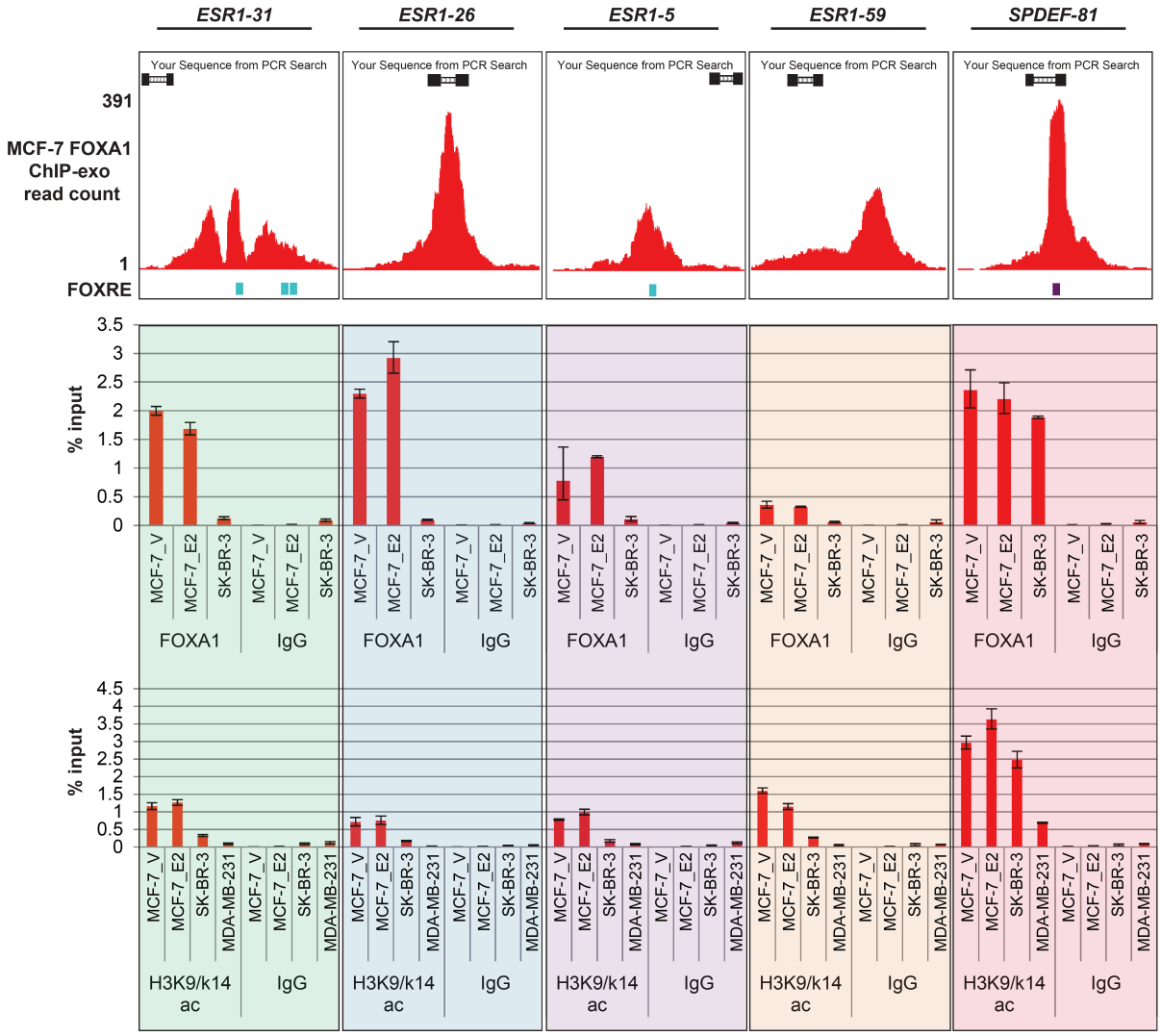


Figure 7. SK-BR-3 cells exhibit reduced FOXA1 binding to *ESR1* regulatory regions. FOXA1 and H3K9/K14Ac association with *ESR1* enhancers and the *SPDEF* promoter region was measured using ChIP-qPCR in MCF-7 (with or without E2 25nM for 45 min), SK-BR-3 and MDA-MB-231 cells. Enrichment is represented as percentage of input of one representative experiment (N=3). Error bars represent standard deviation of the mean of three technical replicates.

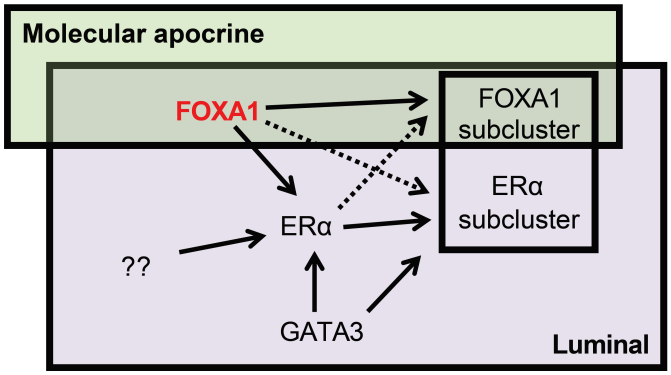
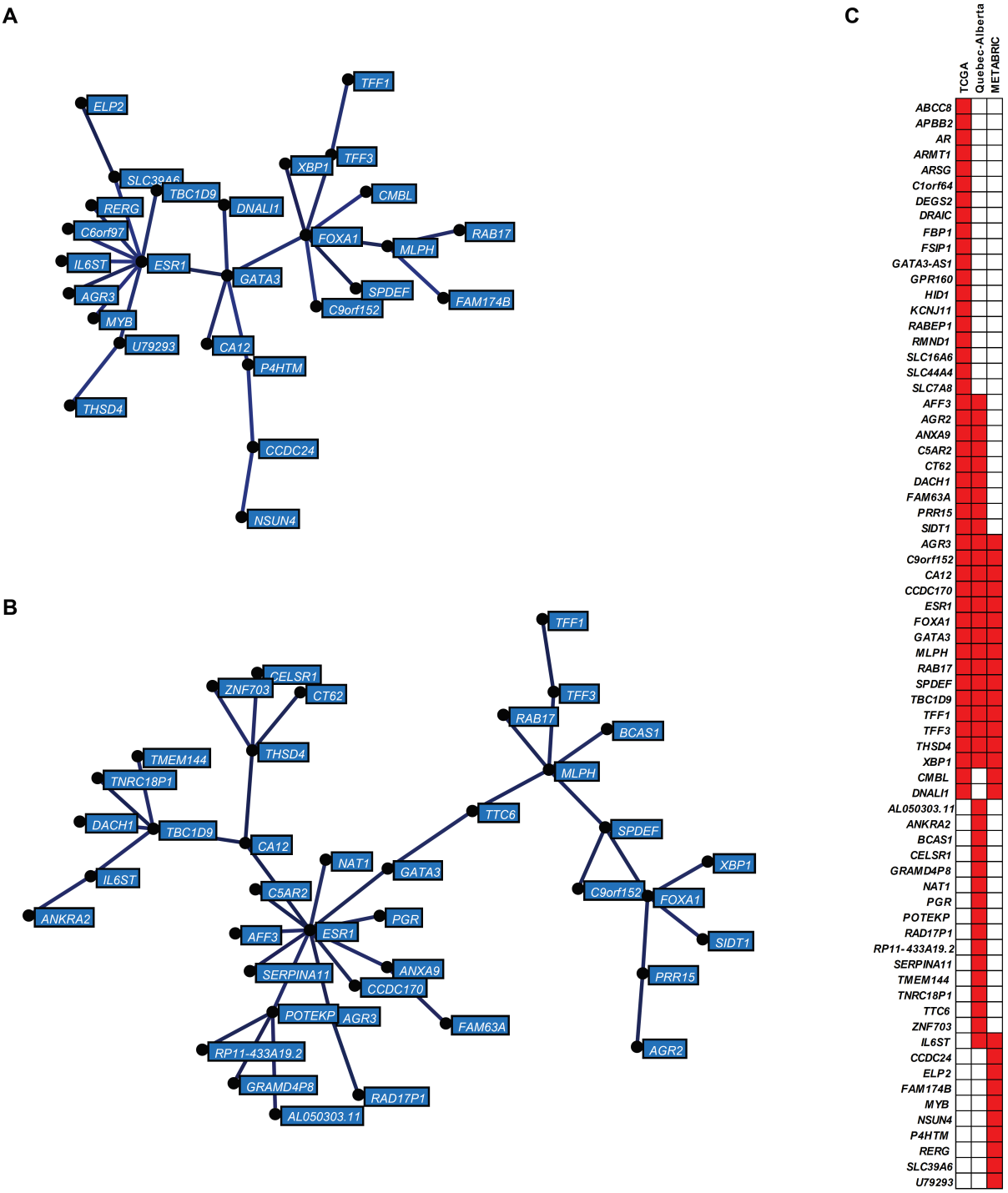
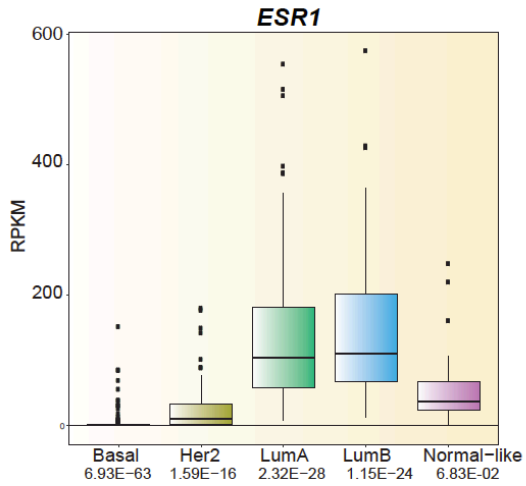
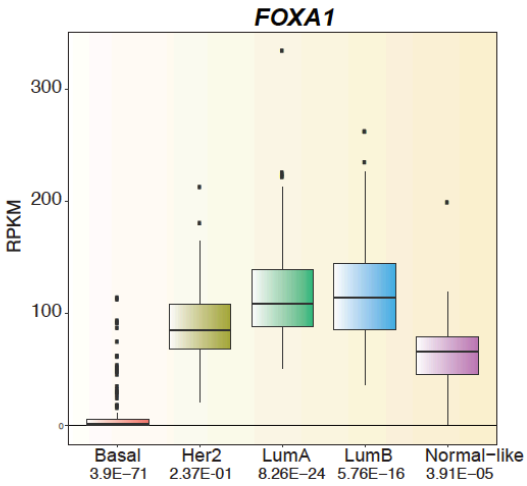


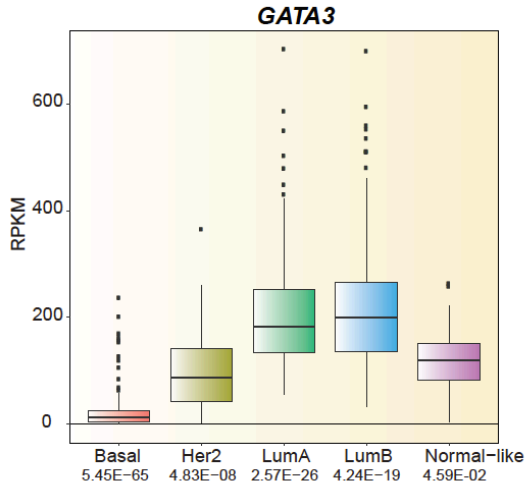
Figure 8. Role of FOXA1 as a master regulator in molecular apocrine and luminal breast tumour subtypes. mRNA levels of *FOXA1*, *ESR1*, *GATA3*, as well as those of their target genes identified herein represent markers for a rationale, TF-based tumour subtype classification.

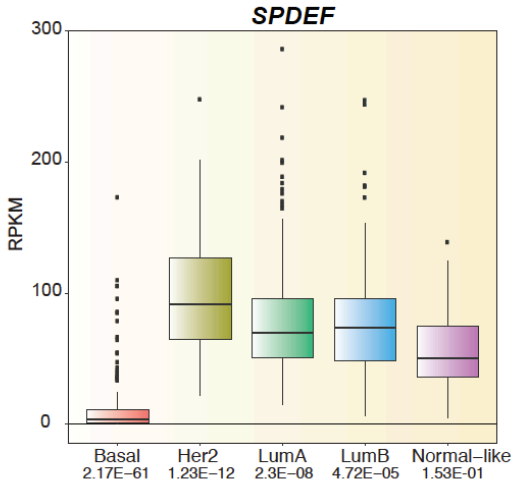


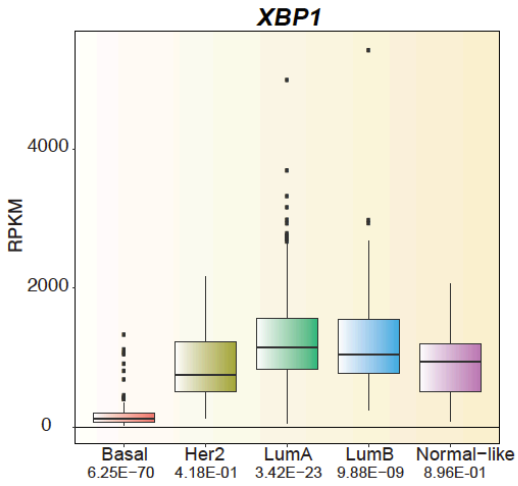
Supplementary Figure 1. The 'luminal' cluster in the METABRIC and Quebec-Alberta datasets. The minimal spanning tree containing genes whose expression is highly correlated with *ESR1* is shown in both the (A) METABRIC (27 genes) and (B) Quebec-Alberta (40 genes) datasets. Genes whose expression is highly correlated with that of *ESR1* in the TCGA, METABRIC and Quebec-Alberta datasets are listed in a table in (C). White boxes indicate absence of a gene in the *ESR1* cluster, whereas red boxes indicate its presence.

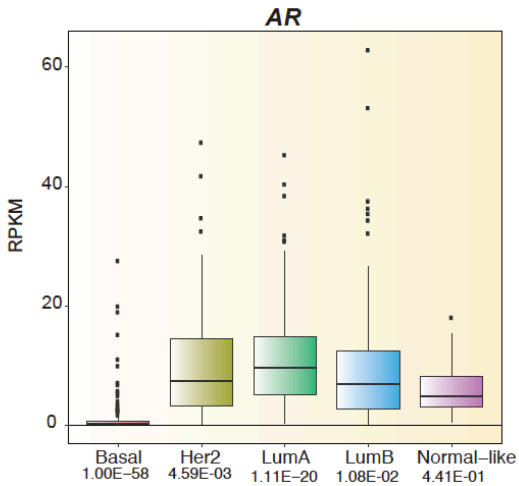

ESR1
RPKM
600
400
200
0
Basal
6.93E−63
Her2
1.59E−16
LumA
2.32E−28
LumB
1.15E−24
Normal−like
6.83E−02


FOXA1
RPKM
300
200
100
0
Basal
3.9E−71
Her2
2.37E−01
LumA
8.26E−24
LumB
5.76E−16
Normal−like
3.91E−05

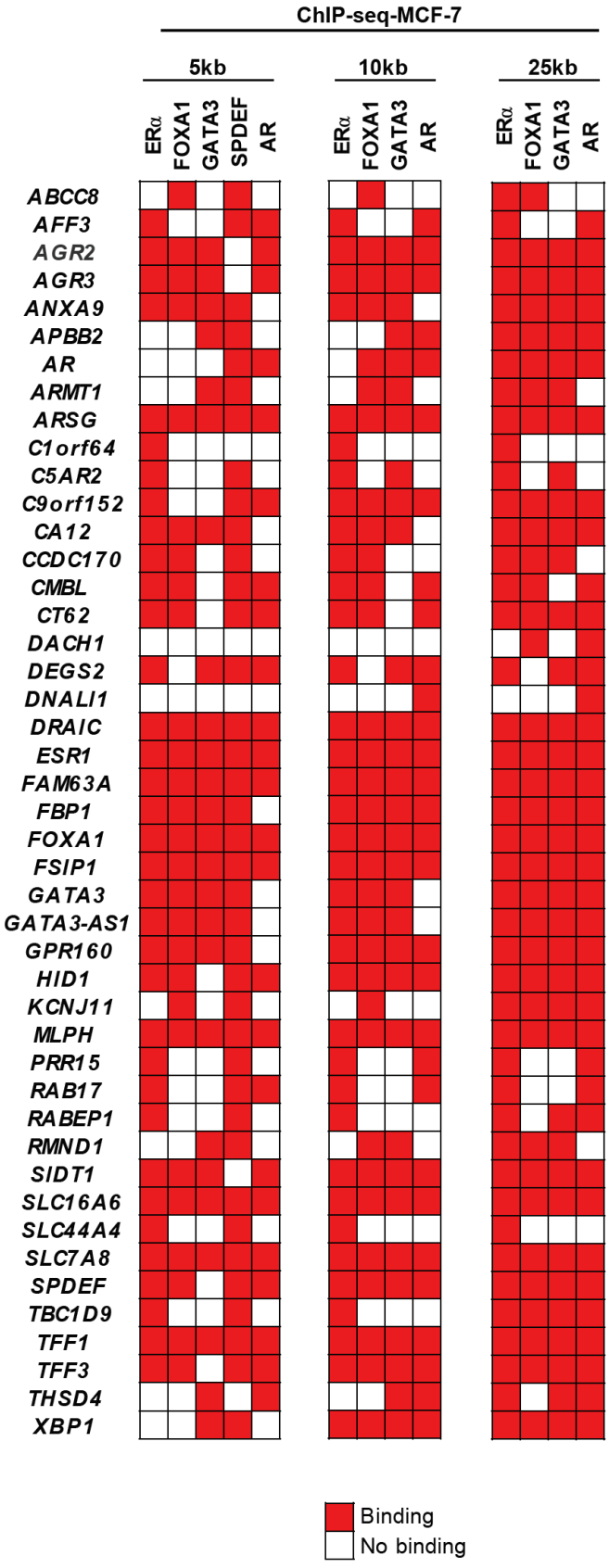

GATA3
RPKM
600
400
200
0
Basal
5.45E−65
Her2
4.83E−08
LumA
2.57E−26
LumB
4.24E−19
Normal−like
4.59E−02


SPDEF
RPKM
300
200
100
0
Basal
2.17E−61
Her2
1.23E−12
LumA
2.3E−08
LumB
4.72E−05
Normal−like
1.53E−01

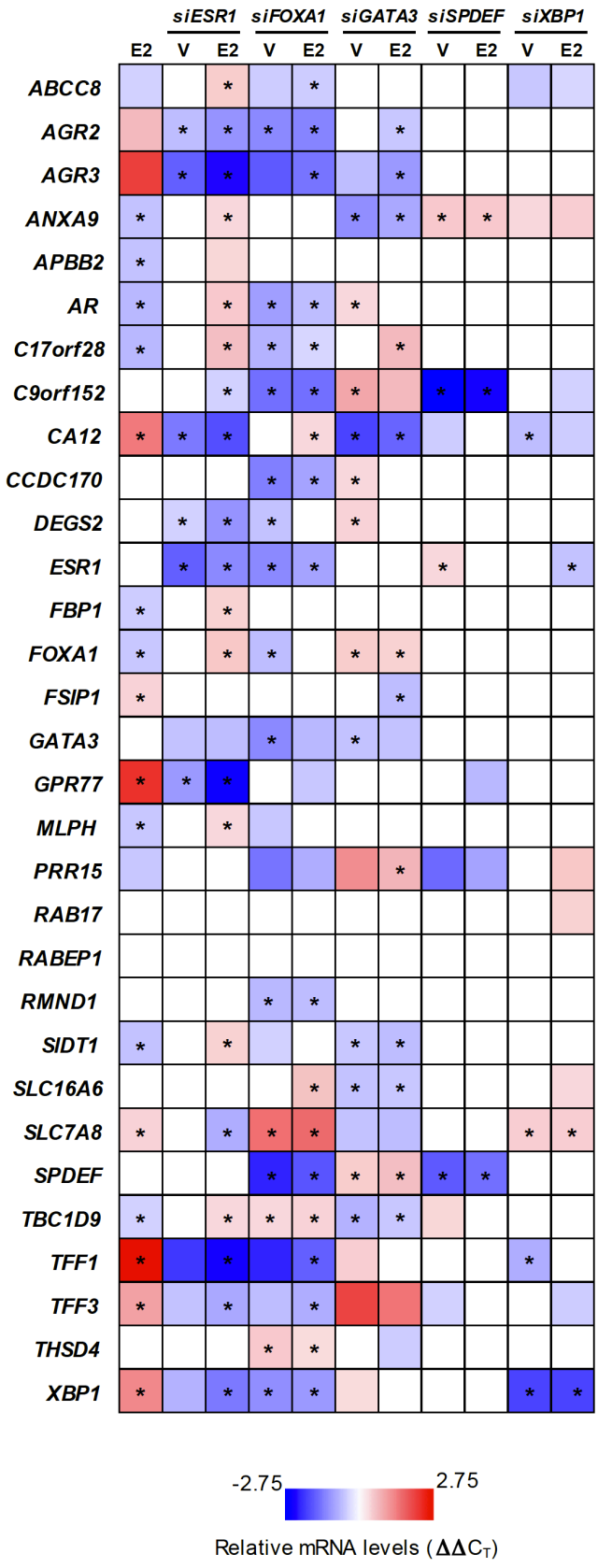

XBP1
RPKM
4000
2000
0
Basal
6.25E−70
Her2
4.18E−01
LumA
3.42E−23
LumB
9.88E−09
Normal−like
8.96E−01


AR
RPKM
60
40
20
0
Basal
1.00E−58
Her2
4.59E−03
LumA
1.11E−20
LumB
1.08E−02
Normal−like
4.41E−01

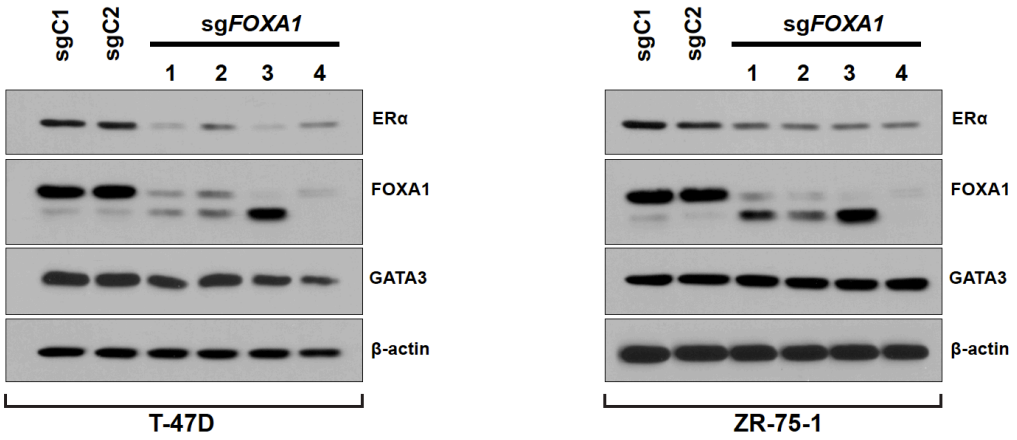
Supplementary Figure 2. *ESR1* cluster genes are highly expressed in luminal tumours. Expression levels (RPKM) of *ESR1, FOXA1, GATA3, SPDEF, XBP1*, and *AR* from the *ESR1* cluster are shown across breast tumours samples representing the basal, HER2-positive, luminal A, luminal B, and normal-like subtypes identified using the PAM50 gene signature classifier. P-values for each boxplot were generated using a Mann–Whitney U test.



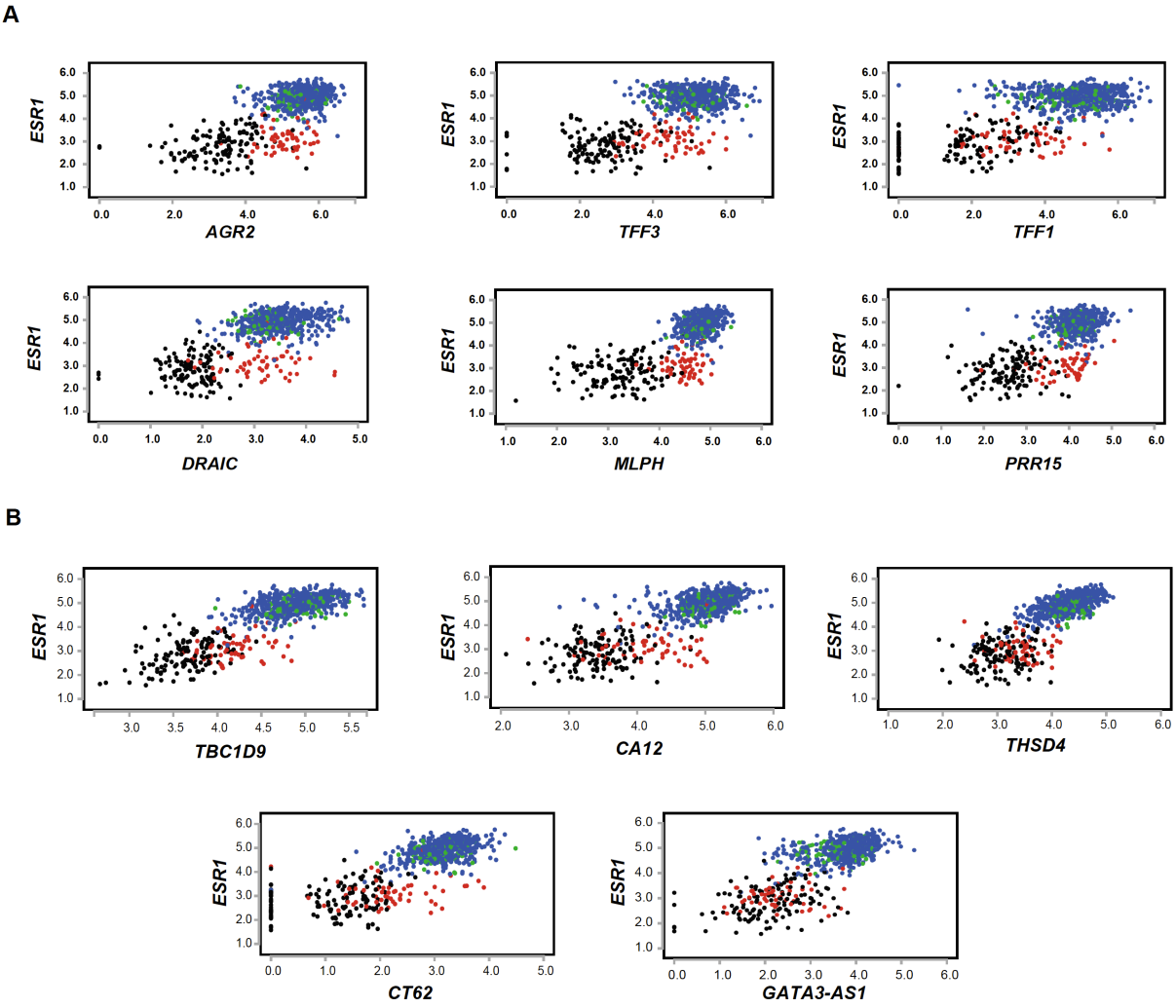
Supplementary Figure 3. The 'luminal' cluster genes are enriched in ChIP binding regions of luminal TFs in MCF-7 cells. ChIP-Seq binding of ERα, FOXA1, GATA3, SPDEF and AR is shown across the flanking regulatory regions (5kb, 10kb, or 25kb window) of genes in the *ESR1* cluster in the TCGA dataset.



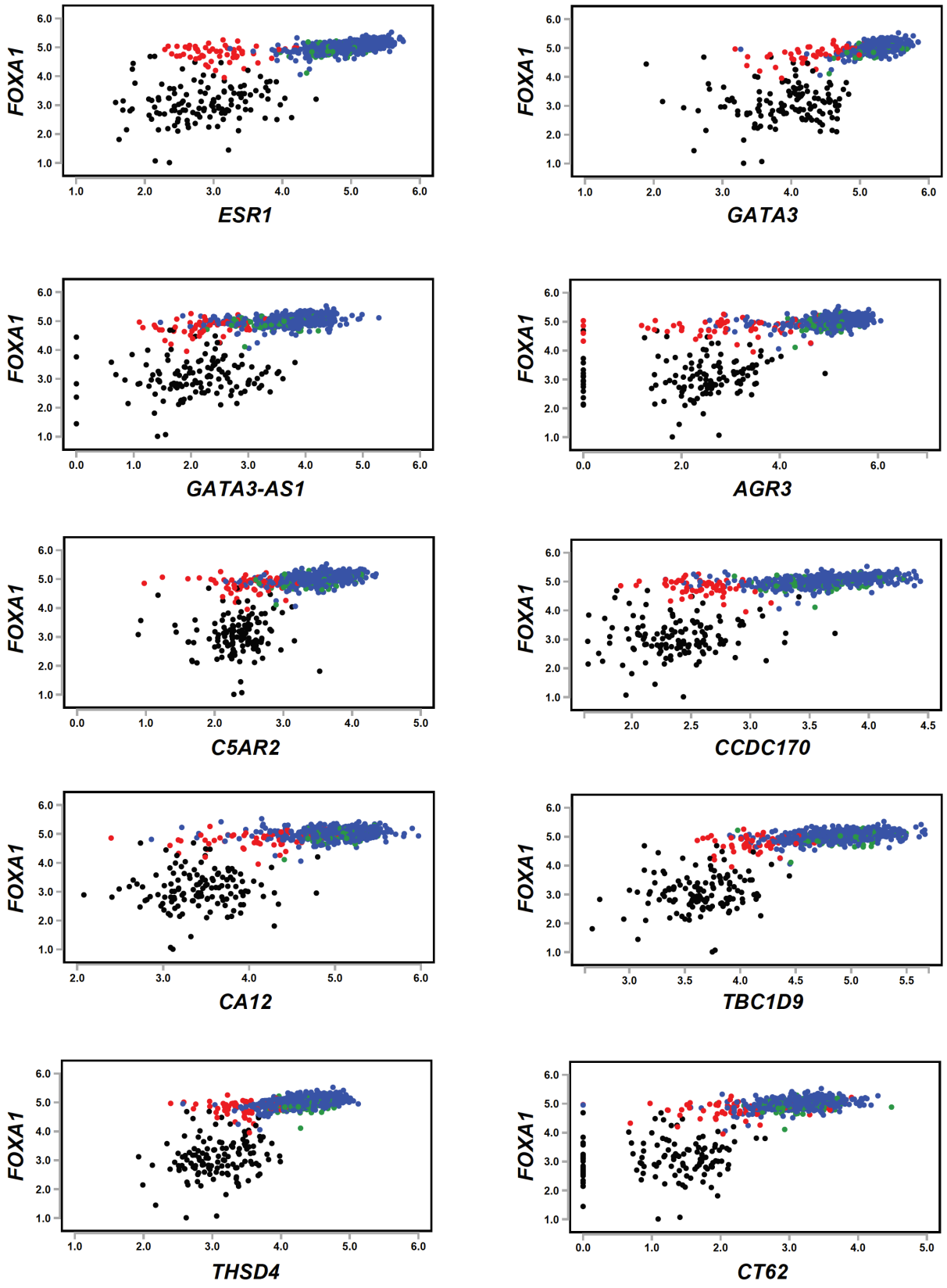
Supplementary Figure 4. The 'luminal' cluster is mainly controlled by ERα, FOXA1 and GATA3.
MCF-7 cells were transfected with two different siRNAs targeting *ESR1, FOXA1, GATA3*, *SPDEF*, or *XBP1* for 72 h and treated with or without E2 (25nM) for the last 24 h of transfection. Relative mRNA levels of a panel of genes in the luminal cluster were measured using RT-qPCR and presented as a heatmap. Upregulation is represented in red and downregulation is represented in blue. White indicates absence of regulation (ratio of siRNA to sicontrol is lower than 1.3 or greater than -1.3). Asterisks indicate statistically significant regulations from three independent experiments ($p<0.05$ in a Student's paired t-test).



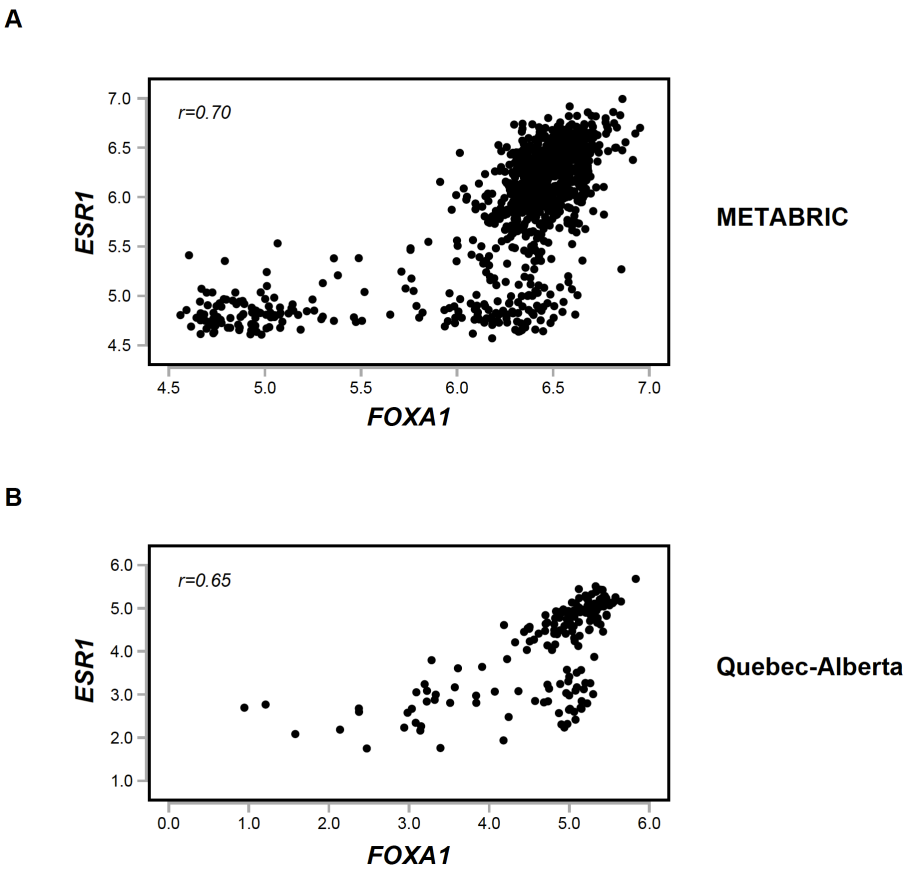
Supplementary Figure 5. SgRNA-mediated depletion of *FOXA1* results in downregulation of ERα expression. T-47D and ZR-75-1 cells were infected with sgRNAs targeting *FOXA1* in hormone-replete media (N=2). Infected cells were selected with puromycin for 7-10 days. Protein lysates were probed by western analysis using ERα, FOXA1, and GATA3 antibodies. β-actin was used as a loading control.



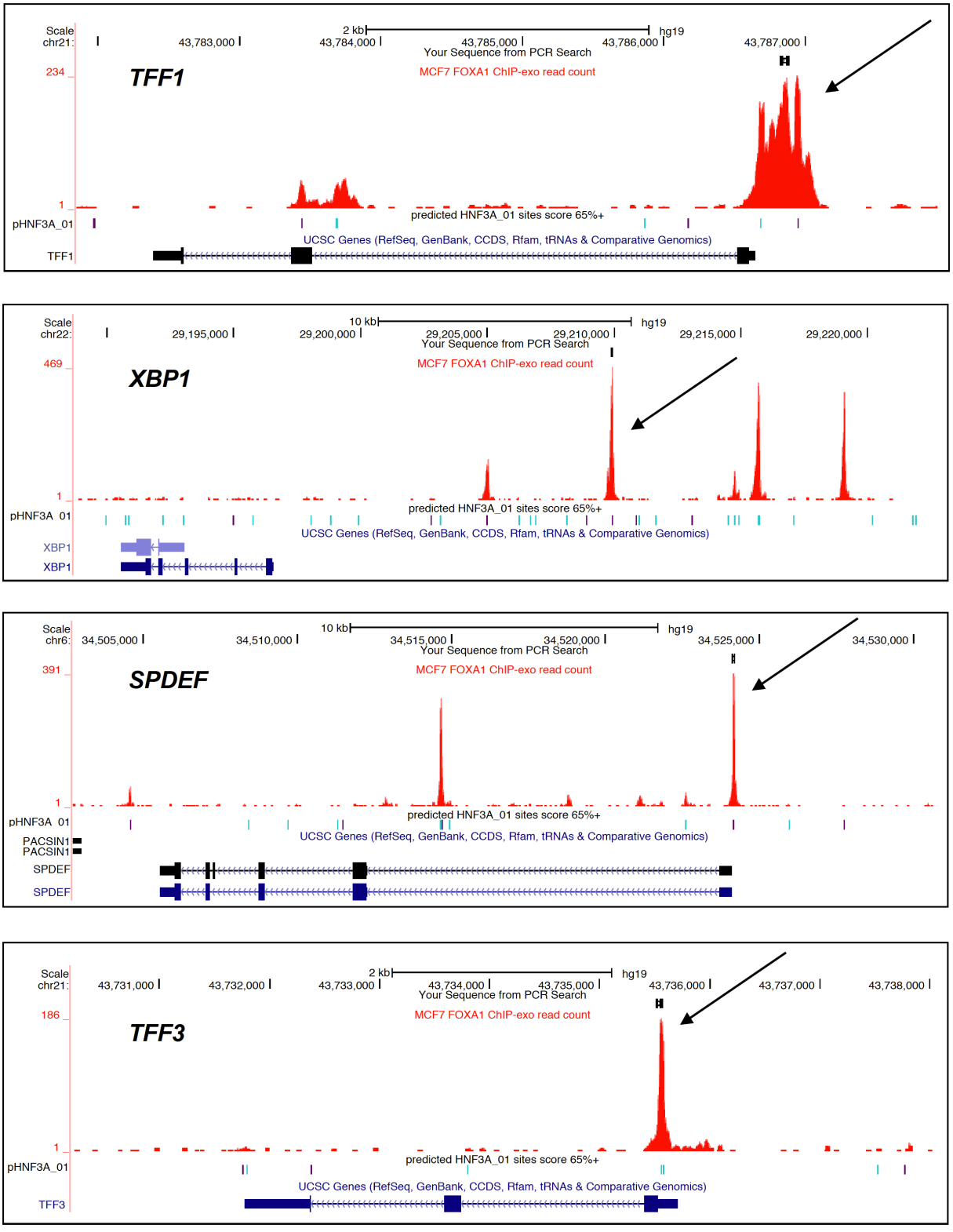
Supplementary Figure 6. Molecular apocrine tumours are *FOXA1* metagene^{high} and *ESR1* metagene^{low}. Tumour samples from the TCGA dataset were sorted according to expression levels of *ESR1* compared to *FOXA1* sub-cluster genes (*AGR2, TFF3, TFF1, DRAIC, MLPH,* and *PRR15*) as shown in (A) or *ESR1* sub-cluster genes *(TBC1D9, CA12, THSD4, CT62,* and *GATA3-AS1)* as shown in (B). Luminal tumours are identified in blue, normal-like in green, molecular apocrine in red and basal-like in black.


FOXA1
ESR1
GATA3
GATA3-AS1
AGR3
C5AR2
CCDC170
CA12
TBC1D9
THSD4
CT62

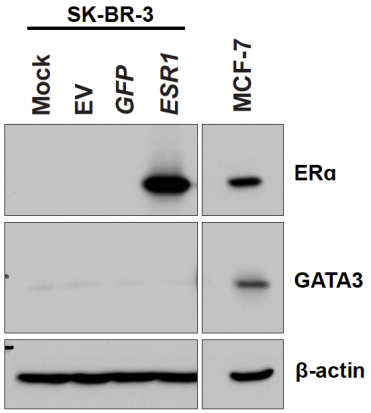
Supplementary Figure 7. Absence of *ESR1* sub-cluster^{high} *FOXA1*^{low} tumours in the TCGA dataset. Tumour samples from the TCGA dataset were sorted according to expression levels of *FOXA1* compared to *ESR1* sub-cluster genes (*ESR1, GATA3, GATA3-AS1, AGR3, C5AR2, CCDC170, CA12, TBC1D9, THSD4*, and *CT62*).



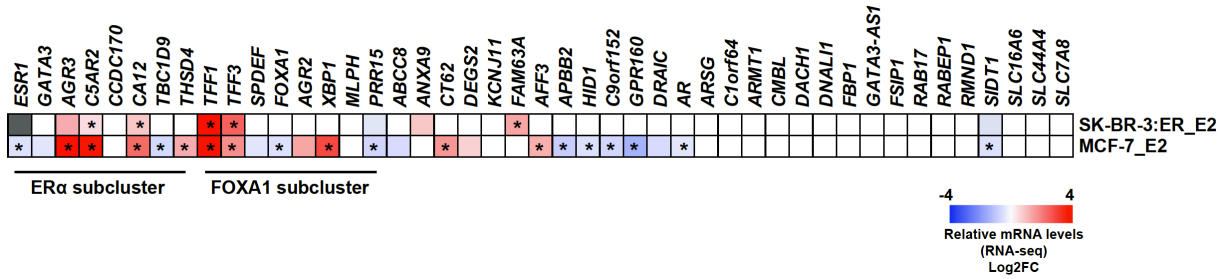
Supplementary Figure 8. *ESR1* and *FOXA1* expression is positively correlated in the METABRIC and Quebec-Alberta breast cancer datasets. Tumour samples from the METABRIC (A) and Quebec-Alberta (B) datasets were sorted based on expression levels of *ESR1* and *FOXA1*. A correlation coefficient *r* is indicated.



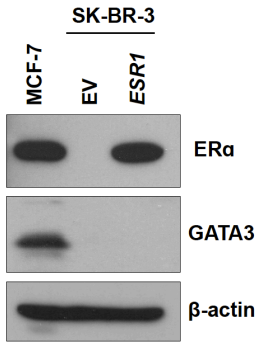
Supplementary Figure 9. UCSC Genome Browser views of FOXREs (pHNF3A_01) as well as occupancy at flanking regulatory regions of *TFF1, XBP1, SPDEF*, and *TFF3.*



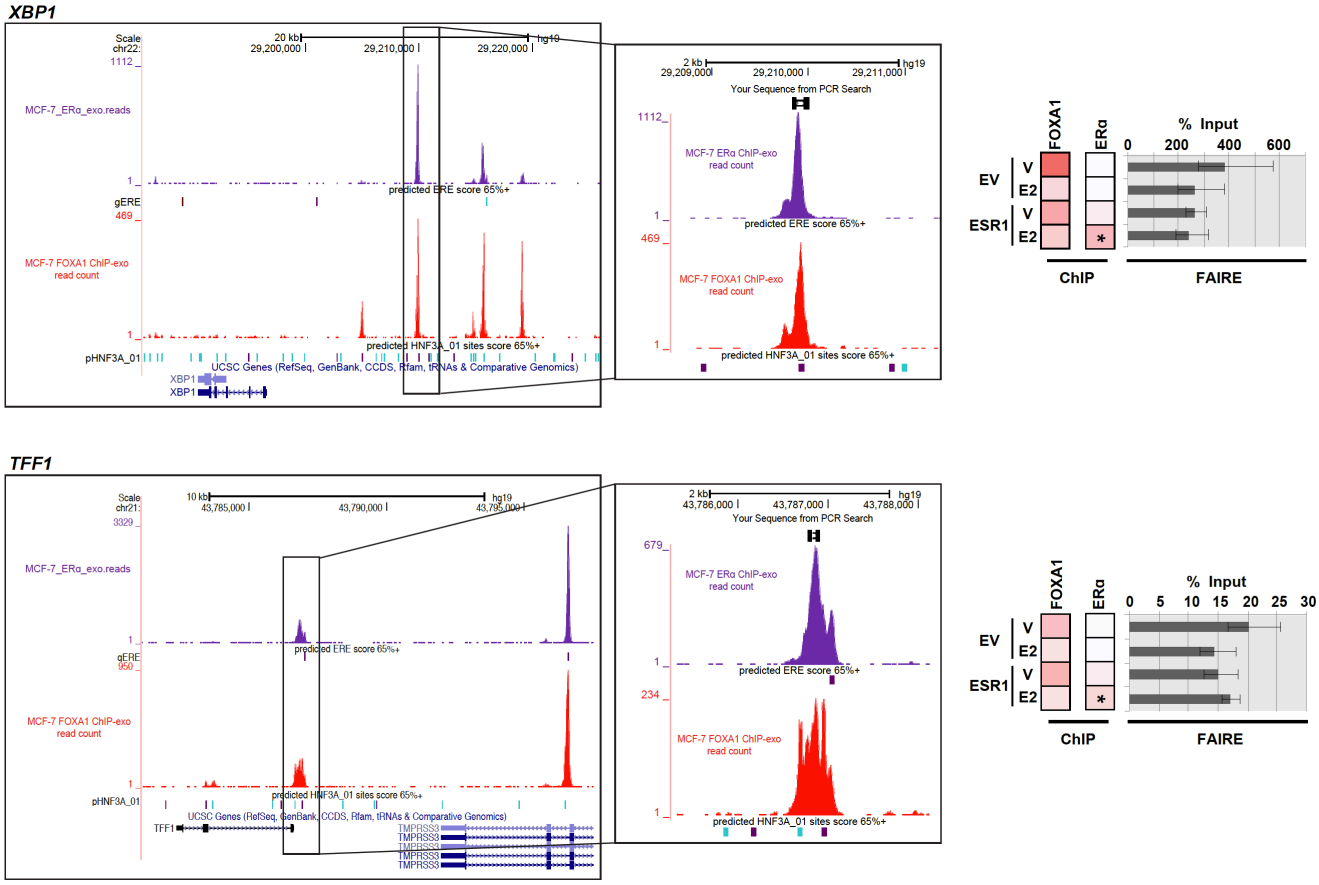
Supplementary Figure 10. Transient ERα expression in SK-BR-3 cells. Protein levels of ERα and GATA3 were assessed following transient ectopic expression of ERα in SK-BR-3 cells. MCF-7 cell lysates are immunoblotted as a control for comparison.



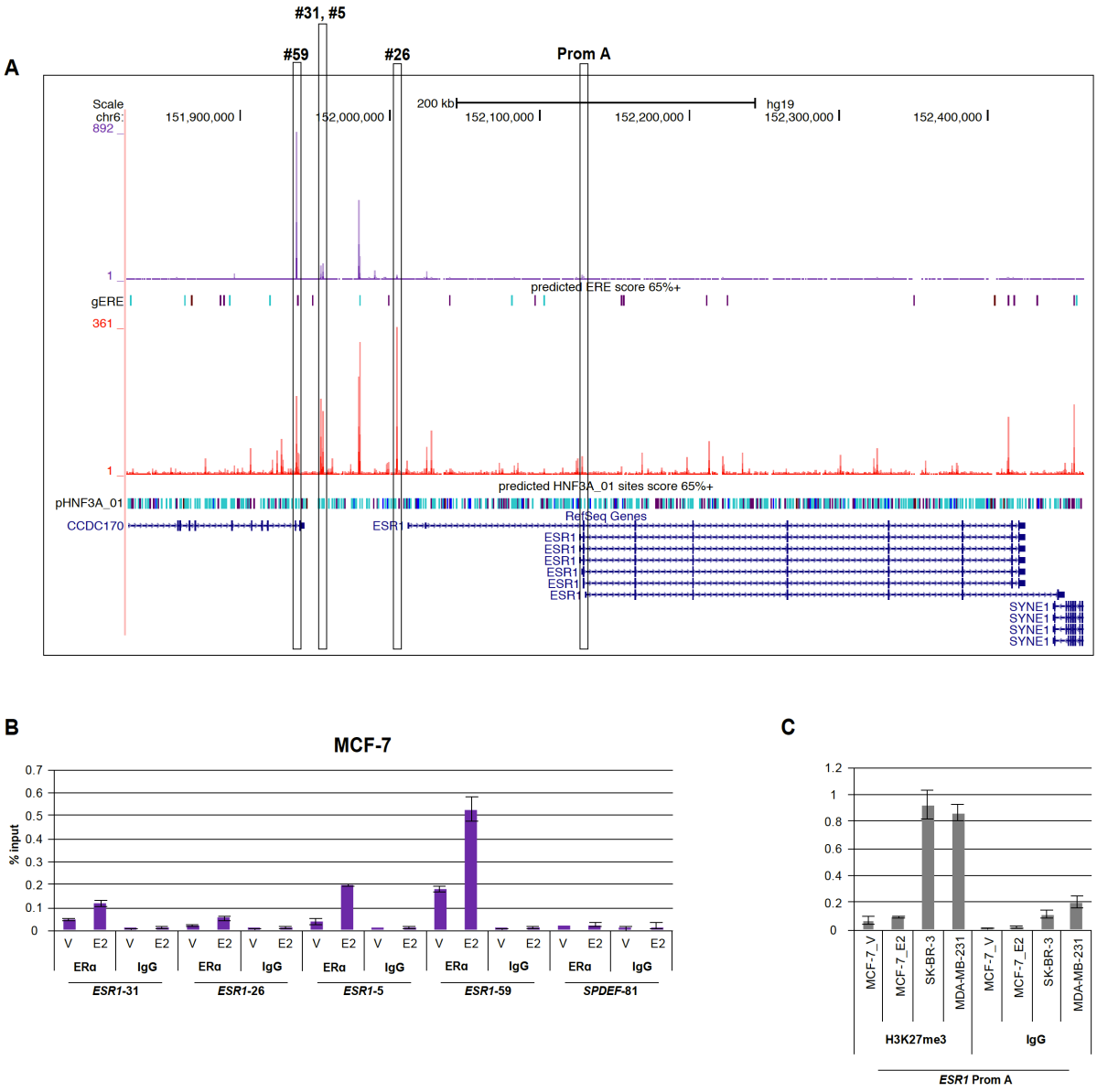
Supplementary Figure 11. Regulation of luminal cluster genes in SK-BR-3:ER cells. Relative expression levels of luminal cluster genes are derived from transcriptome analysis of SK-BR-3_ER cells treated with E2 (25nM, 24 h) (vs SK-BR-3_ER cells treated with vehicle) and MCF7 cells in the presence of E2 (25nM, 16 h) (vs absence). Asterisks denote statistically significant regulations (q<0.05). White boxes indicate absence of significant regulation (-1.4<FC<1.4).



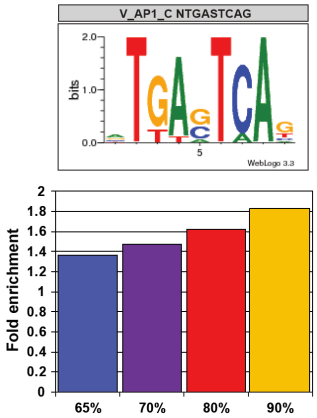
Supplementary Figure 12. Stable ERα expression in SK-BR-3 cells. Protein expression levels of ERα and GATA3 were assessed in SK-BR-3 cells stably expressing HA-ERα by western blots.



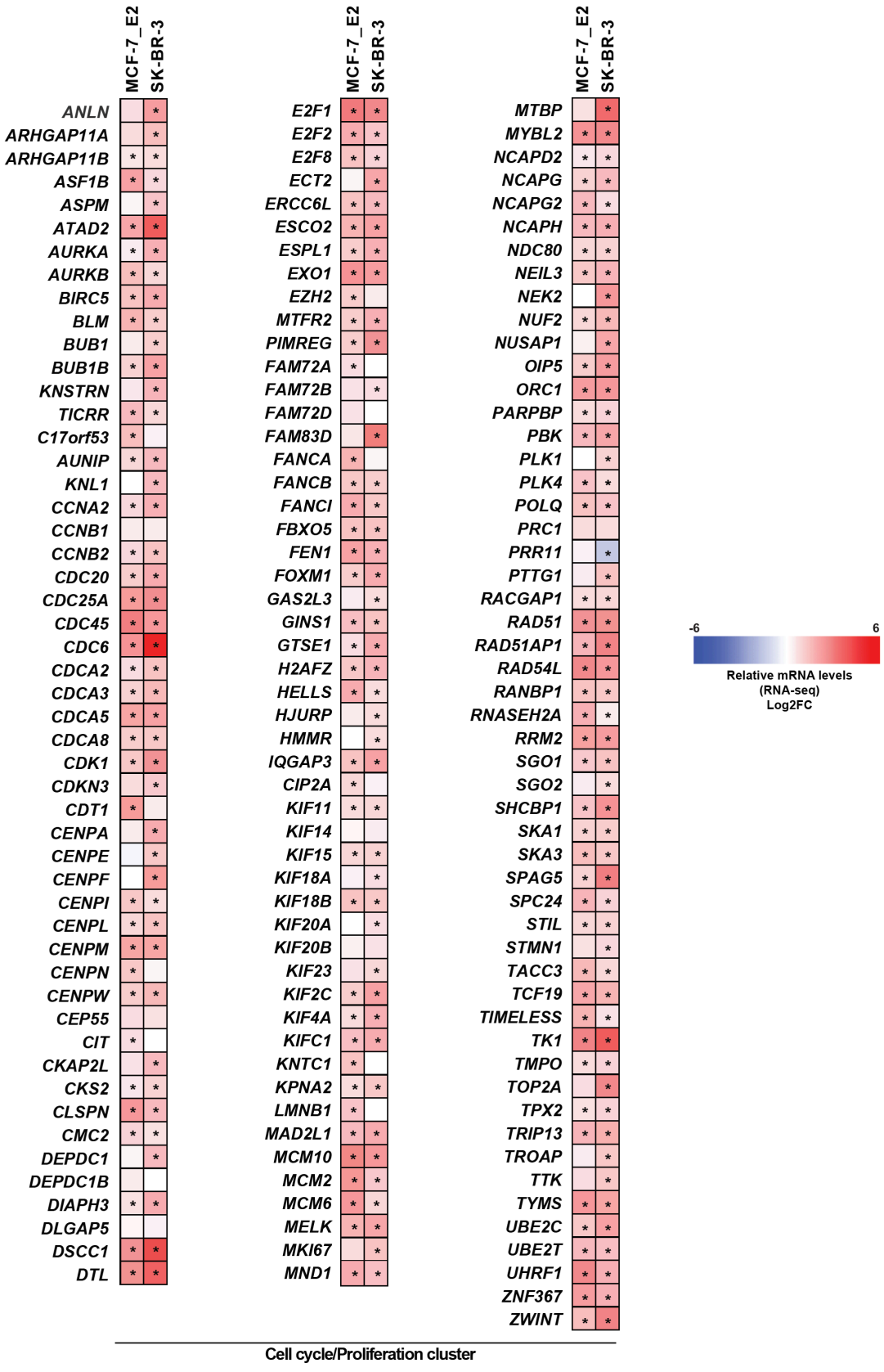
Supplementary Figure 13. FOXA1 pre-occupancy coincides with lack of further chromatin opening by ERα at *XBP1* and *TFF1* upstream regulatory regions. UCSC genome browser views of ERα and FOXA1 consensus sites as well as occupancy at the *XBP1* and *TFF1* regulatory regions in MCF-7 cells are shown. Peaks examined by ChIP-qPCR and FAIRE are shown in close-ups. ChIP-qPCR was performed in SK-BR-3 cells transiently expressing ERα and treated with E2 (25nM) for 24 h. Results are shown as a heatmap of the mean of folds of enrichment over IgG (N=3; * $p<0.05$ of *ESR1*_E2 vs ESR1_V, Student's paired t-test). FAIRE was performed in SK-BR-3 cells stably expressing ERα and treated with E2 for 24 h. Enrichment is shown as percentage of input. (N=3, error bars represent mean ± SEM).



Supplementary Figure 14. ERα and H3K27me3 recruitment at *ESR1* upstream regulatory regions. UCSC genome browser view of ERα and FOXA1 occupancy at the *ESR1* gene in MCF-7 cells is shown in (A). Peaks examined in Figure 7 are shown in boxes with the corresponding RT-qPCR probe number. ERα recruitment was assessed by ChIP-qPCR in MCF-7 cells stimulated with E2 (25nM) for 45 mins and represented as percentage of input in (B). H3K27me3 recruitment was assessed in E2-stimulated MCF-7 as well as SK-BR-3 and MDA-MB-231 cells (C). N=3, one representative experiment is shown.



Supplementary Figure 15. AP1 consensus motifs are enriched in E2-target genes uniquely regulated in SK-BR-3_ER cells.


MCF-7_E2
SK-BR-3
ANLN
ARHGAP11A
ARHGAP11B
ASF1B
ASPM
ATAD2
AURKA
AURKB
BIRC5
BLM
BUB1
BUB1B
KNSTRN
TICRR
C17orf53
AUNIP
KNL1
CCNA2
CCNB1
CCNB2
CDC20
CDC25A
CDC45
CDC6
CDCA2
CDCA3
CDCA5
CDCA8
CDK1
CDKN3
CDT1
CENPA
CENPE
CENPF
CENPI
CENPL
CENPM
CENPN
CENPW
CEP55
CIT
CKAP2L
CKS2
CLSPN
CMC2
DEPDC1
DEPDC1B
DIAPH3
DLGAP5
DSCC1
DTL
E2F1
E2F2
E2F8
ECT2
ERCC6L
ESCO2
ESPL1
EXO1
EZH2
MTFR2
PIMREG
FAM72A
FAM72B
FAM72D
FAM83D
FANCA
FANCB
FANCI
FBXO5
FEN1
FOXM1
GAS2L3
GINS1
GTSE1
H2AFZ
HELLS
HJURP
HMMR
IQGAP3
CIP2A
KIF11
KIF14
KIF15
KIF18A
KIF18B
KIF20A
KIF20B
KIF23
KIF2C
KIF4A
KIFC1
KNTC1
KPNA2
LMNB1
MAD2L1
MCM10
MCM2
MCM6
MELK
MKI67
MND1
MTBP
MYBL2
NCAPD2
NCAPG
NCAPG2
NCAPH
NDC80
NEIL3
NEK2
NUF2
NUSAP1
OIP5
ORC1
PARPBP
PBK
PLK1
PLK4
POLQ
PRC1
PRR11
PTTG1
RACGAP1
RAD51
RAD51AP1
RAD54L
RANBP1
RNASEH2A
RRM2
SGO1
SGO2
SHCBP1
SKA1
SKA3
SPAG5
SPC24
STIL
STMN1
TACC3
TCF19
TIMELESS
TK1
TMPO
TOP2A
TPX2
TRIP13
TROAP
TTK
TYMS
UBE2C
UBE2T
UHRF1
ZNF367
ZWINT
-6
6
Relative mRNA levels
(RNA-seq)
Log2FC
Cell cycle/Proliferation cluster

Supplementary Figure 16. Expression of proliferation cluster in SK-BR-3 and MCF-7 cells. Relative expression levels of the cell cycle/proliferation cluster genes in the TCGA dataset of breast tumours were examined using the MCF-7 and SK-BR-3 transcriptome sequencing datasets. Differential gene expression was generated by comparing SK-BR-3 cells and E2-stimulated MCF-7 cells to vehicle-treated MCF-7 cells. Asterisks denote statistically significant regulations ($q<0.05$).

Supplementary Table 1. SiRNA oligonucleotide sequences used in this study

Gene	Sequence
Scrambled	UGGUUUACAUGUCGACUAA
ESR1#1	GAAUGUGCCUGGCUAGAGAUU
ESR1#2	CAUGAGAGCUGCCAACCUUUU
FOXA1#1	GCACUGCAAUACUCGCCUUUU
FOXA1#2	CCUAAACACUUCCUAGCUCUU
GATA3#1	AAGCCUAAACGCGAUGGAUAUUU
GATA3#2	AACAUCGACGGUCAAGGCAACUU
SPDEF#1	GAGGAGAGCUGGACCGACAUU
SPDEF#2	UUAGGUGGCUCAACAAGGAUU
AR#1	AAGAAGGCCAGUUGUAUGGACUU
AR#2	AAGACGCUUCUACCAGCUCACUU
XBP1#1	CGAAAGAAGGCUCGAAUGAUU
XBP1#2	CAACUUGGACCCAGUCAUGUU

Supplementary Table 2. RT-qPCR primer sequences

Gene	UPL probe number	Sequence
RPLP0 (Reference)	74	5'-TCCCACTTGCTGAAAAGGTC-3'
		5'-AGCAGGAGCAGCTGTGGT-3'
YWHAZ (Reference)	2	5'-GCAATTACTGAGAGACAACTTGACA-3'
		5'-TGGAAGGCCGGTTAATTTT-3'
TBP (Reference)	87	5'-GAACATCATGGATCAGAACAACA-3'
		5'-ATAGGGATTCCGGGAGTCAT-3'
ABCC8	78	5'-GTGGCTCTGTGGCTTTTCTC-3'
		5'-AGACCCTCATGAACCGACAG-3'
AGR2	47	5'-GTAGGAGAGGGCCACAAGG-3'
		5'-GGTGGGTGAGGAAATCCAG-3'
AGR3	11	5'-AGGCTCATATGTGTACAATCTGTTAGA-3'
		5'-TGGGCAATATGTGCCTAGAA-3'
ANXA9	39	5'-CTGCAAGATGCCACTGGTC-3'
		5'-TGCCTGGCAGTCTACAAACA-3'

APBB2	27	5'-GTCGCTATGGGCTCACTGTT-3'
		5'-GCAATGGTGCAGGAAATAGG-3'
AR	14	5'-GCCTTGCTCTCTAGCCTCAA-3'
		5'-GTCGTCCACGTGTAAGTTGC-3'
C17orf28	1	5'-CCAGCCTCCTAACCCAAAG-3'
		5'-AGCCTGGCCACACCTTTT-3'
C9orf152	10	5'-GCCCAGTATGAAGGCTTGAA-3'
		5'-CAGGAGCAGGTGTGTTTCCT-3'
CA12	11	5'-GCTCCTGCTGGTGATCTTAAA-3'
		5'-CCAAAATAAGTCCACTTGGAACC-3'
CCDC170	10	5'-TCACGTTCAAGCCTAGCATC-3'
		5'-GAGAAAAGCAGCTGGCAGAC-3'
DEGS2	28	5'-TTGATGGCCGGGTACTTG-3'
		5'-GCGACTTCGAGTGGGTCTAC-3'
ESR1	24	5'-TTACTGACCAACCTGGCAGA-3'
		5'-ATCATGGAGGGTCAAATCCA-3'
FBP1	6	5'-ACCCTGCCGTCACTGAGTA-3'
		5'-GCCCCATAAGGAGCTGAAT-3'
FOXA1	47	5'-ATCATTGCCATCGTGTGCTT-3'
		5'-CACCATGTCCAACTGTGGAA-3'
FSIP1	2	5'-TCCAAGACTTGAAAATCGGAAT-3'
		5'-CGTTGCTCTTTGGTGTTCCT-3'
GATA3	36	5'-ACTACGGAAACTCGGTCAGG-3'
		5'-GGTAGGGATCCATGAAGCAG-3'
GPR77	25	5'-AAGAAACCGGATGGCAGTC-3'
		5'-CAGTGTGTGGTGGACTACGG-3'
MLPH	67	5'-GGCCTCCTCCTCTACATCG-3'
		5'-GGGCGTCTTCTGAGAGTCA-3'
PRR15	89	5'-CCTGTTAAGCCCGAGCAG-3'
		5'-CCCTTTCTCCACGTGGTCT-3'
RAB17	77	5'-CGACTTCAAGAGTATCCTGCCTA-3'
		5'-CAGCTGTGTCCCAGATCTCA-3'
RABEP1	88	5'-TGTGTCATTACAGCAAGCAGAA-3'
		5'-TTAATACCAACTCCCGCAGTG-3'
RMND1	5	5'-GCGCTTCCTTCTTCTCTTCC-3'

		5'-CTTCGGCACTGATGTGCTT-3'
SIDT1	71	5'-AGCCCAAGTCCTTAAGGTGAG-3'
		5'-CCATACTAAAAGACTGGTAACAGCAG-3'
SLC16A6	20	5'-GCTGGGGTCTACATCTTCATTC-3'
		5'-TTTGGTCCACCAACAAACCT-3'
SLC7A8	4	5'-CTCTGACCACAGGCTGAAGA-3'
		5'-CGCTGGAAGAAGCCTGATA-3'
SPDEF	26	5'-CTCAGCTTGTCGTAGTTCATGG-3'
		5'-AGGTGGCTCAACAAGGAGAA-3'
TBC1D9	37	5'-GAAAAGTGCACGGTATCTCTCA-3'
		5'-AAAAAGTGTCTGCTCTAAAACGTG-3'
TFF1	66	5'-ACCATGGAGACAAGGTGAT-3'
		5'-AAATTCACACTCCTCTTCTG-3'
TFF3	4	5'-TGGAGGTGCCTCAGAAGGT-3'
		5'-GCTGCTGCTTTGACTCCAG-3'
THSD4	19	5'-GCAGGAAGGACTTCATGAGAA-3'
		5'-GAGCTCTCCTTGTGCCCTAGT-3'
XBP1	62	5'-CCCTGGTTGCTGAAGAGG-3'
		5'-TGGAGGGGTGACAACTGG-3'

Supplementary Table 3. ChIP/FAIRE primer sequences

Gene	Probe number	Sequence
TFF1 enhancer	23	TCTTCTCCACGCCCTGTAAA
		GACAGCTGCCAGGTACGG
TFF1 promoter	30	GACGACATGTGGTGAGGTCAT
		CCATGGGAAAGAGGGACTTT
XBP1	85	TGCCAAGTATCTGCCCTGTT
		TCTGTGTGCTCAGCAAATGTT
SPDEF	81	GTCAGACAGCCGCGAGAT
		AGCAAGGAGTGTATTTGTTCAGG
TFF3	79	CAGCTCAGGACTCGCTTCAT
		AGAGCAGCTGTGCAAACAAC
AGR3	8	CCAGACTCGATTAGGCTGACA

		GGAGCCCAGAGAATGGAGA
GREB1-P3	48	GCCTGGGGATATTTGCTTCT
		AGCTCAACATCTGCTTGTATGC
GREB1-P4	26	GAGAGGGTGGTGACACTTGG
		CCGTGCCTACCACAAGGT
CA12	65	AGTTTCAGGCGGGATCAGT
		CTCCTGGCAGAGTCAGAAGC
GPR77	51	GGCCAGGTCACTCTGACATT
		AGCTGTGGGTAGCACAGGAC
ESR1	59	GCAGAGATCAGTTGCAGGAA
		GGTCGGCCATCAAGTCTG
ESR1	26	TCTTTGGTCTTATCCTCTTATCTTCC
		AGATAATGACATCGGAGACAGTACA
ESR1-promF	75	CTGCATTTCCTAATTTCATGGTC
		AAAATCCGTTCTGAGTCGGTAG
ESR1	31	AAGAAACTGCCAAACTGTCTCC
		AACACCAAATGCTGCCAAG
ESR1	5	GGGCTAGTGTATCTTTCTGCTGAT
		TTAAGAGCAAGCGTCCTTGG
ESR1-promA	29	GCTGTGCTCTTTTTCCAGGT
		GGTGTGGAGGGTCATGGTC

Supplementary Table 4. sgRNA sequences used in this study

	Sequences
sgRNA-FOXA1#1	GATGTTAGGAACTGTGAAGA
sgRNA-FOXA1#2	GTAGTAGCTGTTCCAGTCGC
sgRNA-FOXA1#3	GAAGGGCATGAAACCAGCGAC
sgRNA-FOXA1#4	GTTCATGTTGCTGACCGGGA
sgRNA-FOXA1#5	GTTGGACGGCGCGTACGCCA
sgRNA-ctrl#1	GTAGCGAACGTGTCCGGCGT
sgRNA-ctrl#2	GACCGGAACGATCTCGCGTA

Supplementary Table 5. Luminal gene signatures enriched in the luminal TCGA cluster

P-value	Q-value	Odds	Name	Type
8.40E-35	1.44E-32	84.51	CHARAFE_BREAST_CANCER_LUMINAL_VS_BASAL_UP	msigdb
3.98E-24	2.39E-22	44.44	CHARAFE_BREAST_CANCER_LUMINAL_VS_MESENCHYMAL_UP	msigdb
4.21E-23	2.19E-21	78.02	SMID_BREAST_CANCER_LUMINAL_B_UP	msigdb
6.79E-11	1.40E-09	45.1	upegulated in the mature luminal	microarrays
7.81E-11	1.56E-09	44.26	LIM_MAMMARY_LUMINAL_MATURE_UP	msigdb
5.33E-10	9.06E-09	173.44	Breast_Muggerud06_27genes_LuminalvNonLuminal	GeneSigDB
6.32E-10	1.07E-08	79.33	Breast Cancer Luminal A vs. Basal like - Sorlie et al., 2006	microarrays
9.16E-10	1.56E-08	74.15	Breast_Sorlie06_54genes_Luminal_v_Basal	GeneSigDB

Supplementary Table 6. Chromosomal locations and amplification status of the *ESR1* cluster genes in the TCGA dataset of breast tumours.

Gene name	Chromosome	Gene start (bp)	Gene end (bp)	Karyotype band	NUM_CASES_ AMPLIFIED	% CASES_ AMPLIFIED
C1orf64	1	16004236	16008807	p36.13	2	0.20%
ANXA9	1	150982017	150995634	q21.3	54	7%
FAM63A	1	150996086	151008375	q21.3	54	7%
DNALI1	1	37556919	37566857	p34.3	5	0.60%
AFF3	2	99545419	100142739	q11.2	2	0.20%
MLPH	2	237485428	237555318	q37.3	2	0.20%
RAB17	2	237574322	237601614	q37.3	2	0.20%
GPR160	3	170037929	170085403	q26.2	17	2.10%
SIDT1	3	113532296	113629578	q13.2	8	1%
APBB2	4	40810027	41216714	p13	5	0.60%
TBC1D9	4	140620765	140756120	q31.21	4	0.50%
CMBL	5	10275875	10308026	p15.2	9	1.10%
ARMT1	6	151452258	151470101	q25.1	13	1.60%
RMND1	6	151404763	151452181	q25.1	14	1.70%
SLC44A4	6	31863192	31879046	p21.33	2	0.20%
CCDC170	6	151494030	151621193	q25.1	16	1.90%
ESR1	6	151656691	152129619	q25.1	13	1.60%
SPDEF	6	34537802	34556333	p21.31	5	0.60%
AGR2	7	16791811	16833433	p21.1	8	1%
PRR15	7	29563811	29567295	p14.3	6	0.70%
AGR3	7	16859405	16881987	p21.1	8	1%
FBP1	9	94603133	94640249	q22.32	3	0.40%
C9orf152	9	110190048	110208189	q31.3	2	0.20%
GATA3-AS1	10	8050450	8053484	p14	15	1.80%
GATA3	10	8053604	8075198	p14	16	1.90%
ABCC8	11	17392885	17476845	p15.1	7	0.80%
KCNJ11	11	17385859	17389331	p15.1	7	0.80%
DACH1	13	71437966	71867192	q21.33	5	0.60%
DEGS2	14	100143957	100160163	q32.2	5	0.60%
SLC7A8	14	23125295	23183674	q11.2	3	0.40%
FOXA1	14	37589984	37596059	q21.1	10	1.20%
DRAIC	15	69463026	69571440	q23	5	0.60%
FSIP1	15	39600031	39782830	q14	4	0.50%
CT62	15	71110244	71115494	q23	5	0.60%
CA12	15	63321378	63382161	q22.2	3	0.40%
THSD4	15	71096952	71783383	q23	5	0.60%
ARSG	17	68259182	68422731	q24.2	56	7%
HID1	17	74950743	74973166	q25.1	32	4%
RABEP1	17	5282265	5385812	p13.2	4	0.50%
SLC16A6	17	68267026	68291267	q24.2	57	7%
C5AR2	19	47332147	47347327	q13.32	1	0.10%
TFF1	21	42362282	42366594	q22.3	4	0.50%
TFF3	21	42311667	42315651	q22.3	4	0.50%
XBP1	22	28794555	28800597	q12.1	4	0.50%
AR	X	67544032	67730619	q12	0	0%

Chapter Three.
Histone Deacetylase Inhibitors as Regulators of Luminal and Lactogenic Transcriptional Programs in Luminal Breast Cancer Cells

RATIONALE

Antiestrogen therapy has been effective in blocking the proliferative effects of estrogen receptor alpha (ERα), but unfortunately about 50% of patients will relapse due to resistance. Histone deacetylase inhibitors (HDACis) are a class of anticancer agents with anti-proliferative activity in both solid and haematological tumours (283). Inhibition of HDAC activity leads to histone hyperacetylation, causing chromatin opening, enhancing transcription factor accessibility and promoting gene transcription. However, HDACis can also repress a significant percentage of genes, due to changes in acetylation of other substrate proteins including transcription factors. Importantly, HDACis have been shown to suppress expression of ERα, abrogate 17β-estradiol (E2) signalling and induce differentiation and apoptosis in several breast cancer cell lines (413). These properties have led to the use of these inhibitors in clinical trials either as monotherapies or in combination with antiestrogens for treatment of ERα-positive breast tumours (364). However, the mechanisms of antiproliferative actions of HDACis in breast cancer still require investigation, especially as it remains unclear how these drugs impact transcriptional reprogramming and differentiation. Our hypothesis in this study was that the shutdown of E2 signalling and antiproliferative/proapoptotic properties of HDACis in breast cancer cells result in part from the induction of tumour cell differentiation by modulated acetylation of one or several transcriptional regulators in ERα-positive breast cancer cell lines. Our main goal has been to characterize the mechanisms of the ERα-suppressive and antiproliferative effects of HDACis using genome-wide gene expression profiling in breast cancer cells.

HISTONE DEACETYLASE INHIBITORS AS REGULATORS OF LUMINAL AND LACTOGENIC TRANSCRIPTIONAL PROGRAMS IN LUMINAL BREAST CANCER CELLS

Houssam Ismail[1], Martine Bail[2], Justyna Kulpa[1], Celia Li[1], David Laperrière[1] and Sylvie Mader[1,3,*]

[1] Institute for Research in Immunology and Cancer (IRIC), Université de Montréal, Montréal, QC H3C 3J7, Canada.
[2] Chemical Engineering department, Ecole Polytechnique Montréal, Montréal, QC H3C 3A7, Canada.
[3] Department of Biochemistry, Université de Montréal, Montréal, QC H3C 3J7, Canada, and Centre de Recherche du Centre Hospitalier Universitaire de l'Université de Montréal, Montréal, QC H2X 0A9, Canada.

* To whom correspondence should be addressed.
Tel: +1-514-343-7611; Fax: +1-514-343-6843; Email: Sylvie.Mader@umontreal.ca

Author contributions

H.I. and S.M. designed the study and wrote the manuscript. H.I. designed experiments, performed experiments and analysed all data. M.B. designed and performed gene expression microarrays. J.K. designed and performed HDACi proliferation assays. C.L. assisted with sample preparation and validation experiments for HDACi effects in T-47D and ZR-75-1 ERα-positive breast cancer cell lines as well as C646 experiments and acetylation assays in MCF-7 cells. D.L. generated the TFBS pipeline and processed raw transcriptome data.

ABSTRACT

Histone deacetylase inhibitors (HDACis) are a class of drugs with anti-tumourigenic activity in several types of cancers. HDACis have been shown to suppress expression of the estrogen receptor (ERα), which is overexpressed in two thirds of breast tumours. ERα, FOXA1 and GATA3, all luminal-lineage transcription factors (TFs), form a cross-regulatory network that controls estrogen signalling and mammary gland development. Here, we report that HDACis Trichostatin A (TSA) and suberoylanilide hydroxamic acid (SAHA) induce extensive transcriptional reprogramming and abrogate expression of ERα, FOXA1 and GATA3 in several ERα-positive luminal breast cancer cell lines. Moreover, TF protein expression is also suppressed by increased protein turnover, as revealed by cycloheximide and MG132 experiments. On the other hand, HDACi treatment induces expression of several markers of lactogenic differentiation, including the SREBP-1-driven lipid/cholesterol biosynthesis pathway, and reduces cell viability. Further, ERα, FOXA1 and GATA3 protein levels are reduced during lactation in the mouse mammary gland. Finally, we find that HDACis increase SREBP-1 acetylation, which stabilizes the protein by preventing its ubiquitination. Chemical inhibition of the histone acetyltransferase p300 prevents induction of SREBP-1 as well as its acetylation and leads to increased expression of cholesterol/lipid biosynthesis genes. It also reverses the suppressive effects of HDACis on luminal TFs, suggesting that acetylation also plays a role in their regulation. Collectively, our study suggests that treatment of ERα-positive breast tumours by clinically relevant HDACis fully suppresses estrogen-mediated proliferation and induces pathways characteristic of lactogenic differentiation via increased acetylation of multiple protein substrates.

INTRODUCTION

Breast cancer is a heterogeneous disease comprising several molecular subtypes as revealed by transcriptome profiling studies. These are the 'luminal', 'Her2+', 'basal-like' and 'claudin-low'. Luminal tumours express or overexpress the estrogen receptor (ERα) and are targeted with antiestrogens. HER2-positive tumours harbour an amplification in the HER2-encoding gene, *ERBB2*, and are targeted with Herceptin, a monoclonal antibody. The basal-like and claudin-low tumours are ERα-negative and are treated with chemotherapy. In luminal breast cancer, two luminal-specific TFs, GATA3 and FOXA1, play important roles in controlling ERα expression and signalling and mammary gland development. Expression of these TFs positively correlates with the luminal subtype of breast cancer and favourable patient prognosis (140, 206, 414). GATA3, a lineage-restricted member of the GATA family of TFs, is an essential factor for mammary gland ductal branching and alveologenesis (108, 109). It functions concomitantly with ERα where it positively controls its expression and modulates transcription of downstream E2 target genes (214). On the other hand, Forkhead box protein A1 (FOXA1), a member of the forkhead family of winged-helix TFs, has established roles in the development and differentiation of several organs including liver, kidney, and pancreas (91, 415, 416). Importantly, FOXA1 is required for proper ductal branching in the mammary gland, but unlike GATA3, is not needed for alveolar cell differentiation (95). FOXA1 also acts upstream of ERα both *in vivo* and *ex vivo*, regulating the transcription of the *ESR1* gene (95) **(see Chapter 2)**. FOXA1 can also regulate ERα function as it is required for optimal expression of around 95% of E2-regulated genes and E2-induced proliferation (167). However, while ERα, GATA3 and FOXA1 form a master co-regulatory network in breast cancer cell lines and normal mammary epithelial cells, the mechanisms underlying their tightly-controlled expression remain largely unexplored.

Targeting the epigenome for cancer therapy has gained considerable momentum over the past decade, especially as there is a growing body of evidence that chromatin-modifying enzymes are frequently mutated in cancer (417). These enzymes act in concert and could be grouped into three distinct categories that grant the epigenome its plasticity: writers, readers and erasers. Writers modify histones and DNA. Readers recognize and interpret these modifications; erasers remove them, ultimately resulting in transcriptional changes. Several drugs that target enzymes in these three categories are either currently being assessed in clinical trials or are being used in the clinic. For instance, the DNA methyltransferase inhibitor Decitabine has been approved by the European Medicines Agency (EMA) for treatment of

Acute Myeloid Leukemia. In addition, inhibitors of bromodomain-containing proteins that recognize acetylated residues are in Phase I and II clinical trials for treatment of haematological cancers (322). Histone deacetylase inhibitors (HDACis) were originally identified as compounds that induce differentiation in erythroleukemic cell lines (311, 324), and today, three of them are used in the clinic for treatment of peripheral/cutaneous T-cell lymphoma: suberoylanilide hydroxamic acid (SAHA; Vorinostat), Romidepsin, and Belinostat (322). Surprisingly, given that HDACis readily increase overall levels of histone acetylation, their effects on transcription appear to be gene-specific and lead to activation as well as repression of gene expression (333). The identity of genes regulated by HDACis suggests pleiotropic effects on pathways that control proliferation, apoptosis, tumour angiogenesis and immunology, as well as tumour cell differentiation (283, 418, 419). In addition, HDACs can modulate the acetylation of several non-histone substrates including proteins involved in chromatin remodelling, cell cycle regulation, splicing, nuclear transport, and actin nucleation (420). Recent efforts have aimed to further elucidate the mechanisms of action of these drugs, especially as their anticancer activities have been extended to solid cancers, including breast, endometrial and ovarian cancers, where they promote cell differentiation (328, 329, 331). Preclinical studies in breast cancer have found that HDACis have broad-spectrum growth-inhibitory effects where they induce a G1-phase cell cycle arrest via downregulation of cyclin D1 and upregulation of $p21^{WAF1/CIP1}$ (342, 421, 422). Additionally, treatment of ERα-negative breast cancer cells with HDACis was reported to relieve epigenetic silencing of the *ESR1* gene (gene encoding ERα) and restore sensitivity to antiestrogens (359, 360, 423, 424). Previous results from our lab and others have shown that HDACis block expression of ERα both transcriptionally and post-transcriptionally in ERα-positive breast cancer cells (349, 350, 353, 355, 356). These findings subsequently lead to a phase II clinical trial where clinical benefit was observed in 40% of ERα-positive metastatic breast cancer patients following combined treatment with SAHA and tamoxifen (364).

The molecular mechanisms by which HDACis reprogram breast cancer cells and regulate their differentiation remain unclear. Here, we report that HDACis cause extensive transcriptional reprogramming of MCF-7 breast cancer cells, suppressing expression of luminal TFs ERα, FOXA1 and GATA3 through both transcriptional and post-transcriptional mechanisms. In addition, we show that HDACis induce expression of genes that are upregulated during lactogenic differentiation in the human breast, including *SREBF1*, a master TF that controls expression of cholesterol and lipid biosynthesis genes. Consistent

with our observations in HDACi-treated breast cancer cells, we find that expression of ERα, FOXA1 and GATA3 is reduced during lactation in the normal mouse mammary gland. Finally, we identify a role of acetyltransferases p300/CBP in HDACi-mediated suppression of luminal TFs and in stabilizing SREBP-1 protein levels via increased acetylation. Together, these results demonstrate mechanistically how HDACis can transcriptionally reprogram luminal breast cancer cells towards a lactogenic differentiation state that is marked by loss of expression of luminal TFs.

MATERIAL AND METHODS

Cell culture and reagents

All cell lines used in this study were purchased from ATCC. MCF-7 breast cancer cells were maintained in α-Minimal Eagle's Medium (α-MEM; Wisent) supplemented with 10% fetal bovine serum (FBS; Sigma), 1% penicillin/streptomycin (Wisent) and 1% L-glutamine (Wisent). T-47D and ZR-75-1 breast cancer cell lines were maintained in RPMI 1640 (Wisent) supplemented with 10% FBS, 1% penicillin/streptomycin, 1% L-glutamine, 10mM HEPES (Sigma), and 1% sodium pyruvate (wisent). All cell lines were kept in a humidified 37°C incubator with 5% CO_2. Prior to HDACi treatment, all cells were cultured in DMEM without phenol red supplemented with 2% L-glutamine, 1% penicillin/streptomycin, and 10% charcoal-dextran treated FBS (FBST) for 72 h. Cells were then starved in DMEM without phenol red supplemented with 2% L-glutamine, 1% penicillin/streptomycin, and 0.5% FBST for 16 h before being treated with the desired compounds. TSA, SAHA, MS-275 and C646 were all purchased from Sigma and dissolved in DMSO at 10mM each and stored in dark at -20˚C. The mother stock solution was then diluted to the desired working stock concentration in DMSO. Following treatment, cells were washed twice with ice-cold phosphate buffered saline (PBS) and harvested for subsequent protein and RNA extractions.

RNA Extractions, Reverse Transcription and Real-Time Quantitative PCR

Total RNA was extracted using the TRIzol reagent according to the manufacturer's protocol (Invitrogen). 2 µg of RNA were reverse transcribed to cDNA following specifications of the RevertAid H first minus strand cDNA synthesis kit (Thermo Fisher Scientific) using oligo(dT)18 primers. cDNA was subsequently diluted one in ten and expression levels of target genes were assessed by real-time quantitative PCR (RT-qPCR) using the Universal ProbeLibrary system (Roche). *YWHAZ* and *RPLP0* were used as reference genes. mRNA levels were determined using the $2^{-\Delta\Delta C}{}_T$ (RQ) method (425). First, ΔC_T is calculated by determining the difference between the C_T values of the target gene and the reference genes. $\Delta\Delta C_T$ is then calculated as the difference between ΔC_T of each condition and that of the calibrator control. RQ<1 indicates repression and RQ>1 indicates induction. RT-qPCR primer sequences of genes tested in this study are provided in **Supplementary Table 1**.

Protein Extraction and Western Blot Analysis

Whole cell extracts were prepared using a total lysis buffer (50 mM Tris-HCl pH 7.4, 5.0 mM EDTA, 150 mM NaCl, 0.5% Triton, 1.0% NP40. 2% SDS and freshly-added protease inhibitors (PMSF at 10mM; leupeptin, pepstatin, and aprotinin at 1 µg/ml). Extracts were homogenized by sonication and quantified using Lowry assay (BioRad). Equal amounts of proteins (20-60µg) were electrophoresed on an 8-14% SDS-polyacrylamide gel. Proteins were transferred onto polyvinylidene difluoride membranes (Millipore). Membranes were blotted with a rabbit anti-ERα (1 in 1000; Clone 60C from Millipore), rabbit anti-FOXA1 (1:3000; ab23738 from Abcam), anti-GATA3 (1 in 3000; Clone 16E10A23 from BioLegend), anti-SREBP-1 (1 in 1000; SC-13551), mouse monoclonal anti-β-actin (1 in 10000; Sigma). HRP-conjugated anti-mouse and rabbit IgG were used as secondary antibodies (Cedarlane). Immunodetection was performed using enhanced chemiluminescence (PerkinElmer Life and Analytical Sciences) as per manufacturer's instructions.

Microarray gene expression profiling

MCF-7 cells cultured in hormone-depleted medium were treated with E2 (25nM), TSA (300nM) or E2+TSA for 24 h. RNA was extracted as mentioned earlier and purified using the RNeasy MinElute Cleanup Kit (QIAgen). cRNA synthesis from total RNA, labelling and hybridization to Illumina WG-6 expression microarrays were performed at the Genome Quebec and McGill University Innovation Center using standard protocols (http://www.genomequebec.mcgill.ca/). Three biological replicates per condition were hybridized and analysed on 12 chips. Gene expression analysis was performed using the lumi and limma Bioconductor packages (426, 427). Genes deemed significantly regulated were those with 1.4-fold change between vehicle and treatments, average log2-expression levels greater than 5 across all samples (A-value) and a p-value for moderated statistics (428) smaller than 0.01. P-values were adjusted to control the false discovery rate with the Benjamini and Hochberg method (429).

RNA sequencing and analysis

MCF-7 cells cultured in hormone-depleted medium were treated with TSA (300nM), SAHA (5µM), or MS-275 (5µM) for 8 h. RNA was extracted as previously mentioned. RNA purity was assessed by nanodrop using 260/280 and 260/230 ratios. Quality of total RNA was assessed with the BioAnalyzer Nano (Agilent) and all samples had a RIN above 9. Library

preparation was done with the Truseq RNA stranded kit (Illumina). Libraries were quantified by QuBit and BioAnalyzer. All libraries were diluted to 10 nM and normalized by qPCR using the KAPA library quantification kit (KAPA; Cat no. KK4973). Libraries were pooled to equimolar concentration. Sequencing was performed with the Illumina Hiseq2000 using the SBS v3 kit (200 cycles, paired-end). Around 60M paired-end PF reads was generated per sample. Library preparation and sequencing was performed at the Institute for Research in Immunology and Cancer's Genomics Platform (IRIC). The standard MUGQIC RNA-Seq pipeline (version 2.1.1) was used. Briefly, reads were trimmed from the 3' end to have a phred score of at least 30. Illumina sequencing adapters were removed from the reads, and all reads were required to have a length of at least 32 bp. Trimming and clipping were performed using Trimmomatic (430). The filtered reads were aligned to a reference genome. The genome used in this analysis is Homo_sapiens assembly GRCh37. Each readset is aligned using STAR (431). Then, all readset BAM files from the same sample are merged into a single global BAM file using Picard. The differential gene expression analysis is done using DESeq (432) and edgeR (433) Bioconductor packages and a nominal DESeq p-value is generated using the Wald statistical test.

GO term enrichment analysis

GO term enrichment analysis in this chapter was performed using the PANTHER classification system available at http://pantherdb.org/. The statistical test used to calculate p-values is the binomial test (434).

Transcription factor binding site enrichment analysis (TFBS)

TFBS analysis was performed as described previously (435).

Immunohistochemistry

Mammary glands from C57BL/6 mice were formalin fixed and paraffin embedded at the IRIC histology core facility. For staining, deparaffinization and antigen retrieval were performed using a Discovery XT automatic stainer (Ventana Medical Systems) using protocols recommended by the manufacturer. Primary antibodies used were directed against ERα (C-60, Millipore Technologies, 1:50 dilution for 180 min at room temperature), FOXA1 (ab23738, 1:200 dilution for 60 min at room temperature) and GATA3 (Clone 16E10A23 from BioLegend, 1:20 for 180 min at room temperature). Primary antibody incubation was followed

by incubation with appropriate biotin conjugated secondary antibodies (1:100, Jackson ImmunoResearch Laboratories). For staining, streptavidin-horseradish peroxidase, and 3,3 diaminobenzidine were used according to the manufacturer's instructions (DABmap detection Kit, Ventana Medical Systems). The sections were then counterstained with Gill hematoxylin and a bluing reagent was applied for postcounterstaining. Sections were scanned at X40 magnification using the C9600 NanoZoomer System (Hamamatsu Corporation, Bridgewater, NJ, USA), which can adjust focus on any part of the slide. NDP Scan software (version 2.2.17) was used to visualize virtual slides and extract images.

Immunoprecipitation

MCF-7 cells were treated with TSA (300nM) or SAHA (5μM) or MS-275 (5μM) for 3 h and collected on ice. For C646 acetylation experiments, MCF-7 cells were pre-treated with C646 (20μM), MG132 (10μM) or both for 3 h and then treated with vehicle or TSA (300nM) for 3 h. DMSO served as the vehicle control for all treatments. Harvested cells were lysed in IP buffer (50mM tris-HCl pH 7.5, 150mM NaCl, 5mM EDTA, and 0.1% NP40) for 10 minutes on ice. The lysis buffer was supplemented with TSA (300nM) or SAHA (5μM) or MS-275 (5μM) to prevent loss of acetylation during extraction steps. Protein lysates were subsequently sonicated for 1 minute and quantified using the lowry method. 1 mg of protein lysates were incubated with an anti-acetyl-lysine antibody (sc-32268, Santa Cruz) or mouse IgG (ChromPure Mouse IgG, Jackson laboratories) (9μg of antibody per 1mg of protein lysate) overnight on a rotor at 4°C. The following day, beads were washed twice with PBS and blocked with PBS containing 0.5% BSA for 10 minutes on a rotor at room temperature. Blocked beads were then washed once with PBS and once with IP buffer and added to immunoprecipitated protein lysates for 4 h on a rotor at 4°C. Beads were subsequently centrifuged and washed six times with ice-cold IP buffer and proteins were eluted in 60μl of 2X Laemmli sample buffer. 60μg of input sample and 30μl of immunoprecipitated samples were resolved by SDS-PAGE and immunoblotted for SREBP-1 (1 in 1000, sc-8984, Santa Cruz) and histone H4 (ab10158, abcam).

DNA transfection

An electroporation strategy was used for DNA transfections in MCF-7 cells. Five million cells were counted and resuspended in 200μl of culture media. A 40μg DNA mix was prepared containing 6μg of plasmid DNA, 34 μg of salmon sperm DNA (Invitrogen) and 0.2M NaCl.

PSG5-*ESR1*, pCMV6-XL5-*FOXA1*, pCMV6-XL5-*GATA3*, and pCMV6-XL5-EGFP plasmids were used. pCMV6-XL5 empty vector was utilized as a negative control. The cell and DNA mixture was then placed in a cuvette and pulsed using a Gene Pulser (Bio-Rad) at 250V and a capacitance of 975μF. Electroporated cells were then washed with fresh medium and allowed to adhere overnight. The following day, cells were treated with HDACis and RNA and proteins were extracted for subsequent assays.

Proliferation assays

MCF-7 cells were seeded at a density of 2000 cells per well in 96-well plates. The next day, cells were incubated with media containing the alamarBlue® reagent (1:50 dilution, Invitrogen) for 3 h at 37°C and absorbance was read at 590nm. Afterwards, cells were treated for 3 consecutive 72 h with 1:2 serial dilutions of a starting concentration of 20μM of TSA, SAHA or MS-275. AlamarBlue® absorption signal was measured at the end of the assay.

RESULTS

HDACis suppress estrogenic signalling including activation of proliferative target genes in MCF-7 breast cancer cells

Previous results from our lab have shown that HDACis Trichostatin A (TSA), a pan-specific hydroxamic acid, and sodium butyrate inhibit E2-mediated transcriptional activation of an ERE-driven luciferase reporter in Ishikawa endometrial cancer cells (349). Suppression of E2-mediated proliferation and signalling by HDACis has also been reported in MCF-7 cells (342, 350). These findings prompted us to examine the impact of HDACis on E2 signalling in a genome-wide manner. Hormone-depleted MCF-7 cells were treated with E2 (25nM), TSA (300nM), or E2+TSA for 24 h and gene expression profiles were assessed by Illumina WG-6 expression microarrays **(Figure 1).** A fold change (FC) of 1.4 and $p<0.01$ were chosen to select significantly regulated genes. Regulation by E2 treatment was biased towards activation (369 genes UP and 179 DOWN), whereas that of TSA (2478 UP, 2359 DOWN) and E2+TSA (2483 UP, 2387 DOWN) was equally distributed between up- and downregulation **(Figure 1A)**. The greater number of regulated genes by TSA can be explained by the pleiotropic effects HDACis have on global histone acetylation. Strikingly, the majority of E2-induced genes were either not significantly regulated (231/369, grey dots) or repressed (59/369, green dots) following co-treatment with TSA (**Figure 1B**), suggesting shutdown of E2 signalling. The remaining 79 genes that were induced by both E2 and E2+TSA (orange dots in **Figure 1B)** were also induced by TSA treatment alone **(Figure 1A)**. Conversely, the majority of E2-repressed genes were also repressed by TSA or E2+TSA (**Figure 1A and 1B**). Plotting the folds of gene regulation in the presence of TSA vs E2+TSA revealed a near-perfect correlation (**Figure 1C**), indeed corresponding to an almost complete suppression of the impact of E2 on gene transcription. This result was not surprising given previous findings of suppression of ERα expression by HDACis (349, 350, 353, 355, 356), which we confirmed as shown in **Figure 1D.** TSA-mediated suppression of E2 regulation of two E2-target genes, *GREB1* and *XBP1,* was also confirmed by RT-qPCR (**Figure 1E)**. Lastly, given the antiproliferative nature of HDACis, we examined the impact of TSA on E2-mediated induction of a cluster of 99 cell cycle genes (366). As expected, the vast majority of these genes (81/99), induced by E2 alone, were not significantly regulated following E2+TSA treatment **(Figure 1F)**. Collectively, these results indicate that HDACis can efficiently suppress the E2-dependent response in MCF-7 cells including proliferation genes.

HDACis induce transcriptional reprogramming of MCF-7 breast cancer cells

To explore the molecular basis underlying the gene regulation properties of HDACis in breast cancer cells, we sequenced the transcriptomes of luminal ERα-positive MCF-7 cells following treatment with one of three different HDACis: TSA (300 nM), SAHA (5 μM), or MS-275 (5 μM) for 8 h in hormone-depleted media. Like TSA, SAHA also belongs to the hydroxamic acid class of HDACis and is pan-specific targeting class I (HDAC1, 2, 3 and 8), II (HDAC4, 5, 6, 7, 9 and 10) and IV (HDAC11) HDACs. MS-275, on the other hand, is a benzamide and preferentially inhibits HDAC1 and 3 (436). About the same numbers of genes were up- or downregulated (compared to vehicle control; FC>1.4 or <-1.4 and p<0.01) in all three datasets (TSA: 3873 UP and 3924 DOWN genes; SAHA: 3514 UP and 3592 DOWN; MS-275: 3926 UP and 3878 DOWN) **(Figure 2A),** consistent with previous HDACi gene expression profiling studies (333, 437, 438). Notably, up to 18% of genes regulated by HDACis were TFs with experimentally validated DNA-binding capacities (439) (TSA: 465 UP, 930 DOWN; SAHA: 422 UP, 844 DOWN; MS-275: 442 UP, 880 DOWN) (**Figure 2A)**, the majority of these TFs being repressed by HDACi treatment.

Genes significantly regulated by TSA were selected and their folds of regulation were plotted against those in the SAHA condition (Figure 2B). This revealed a high correlation between regulation by both drugs (r=0.96 for all genes; r=0.93 and r=0.79 for upregulated and downregulated genes, respectively) (Figure 2B), as expected given the congruence of their HDAC targets. Of note, a large proportion of upregulated genes were highly induced by both drugs (10-1000 fold), suggestive of derepression of silenced genes. Indeed, examination of the top 100 induced genes by TSA/SAHA reveals their lack of expression in the vehicle condition (**Supplementary Figure 1)**. However, examination of regulation by MS-275 treatment of TSA/SAHA commonly regulated genes (3082 UP and 3268 DOWN), or conversely, by TSA/SAHA of MS-275-regulated genes, revealed a lower degree of similarity between these two classes of HDACis, and identified four categories of genes significantly regulated by all three drugs **(Figure 2C):** (i) genes regulated in the same direction by TSA/SAHA and MS-275 (1024 UP and 1147 DOWN, representing about 33% and 35% of TSA/SAHA regulated genes, respectively; red and light blue dots in **Figure 2C**); (ii) genes regulated in opposite directions by TSA/SAHA and MS-275 (544 genes UP with TSA/SAHA and DOWN with MS-275, representing 31.5% of the TSA/SAHA UP genes; 364 genes DOWN by TSA/SAHA and UP by MS-275, representing 42.5% of TSA/SAHA DOWN genes; purple and yellow dots in **Figure 2C**) (iii) genes uniquely regulated by TSA/SAHA (972 UP

and 1394 DOWN; dark blue dots in **Figure 2C**); (iv) genes uniquely regulated by MS-275 (638 UP and 1619 DOWN, representing almost 16% and 42% of UP and DOWN genes regulated by this drug, respectively; green dots in **Figure 2C**). Together, these results indicate that while the two classes of drugs share common target enzymes and genes, they also exert unique and divergent regulatory effects on gene expression.

Next, we performed GO term enrichment analysis to identify biological processes over-represented in our datasets (440, 441). The top 20 GO terms are presented in **Figure 2D** and **2E** for TSA/SAHA and MS-275-responsive genes, respectively. Processes overrepresented in the TSA/SAHA upregulated genes included cilium organization (fold enrichment=2.26, *p*=6.73E-10) and assembly (fold enrichment=2.3, *p*=5.09E-10), cytoskeleton-related processes such as membrane docking (fold enrichment=2.58, *p*=1.37E-06) and microtubule-based protein transport (fold enrichment=3.41, *p*=5.94E-03) (**Figure 2D)**. This result is in line with previous reports highlighting the role of TSA and HDAC6-specific inhibitors, tubacin and tubastatin A, in preventing ciliary disassembly by promoting tubulin acetylation (442, 443). In addition, we found that multiple downstream effector proteins of smoothened signalling were upregulated by TSA and SAHA (*PTCH1, SUFU, GLI1;* smoothened signalling: fold enrichment=2.79, *p*=1.97E-02) **(Supplementary Figure 2A)**. Notably, these proteins have been shown to concentrate in primary cilia where hedgehog signalling takes place (444–447).

Although a role of HDACis in inducing lipid droplet formation has been reported in breast cancer cells (331, 344, 448), the mechanism by which this happens is not known. Interestingly, our results reveal that several metabolic processes including lipid (fold enrichment=2.42, *p*=4.94E-05) and sterol biosynthesis (fold enrichment=3.16, *p*=9.16E-02) were enriched in genes upregulated by TSA and SAHA but not MS-275. Upregulated genes included ones involved in cholesterol biosynthesis such as mevalonate kinase (*MVK)*, lanosterol synthase (*LSS),* 3-hydroxy-3-methylglutaryl-Coenzyme A reductase *(HMGCR),* farnesyl diphosphate farnesyl transferase 1 *(FDFT1),* as well as the master TF sterol regulatory element-binding factor 1, *SREBF1* **(Supplementary Figure 2A)**. When cholesterol/lipid levels are low, SREBP-1, which resides in the endoplasmic reticulum, is cleaved and its active TF domain is translocated to the nucleus to activate transcription of fatty acid and cholesterol biosynthesis enzymes (449).

Conversely, GO terms enriched in genes downregulated by TSA/SAHA (**Figure 2D)** included genes involved in processes of transcription termination by RNA polymerase II (fold

enrichment=3.02, *p*=2.75E-03), mRNA 3'end processing (fold enrichment=2.8, *p*=8.07E-03), RNA splicing (fold enrichment=2.48, *p*=4.28E-03) as well as translation (fold enrichment=2.81, *p*=5.84E-08). Furthermore, we found enrichment of processes involving chromatin remodelling as well as histone acetylation and methylation. Closer inspection of genes involved with these processes uncovered repression of several chromatin remodellers (*CHRAC1, PBRM1, SMARCB1, SMARCD1, SMARCC1, SMARCD2, SMARCAD1, SMARCC2, SMARCA4, RSF1, BPTF, SCMH1*) and histone acetyltransferases (*EP300, CREBBP*, *JADE1-3, KAT2A, KAT5, KAT6B, KAT7, KAT8*), and methyltransferases (*SETD7, SETDB1, SETD2, SETD1B, SETD1A, PRDM6, MECP2, ASH1L, KMT2D, NSD1, EZH1, WHSC1 WHSC1L1, SUV420H1, EHMT1, EHMT2, PRMT2, PRMT6*) **(Supplementary Figure 2A)**, suggesting indirect mechanisms by which HDACis can activate or repress gene transcription.

GO terms enriched in MS-275-upregulated genes included ones related to epithelial cell differentiation (fold enrichment=2.65, *p*=4.37E-05) and development (fold enrichment=2.16, *p*=1.56E-03), mesenchymal cell differentiation (fold enrichment=2.19, *p*=4.52E-02), and mammary gland development (fold enrichment=2.52, *p*=3.92E-04) (**Figure 2E).** Genes related to epithelial cell differentiation included *GRHL1* and *2, OVOL2, TP63, PRLR,* and *LIF* **(see heatmap in Supplementary Figure 2B)**. Intriguingly, the angiogenesis GO term was also enriched in MS-275 upregulated genes (fold enrichment=2.21, *p*=6.94E-08) after 8 h of treatment. It will be essential to examine the sustainability of induction of these genes following prolonged MS-275 treatment, especially as this inhibitor has been shown to inhibit angiogenesis and metastasis in a breast cancer xenograft mouse model (450). MS-275-downregulated genes were enriched in GO terms associated with nuclear pore organization (fold enrichment= 4.92, *p*=3.66E-02), tRNA (fold enrichment=4.3, *p*=2.65E-05) and mRNA export from the nucleus (fold enrichment=3.85, *p*=3.51E-17) as well as regulation of transcription elongation by RNA polymerase II (fold enrichment= 3.86, *p*=3.39E-02) (**Figure 2E and Supplementary Figure 2B)**. These data suggested that HDACs may regulate the processes of transcription and translation by affecting mRNA and tRNA export to the cytoplasm further emphasizing the extensive degree of transcriptional reprogramming as a result of their inhibition.

Consistent with the antiproliferative properties of these inhibitors in MCF-7 cells **(Supplementary Figure 3A)**, Ingenuity Pathway Analysis (IPA) (451, 452) predicted an increase in apoptosis (z-score=2.897, *p*=5.86E-12) and decrease in cell viability of tumour

cells (z-score=-2.547, *p*=3.5E-12) by TSA and SAHA. Indeed, several anti-apoptotic (*CFLAR, BCL2, BCL2L1, XIAP*) and pro-apoptotic (*APAF1, CYCS AIFM1, BCL6 CASP9* and *10*) genes, were downregulated and upregulated, respectively. Furthermore, IPA predicted activation of apoptosis by TSA/SAHA via inhibition of Remodeling and Spacing Factor 1 (RSF1) function (activation z-score= -2, *p*=1.55E-02) **(Supplementary Table 2)**. *RSF1* expression is inhibited by TSA/SAHA and this in turn is predicted to prevent expression of anti-apoptotic RSF1 targets *CFLAR, BCL2, BCL2L1,* and *XIAP* **(Supplementary Figure 3B)** (453)**.** While RSF1 expression was also repressed by MS-275, its downstream anti-apoptotic targets were not significantly regulated under these conditions, with the exception of *BCL2,* which was increased. However, it is possible that prolonged treatment with MS-275 is required to observe changes in expression of genes related to apoptosis. Transcriptional regulators predicted to be commonly activated by TSA, SAHA and MS-275 included *HIF1A* (activation z-score= 2.425, *p*=1.88E-05 for TSA/SAHA; activation z-score= 4.21, *p*=3.17E-07 for MS-275) and *EGR1* (activation z-score= 2.019, *p*=5.57E-05 for TSA/SAHA; activation z-score= 2.437, *p*=2.95E-03 for MS-275) **(Supplementary Table 2 and 3)**. This is compatible with previous reports that acetylation of both factors by p300 increases their protein stability and transcription of their downstream target genes (454, 455).

Lastly, because HDACis resulted in loss of expression of many TFs and chromatin-modifying enzymes, we reasoned that screening regulatory sequences of HDACi-target genes for enrichment in TF binding may provide insight into their mechanism of action. We performed TF binding site (TFBS) enrichment analysis surveying the regulatory sequences within 2.5 kb, 5 kb or 10 kb upstream or downstream of the transcription start site (TSS) of HDACi-responsive genes (**Supplementary Figure 4A and 4B)**. We found enrichment of MAF1 (V_MAF_Q6, fold enrichment=1.556 at 2.5 kb), EGR3 (V_EGR3_01, fold enrichment=1.553 at 2.5 kb), and ATF (V_ATF_01, fold enrichment=1.542 at 2.5 kb) binding motifs in genes upregulated by TSA/SAHA whereas HOXA3 binding motifs (V_HOXA3_01, fold enrichment=2.05 at 2.5 kb) were enriched in MS-275-upregulated genes. On the other hand, YY1 (V_YY1_Q6_02, fold enrichment=2.78 at 10 kb) and SMAD binding motifs (V_SMAD_Q6, fold enrichment=1.523 at 5 kb) were enriched in genes downregulated by TSA/SAHA. Binding motif enrichment of YY1 (fold enrichment=2.064 at 2.5 kb) as well as members of the ETS TF family (C-ets-1 (V_CETS1P54_01, fold enrichment=2.011 at 10 kb) and ELK1 (V_ELK1_04, fold enrichment=1.817 at 10 kb)) was observed in genes downregulated by MS-275**.** Interestingly, HDACi treatment modulated expression of several of

these enriched TFs (**Supplementary Figure 4C**), which may contribute to transcriptional reprogramming by these inhibitors. In addition, acetylation of members of the ATF, EGR, SMAD and ETS protein families has been reported. Acetylation of ATF2, ATF4, ATF5 and EGR1 can increase their transcriptional activity (455–458), consistent with enrichment of their binding sites in genes upregulated by TSA/SAHA. Acetylation of SMAD2 and SMAD3 as well as ETS1 and ETV1 has been reported to increase their transcriptional activity (459–462). This is inconsistent with enrichment of their binding in downregulated genes, although it remains unclear whether HDACis indeed modulate their acetylation in breast cancer cells. YY1, whose binding sites were most enriched in HDACi-downregulated genes and is repressed by HDACi treatment, can activate or repress transcription by recruiting HAT or HDAC activities (463, 464). Furthermore, p300-mediated acetylation of YY1 at its central glycine-lysine rich domain is necessary for its transcriptional repressor function (465). PCAF can also acetylate YY1 at its C-terminal zinc finger domain and this results in reduced binding to DNA (465). It remains necessary to assess whether YY1 acetylation is modulated following HDACi treatment at early time points when its expression is not yet altered and determine if expression of an acetylation-defective mutant may impede HDACi-mediated repression of transcription.

HDACis abolish expression of luminal TFs in breast cancer cells

Our transcriptome data revealed repression of several luminal-lineage TFs including *ESR1, FOXA1,* and *GATA3* irrespective of hormonal supplementation (RNA-Seq in hormone-depleted media and microarray analyses with or without E2 supplementation) **(Supplementary Figure 5A and 5B)**. Monitoring mRNA expression of these TFs at 1, 2, 3, 6, 8, 12, 16 and 24 h after treatment with HDACis in MCF-7 cells maintained in hormone-depleted media (**Figure 3A)** revealed that *GATA3* expression was repressed 2-fold or more as early as 1 h of treatment with TSA or SAHA. *FOXA1* relative mRNA levels were downregulated at 2-3 h and this was followed by a similar decline in *ESR1* mRNA levels at 3 h of treatment. Transcriptional repression by TSA and SAHA was maximal at 12 h of treatment for *FOXA1* and *GATA3,* and 24 h for *ESR1* **(Figure 3A)**. MS-275, on the other hand, exhibited slower kinetics for downregulation of *ESR1,* progressing from 6 to 24 h, whereas GATA3 expression was transiently induced at 6 h and repressed at later time points, and *FOXA1* expression remained unaffected at all time points **(Figure 3A)**. Lastly, rapid transcriptional repression of *SPDEF* and *XBP1*, two TFs whose expression is strongly associated with that of *FOXA1*, *ESR1* and *GATA3* (see Chapter 2), was also observed

following treatment with TSA or SAHA with decreased expression as early as 1 h **(Supplementary Figure 6)**. Of note, MS275 displayed a different pattern of effects, with transient induction of expression of *XBP1* and *SPDEF* at 2-6 h, and weak repressive effects at longer time points.

ERα, FOXA1 and GATA3 protein levels were also repressed by TSA and SAHA (**Figure 3B)**. GATA3 protein levels were markedly reduced between 6-24 h of treatment, and a similar impact on ERα protein levels was observed with a slight delay. FOXA1 protein levels were also repressed but less markedly and more transiently. Downregulation of ERα and GATA3 by MS-275 was only observed at 24 h of treatment, and with only minor impact on FOXA1 expression. We further extended our observations in MCF-7 cells to two other ERα-positive breast cancer cell lines: T-47D and ZR-75-1, where we similarly observed HDACi-mediated repression of these TFs at the RNA and protein levels at 16 h of treatment (**Figure 3C and 3D)**. However, unlike MCF-7 cells, MS-275 treatment resulted in downregulation of *FOXA1* mRNA levels, albeit with no significant change in its protein levels at this time point. It remains possible however that prolonged treatment is necessary to observe an impact of MS-275 on FOXA1 expression.

SiRNA-mediated downregulation of FOXA1 in MCF-7 cells has been previously reported to induce expression of basal markers and increase metastasis and invasion (212). We therefore wondered whether HDACi treatment may affect expression of basal markers given the downregulation of luminal TFs. Transcriptome profiling of breast cancer cell lines has led to identification of two basal subtypes: basal A and B. Basal A cell lines are more representative of basal-like tumours whereas basal B like the claudin-low subtype exhibit characteristics of epithelial to mesenchymal transition (466). We examined the impact of HDACi treatment on expression of the top 30 luminal, basal A and basal B markers generated by the Neve et al. classifier (466) **(Supplementary Figure 7)**. Genes downregulated by TSA or SAHA were significantly enriched in luminal markers (fold enrichment= 1.95 and 2.5, $p<0.05$, respectively) and basal A markers (fold enrichment= 2.32 and 1.95, $p<0.05$, respectively). Basal B markers were significantly enriched in genes both up- and downregulated by TSA and SAHA. On the other hand, MS-275 overwhelmingly resulted in significant enrichment of luminal, basal A and basal B markers in induced (fold enrichment= 3.03, 3.03 and 3.57, $p<0.05$, respectively), but not repressed genes. These results indicate that HDACis have a differential impact on expression of subtype markers, but do not favour reprogramming towards one specific subtype.

HDACis abolish expression of luminal TFs via both transcriptional and post-transcriptional mechanisms

To explore the mechanism(s) of repression of *ESR1*, *FOXA1* and *GATA3*, we addressed the possibility that their downregulation may be the result of loss of one or several upstream regulators. Although FOXA1 is a positive regulator of *ESR1* and *GATA3* transcription, as well as of *XBP1* and *SPDEF* **(see Chapter 2)**, the kinetics of its HDACi-mediated loss at the protein level **(Figure 3B)** do not support the hypothesis that its reduction triggers HDACi-mediated downregulation of ERα and GATA3. However, it remains possible that loss of FOXA1 expression may contribute to sustained repression of ERα and GATA3 expression at later time points. To examine this possibility, we ectopically expressed each of these TFs individually in MCF-7 cells in the presence or absence of TSA treatment **(Supplementary Figure 8)**. We found that overexpression of *FOXA1* did not alleviate transcriptional repression by TSA, nor did it restore protein levels of ERα or GATA3 **(Supplementary Figure 8A and 8B)**. Notably, overexpression of GATA3 or ERα did not prevent TSA-mediated repression of *FOXA1* mRNA or protein levels.

To further dissect the mechanisms underlying repression of luminal TFs, we asked whether *de novo* protein synthesis is necessary for their HDACi-mediated repression at the RNA and/or protein levels. MCF-7 cells were pre-treated with cycloheximide, a translation elongation inhibitor, for 1 h and then treated with TSA, SAHA or MS-275 for 16 h. Interestingly, cycloheximide alleviated repression of ERα protein expression by TSA, SAHA and MS-275 **(Figure 4A)**. While protein levels of FOXA1 and GATA3 were already reduced with cycloheximide treatment alone, we did not observe any further reduction in their protein levels following HDACi co-treatment **(Figure 4A)**. Intriguingly however, we found that mRNA levels of *ESR1, FOXA1* and *GATA3* were still reduced following co-treatment with cycloheximide **(Figure 4B)**. These results indicate that HDACi-mediated repression of protein expression of these TFs is cycloheximide-sensitive, whereas their transcriptional repression is not, suggesting different underlying mechanisms.

It was previously reported that SAHA enhances ERα ubiquitination and degradation through the proteasome via CHIP (encoded by *STUB1*), an E3 ubiquitin ligase (355). We therefore assessed whether HDACis regulate FOXA1 and GATA3 expression similarly through the proteasome pathway. Indeed, co-treatment with MG132, a proteasome inhibitor, prevented the repressive actions of TSA, SAHA and MS-275 and stabilized ERα, FOXA1 and

GATA3 protein levels **(Figure 4C)**. Together with the cycloheximide experiment, these data suggest that HDACis may target luminal TFs for degradation by the proteasome through inducing expression of one or several effector protein(s). While the identity of such proteins remains to be determined by mapping global changes in protein expression in the presence or absence of cycloheximide and HDACis, the possibility of ubiquitin ligases playing a role is a plausible one. Indeed, our transcriptome sequencing revealed upregulation of several genes encoding proteins involved in the ubiquitin-proteasome pathway by TSA and SAHA **(Figure 4D).** This included ubiquitin ligases (*STUB1*, *CBL*, *CCNB1IP1*, *MARCH2, HERC5, NEDD4L, RNF41, RNF123, TNFAIP3, TRIM2, UBE4A,* and *USP20*) as well as several members of the F-Box protein family which together with SKP1 and Cullin form the SCF ubiquitin ligase complex (*FBXO10, 36, 43,* and *FBXL18*) **(Figure 4D)**.

Collectively, these results indicated that HDACis can repress expression of luminal TFs simultaneously through both transcriptional and post-transcriptional mechanisms.

HDACis activate processes implicated in lactogenic differentiation in breast cancer cells

HDACis have been previously reported to induce specific lactogenic differentiation traits in breast cancer cells, such as lipid droplet formation or induction of milk proteins (331, 343, 344). Milk produced in the mouse mammary gland is primarily composed of lipids (30%; 29% of which are triglycerides), proteins (12%) and lactose (5%) (13, 467). The mammary gland has the capacity to synthesize fatty acids (FAs) *de novo*, in addition to obtaining them through maternal diet or mobilization of fat reserves. Consistent with these data, lactogenic differentiation markers were shown to encompass genes involved in FA synthesis, elongation, desaturation, uptake, and transport. Additionally, genes associated with triglyceride synthesis, cholesterol synthesis, as well as formation of lipid droplets are also regulated during lactation (468).

Our finding that lipid and sterol biosynthesis GO terms were enriched in genes upregulated by TSA and SAHA in MCF-7 cells **(Figure 2D and Supplementary Figure 2A)** prompted us to further examine the underlying transcriptional programs that govern this HDACi regulation. Reanalysis of a previously published gene expression microarray dataset by Mohammad et al. (468) identified genes that were differentially expressed at 72 h postpartum (compared to 6 h postpartum, FC>1.4 or <-1.4 and $p<0.05$; n=7). Expectedly, GO term analysis of upregulated genes identified enrichment of several metabolic processes that

were also regulated by HDACis **(Supplementary Figure 9A and 9B)**. Acetyl-CoA biosynthesis (fold enrichment=7.15, *p*=1.87E-03) was marked by induction of several enzymes including ATP-citrate synthase (*ACLY*) and Acetyl-coenzyme A synthetase (*ACSS2*)**,** both of which can drive Acetyl-CoA synthesis through conversion of citrate or acetate, respectively. This is additionally necessary for fatty acid synthesis as well as oxidation of pyruvate for energy production by the tricarboxyllic acid (TCA) cycle, which was also found enriched (TCA metabolism, fold enrichment=3.58, *p*=4.93E-02). Branched-chain amino acid (BCAA) catabolism serves as a source of energy and can additionally contribute to FA synthesis (fold enrichment=5.9, *p*=2.66E-04) (469). Genes involved in cholesterol biosynthesis (fold enrichment=3.69, *p*=9.66E-03) were also upregulated during lactation. On the other hand, several 40S and 60S ribosomal proteins as well as eukaryotic translation initiation factor subunits were induced during lactation suggesting a potential elevation in protein synthesis (cytoplasmic translation, fold enrichment=3.63, *p*=3.83E-03). Strikingly, the majority of genes in each of these processes were also induced by TSA and SAHA treatment (fold enrichment >2, p<0.05), with the exception of the cytoplasmic translation GO term, which was largely unregulated **(Supplementary Figure 9B)**.

GO terms analysis of downregulated genes at 72 h postpartum revealed enrichment of interferon response processes, particularly ones associated with type I interferon IFN-alpha, which had the highest fold of enrichment (response to interferon-alpha, fold enrichment=6.16, *p*=1.13E-03) **(Supplementary Figure 9C).** Interestingly, several genes in this GO term were also downregulated by TSA and/or SAHA (fold enrichment >2, p<0.05) **(Supplementary Figure 9D),** including interferon-induced transmembrane proteins (*IFITM1, 2* and *3*), interferon-induced protein with tetratricopeptide repeats (*IFIT2* and *3*) and the 2'-5'-oligoadenylate synthetase 1, *OAS1*.

IPA identified sterol regulatory element-binding protein 1 (SREBP-1 encoded by *SREBF1*) as the most activated TF by TSA and SAHA (z-score= 4.347, overlap p-value= 6.96E-03) based on the direction of regulation of its downstream target genes (**Supplementary Table 2)**. MS-275 treatment was also predicted to result in activation of SREBP-1, although this was not statistically significant (z-score=1.98, overlap p-value= 1). These data are consistent with an increase in *SREBF1* expression by TSA, SAHA and MS-275 **(Lipid biosynthetic process, Supplementary Figure 2A)**. Similarly, *SREBF1* expression was found increased during lactation and IPA predicted activation of this TF at 48 h (z-score=2.498, overlap p-value=1.14E-07), 72 h (z-score=2.802, overlap p-value=2.25E-

10) and 7 days postpartum (z-score=2.967, overlap p-value= 4.88E-11) as previously reported (468). Using IPA, we generated a list of 44 known SREBP-1 target genes whose expression was regulated in a direction consistent with activation of SREBP-1 at 7 days postpartum (38 UP and 6 DOWN), and examined how their expression changes in response to HDACi treatment in MCF-7 cells **(Figure 5A)**. Remarkably, we observed that many SREBP-1 target genes induced during lactation were also significantly upregulated by TSA (28 genes, fold enrichment=3.99, p-value=4.27E-13) and SAHA (26 genes, fold enrichment=4.08, p-value=6.03E-12), and to a weaker degree by MS-275 (13 genes, fold enrichment=1.83, p-vaue=0.017) **(Figure 5A)**. Functions of overlapping genes included FA activation (*ACSS2, ACSL1, ACSL4*), FA modification (*ELOVL6*), FA oxidation (*EHHADH*), FA desaturation (*SC5D*), FA binding (*DBI*), triglyceride synthesis (*LPIN1, GPAM*), FA transport (*SCARB1*), as well as the cholesterol biosynthesis enzymes (*DHCR7, FDFT1, FDPS, HMGCR, HMGCS1, IDI1, LSS, MVD, SQLE*).

Next, we confirmed HDACi-mediated induction of expression of a panel of SREBP-1 target genes, including *SREBF1,* which positively autoregulates its own expression (470, 471), by RT-qPCRs in all MCF-7, T-47D and ZR-75-1 cells **(Figure 5B)**. Induction of SREBP-1 protein levels by HDACis was also confirmed in all three cell lines **(Figure 5C).** Because acetylation has been reported to increase SREBP-1 protein stability by competing with ubiquitination (294), we asked whether HDACis modulated acetylation of this TF. We performed acetyl-lysine immunoprecipitations followed by SREBP-1 immunoblots in MCF-7 cells after 3 h of treatment with HDACis, when protein expression of this TF was not yet induced, and indeed observed an increase in its acetylation **(Figure 5D)**. This result indicated that in addition to increasing mRNA expression of *SREBF1*, HDACis may also post-transcriptionally increase expression of this TF by enhancing its acetylation.

Together, these findings suggest that the global changes in metabolism of mammary epithelial cells during lactation, most notably activation of lipid and cholesterol biosynthesis, can be partially recapitulated by hydroxamic acid HDACis, suggesting that these inhibitors can propagate lactogenic differentiation signals in breast cancer cells.

Expression of FOXA1, ERα and GATA3 is reduced during lactation in the mouse mammary gland

Given the roles of ERα, FOXA1 and GATA3 in ductal cell proliferation and mammary gland duct development (80, 95, 108), we assessed whether their expression is suppressed during

lactogenic differentiation of the mammary gland as well as upon HDACi treatment in cultured cells. Expression levels of the corresponding proteins was assessed during different stages of mouse mammary gland development: puberty (6 weeks of age), pregnancy (15 days through:dP15), and lactation (dL5, 10, and 15) **(Figure 6)**. As expected, in the pubertal mammary gland of virgin mice, most luminal epithelial cells lining ducts showed positive nuclear immunostaining for ERα, FOXA1 and GATA3. We note that a small fraction of luminal ductal cells, however, stained negatively for these TFs. Strikingly, though, we found that expression levels of ERα, FOXA1 and GATA3 are reduced in luminal cells at different time points during pregnancy and lactation (**Figure 6)**. Together, these results indicate that the similitudes between phenotypes observed in HDACi-treated breast cancer cells and during lactogenic differentiation in the normal mouse mammary gland extend to reduction of expression of luminal TFs and suggest a potential role for HDAC proteins in the suppression of lactogenic differentiation.

HDACis exert part of their actions through p300/CBP

HDACi treatment is expected to increase protein acetylation via histone acetyltransferase proteins, which belong to 5 main families, GNAT, MYST, p300/CBP, basal TFs, and steroid receptor coactivators. Interestingly, several of these HATs were downregulated by HDACi treatment in MCF-7 cells **(Figure 7A).** We focused our attention on p300/CBP as potential candidates for mediating the changes observed by HDACi treatment for two reasons: (i) p300/CBP act as cofactors of ERα allowing transcriptional activation of E2-target genes in breast cancer cells (43, 44, 241), and (ii) a small molecule chemical inhibitor of p300/CBP (C646) has been shown to reduce TSA-mediated histone H3 hyperacetylation (472). We subjected MCF-7 cells to DMSO or C646 pre-treatment for 6 h followed by treatment with HDACis for 16 h and monitored expression levels of *ESR1, FOXA1, GATA3* as well as *SREBF1* by RT-qPCR and immunoblots **(Figure 7B and 7C)**. While C646 pre-treatment alone resulted in downregulation of expression of all four TFs, it partially alleviated transcriptional repression of *ESR1* and *GATA3* and reduced induction of *SREBF1* transcript upon co-treatment with HDACis **(Figure 7B and 7C).** Interestingly, repression of *FOXA1* transcription by HDACis was still detectable with C646 pre-treatment, suggesting a lack of requirement for p300/CBP in HDACi-mediated transcriptional repression of this TF.

C646 co-treatment also relieved repression of ERα and GATA3 protein expression by HDACis, whereas FOXA1 protein levels were strongly repressed irrespective of HDACi

treatment with C646 co-treatment **(Figure 7C)**. In addition, HDACi-mediated induction of SREBP-1 protein expression was abrogated by C646 co-treatment **(Figure 7C)**. Intriguingly, these data suggested that p300/CBP act as positive regulators of expression of ERα and GATA3, a mode of action that was counterintuitive to their seemingly inhibitory role in the presence of HDACis. This result suggests that increased acetylation of protein substrates (1) mediates the actions of HDACis and (2) has a different outcome on gene expression than when acetylation is kept in a state of equilibrium.

Lastly, reduced expression of *SREBF1* in the presence of C646 prompted us to examine whether this was due to loss of acetylation of this TF, which normally stabilizes the protein and prevents its degradation (294). Protein acetylation levels were monitored in the absence or presence of C646 and/or MG132 to stabilize SREBP-1 protein expression **(Figure 7D)**. We found that C646 pre-treatment prevented TSA-mediated induction of SREBP-1 acetylation in the presence of MG132 **(Figure 7D).** This result suggests that HDACis induce SREBF1 acetylation in a p300/CBP-dependent manner.

DISCUSSION

HDACis have been reported to alter gene expression patterns and inhibit cell proliferation in breast cancer cells (342, 421, 422). Using gene expression microarrays, Reid et al. originally reported that 88% of previously identified E2-upregulated genes are repressed by TSA or valproic acid treatment in MCF-7 cells cultured in hormone-replete media (350). A subsequent study revealed through GRO-seq and microarray analysis that HDACi-repressed genes in HER2-positive breast cancer cells display higher RNA Pol II pausing indices resulting in a block in transcription elongation (473). Additionally, HDACi-repressed genes were found to be associated with chromatin organization and transcriptional regulation processes (473). The work presented herein complements previous studies in documenting global suppression of E2 signalling by TSA in the context of combined treatment. Indeed, we observed that the majority of E2 target genes lose E2 regulation in the presence of TSA, consistent with downregulation of ERα expression. Our study additionally provides new insights into the molecular basis of HDACi-mediated transcriptional reprogramming of ERα-positive luminal breast cancer cells via comparative RNA-Seq analysis of the effects of three different HDACis, i.e. TSA, SAHA and MS-275. In keeping with previous reports in different cell types (333, 437, 438), we observed a balanced impact of all three inhibitors on the numbers of genes activated or repressed. HDACis inhibit HDAC enzymes and therefore shift the steady state balance toward elevated histone acetylation levels. This is proposed to relieve compacted DNA, thereby enhancing TF accessibility and promoting gene transcription. However, HDACi treatment also alters the expression of a large number of TFs and chromatin-modifying factors. Our observation of loss of expression of several HATs at 8 h after treatment with HDACis, suggested the existence of negative feed-back mechanisms controlling protein acetylation. We note that HDACis have been previously reported to cause reduced expression of several HAT proteins, including p300 and CBP, while simultaneously resulting in H3K9/K14 deacetylation which correlates with transcriptional repression (474). As expected, TSA and SAHA, two hydroxamic acid inhibitors, shared similar mechanisms of action whereas MS-275 exerted divergent effects on gene regulation in MCF-7 cells, as illustrated by differences in GO term enrichment and in TFBS enrichment in regulated genes.

While loss of expression of ERα was reported before (342, 349, 350, 353, 355, 356), our transcriptome analysis revealed reduced expression of several luminal TFs, including luminal master regulators GATA3, FOXA1, XBP1 and SPDEF **(Chapter 2)**. Our finding that *de novo* protein synthesis was not required for HDACi-mediated transcriptional repression of

these TFs refutes the possibility of increased expression of (a) transcriptional repressor(s) by HDACis. It remains possible, however, that HDACis repress transcription of luminal TFs through decreased expression of upstream transcriptional activators and/or modulation of activity of upstream regulatory factors through changes in their acetylation. Epigenetic regulation of the genes encoding these TFs by HDACis is a plausible hypothesis given our previous findings that transcriptional repression of *ESR1* was associated with a decrease in histone acetylation at promoters A, B and C (349). In addition, our finding that expression of a large number of histone acetyltransferases (*EP300, CREBBP, KAT2A, KAT5, KAT6B, KAT7, KAT8*) and chromatin remodelling factors (SWI-SNF subunits: *SMARCA4, SMARCAD1, SMARCB1, SMARCC1, SMARCC2, SMARCD1, SMARCD2, SMARCE1*; ISWI complexes *BAZ1A, BAZ1B, CHRAC1, RSF1*; INO80/SWR complexes: *INO80, YY1*) was suppressed by HDACi treatment **(Supplementary Figure 2A)** is compatible with loss of expression of genes dependent on activating signals for expression. In addition, links between HDACi treatment and DNA methylation have been previously described. Global DNA methylation levels were found reduced following HDACi treatment due to reduced expression/increased degradation of DNMT1 (475, 476). Conversely, inhibition of H3K27 methylating enzymes EZH1/2 was found to partially relieve transcriptional repression by HDACis (477). In our dataset, *DNMT1* and *DNMT3A* were repressed upon treatment with all three HDACis, whereas TSA and SAHA (but not MS-275 under our experimental conditions) induced expression of *DNMT3B* and repressed *EZH1*. It remains to be determined whether histone/DNA methylation is altered at the promoters of luminal TFs, and more globally, of HDACi-responsive genes in breast cancer cells. Of note, expression of the insulator protein CTCF was repressed by all three inhibitors. Although CTCF acetylation has not been reported, HDAC activity is recruited by CTCF and contributes to its insulator function (478). Additional experiments are needed to examine whether transcriptional repression by HDACis is mediated via inhibition of CTCF expression and/or function.

Given the positive regulatory role of FOXA1 in *ESR1* and *GATA3* expression **(chapter 2)**, we reasoned that its downregulation by HDACis may result in their decreased expression. However, reduction of FOXA1 protein expression occurred only after reduction of ERα and GATA3 protein levels **(Figure 3B)**. Furthermore, overexpression of FOXA1 did not restore ERα or GATA3 mRNA or protein levels in the presence of HDACis **(Supplementary Figure 8)**. While these results argued against this hypothesis, it remains possible that acetylation of FOXA1 is altered by HDACi treatment at early time points when its expression is not yet

affected, and this in turn would alter its binding to promoters of *ESR1* and *GATA3*. Indeed, FOXA1 acetylation has been reported to reduce its binding to DNA as well as its capacity to remodel chromatin (298). Because ERα, FOXA1 and GATA3 are co-expressed in luminal breast cancer cells, and their HDACi-mediated transcriptional repression occurred concomitantly, it remains possible that HDAC proteins regulate their expression through (a) common upstream regulator(s) to be identified. Future shRNA/CRISPR-cas9 targeted or genome-wide screens will be essential to identify regulators that mediate the repressive effects of HDACis on expression of luminal TFs.

The rationale behind the use of epigenetic therapy for cancer treatment has been its capacity to rewire cancer cells towards differentiation (322). Our finding that several metabolic processes characteristic of lactation in the human mammary gland, including SREBP-1-driven lipid and cholesterol biosynthesis, are also induced by HDACis suggested that these inhibitors can induce features of lactogenic differentiation. It remains unclear whether marked reduction of ERα, FOXA1 and GATA3 protein on one hand and induction of SREBP-1 on the other hand happen in parallel or are intricately connected. In the case of the latter, is loss of expression of luminal TFs a prerequisite for initiation of lactogenic differentiation, or is it a consequence? Using immunohistochemical analysis, Bernardo et al. showed that expression of FOXA1 and ERα is reduced in alveoli of virgin, pregnant and lactating mice (95). Importantly, conditional knockouts of *FOXA1* in the mouse mammary gland did not preclude its capacity to produce alveoli even though ductal invasion was impaired (95). In fact, *FOXA1* heterozygous null glands, which are consequently ERα^{low}, produced more alveoli, suggesting that FOXA1 might inhibit alveologenesis (95). Future experiments will address the possibility that HDACis enrich for an alveolar lineage subpopulation following treatment of breast cancer cells. It will also be important to examine whether HDAC proteins are differentially expressed during different stages of mammary gland development.

High expression of cholesterol biosynthesis enzymes has been shown to be associated with poor prognosis in breast cancer (479, 480). In addition, a recent study has shown that this pathway is upregulated in breast cancer cell lines resistant to E2-deprivation therapy. These cells continue to proliferate through the actions of the oxysterols 25- and 27-hydroxycholesterol, which have a steroid structure similar to that of E2 and can activate ERα signalling (481). In the context of HDACi treatment, increased cholesterol synthesis is not expected to increase proliferation as these inhibitors also repress expression of ERα, which is

their main survival signal, and eventually undergo apoptosis. Additional phenotypic assays are needed to confirm that breast cancer cells undergo lactogenic differentiation. Particularly, a differentiation assay will be optimised whereby breast cancer cells are treated with a dose low enough to allow the cells to survive for an extended period of time so that their differentiation state can be examined. Moreover, to examine whether there is a link between loss of luminal TFs and induction of lactogenic differentiation markers, we will overexpress these factors together or in different combinations and determine whether that abrogates the HDACi-mediated differentiation response.

While the impact of MS-275 on expression of SREBP-1-driven lipid/cholesterol genes was weaker than that of TSA and SAHA at 8 h of treatment, this result indicated either the implication of additional HDACs than HDAC1 and 3 in regulation of expression of differentiation genes or the need for prolonged treatment with MS-275. We note that at the drug concentrations used in this study, both TSA and SAHA fully inhibit HDAC1 and 3 activities, whereas activities of other HDACs are only partially affected, based on a previous *in vitro* study (327). The finding that many genes are uniquely regulated by MS-275 but not TSA or SAHA possibly suggests a different impact of inhibition of additional HDACs. Thus, it will be interesting to examine expression of MS-275 unique genes using TSA and SAHA concentrations at which only HDAC1 and 3 are inhibited. It is noteworthy that HDACi selectivity *ex vivo* might differ from that reported in *in vitro* studies, which could explain differences in gene regulation. Ultimately, it will be critical to investigate the contribution of inhibition of each HDAC using shRNA or CRISPR-cas9 technology.

Lastly, because chemical inhibition of p300/CBP partially relieved HDACi-mediated repression of ERα and GATA3 and prevented increase of SREBP-1 acetylation, this indicated that (1) additional HATs are potentially involved in regulation of luminal TFs, (2) modulated acetylation of HDAC targets contributes to expression of luminal TFs and (3) p300/CBP acetylates SREBP-1 to stabilize it thereby contributing to its lactogenic transcriptional program. We propose a two-edged-sword mechanism whereby HDACis can regulate luminal TFs at both the transcriptional and post-transcriptional levels partly through p300/CBP activity. In the presence of HDACis, the balance is shifted resulting in hyperacetylation **(Figure 8)**. This may involve (1) increased acetylation of an activator of luminal TF expression that inhibits its function, (2) increased acetylation of a repressor of luminal TF expression that activates its function, (3) increased expression of inhibitory miRNAs, (4) increased expression of proteins that can regulate turn-over of luminal TFs such as ubiquitin ligases, and (5)

decreased expression of upstream activators. On the other hand, in the presence of a p300/CBP inhibitor alone, the balance is shifted resulting in hypoacetylation. This results in repression of luminal TFs, potentially through deacetylation of their promoters. Counterintuitively, concomitant inhibition of HDACs and p300/CBP does not result in further repression of luminal TFs as one would expect, likely because under these conditions, acetylation/deacetylation is in steady state.

The picture that emerges from this study highlights HDACs as regulators of luminal TFs and lactogenic differentiation genes in breast cancer cells. Ultimate identification of upstream regulators and perturbing their function will help determine whether they can transcriptionally reprogram breast cancer cells from one subtype to another or alter their differentiation state into a less tumourigenic phenotype. Identification of the mechanisms of action of HDACis will also help determine whether ERα-positive breast tumours may benefit from combining HDACis with hormonal treatment for a more effective suppression of estrogen signalling and tumour growth.

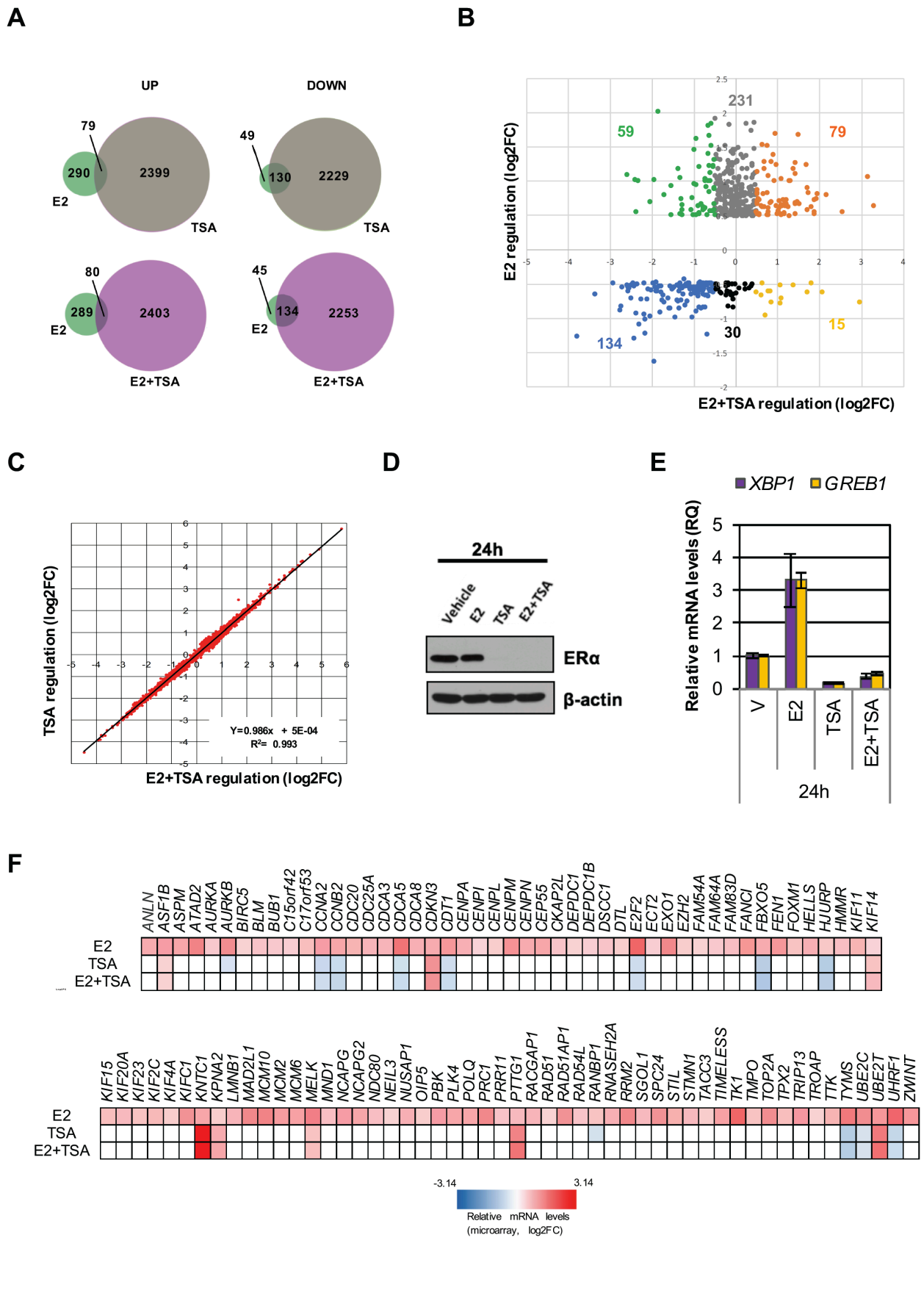

A
UP
DOWN
79
290
2399
E2
TSA
49
130
2229
TSA
80
289
2403
E2
E2+TSA
45
134
2253
E2
E2+TSA
B
231
59
79
134
30
15
E2 regulation (log2FC)
E2+TSA regulation (log2FC)
C
TSA regulation (log2FC)
Y=0.986x + 5E-04
R²= 0.993
E2+TSA regulation (log2FC)
D
24h
Vehicle
E2
TSA
E2+TSA
ERα
β-actin
E
XBP1
GREB1
Relative mRNA levels (RQ)
V
E2
TSA
E2+TSA
24h
F
ANLN ASF1B ASPM ATAD2 AURKA AURKB BIRC5 BLM BUB1 C15orf42 C17orf53 CCNA2 CCNB2 CDC20 CDC25A CDCA3 CDCA5 CDCA8 CDKN3 CDT1 CENPA CENPI CENPL CENPM CENPN CEP55 CKAP2L DEPDC1 DEPDC1B DSCC1 DTL E2F2 ECT2 EXO1 EZH2 FAM54A FAM64A FAM83D FANCI FBXO5 FEN1 FOXM1 HELLS HJURP HMMR KIF11 KIF14
E2
TSA
E2+TSA
KIF15 KIF20A KIF23 KIF2C KIF4A KIFC1 KNTC1 KPNA2 LMNB1 MAD2L1 MCM10 MCM2 MCM6 MELK MND1 NCAPG NCAPG2 NDC80 NEIL3 NUSAP1 OIP5 PBK PLK4 POLQ PRC1 PRR11 PTTG1 RACGAP1 RAD51 RAD51AP1 RAD54L RANBP1 RNASEH2A RRM2 SGOL1 SPC24 STIL STMN1 TACC3 TIMELESS TK1 TMPO TOP2A TPX2 TRIP13 TROAP TTK TYMS UBE2C UBE2T UHRF1 ZWINT
E2
TSA
E2+TSA
-3.14
3.14
Relative mRNA levels
(microarray, log2FC)

Figure 1. TSA suppresses E2 signalling in MCF-7 cells.
MCF-7 cells cultured in hormone-depleted media were treated with vehicle, E2 (25nM), TSA (300nM), or E2+TSA for 24 hours.
(A) Microarray gene expression profiling was performed (N=3). Venn diagrams showing the overlap of upregulated/downregulated genes (FC>1.4 or <-1.4 and p<0.01 compared to vehicle control) between the E2 and TSA/E2+TSA datasets.
(B) Folds of gene regulation in the E2 condition (FC>1.4 or <-1.4 and p<0.01) were plotted against folds of gene regulation in the E2+TSA condition. Six categories of genes can be identified: (i) Genes upregulated by E2 and E2+TSA (orange dots, 79 genes), (ii) genes upregulated by E2 and downregulated by E2+TSA (green dots, 59 genes), (iii) genes downregulated by E2 and upregulated by E2+TSA (yellow dots, 15 genes), (iv) genes downregulated by E2 and E2+TSA (blue dots, 134 genes), (v) genes upregulated by E2 but not significantly regulated by E2+TSA (grey dots, 231 genes), and (vi) genes downregulated by E2 but not significantly regulated by E2+TSA (black dots, 30 genes).
(C) Folds of gene regulation by TSA were plotted against folds of genes regulation by E2+TSA. A correlation coefficient R^2 is calculated.
(D) Western analysis showing ERα protein levels (N=3). β-actin serves as a loading control.
(E) Relative mRNA levels of *XBP1* and *GREB1* were assessed by RT-qPCRs. Error bars represent standard deviation of three technical replicates (N=1).
(F) Relative mRNA levels (microarray data) of 99 cell cycle/proliferation genes is shown as a heatmap. All gene regulations with a FC>1.4 or <-1.4 are statistically significant (N=3, p<0.01 compared to vehicle control using Fisher's exact test)

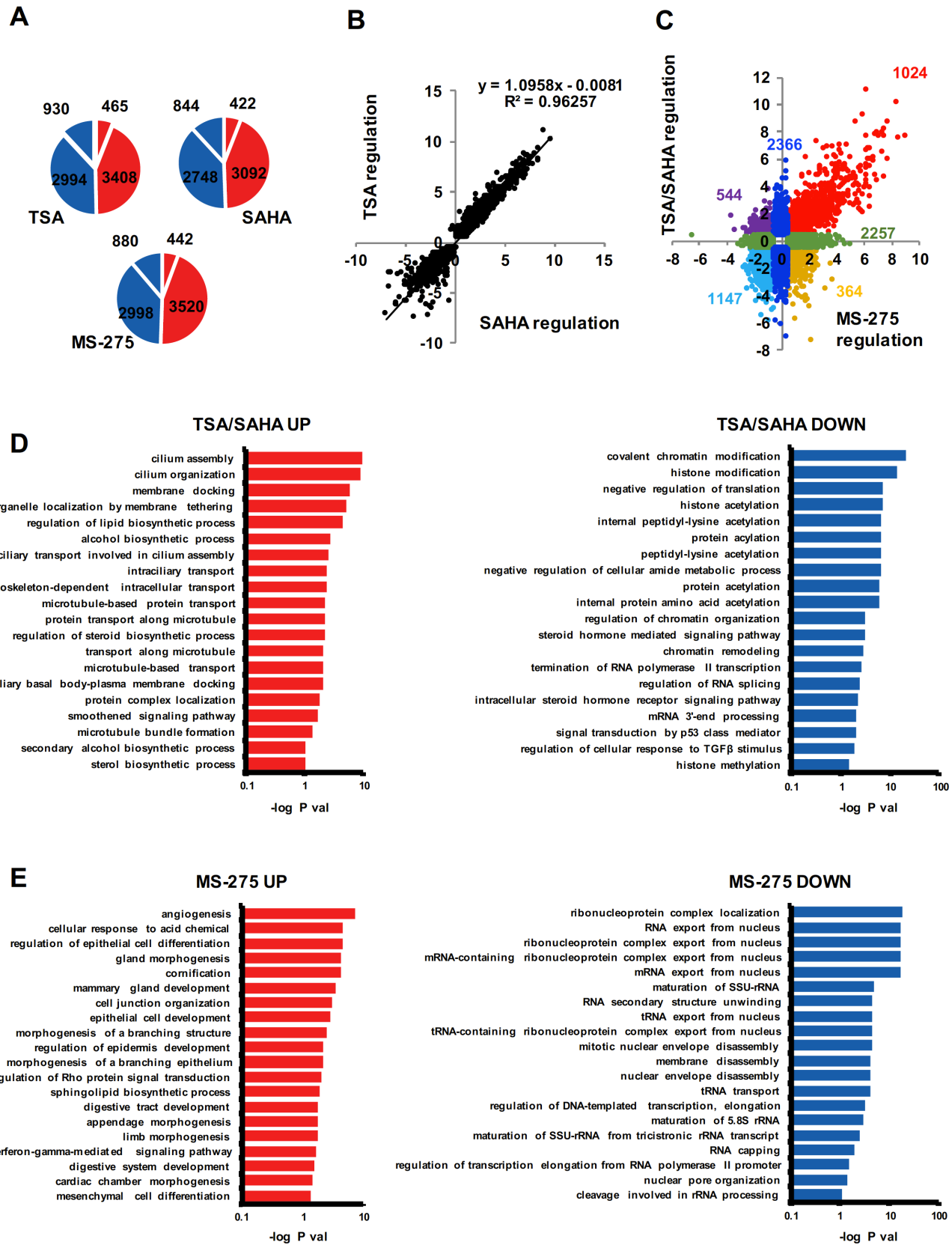

A
930
465
2994
3408
TSA
844
422
2748
3092
SAHA
880
442
2998
3520
MS-275
B
y = 1.0958x - 0.0081
R² = 0.96257
TSA regulation
SAHA regulation
C
TSA/SAHA regulation
MS-275 regulation
1024
2366
544
2257
1147
364
D
TSA/SAHA UP
cilium assembly
cilium organization
membrane docking
organelle localization by membrane tethering
regulation of lipid biosynthetic process
alcohol biosynthetic process
intraciliary transport involved in cilium assembly
intraciliary transport
cytoskeleton-dependent intracellular transport
microtubule-based protein transport
protein transport along microtubule
regulation of steroid biosynthetic process
transport along microtubule
microtubule-based transport
ciliary basal body-plasma membrane docking
protein complex localization
smoothened signaling pathway
microtubule bundle formation
secondary alcohol biosynthetic process
sterol biosynthetic process
-log P val
TSA/SAHA DOWN
covalent chromatin modification
histone modification
negative regulation of translation
histone acetylation
internal peptidyl-lysine acetylation
protein acylation
peptidyl-lysine acetylation
negative regulation of cellular amide metabolic process
protein acetylation
internal protein amino acid acetylation
regulation of chromatin organization
steroid hormone mediated signaling pathway
chromatin remodeling
termination of RNA polymerase II transcription
regulation of RNA splicing
intracellular steroid hormone receptor signaling pathway
mRNA 3'-end processing
signal transduction by p53 class mediator
regulation of cellular response to TGFβ stimulus
histone methylation
-log P val
E
MS-275 UP
angiogenesis
cellular response to acid chemical
regulation of epithelial cell differentiation
gland morphogenesis
cornification
mammary gland development
cell junction organization
epithelial cell development
morphogenesis of a branching structure
regulation of epidermis development
morphogenesis of a branching epithelium
regulation of Rho protein signal transduction
sphingolipid biosynthetic process
digestive tract development
appendage morphogenesis
limb morphogenesis
interferon-gamma-mediated signaling pathway
digestive system development
cardiac chamber morphogenesis
mesenchymal cell differentiation
-log P val
MS-275 DOWN
ribonucleoprotein complex localization
RNA export from nucleus
ribonucleoprotein complex export from nucleus
mRNA-containing ribonucleoprotein complex export from nucleus
mRNA export from nucleus
maturation of SSU-rRNA
RNA secondary structure unwinding
tRNA export from nucleus
tRNA-containing ribonucleoprotein complex export from nucleus
mitotic nuclear envelope disassembly
membrane disassembly
nuclear envelope disassembly
tRNA transport
regulation of DNA-templated transcription, elongation
maturation of 5.8S rRNA
maturation of SSU-rRNA from tricistronic rRNA transcript
RNA capping
regulation of transcription elongation from RNA polymerase II promoter
nuclear pore organization
cleavage involved in rRNA processing
-log P val

Figure 2. Transcriptome profiling of HDACi-treated MCF-7 cells.
(A) Transcriptome profiles of MCF-7 cells treated with TSA (300nM), SAHA (5μM), or MS-275 (5μM) for 8 h were assessed by RNA-Seq (N=2). Genes were considered significantly regulated (UP in red and DOWN in blue) if they met a fold change cutoff of 1.4 (as compared to the vehicle control) and a q value cutoff of 0.01. Broken pie chart slices represent numbers of regulated TFs across each dataset.
(B) Folds of gene regulation in the TSA condition (FC of 1.4, q<0.01) were plotted against folds of gene regulation in the SAHA condition to assess the degree of similitude.
(C) Folds of regulation of TSA/SAHA responsive genes (FC of 1.4 and q<0.01 for both datasets) were plotted against their folds of regulation by MS-275 (no cutoffs). Folds of regulation of MS-275-responsive genes (FC of 1.4 and q<0.01) were also plotted against those by TSA/SAHA (no cutoffs) to represent genes uniquely regulated by MS-275. This analysis reveals four different gene categories: (i) Genes that are regulated in the same direction by TSA, SAHA and MS-275 (UP in red (1024) and DOWN in light blue (1147)); (ii) Genes that are regulated in opposite directions by TSA/SAHA on one hand and MS-275 on the other hand (UP TSA/SAHA and DOWN MS-275 in purple (544), DOWN TSA/SAHA and UP MS-275 in orange (364)); (iii) genes that are uniquely significantly regulated by TSA/SAHA are shown in dark blue (2366); (iv) genes that are uniquely significantly regulated by MS-275 shown in green (2257).
GO term enrichment analysis for TSA/SAHA (D) and MS-275 (E) up- and downregulated genes was performed. The top 20 enriched biological processes (fold enrichment >2) are shown on the y-axis vs their –log *p* value on the x-axis.

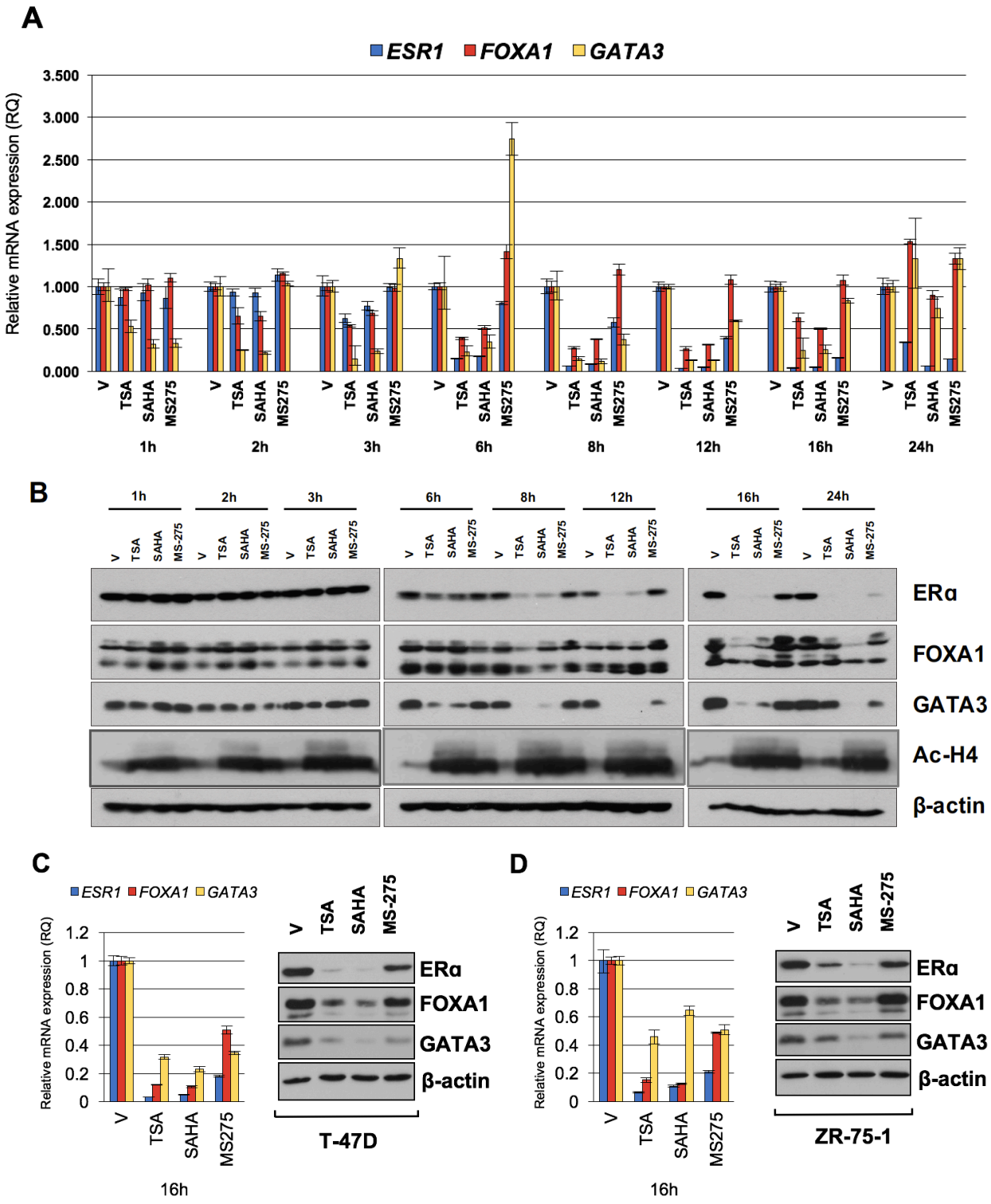

A
ESR1 FOXA1 GATA3
Relative mRNA expression (RQ)
3.500
3.000
2.500
2.000
1.500
1.000
0.500
0.000
V TSA SAHA MS275
1h 2h 3h 6h 8h 12h 16h 24h
B
1h 2h 3h 6h 8h 12h 16h 24h
V TSA SAHA MS-275
ERα
FOXA1
GATA3
Ac-H4
β-actin
C
ESR1 FOXA1 GATA3
Relative mRNA expression (RQ)
1.2
1
0.8
0.6
0.4
0.2
0
V TSA SAHA MS275
16h
V TSA SAHA MS-275
ERα
FOXA1
GATA3
β-actin
T-47D
D
ESR1 FOXA1 GATA3
Relative mRNA expression (RQ)
1.2
1
0.8
0.6
0.4
0.2
0
V TSA SAHA MS275
16h
V TSA SAHA MS-275
ERα
FOXA1
GATA3
β-actin
ZR-75-1

Figure 3. HDACis repress expression of ERα, FOXA1 and GATA3 in luminal breast cancer cell lines.
(A) MCF-7 cells were treated with vehicle (V), TSA (300nM), SAHA (5μM), or MS-275 (5μM) for 1, 2, 3, 6, 8, 12, 16, and 24 h. Relative mRNA levels of *ESR1*, *FOXA1* and *GATA3* were examined by RT-qPCRs.
(B) MCF-7 protein lysates from the same experiment shown in (A) were probed for ERα, FOXA1 and GATA3 expression. Histone H4 acetylation served as a positive control for inhibition of HDAC activity. β-actin serves as a loading control.
Relative mRNA and protein levels of ERα, FOXA1 and GATA3 were examined in T-47D (C) and ZR-75-1 (D) cells treated with TSA (300nM), SAHA (5μM), or MS-275 (5μM) for 16 h.
Error bars represent standard deviation of three technical replicates (N=1).

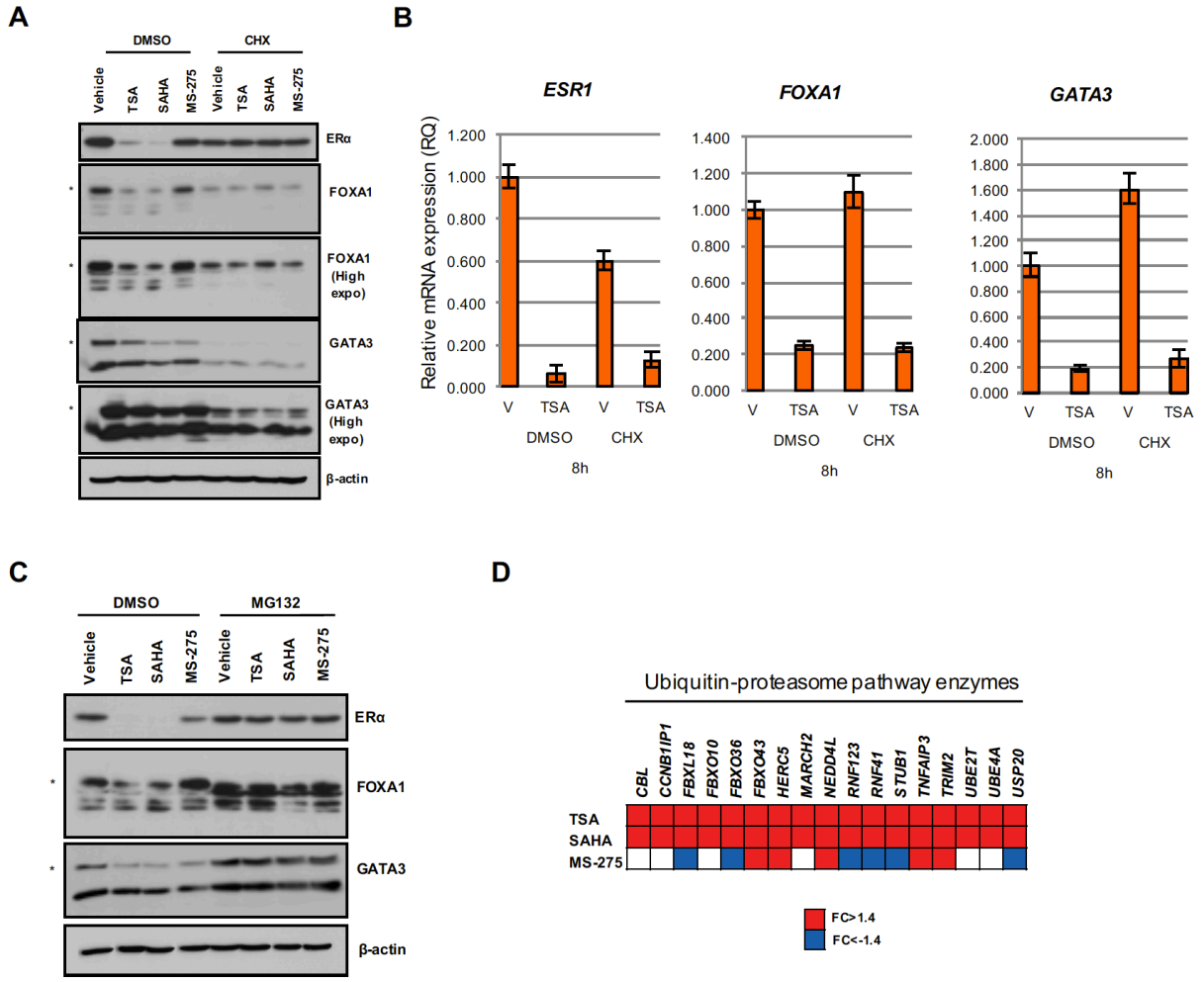


Figure 4. HDACi-mediated repression of ERα, FOXA1 and GATA3 is dependent on *de novo* protein synthesis and degradation by the proteasome.

(A) MCF-7 cells were pre-treated with the de-novo protein synthesis inhibitor, cycloheximide (CHX; 10 μg/ml) or vehicle (DMSO) for one hour, then treated with vehicle, TSA (300nM), SAHA (5μM), or MS-275 (5μM) for 16 hours. Protein levels of ERα, FOXA1 and GATA3 were assessed by western blots. β-actin serves as a loading control. The band corresponding to the correct molecular weight for the probed protein is depicted by an asterisk (*).

(B) MCF-7 cells were pre-treated with cycloheximide (10 μg/ml) for one hour then treated with vehicle or TSA (300nM) for 8 hours. Relative mRNA levels of *ESR1*, *FOXA1* and *GATA3* were examined by RT-qPCRs. Error bars represent standard deviation of three technical replicates (N=1).

(C) MCF-7 cells were pre-treated with MG132 (10 μM) for three hours then treated with vehicle, TSA (300nM), SAHA (5μM), or MS-275 (5μM) for 16 hours. Protein expression of ERα, FOXA1 and GATA3 was examined by western analysis.

(D) Relative mRNA levels of several ubiquitin-proteasome pathway enzymes were assessed using the HDACi-treated MCF-7 RNA-seq datasets (N=2, q<0.01).

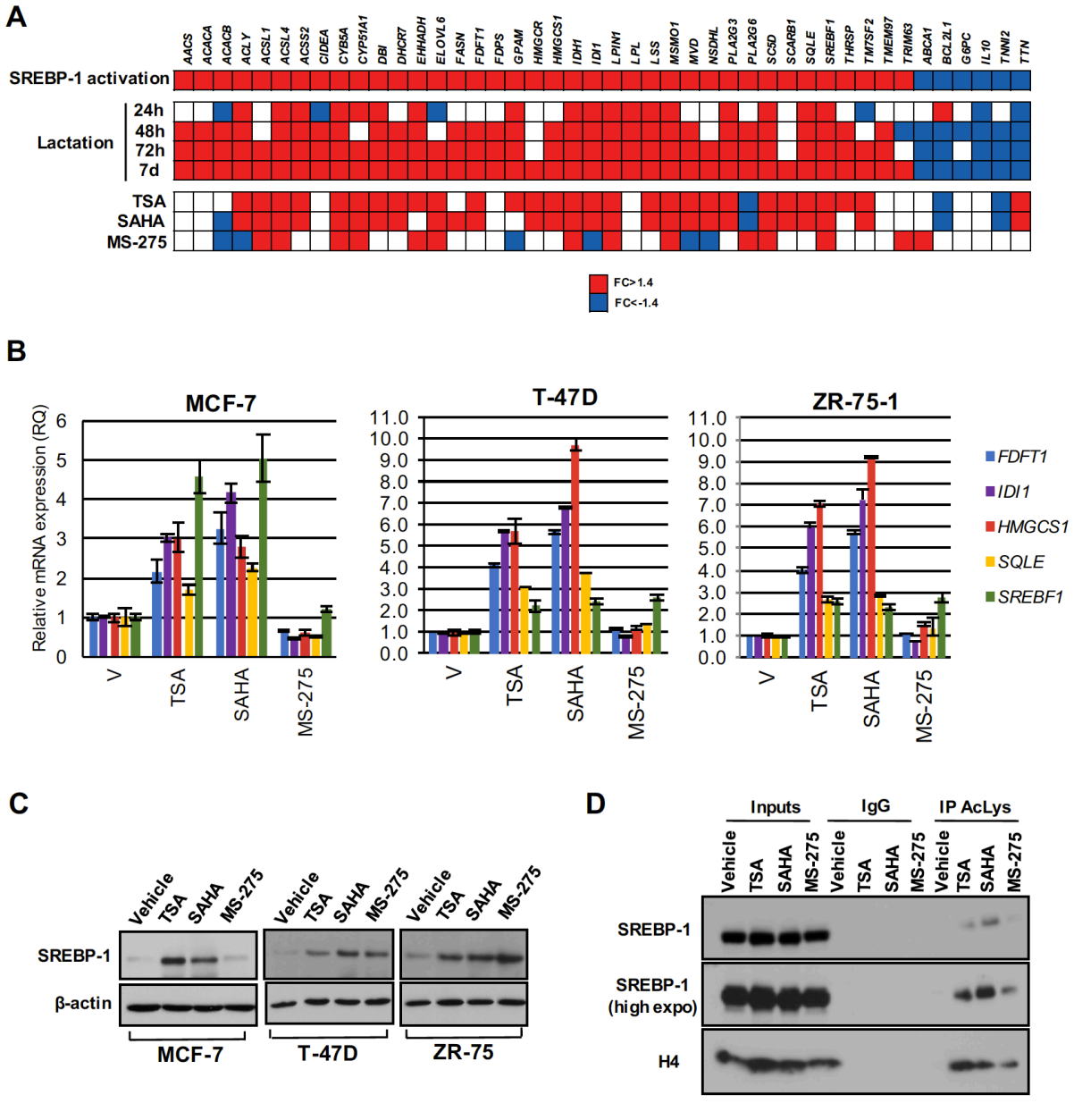

A
AACS ACACA ACACB ACLY ACSL1 ACSL4 ACSS2 CIDEA CYB5A CYP51A1 DBI DHCR7 EHHADH ELOVL6 FASN FDFT1 FDPS GPAM HMGCR HMGCS1 IDH1 IDI1 LPIN1 LPL LSS MSMO1 MVD NSDHL PLA2G3 PLA2G6 SCD SCARB1 SQLE SREBF1 THRSP TM7SF2 TMEM97 TRIM63 ABCA1 BCL2L1 G6PC IL10 TNNI2 TTN
SREBP-1 activation
Lactation
24h
48h
72h
7d
TSA
SAHA
MS-275
FC>1.4
FC<-1.4
B
MCF-7
T-47D
ZR-75-1
Relative mRNA expression (RQ)
V
TSA
SAHA
MS-275
FDFT1
IDI1
HMGCS1
SQLE
SREBF1
C
Vehicle
TSA
SAHA
MS-275
SREBP-1
β-actin
MCF-7
T-47D
ZR-75
D
Inputs
IgG
IP AcLys
SREBP-1
SREBP-1 (high expo)
H4

Figure 5. HDACis induce expression of the lactogenic differentiation marker SREBP-1 and its downstream targets.
(A) Ingenuity pathway analysis reveals activation of the TF SREBP-1 at 48h, 72h, and 7 days postpartum. Top row shows the direction of regulation of SREBP-1 targets that predict activation of this TF. Relative expression levels of 44 SREBP-1 target genes are shown at indicated time points during lactation and following HDACi treatment in MCF-7 cells (RNA-seq, N=2, $q<0.01$).
(B) Relative mRNA levels of a panel of SREBP-1 target genes were examined by RT-qPCRs in MCF-7, T-47D and ZR-75 cells treated with vehicle, TSA (300nM), SAHA (5μM), or MS-275 (5μM) for 16 hours. Error bars represent standard deviation of three technical replicates (N=1).
(C) Protein levels of SREBP-1 were examined by western blots in MCF-7, T-47D and ZR-75-1 cells treated with HDACis for 16 hours.
(D) MCF-7 cells were treated with the indicated HDACis for three hours. Cells lysates were immunoprecipitated using an anti-acetyl-lysine (AcLys) antibody. The IP was then probed for SREBP-1 by immunoblotting. Histone H4 serves as a positive control for increased acetylation by HDACis.

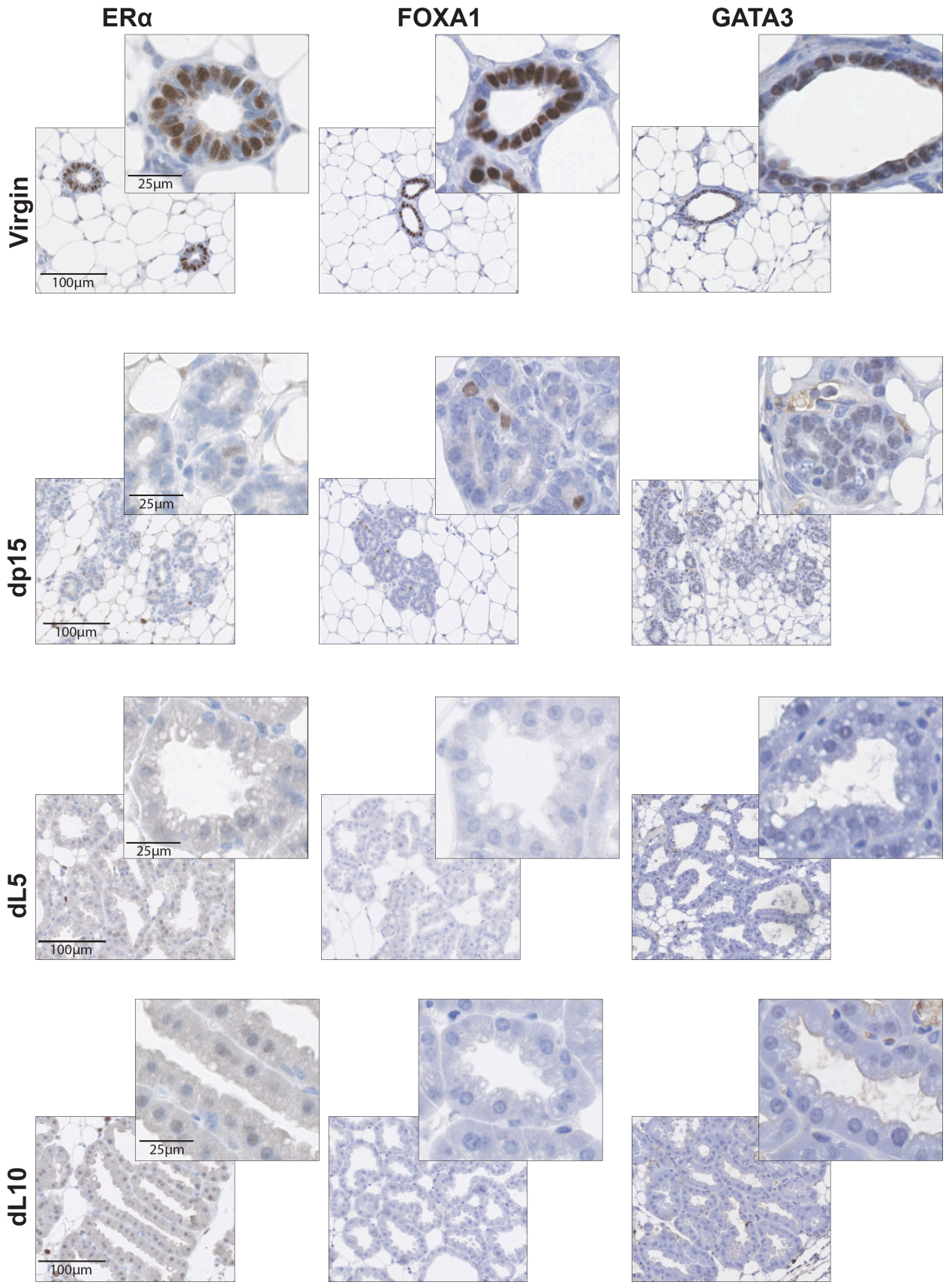

ERα
FOXA1
GATA3
Virgin
dp15
dL5
dL10
25µm
100µm

Figure 6. ERα, FOXA1 and GATA3 expression is reduced during lactation in the mouse mammary gland.
ERα, FOXA1 and GATA3 immunohistochemical staining was performed in normal mouse mammary glands during puberty (virgin), 15 days through pregnancy (dP15), and five and 10 days through lactation (5dL, 10dL). Representative images are shown at 25μm and 100μm magnifications (N=5).

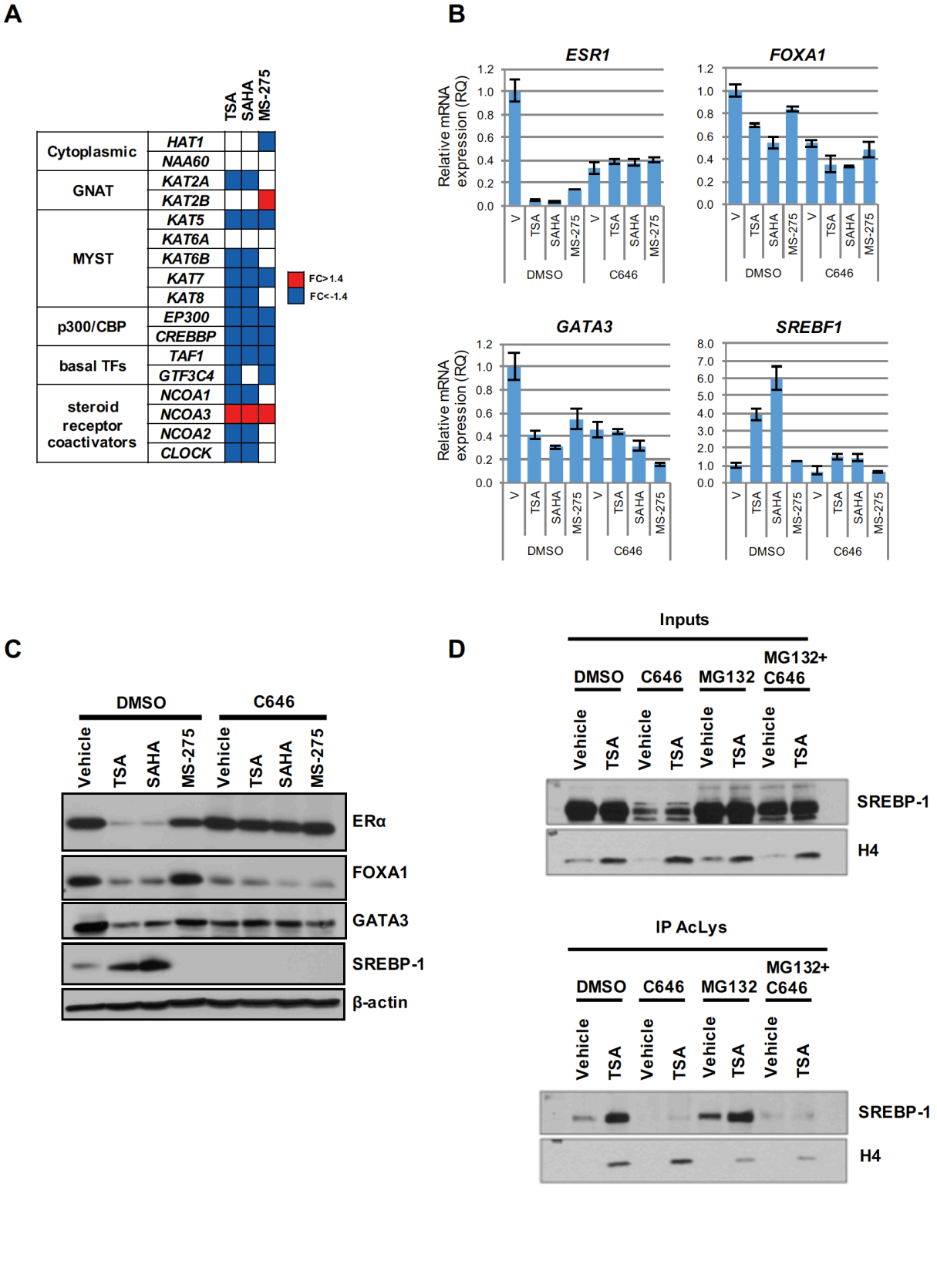

A
TSA
SAHA
MS-275
Cytoplasmic
HAT1
NAA60
GNAT
KAT2A
KAT2B
MYST
KAT5
KAT6A
KAT6B
KAT7
KAT8
p300/CBP
EP300
CREBBP
basal TFs
TAF1
GTF3C4
steroid receptor coactivators
NCOA1
NCOA3
NCOA2
CLOCK
FC>1.4
FC<-1.4
B
ESR1
FOXA1
GATA3
SREBF1
Relative mRNA expression (RQ)
V
TSA
SAHA
MS-275
DMSO
C646
C
DMSO
C646
Vehicle
TSA
SAHA
MS-275
ERα
FOXA1
GATA3
SREBP-1
β-actin
D
Inputs
DMSO
C646
MG132
MG132+ C646
Vehicle
TSA
SREBP-1
H4
IP AcLys

Figure 7. Chemical inhibition of p300/CBP partially alleviates HDACi-mediated actions.
(A) Relative mRNA levels (RNA-seq) of histone acetyltransferases belonging to different HAT families following HDACi treatment in MCF-7 cells (RNA-seq, N=2, q<0.01) .
MCF-7 cells were pretreated with the p300/CBP inhibitor C646 (20μM) or DMSO for 6 hours, then treated with indicated HDACis for 16 hours. (B) Relative expression levels of *ESR1, FOXA1, GATA3* and *SREBF1* were assessed by RT-qPCRs (compared to vehicle control). Error bars represent standard deviation of three technical replicates. One representative experiment is shown (N=2).
(C) Protein expression of ERα, FOXA1, GATA3 and SREBP-1 were examined by immunoblotting (N=3).
(D) MCF-7 cells were pretreated with DMSO, C646 (20μM), MG132 (10μM) or C646+MG132 for 3 hours, then treated with vehicle or TSA (300nM) for 3 hours. Cells lysates were immunoprecipitated using an anti-acetyl-lysine (AcLys) antibody then probed for SREBP-1 to assess its acetylation. Histone H4 serves as a positive control for TSA-induced acetylation.

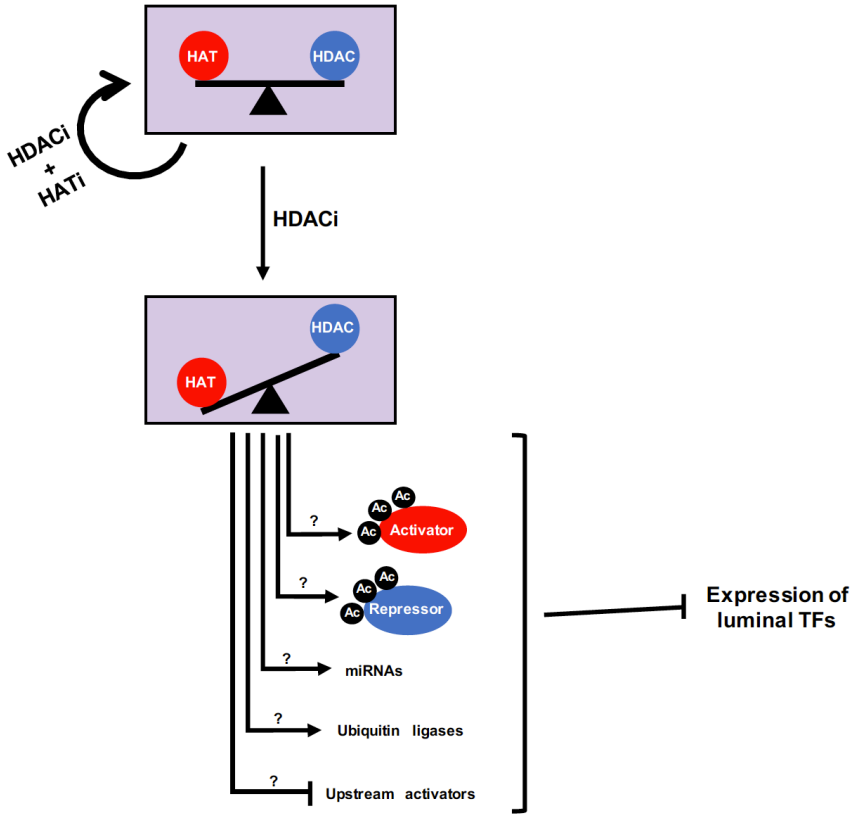
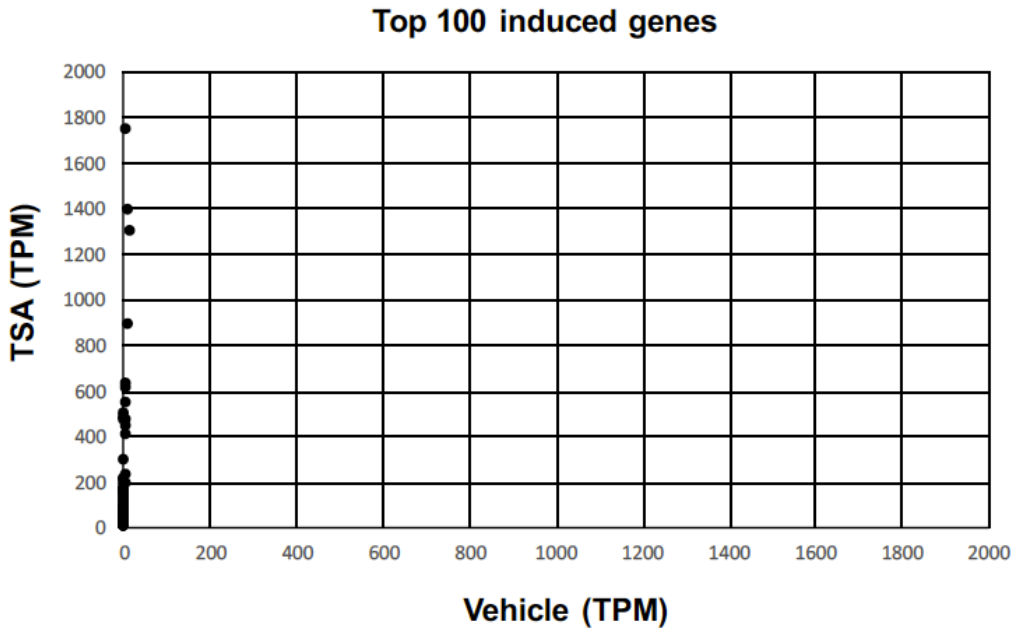
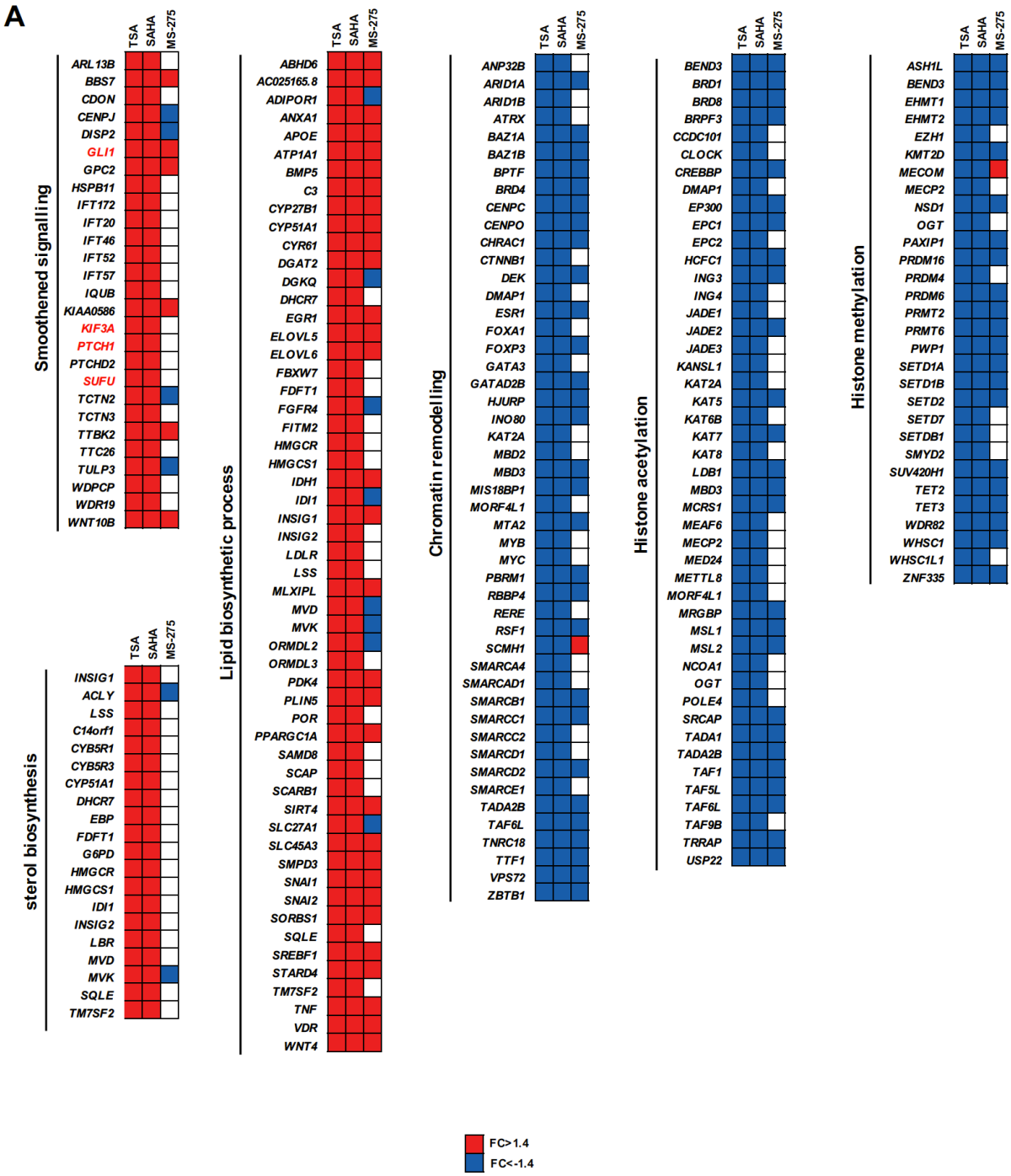
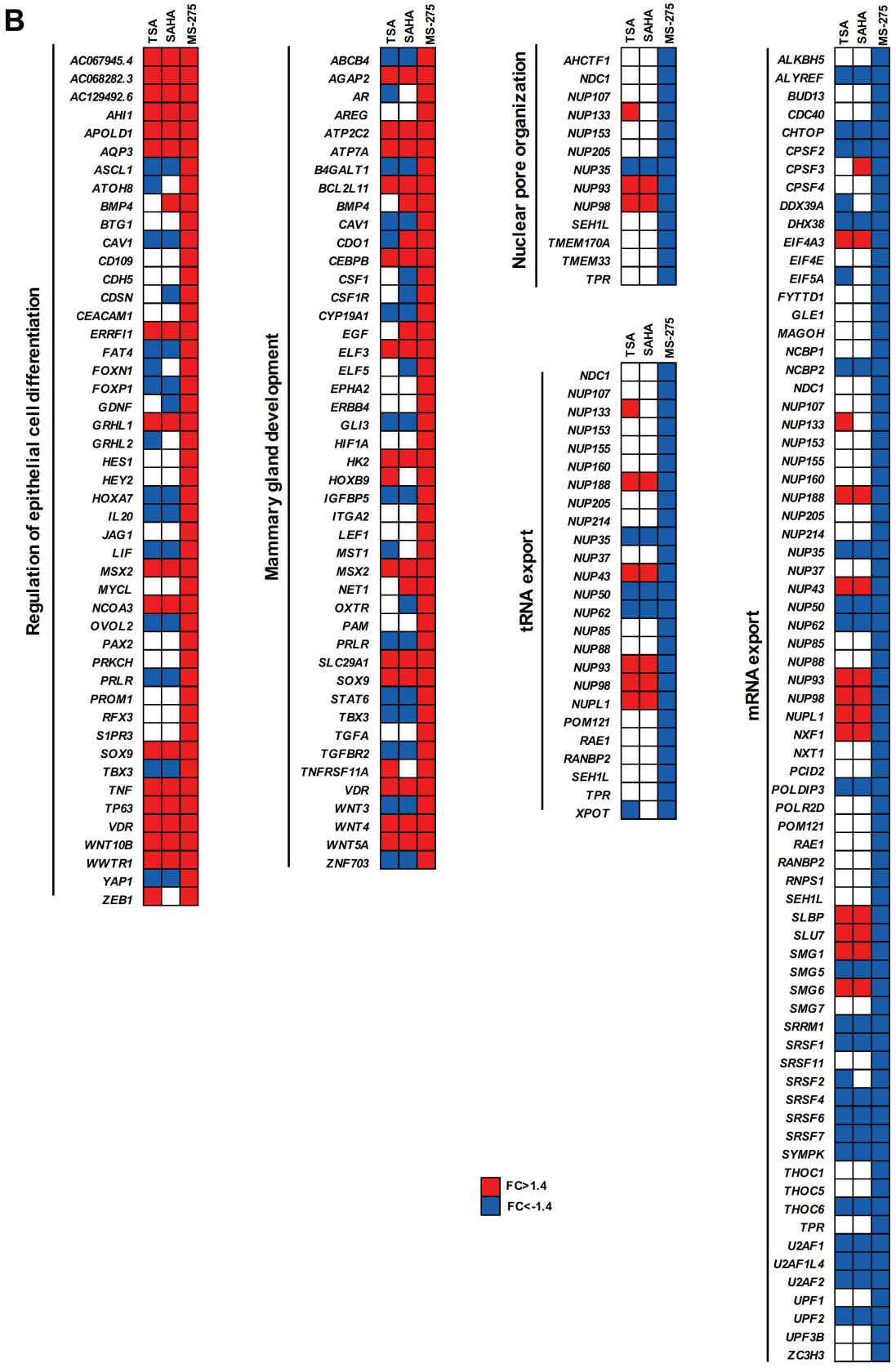


Figure 8. Model depicting potential mechanisms of action of HDACis in regulation of expression of luminal TFs.



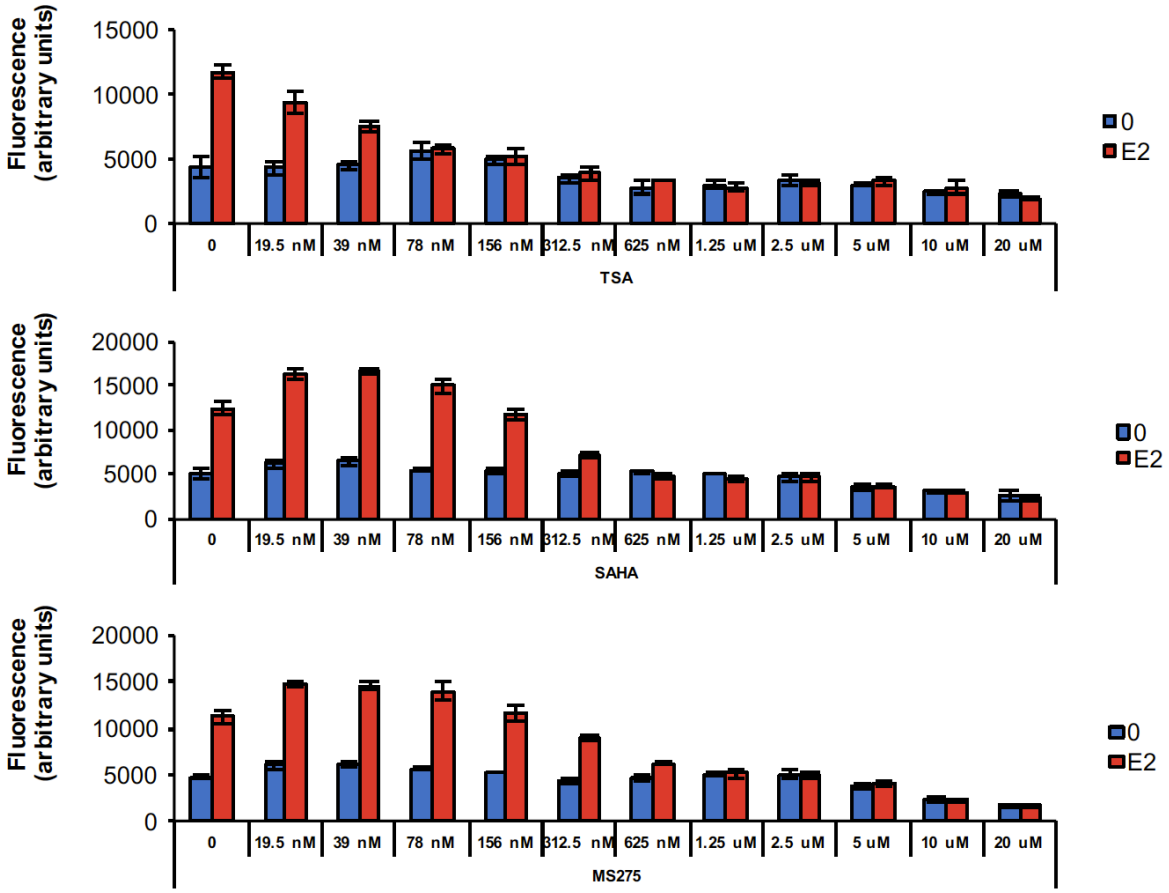
Supplementary Figure 1. TSA/SAHA treatment results in transcriptional derepression.
TPMs of the top 100 TSA/SAHA induced genes in the presence of TSA treatment are plotted against their respective TPMs in the vehicle control.


A
TSA
SAHA
MS-275
Smoothened signalling
ARL13B
BBS7
CDON
CENPJ
DISP2
GLI1
GPC2
HSPB11
IFT172
IFT20
IFT46
IFT52
IFT57
IQUB
KIAA0586
KIF3A
PTCH1
PTCHD2
SUFU
TCTN2
TCTN3
TTBK2
TTC26
TULP3
WDPCP
WDR19
WNT10B
sterol biosynthesis
INSIG1
ACLY
LSS
C14orf1
CYB5R1
CYB5R3
CYP51A1
DHCR7
EBP
FDFT1
G6PD
HMGCR
HMGCS1
IDI1
INSIG2
LBR
MVD
MVK
SQLE
TM7SF2
Lipid biosynthetic process
ABHD6
AC025165.8
ADIPOR1
ANXA1
APOE
ATP1A1
BMP5
C3
CYP27B1
CYP51A1
CYR61
DGAT2
DGKQ
DHCR7
EGR1
ELOVL5
ELOVL6
FBXW7
FDFT1
FGFR4
FITM2
HMGCR
HMGCS1
IDH1
IDI1
INSIG1
INSIG2
LDLR
LSS
MLXIPL
MVD
MVK
ORMDL2
ORMDL3
PDK4
PLIN5
POR
PPARGC1A
SAMD8
SCAP
SCARB1
SIRT4
SLC27A1
SLC45A3
SMPD3
SNAI1
SNAI2
SORBS1
SQLE
SREBF1
STARD4
TM7SF2
TNF
VDR
WNT4
Chromatin remodelling
ANP32B
ARID1A
ARID1B
ATRX
BAZ1A
BAZ1B
BPTF
BRD4
CENPC
CENPO
CHRAC1
CTNNB1
DEK
DMAP1
ESR1
FOXA1
FOXP3
GATA3
GATAD2B
HJURP
INO80
KAT2A
MBD2
MBD3
MIS18BP1
MORF4L1
MTA2
MYB
MYC
PBRM1
RBBP4
RERE
RSF1
SCMH1
SMARCA4
SMARCAD1
SMARCB1
SMARCC1
SMARCC2
SMARCD1
SMARCD2
SMARCE1
TADA2B
TAF6L
TNRC18
TTF1
VPS72
ZBTB1
Histone acetylation
BEND3
BRD1
BRD8
BRPF3
CCDC101
CLOCK
CREBBP
DMAP1
EP300
EPC1
EPC2
HCFC1
ING3
ING4
JADE1
JADE2
JADE3
KANSL1
KAT2A
KAT5
KAT6B
KAT7
KAT8
LDB1
MBD3
MCRS1
MEAF6
MECP2
MED24
METTL8
MORF4L1
MRGBP
MSL1
MSL2
NCOA1
OGT
POLE4
SRCAP
TADA1
TADA2B
TAF1
TAF5L
TAF6L
TAF9B
TRRAP
USP22
Histone methylation
ASH1L
BEND3
EHMT1
EHMT2
EZH1
KMT2D
MECOM
MECP2
NSD1
OGT
PAXIP1
PRDM16
PRDM4
PRDM6
PRMT2
PRMT6
PWP1
SETD1A
SETD1B
SETD2
SETD7
SETDB1
SMYD2
SUV420H1
TET2
TET3
WDR82
WHSC1
WHSC1L1
ZNF335
FC>1.4
FC<-1.4

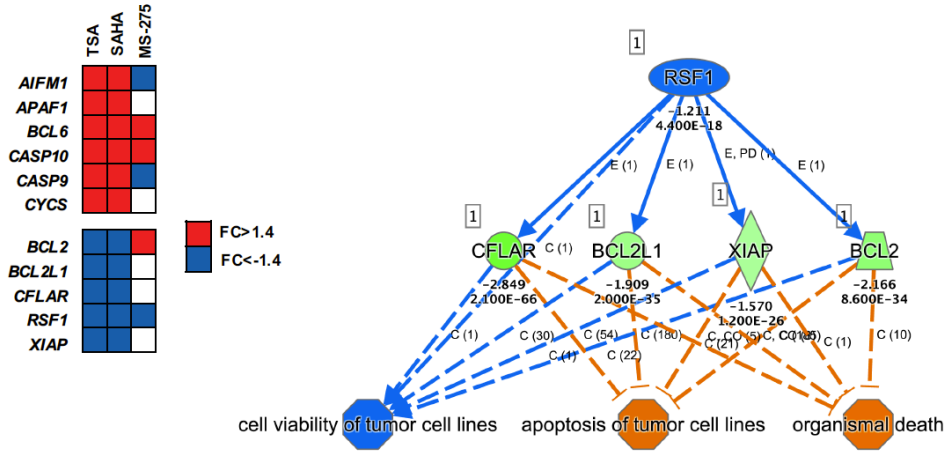

B
TSA
SAHA
MS-275
Regulation of epithelial cell differentiation
AC067945.4
AC068282.3
AC129492.6
AHI1
APOLD1
AQP3
ASCL1
ATOH8
BMP4
BTG1
CAV1
CD109
CDH5
CDSN
CEACAM1
ERRFI1
FAT4
FOXN1
FOXP1
GDNF
GRHL1
GRHL2
HES1
HEY2
HOXA7
IL20
JAG1
LIF
MSX2
MYCL
NCOA3
OVOL2
PAX2
PRKCH
PRLR
PROM1
RFX3
S1PR3
SOX9
TBX3
TNF
TP63
VDR
WNT10B
WWTR1
YAP1
ZEB1
Mammary gland development
ABCB4
AGAP2
AR
AREG
ATP2C2
ATP7A
B4GALT1
BCL2L11
BMP4
CAV1
CDO1
CEBPB
CSF1
CSF1R
CYP19A1
EGF
ELF3
ELF5
EPHA2
ERBB4
GLI3
HIF1A
HK2
HOXB9
IGFBP5
ITGA2
LEF1
MST1
MSX2
NET1
OXTR
PAM
PRLR
SLC29A1
SOX9
STAT6
TBX3
TGFA
TGFBR2
TNFRSF11A
VDR
WNT3
WNT4
WNT5A
ZNF703
Nuclear pore organization
AHCTF1
NDC1
NUP107
NUP133
NUP153
NUP205
NUP35
NUP93
NUP98
SEH1L
TMEM170A
TMEM33
TPR
tRNA export
NDC1
NUP107
NUP133
NUP153
NUP155
NUP160
NUP188
NUP205
NUP214
NUP35
NUP37
NUP43
NUP50
NUP62
NUP85
NUP88
NUP93
NUP98
NUPL1
POM121
RAE1
RANBP2
SEH1L
TPR
XPOT
mRNA export
ALKBH5
ALYREF
BUD13
CDC40
CHTOP
CPSF2
CPSF3
CPSF4
DDX39A
DHX38
EIF4A3
EIF4E
EIF5A
FYTTD1
GLE1
MAGOH
NCBP1
NCBP2
NDC1
NUP107
NUP133
NUP153
NUP155
NUP160
NUP188
NUP205
NUP214
NUP35
NUP37
NUP43
NUP50
NUP62
NUP85
NUP88
NUP93
NUP98
NUPL1
NXF1
NXT1
PCID2
POLDIP3
POLR2D
POM121
RAE1
RANBP2
RNPS1
SEH1L
SLBP
SLU7
SMG1
SMG5
SMG6
SMG7
SRRM1
SRSF1
SRSF11
SRSF2
SRSF4
SRSF6
SRSF7
SYMPK
THOC1
THOC5
THOC6
TPR
U2AF1
U2AF1L4
U2AF2
UPF1
UPF2
UPF3B
ZC3H3
FC>1.4
FC<-1.4

Supplementary Figure 2. Enriched biological processes in HDACi-responsive genes reveal extensive transcriptional reprogramming.
Enriched biological processes in (A) TSA/SAHA and (B) MS-275 significantly up and downregulated genes are shown as heatmaps. Red denotes induction (FC>1.4) and blue denotes repression (FC<-1.4). All regulations shown are statistically significant (RNA-seq, N=2, q<0.01).

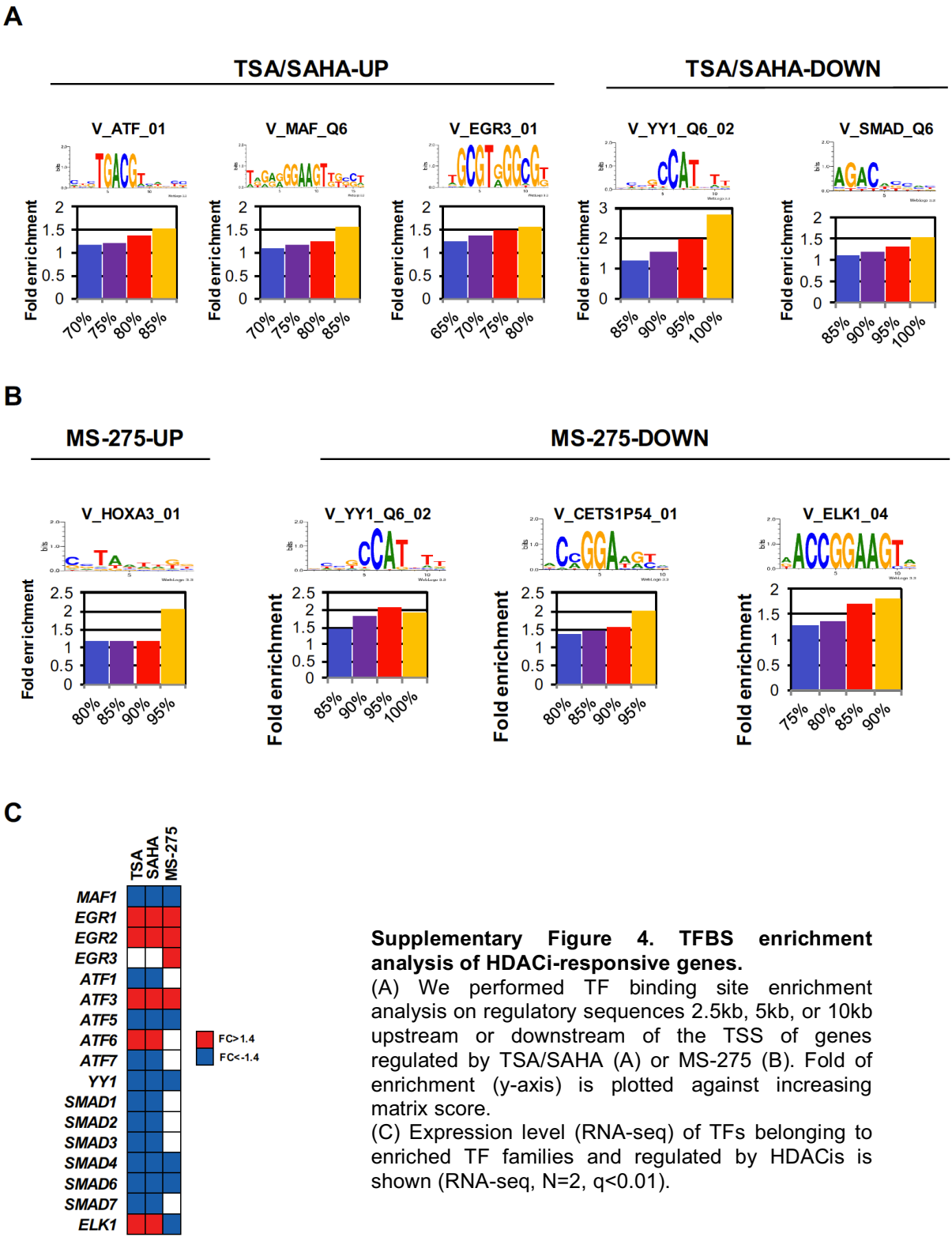
A


Fluorescence (arbitrary units)
15000
10000
5000
0
0
19.5 nM
39 nM
78 nM
156 nM
312.5 nM
625 nM
1.25 uM
2.5 uM
5 uM
10 uM
20 uM
TSA
0
E2
Fluorescence (arbitrary units)
20000
15000
10000
5000
0
SAHA
Fluorescence (arbitrary units)
MS275

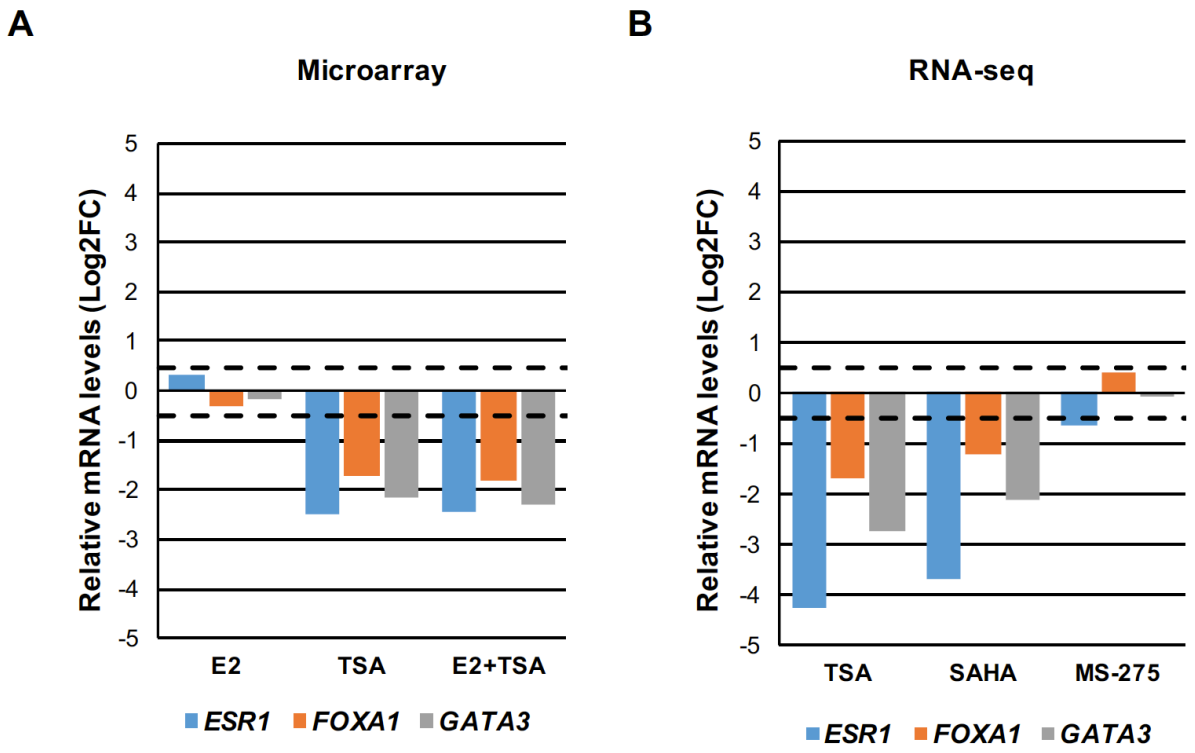
B


TSA
SAHA
MS-275
AIFM1
APAF1
BCL6
CASP10
CASP9
CYCS
BCL2
BCL2L1
CFLAR
RSF1
XIAP
FC>1.4
FC<-1.4
RSF1
CFLAR
BCL2L1
XIAP
BCL2
cell viability of tumor cell lines
apoptosis of tumor cell lines
organismal death

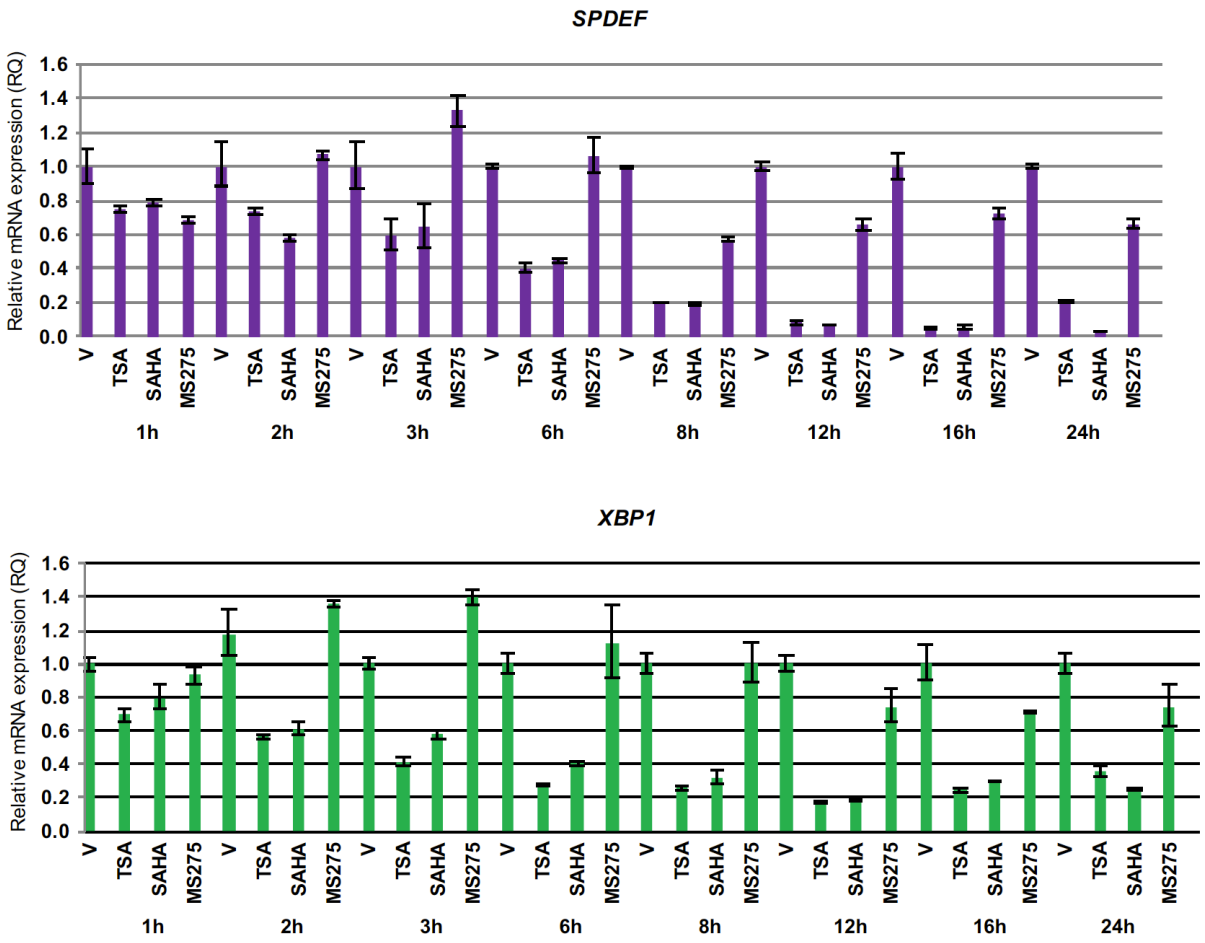
Supplementary Figure 3. HDACis inhibit E2-dependent proliferation of MCF-7 cells.
(A) MCF-7 cell proliferation was assessed by alamarBlue® assay following three consecutive 72-hour treatments with TSA, SAHA, or MS-275 at the indicated concentrations in the presence or absence of E2 (25nM).
(B) TSA and SAHA induce expression of pro-apoptotic genes (*APAF1, CYCS AIFM1, BCL6 CASP9* and *10*) and repress expression of anti-apoptotic genes (*CFLAR, BCL2, BCL2L1, XIAP)* (RNA-seq, N=2, $q<0.01$). Ingenuity pathway analysis (IPA) of TSA/SAHA responsive genes predicts an increase in apoptosis and a decrease in cell viability based on predicted inhibition of RSF1, which positively regulates expression of anti-apoptotic genes.



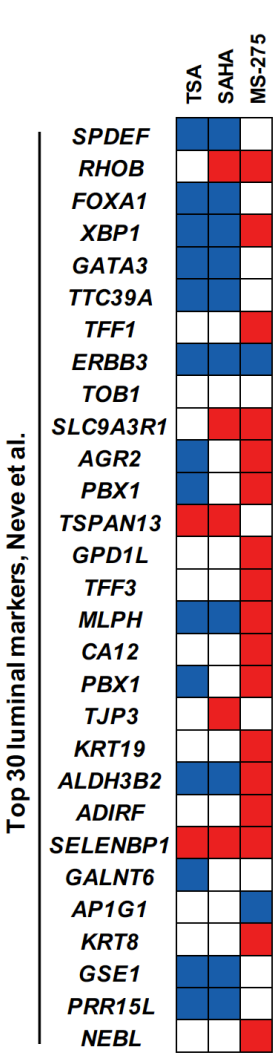
Supplementary Figure 4. TFBS enrichment analysis of HDACi-responsive genes.
(A) We performed TF binding site enrichment analysis on regulatory sequences 2.5kb, 5kb, or 10kb upstream or downstream of the TSS of genes regulated by TSA/SAHA (A) or MS-275 (B). Fold of enrichment (y-axis) is plotted against increasing matrix score.
(C) Expression level (RNA-seq) of TFs belonging to enriched TF families and regulated by HDACis is shown (RNA-seq, N=2, q<0.01).



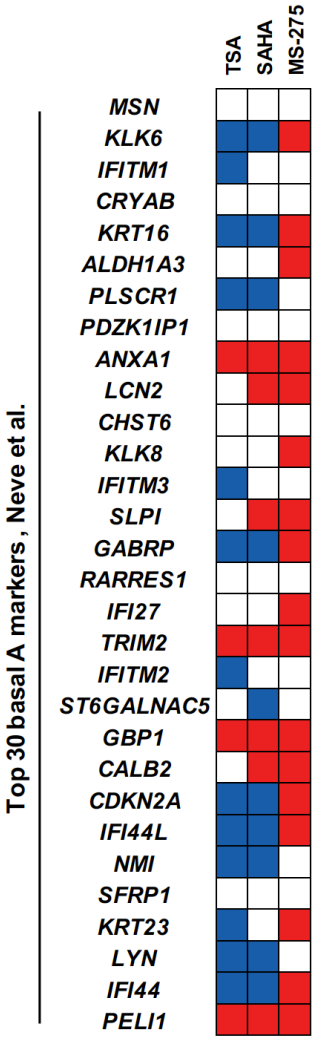
Supplementary Figure 5. HDACis repress expression of *ESR1*, *FOXA1*, and *GATA3* in the presence or absence of E2.
(A). MCF-7 cells cultured in hormone-depleted media were treated with E2 (25nM), TSA (300nM), or E2+TSA for 24 h. Relative mRNA levels (compared to vehicle control; Log2FC) were assessed by microarray profiling.
(B). MCF-7 cells cultured in hormone-depleted media were treated with TSA (300nM), SAHA (5μM) or MS-275 (5μM) for 8 h. Relative mRNA levels (compared to vehicle control; Log2FC) were assessed by RNA-seq.
Dotted lines mark the 1.4 fold change threshold (Log2FC=0.485).

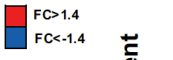


Supplementary Figure 6. HDACis repress expression of luminal-lineage TFs *SPDEF* and *XBP1*. MCF-7 cells were treated with vehicle (V), TSA (300nM), SAHA (5μM), or MS-275 (5μM) for 1, 2, 3, 6, 8, 12, 16, and 24 h. Relative mRNA levels of *SPDEF* and *XBP1* were examined by RT-qPCRs. Error bars represent standard deviation of three technical replicates (N=1).

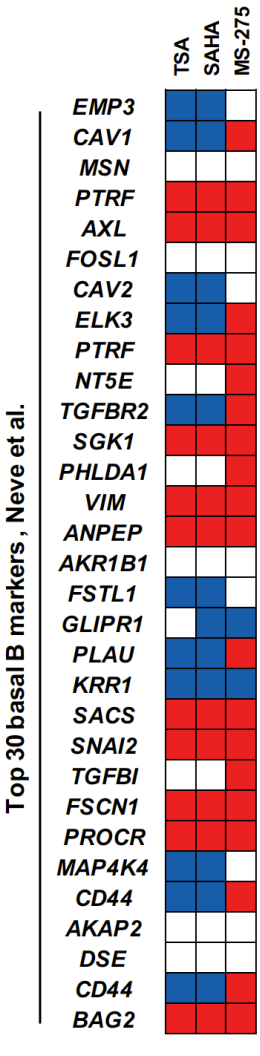


		Fold enrichment	
luminal	TSA-UP	0.36	
	TSA-DOWN	2.50	*
	SAHA-UP	1.00	
	SAHA-DOWN	1.95	*
	MS-275 UP	3.03	*
	MS-275 DOWN	0.36	





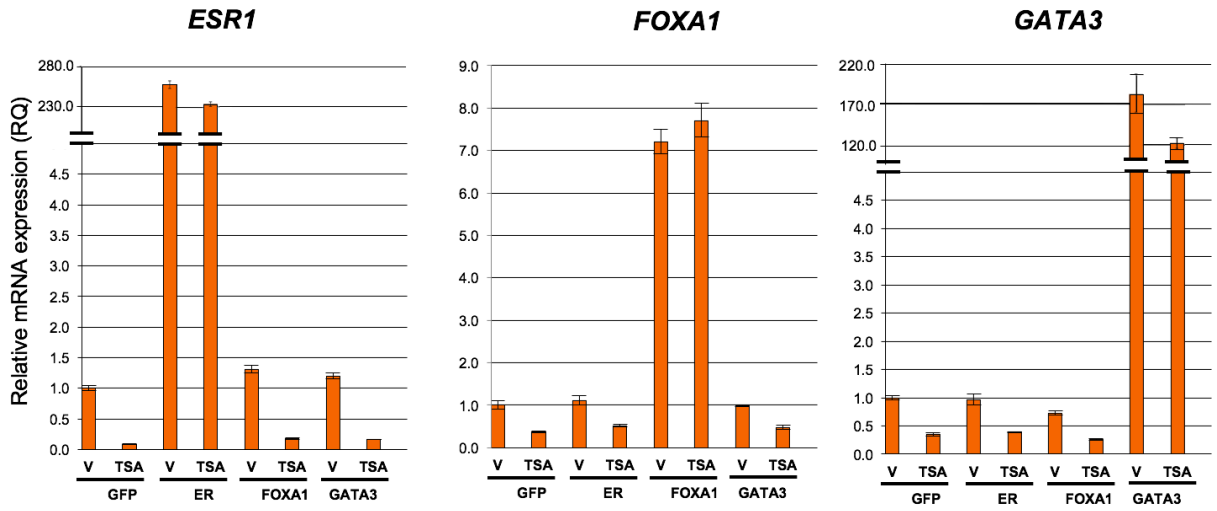
		Fold enrichment	
Basal A	TSA-UP	0.72	
	TSA-DOWN	2.32	*
	SAHA-UP	1.39	
	SAHA-DOWN	1.95	*
	MS-275 UP	3.03	*
	MS-275 DOWN	0.00	



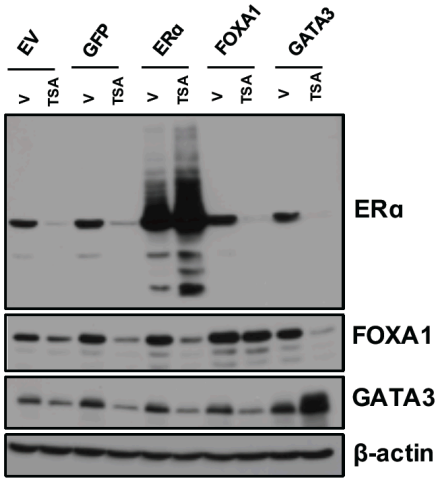
		Fold enrichment	
Basal B	TSA-UP	1.99	*
	TSA-DOWN	1.96	*
	SAHA-UP	2.19	*
	SAHA-DOWN	2.34	*
	MS-275 UP	3.57	*
	MS-275 DOWN	0.36	

Supplementary Figure 7. Impact of HDACis on expression of basal and luminal markers.
The top 30 luminal, basal A and basal B markers were obtained from the Neve et al. breast cancer cell line classifier. The impact of HDACi treatment on expression of these genes was determined using the RNA-Seq HDACi datasets in MCF-7 cells. A fold enrichment is calculated as a ratio of the number of observed vs expected genes regulated by HDACis. Fold enrichment >1 indicates over-representation. Asterisks denote statistical significance, $p<0.05$, Fishers exact test.

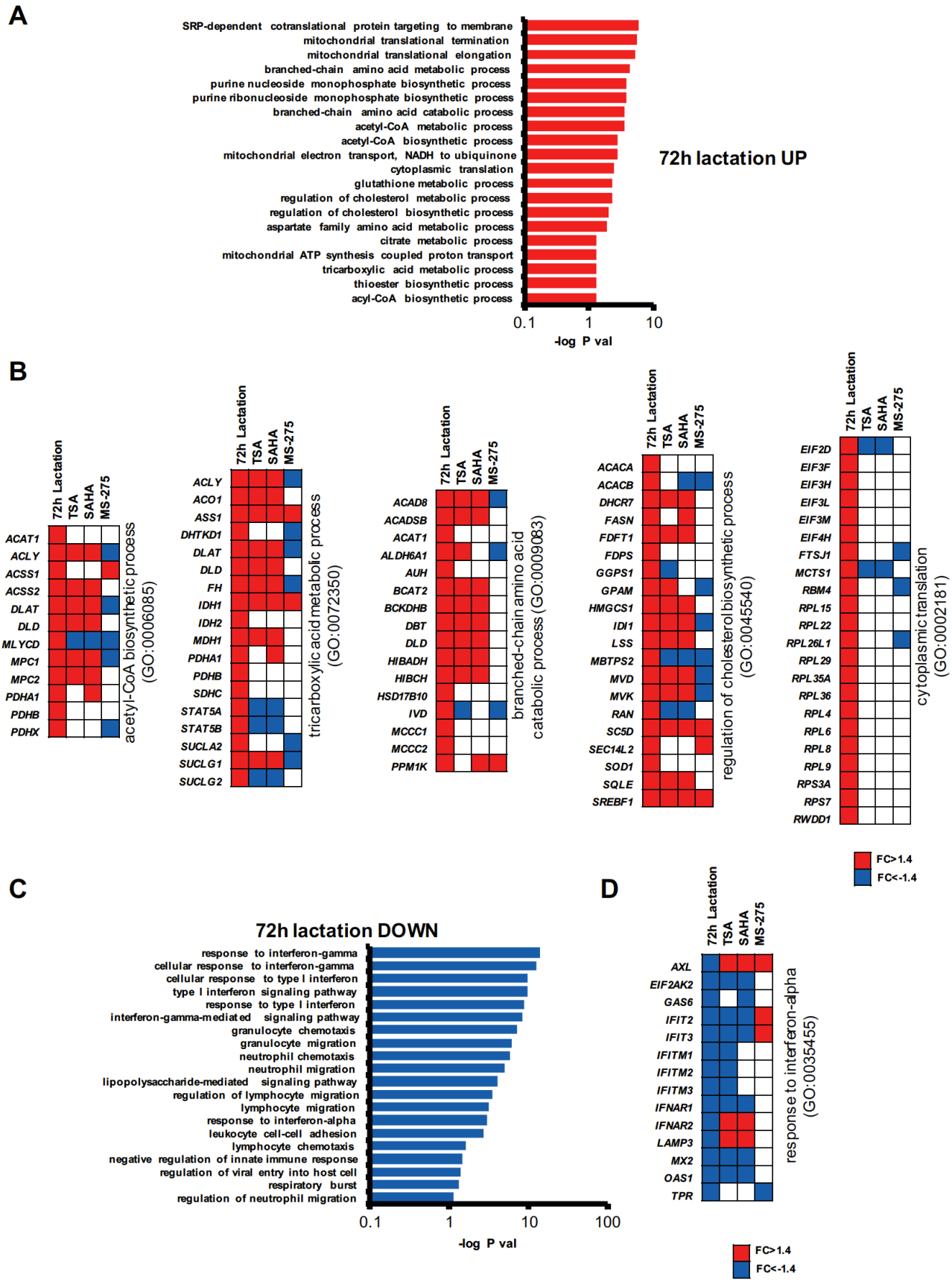
A



B



Supplementary Figure 8. Ectopic overexpression of FOXA1 does not prevent HDACi-mediated repression of ERα and GATA3.
MCF-7 cells were transiently transfected by electroporation with ERα, FOXA1 and GATA3 expression plasmids (48 h). Transfection of an empty vector or GFP expression plasmid served as a negative control. Electroporated cells were treated with vehicle or TSA (300nM) for the last 24 h of transfection. (A) Relative mRNA levels of *ESR1, FOXA1* and *GATA3* were assessed by RT-qPCRs and compared to the GFP-vehicle condition. Error bars represent standard deviation of three technical replicates (N=1). (B) Protein lysates from the same experiments were probed for ERα, FOXA1 and GATA3 expression by western analysis. β-actin serves as a loading control.


A
SRP-dependent cotranslational protein targeting to membrane
mitochondrial translational termination
mitochondrial translational elongation
branched-chain amino acid metabolic process
purine nucleoside monophosphate biosynthetic process
purine ribonucleoside monophosphate biosynthetic process
branched-chain amino acid catabolic process
acetyl-CoA metabolic process
acetyl-CoA biosynthetic process
mitochondrial electron transport, NADH to ubiquinone
cytoplasmic translation
glutathione metabolic process
regulation of cholesterol metabolic process
regulation of cholesterol biosynthetic process
aspartate family amino acid metabolic process
citrate metabolic process
mitochondrial ATP synthesis coupled proton transport
tricarboxylic acid metabolic process
thioester biosynthetic process
acyl-CoA biosynthetic process
0.1
1
10
-log P val
72h lactation UP
B
72h Lactation
TSA
SAHA
MS-275
ACAT1
ACLY
ACSS1
ACSS2
DLAT
DLD
MLYCD
MPC1
MPC2
PDHA1
PDHB
PDHX
acetyl-CoA biosynthetic process (GO:0006085)
ACLY
ACO1
ASS1
DHTKD1
DLAT
DLD
FH
IDH1
IDH2
MDH1
PDHA1
PDHB
SDHC
STAT5A
STAT5B
SUCLA2
SUCLG1
SUCLG2
tricarboxylic acid metabolic process (GO:0072350)
ACAD8
ACADSB
ACAT1
ALDH6A1
AUH
BCAT2
BCKDHB
DBT
DLD
HIBADH
HIBCH
HSD17B10
IVD
MCCC1
MCCC2
PPM1K
branched-chain amino acid catabolic process (GO:0009083)
ACACA
ACACB
DHCR7
FASN
FDFT1
FDPS
GGPS1
GPAM
HMGCS1
IDI1
LSS
MBTPS2
MVD
MVK
RAN
SC5D
SEC14L2
SOD1
SQLE
SREBF1
regulation of cholesterol biosynthetic process (GO:0045540)
EIF2D
EIF3F
EIF3H
EIF3L
EIF3M
EIF4H
FTSJ1
MCTS1
RBM4
RPL15
RPL22
RPL26L1
RPL29
RPL35A
RPL36
RPL4
RPL6
RPL8
RPL9
RPS3A
RPS7
RWDD1
cytoplasmic translation (GO:0002181)
FC>1.4
FC<-1.4
C
72h lactation DOWN
response to interferon-gamma
cellular response to interferon-gamma
cellular response to type I interferon
type I interferon signaling pathway
response to type I interferon
interferon-gamma-mediated signaling pathway
granulocyte chemotaxis
granulocyte migration
neutrophil chemotaxis
neutrophil migration
lipopolysaccharide-mediated signaling pathway
regulation of lymphocyte migration
lymphocyte migration
response to interferon-alpha
leukocyte cell-cell adhesion
lymphocyte chemotaxis
negative regulation of innate immune response
regulation of viral entry into host cell
respiratory burst
regulation of neutrophil migration
0.1
1
10
100
-log P val
D
72h Lactation
TSA
SAHA
MS-275
AXL
EIF2AK2
GAS6
IFIT2
IFIT3
IFITM1
IFITM2
IFITM3
IFNAR1
IFNAR2
LAMP3
MX2
OAS1
TPR
response to interferon-alpha (GO:0035455)
FC>1.4
FC<-1.4

Supplementary Figure 9. Impact of HDACis on pathways enriched in genes significantly regulated during lactation in the human mammary gland.
(A) GO term enrichment analysis of significantly upregulated (FC>1.4 and p<0.05) genes at 72 hours postpartum (compared to 6 hours postpartum) was performed. The top 20 enriched biological processes (fold enrichment >2) are shown on the y axis and plotted against their respective –log p values (x-axis).
(B) Pathways enriched in significantly upregulated genes at 72 hours postpartum are shown. The impact of HDACis on expression of genes in these pathways was examined using RNA-seq data of HDACi-treated MCF-7 cells (RNA-seq, N=2, q<0.01).
(C) GO term enrichment analysis was performed on genes significantly downregulated (FC<-1.4 and p<0.05) at 72 hours postpartum. Top 20 enriched biological processes are shown.
(D) The impact of HDACis on genes in the 'response to interferon alpha' process is shown.

Supplementary Table 1. RT-qPCR oligo sequences and their corresponding UPL probe number are shown.

Gene	UPL probe number	Sequence
RPLP0 (reference gene)	74	5'-TCCCACTTGCTGAAAAGGTC-3'
		5'-AGCAGGAGCAGCTGTGGT-3'
YWHAZ (reference gene)	2	5'-GCAATTACTGAGAGACAACTTGACA-3'
		5'-TGGAAGGCCGGTTAATTTT-3'
ESR1	24	5'-TTACTGACCAACCTGGCAGA-3'
		5'-ATCATGGAGGGTCAAATCCA-3'
FOXA1	47	5'-ATCATTGCCATCGTGTGCTT-3'
		5'-CACCATGTCCAACTGTGGAA-3'
GATA3	36	5'-ACTACGGAAACTCGGTCAGG-3'
		5'-GGTAGGGATCCATGAAGCAG-3'
SPDEF	26	5'-CTCAGCTTGTCGTAGTTCATGG-3'
		5'-AGGTGGCTCAACAAGGAGAA-3'
XBP1	62	5'-CCCTGGTTGCTGAAGAGG-3'
		5'-TGGAGGGGTGACAACTGG-3'
YY1	10	5’-GAGGTGATCCTGGTGCAGA-3’
		5’-GATCCTCGAAGCCGTCCT-3’
FDFT1	27	5’-ATCTTCACTGCCCCTTTGAA-3’
		5’-TGTGTTTAACTTCTGTGCTATTCCA-3’
HMGCS1	74	5’-GTTCACCCCAGATACCATCAG-3’
		5’-GCTAGGAATTCCCTTGGAAGA-3’
SQLE	61	5’-CACGACAAAGGAATGAGATGTT-3’
		5’-TGATGATGCAGCTATTTTCGAG-3’
SREBF1	77	5’-CGCTCCTCCATCAATGACA-3’
		5’-TGCGCAAGACAGCAGATTTA-3’
DDX58	13	5’-AAGCACTTGCTACCTCTTGCTC-3’
		5’-ATGTGGGCAATGTCATCAAA-3’
IFIT1	9	5’-GCTCCAGACTATCCTTGACCTG-3’
		5’-AGAACGGCTGCCTAATTTACAG-3’
IFIT2	27	5’-GTAGGCTGCTCTCCAAGGAA-3’
		5’-TGGTGGCAGAAGAGGAAGAT-3’
IFIT3	80	5’-GCATTTCAGCTGTGGAAGG-3’
		5’-CAGAACTGCAGGGAAACAGC-3’
OAS1	27	5’-TCCAGTCCTCTTCTGCCTGT-3’
		5’-GTGAGCTCCTGGATTCTGCT-3’
OAS2	36	5’-ATGAGCCCTGCATAAACCTC-3’
		5’-CCTGCCTTTAATGCACTGG-3’
IFIH1	36	5’-ATTTGGTAAGGCCTGAGCTG-3’

		5’-AGGCACCATGGGAAGTGAT-3’
RSAD2	39	5’-AGGTATTCTCCCCGGTCTTG-3’
		5’-TGCTTTTGCTTAAGGAAGCTG-3’

Supplementary Table 2. TSA/SAHA responsive genes were subjected to IPA analysis. Upstream transcriptional regulators predicted to be activated or inhibited by TSA/SAHA treatment based on the direction of regulation of their target genes as determined by IPA. Log2FC and p values (<0.05) are shown for the TSA condition.

Upstream Transcriptional Regulator	Expr Log Ratio	Predicted Activation State	Activation z-score	p-value of overlap
SREBF2		Activated	4.378	0.00451
SREBF1	2.446	Activated	4.347	0.00696
GMNN	1.544	Activated	2.646	0.00535
SOX1		Activated	2.502	0.0147
CREB1		Activated	2.46	0.0038
HIF1A		Activated	2.425	0.0000188
SOX3		Activated	2.414	0.00275
XBP1	-1.914	Activated	2.329	0.014
TCF7L2	-0.583	Activated	2.29	0.00203
ATF4		Activated	2.279	0.000444
FOXO1	-0.897	Activated	2.19	0.0105
NFYA		Activated	2.132	0.0225
EGR1	0.888	Activated	2.019	0.0000557
RSF1	-1.211	Inhibited	-2	0.0155
STAG1	-1.775	Inhibited	-2	0.027
SOX4		Inhibited	-2.138	0.0332
HDAC5		Inhibited	-2.206	0.0124
POU5F1	2.582	Inhibited	-2.95	0.000154

Supplementary Table 3. Upstream transcriptional regulators predicted to be activated or inhibited by MS-275 treatment based on the direction of regulation of their target genes as determined by IPA.

Upstream Transcriptional Regulator	Expr Log Ratio	Predicted Activation State	Activation z-score	p-value of overlap
SMARCA4		Activated	7.436	0.00781
CTNNB1		Activated	5.471	0.0098
EPAS1	1.326	Activated	4.609	0.0000224
STAT4	3.032	Activated	4.225	0.00428
HIF1A	0.553	Activated	4.21	0.000000317
FOXL2		Activated	3.993	0.00953
STAT3	-0.696	Activated	3.837	0.0195
SP1	-0.503	Activated	3.777	0.000264
CDKN2A		Activated	3.698	0.0254
SMAD4	-0.588	Activated	3.652	0.00818
PAX7		Activated	3.615	0.0411
NFKBIA	1.368	Activated	3.587	0.00063
CREBBP	-0.779	Activated	3.389	0.00109
EP300	-1.167	Activated	3.369	0.000233
HOXA9		Activated	3.328	0.00864
CEBPA		Activated	3.285	0.0202
GLI1	3.911	Activated	3.105	0.0153
ECSIT		Activated	2.915	0.00962
EHF	5.798	Activated	2.844	0.0291
TP53	-1.893	Activated	2.564	1.33E-12
NUPR1		Activated	2.44	4.02E-08
EGR1	1.345	Activated	2.437	0.00295
ETS1		Activated	2.417	0.0229
ZBTB17	-0.625	Activated	2.4	0.00322
FOXO3	0.838	Activated	2.397	0.00172
ELF3	1.844	Activated	2.382	0.00766
TWIST1		Activated	2.273	0.0215
KLF6	0.523	Activated	2.098	0.028
ERG		Activated	2.088	0.000796
TCF7L2	0.687	Activated	2.014	0.01
CLOCK		Inhibited	-2	0.000517
SOX2		Inhibited	-2.063	0.0297
HLX		Inhibited	-2.079	0.0305
NFE2L2		Inhibited	-2.1	0.0208
DACH1		Inhibited	-2.284	0.0261
E2F1		Inhibited	-2.294	0.000959
MYT1		Inhibited	-2.333	0.0309
PRDM8	7.712	Inhibited	-2.449	0.0308

MEOX2		Inhibited	-2.491	0.0341
GLIS2		Inhibited	-2.525	0.0123
HSF2		Inhibited	-2.54	0.0115
MYC		Inhibited	-2.559	0.0122
ZNF217	0.891	Inhibited	-3.9	0.0128
XBP1	0.678	Inhibited	-4.771	0.00214
CBX5	-0.804	Inhibited	-6.856	0.000846

Chapter Four.
Discussion and Perspectives

The advent of next-generation sequencing has greatly increased our understanding of the heterogeneity of breast cancer, a disease that remains the second leading cause of death in women in Canada. This has led to the identification of gene signatures distinguishing several molecular subtypes of breast cancer (139, 140). Luminal tumours represent about 70% of breast cancers and they are characterized by ERα expression or overexpression, which drives their proliferation. Identifying factors that act upstream of ERα can help determine whether their targeting may be beneficial in patients who suffer from resistance to antiestrogen therapy, especially as most distant metastases retain ERα expression (482).

The original finding by Sorlie et al. and Perou et al. that expression of ERα is strongly correlated with that of FOXA1 and GATA3 in luminal tumours led to a whole body of literature on the biology of these TFs in breast tissue and how they can impact ERα function and expression. Do FOXA1 and GATA3 regulate ERα expression, which could explain why their expression is positively correlated in tumour samples? Two knockout studies in the normal mouse mammary gland revealed that both FOXA1 and GATA3 are positive regulators of ERα expression (95, 108). Studies in human breast cancer cell lines MCF-7 and T-47D, however, have reported conflicting results. Bernardo et al. found that siRNA-mediated knockdown of FOXA1 in MCF-7 cells reduced ERα expression (95), whereas Hurtado et al. found that it has no effect (167). GATA3 has been shown to positively regulate expression of ERα in MCF-7 and T-47D cells (214), although a subsequent study found that it has no impact (244). While these contrasting results may be explained by differences in cell culture or transfection conditions, it remains important to document whether ERα is a FOXA1 and GATA3 target because this may dictate whether patients could benefit from therapy targeting FOXA1 and/or GATA3. Interestingly, siRNA-mediated targeting of both FOXA1 and GATA3 has revealed that they can impact ERα binding across the genome. 50% of ERα binding events overlap with a FOXA1 binding event and depletion of FOXA1 was reported to reduce the signal intensity at 90% of ERα binding sites (167). In addition, GATA3 depletion results in redistribution of one third of ERα binding events (244) and 30% of ERα-bound sites are also bound by FOXA1 and GATA3 (245). Thus, given their capacity to bind heterochromatin, the current paradigm is that both FOXA1 and GATA3 act as pioneer factors that can open chromatin and dictate ERα binding across the genome either directly or indirectly. However, it has not been possible to address the hypothesis that ‘pioneer’ FOXA1 and GATA3 binding to DNA precedes that of ERα because all three TFs are co-expressed in luminal breast cancer cells. In the study presented in Chapter 2, we examined large transcriptome datasets from

patient tumour samples and identified transcriptional networks governed by ERα, FOXA1 and GATA3 that can differentiate between different tumour subtypes. Additionally, we engineered an SK-BR-3 molecular apocrine breast cancer cell line (FOXA1-positive, ERα-negative, GATA3-negative) that stably expresses ERα and found that the receptor can actually open chromatin on sites weakly pre-bound by FOXA1. We further documented a positive regulatory role of FOXA1 on *ESR1* expression in several luminal breast cancer cell lines. The pioneer factor/upstream regulatory role of GATA3 and FOXA1 will be further discussed in **sections 4.1 and 4.2, respectively.**

In the study outlined in Chapter 3, we investigated the mechanisms of action of HDACis in breast cancer cells in light of recent evidence that these inhibitors can synergize with antiestrogens for breast cancer treatment. We found that HDAC proteins act as upstream regulators of not just ERα, but also FOXA1 and GATA3, as well as genes involved in lactogenic differentiation processes. The impact of HDACis on the luminal gene cluster (described in Chapter 2), as well as E2 signalling will be discussed in **sections 4.3 and 4.4**, respectively. Lastly, the question of whether FOXA1 can be a therapeutic target in breast cancer will be addressed in **section 4.5**.

4.1 On the role of GATA3 in ERα expression and signalling

Previous studies have relied on siRNA-mediated knockdown of GATA3 to address its transcriptional roles in ERα-positive cancer cell lines. However, because siRNA targeting does not result in complete protein depletion, residual protein can impact interpretation of results. A better approach to examine a potential 'pioneer factor' role of a TF is to express it in a relevant cell line where it is not normally expressed. By doing so, chromatin opening can be examined before and after expression of this TF. Takaku et al. have utilized this approach to examine the function of GATA3 in chromatin opening using Assay for Transposase-Accessible Chromatin using sequencing (ATAC-Seq) (483) in ERα-negative, GATA3-negative, and FOXA1-negative MDA-MB-231 breast cancer cells (104). Interestingly, 25% of GATA3 ChIP peaks became sensitive to transposase treatment following GATA3 expression, i.e. chromatin was more open at these sites. These peaks were associated with increased H3K4me1 and H3K27Ac marks, and transcriptional activation of epithelial identity genes *KRT8*, *KRT18*, and *CDH1*, suggesting that GATA3 induces MET features in MDA-MB-231 breast cancer cells. Of note, 3% of GATA3 peaks in this study were less open following GATA3 expression and harboured reduced H3K4me1 and H3K27Ac marks, suggesting that

GATA3 can also repress transcription, although to a minor degree. GATA3 was able to remodel chromatin at sites that became more accessible and this was associated with recruitment of the chromatin remodelling protein BRG1 that interacted with GATA3 through its TA1 domain. Although this study suggests that GATA3 acts as a pioneer factor, given its capacity to bind its motif in reconstituted nucleosomal DNA *in vitro* (104), several important questions need to be answered to support this conclusion. What is the impact of BRG1 depletion on DNA accessibility at GATA3-bound sites? Is it possible that other upstream factors, including BRG1, could open chromatin at GATA3-bound sites to facilitate its subsequent binding? Mere opening of chromatin following expression of a given factor may not necessarily be a consequence of direct binding of this factor to DNA. Thus, what would be the impact of expression of a GATA3 mutant that lacks any DNA-binding capacity on chromatin accessibility? Another potential mechanism by which GATA3 can open chromatin is regulation of expression or modulation of activity via protein-protein interactions of other factors that might play a role in chromatin opening at sites where it is bound.

The pioneer factor roles of FOXA1 and ERα can also be addressed using the MDA-MB-231 cell line model, representative of the claudin-low subtype of triple-negative tumours. It will thus be interesting to express each of these three factors (WT or non-DNA-binding mutants) separately or in various combinations and perform RNA-Seq to study downstream targets, ChIP-Seq against against the three TFs and various activating or repressive histone marks to study recruitment to DNA and ATAC-Seq to study chromatin accessibility in the presence and absence of E2 stimulation. Such an experiment would help determine the contribution of each TF to other TF recruitment and activation of the E2 response. Two approaches can be taken to examine whether additional factors are involved in chromatin opening of sites after TF expression: (i) bioinformatic mining of TF-bound sites for enrichment in motifs of other factors and (ii) identification of interacting partners of each of these TFs on DNA using rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) (484), a technique that has allowed the identification of GREB1 as an ERα cofactor (485). Additionally, enrichment of ERα, GATA3, and FOXA1 binding motifs will be preferentially examined on sites where they are co-bound. Is ERα mostly recruited to these sites via tethering to FOXA1 or GATA3? In that case, is there enrichment in FOXREs and/or GATA motifs at these sites or are these factors recruited via chromatin loops or tethering to other factors? It is also possible that ERα is recruited to these sites by directly binding to EREs, which can be further stabilized by binding of FOXA1 and GATA3. One potential drawback of

this system is that MDA-MB-231 cells have undergone EMT and may not express all E2-response components present in luminal cells. This would not allow for full recapitulation of mechanisms by which binding of these factors is regulated. Indeed, reanalysis of transcriptome profiles of MDA-MB-231 cells co-expressing ERα, FOXA1 and GATA3 (245) revealed that this was not sufficient to reconstitute the E2 response observed in luminal cells. The BT-20 triple-negative cell line, which maintains its epithelial features, and is representative of the basal-like subtype, could be used instead.

We have employed a similar approach in SK-BR-3 cells to examine the pioneer factor role of FOXA1. Indeed, we observed activation of an E2-response following transient ectopic expression of ERα in these cells, albeit with a smaller number of regulated genes than in MCF-7 cells. It remains to be determined whether a broader response would be observed in SK-BR-3 cells stably expressing ERα, which could be the case if cascades of chromatin remodelling events are initiated via ERα but also its downstream TFs, leading to opening of new EREs. However, one of the key differences between MCF-7 and SK-BR-3 cells is that the latter also lack GATA3 expression. Thus, we decided to examine the impact of transient ectopic co-expression of GATA3 with ERα on the E2 response in SK-BR-3 cells through transcriptome profiling. Notably, despite its positive regulatory role in *ESR1* expression in the ERα-positive breast cancer cell lines T-47D and ZR-75-1 (**Figure 2, Chapter 2)**, transient expression of WT *GATA3* was not sufficient to restore ERα protein levels in SK-BR-3 cells (**Figure 1A).** The *ESR1* gene promoter is epigenetically silenced via DNA hypermethylation in ERα-negative breast cancer cell lines (357, 358). This result indicates that expression of two upstream regulators of *ESR1* is insufficient in SK-BR-3 cells to restore its expression, implying the existence of as yet unidentified additional factors or cofactors. In this respect, it would be of interest to treat SK-BR-3 cells expressing GATA3 (or not) with epigenetic modulators such as HDACis or DNMT inhibitors to determine whether this may reverse the epigenetic inactivation of *ESR1* regulatory sequences, and enable binding of FOXA1 and GATA3 to these sequences as observed in MCF-7 cells.

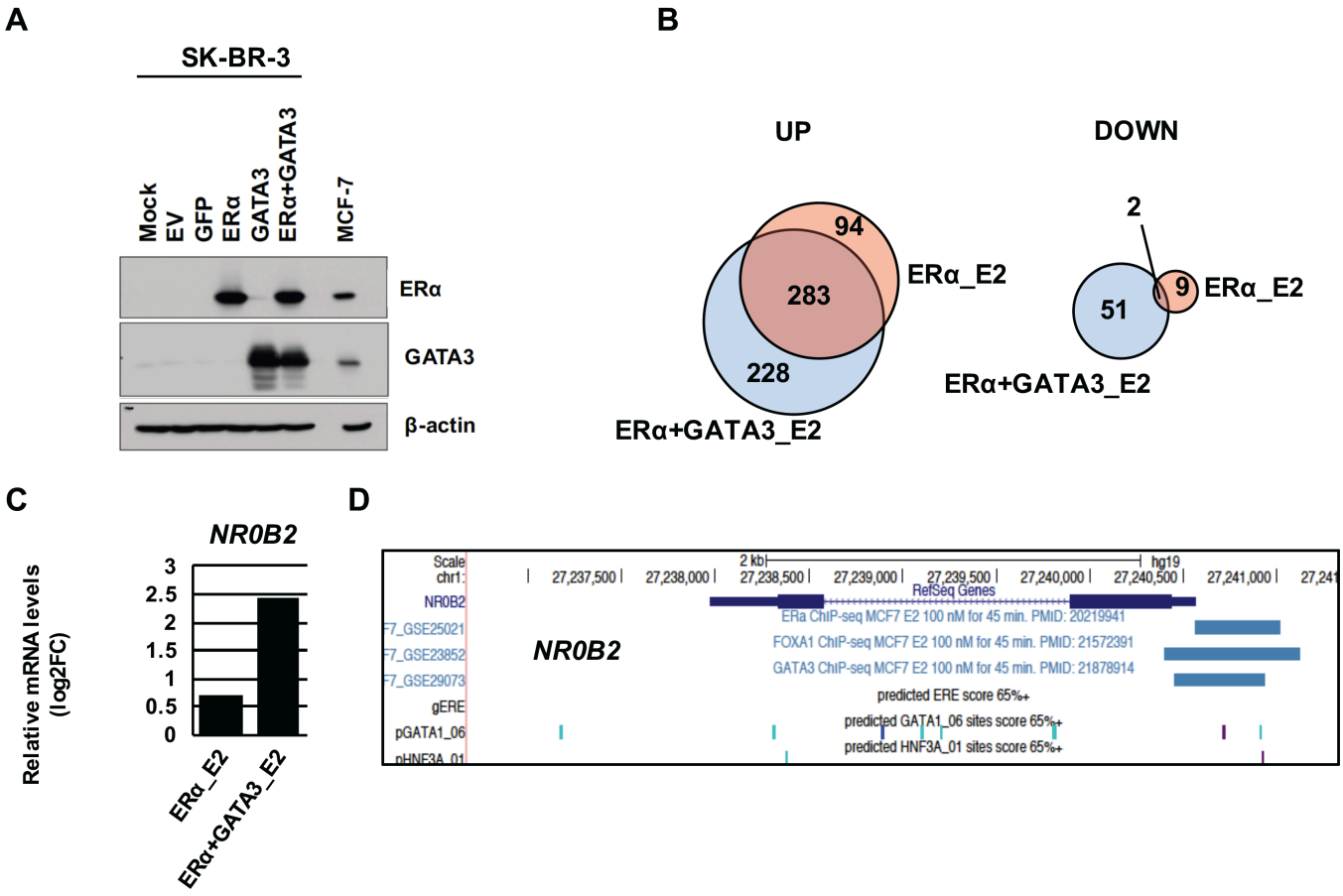
We identified 511 upregulated and 53 downregulated E2 target genes in the ERα+GATA3 co-expression condition (FC>1.4 or FC<-1.4 and q<0.05; ERα+GATA3_E2 vs ERα+GATA3_vehicle). While most genes in the ERα_E2 dataset (**Chapter 2)** were also upregulated by ERα+GATA3_E2, a subset of genes were uniquely significantly regulated by ERα+GATA3_E2 **(Figure 1B).** Although the mechanism(s) by which these genes become significantly regulated by E2 remain(s) unclear, our data is consistent with the previous

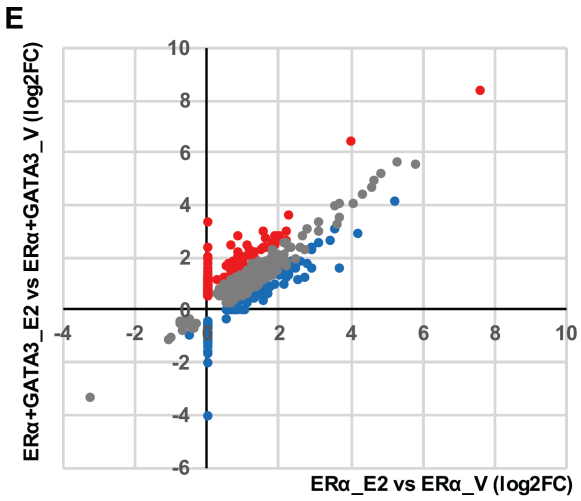
finding that ~ 45% of E2-dependent ERα binding sites in MCF-7 cells overlap with a GATA3 binding site (245). In addition, the number of E2-downregulated genes increased following GATA3 coexpression **(Figure 1B)**. GATA3 binding may result in chromatin opening and increase accessibility of novel EREs, while the gain in transcriptional repression may result from overlapping GATA3 and ERE target sites leading to competition between the two factors for binding to DNA. We postulated that GATA3 coexpression may regulate expression of factors involved in ERα signalling. Indeed, GATA3 coexpression led to increased induction of nuclear receptor subfamily 0 group B member 2 (NR0B2) **(Figure 1C)**, an orphan nuclear receptor that has been shown to inhibit ERα's transcriptional activity through competition for coactivator binding to the AF2 domain of the receptor (486, 487). Notably, *NR0B2* is bound by ERα, FOXA1 and GATA3 in MCF-7 cells **(Figure 1D)**. Additional experiments are needed to examine whether depletion of NR0B2 in SK-BR-3 cells contributes to transcriptional repression of E2 target genes in the presence of GATA3.

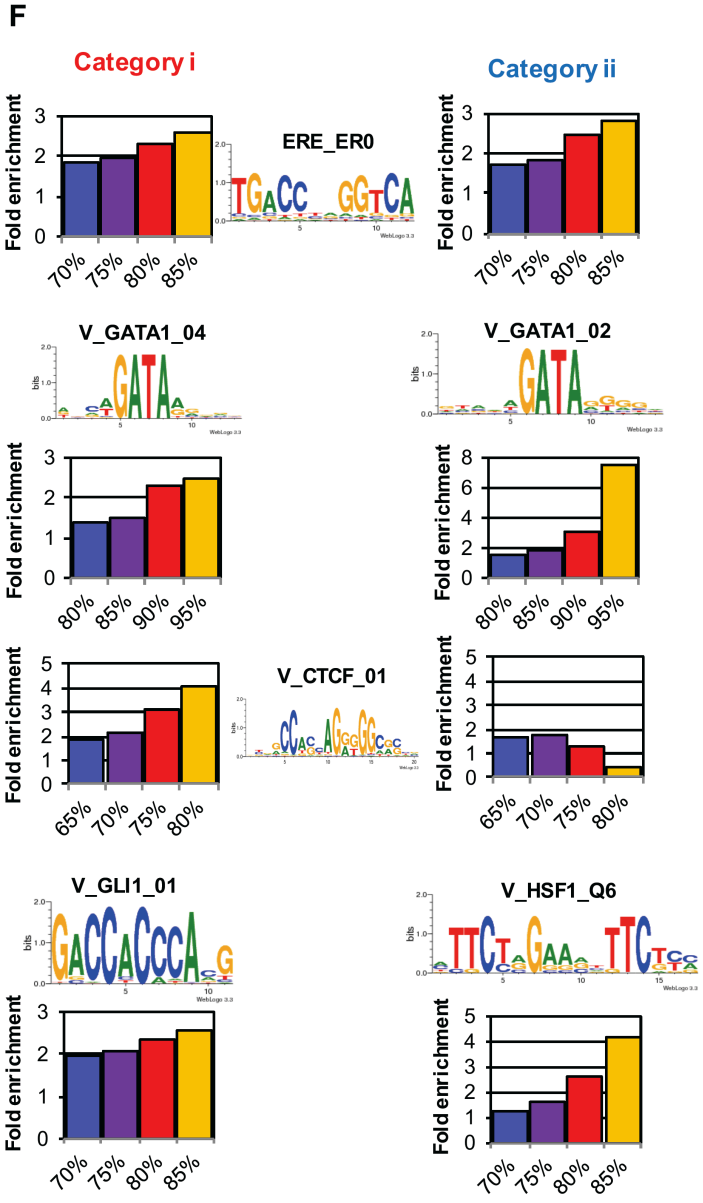
By plotting folds of gene regulation in the ERα+GATA3_E2 condition against those in the ERα_E2 condition for all E2 target genes that are significantly regulated by either condition (667 genes; ERα+GATA3_E2: FC of 1.4, q<0.05 OR ERα_E2: FC of 1.4, q<0.05), we identified three main categories of E2 target genes **(Figure 1E)**: (i) genes whose regulation by E2 is increased (>1.4 fold) upon GATA3 co-expression (145 genes), (ii) genes whose regulation by E2 is decreased (>1.4 fold) upon GATA3 co-expression (114 genes), and (iii) genes whose E2 response was not significantly affected by GATA3 co-expression (408 genes). Our observation that GATA3 coexpression had no significant impact on the majority of E2 target genes indicated that either (1) this TF exerts a minor role in ERα signalling or (2) mechanisms underlying cooperativity between these two TFs are lacking in SK-BR-3 cells. GATA3's ability to increase or decrease E2 regulation on a subset of targets is consistent with the previous report by Theodorou et al. where GATA3 knockdown enhanced ERα binding at 17% of ERα binding sites, and reduced it at 16% of these sites while the remaining 66% were unaffected (244). The mechanism by which GATA3 can increase or decrease ERα binding and signalling remains unclear, but we postulated that this involved cooperativity with other cofactors. TFBS analysis of genes in categories (i) and (ii) revealed enrichment of EREs and GATA motifs, as expected **(Figure 1F)**. Interestingly, regulatory sequences of category (i) genes were enriched in CTCF and GLI1 binding motifs, which were weakly enriched in category (ii) genes. Although a small fraction of ERα binding sites (~8%) overlap with CTCF binding, co-bound sites have been reported to be enriched near E2-target

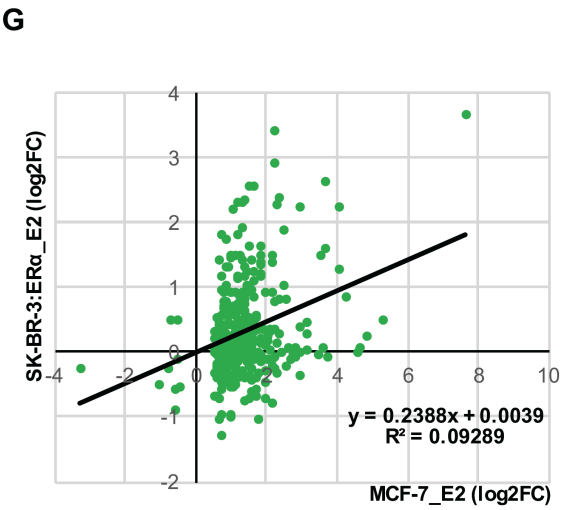
genes, with a bias towards downregulated ones (488). Enrichment of FOXREs and GATA motifs at E2-stimulated CTCF binding sites has also been reported (489). In addition, CTCF knockdown by siRNA can increase ERα association with the chromatin fraction resulting in upregulation and downregulation of 20% and 12% of E2 target genes, respectively (489). It will be of interest to deplete GATA3 and/or CTCF and determine how this impacts ERα binding and transcriptional activity. It is possible that, at a subset of E2 target genes, CTCF acts as an insulator to prevent propagation of heterochromatin and enhance the cooperativity between ERα and GATA3 to facilitate transcription. Enrichment of GLI1 motifs is consistent with the previous finding that GLI1 knockdown reduces expression of ERα as well as transcriptional activation of its downstream targets (490). Category (ii) genes were strongly enriched in binding motifs of heat shock factor 1 (HSF1) **(Figure 1F)**. Interestingly, HSF1 has been shown to repress E2-dependent transcription from a luciferase reporter by interacting with the corepressor protein MTA1, a member of the NuRD complex (491). A potential role of GATA3 in recruitment of HSF1 and MTA1 to repress E2-dependent transcription remains to be examined.

Lastly, we examined the degree of correlation between the transcriptomes of SK-BR-3 cells expressing ERα or ERα+GATA3 to those of MCF-7 cells in the presence of E2 **(Figure 1G)**. Although the correlation between both cell lines was weak, it increased from r=0.3 (ERα alone) to r=0.37 (ERα+GATA3), suggesting that reintroduction of luminal TFs together can better recapitulate the E2 response observed in luminal cells. Nevertheless, this also points to significant divergence between the epigenetic landscape of SK-BR-3 cells and luminal cells, which will be interesting to study using approaches to map chromatin status such as ATAC-Seq.


A
SK-BR-3
Mock
EV
GFP
ERα
GATA3
ERα+GATA3
MCF-7
ERα
GATA3
β-actin
B
UP
DOWN
94
283
228
ERα_E2
ERα+GATA3_E2
2
9
51
ERα_E2
ERα+GATA3_E2
C
NR0B2
Relative mRNA levels (log2FC)
3
2.5
2
1.5
1
0.5
0
ERα_E2
ERα+GATA3_E2
D
Scale
chr1:
2 kb
hg19
27,237,500
27,238,000
27,238,500
27,239,000
27,239,500
27,240,000
27,240,500
27,241,000
NR0B2
RefSeq Genes
F7_GSE25021
F7_GSE23852
F7_GSE29073
NR0B2
ERa ChIP-seq MCF7 E2 100 nM for 45 min. PMID: 20219941
FOXA1 ChIP-seq MCF7 E2 100 nM for 45 min. PMID: 21572391
GATA3 ChIP-seq MCF7 E2 100 nM for 45 min. PMID: 21878914
gERE
predicted ERE score 65%+
pGATA1_06
predicted GATA1_06 sites score 65%+
predicted HNF3A_01 sites score 65%+


E
ERα+GATA3_E2 vs ERα+GATA3_V (log2FC)
ERα_E2 vs ERα_V (log2FC)


F
Category i
Category ii
ERE_ER0
V_GATA1_04
V_GATA1_02
V_CTCF_01
V_GLI1_01
V_HSF1_Q6
Fold enrichment


G
SK-BR-3:ERα_E2 (log2FC)
MCF-7_E2 (log2FC)
y = 0.2388x + 0.0039
R² = 0.09289

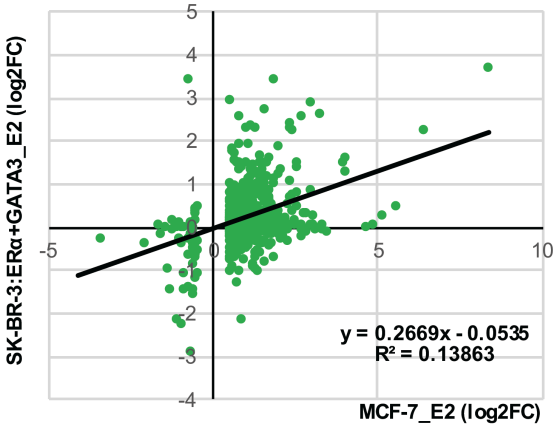

SK-BR-3:ERα+GATA3_E2 (log2FC)
MCF-7_E2 (log2FC)
y = 0.2669x - 0.0535
R² = 0.13863

Figure 1. Impact of GATA3 on ERα signalling.
(A) SK-BR-3 cells were transiently transfected with ERα, GATA3 or both for 72 h and treated with E2 for the last 24 h of transfection. Mock, empty vector and GFP transfections served as negative controls. Protein levels of ERα and GATA3 were assessed by western analysis. β-actin serves as a loading control.
(B) Venn diagrams showing overlap between regulation by ERα_E2 (vs ERα_V) and ERα+GATA3_E2 (vs ERα+GATA3_V) (0.485<log2FC<-0.485 and q<0.05).
(C) *NR0B2* relative mRNA expression levels (RNA-Seq) are shown for the ERα_E2 and ERα+GATA3_E2 conditions.
(D) UCSC genome browser views showing ERα, FOXA1 and GATA3 binding and their respective motifs in E2-stimulated MCF-7 cells.
(E) Folds of regulation in the ERα+GATA3_E2 condition are plotted against folds of regulation in the ERα_E2 condition. Genes whose expression was enhanced by E2 treatment in the GATA3-coexpression condition were defined as: difference between log2FC ERα+GATA3_E2 and log2FC ERα_E2 is greater than +0.485 (red, category i). Genes whose expression was decreased by E2 treatment in the GATA3-coexpression condition were defined as: difference between log2FC ERα+GATA3_E2 and log2FC ERα_E2 is smaller than -0.485 (blue, category ii). Genes whose expression was not significantly altered by E2 treatment in the GATA3-coexpression condition were defined as: difference between log2FC ERα+GATA3_E2 and log2FC ERα_E2 (grey, category iii) is between -0.485 and +0.485.
(F) TFBS enrichment analysis for genes in categories (i) and (ii) is shown.
(G) Degree of correlation between transcriptomes of SK-BR-3 cells expressing either ERα alone or ERα+GATA3 (y-axis; 0.485<log2FC<-0.485 and q<0.05) and MCF-7 cells (x-axis; no cutoffs) in the presence of E2. A correlation coefficient *r* is calculated.

4.2 Is FOXA1 a pioneer factor for ERα recruitment?

In breast cancer, FOXA1 has been traditionally described as a factor that pioneers ERα recruitment to its sites across the genome (**section 3.2.1)**. However, more recent studies indicate that this model is more symmetric than once thought, as nuclear receptors can also pioneer FOXA1 recruitment and chromatin opening. An early study by Rigaud et al. proposed a hit-and-run model whereby liganded glucocorticoid receptor (GR) binds a glucocorticoid response element in the rat tyrosine aminotransferase gene, opens chromatin at that site, and allows subsequent binding of FOXA1 (492). It was also reported that the retinoic acid receptor can bind highly compacted chromatin and mediate its remodelling through recruitment of p300 and SWI/SNF complexes *in vitro* (493). More recently, Swinstead et al. revealed that hormone activation of both ERα and GR can result in *de novo* recruitment of FOXA1 at sites where it was not present in the vehicle treatment condition (494). These sites can be divided into two categories: FOXA1 sites that are co-bound by ERα or GR (and thus sites where ERα or GR can pioneer), and interestingly, FOXA1 sites that are gained by hormonal treatment but do not overlap with ERα or GR binding. More than 90% of FOXA1 sites in the first category were within 100bp of an ERα or GR binding site. In addition, these

FOXA1 induced sites co-bound by ERα were enriched in EREs and AP-1 motifs but weak FOXA1 motifs, suggesting that both ERα and AP-1 may cooperate to recruit FOXA1. They also exhibited an increase in DNase hypersensitivity, indicating that activation of ERα or GR may mediate chromatin opening allowing subsequent FOXA1 binding. On the other hand, while it remains unclear how hormonal stimulation can influence FOXA1 binding at sites where the cognate receptor is absent, two possibilities may explain this: (i) ERα or GR modulate expression of secondary factors that can dictate FOXA1 binding or (ii) ERα or GR bind transiently in a hit-and-run mechanism to open chromatin at these sites and allow FOXA1 binding. Importantly, in this study, the authors challenge the traditional pioneer factor model whereby FOXA1 can stably bind for long durations to keep chromatin open for subsequent TF recruitment. The authors instead demonstrate through single-molecule tracking experiments that FOXA1 binding is rather dynamic with a DNA residence time of approximately 9 seconds. In this newly proposed model referred to as 'Dynamic Assisted Loading', pioneer factors can engage chromatin for a very short period of time to open it and allow binding of other factors. Collectively, these findings suggest that ERα may also act as a pioneer factor. This is consistent with our finding that while prior presence of FOXA1 seems to correlate with chromatin opening that is not affected by ectopic ERα expression in SK-BR-3 cells, ERα can increase chromatin opening on sites containing EREs and weakly bound by FOXA1 (*C5AR2* and *GREB1*). As we have only demonstrated such as role of ERα on two tested sites, it remains important to demonstrate this on a genome-wide level through ChIP-Seq experiments coupled with ATAC-Seq/FAIRE-Seq and TFBS analysis. This will also be addressed in experiments described in the previous section in MDA-MB-231/BT-20 cells. Of note, we have attempted to deplete FOXA1 to ascertain that ERα-mediated chromatin opening at these sites is independent of FOXA1, however, FOXA1 depletion led to reduced expression of the ectopically expressed receptor (data not shown), which precludes a correct interpretation of subsequent ChIP/FAIRE results.

The hypothesis that chromatin remodelling factors are involved in mediating transcriptional activation by FOXA1 has been largely unexplored. Additional studies are needed to address this as well as the possibility that ERα can also recruit chromatin remodelling factors to sites with high-affinity EREs, similar to recruitment of BRG1 to GATA3-bound sites (104). This is plausible as a role for chromatin remodelling factors in assisting TFs in engaging their target genes has been shown for iPSC reprogramming factors, which act both as pioneer and master TFs. For instance, OCT4 mediates chromatin opening at 76%

of its binding sites through BRG1 recruitment (495). Furthermore, BAF155, a BRG-1 associated factor, can increase OCT4 binding to pluripotency genes and enhance reprogramming efficiency (496). Another chromatin remodelling factor, INO80, was shown to be co-bound with OCT4, SOX2 and NANOG at promoters of pluripotency genes to increase chromatin opening and transcriptional activation (497).

In **chapter 2,** using siRNAs and sgRNA-mediated depletion, we show that FOXA1 is indeed a positive regulator of ERα expression in several luminal breast cancer cell lines, consistent with the positive correlation between their expression in breast tumour samples. We did not identify tumour samples that were *ESR1*^{high} and *FOXA1*^{low}, but the opposite is true: molecular apocrine tumours were *ESR1*^{low} and *FOXA1*^{high}. This result indicates that FOXA1 is not sufficient to drive ERα expression in molecular apocrine tumours as in SK-BR-3 cells, similar to the case of GATA3 (**Figure 1B)**. In fact, we find that FOXA1 recruitment is reduced at the *ESR1* upstream regulatory regions in SK-BR-3 cells in comparison to that in MCF-7 cells. A potential explanation for this reduced recruitment is that FOXA1 requires the assistance of (a) chromatin remodelling enzyme(s) or (a) cofactor(s), that is/are not expressed or not functional in SK-BR-3 cells, to facilitate chromatin opening and transcription of *ESR1*. FOXA1 interacting partners on the *ESR1* gene promoter in MCF-7 cells could be identified by performing a DNA-binding assay *in vitro* where MCF-7 nuclear cell extracts are incubated with immobilized *ESR1* promoter sequences (WT or mutant for FOXA1 binding) followed by elution of protein complexes and mass spectrometry, as previously reported for TAL-1 interacting partners in erythroid cells (498). A potential interacting partner of FOXA1 is p53. It has been reported that knockdown of WT p53 in LNCaP cells results in redistribution of AR binding genome-wide. On a number of sites, FOXA1 and AR binding was diminished, suggesting that p53 might play a role in their recruitment (499). Additionally, about 33% of FOXA1 binding sites in the adult mouse liver are enriched in p53 binding motifs and p53 and FOXA1 co-occupancy was validated on a number of these sites (500). Although a direct physical interaction between FOXA1 and p53 has not been reported, both proteins have been shown to physically interact with SMAD3 (501, 502).

As mentioned earlier, *TP53* mutations occur at a much higher frequency in basal-like (80%) and HER2+ (72%) tumours than in luminal tumours (12% of luminal A and 29% of luminal B) (147). These mutations result in loss of tumour suppressor function and act as dominant negative causing inhibition of the DNA-binding activity of WT proteins (503–505). In addition, *TP53* mutations increase the protein's stability resulting in its accumulation in cancer

cells and this is associated with gaining new functions that promote cell growth and tumourigenesis (506, 507). Mutant p53 can exert this effect by regulating activity of other TFs. Di Agostino et al. found that the interaction of mutant p53 with NF-Y, a TF that regulates expression of cell cycle genes, promotes expression of these genes in response to DNA damage contrary to the inhibitory role that WT p53 has (508). In fact, mutant p53 induces transcriptional activation by co-occupying sites bound by NF-Y at promoters of its target genes and increasing recruitment of p300 (508). Through ChIP-on-chip experiments, Stambolsky et al. reported enrichment of vitamin D receptor (VDR) binding motifs in sites bound by mutant p53 in SK-BR-3 cells (509). The authors found that mutant p53 physically interacts with VDR and allows its recruitment to its target genes along with p300, thereby enhancing the vitamin D transcriptional response (509). While vitamin D is reported to promote cell death, it resulted in reduction of cisplatin- and etoposide-induced apoptosis in SK-BR-3 cells in a mutant p53-dependent manner (509).

It is plausible to speculate based on the aforementioned studies that p53 may act as a licensing factor for FOXA1 binding and that targeting mutations may alter this role. Analysis of the mutational status of MCF-7 and SK-BR-3 cells using cBioPortal reveals that *TP53* is mutated in SK-BR-3 (missense mutation R175H) but WT in MCF-7 cells (160, 161, 510). The above hypothesis could explain reduced FOXA1 binding to the *ESR1* promoter in SK-BR-3 cells, especially as (i) p53 positively regulates ERα expression by directly binding the *ESR1* gene promoter in luminal breast cancer cells (**section 3.3.2)** and (ii) mutant p53 has been shown to bind the *ESR1* gene promoter along with HDAC1, MeCP2, and DNMT1 in ERα-negative MDA-MB-468 cells (511, 512). Because it remains unclear how FOXA1 and p53 influence each other's binding on a genome-wide scale, their binding profiles will be examined in ChIP-Seq experiments following depletion (or not) of either TF. Importantly, the impact of the R175H mutant p53 (as well as other p53 mutants that have been characterized in basal-like tumours) on gene expression profiles and FOXA1 binding, will be assessed by RNA-Seq and ChIP-Seq, respectively, following their expression in MCF-7 cells. In particular, FOXA1 binding to *ESR1* and *ESR1* gene expression will be examined. Conversely, the p53 mutation in SK-BR-3 cells could be corrected using the CRISPR-cas9 gene editing system, although recombination of all *TP53* copies is not guaranteed, and residual mutant proteins will inhibit the WT form. Alternatively, we could utilize NSC319726, a mutant-specific p53 reactivator, which can restore WT structure and function to the R175H mutant (513).

4.3 Hydroxamate HDACis repress a luminal transcriptional program

Our observation in Chapter 3 that HDACis repressed expression of luminal-lineage TFs prompted us to examine how these inhibitors affect expression of the highly correlated luminal gene cluster we identified in Chapter 2. As expected, most genes within the luminal cluster are repressed by both TSA and SAHA, whereas in contrast, MS-275 treatment largely resulted in transcriptional activation after 8 h of treatment **(Figure 2)**.

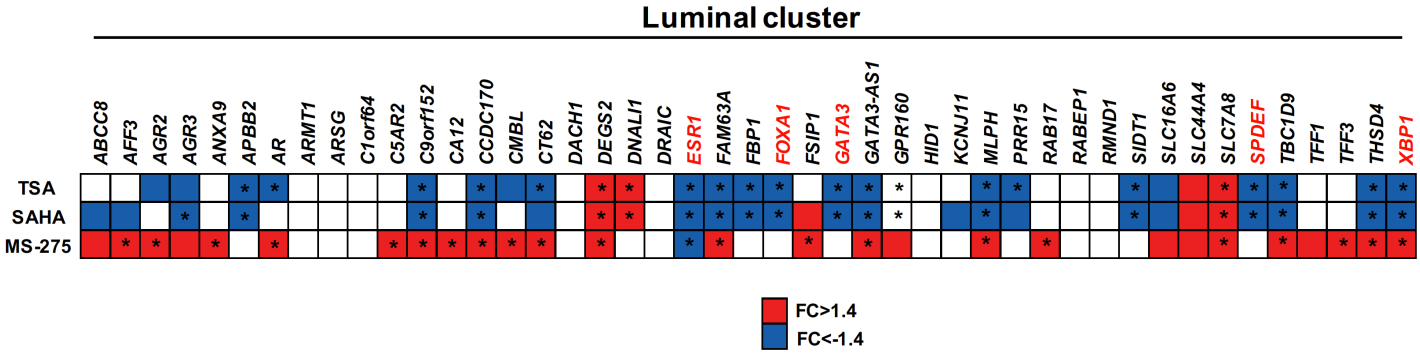


Figure 2. TSA and SAHA repress expression of the luminal cluster genes.
MCF-7 cells were treated with TSA (300nM), SAHA (5μM) or MS-275 (5μM) for 8 h in hormone-depleted media. Relative mRNA levels (RNA-Seq HDACi datasets reported in Chapter 3) of genes in the luminal cluster identified in Chapter 2 are represented as a heatmap.
* denotes statistically significant regulations, q<0.01, N=2

A potential mechanism for HDACi-mediated repression of expression of luminal genes is that it is a consequence of loss of expression of luminal TFs. However, this could not fully explain the observed regulation, as unlike ERα and GATA3, FOXA1 protein expression is not altered after 8 h of treatment with HDACis **(Figure 3A in Chapter 3; also see Figure 5A below)**, although the protein may have lost its function by increased acetylation/loss of expression of cofactors. To test this, FOXA1 acetylation levels will be examined by acetyl-lysine IPs as performed in **Chapter 3** for SREBP-1. Additionally, FOXA1 binding to luminal genes will be examined by ChIP following TSA and SAHA treatment. If indeed FOXA1 binding is altered, this experiment can then be performed on a genome-wide level. Loss of binding of FOXA1 to these genes may also be associated with changes in chromatin accessibility, especially as several chromatin remodelling factors (including *BRG1* (*SMARCA4*)) and HATs (including p300, whose binding was enriched in luminal genes) are repressed by TSA and SAHA treatment. Thus, ChIP against active (H3K9ac, H3K14ac) and repressive (H3K27me3) histone marks, p300, and RNA pol II along with FAIRE- or ATAC-RT-qPCR experiments will be performed to determine whether transcriptional repression of

luminal genes (and increased transcription of lactogenic differentiation genes) correspond to differential accessibility of their enhancers and/or promoters.

In line with the previous hypothesis that p53 may affect FOXA1 binding, we also observed that its expression is repressed following HDACi treatment (**Figure 3).** Transcriptional downregulation of *TP53* occurred concomitantly with downregulation of luminal TFs (**Chapter 3**), beginning as early as 2 h of treatment with maximum repression reached at 16 h. It will be of interest to also test the impact of HDACis on p53 protein expression, to determine whether loss of p53 function occurs early and may contribute to the observed repression of luminal gene expression. Of note, SAHA was reported to induce loss of mutant (but not WT) p53 mRNA and protein in triple-negative breast cancer cell lines via acetylation of transcription factor YY1 and loss of its binding to the the *TP53* promoter (514). Whether a similar mechanism is at play in ERα-positive breast cancer cell lines remains to be investigated.

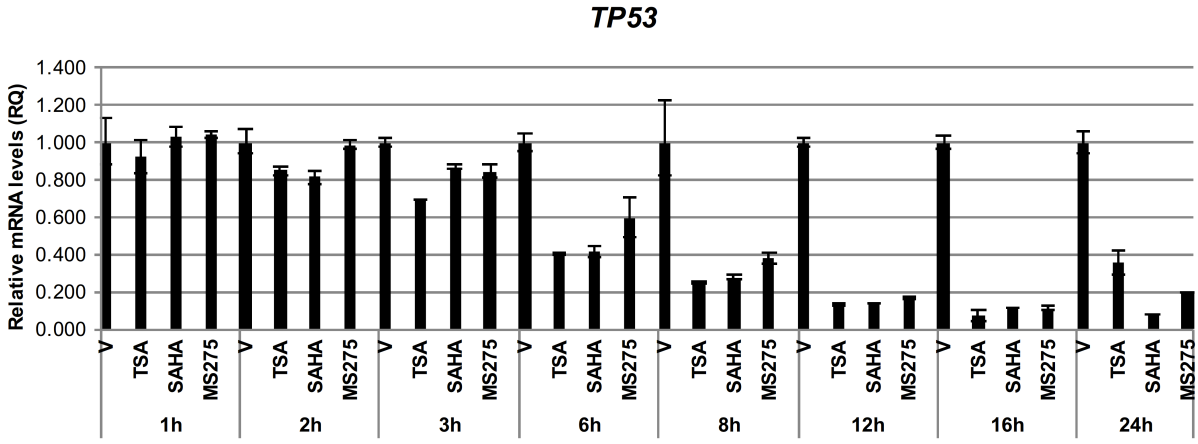


Figure 3. HDACis repress expression of *TP53*.
MCF-7 cells were treated with TSA (300nM), SAHA (5μM) or MS-275 (5μM) for for the indicated time points in hormone-depleted media. Relative mRNA levels (RQ=2^ddCt) were measured by RT-qPCR (N=1).

One question arises from these observations: Does p53 regulate expression of the luminal cluster, and thus, does its loss contribute to HDACi-mediated reprogramming of luminal cells? We addressed this using RNA-Seq datasets of ERα-positive luminal T47Dkbluc breast cancer cells infected with shRNA against luminal TFs *ESR1, FOXA1, GATA3, XBP1* as well as *TP53* (J. Kulpa and S. Mader, manuscript in preparation). Similar to our findings with siFOXA1 in MCF-7 cells **(Chapter 2)**, shFOXA1 resulted in downregulation of expression of most luminal cluster genes in T-47D cells **(Figure 4)**. Strikingly, though, *TP53* knockdown resulted in downregulation of many luminal cluster genes including all luminal TFs **(Figure 4)**.

This result needs to be verified in other ERα-positive breast cancer cell lines including MCF-7 cells. To further test whether downregulation of p53 by HDACis in MCF-7 cells plays a role in the loss of luminal TF expression in the presence of HDACis, p53 can be overexpressed in the presence of HDACis and expression of luminal TFs can be examined through RT-qPCRs and western analysis. Levels of p53 acetylation will also be examined following HDACi treatment. p53 has been one of the first non-histone proteins shown to be acetylated by HATs, including MOZ/MOF/Tip60 and p300/CBP/PCAF (515). Although acetylation of p53 has been shown to contribute to its transcriptional activation and tumour-suppressive properties, and protect it against degradation by MDM2 (515), its impact can be context-, residue- and target promoter-specific. Ultimately, it will also be important to examine whether p53 ChIP binding regions are enriched in regulatory sequences of luminal cluster genes. This will help determine whether transcriptional regulation of the luminal cluster by p53 occurs directly, or indirectly via downregulation of luminal TFs. Finally, it will be interesting to examine how various p53 mutations affect expression of the luminal cluster.

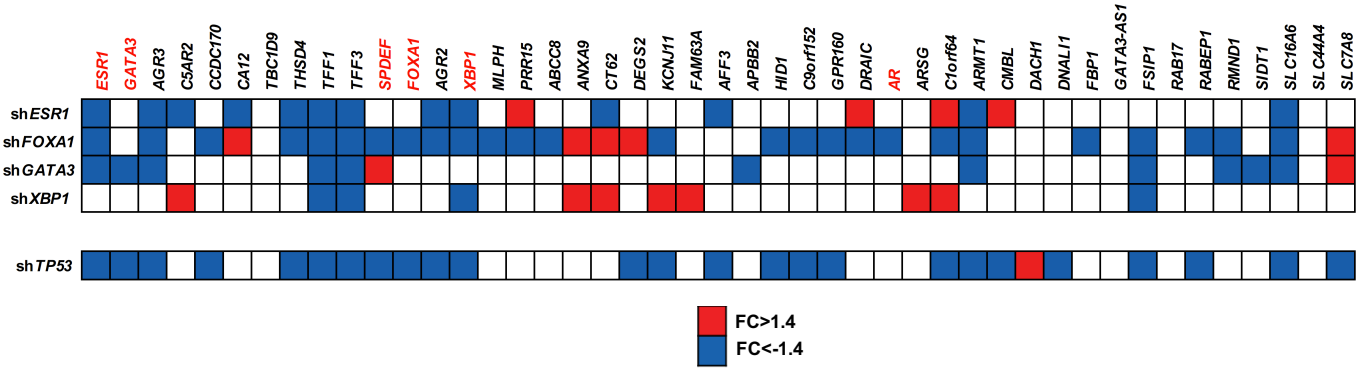


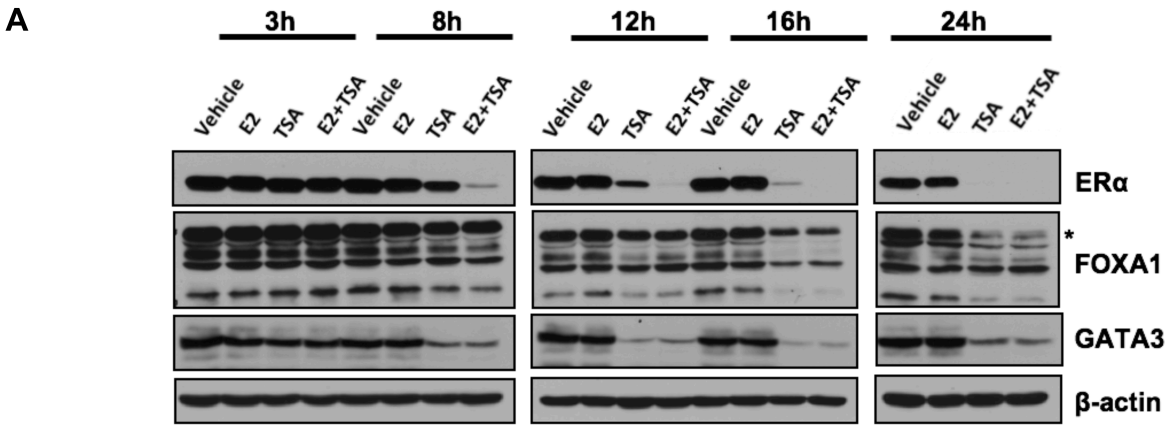
Figure 4. p53 positively regulates expression of luminal cluster genes.
T-47Dkbluc cells were infected with shRNA against *ESR1, FOXA1, GATA3, XBP1* (luminal TFs) or *TP53* and RNA-Seq profiles were generated (N=2). Relative expression of genes in the luminal cluster identified in Chapter 2 is shown as a heatmap.

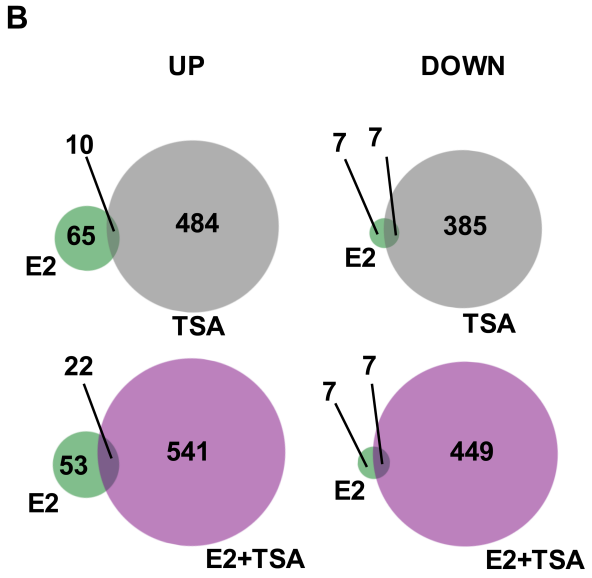
4.4 TSA represses E2 signalling in MCF-7 cells independently of ERα expression

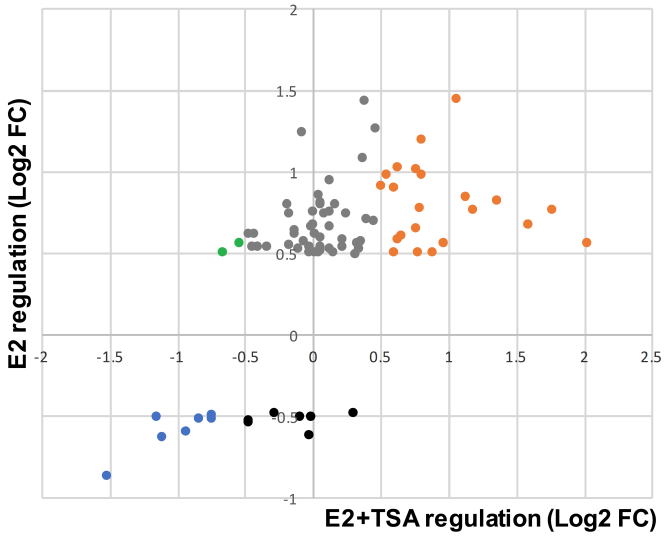
In Chapter 3, we show that treatment of MCF-7 cells with TSA for 24 h abrogated the E2 response consistent with suppression of ERα protein levels. Notably, TSA-mediated downregulation of ERα protein levels was enhanced by hormone treatment at 8, 12 and 16 h, consistent with the previously reported role of E2 in increasing the receptor's turnover (516), while FOXA1 and GATA3 protein levels were downregulated by TSA irrespective of hormone

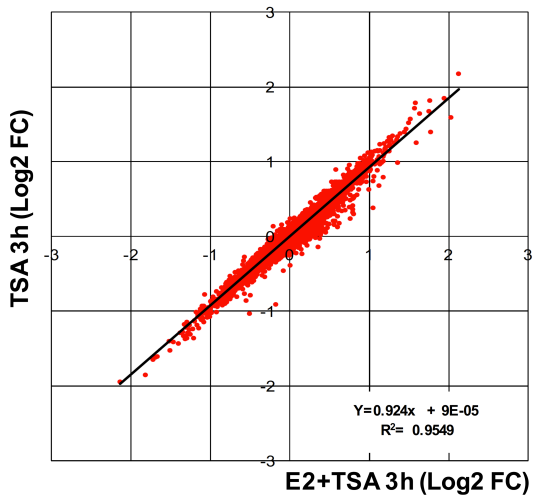
treatment **(Figure 5A)**. We have additionally examined gene expression profiles of MCF-7 cells following a short treatment time (3h) with E2, TSA or E2+TSA. ERα, FOXA1 and GATA3 protein expression levels were unaffected by TSA treatment at this time as shown in **Figure 5A.** Similar to regulation observed at 24 h **(Figure 1, Chapter 3)**, regulation by E2 treatment at 3 h **(Figure 5B)** was biased towards transcriptional activation (75 genes UP and 14 DOWN, FC>1.4 or <-1.4, and p<0.01), whereas TSA exerted almost equal influence on gene transcription in the absence (494 UP, 392 DOWN) or presence (563 UP, 456 DOWN) of E2, albeit with a smaller number of regulated genes under all three conditions. Plotting folds of regulation of the 89 E2 targets against their folds of regulation by E2+TSA reveals that the majority of these genes (54/89) become unregulated upon co-treatment with TSA (grey (47 upregulated genes) and black dots (7 downregulated genes)) (Figure 5C). Consistent with this result, a strong linear correlation is observed in the TSA vs E2+TSA log2FC plot in **Figure 5D**, suggesting that E2 exerts no pronounceable impact in the presence of TSA. Shutdown of E2 signalling was confirmed by RT-qPCRs on *XBP1* and *GREB1*, two E2 target genes as shown in **Figure 5E**. This intriguing result suggested that ERα or any of its signalling cascade components are not functional in the presence of TSA at early time points of treatment. Interestingly, TSA treatment of LNCaP cells has been previously shown to block activation of AR target genes in the presence of an AR ligand, R1881 (517). While TSA treatment did not alter AR binding to its target genes, it prevented recruitment of its coactivator, SRC1, as well as p300 and RNA pol II (517). Thus, we will examine whether recruitment of ERα, FOXA1, the p160 family of coactivators (SRC1, SRC2, SRC3), as well as p300 and RNA pol II is affected by TSA treatment on promoters/enhancers of ERα targets in MCF-7 cells. As mentioned earlier, FOXA1 binding to DNA can be negatively affected by acetylation (298). In the case of ERα, its acetylation on residues K268 and K288 has been shown to increase DNA binding and transcriptional activity of the receptor (274), whereas acetylation of its K302 and K303 residues reduces transcriptional activation in luciferase assays (518). Moreover, acetylation of two ERα coactivators, PGC-1α and SRC3, represses their function (519, 520). Therefore, we will also determine whether acetylation of these factors is induced in the presence of E2+TSA after 3 h of treatment and whether that might contribute to suppression of E2 signalling. FAIRE- or ATAC-Seq experiments can also be performed to study any changes in chromatin accessibility by TSA treatment. Ultimately, it will be crucial to identify the specific HDAC proteins responsible for this phenotype as well as loss of expression of luminal genes described in the previous section through shRNA or siRNA

targeting of individual HDACs. This would provide a strong foundation for the development of more selective inhibitors.


A
3h
8h
12h
16h
24h
Vehicle
E2
TSA
E2+TSA
ERα
FOXA1
GATA3
β-actin
*


B
UP
DOWN
10
65
484
E2
TSA
7
7
385
E2
TSA
22
53
541
E2
E2+TSA
7
7
449
E2
E2+TSA


C
E2 regulation (Log2 FC)
E2+TSA regulation (Log2 FC)


D
TSA 3h (Log2 FC)
E2+TSA 3h (Log2 FC)
Y=0.924x + 9E-05
R²= 0.9549

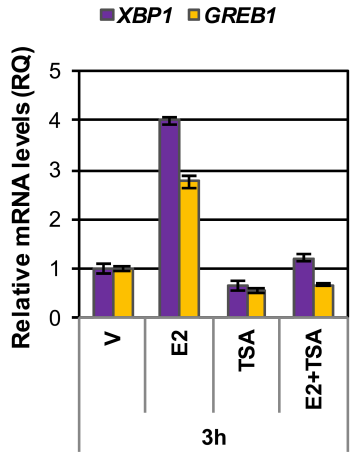

E
XBP1
GREB1
Relative mRNA levels (RQ)
V
E2
TSA
E2+TSA
3h

Figure 5. TSA suppresses E2 signalling in MCF-7 cells independently of ERα expression.
MCF-7 cells cultured in hormone-depleted media were treated with E2 (25 nM), TSA (300 nM), or E2+TSA for 3, 8, 12, 16 and 24 h. (A) Protein expression of ERα, FOXA1 and GATA3 was examined by western blots.
(B) Microarray gene expression profiles of MCF-7 cells treated for 3 h with E2, TSA, or E2+TSA were generated (compared to vehicle control). The overlap between the E2 and TSA or E2+TSA genesets is shown as Venn diagrams for the UP and DOWN categories (FC>1.4 or FC<-1.4, $p<0.01$).
(C) Fold change values of significantly regulated E2 target genes (FC>1.4 or FC<-1.4, $p<0.01$) were plotted against those following E2+TSA treatment.
(D) Fold change values of genes regulated in the TSA condition were plotted against those in the E2+TSA condition.
(E) Relative mRNA levels of *XBP1* and *GREB1*, two E2 target genes, were measured by RT-qPCR.

4.5 Is FOXA1 a therapeutic target in breast cancer?

Because of the role that FOXA1 plays in activating ERα signalling, it has been postulated that its targeting may be beneficial for treatment of luminal breast cancers. The rationale for this hypothesis is also based on previous findings that knockdown of FOXA1 in breast cancer cells inhibits their proliferation (167, 202, 232, 521). Efforts are currently ongoing to identify small molecule inhibitors of kinases that positively regulate FOXA1 activity (522). However, a recent study by Bernardo et al. has shown that even though knockdown of FOXA1 in MCF-7 luminal breast cancer cells reduces cell proliferation, it also induces expression of basal-like markers, which is accompanied by increased migration and invasion (212). It was postulated that because MCF-7 cells are a heterogeneous cell population, FOXA1 silencing may result in growth arrest of more differentiated cells but at the same time enrich for cells with more plasticity to de-differentiate to a more basal-like phenotype (212). Another report showed that ectopic expression of FOXA1 in MDA-MB-231 cells increases expression of E-cadherin and reduces their migratory potential (523). Our results that FOXA1 knockdown results in reduced expression of luminal markers is compatible with these studies, although alternative fates include basal-like and possibly also lactogenic differentiation, since suppression of FOXA1 is expected to also lead to suppression of ERα expression. Clearly, a FOXA1-targetting approach is unlikely to be applied in ERα-positive breast tumours unless they have already developed resistance to hormonal therapy.

Undoubtedly, additional studies are needed to examine whether deletion or overexpression of FOXA1 could serve as a potential therapy in breast cancer. The role of FOXA1 in tumourigenesis and metastasis will need to be assessed *in vivo* by overexpressing or knocking down FOXA1 in mouse models of basal-like (MMTV-*Wnt1*) or luminal (MMTV-PyMT) breast tumours, respectively. In addition, it will be important to examine whether FOXA1 deletion or overexpression in the normal mouse mammary gland results in tumour formation as mice age.

Concluding Remarks

It is becoming more apparent that breast cancer is a disease of transcriptional addiction with each molecular subtype characterized by a distinct set of deregulated transcriptional programs. In studies presented herein, we have identified ERα, FOXA1 and GATA3 as three TFs whose expression or lack thereof can largely recapitulate three main breast tumour subtypes. We further demonstrate that FOXA1 governs this transcription factor hierarchy, positively controlling expression of the luminal transcriptional network in breast cancer. Our novel finding that molecular apocrine tumours express only a subset of FOXA1-dependent luminal genes sheds light on the master regulator role of this TF in specifying the phenotype of this understudied subtype. While FOXA1 is thought to be a pioneer factor for ERα recruitment, we show that ERα can also open chromatin on sites weakly bound by FOXA1, suggesting that the receptor may open chromatin and possibly act as a pioneer factor for other TFs, although these conclusions will need to be verified on a larger scale. Our studies additionally indicate that transcriptional shutoff of the *ESR1* gene promoter in molecular apocrine cells correlates with reduced binding of FOXA1, which positively regulates expression of the receptor in luminal breast cancer cells. Gaining insight into the function of TFs that are upstream of ERα will determine whether targeting them is beneficial for treatment of ERα-positive luminal breast cancers and whether they could be used as biomarkers for subtype classification. The unravelling of transcriptional programs characteristic of different tumour subtypes may help determine whether their manipulation can convert an aggressive tumour to a more differentiated one.

For the first time, we demonstrate that chemical inhibition of HDAC proteins represses a luminal transcriptional program comprising TFs ERα, FOXA1 and GATA3. This was associated with induction of differentiation markers that are normally upregulated during lactation in the human breast. Our study also identifies a role of p300 in mediating HDACi effects. Experiments exploring the mechanisms of action of HDACis are needed, as these will contribute to the rationale of using them in combined therapies with antiestrogens for treatment of breast cancer or designing of more specific drugs.

Finally, characterization of epigenetic regulators and TFs orchestrating the luminal phenotype and governing *ESR1* expression will not only provide insight into mechanisms of breast tumour subtype specification, but should upon characterization of the patterns of expression of these TFs in normal human epithelial cells clarify the relationship between these subtypes and the normal epithelial differentiation process during mammary gland development and lactation.

Bibliography

1. Hennighausen,L. and Robinson,G.W. (2005) Information networks in the mammary gland. *Nat. Rev. Mol. Cell Biol.*, **6**, 715–725.

2. Paine,I.S. and Lewis,M.T. (2017) The Terminal End Bud: the Little Engine that Could. *J. Mammary Gland Biol. Neoplasia*, **22**, 93–108.

3. Williams,J.M. and Daniel,C.W. (1983) Mammary ductal elongation: differentiation of myoepithelium and basal lamina during branching morphogenesis. *Dev. Biol.*, **97**, 274–290.

4. Humphreys,R.C., Krajewska,M., Krnacik,S., Jaeger,R., Weiher,H., Krajewski,S., Reed,J.C. and Rosen,J.M. (1996) Apoptosis in the terminal endbud of the murine mammary gland: a mechanism of ductal morphogenesis. *Dev. Camb. Engl.*, **122**, 4013–4022.

5. Nandi,S. (1958) Endocrine control of mammarygland development and function in the C3H/ He Crgl mouse. *J. Natl. Cancer Inst.*, **21**, 1039–1063.

6. Lyons,W.R. (1958) Hormonal synergism in mammary growth. *Proc. R. Soc. Lond. B Biol. Sci.*, **149**, 303–325.

7. Daniel,C.W., Silberstein,G.B. and Strickland,P. (1987) Direct action of 17 beta-estradiol on mouse mammary ducts analyzed by sustained release implants and steroid autoradiography. *Cancer Res.*, **47**, 6052–6057.

8. Silberstein,G.B., Van Horn,K., Shyamala,G. and Daniel,C.W. (1994) Essential role of endogenous estrogen in directly stimulating mammary growth demonstrated by implants containing pure antiestrogens. *Endocrinology*, **134**, 84–90.

9. Atwood,C.S., Hovey,R.C., Glover,J.P., Chepko,G., Ginsburg,E., Robison,W.G. and Vonderhaar,B.K. (2000) Progesterone induces side-branching of the ductal epithelium in the mammary glands of peripubertal mice. *J. Endocrinol.*, **167**, 39–52.

10. Macias,H. and Hinck,L. (2012) Mammary gland development. *Wiley Interdiscip. Rev. Dev. Biol.*, **1**, 533–557.

11. McManaman,J.L. and Neville,M.C. (2003) Mammary physiology and milk secretion. *Adv. Drug Deliv. Rev.*, **55**, 629–641.

12. Treuting,P.M., Dintzis,S.M., Frevert,C.W., Liggitt,H.D. and Montine,K.S. (2012) Comparative Anatomy and Histology: A Mouse and Human Atlas Academic Press.

13. Anderson,S.M., Rudolph,M.C., McManaman,J.L. and Neville,M.C. (2007) Key stages in mammary gland development. Secretory activation in the mammary gland: it's not just about milk protein synthesis! *Breast Cancer Res. BCR*, **9**, 204.

14. Mills,E.S. and Topper,Y.J. (1970) Some ultrastructural effects of insulin, hydrocortisone, and prolactin on mammary gland explants. *J. Cell Biol.*, **44**, 310–328.

15. Rosano,T.G. and Jones,D.H. (1976) Developmental changes in mitochondria during the transition into lactation in the mouse mammary gland. *J. Cell Biol.*, **69**, 573–580.

16. Jerry,D.J., Kuperwasser,C., Downing,S.R., Pinkas,J., He,C., Dickinson,E., Marconi,S. and Naber,S.P. (1998) Delayed involution of the mammary epithelium in BALB/c-p53null mice. *Oncogene*, **17**, 2305–2312.

17. Lund,L.R., Rømer,J., Thomasset,N., Solberg,H., Pyke,C., Bissell,M.J., Danø,K. and Werb,Z. (1996) Two distinct phases of apoptosis in mammary gland involution: proteinase-independent and -dependent pathways. *Dev. Camb. Engl.*, **122**, 181–193.

18. Alexander,C.M., Selvarajan,S., Mudgett,J. and Werb,Z. (2001) Stromelysin-1 regulates adipogenesis during mammary gland involution. *J. Cell Biol.*, **152**, 693–703.

19. Clarkson,R.W.E., Boland,M.P., Kritikou,E.A., Lee,J.M., Freeman,T.C., Tiffen,P.G. and Watson,C.J. (2006) The genes induced by signal transducer and activators of transcription (STAT)3 and STAT5 in mammary epithelial cells define the roles of these STATs in mammary development. *Mol. Endocrinol. Baltim. Md*, **20**, 675–685.

20. Kritikou,E.A., Sharkey,A., Abell,K., Came,P.J., Anderson,E., Clarkson,R.W.E. and Watson,C.J. (2003) A dual, non-redundant, role for LIF as a regulator of development and STAT3-mediated cell death in mammary gland. *Dev. Camb. Engl.*, **130**, 3459–3468.

21. Abell,K., Bilancio,A., Clarkson,R.W.E., Tiffen,P.G., Altaparmakov,A.I., Burdon,T.G., Asano,T., Vanhaesebroeck,B. and Watson,C.J. (2005) Stat3-induced apoptosis requires a molecular switch in PI(3)K subunit composition. *Nat. Cell Biol.*, **7**, 392–398.

22. Tonner,E., Barber,M.C., Travers,M.T., Logan,A. and Flint,D.J. (1997) Hormonal control of insulin-like growth factor-binding protein-5 production in the involuting mammary gland of the rat. *Endocrinology*, **138**, 5101–5107.

23. Deome,K.B., Faulkin,L.J., Bern,H.A. and Blair,P.B. (1959) Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. *Cancer Res.*, **19**, 515–520.

24. Hoshino,K. (1964) REGENERATION AND GROWTH OF QUANTITATIVELY TRANSPLANTED MAMMARY GLANDS OF NORMAL FEMALE MICE. *Anat. Rec.*, **150**, 221–235.

25. Ehmann,U.K., Guzman,R.C., Osborn,R.C., Young,J.T., Cardiff,R.D. and Nandi,S. (1987) Cultured mouse mammary epithelial cells: normal phenotype after implantation. *J. Natl. Cancer Inst.*, **78**, 751–757.

26. Kordon,E.C. and Smith,G.H. (1998) An entire functional mammary gland may comprise the progeny from a single cell. *Dev. Camb. Engl.*, **125**, 1921–1930.

27. Sleeman,K.E., Kendrick,H., Ashworth,A., Isacke,C.M. and Smalley,M.J. (2006) CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. *Breast Cancer Res. BCR*, **8**, R7.

28. Shackleton,M., Vaillant,F., Simpson,K.J., Stingl,J., Smyth,G.K., Asselin-Labat,M.-L., Wu,L., Lindeman,G.J. and Visvader,J.E. (2006) Generation of a functional mammary gland from a single stem cell. *Nature*, **439**, 84–88.

29. Van Keymeulen,A., Rocha,A.S., Ousset,M., Beck,B., Bouvencourt,G., Rock,J., Sharma,N., Dekoninck,S. and Blanpain,C. (2011) Distinct stem cells contribute to mammary gland development and maintenance. *Nature*, **479**, 189–193.

30. Keller,P.J., Arendt,L.M. and Kuperwasser,C. (2011) Stem cell maintenance of the mammary gland: it takes two. *Cell Stem Cell*, **9**, 496–497.

31. Shehata,M., Teschendorff,A., Sharp,G., Novcic,N., Russell,I.A., Avril,S., Prater,M., Eirew,P., Caldas,C., Watson,C.J., *et al.* (2012) Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. *Breast Cancer Res. BCR*, **14**, R134.

32. Takahashi,K. and Yamanaka,S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, **126**, 663–676.

33. Prossnitz,E.R., Arterburn,J.B., Smith,H.O., Oprea,T.I., Sklar,L.A. and Hathaway,H.J. (2008) Estrogen Signaling through the Transmembrane G Protein–Coupled Receptor GPR30. *Annu. Rev. Physiol.*, **70**, 165–190.

34. Babiker,F.A., De Windt,L.J., van Eickels,M., Grohe,C., Meyer,R. and Doevendans,P.A. (2002) Estrogenic hormone action in the heart: regulatory network and function. *Cardiovasc. Res.*, **53**, 709–719.

35. Arnal,J.-F., Lenfant,F., Metivier,R., Flouriot,G., Henrion,D., Adlanmerini,M., Fontaine,C., Gourdy,P., Chambon,P., Katzenellenbogen,B., *et al.* (2017) Membrane and Nuclear Estrogen Receptor Alpha Actions: From Tissue Specificity to Medical Implications. *Physiol. Rev.*, **97**, 1045–1087.

36. Driscoll,M.D., Sathya,G., Muyan,M., Klinge,C.M., Hilf,R. and Bambara,R.A. (1998) Sequence requirements for estrogen receptor binding to estrogen response elements. *J. Biol. Chem.*, **273**, 29321–29330.

37. Traboulsi,T., El Ezzy,M., Gleason,J.L. and Mader,S. (2017) Antiestrogens: structure-activity relationships and use in breast cancer treatment. *J. Mol. Endocrinol.*, **58**, R15–R31.

38. Liang,J. and Shang,Y. (2013) Estrogen and cancer. *Annu. Rev. Physiol.*, **75**, 225–240.

39. Spencer,T.E., Jenster,G., Burcin,M.M., Allis,C.D., Zhou,J., Mizzen,C.A., McKenna,N.J., Onate,S.A., Tsai,S.Y., Tsai,M.J., *et al.* (1997) Steroid receptor coactivator-1 is a histone acetyltransferase. *Nature*, **389**, 194–198.

40. Chen,H., Lin,R.J., Schiltz,R.L., Chakravarti,D., Nash,A., Nagy,L., Privalsky,M.L., Nakatani,Y. and Evans,R.M. (1997) Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell*, **90**, 569–580.

41. Heery,D.M., Kalkhoven,E., Hoare,S. and Parker,M.G. (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature*, **387**, 733–736.

42. Mak,H.Y., Hoare,S., Henttu,P.M. and Parker,M.G. (1999) Molecular determinants of the estrogen receptor-coactivator interface. *Mol. Cell. Biol.*, **19**, 3895–3903.

43. Hanstein,B., Eckner,R., DiRenzo,J., Halachmi,S., Liu,H., Searcy,B., Kurokawa,R. and Brown,M. (1996) p300 is a component of an estrogen receptor coactivator complex. *Proc. Natl. Acad. Sci. U. S. A.*, **93**, 11540–11545.

44. Kraus,W.L. and Kadonaga,J.T. (1998) p300 and estrogen receptor cooperatively activate transcription via differential enhancement of initiation and reinitiation. *Genes Dev.*, **12**, 331–342.

45. Chen,D., Ma,H., Hong,H., Koh,S.S., Huang,S.M., Schurter,B.T., Aswad,D.W. and Stallcup,M.R. (1999) Regulation of transcription by a protein methyltransferase. *Science*, **284**, 2174–2177.

46. Chen,D., Huang,S.M. and Stallcup,M.R. (2000) Synergistic, p160 coactivator-dependent enhancement of estrogen receptor function by CARM1 and p300. *J. Biol. Chem.*, **275**, 40810–40816.

47. Teyssier,C., Chen,D. and Stallcup,M.R. (2002) Requirement for multiple domains of the protein arginine methyltransferase CARM1 in its transcriptional coactivator function. *J. Biol. Chem.*, **277**, 46066–46072.

48. Chiba,H., Muramatsu,M., Nomoto,A. and Kato,H. (1994) Two human homologues of Saccharomyces cerevisiae SWI2/SNF2 and Drosophila brahma are transcriptional coactivators cooperating with the estrogen receptor and the retinoic acid receptor. *Nucleic Acids Res.*, **22**, 1815–1820.

49. Belandia,B., Orford,R.L., Hurst,H.C. and Parker,M.G. (2002) Targeting of SWI/SNF chromatin remodelling complexes to estrogen-responsive genes. *EMBO J.*, **21**, 4094–4103.

50. Brown,A.M., Jeltsch,J.M., Roberts,M. and Chambon,P. (1984) Activation of pS2 gene transcription is a primary response to estrogen in the human breast cancer cell line MCF-7. *Proc. Natl. Acad. Sci. U. S. A.*, **81**, 6344–6348.

51. Métivier,R., Penot,G., Hübner,M.R., Reid,G., Brand,H., Kos,M. and Gannon,F. (2003) Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell*, **115**, 751–763.

52. Fernandes,I., Bastien,Y., Wai,T., Nygard,K., Lin,R., Cormier,O., Lee,H.S., Eng,F., Bertos,N.R., Pelletier,N., *et al.* (2003) Ligand-Dependent Nuclear Receptor Corepressor LCoR Functions by Histone Deacetylase-Dependent and -Independent Mechanisms. *Mol. Cell*, **11**, 139–150.

53. Delage-Mourroux,R., Martini,P.G., Choi,I., Kraichely,D.M., Hoeksema,J. and Katzenellenbogen,B.S. (2000) Analysis of estrogen receptor interaction with a repressor of

estrogen receptor activity (REA) and the regulation of estrogen receptor transcriptional activity by REA. *J. Biol. Chem.*, **275**, 35848–35856.

54. Kurtev,V., Margueron,R., Kroboth,K., Ogris,E., Cavailles,V. and Seiser,C. (2004) Transcriptional regulation by the repressor of estrogen receptor activity via recruitment of histone deacetylases. *J. Biol. Chem.*, **279**, 24834–24843.

55. Umayahara,Y., Kawamori,R., Watada,H., Imano,E., Iwama,N., Morishima,T., Yamasaki,Y., Kajimoto,Y. and Kamada,T. (1994) Estrogen regulation of the insulin-like growth factor I gene transcription involves an AP-1 enhancer. *J. Biol. Chem.*, **269**, 16433–16442.

56. Safe,S. (2001) Transcriptional activation of genes by 17 beta-estradiol through estrogen receptor-Sp1 interactions. *Vitam. Horm.*, **62**, 231–252.

57. Tzagarakis-Foster,C., Geleziunas,R., Lomri,A., An,J. and Leitman,D.C. (2002) Estradiol represses human T-cell leukemia virus type 1 Tax activation of tumor necrosis factor-alpha gene transcription. *J. Biol. Chem.*, **277**, 44772–44777.

58. Lalmansingh,A.S. and Uht,R.M. (2008) Estradiol regulates corticotropin-releasing hormone gene (crh) expression in a rapid and phasic manner that parallels estrogen receptor-alpha and -beta recruitment to a 3',5'-cyclic adenosine 5'-monophosphate regulatory region of the proximal crh promoter. *Endocrinology*, **149**, 346–357.

59. Liu,W., Konduri,S.D., Bansal,S., Nayak,B.K., Rajasekaran,S.A., Karuppayil,S.M., Rajasekaran,A.K. and Das,G.M. (2006) Estrogen receptor-alpha binds p53 tumor suppressor protein directly and represses its function. *J. Biol. Chem.*, **281**, 9837–9840.

60. Stender,J.D., Kim,K., Charn,T.H., Komm,B., Chang,K.C.N., Kraus,W.L., Benner,C., Glass,C.K. and Katzenellenbogen,B.S. (2010) Genome-wide analysis of estrogen receptor alpha DNA binding and tethering mechanisms identifies Runx1 as a novel tethering factor in receptor-mediated transcriptional activation. *Mol. Cell. Biol.*, **30**, 3943–3955.

61. Kato,S., Endoh,H., Masuhiro,Y., Kitamoto,T., Uchiyama,S., Sasaki,H., Masushige,S., Gotoh,Y., Nishida,E., Kawashima,H., *et al.* (1995) Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. *Science*, **270**, 1491–1494.

62. Bunone,G., Briand,P.A., Miksicek,R.J. and Picard,D. (1996) Activation of the unliganded estrogen receptor by EGF involves the MAP kinase pathway and direct phosphorylation. *EMBO J.*, **15**, 2174–2183.

63. Thomas,R.S., Sarwar,N., Phoenix,F., Coombes,R.C. and Ali,S. (2008) Phosphorylation at serines 104 and 106 by Erk1/2 MAPK is important for estrogen receptor-alpha activity. *J. Mol. Endocrinol.*, **40**, 173–184.

64. Neuman,E., Ladha,M.H., Lin,N., Upton,T.M., Miller,S.J., DiRenzo,J., Pestell,R.G., Hinds,P.W., Dowdy,S.F., Brown,M., *et al.* (1997) Cyclin D1 stimulation of estrogen receptor transcriptional activity independent of cdk4. *Mol. Cell. Biol.*, **17**, 5338–5347.

65. Zwijsen,R.M., Buckle,R.S., Hijmans,E.M., Loomans,C.J. and Bernards,R. (1998) Ligand-independent recruitment of steroid receptor coactivators to estrogen receptor by cyclin D1. *Genes Dev.*, **12**, 3488–3498.

66. Wyckoff,M.H., Chambliss,K.L., Mineo,C., Yuhanna,I.S., Mendelsohn,M.E., Mumby,S.M. and Shaul,P.W. (2001) Plasma membrane estrogen receptors are coupled to endothelial nitric-oxide synthase through Galpha(i). *J. Biol. Chem.*, **276**, 27071–27076.

67. Migliaccio,A., Di Domenico,M., Castoria,G., de Falco,A., Bontempo,P., Nola,E. and Auricchio,F. (1996) Tyrosine kinase/p21ras/MAP-kinase pathway activation by estradiol-receptor complex in MCF-7 cells. *EMBO J.*, **15**, 1292–1300.

68. Acconcia,F., Ascenzi,P., Bocedi,A., Spisni,E., Tomasi,V., Trentalance,A., Visca,P. and Marino,M. (2005) Palmitoylation-dependent estrogen receptor alpha membrane localization: regulation by 17beta-estradiol. *Mol. Biol. Cell*, **16**, 231–237.

69. Shyamala,G., Chou,Y.-C., Louie,S.G., Guzman,R.C., Smith,G.H. and Nandi,S. (2002) Cellular expression of estrogen and progesterone receptors in mammary glands: regulation by hormones, development and aging. *J. Steroid Biochem. Mol. Biol.*, **80**, 137–148.

70. Eddy,E.M., Washburn,T.F., Bunch,D.O., Goulding,E.H., Gladen,B.C., Lubahn,D.B. and Korach,K.S. (1996) Targeted disruption of the estrogen receptor gene in male mice causes alteration of spermatogenesis and infertility. *Endocrinology*, **137**, 4796–4805.

71. Bocchinfuso,W.P. and Korach,K.S. (1997) Mammary gland development and tumorigenesis in estrogen receptor knockout mice. *J. Mammary Gland Biol. Neoplasia*, **2**, 323–334.

72. Asselin-Labat,M.-L., Shackleton,M., Stingl,J., Vaillant,F., Forrest,N.C., Eaves,C.J., Visvader,J.E. and Lindeman,G.J. (2006) Steroid hormone receptor status of mouse mammary stem cells. *J. Natl. Cancer Inst.*, **98**, 1011–1014.

73. Krege,J.H., Hodgin,J.B., Couse,J.F., Enmark,E., Warner,M., Mahler,J.F., Sar,M., Korach,K.S., Gustafsson,J.A. and Smithies,O. (1998) Generation and reproductive phenotypes of mice lacking estrogen receptor beta. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 15677–15682.

74. Cunha,G.R., Young,P., Hom,Y.K., Cooke,P.S., Taylor,J.A. and Lubahn,D.B. (1997) Elucidation of a role for stromal steroid hormone receptors in mammary gland growth and development using tissue recombinants. *J. Mammary Gland Biol. Neoplasia*, **2**, 393–402.

75. Mueller,S.O., Clark,J.A., Myers,P.H. and Korach,K.S. (2002) Mammary gland development in adult mice requires epithelial and stromal estrogen receptor alpha. *Endocrinology*, **143**, 2357–2365.

76. Couse,J.F., Curtis,S.W., Washburn,T.F., Lindzey,J., Golding,T.S., Lubahn,D.B., Smithies,O. and Korach,K.S. (1995) Analysis of transcription and estrogen insensitivity in the female mouse after targeted disruption of the estrogen receptor gene. *Mol. Endocrinol. Baltim. Md*, **9**, 1441–1454.

77. Kos,M., Denger,S., Reid,G., Korach,K.S. and Gannon,F. (2002) Down but not out? A novel protein isoform of the estrogen receptor alpha is expressed in the estrogen receptor alpha knockout mouse. *J. Mol. Endocrinol.*, **29**, 281–286.

78. Dupont,S., Krust,A., Gansmuller,A., Dierich,A., Chambon,P. and Mark,M. (2000) Effect of single and compound knockouts of estrogen receptors alpha (ERalpha) and beta (ERbeta) on mouse reproductive phenotypes. *Dev. Camb. Engl.*, **127**, 4277–4291.

79. Mallepell,S., Krust,A., Chambon,P. and Brisken,C. (2006) Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 2196–2201.

80. Feng,Y., Manka,D., Wagner,K.-U. and Khan,S.A. (2007) Estrogen receptor-alpha expression in the mammary epithelium is required for ductal and alveolar morphogenesis in mice. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 14718–14723.

81. Brisken,C., Park,S., Vass,T., Lydon,J.P., O'Malley,B.W. and Weinberg,R.A. (1998) A paracrine role for the epithelial progesterone receptor in mammary gland development. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 5076–5081.

82. Lam,E.W.-F., Brosens,J.J., Gomes,A.R. and Koo,C.-Y. (2013) Forkhead box proteins: tuning forks for transcriptional harmony. *Nat. Rev. Cancer*, **13**, 482–495.

83. Weigel,D., Jürgens,G., Küttner,F., Seifert,E. and Jäckle,H. (1989) The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo. *Cell*, **57**, 645–658.

84. Costa,R.H., Grayson,D.R. and Darnell,J.E. (1989) Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and alpha 1-antitrypsin genes. *Mol. Cell. Biol.*, **9**, 1415–1425.

85. Clark,K.L., Halay,E.D., Lai,E. and Burley,S.K. (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*, **364**, 412–420.

86. Overdier,D.G., Porcella,A. and Costa,R.H. (1994) The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino-acid residues adjacent to the recognition helix. *Mol. Cell. Biol.*, **14**, 2755–2766.

87. Goytisolo,F.A., Gerchman,S.E., Yu,X., Rees,C., Graziano,V., Ramakrishnan,V. and Thomas,J.O. (1996) Identification of two DNA-binding sites on the globular domain of histone H5. *EMBO J.*, **15**, 3421–3429.

88. Cirillo,L.A., McPherson,C.E., Bossard,P., Stevens,K., Cherian,S., Shim,E.Y., Clark,K.L., Burley,S.K. and Zaret,K.S. (1998) Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO J.*, **17**, 244–254.

89. Cirillo,L.A., Lin,F.R., Cuesta,I., Friedman,D., Jarnik,M. and Zaret,K.S. (2002) Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell*, **9**, 279–289.

90. Shim,E.Y., Woodcock,C. and Zaret,K.S. (1998) Nucleosome positioning by the winged helix transcription factor HNF3. *Genes Dev.*, **12**, 5–10.

91. Lee,C.S., Friedman,J.R., Fulmer,J.T. and Kaestner,K.H. (2005) The initiation of liver development is dependent on Foxa transcription factors. *Nature*, **435**, 944–947.

92. Kaestner,K.H., Katz,J., Liu,Y., Drucker,D.J. and Schütz,G. (1999) Inactivation of the winged helix transcription factor HNF3alpha affects glucose homeostasis and islet glucagon gene expression in vivo. *Genes Dev.*, **13**, 495–504.

93. Ferri,A.L.M., Lin,W., Mavromatakis,Y.E., Wang,J.C., Sasaki,H., Whitsett,J.A. and Ang,S.-L. (2007) Foxa1 and Foxa2 regulate multiple phases of midbrain dopaminergic neuron development in a dosage-dependent manner. *Dev. Camb. Engl.*, **134**, 2761–2769.

94. Besnard,V., Wert,S.E., Kaestner,K.H. and Whitsett,J.A. (2005) Stage-specific regulation of respiratory epithelial cell differentiation by Foxa1. *Am. J. Physiol. Lung Cell. Mol. Physiol.*, **289**, L750-759.

95. Bernardo,G.M., Lozada,K.L., Miedler,J.D., Harburg,G., Hewitt,S.C., Mosley,J.D., Godwin,A.K., Korach,K.S., Visvader,J.E., Kaestner,K.H., *et al.* (2010) FOXA1 is an essential determinant of ERalpha expression and mammary ductal morphogenesis. *Dev. Camb. Engl.*, **137**, 2045–2054.

96. Lentjes,M.H.F.M., Niessen,H.E.C., Akiyama,Y., de Bruïne,A.P., Melotte,V. and van Engeland,M. (2016) The emerging role of GATA transcription factors in development and disease. *Expert Rev. Mol. Med.*, **18**, e3.

97. Linnemann,A.K., O'Geen,H., Keles,S., Farnham,P.J. and Bresnick,E.H. (2011) Genetic framework for GATA factor function in vascular biology. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 13641–13646.

98. Bresnick,E.H., Katsumura,K.R., Lee,H.-Y., Johnson,K.D. and Perkins,A.S. (2012) Master regulatory GATA transcription factors: mechanistic principles and emerging links to hematologic malignancies. *Nucleic Acids Res.*, **40**, 5819–5831.

99. Vakoc,C.R., Letting,D.L., Gheldof,N., Sawado,T., Bender,M.A., Groudine,M., Weiss,M.J., Dekker,J. and Blobel,G.A. (2005) Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol. Cell*, **17**, 453–462.

100. Kim,S.-I., Bultman,S.J., Kiefer,C.M., Dean,A. and Bresnick,E.H. (2009) BRG1 requirement for long-range interaction of a locus control region with a downstream promoter. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 2259–2264.

101. Pal,S., Cantor,A.B., Johnson,K.D., Moran,T.B., Boyer,M.E., Orkin,S.H. and Bresnick,E.H. (2004) Coregulator-dependent facilitation of chromatin occupancy by GATA-1. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 980–985.

102. Lamonica,J.M., Vakoc,C.R. and Blobel,G.A. (2006) Acetylation of GATA-1 is required for chromatin occupancy. *Blood*, **108**, 3736–3738.

103. Bossard,P. and Zaret,K.S. (1998) GATA transcription factors as potentiators of gut endoderm differentiation. *Dev. Camb. Engl.*, **125**, 4909–4917.

104. Takaku,M., Grimm,S.A., Shimbo,T., Perera,L., Menafra,R., Stunnenberg,H.G., Archer,T.K., Machida,S., Kurumizaka,H. and Wade,P.A. (2016) GATA3-dependent cellular reprogramming requires activation-domain dependent recruitment of a chromatin remodeler. *Genome Biol.*, **17**, 36.

105. Pandolfi,P.P., Roth,M.E., Karis,A., Leonard,M.W., Dzierzak,E., Grosveld,F.G., Engel,J.D. and Lindenbaum,M.H. (1995) Targeted disruption of the GATA3 gene causes severe abnormalities in the nervous system and in fetal liver haematopoiesis. *Nat. Genet.*, **11**, 40–44.

106. Lim,K.C., Lakshmanan,G., Crawford,S.E., Gu,Y., Grosveld,F. and Engel,J.D. (2000) Gata3 loss leads to embryonic lethality due to noradrenaline deficiency of the sympathetic nervous system. *Nat. Genet.*, **25**, 209–212.

107. Van Esch,H. and Devriendt,K. (2001) Transcription factor GATA3 and the human HDR syndrome. *Cell. Mol. Life Sci. CMLS*, **58**, 1296–1300.

108. Kouros-Mehr,H., Slorach,E.M., Sternlicht,M.D. and Werb,Z. (2006) GATA-3 Maintains the Differentiation of the Luminal Cell Fate in the Mammary Gland. *Cell*, **127**, 1041–1055.

109. Asselin-Labat,M.-L., Sutherland,K.D., Barker,H., Thomas,R., Shackleton,M., Forrest,N.C., Hartley,L., Robb,L., Grosveld,F.G., van der Wees,J., *et al.* (2007) Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nat. Cell Biol.*, **9**, 201–209.

110. Sizemore,G.M., Pitarresi,J.R., Balakrishnan,S. and Ostrowski,M.C. (2017) The ETS family of oncogenic transcription factors in solid tumours. *Nat. Rev. Cancer*, **17**, 337–351.

111. Oakes,S.R., Naylor,M.J., Asselin-Labat,M.-L., Blazek,K.D., Gardiner-Garden,M., Hilton,H.N., Kazlauskas,M., Pritchard,M.A., Chodosh,L.A., Pfeffer,P.L., *et al.* (2008) The Ets transcription factor Elf5 specifies mammary alveolar cell fate. *Genes Dev.*, **22**, 581–586.

112. Choi,Y.S., Chakrabarti,R., Escamilla-Hernandez,R. and Sinha,S. (2009) Elf5 conditional knockout mice reveal its role as a master regulator in mammary alveolar development: failure of Stat5 activation and functional differentiation in the absence of Elf5. *Dev. Biol.*, **329**, 227–241.

113. Chakrabarti,R., Wei,Y., Romano,R.-A., DeCoste,C., Kang,Y. and Sinha,S. (2012) Elf5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. *Stem Cells Dayt. Ohio*, **30**, 1496–1508.

114. Oettgen,P., Finger,E., Sun,Z., Akbarali,Y., Thamrongsak,U., Boltax,J., Grall,F., Dube,A., Weiss,A., Brown,L., *et al.* (2000) PDEF, a novel prostate epithelium-specific ets transcription factor, interacts with the androgen receptor and activates prostate-specific antigen gene expression. *J. Biol. Chem.*, **275**, 1216–1225.

115. Chen,G., Korfhagen,T.R., Xu,Y., Kitzmiller,J., Wert,S.E., Maeda,Y., Gregorieff,A., Clevers,H. and Whitsett,J.A. (2009) SPDEF is required for mouse pulmonary goblet cell differentiation and regulates a network of genes associated with mucus production. *J. Clin. Invest.*, **119**, 2914–2924.

116. Horst,D., Gu,X., Bhasin,M., Yang,Q., Verzi,M., Lin,D., Joseph,M., Zhang,X., Chen,W., Li,Y.-P., *et al.* (2010) Requirement of the epithelium-specific Ets transcription factor Spdef for mucous gland cell function in the gastric antrum. *J. Biol. Chem.*, **285**, 35047–35055.

117. Gregorieff,A., Stange,D.E., Kujala,P., Begthel,H., van den Born,M., Korving,J., Peters,P.J. and Clevers,H. (2009) The ets-domain transcription factor Spdef promotes maturation of goblet and paneth cells in the intestinal epithelium. *Gastroenterology*, **137**, 1333-1345–3.

118. Buchwalter,G., Hickey,M.M., Cromer,A., Selfors,L.M., Gunawardane,R.N., Frishman,J., Jeselsohn,R., Lim,E., Chi,D., Fu,X., *et al.* (2013) PDEF promotes luminal differentiation and acts as a survival factor for ER-positive breast cancer cells. *Cancer Cell*, **23**, 753–767.

119. Liou,H.C., Boothby,M.R., Finn,P.W., Davidon,R., Nabavi,N., Zeleznik-Le,N.J., Ting,J.P. and Glimcher,L.H. (1990) A new member of the leucine zipper class of proteins that binds to the HLA DR alpha promoter. *Science*, **247**, 1581–1584.

120. Yoshida,H., Okada,T., Haze,K., Yanagi,H., Yura,T., Negishi,M. and Mori,K. (2000) ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Mol. Cell. Biol.*, **20**, 6755–6767.

121. Yoshida,H., Matsui,T., Yamamoto,A., Okada,T. and Mori,K. (2001) XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell*, **107**, 881–891.

122. Lee,A.-H., Chu,G.C., Iwakoshi,N.N. and Glimcher,L.H. (2005) XBP-1 is required for biogenesis of cellular secretory machinery of exocrine glands. *EMBO J.*, **24**, 4368–4380.

123. Reimold,A.M., Etkin,A., Clauss,I., Perkins,A., Friend,D.S., Zhang,J., Horton,H.F., Scott,A., Orkin,S.H., Byrne,M.C., *et al.* (2000) An essential role in liver development for transcription factor XBP-1. *Genes Dev.*, **14**, 152–157.

124. Reimold,A.M., Iwakoshi,N.N., Manis,J., Vallabhajosyula,P., Szomolanyi-Tsuda,E., Gravallese,E.M., Friend,D., Grusby,M.J., Alt,F. and Glimcher,L.H. (2001) Plasma cell differentiation requires the transcription factor XBP-1. *Nature*, **412**, 300–307.

125. Hasegawa,D., Calvo,V., Avivar-Valderas,A., Lade,A., Chou,H.-I., Lee,Y.A., Farias,E.F., Aguirre-Ghiso,J.A. and Friedman,S.L. (2015) Epithelial Xbp1 Is Required for Cellular Proliferation and Differentiation during Mammary Gland Development. *Mol. Cell. Biol.*, **35**, 1543–1556.

126. Davis,K.R., Giesy,S.L., Long,Q., Krumm,C.S., Harvatine,K.J. and Boisclair,Y.R. (2016) XBP1 Regulates the Biosynthetic Capacity of the Mammary Gland During Lactation by

Controlling Epithelial Expansion and Endoplasmic Reticulum Formation. *Endocrinology*, **157**, 417–428.

127. Notini,A.J., Davey,R.A., McManus,J.F., Bate,K.L. and Zajac,J.D. (2005) Genomic actions of the androgen receptor are required for normal male sexual differentiation in a mouse model. *J. Mol. Endocrinol.*, **35**, 547–555.

128. Riegman,P.H., Vlietstra,R.J., van der Korput,J.A., Brinkmann,A.O. and Trapman,J. (1991) The promoter of the prostate-specific antigen gene contains a functional androgen responsive element. *Mol. Endocrinol. Baltim. Md*, **5**, 1921–1930.

129. Tan,M.H.E., Li,J., Xu,H.E., Melcher,K. and Yong,E. (2015) Androgen receptor: structure, role in prostate cancer and drug discovery. *Acta Pharmacol. Sin.*, **36**, 3–23.

130. Tomlins,S.A., Rhodes,D.R., Perner,S., Dhanasekaran,S.M., Mehra,R., Sun,X.-W., Varambally,S., Cao,X., Tchinda,J., Kuefer,R., *et al.* (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*, **310**, 644–648.

131. Gottlieb,B., Beitel,L.K., Nadarajah,A., Paliouras,M. and Trifiro,M. (2012) The androgen receptor gene mutations database: 2012 update. *Hum. Mutat.*, **33**, 887–894.

132. Leung,J.K. and Sadar,M.D. (2017) Non-Genomic Actions of the Androgen Receptor in Prostate Cancer. *Front. Endocrinol.*, **8**, 2.

133. Li,S., Han,B., Liu,G., Li,S., Ouellet,J., Labrie,F. and Pelletier,G. (2010) Immunocytochemical localization of sex steroid hormone receptors in normal human mammary gland. *J. Histochem. Cytochem. Off. J. Histochem. Soc.*, **58**, 509–515.

134. Yeh,S., Hu,Y.-C., Wang,P.-H., Xie,C., Xu,Q., Tsai,M.-Y., Dong,Z., Wang,R.-S., Lee,T.-H. and Chang,C. (2003) Abnormal mammary gland development and growth retardation in female mice and MCF7 breast cancer cells lacking androgen receptor. *J. Exp. Med.*, **198**, 1899–1908.

135. Canadian Cancer Society's Advisory Committee on Cancer Statistics. Canadian Cancer Statistics 2017. Toronto, ON: Canadian Cancer Society; 2017. Available at: cancer.ca/Canadian-CancerStatistics-2017-EN.pdf.

136. De,P., Neutel,C.I., Olivotto,I. and Morrison,H. (2010) Breast cancer incidence and hormone replacement therapy in Canada. *J. Natl. Cancer Inst.*, **102**, 1489–1495.

137. Neutel,C.I. and Morrison,H. (2010) Could recent decreases in breast cancer incidence really be due to lower HRT use? Trends in attributable risk for modifiable breast cancer risk factors in Canadian women. *Can. J. Public Health Rev. Can. Sante Publique*, **101**, 405–409.

138. Wesolowski,R. and Ramaswamy,B. (2011) Gene expression profiling: changing face of breast cancer classification and management. *Gene Expr.*, **15**, 105–115.

139. Perou,C.M., Sørlie,T., Eisen,M.B., van de Rijn,M., Jeffrey,S.S., Rees,C.A., Pollack,J.R., Ross,D.T., Johnsen,H., Akslen,L.A., *et al.* (2000) Molecular portraits of human breast tumours. *Nature*, **406**, 747–752.

140. Sørlie,T., Perou,C.M., Tibshirani,R., Aas,T., Geisler,S., Johnsen,H., Hastie,T., Eisen,M.B., van de Rijn,M., Jeffrey,S.S., *et al.* (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 10869–10874.

141. Sorlie,T., Tibshirani,R., Parker,J., Hastie,T., Marron,J.S., Nobel,A., Deng,S., Johnsen,H., Pesich,R., Geisler,S., *et al.* (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 8418–8423.

142. Charafe-Jauffret,E., Ginestier,C., Monville,F., Finetti,P., Adélaïde,J., Cervera,N., Fekairi,S., Xerri,L., Jacquemier,J., Birnbaum,D., *et al.* (2006) Gene expression profiling of breast cell lines identifies potential new basal markers. *Oncogene*, **25**, 2273–2284.

143. Parker,J.S., Mullins,M., Cheang,M.C.U., Leung,S., Voduc,D., Vickery,T., Davies,S., Fauron,C., He,X., Hu,Z., *et al.* (2009) Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, **27**, 1160–1167.

144. Herschkowitz,J.I., Simin,K., Weigman,V.J., Mikaelian,I., Usary,J., Hu,Z., Rasmussen,K.E., Jones,L.P., Assefnia,S., Chandrasekharan,S., *et al.* (2007) Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol.*, **8**, R76.

145. Farmer,P., Bonnefoi,H., Becette,V., Tubiana-Hulin,M., Fumoleau,P., Larsimont,D., Macgrogan,G., Bergh,J., Cameron,D., Goldstein,D., *et al.* (2005) Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene*, **24**, 4660–4671.

146. Curtis,C., Shah,S.P., Chin,S.-F., Turashvili,G., Rueda,O.M., Dunning,M.J., Speed,D., Lynch,A.G., Samarajiwa,S., Yuan,Y., *et al.* (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, **486**, 346–352.

147. Network,T.C.G.A. (2012) Comprehensive molecular portraits of human breast tumours. *Nature*, **490**, 61–70.

148. Ciriello,G., Gatza,M.L., Beck,A.H., Wilkerson,M.D., Rhie,S.K., Pastore,A., Zhang,H., McLellan,M., Yau,C., Kandoth,C., *et al.* (2015) Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell*, **163**, 506–519.

149. Visvader,J.E. and Stingl,J. (2014) Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev.*, **28**, 1143–1158.

150. Guedj,M., Marisa,L., de Reynies,A., Orsetti,B., Schiappa,R., Bibeau,F., MacGrogan,G., Lerebours,F., Finetti,P., Longy,M., *et al.* (2012) A refined molecular taxonomy of breast cancer. *Oncogene*, **31**, 1196–1206.

151. Beatson,G. (1896) ON THE TREATMENT OF INOPERABLE CASES OF CARCINOMA OF THE MAMMA: SUGGESTIONS FOR A NEW METHOD OF TREATMENT, WITH ILLUSTRATIVE CASES. *The Lancet*, **148**, 162–165.

152. JENSEN,E.V. and JACOBSON,H.I. (1960) Fate of Steroid Estrogens in Target Tissues. In *Biological Activities of Steroids in Relation to Cancer*. Academic Press, pp. 161–178.

153. Kamińska,M., Ciszewski,T., Łopacka-Szatan,K., Miotła,P. and Starosławska,E. (2015) Breast cancer risk factors. *Przeglad Menopauzalny Menopause Rev.*, **14**, 196–202.

154. García-Becerra,R., Santos,N., Díaz,L. and Camacho,J. (2012) Mechanisms of resistance to endocrine therapy in breast cancer: focus on signaling pathways, miRNAs and genetically based resistance. *Int. J. Mol. Sci.*, **14**, 108–145.

155. Ring,A. and Dowsett,M. (2004) Mechanisms of tamoxifen resistance. *Endocr. Relat. Cancer*, **11**, 643–658.

156. Viedma-Rodríguez,R., Baiza-Gutman,L., Salamanca-Gómez,F., Diaz-Zaragoza,M., Martínez-Hernández,G., Ruiz Esparza-Garrido,R., Velázquez-Flores,M.A. and Arenas-Aranda,D. (2014) Mechanisms associated with resistance to tamoxifen in estrogen receptor-positive breast cancer (review). *Oncol. Rep.*, **32**, 3–15.

157. Jeselsohn,R., Buchwalter,G., De Angelis,C., Brown,M. and Schiff,R. (2015) ESR1 mutations—a mechanism for acquired endocrine resistance in breast cancer. *Nat. Rev. Clin. Oncol.*, **12**, 573–583.

158. Hilmi,K., Hussein,N., Mendoza-Sanchez,R., El-Ezzy,M., Ismail,H., Durette,C., Bail,M., Rozendaal,M.J., Bouvier,M., Thibault,P., *et al.* (2012) Role of SUMOylation in full antiestrogenicity. *Mol. Cell. Biol.*, **32**, 3823–3837.

159. Fabian,C.J. (2007) The what, why and how of aromatase inhibitors: hormonal agents for treatment and prevention of breast cancer. *Int. J. Clin. Pract.*, **61**, 2051–2063.

160. Cerami,E., Gao,J., Dogrusoz,U., Gross,B.E., Sumer,S.O., Aksoy,B.A., Jacobsen,A., Byrne,C.J., Heuer,M.L., Larsson,E., *et al.* (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.*, **2**, 401–404.

161. Gao,J., Aksoy,B.A., Dogrusoz,U., Dresdner,G., Gross,B., Sumer,S.O., Sun,Y., Jacobsen,A., Sinha,R., Larsson,E., *et al.* (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.*, **6**, pl1.

162. Toy,W., Shen,Y., Won,H., Green,B., Sakr,R.A., Will,M., Li,Z., Gala,K., Fanning,S., King,T.A., *et al.* (2013) ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat. Genet.*, **45**, 1439–1445.

163. Robinson,D.R., Wu,Y.-M., Vats,P., Su,F., Lonigro,R.J., Cao,X., Kalyana-Sundaram,S., Wang,R., Ning,Y., Hodges,L., *et al.* (2013) Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat. Genet.*, **45**, 1446–1451.

164. Li,S., Shen,D., Shao,J., Crowder,R., Liu,W., Prat,A., He,X., Liu,S., Hoog,J., Lu,C., *et al.* (2013) Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. *Cell Rep.*, **4**, 1116–1130.

165. Jeselsohn,R., Yelensky,R., Buchwalter,G., Frampton,G., Meric-Bernstam,F., Gonzalez-Angulo,A.M., Ferrer-Lozano,J., Perez-Fidalgo,J.A., Cristofanilli,M., Gómez,H., *et al.* (2014) Emergence of constitutively active estrogen receptor-α mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, **20**, 1757–1767.

166. Merenbakh-Lamin,K., Ben-Baruch,N., Yeheskel,A., Dvir,A., Soussan-Gutman,L., Jeselsohn,R., Yelensky,R., Brown,M., Miller,V.A., Sarid,D., *et al.* (2013) D538G mutation in estrogen receptor-α: A novel mechanism for acquired endocrine resistance in breast cancer. *Cancer Res.*, **73**, 6856–6864.

167. Hurtado,A., Holmes,K.A., Ross-Innes,C.S., Schmidt,D. and Carroll,J.S. (2011) FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat. Genet.*, **43**, 27–33.

168. Ross-Innes,C.S., Stark,R., Teschendorff,A.E., Holmes,K.A., Ali,H.R., Dunning,M.J., Brown,G.D., Gojis,O., Ellis,I.O., Green,A.R., *et al.* (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*, **481**, 389–393.

169. Fu,X., Jeselsohn,R., Pereira,R., Hollingsworth,E.F., Creighton,C.J., Li,F., Shea,M., Nardone,A., De Angelis,C., Heiser,L.M., *et al.* (2016) FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8 expression in ER-positive breast cancer. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, E6600–E6609.

170. Rheinbay,E., Parasuraman,P., Grimsby,J., Tiao,G., Engreitz,J.M., Kim,J., Lawrence,M.S., Taylor-Weiner,A., Rodriguez-Cuevas,S., Rosenberg,M., *et al.* (2017) Recurrent and functional regulatory mutations in breast cancer. *Nature*, **547**, 55–60.

171. Usary,J., Llaca,V., Karaca,G., Presswala,S., Karaca,M., He,X., Langerød,A., Kåresen,R., Oh,D.S., Dressler,L.G., *et al.* (2004) Mutation of GATA3 in human breast tumors. *Oncogene*, **23**, 7669–7678.

172. Smith,V.M., Lee,P.P., Szychowski,S. and Winoto,A. (1995) GATA-3 dominant negative mutant. Functional redundancy of the T cell receptor alpha and beta enhancers. *J. Biol. Chem.*, **270**, 1515–1520.

173. Bates,D.L., Chen,Y., Kim,G., Guo,L. and Chen,L. (2008) Crystal structures of multiple GATA zinc fingers bound to DNA reveal new insights into DNA recognition and self-association by GATA. *J. Mol. Biol.*, **381**, 1292–1306.

174. Gaynor,K.U., Grigorieva,I.V., Allen,M.D., Esapa,C.T., Head,R.A., Gopinath,P., Christie,P.T., Nesbit,M.A., Jones,J.L. and Thakker,R.V. (2013) GATA3 mutations found in breast cancers may be associated with aberrant nuclear localization, reduced transactivation and cell invasiveness. *Horm. Cancer*, **4**, 123–139.

175. Adomas,A.B., Grimm,S.A., Malone,C., Takaku,M., Sims,J.K. and Wade,P.A. (2014) Breast tumor specific mutation in GATA3 affects physiological mechanisms regulating transcription factor turnover. *BMC Cancer*, **14**, 278.

176. Shinnakasu,R., Yamashita,M., Shinoda,K., Endo,Y., Hosokawa,H., Hasegawa,A., Ikemizu,S. and Nakayama,T. (2006) Critical YxKxHxxxRP motif in the C-terminal region of GATA3 for its DNA binding and function. *J. Immunol. Baltim. Md 1950*, **177**, 5801–5810.

177. Crossley,M., Merika,M. and Orkin,S.H. (1995) Self-association of the erythroid transcription factor GATA-1 mediated by its zinc finger domains. *Mol. Cell. Biol.*, **15**, 2448–2456.

178. Jiang,Y.-Z., Yu,K.-D., Zuo,W.-J., Peng,W.-T. and Shao,Z.-M. (2014) GATA3 mutations define a unique subtype of luminal-like breast cancer with improved survival. *Cancer*, **120**, 1329–1337.

179. Ellis,M.J., Ding,L., Shen,D., Luo,J., Suman,V.J., Wallis,J.W., Van Tine,B.A., Hoog,J., Goiffon,R.J., Goldstein,T.C., *et al.* (2012) Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature*, **486**, 353–360.

180. Eroles,P., Bosch,A., Pérez-Fidalgo,J.A. and Lluch,A. (2012) Molecular biology in breast cancer: intrinsic subtypes and signaling pathways. *Cancer Treat. Rev.*, **38**, 698–707.

181. Slamon,D.J., Clark,G.M., Wong,S.G., Levin,W.J., Ullrich,A. and McGuire,W.L. (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*, **235**, 177–182.

182. Borgquist,S., Zhou,W., Jirström,K., Amini,R.-M., Sollie,T., Sørlie,T., Blomqvist,C., Butt,S. and Wärnberg,F. (2015) The prognostic role of HER2 expression in ductal breast carcinoma in situ (DCIS); a population-based cohort study. *BMC Cancer*, **15**, 468.

183. Carter,P., Presta,L., Gorman,C.M., Ridgway,J.B., Henner,D., Wong,W.L., Rowland,A.M., Kotts,C., Carver,M.E. and Shepard,H.M. (1992) Humanization of an anti-p185HER2 antibody for human cancer therapy. *Proc. Natl. Acad. Sci. U. S. A.*, **89**, 4285–4289.

184. Klapper,L.N., Waterman,H., Sela,M. and Yarden,Y. (2000) Tumor-inhibitory antibodies to HER-2/ErbB-2 may act by recruiting c-Cbl and enhancing ubiquitination of HER-2. *Cancer Res.*, **60**, 3384–3388.

185. Clynes,R.A., Towers,T.L., Presta,L.G. and Ravetch,J.V. (2000) Inhibitory Fc receptors modulate in vivo cytotoxicity against tumor targets. *Nat. Med.*, **6**, 443–446.

186. Junttila,T.T., Akita,R.W., Parsons,K., Fields,C., Lewis Phillips,G.D., Friedman,L.S., Sampath,D. and Sliwkowski,M.X. (2009) Ligand-independent HER2/HER3/PI3K complex is disrupted by trastuzumab and is effectively inhibited by the PI3K inhibitor GDC-0941. *Cancer Cell*, **15**, 429–440.

187. Foulkes,W.D., Stefansson,I.M., Chappuis,P.O., Bégin,L.R., Goffin,J.R., Wong,N., Trudel,M. and Akslen,L.A. (2003) Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. *J. Natl. Cancer Inst.*, **95**, 1482–1485.

188. Proia,T.A., Keller,P.J., Gupta,P.B., Klebba,I., Jones,A.D., Sedic,M., Gilmore,H., Tung,N., Naber,S.P., Schnitt,S., *et al.* (2011) Genetic predisposition directs breast cancer phenotype by dictating progenitor cell fate. *Cell Stem Cell*, **8**, 149–163.

189. Carey,L.A., Perou,C.M., Livasy,C.A., Dressler,L.G., Cowan,D., Conway,K., Karaca,G., Troester,M.A., Tse,C.K., Edmiston,S., *et al.* (2006) Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA*, **295**, 2492–2502.

190. Rakha,E.A. and Ellis,I.O. (2009) Triple-negative/basal-like breast cancer: review. *Pathology (Phila.)*, **41**, 40–47.

191. Rouzier,R., Perou,C.M., Symmans,W.F., Ibrahim,N., Cristofanilli,M., Anderson,K., Hess,K.R., Stec,J., Ayers,M., Wagner,P., *et al.* (2005) Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, **11**, 5678–5685.

192. Shakya,R., Szabolcs,M., McCarthy,E., Ospina,E., Basso,K., Nandula,S., Murty,V., Baer,R. and Ludwig,T. (2008) The basal-like mammary carcinomas induced by Brca1 or Bard1 inactivation implicate the BRCA1/BARD1 heterodimer in tumor suppression. *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 7040–7045.

193. Xu,X., Wagner,K.U., Larson,D., Weaver,Z., Li,C., Ried,T., Hennighausen,L., Wynshaw-Boris,A. and Deng,C.X. (1999) Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. *Nat. Genet.*, **22**, 37–43.

194. Lim,E., Vaillant,F., Wu,D., Forrest,N.C., Pal,B., Hart,A.H., Asselin-Labat,M.-L., Gyorki,D.E., Ward,T., Partanen,A., *et al.* (2009) Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat. Med.*, **15**, 907–913.

195. Molyneux,G., Geyer,F.C., Magnay,F.-A., McCarthy,A., Kendrick,H., Natrajan,R., Mackay,A., Grigoriadis,A., Tutt,A., Ashworth,A., *et al.* (2010) BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell*, **7**, 403–417.

196. Prat,A., Parker,J.S., Karginova,O., Fan,C., Livasy,C., Herschkowitz,J.I., He,X. and Perou,C.M. (2010) Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res. BCR*, **12**, R68.

197. Mani,S.A., Guo,W., Liao,M.-J., Eaton,E.N., Ayyanan,A., Zhou,A.Y., Brooks,M., Reinhard,F., Zhang,C.C., Shipitsin,M., *et al.* (2008) The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*, **133**, 704–715.

198. Creighton,C.J., Li,X., Landis,M., Dixon,J.M., Neumeister,V.M., Sjolund,A., Rimm,D.L., Wong,H., Rodriguez,A., Herschkowitz,J.I., *et al.* (2009) Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 13820–13825.

199. Lehmann-Che,J., Hamy,A.-S., Porcher,R., Barritault,M., Bouhidel,F., Habuellelah,H., Leman-Detours,S., de Roquancourt,A., Cahen-Doidy,L., Bourstyn,E., *et al.* (2013)

Molecular apocrine breast cancers are aggressive estrogen receptor negative tumors overexpressing either HER2 or GCDFP15. *Breast Cancer Res. BCR*, **15**, R37.

200. Baniwal,S.K., Little,G.H., Chimge,N.-O. and Frenkel,B. (2012) Runx2 controls a feed-forward loop between androgen and prolactin-induced protein (PIP) in stimulating T47D cell proliferation. *J. Cell. Physiol.*, **227**, 2276–2282.

201. Doane,A.S., Danso,M., Lal,P., Donaton,M., Zhang,L., Hudis,C. and Gerald,W.L. (2006) An estrogen receptor-negative breast cancer subset characterized by a hormonally regulated transcriptional program and response to androgen. *Oncogene*, **25**, 3994–4008.

202. Robinson,J.L.L., Macarthur,S., Ross-Innes,C.S., Tilley,W.D., Neal,D.E., Mills,I.G. and Carroll,J.S. (2011) Androgen receptor driven transcription in molecular apocrine breast cancer is mediated by FoxA1. *EMBO J.*, **30**, 3019–3027.

203. Sahu,B., Laakso,M., Ovaska,K., Mirtti,T., Lundin,J., Rannikko,A., Sankila,A., Turunen,J.-P., Lundin,M., Konsti,J., *et al.* (2011) Dual role of FoxA1 in androgen receptor binding to chromatin, androgen signalling and prostate cancer. *EMBO J.*, **30**, 3962–3976.

204. Badve,S., Turbin,D., Thorat,M.A., Morimiya,A., Nielsen,T.O., Perou,C.M., Dunn,S., Huntsman,D.G. and Nakshatri,H. (2007) FOXA1 expression in breast cancer--correlation with luminal subtype A and survival. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, **13**, 4415–4421.

205. Thorat,M.A., Marchio,C., Morimiya,A., Savage,K., Nakshatri,H., Reis-Filho,J.S. and Badve,S. (2008) Forkhead box A1 expression in breast cancer is associated with luminal subtype and good prognosis. *J. Clin. Pathol.*, **61**, 327–332.

206. Mehta,R.J., Jain,R.K., Leung,S., Choo,J., Nielsen,T., Huntsman,D., Nakshatri,H. and Badve,S. (2012) FOXA1 is an independent prognostic marker for ER-positive breast cancer. *Breast Cancer Res. Treat.*, **131**, 881–890.

207. Yoon,N.K., Maresh,E.L., Shen,D., Elshimali,Y., Apple,S., Horvath,S., Mah,V., Bose,S., Chia,D., Chang,H.R., *et al.* (2010) Higher levels of GATA3 predict better survival in women with breast cancer. *Hum. Pathol.*, **41**, 1794–1801.

208. Kouros-Mehr,H., Bechis,S.K., Slorach,E.M., Littlepage,L.E., Egeblad,M., Ewald,A.J., Pai,S.-Y., Ho,I.-C. and Werb,Z. (2008) GATA-3 links tumor differentiation and dissemination in a luminal breast cancer model. *Cancer Cell*, **13**, 141–152.

209. Asselin-Labat,M.-L., Sutherland,K.D., Vaillant,F., Gyorki,D.E., Wu,D., Holroyd,S., Breslin,K., Ward,T., Shi,W., Bath,M.L., *et al.* (2011) Gata-3 negatively regulates the tumor-initiating capacity of mammary luminal progenitor cells and targets the putative tumor suppressor caspase-14. *Mol. Cell. Biol.*, **31**, 4609–4622.

210. Chou,J., Lin,J.H., Brenot,A., Kim,J., Provot,S. and Werb,Z. (2013) GATA3 suppresses metastasis and modulates the tumour microenvironment by regulating microRNA-29b expression. *Nat. Cell Biol.*, **15**, 201–213.

211. McCune,K., Mehta,R., Thorat,M.A., Badve,S. and Nakshatri,H. (2010) Loss of ERα and FOXA1 expression in a progression model of luminal type breast cancer: insights from PyMT transgenic mouse model. *Oncol. Rep.*, **24**, 1233–1239.

212. Bernardo,G.M., Bebek,G., Ginther,C.L., Sizemore,S.T., Lozada,K.L., Miedler,J.D., Anderson,L.A., Godwin,A.K., Abdul-Karim,F.W., Slamon,D.J., *et al.* (2013) FOXA1 represses the molecular phenotype of basal breast cancer cells. *Oncogene*, **32**, 554–563.

213. Xu,Y., Qin,L., Sun,T., Wu,H., He,T., Yang,Z., Mo,Q., Liao,L. and Xu,J. (2017) Twist1 promotes breast cancer invasion and metastasis by silencing Foxa1 expression. *Oncogene*, **36**, 1157–1166.

214. Eeckhoute,J., Keeton,E.K., Lupien,M., Krum,S.A., Carroll,J.S. and Brown,M. (2007) Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. *Cancer Res.*, **67**, 6477–6483.

215. Gaughan,L., Stockley,J., Coffey,K., O'Neill,D., Jones,D.L., Wade,M., Wright,J., Moore,M., Tse,S., Rogerson,L., *et al.* (2013) KDM4B is a master regulator of the estrogen receptor signalling cascade. *Nucleic Acids Res.*, **41**, 6892–6904.

216. Caizzi,L., Ferrero,G., Cutrupi,S., Cordero,F., Ballaré,C., Miano,V., Reineri,S., Ricci,L., Friard,O., Testori,A., *et al.* (2014) Genome-wide activity of unliganded estrogen receptor-α in breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 4892–4897.

217. Berthois,Y., Katzenellenbogen,J.A. and Katzenellenbogen,B.S. (1986) Phenol red in tissue culture media is a weak estrogen: implications concerning the study of estrogen-responsive cells in culture. *Proc. Natl. Acad. Sci. U. S. A.*, **83**, 2496–2500.

218. Wijayaratne,A.L. and McDonnell,D.P. (2001) The human estrogen receptor-alpha is a ubiquitinated protein whose stability is affected differentially by agonists, antagonists, and selective estrogen receptor modulators. *J. Biol. Chem.*, **276**, 35684–35692.

219. Laganière,J., Deblois,G., Lefebvre,C., Bataille,A.R., Robert,F. and Giguère,V. (2005) From the Cover: Location analysis of estrogen receptor alpha target promoters reveals that FOXA1 defines a domain of the estrogen response. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 11651–11656.

220. Wolf,I., Bose,S., Williamson,E.A., Miller,C.W., Karlan,B.Y. and Koeffler,H.P. (2007) FOXA1: Growth inhibitor and a favorable prognostic factor in human breast cancer. *Int. J. Cancer*, **120**, 1013–1022.

221. Frasor,J., Danes,J.M., Komm,B., Chang,K.C.N., Lyttle,C.R. and Katzenellenbogen,B.S. (2003) Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype. *Endocrinology*, **144**, 4562–4574.

222. Oh,D.S., Troester,M.A., Usary,J., Hu,Z., He,X., Fan,C., Wu,J., Carey,L.A. and Perou,C.M. (2006) Estrogen-regulated genes predict survival in hormone receptor-positive breast cancers. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, **24**, 1656–1664.

223. Sengupta,S., Sharma,C.G.N. and Jordan,V.C. (2010) Estrogen regulation of X-box binding protein-1 and its role in estrogen induced growth of breast and endometrial cancer cells. *Horm. Mol. Biol. Clin. Investig.*, **2**, 235–243.

224. Perez-Stable,C.M., Pozas,A. and Roos,B.A. (2000) A role for GATA transcription factors in the androgen regulation of the prostate-specific antigen gene enhancer. *Mol. Cell. Endocrinol.*, **167**, 43–53.

225. Al-azzeh,E.D., Fegert,P., Blin,N. and Gött,P. (2000) Transcription factor GATA-6 activates expression of gastroprotective trefoil genes TFF1 and TFF2. *Biochim. Biophys. Acta*, **1490**, 324–332.

226. Robyr,D., Gegonne,A., Wolffe,A.P. and Wahli,W. (2000) Determinants of vitellogenin B1 promoter architecture. HNF3 and estrogen responsive transcription within chromatin. *J. Biol. Chem.*, **275**, 28291–28300.

227. Gibson,G.G., el-Sankary,W. and Plant,N.J. (2002) Receptor-dependent regulation of the CYP3A4 gene. *Toxicology*, **181**–**182**, 199–202.

228. Beck,S., Sommer,P., dos Santos Silva,E., Blin,N. and Gött,P. (1999) Hepatocyte nuclear factor 3 (winged helix domain) activates trefoil factor gene TFF1 through a binding motif adjacent to the TATAA box. *DNA Cell Biol.*, **18**, 157–164.

229. Lacroix,M. and Leclercq,G. (2004) About GATA3, HNF3A, and XBP1, three genes co-expressed with the oestrogen receptor-alpha gene (ESR1) in breast cancer. *Mol. Cell. Endocrinol.*, **219**, 1–7.

230. Berry,M., Nunez,A.M. and Chambon,P. (1989) Estrogen-responsive element of the human pS2 gene is an imperfectly palindromic sequence. *Proc. Natl. Acad. Sci. U. S. A.*, **86**, 1218–1222.

231. Carroll,J.S., Liu,X.S., Brodsky,A.S., Li,W., Meyer,C.A., Szary,A.J., Eeckhoute,J., Shao,W., Hestermann,E.V., Geistlinger,T.R., *et al.* (2005) Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell*, **122**, 33–43.

232. Eeckhoute,J., Carroll,J.S., Geistlinger,T.R., Torres-Arzayus,M.I. and Brown,M. (2006) A cell-type-specific transcriptional network required for estrogen regulation of cyclin D1 and cell cycle progression in breast cancer. *Genes Dev.*, **20**, 2513–2526.

233. Deschênes,J., Bourdeau,V., White,J.H. and Mader,S. (2007) Regulation of GREB1 transcription by estrogen receptor alpha through a multipartite enhancer spread over 20 kb of upstream flanking sequences. *J. Biol. Chem.*, **282**, 17335–17339.

234. Fullwood,M.J., Liu,M.H., Pan,Y.F., Liu,J., Xu,H., Mohamed,Y.B., Orlov,Y.L., Velkov,S., Ho,A., Mei,P.H., *et al.* (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature*, **462**, 58–64.

235. Carroll,J.S., Meyer,C.A., Song,J., Li,W., Geistlinger,T.R., Eeckhoute,J., Brodsky,A.S., Keeton,E.K., Fertuck,K.C., Hall,G.F., *et al.* (2006) Genome-wide analysis of estrogen receptor binding sites. *Nat. Genet.*, **38**, 1289–1297.

236. Lin,C.-Y., Vega,V.B., Thomsen,J.S., Zhang,T., Kong,S.L., Xie,M., Chiu,K.P., Lipovich,L., Barnett,D.H., Stossi,F., *et al.* (2007) Whole-genome cartography of estrogen receptor alpha binding sites. *PLoS Genet.*, **3**, e87.

237. Lupien,M., Eeckhoute,J., Meyer,C.A., Wang,Q., Zhang,Y., Li,W., Carroll,J.S., Liu,X.S. and Brown,M. (2008) FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell*, **132**, 958–970.

238. Sérandour,A.A., Avner,S., Percevault,F., Demay,F., Bizot,M., Lucchetti-Miganeh,C., Barloy-Hubler,F., Brown,M., Lupien,M., Métivier,R., *et al.* (2011) Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. *Genome Res.*, **21**, 555–565.

239. Giresi,P.G., Kim,J., McDaniell,R.M., Iyer,V.R. and Lieb,J.D. (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res.*, **17**, 877–885.

240. Eeckhoute,J., Lupien,M., Meyer,C.A., Verzi,M.P., Shivdasani,R.A., Liu,X.S. and Brown,M. (2009) Cell-type selective chromatin remodeling defines the active subset of FOXA1-bound enhancers. *Genome Res.*, **19**, 372–380.

241. Zwart,W., Theodorou,V., Kok,M., Canisius,S., Linn,S. and Carroll,J.S. (2011) Oestrogen receptor-co-factor-chromatin specificity in the transcriptional regulation of breast cancer. *EMBO J.*, **30**, 4764–4776.

242. Liu,Z., Merkurjev,D., Yang,F., Li,W., Oh,S., Friedman,M.J., Song,X., Zhang,F., Ma,Q., Ohgi,K.A., *et al.* (2014) Enhancer activation requires trans-recruitment of a mega transcription factor complex. *Cell*, **159**, 358–373.

243. Li,W., Notani,D., Ma,Q., Tanasa,B., Nunez,E., Chen,A.Y., Merkurjev,D., Zhang,J., Ohgi,K., Song,X., *et al.* (2013) Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature*, **498**, 516–520.

244. Theodorou,V., Stark,R., Menon,S. and Carroll,J.S. (2013) GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Res.*, **23**, 12–22.

245. Kong,S.L., Li,G., Loh,S.L., Sung,W.-K. and Liu,E.T. (2011) Cellular reprogramming by the conjoint action of ERα, FOXA1, and GATA3 to a ligand-inducible growth state. *Mol. Syst. Biol.*, **7**, 526.

246. Ray,P.S., Wang,J., Qu,Y., Sim,M.-S., Shamonki,J., Bagaria,S.P., Ye,X., Liu,B., Elashoff,D., Hoon,D.S., *et al.* (2010) FOXC1 is a potential prognostic biomarker with functional significance in basal-like breast cancer. *Cancer Res.*, **70**, 3870–3876.

247. Yu-Rice,Y., Jin,Y., Han,B., Qu,Y., Johnson,J., Watanabe,T., Cheng,L., Deng,N., Tanaka,H., Gao,B., *et al.* (2016) FOXC1 is involved in ERα silencing by counteracting GATA3 binding and is implicated in endocrine resistance. *Oncogene*, **35**, 5400–5411.

248. Guo,S. and Sonenshein,G.E. (2004) Forkhead box transcription factor FOXO3a regulates estrogen receptor alpha expression and is repressed by the Her-2/neu/phosphatidylinositol 3-kinase/Akt signaling pathway. *Mol. Cell. Biol.*, **24**, 8681–8690.

249. Brunet,A., Bonni,A., Zigmond,M.J., Lin,M.Z., Juo,P., Hu,L.S., Anderson,M.J., Arden,K.C., Blenis,J. and Greenberg,M.E. (1999) Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell*, **96**, 857–868.

250. Oh,A.S., Lorant,L.A., Holloway,J.N., Miller,D.L., Kern,F.G. and El-Ashry,D. (2001) Hyperactivation of MAPK induces loss of ERalpha expression in breast cancer cells. *Mol. Endocrinol. Baltim. Md*, **15**, 1344–1359.

251. Madureira,P.A., Varshochi,R., Constantinidou,D., Francis,R.E., Coombes,R.C., Yao,K.-M. and Lam,E.W.-F. (2006) The Forkhead box M1 protein regulates the transcription of the estrogen receptor alpha in breast cancer cells. *J. Biol. Chem.*, **281**, 25167–25176.

252. Millour,J., Constantinidou,D., Stavropoulou,A.V., Wilson,M.S.C., Myatt,S.S., Kwok,J.M.-M., Sivanandan,K., Coombes,R.C., Medema,R.H., Hartman,J., *et al.* (2010) FOXM1 is a transcriptional target of ERalpha and has a critical role in breast cancer endocrine sensitivity and resistance. *Oncogene*, **29**, 2983–2995.

253. Laoukili,J., Kooistra,M.R.H., Brás,A., Kauw,J., Kerkhoven,R.M., Morrison,A., Clevers,H. and Medema,R.H. (2005) FoxM1 is required for execution of the mitotic programme and chromosome stability. *Nat. Cell Biol.*, **7**, 126–136.

254. Sanders,D.A., Ross-Innes,C.S., Beraldi,D., Carroll,J.S. and Balasubramanian,S. (2013) Genome-wide mapping of FOXM1 binding reveals co-binding with estrogen receptor alpha in breast cancer cells. *Genome Biol.*, **14**, R6.

255. Angeloni,S.V., Martin,M.B., Garcia-Morales,P., Castro-Galache,M.D., Ferragut,J.A. and Saceda,M. (2004) Regulation of estrogen receptor-alpha expression by the tumor suppressor gene p53 in MCF-7 cells. *J. Endocrinol.*, **180**, 497–504.

256. Shirley,S.H., Rundhaug,J.E., Tian,J., Cullinan-Ammann,N., Lambertz,I., Conti,C.J. and Fuchs-Young,R. (2009) Transcriptional regulation of estrogen receptor-alpha by p53 in human breast cancer cells. *Cancer Res.*, **69**, 3405–3414.

257. Duong,V., Boulle,N., Daujat,S., Chauvet,J., Bonnet,S., Neel,H. and Cavaillès,V. (2007) Differential regulation of estrogen receptor alpha turnover and transactivation by Mdm2 and stress-inducing agents. *Cancer Res.*, **67**, 5513–5521.

258. Fuchs-Young,R., Shirley,S.H., Lambertz,I., Colby,J.K.L., Tian,J., Johnston,D., Gimenez-Conti,I.B., Donehower,L.A., Conti,C.J. and Hursting,S.D. (2011) P53 genotype as a determinant of ER expression and tamoxifen response in the MMTV-Wnt-1 model of mammary carcinogenesis. *Breast Cancer Res. Treat.*, **130**, 399–408.

259. Qin,C., Nguyen,T., Stewart,J., Samudio,I., Burghardt,R. and Safe,S. (2002) Estrogen up-regulation of p53 gene expression in MCF-7 breast cancer cells is mediated by calmodulin kinase IV-dependent activation of a nuclear factor kappaB/CCAAT-binding transcription factor-1 complex. *Mol. Endocrinol. Baltim. Md*, **16**, 1793–1809.

260. Hurd,C., Dinda,S., Khattree,N. and Moudgil,V.K. (1999) Estrogen-dependent and independent activation of the P1 promoter of the p53 gene in transiently transfected breast cancer cells. *Oncogene*, **18**, 1067–1072.

261. Berger,C.E., Qian,Y., Liu,G., Chen,H. and Chen,X. (2012) p53, a target of estrogen receptor (ER) α, modulates DNA damage-induced growth suppression in ER-positive breast cancer cells. *J. Biol. Chem.*, **287**, 30117–30127.

262. Frasor,J., El-Shennawy,L., Stender,J.D. and Kastrati,I. (2015) NFκB affects estrogen receptor expression and activity in breast cancer through multiple mechanisms. *Mol. Cell. Endocrinol.*, **418 Pt 3**, 235–239.

263. Nakshatri,H., Bhat-Nakshatri,P., Martin,D.A., Goulet,R.J. and Sledge,G.W. (1997) Constitutive activation of NF-kappaB during progression of breast cancer to hormone-independent growth. *Mol. Cell. Biol.*, **17**, 3629–3639.

264. Van Laere,S.J., Van der Auwera,I., Van den Eynden,G.G., Elst,H.J., Weyler,J., Harris,A.L., van Dam,P., Van Marck,E.A., Vermeulen,P.B. and Dirix,L.Y. (2006) Nuclear factor-kappaB signature of inflammatory breast cancer by cDNA microarray validated by quantitative real-time reverse transcription-PCR, immunohistochemistry, and nuclear factor-kappaB DNA-binding. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, **12**, 3249–3256.

265. Biswas,D.K., Shi,Q., Baily,S., Strickland,I., Ghosh,S., Pardee,A.B. and Iglehart,J.D. (2004) NF-kappa B activation in human breast cancer specimens and its role in cell proliferation and apoptosis. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 10137–10142.

266. Sarkar,D.K., Jana,D., Patil,P.S., Chaudhari,K.S., Chattopadhyay,B.K., Chikkala,B.R., Mandal,S. and Chowdhary,P. (2013) Role of NF-κB as a Prognostic Marker in Breast Cancer : A Pilot Study in Indian Patients. *Indian J. Surg. Oncol.*, **4**, 242–247.

267. Belguise,K. and Sonenshein,G.E. (2007) PKCtheta promotes c-Rel-driven mammary tumorigenesis in mice and humans by repressing estrogen receptor alpha synthesis. *J. Clin. Invest.*, **117**, 4009–4021.

268. Wang,X., Belguise,K., Kersual,N., Kirsch,K.H., Mineva,N.D., Galtier,F., Chalbos,D. and Sonenshein,G.E. (2007) Oestrogen signalling inhibits invasive phenotype by repressing RelB and its target BCL2. *Nat. Cell Biol.*, **9**, 470–478.

269. Wang,X., Belguise,K., O'Neill,C.F., Sánchez-Morgan,N., Romagnoli,M., Eddy,S.F., Mineva,N.D., Yu,Z., Min,C., Trinkaus-Randall,V., *et al.* (2009) RelB NF-kappaB represses estrogen receptor alpha expression via induction of the zinc finger protein Blimp1. *Mol. Cell. Biol.*, **29**, 3832–3844.

270. Reijm,E.A., Jansen,M.P.H.M., Ruigrok-Ritstier,K., van Staveren,I.L., Look,M.P., van Gelder,M.E.M., Sieuwerts,A.M., Sleijfer,S., Foekens,J.A. and Berns,E.M.J.J. (2011) Decreased expression of EZH2 is associated with upregulation of ER and favorable outcome to tamoxifen in advanced breast cancer. *Breast Cancer Res. Treat.*, **125**, 387–394.

271. Acharyya,S., Sharma,S.M., Cheng,A.S., Ladner,K.J., He,W., Kline,W., Wang,H., Ostrowski,M.C., Huang,T.H. and Guttridge,D.C. (2010) TNF inhibits Notch-1 in skeletal muscle cells by Ezh2 and DNA methylation mediated repression: implications in duchenne muscular dystrophy. *PloS One*, **5**, e12479.

272. Hosey,A.M., Gorski,J.J., Murray,M.M., Quinn,J.E., Chung,W.Y., Stewart,G.E., James,C.R., Farragher,S.M., Mulligan,J.M., Scott,A.N., *et al.* (2007) Molecular basis for estrogen receptor alpha deficiency in BRCA1-linked breast cancer. *J. Natl. Cancer Inst.*, **99**, 1683–1694.

273. Ma,Y., Fan,S., Hu,C., Meng,Q., Fuqua,S.A., Pestell,R.G., Tomita,Y.A. and Rosen,E.M. (2010) BRCA1 regulates acetylation and ubiquitination of estrogen receptor-alpha. *Mol. Endocrinol. Baltim. Md*, **24**, 76–90.

274. Kim,M.Y., Woo,E.M., Chong,Y.T.E., Homenko,D.R. and Kraus,W.L. (2006) Acetylation of estrogen receptor alpha by p300 at lysines 266 and 268 enhances the deoxyribonucleic acid binding and transactivation activities of the receptor. *Mol. Endocrinol. Baltim. Md*, **20**, 1479–1493.

275. Castles,C.G., Oesterreich,S., Hansen,R. and Fuqua,S.A. (1997) Auto-regulation of the estrogen receptor promoter. *J. Steroid Biochem. Mol. Biol.*, **62**, 155–163.

276. Larsen,M.J., Kruse,T.A., Tan,Q., Lænkholm,A.-V., Bak,M., Lykkesfeldt,A.E., Sørensen,K.P., Hansen,T.V.O., Ejlertsen,B., Gerdes,A.-M., *et al.* (2013) Classifications within molecular subtypes enables identification of BRCA1/BRCA2 mutation carriers by RNA tumor profiling. *PloS One*, **8**, e64268.

277. Dreijerink,K.M.A., Mulder,K.W., Winkler,G.S., Höppener,J.W.M., Lips,C.J.M. and Timmers,H.T.M. (2006) Menin links estrogen receptor activation to histone H3K4 trimethylation. *Cancer Res.*, **66**, 4929–4935.

278. Hughes,C.M., Rozenblatt-Rosen,O., Milne,T.A., Copeland,T.D., Levine,S.S., Lee,J.C., Hayes,D.N., Shanmugam,K.S., Bhattacharjee,A., Biondi,C.A., *et al.* (2004) Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. *Mol. Cell*, **13**, 587–597.

279. Yokoyama,A., Wang,Z., Wysocka,J., Sanyal,M., Aufiero,D.J., Kitabayashi,I., Herr,W. and Cleary,M.L. (2004) Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol. Cell. Biol.*, **24**, 5639–5649.

280. Dreijerink,K.M.A., Groner,A.C., Vos,E.S.M., Font-Tello,A., Gu,L., Chi,D., Reyes,J., Cook,J., Lim,E., Lin,C.Y., *et al.* (2017) Enhancer-Mediated Oncogenic Function of the Menin Tumor Suppressor in Breast Cancer. *Cell Rep.*, **18**, 2359–2372.

281. Nakata,Y., Brignier,A.C., Jin,S., Shen,Y., Rudnick,S.I., Sugita,M. and Gewirtz,A.M. (2010) c-Myb, Menin, GATA-3, and MLL form a dynamic transcription complex that plays a pivotal role in human T helper type 2 cell development. *Blood*, **116**, 1280–1290.

282. Dreijerink,K.M.A., Goudet,P., Burgess,J.R., Valk,G.D. and International Breast Cancer in MEN1 Study Group (2014) Breast-cancer predisposition in multiple endocrine neoplasia type 1. *N. Engl. J. Med.*, **371**, 583–584.

283. Bolden,J.E., Peart,M.J. and Johnstone,R.W. (2006) Anticancer activities of histone deacetylase inhibitors. *Nat. Rev. Drug Discov.*, **5**, 769–784.

284. Gao,L., Cueto,M.A., Asselbergs,F. and Atadja,P. (2002) Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family. *J. Biol. Chem.*, **277**, 25748–25755.

285. Allfrey,V.G., Faulkner,R. and Mirsky,A.E. (1964) ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proc. Natl. Acad. Sci. U. S. A.*, **51**, 786–794.

286. Hebbes,T.R., Thorne,A.W. and Crane-Robinson,C. (1988) A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.*, **7**, 1395–1402.

287. Roth,S.Y., Denu,J.M. and Allis,C.D. (2001) Histone acetyltransferases. *Annu. Rev. Biochem.*, **70**, 81–120.

288. Fujisawa,T. and Filippakopoulos,P. (2017) Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. *Nat. Rev. Mol. Cell Biol.*, **18**, 246–262.

289. Verdin,E. and Ott,M. (2015) 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat. Rev. Mol. Cell Biol.*, **16**, 258–264.

290. Yamagata,T., Mitani,K., Oda,H., Suzuki,T., Honda,H., Asai,T., Maki,K., Nakamoto,T. and Hirai,H. (2000) Acetylation of GATA-3 affects T-cell survival and homing to secondary lymphoid organs. *EMBO J.*, **19**, 4676–4687.

291. Barlev,N.A., Liu,L., Chehab,N.H., Mansfield,K., Harris,K.G., Halazonetis,T.D. and Berger,S.L. (2001) Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. *Mol. Cell*, **8**, 1243–1254.

292. Jin,Y.-H., Jeon,E.-J., Li,Q.-L., Lee,Y.H., Choi,J.-K., Kim,W.-J., Lee,K.-Y. and Bae,S.-C. (2004) Transforming growth factor-beta stimulates p300-dependent RUNX3 acetylation, which inhibits ubiquitination-mediated degradation. *J. Biol. Chem.*, **279**, 29409–29417.

293. Martínez-Balbás,M.A., Bauer,U.M., Nielsen,S.J., Brehm,A. and Kouzarides,T. (2000) Regulation of E2F1 activity by acetylation. *EMBO J.*, **19**, 662–671.

294. Giandomenico,V., Simonsson,M., Grönroos,E. and Ericsson,J. (2003) Coactivator-dependent acetylation stabilizes members of the SREBP family of transcription factors. *Mol. Cell. Biol.*, **23**, 2587–2599.

295. Gaughan,L., Logan,I.R., Cook,S., Neal,D.E. and Robson,C.N. (2002) Tip60 and histone deacetylase 1 regulate androgen receptor activity through changes to the acetylation status of the receptor. *J. Biol. Chem.*, **277**, 25904–25913.

296. Bereshchenko,O.R., Gu,W. and Dalla-Favera,R. (2002) Acetylation inactivates the transcriptional repressor BCL6. *Nat. Genet.*, **32**, 606–613.

297. Caillaud,A., Prakash,A., Smith,E., Masumi,A., Hovanessian,A.G., Levy,D.E. and Marié,I. (2002) Acetylation of Interferon Regulatory Factor-7 by p300/CREB-binding Protein (CBP)-associated Factor (PCAF) Impairs its DNA Binding. *J. Biol. Chem.*, **277**, 49417–49421.

298. Kohler,S. and Cirillo,L.A. (2010) Stable chromatin binding prevents FoxA acetylation, preserving FoxA chromatin remodeling. *J. Biol. Chem.*, **285**, 464–472.

299. Grönroos,E., Hellman,U., Heldin,C.-H. and Ericsson,J. (2002) Control of Smad7 stability by competition between acetylation and ubiquitination. *Mol. Cell*, **10**, 483–493.

300. Patel,J.H., Du,Y., Ard,P.G., Phillips,C., Carella,B., Chen,C.-J., Rakowski,C., Chatterjee,C., Lieberman,P.M., Lane,W.S., *et al.* (2004) The c-MYC oncoprotein is a substrate of the acetyltransferases hGCN5/PCAF and TIP60. *Mol. Cell. Biol.*, **24**, 10826–10834.

301. Ito,A., Kawaguchi,Y., Lai,C.-H., Kovacs,J.J., Higashimoto,Y., Appella,E. and Yao,T.-P. (2002) MDM2-HDAC1-mediated deacetylation of p53 is required for its degradation. *EMBO J.*, **21**, 6236–6245.

302. Jeong,J.W., Bae,M.K., Ahn,M.Y., Kim,S.H., Sohn,T.K., Bae,M.H., Yoo,M.A., Song,E.J., Lee,K.J. and Kim,K.W. (2002) Regulation and destabilization of HIF-1alpha by ARD1-mediated acetylation. *Cell*, **111**, 709–720.

303. Mateo,F., Vidal-Laliena,M., Canela,N., Busino,L., Martinez-Balbas,M.A., Pagano,M., Agell,N. and Bachs,O. (2009) Degradation of cyclin A is regulated by acetylation. *Oncogene*, **28**, 2654–2666.

304. Yuan,Z.-L., Guan,Y.-J., Chatterjee,D. and Chin,Y.E. (2005) Stat3 dimerization regulated by reversible acetylation of a single lysine residue. *Science*, **307**, 269–273.

305. Scott,D.C., Monda,J.K., Bennett,E.J., Harper,J.W. and Schulman,B.A. (2011) N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. *Science*, **334**, 674–678.

306. Bali,P., Pranpat,M., Bradner,J., Balasis,M., Fiskus,W., Guo,F., Rocha,K., Kumaraswamy,S., Boyapalle,S., Atadja,P., *et al.* (2005) Inhibition of histone deacetylase 6 acetylates and disrupts the chaperone function of heat shock protein 90: a novel basis for antileukemia activity of histone deacetylase inhibitors. *J. Biol. Chem.*, **280**, 26729–26734.

307. Cohen,H.Y., Lavu,S., Bitterman,K.J., Hekking,B., Imahiyerobo,T.A., Miller,C., Frye,R., Ploegh,H., Kessler,B.M. and Sinclair,D.A. (2004) Acetylation of the C terminus of Ku70 by CBP and PCAF controls Bax-mediated apoptosis. *Mol. Cell*, **13**, 627–638.

308. Chen Lf, null, Fischle,W., Verdin,E. and Greene,W.C. (2001) Duration of nuclear NF-kappaB action regulated by reversible acetylation. *Science*, **293**, 1653–1657.

309. Kim,H.-J. and Bae,S.-C. (2011) Histone deacetylase inhibitors: molecular mechanisms of action and clinical trials as anti-cancer drugs. *Am. J. Transl. Res.*, **3**, 166–179.

310. Riggs,M.G., Whittaker,R.G., Neumann,J.R. and Ingram,V.M. (1977) n-Butyrate causes histone modification in HeLa and Friend erythroleukaemia cells. *Nature*, **268**, 462–464.

311. Leder,A. and Leder,P. (1975) Butyric acid, a potent inducer of erythroid differentiation in cultured erythroleukemic cells. *Cell*, **5**, 319–322.

312. Candido,E.P., Reeves,R. and Davie,J.R. (1978) Sodium butyrate inhibits histone deacetylation in cultured cells. *Cell*, **14**, 105–113.

313. Vidali,G., Boffa,L.C., Bradbury,E.M. and Allfrey,V.G. (1978) Butyrate suppression of histone deacetylation leads to accumulation of multiacetylated forms of histones H3 and H4 and increased DNase I sensitivity of the associated DNA sequences. *Proc. Natl. Acad. Sci. U. S. A.*, **75**, 2239–2243.

314. Zhang,Z., Yamashita,H., Toyama,T., Sugiura,H., Ando,Y., Mita,K., Hamaguchi,M., Hara,Y., Kobayashi,S. and Iwase,H. (2005) Quantitation of HDAC1 mRNA expression in invasive carcinoma of the breast*. *Breast Cancer Res. Treat.*, **94**, 11–16.

315. Halkidou,K., Gaughan,L., Cook,S., Leung,H.Y., Neal,D.E. and Robson,C.N. (2004) Upregulation and nuclear recruitment of HDAC1 in hormone refractory prostate cancer. *The Prostate*, **59**, 177–189.

316. Choi,J.H., Kwon,H.J., Yoon,B.I., Kim,J.H., Han,S.U., Joo,H.J. and Kim,D.Y. (2001) Expression profile of histone deacetylase 1 in gastric cancer tissues. *Jpn. J. Cancer Res. Gann*, **92**, 1300–1304.

317. Wilson,A.J., Byun,D.-S., Popova,N., Murray,L.B., L'Italien,K., Sowa,Y., Arango,D., Velcich,A., Augenlicht,L.H. and Mariadason,J.M. (2006) Histone deacetylase 3 (HDAC3) and other class I HDACs regulate colon cell maturation and p21 expression and are deregulated in human colon cancer. *J. Biol. Chem.*, **281**, 13548–13558.

318. Song,J., Noh,J.H., Lee,J.H., Eun,J.W., Ahn,Y.M., Kim,S.Y., Lee,S.H., Park,W.S., Yoo,N.J., Lee,J.Y., *et al.* (2005) Increased expression of histone deacetylase 2 is found in human gastric cancer. *APMIS Acta Pathol. Microbiol. Immunol. Scand.*, **113**, 264–268.

319. Huang,B.H., Laban,M., Leung,C.H.-W., Lee,L., Lee,C.K., Salto-Tellez,M., Raju,G.C. and Hooi,S.C. (2005) Inhibition of histone deacetylase 2 increases apoptosis and p21Cip1/WAF1 expression, independent of histone deacetylase 1. *Cell Death Differ.*, **12**, 395–404.

320. Zhang,Z., Yamashita,H., Toyama,T., Sugiura,H., Omoto,Y., Ando,Y., Mita,K., Hamaguchi,M., Hayashi,S.-I. and Iwase,H. (2004) HDAC6 expression is correlated with better survival in breast cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, **10**, 6962–6968.

321. Marks,P.A. (2010) The clinical development of histone deacetylase inhibitors as targeted anticancer drugs. *Expert Opin. Investig. Drugs*, **19**, 1049–1066.

322. Jones,P.A., Issa,J.-P.J. and Baylin,S. (2016) Targeting the cancer epigenome for therapy. *Nat. Rev. Genet.*, **17**, 630–641.

323. Tsuji,N., Kobayashi,M., Nagashima,K., Wakisaka,Y. and Koizumi,K. (1976) A new antifungal antibiotic, trichostatin. *J. Antibiot. (Tokyo)*, **29**, 1–6.

324. Yoshida,M., Nomura,S. and Beppu,T. (1987) Effects of trichostatins on differentiation of murine erythroleukemia cells. *Cancer Res.*, **47**, 3688–3691.

325. Yoshida,M., Kijima,M., Akita,M. and Beppu,T. (1990) Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. *J. Biol. Chem.*, **265**, 17174–17179.

326. Shirakawa,K., Chavez,L., Hakre,S., Calvanese,V. and Verdin,E. (2013) Reactivation of latent HIV by histone deacetylase inhibitors. *Trends Microbiol.*, **21**, 277–285.

327. Bradner,J.E., West,N., Grachan,M.L., Greenberg,E.F., Haggarty,S.J., Warnow,T. and Mazitschek,R. (2010) Chemical phylogenetics of histone deacetylases. *Nat. Chem. Biol.*, **6**, 238–243.

328. Uchida,H., Maruyama,T., Nagashima,T., Asada,H. and Yoshimura,Y. (2005) Histone deacetylase inhibitors induce differentiation of human endometrial adenocarcinoma cells through up-regulation of glycodelin. *Endocrinology*, **146**, 5365–5373.

329. Strait,K.A., Dabbas,B., Hammond,E.H., Warnick,C.T., Iistrup,S.J. and Ford,C.D. (2002) Cell cycle blockade and differentiation of ovarian cancer cells by the histone deacetylase inhibitor trichostatin A are associated with changes in p21, Rb, and Id proteins. *Mol. Cancer Ther.*, **1**, 1181–1190.

330. Cain,J.E., McCaw,A., Jayasekara,W.S.N., Rossello,F.J., Marini,K.D., Irving,A.T., Kansara,M., Thomas,D.M., Ashley,D.M. and Watkins,D.N. (2013) Sustained Low-Dose Treatment with the Histone Deacetylase Inhibitor LBH589 Induces Terminal Differentiation of Osteosarcoma Cells. *Sarcoma*, **2013**, 608964.

331. Munster,P.N., Troso-Sandoval,T., Rosen,N., Rifkind,R., Marks,P.A. and Richon,V.M. (2001) The histone deacetylase inhibitor suberoylanilide hydroxamic acid induces differentiation of human breast cancer cells. *Cancer Res.*, **61**, 8492–8497.

332. Sandor,V., Senderowicz,A., Mertins,S., Sackett,D., Sausville,E., Blagosklonny,M.V. and Bates,S.E. (2000) P21-dependent g(1)arrest with downregulation of cyclin D1 and upregulation of cyclin E by the histone deacetylase inhibitor FR901228. *Br. J. Cancer*, **83**, 817–825.

333. Glaser,K.B., Staver,M.J., Waring,J.F., Stender,J., Ulrich,R.G. and Davidsen,S.K. (2003) Gene expression profiling of multiple histone deacetylase (HDAC) inhibitors: defining a common gene set produced by HDAC inhibition in T24 and MDA carcinoma cell lines. *Mol. Cancer Ther.*, **2**, 151–163.

334. Insinga,A., Monestiroli,S., Ronzoni,S., Gelmetti,V., Marchesi,F., Viale,A., Altucci,L., Nervi,C., Minucci,S. and Pelicci,P.G. (2005) Inhibitors of histone deacetylases induce tumor-selective apoptosis through activation of the death receptor pathway. *Nat. Med.*, **11**, 71–76.

335. Zhang,X.D., Gillespie,S.K., Borrow,J.M. and Hersey,P. (2004) The histone deacetylase inhibitor suberic bishydroxamate regulates the expression of multiple apoptotic mediators and induces mitochondria-dependent apoptosis of melanoma cells. *Mol. Cancer Ther.*, **3**, 425–435.

336. Rahmani,M., Reese,E., Dai,Y., Bauer,C., Payne,S.G., Dent,P., Spiegel,S. and Grant,S. (2005) Coadministration of histone deacetylase inhibitors and perifosine synergistically induces apoptosis in human leukemia cells through Akt and ERK1/2 inactivation and the generation of ceramide and reactive oxygen species. *Cancer Res.*, **65**, 2422–2432.

337. Butler,L.M., Zhou,X., Xu,W.-S., Scher,H.I., Rifkind,R.A., Marks,P.A. and Richon,V.M. (2002) The histone deacetylase inhibitor SAHA arrests cancer cell growth, up-regulates thioredoxin-binding protein-2, and down-regulates thioredoxin. *Proc. Natl. Acad. Sci. U. S. A.*, **99**, 11700–11705.

338. Maeda,T., Towatari,M., Kosugi,H. and Saito,H. (2000) Up-regulation of costimulatory/adhesion molecules by histone deacetylase inhibitors in acute myeloid leukemia cells. *Blood*, **96**, 3847–3856.

339. Skov,S., Pedersen,M.T., Andresen,L., Straten,P.T., Woetmann,A. and Odum,N. (2005) Cancer cells become susceptible to natural killer cell killing after exposure to histone deacetylase inhibitors due to glycogen synthase kinase-3-dependent expression of MHC class I-related chain A and B. *Cancer Res.*, **65**, 11136–11145.

340. Reddy,P., Maeda,Y., Hotary,K., Liu,C., Reznikov,L.L., Dinarello,C.A. and Ferrara,J.L.M. (2004) Histone deacetylase inhibitor suberoylanilide hydroxamic acid reduces acute graft-versus-host disease and preserves graft-versus-leukemia effect. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 3921–3926.

341. Margueron,R., Licznar,A., Lazennec,G., Vignon,F. and Cavaillès,V. (2003) Oestrogen receptor alpha increases p21(WAF1/CIP1) gene expression and the antiproliferative activity of histone deacetylase inhibitors in human breast cancer cells. *J. Endocrinol.*, **179**, 41–53.

342. Alao,J.P., Lam,E.W.-F., Ali,S., Buluwela,L., Bordogna,W., Lockey,P., Varshochi,R., Stavropoulou,A.V., Coombes,R.C. and Vigushin,D.M. (2004) Histone deacetylase inhibitor trichostatin A represses estrogen receptor alpha-dependent transcription and promotes proteasomal degradation of cyclin D1 in human breast carcinoma cell lines. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, **10**, 8094–8104.

343. Travaglini,L., Vian,L., Billi,M., Grignani,F. and Nervi,C. (2009) Epigenetic reprogramming of breast cancer cells by valproic acid occurs regardless of estrogen receptor status. *Int. J. Biochem. Cell Biol.*, **41**, 225–234.

344. Davis,T., Kennedy,C., Chiew,Y.E., Clarke,C.L. and deFazio,A. (2000) Histone deacetylase inhibitors decrease proliferation and modulate cell cycle gene expression in normal mammary epithelial cells. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, **6**, 4334–4342.

345. Palmieri,D., Lockman,P.R., Thomas,F.C., Hua,E., Herring,J., Hargrave,E., Johnson,M., Flores,N., Qian,Y., Vega-Valle,E., *et al.* (2009) Vorinostat inhibits brain metastatic colonization in a model of triple-negative breast cancer and induces DNA double-strand breaks. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, **15**, 6148–6157.

346. Hirokawa,Y., Arnold,M., Nakajima,H., Zalcberg,J. and Maruta,H. (2005) Signal therapy of breast cancers by the HDAC inhibitor FK228 that blocks the activation of PAK1 and abrogates the tamoxifen-resistance. *Cancer Biol. Ther.*, **4**, 956–960.

347. Vigushin,D.M., Ali,S., Pace,P.E., Mirsaidi,N., Ito,K., Adcock,I. and Coombes,R.C. (2001) Trichostatin A is a histone deacetylase inhibitor with potent antitumor activity against breast cancer in vivo. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, **7**, 971–976.

348. Fan,J., Yin,W.-J., Lu,J.-S., Wang,L., Wu,J., Wu,F.-Y., Di,G.-H., Shen,Z.-Z. and Shao,Z.-M. (2008) ER alpha negative breast cancer cells restore response to endocrine therapy by combination treatment with both HDAC inhibitor and DNMT inhibitor. *J. Cancer Res. Clin. Oncol.*, **134**, 883–890.

349. Rocha,W., Sanchez,R., Deschênes,J., Auger,A., Hébert,E., White,J.H. and Mader,S. (2005) Opposite effects of histone deacetylase inhibitors on glucocorticoid and estrogen signaling in human endometrial Ishikawa cells. *Mol. Pharmacol.*, **68**, 1852–1862.

350. Reid,G., Métivier,R., Lin,C.-Y., Denger,S., Ibberson,D., Ivacevic,T., Brand,H., Benes,V., Liu,E.T. and Gannon,F. (2005) Multiple mechanisms induce transcriptional silencing of a subset of genes, including oestrogen receptor alpha, in response to deacetylase inhibition by valproic acid and trichostatin A. *Oncogene*, **24**, 4894–4907.

351. Biçaku,E., Marchion,D.C., Schmitt,M.L. and Münster,P.N. (2008) Selective inhibition of histone deacetylase 2 silences progesterone receptor-mediated signaling. *Cancer Res.*, **68**, 1513–1519.

352. Yao,Y., Li,H., Gu,Y., Davidson,N.E. and Zhou,Q. (2010) Inhibition of SIRT1 deacetylase suppresses estrogen receptor signaling. *Carcinogenesis*, **31**, 382–387.

353. Pryzbylkowski,P., Obajimi,O. and Keen,J.C. (2008) Trichostatin A and 5 Aza-2' deoxycytidine decrease estrogen receptor mRNA stability in ER positive MCF7 cells through modulation of HuR. *Breast Cancer Res. Treat.*, **111**, 15–25.

354. Gajulapalli,V.N.R., Malisetty,V.L., Chitta,S.K. and Manavathi,B. (2016) Oestrogen receptor negativity in breast cancer: a cause or consequence? *Biosci. Rep.*, **36**.

355. Yi,X., Wei,W., Wang,S.-Y., Du,Z.-Y., Xu,Y.-J. and Yu,X.-D. (2008) Histone deacetylase inhibitor SAHA induces ERalpha degradation in breast cancer MCF-7 cells by CHIP-mediated ubiquitin pathway and inhibits survival signaling. *Biochem. Pharmacol.*, **75**, 1697–1705.

356. Fiskus,W., Ren,Y., Mohapatra,A., Bali,P., Mandawat,A., Rao,R., Herger,B., Yang,Y., Atadja,P., Wu,J., *et al.* (2007) Hydroxamic acid analogue histone deacetylase inhibitors attenuate estrogen receptor-alpha levels and transcriptional activity: a result of hyperacetylation and inhibition of chaperone function of heat shock protein 90. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, **13**, 4882–4890.

357. Ottaviano,Y.L., Issa,J.P., Parl,F.F., Smith,H.S., Baylin,S.B. and Davidson,N.E. (1994) Methylation of the estrogen receptor gene CpG island marks loss of estrogen receptor expression in human breast cancer cells. *Cancer Res.*, **54**, 2552–2555.

358. Ferguson,A.T., Lapidus,R.G., Baylin,S.B. and Davidson,N.E. (1995) Demethylation of the estrogen receptor gene in estrogen receptor-negative breast cancer cells can reactivate estrogen receptor gene expression. *Cancer Res.*, **55**, 2279–2283.

359. Sharma,D., Saxena,N.K., Davidson,N.E. and Vertino,P.M. (2006) Restoration of tamoxifen sensitivity in estrogen receptor-negative breast cancer cells: tamoxifen-bound reactivated ER recruits distinctive corepressor complexes. *Cancer Res.*, **66**, 6370–6378.

360. Zhou,Q., Atadja,P. and Davidson,N.E. (2007) Histone deacetylase inhibitor LBH589 reactivates silenced estrogen receptor alpha (ER) gene expression without loss of DNA hypermethylation. *Cancer Biol. Ther.*, **6**, 64–69.

361. Jang,E.R., Lim,S.-J., Lee,E.S., Jeong,G., Kim,T.-Y., Bang,Y.-J. and Lee,J.-S. (2004) The histone deacetylase inhibitor trichostatin A sensitizes estrogen receptor alpha-negative breast cancer cells to tamoxifen. *Oncogene*, **23**, 1724–1736.

362. Zhou,Q., Shaw,P.G. and Davidson,N.E. (2009) Inhibition of histone deacetylase suppresses EGF signaling pathways by destabilizing EGFR mRNA in ER-negative human breast cancer cells. *Breast Cancer Res. Treat.*, **117**, 443–451.

363. Hodges-Gallagher,L., Valentine,C.D., Bader,S.E. and Kushner,P.J. (2007) Inhibition of histone deacetylase enhances the anti-proliferative action of antiestrogens on breast cancer cells and blocks tamoxifen-induced proliferation of uterine cells. *Breast Cancer Res. Treat.*, **105**, 297–309.

364. Munster,P.N., Thurn,K.T., Thomas,S., Raha,P., Lacevic,M., Miller,A., Melisko,M., Ismail-Khan,R., Rugo,H., Moasser,M., *et al.* (2011) A phase II study of the histone deacetylase inhibitor vorinostat combined with tamoxifen for the treatment of patients with hormone therapy-resistant breast cancer. *Br. J. Cancer*, **104**, 1828–1835.

365. Chia,S., Gradishar,W., Mauriac,L., Bines,J., Amant,F., Federico,M., Fein,L., Romieu,G., Buzdar,A., Robertson,J.F.R., *et al.* (2008) Double-blind, randomized placebo controlled trial of fulvestrant compared with exemestane after prior nonsteroidal aromatase inhibitor therapy in postmenopausal women with hormone receptor-positive, advanced breast cancer: results from EFECT. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.*, **26**, 1664–1670.

366. Lemieux,S., Sargeant,T., Laperrière,D., Ismail,H., Boucher,G., Rozendaal,M., Lavallée,V.-P., Ashton-Beaucage,D., Wilhelm,B., Hébert,J., *et al.* (2017) MiSTIC, an integrated platform for the analysis of heterogeneity in large tumour transcriptome datasets. *Nucleic Acids Res.*, 10.1093/nar/gkx338.

367. Prat,A., Parker,J.S., Karginova,O., Fan,C., Livasy,C., Herschkowitz,J.I., He,X. and Perou,C.M. (2010) Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res. BCR*, **12**, R68.

368. Puhalla,S., Bhattacharya,S. and Davidson,N.E. (2012) Hormonal therapy in breast cancer: a model disease for the personalization of cancer care. *Mol. Oncol.*, **6**, 222–236.

369. Callahan,R. and Hurvitz,S. (2011) Human epidermal growth factor receptor-2-positive breast cancer: Current management of early, advanced, and recurrent disease. *Curr. Opin. Obstet. Gynecol.*, **23**, 37–43.

370. Robson,M., Im,S.-A., Senkus,E., Xu,B., Domchek,S.M., Masuda,N., Delaloge,S., Li,W., Tung,N., Armstrong,A., *et al.* (2017) Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N. Engl. J. Med.*, **377**, 523–533.

371. Lim,E., Wu,D., Pal,B., Bouras,T., Asselin-Labat,M.-L., Vaillant,F., Yagita,H., Lindeman,G.J., Smyth,G.K. and Visvader,J.E. (2010) Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. *Breast Cancer Res. BCR*, **12**, R21.

372. Visvader,J.E. and Stingl,J. (2014) Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev.*, **28**, 1143–1158.

373. Takahashi,K. and Yamanaka,S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, **126**, 663–676.

374. Lamouille,S., Xu,J. and Derynck,R. (2014) Molecular mechanisms of epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.*, **15**, 178–196.

375. Onitilo,A.A., Engel,J.M., Greenlee,R.T. and Mukesh,B.N. (2009) Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. *Clin. Med. Res.*, **7**, 4–13.

376. Magnani,L., Eeckhoute,J. and Lupien,M. (2011) Pioneer factors: directing transcriptional regulators within the chromatin environment. *Trends Genet. TIG*, **27**, 465–474.

377. Hurtado,A., Holmes,K.A., Ross-Innes,C.S., Schmidt,D. and Carroll,J.S. (2011) FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat. Genet.*, **43**, 27–33.

378. Theodorou,V., Stark,R., Menon,S. and Carroll,J.S. (2013) GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Res.*, **23**, 12–22.

379. Asselin-Labat,M.-L., Sutherland,K.D., Barker,H., Thomas,R., Shackleton,M., Forrest,N.C., Hartley,L., Robb,L., Grosveld,F.G., van der Wees,J., *et al.* (2007) Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. *Nat. Cell Biol.*, **9**, 201–209.

380. Eeckhoute,J., Keeton,E.K., Lupien,M., Krum,S.A., Carroll,J.S. and Brown,M. (2007) Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. *Cancer Res.*, **67**, 6477–6483.

381. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods San Diego Calif*, **25**, 402–408.

382. Bray,N.L., Pimentel,H., Melsted,P. and Pachter,L. (2016) Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.*, **34**, 525–527.

383. Wang,T., Wei,J.J., Sabatini,D.M. and Lander,E.S. (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, **343**, 80–84.

384. Smid,M., Wang,Y., Zhang,Y., Sieuwerts,A.M., Yu,J., Klijn,J.G.M., Foekens,J.A. and Martens,J.W.M. (2008) Subtypes of breast cancer show preferential site of relapse. *Cancer Res.*, **68**, 3108–3114.

385. Muggerud,A.A., Johnsen,H., Barnes,D.A., Steel,A., Lønning,P.E., Naume,B., Sørlie,T. and Børresen-Dale,A.-L. (2006) Evaluation of MetriGenix custom 4D arrays applied for detection of breast cancer subtypes. *BMC Cancer*, **6**, 59.

386. Sørlie,T., Wang,Y., Xiao,C., Johnsen,H., Naume,B., Samaha,R.R. and Børresen-Dale,A.-L. (2006) Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms. *BMC Genomics*, **7**, 127.

387. Schmidt,D., Schwalie,P.C., Ross-Innes,C.S., Hurtado,A., Brown,G.D., Carroll,J.S., Flicek,P. and Odom,D.T. (2010) A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.*, **20**, 578–588.

388. Tan,S.K., Lin,Z.H., Chang,C.W., Varang,V., Chng,K.R., Pan,Y.F., Yong,E.L., Sung,W.K., Sung,W.K. and Cheung,E. (2011) AP-2γ regulates oestrogen receptor-mediated long-range chromatin interaction and gene transcription. *EMBO J.*, **30**, 2569–2581.

389. Fletcher,M.N.C., Castro,M.A.A., Wang,X., de Santiago,I., O'Reilly,M., Chin,S.-F., Rueda,O.M., Caldas,C., Ponder,B.A.J., Markowetz,F., *et al.* (2013) Master regulators of FGFR2 signalling and breast cancer risk. *Nat. Commun.*, **4**, 2464.

390. Kittler,R., Zhou,J., Hua,S., Ma,L., Liu,Y., Pendleton,E., Cheng,C., Gerstein,M. and White,K.P. (2013) A comprehensive nuclear receptor network for breast cancer cells. *Cell Rep.*, **3**, 538–551.

391. Massie,C.E., Lynch,A., Ramos-Montoya,A., Boren,J., Stark,R., Fazli,L., Warren,A., Scott,H., Madhu,B., Sharma,N., *et al.* (2011) The androgen receptor fuels prostate cancer by regulating central metabolism and biosynthesis. *EMBO J.*, **30**, 2719–2733.

392. Usary,J., Llaca,V., Karaca,G., Presswala,S., Karaca,M., He,X., Langerød,A., Kåresen,R., Oh,D.S., Dressler,L.G., *et al.* (2004) Mutation of GATA3 in human breast tumors. *Oncogene*, **23**, 7669–7678.

393. Adomas,A.B., Grimm,S.A., Malone,C., Takaku,M., Sims,J.K. and Wade,P.A. (2014) Breast tumor specific mutation in GATA3 affects physiological mechanisms regulating transcription factor turnover. *BMC Cancer*, **14**, 278.

394. Guedj,M., Marisa,L., de Reynies,A., Orsetti,B., Schiappa,R., Bibeau,F., MacGrogan,G., Lerebours,F., Finetti,P., Longy,M., *et al.* (2012) A refined molecular taxonomy of breast cancer. *Oncogene*, **31**, 1196–1206.

395. Farmer,P., Bonnefoi,H., Becette,V., Tubiana-Hulin,M., Fumoleau,P., Larsimont,D., Macgrogan,G., Bergh,J., Cameron,D., Goldstein,D., *et al.* (2005) Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene*, **24**, 4660–4671.

396. Robinson,J.L.L., Macarthur,S., Ross-Innes,C.S., Tilley,W.D., Neal,D.E., Mills,I.G. and Carroll,J.S. (2011) Androgen receptor driven transcription in molecular apocrine breast cancer is mediated by FoxA1. *EMBO J.*, **30**, 3019–3027.

397. Grigoriadis,A., Mackay,A., Noel,E., Wu,P.J., Natrajan,R., Frankum,J., Reis-Filho,J.S. and Tutt,A. (2012) Molecular characterisation of cell line models for triple-negative breast cancers. *BMC Genomics*, **13**, 619.

398. Daemen,A., Griffith,O.L., Heiser,L.M., Wang,N.J., Enache,O.M., Sanborn,Z., Pepin,F., Durinck,S., Korkola,J.E., Griffith,M., *et al.* (2013) Modeling precision treatment of breast cancer. *Genome Biol.*, **14**, R110.

399. Serandour,A.A., Brown,G.D., Cohen,J.D. and Carroll,J.S. (2013) Development of an Illumina-based ChIP-exonuclease method provides insight into FoxA1-DNA binding properties. *Genome Biol.*, **14**, R147.

400. Quintin,J., Le Péron,C., Palierne,G., Bizot,M., Cunha,S., Sérandour,A.A., Avner,S., Henry,C., Percevault,F., Belaud-Rotureau,M.-A., *et al.* (2014) Dynamic estrogen receptor interactomes control estrogen-responsive trefoil Factor (TFF) locus cell-specific activities. *Mol. Cell. Biol.*, **34**, 2418–2436.

401. Ghosh,M.G., Thompson,D.A. and Weigel,R.J. (2000) PDZK1 and GREB1 are estrogen-regulated genes expressed in hormone-responsive breast cancer. *Cancer Res.*, **60**, 6367–6375.

402. Bourdeau,V., Deschênes,J., Métivier,R., Nagai,Y., Nguyen,D., Bretschneider,N., Gannon,F., White,J.H. and Mader,S. (2004) Genome-wide identification of high-affinity estrogen response elements in human and mouse. *Mol. Endocrinol. Baltim. Md*, **18**, 1411–1427.

403. Matsuda,T., Yamamoto,T., Muraguchi,A. and Saatcioglu,F. (2001) Cross-talk between transforming growth factor-beta and estrogen receptor signaling through Smad3. *J. Biol. Chem.*, **276**, 42908–42914.

404. Wang,S.E., Wu,F.Y., Shin,I., Qu,S. and Arteaga,C.L. (2005) Transforming growth factor {beta} (TGF-{beta})-Smad target gene protein tyrosine phosphatase receptor type kappa is required for TGF-{beta} function. *Mol. Cell. Biol.*, **25**, 4703–4715.

405. Giresi,P.G., Kim,J., McDaniell,R.M., Iyer,V.R. and Lieb,J.D. (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res.*, **17**, 877–885.

406. Sanga,S., Broom,B.M., Cristini,V. and Edgerton,M.E. (2009) Gene expression meta-analysis supports existence of molecular apocrine breast cancer with a role for androgen receptor and implies interactions with ErbB family. *BMC Med. Genomics*, **2**, 59.

407. Naderi,A. and Hughes-Davies,L. (2008) A functionally significant cross-talk between androgen receptor and ErbB2 pathways in estrogen receptor negative breast cancer. *Neoplasia N. Y. N*, **10**, 542–548.

408. Ni,M., Chen,Y., Lim,E., Wimberly,H., Bailey,S.T., Imai,Y., Rimm,D.L., Liu,X.S. and Brown,M. (2011) Targeting androgen receptor in estrogen receptor-negative breast cancer. *Cancer Cell*, **20**, 119–131.

409. Robinson,J.L.L. and Carroll,J.S. (2012) FoxA1 is a key mediator of hormonal response in breast and prostate cancer. *Front. Endocrinol.*, **3**, 68.

410. Network,T.C.G.A. (2012) Comprehensive molecular portraits of human breast tumours. *Nature*, **490**, 61–70.

411. Gong,C., Fujino,K., Monteiro,L.J., Gomes,A.R., Drost,R., Davidson-Smith,H., Takeda,S., Khoo,U.S., Jonkers,J., Sproul,D., *et al.* (2015) FOXA1 repression is associated with loss of BRCA1 and increased promoter methylation and chromatin silencing in breast cancer. *Oncogene*, **34**, 5012–5024.

412. Williamson,E.A., Wolf,I., O'Kelly,J., Bose,S., Tanosaki,S. and Koeffler,H.P. (2006) BRCA1 and FOXA1 proteins coregulate the expression of the cell cycle-dependent kinase inhibitor p27(Kip1). *Oncogene*, **25**, 1391–1399.

413. Linares,A., Dalenc,F., Balaguer,P., Boulle,N. and Cavailles,V. (2011) Manipulating protein acetylation in breast cancer: a promising approach in combination with hormonal therapies? *J. Biomed. Biotechnol.*, **2011**, 856985.

414. Mehra,R., Varambally,S., Ding,L., Shen,R., Sabel,M.S., Ghosh,D., Chinnaiyan,A.M. and Kleer,C.G. (2005) Identification of GATA3 as a Breast Cancer Prognostic Marker by Global Gene Expression Meta-analysis. *Cancer Res.*, **65**, 11259–11264.

415. Behr,R., Brestelli,J., Fulmer,J.T., Miyawaki,N., Kleyman,T.R. and Kaestner,K.H. (2004) Mild nephrogenic diabetes insipidus caused by Foxa1 deficiency. *J. Biol. Chem.*, **279**, 41936–41941.

416. Gao,N., Ishii,K., Mirosevich,J., Kuwajima,S., Oppenheimer,S.R., Roberts,R.L., Jiang,M., Yu,X., Shappell,S.B., Caprioli,R.M., *et al.* (2005) Forkhead box A1 regulates prostate ductal morphogenesis and promotes epithelial cell maturation. *Dev. Camb. Engl.*, **132**, 3431–3443.

417. Huether,R., Dong,L., Chen,X., Wu,G., Parker,M., Wei,L., Ma,J., Edmonson,M.N., Hedlund,E.K., Rusch,M.C., *et al.* (2014) The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nat. Commun.*, **5**, 3630.

418. Mitsiades,C.S., Mitsiades,N.S., McMullan,C.J., Poulaki,V., Shringarpure,R., Hideshima,T., Akiyama,M., Chauhan,D., Munshi,N., Gu,X., *et al.* (2004) Transcriptional signature of histone deacetylase inhibition in multiple myeloma: biological and clinical implications. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 540–545.

419. Xu,W.S., Parmigiani,R.B. and Marks,P.A. (2007) Histone deacetylase inhibitors: molecular mechanisms of action. *Oncogene*, **26**, 5541–5552.

420. Choudhary,C., Kumar,C., Gnad,F., Nielsen,M.L., Rehman,M., Walther,T.C., Olsen,J.V. and Mann,M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*, **325**, 834–840.

421. Huang,L. and Pardee,A.B. (2000) Suberoylanilide hydroxamic acid as a potential therapeutic agent for human breast cancer treatment. *Mol. Med. Camb. Mass*, **6**, 849–866.

422. Huang,L., Sowa,Y., Sakai,T. and Pardee,A.B. (2000) Activation of the p21WAF1/CIP1 promoter independent of p53 by the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) through the Sp1 sites. *Oncogene*, **19**, 5712–5719.

423. Ottaviano,Y.L., Issa,J.P., Parl,F.F., Smith,H.S., Baylin,S.B. and Davidson,N.E. (1994) Methylation of the estrogen receptor gene CpG island marks loss of estrogen receptor expression in human breast cancer cells. *Cancer Res.*, **54**, 2552–2555.

424. Sharma,D., Blum,J., Yang,X., Beaulieu,N., Macleod,A.R. and Davidson,N.E. (2005) Release of methyl CpG binding proteins and histone deacetylase 1 from the Estrogen receptor alpha (ER) promoter upon reactivation in ER-negative human breast cancer cells. *Mol. Endocrinol. Baltim. Md*, **19**, 1740–1751.

425. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods San Diego Calif*, **25**, 402–408.

426. Du,P., Kibbe,W.A. and Lin,S.M. (2008) lumi: a pipeline for processing Illumina microarray. *Bioinforma. Oxf. Engl.*, **24**, 1547–1548.

427. Smyth,G.K. (2005) limma: Linear Models for Microarray Data. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, Statistics for Biology and Health. Springer, New York, NY, pp. 397–420.

428. Smyth,G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, Article3.

429. Benjamini,Y. and Hochberg,Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.*, **57**, 289–300.

430. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinforma. Oxf. Engl.*, **30**, 2114–2120.

431. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.*, **29**, 15–21.

432. Anders,S. and Huber,W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106.

433. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.*, **26**, 139–140.

434. Mi,H., Muruganujan,A., Casagrande,J.T. and Thomas,P.D. (2013) Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.*, **8**, 1551–1566.

435. Bourdeau,V., Deschênes,J., Laperrière,D., Aid,M., White,J.H. and Mader,S. (2008) Mechanisms of primary and secondary estrogen target gene regulation in breast cancer cells. *Nucleic Acids Res.*, **36**, 76–93.

436. Hu,E., Dul,E., Sung,C.-M., Chen,Z., Kirkpatrick,R., Zhang,G.-F., Johanson,K., Liu,R., Lago,A., Hofmann,G., *et al.* (2003) Identification of novel isoform-selective inhibitors within class I histone deacetylases. *J. Pharmacol. Exp. Ther.*, **307**, 720–728.

437. Peart,M.J., Smyth,G.K., van Laar,R.K., Bowtell,D.D., Richon,V.M., Marks,P.A., Holloway,A.J. and Johnstone,R.W. (2005) Identification and functional significance of genes regulated by structurally different histone deacetylase inhibitors. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 3697–3702.

438. Sasakawa,Y., Naoe,Y., Sogo,N., Inoue,T., Sasakawa,T., Matsuo,M., Manda,T. and Mutoh,S. (2005) Marker genes to predict sensitivity to FK228, a histone deacetylase inhibitor. *Biochem. Pharmacol.*, **69**, 603–616.

439. Chawla,K., Tripathi,S., Thommesen,L., Lægreid,A. and Kuiper,M. (2013) TFcheckpoint: a curated compendium of specific DNA-binding RNA polymerase II transcription factors. *Bioinforma. Oxf. Engl.*, **29**, 2519–2520.

440. Mi,H., Huang,X., Muruganujan,A., Tang,H., Mills,C., Kang,D. and Thomas,P.D. (2017) PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.*, **45**, D183–D189.

441. Gene Ontology Consortium (2015) Gene Ontology Consortium: going forward. *Nucleic Acids Res.*, **43**, D1049-1056.

442. Gradilone,S.A., Radtke,B.N., Bogert,P.S., Huang,B.Q., Gajdos,G.B. and LaRusso,N.F. (2013) HDAC6 inhibition restores ciliary expression and decreases tumor growth. *Cancer Res.*, **73**, 2259–2270.

443. Pugacheva,E.N., Jablonski,S.A., Hartman,T.R., Henske,E.P. and Golemis,E.A. (2007) HEF1-dependent Aurora A activation induces disassembly of the primary cilium. *Cell*, **129**, 1351–1363.

444. Corbit,K.C., Aanstad,P., Singla,V., Norman,A.R., Stainier,D.Y.R. and Reiter,J.F. (2005) Vertebrate Smoothened functions at the primary cilium. *Nature*, **437**, 1018–1021.

445. Haycraft,C.J., Banizs,B., Aydin-Son,Y., Zhang,Q., Michaud,E.J. and Yoder,B.K. (2005) Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. *PLoS Genet.*, **1**, e53.

446. Kovacs,J.J., Whalen,E.J., Liu,R., Xiao,K., Kim,J., Chen,M., Wang,J., Chen,W. and Lefkowitz,R.J. (2008) Beta-arrestin-mediated localization of smoothened to the primary cilium. *Science*, **320**, 1777–1781.

447. Rohatgi,R., Milenkovic,L. and Scott,M.P. (2007) Patched1 regulates hedgehog signaling at the primary cilium. *Science*, **317**, 372–376.

448. Drabsch,Y., Ramsay,R.G. and Gonda,T.J. (2010) MYBsuppresses differentiation and apoptosis of human breast cancer cells. *Breast Cancer Res.*, **12**, R55.

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

450. Srivastava,R.K., Kurzrock,R. and Shankar,S. (2010) MS-275 sensitizes TRAIL-resistant breast cancer cells, inhibits angiogenesis and metastasis, and reverses epithelial-mesenchymal transition in vivo. *Mol. Cancer Ther.*, **9**, 3254–3266.

451. Krämer,A., Green,J., Pollard,J. and Tugendreich,S. (2014) Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinforma. Oxf. Engl.*, **30**, 523–530.

452. Ingenuity Pathway Analysis *QIAGEN Bioinforma.*

453. Yang,Y.-I., Ahn,J.-H., Lee,K.-T., Shih,I.-M. and Choi,J.-H. (2014) RSF1 is a positive regulator of NF-κB-induced gene expression required for ovarian cancer chemoresistance. *Cancer Res.*, **74**, 2258–2269.

454. Geng,H., Liu,Q., Xue,C., David,L.L., Beer,T.M., Thomas,G.V., Dai,M.-S. and Qian,D.Z. (2012) HIF1α protein stability is increased by acetylation at lysine 709. *J. Biol. Chem.*, **287**, 35496–35505.

455. Yu,J., de Belle,I., Liang,H. and Adamson,E.D. (2004) Coactivating factors p300 and CBP are transcriptionally crossregulated by Egr1 in prostate cells, leading to divergent responses. *Mol. Cell*, **15**, 83–94.

456. Karanam,B., Wang,L., Wang,D., Liu,X., Marmorstein,R., Cotter,R. and Cole,P.A. (2007) Multiple roles for acetylation in the interaction of p300 HAT with ATF-2. *Biochemistry (Mosc.)*, **46**, 8207–8216.

457. Liu,D.X., Qian,D., Wang,B., Yang,J.-M. and Lu,Z. (2011) p300-Dependent ATF5 acetylation is essential for Egr-1 gene activation and cell proliferation and survival. *Mol. Cell. Biol.*, **31**, 3906–3916.

458. Gachon,F., Devaux,C. and Mesnard,J.-M. (2002) Activation of HTLV-I transcription in the presence of Tax is independent of the acetylation of CREB-2 (ATF-4). *Virology*, **299**, 271–278.

459. Simonsson,M., Kanduri,M., Grönroos,E., Heldin,C.-H. and Ericsson,J. (2006) The DNA binding activities of Smad2 and Smad3 are regulated by coactivator-mediated acetylation. *J. Biol. Chem.*, **281**, 39870–39880.

460. Inoue,Y., Itoh,Y., Abe,K., Okamoto,T., Daitoku,H., Fukamizu,A., Onozaki,K. and Hayashi,H. (2007) Smad3 is acetylated by p300/CBP to regulate its transactivation activity. *Oncogene*, **26**, 500–508.

461. Chen,J., Fu,Y., Day,D.S., Sun,Y., Wang,S., Liang,X., Gu,F., Zhang,F., Stevens,S.M., Zhou,P., *et al.* (2017) VEGF amplifies transcription through ETS1 acetylation to enable angiogenesis. *Nat. Commun.*, **8**, 383.

462. Goel,A. and Janknecht,R. (2003) Acetylation-mediated transcriptional activation of the ETS protein ER81 by p300, P/CAF, and HER2/Neu. *Mol. Cell. Biol.*, **23**, 6243–6254.

463. Lee,J.S., Galvin,K.M., See,R.H., Eckner,R., Livingston,D., Moran,E. and Shi,Y. (1995) Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300. *Genes Dev.*, **9**, 1188–1198.

464. Yang,W.M., Inouye,C., Zeng,Y., Bearss,D. and Seto,E. (1996) Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. *Proc. Natl. Acad. Sci. U. S. A.*, **93**, 12845–12850.

465. Yao,Y.L., Yang,W.M. and Seto,E. (2001) Regulation of transcription factor YY1 by acetylation and deacetylation. *Mol. Cell. Biol.*, **21**, 5979–5991.

466. Neve,R.M., Chin,K., Fridlyand,J., Yeh,J., Baehner,F.L., Fevr,T., Clark,L., Bayani,N., Coppe,J.-P., Tong,F., *et al.* (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell*, **10**, 515–527.

467. McManaman,J.L. (2009) Formation of milk lipids: a molecular perspective. *Clin. Lipidol.*, **4**, 391–401.

468. Mohammad,M.A.. H.,M.W. (2013) Regulation of lipid synthesis genes and milk fat production in human mammary epithelial cells during secretory activation. *Am J Physiol Endocrinol Metab*, **305**, E700-16.

469. Crown,S.B., Marze,N. and Antoniewicz,M.R. (2015) Catabolism of Branched Chain Amino Acids Contributes Significantly to Synthesis of Odd-Chain and Even-Chain Fatty Acids in 3T3-L1 Adipocytes. *PloS One*, **10**, e0145850.

470. Amemiya-Kudo,M., Shimano,H., Yoshikawa,T., Yahagi,N., Hasty,A.H., Okazaki,H., Tamura,Y., Shionoiri,F., Iizuka,Y., Ohashi,K., *et al.* (2000) Promoter Analysis of the Mouse Sterol Regulatory Element-binding Protein-1c Gene. *J. Biol. Chem.*, **275**, 31078–31085.

471. Horton,J.D., Shah,N.A., Warrington,J.A., Anderson,N.N., Park,S.W., Brown,M.S. and Goldstein,J.L. (2003) Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. *Proc. Natl. Acad. Sci.*, **100**, 12027–12032.

472. Bowers,E.M., Yan,G., Mukherjee,C., Orry,A., Wang,L., Holbert,M.A., Crump,N.T., Hazzalin,C.A., Liszczak,G., Yuan,H., *et al.* (2010) Virtual ligand screening of the p300/CBP histone acetyltransferase: identification of a selective small molecule inhibitor. *Chem. Biol.*, **17**, 471–482.

473. Kim,Y.J., Greer,C.B., Cecchini,K.R., Harris,L.N., Tuck,D.P. and Kim,T.H. (2013) HDAC inhibitors induce transcriptional repression of high copy number genes in breast cancer through elongation blockade. *Oncogene*, **32**, 2828–2835.

474. Rafehi,H., Balcerczyk,A., Lunke,S., Kaspi,A., Ziemann,M., Kn,H., Okabe,J., Khurana,I., Ooi,J., Khan,A.W., *et al.* (2014) Vascular histone deacetylation by pharmacological HDAC inhibition. *Genome Res.*, **24**, 1271–1284.

475. Ou,J.-N., Torrisani,J., Unterberger,A., Provençal,N., Shikimi,K., Karimi,M., Ekström,T.J. and Szyf,M. (2007) Histone deacetylase inhibitor Trichostatin A induces global and gene-specific DNA demethylation in human cancer cell lines. *Biochem. Pharmacol.*, **73**, 1297–1307.

476. Zhou,Q., Agoston,A.T., Atadja,P., Nelson,W.G. and Davidson,N.E. (2008) Inhibition of histone deacetylases promotes ubiquitin-dependent proteasomal degradation of DNA methyltransferase 1 in human breast cancer cells. *Mol. Cancer Res. MCR*, **6**, 873–883.

477. Halsall,J.A., Turan,N., Wiersma,M. and Turner,B.M. (2015) Cells adapt to the epigenomic disruption caused by histone deacetylase inhibitors through a coordinated, chromatin-mediated transcriptional response. *Epigenetics Chromatin*, **8**, 29.

478. Lutz,M., Burke,L.J., Barreto,G., Goeman,F., Greb,H., Arnold,R., Schultheiss,H., Brehm,A., Kouzarides,T., Lobanenkov,V., *et al.* (2000) Transcriptional repression by the insulator protein CTCF involves histone deacetylases. *Nucleic Acids Res.*, **28**, 1707–1713.

479. Kimbung,S., Lettiero,B., Feldt,M., Bosch,A. and Borgquist,S. (2016) High expression of cholesterol biosynthesis genes is associated with resistance to statin treatment and inferior survival in breast cancer. *Oncotarget*, **7**, 59640–59651.

480. Clendening,J.W., Pandyra,A., Boutros,P.C., El Ghamrasni,S., Khosravi,F., Trentin,G.A., Martirosyan,A., Hakem,A., Hakem,R., Jurisica,I., *et al.* (2010) Dysregulation of the mevalonate pathway promotes transformation. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 15051–15056.

481. Simigdala,N., Gao,Q., Pancholi,S., Roberg-Larsen,H., Zvelebil,M., Ribas,R., Folkerd,E., Thompson,A., Bhamra,A., Dowsett,M., *et al.* (2016) Cholesterol biosynthesis pathway as a

novel mechanism of resistance to estrogen deprivation in estrogen receptor-positive breast cancer. *Breast Cancer Res. BCR*, **18**, 58.

482. Hoefnagel,L.D.C., van de Vijver,M.J., van Slooten,H.-J., Wesseling,P., Wesseling,J., Westenend,P.J., Bart,J., Seldenrijk,C.A., Nagtegaal,I.D., Oudejans,J., *et al.* (2010) Receptor conversion in distant breast cancer metastases. *Breast Cancer Res. BCR*, **12**, R75.

483. Buenrostro,J.D., Giresi,P.G., Zaba,L.C., Chang,H.Y. and Greenleaf,W.J. (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods*, **10**, 1213–1218.

484. Mohammed,H., Taylor,C., Brown,G.D., Papachristou,E.K., Carroll,J.S. and D'Santos,C.S. (2016) Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat. Protoc.*, **11**, 316–326.

485. Mohammed,H., D'Santos,C., Serandour,A.A., Ali,H.R., Brown,G.D., Atkins,A., Rueda,O.M., Holmes,K.A., Theodorou,V., Robinson,J.L.L., *et al.* (2013) Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. *Cell Rep.*, **3**, 342–349.

486. Seol,W., Hanstein,B., Brown,M. and Moore,D.D. (1998) Inhibition of estrogen receptor action by the orphan receptor SHP (short heterodimer partner). *Mol. Endocrinol. Baltim. Md*, **12**, 1551–1557.

487. Johansson,L., Thomsen,J.S., Damdimopoulos,A.E., Spyrou,G., Gustafsson,J.A. and Treuter,E. (1999) The orphan nuclear receptor SHP inhibits agonist-dependent transcriptional activity of estrogen receptors ERalpha and ERbeta. *J. Biol. Chem.*, **274**, 345–353.

488. Ross-Innes,C.S., Brown,G.D. and Carroll,J.S. (2011) A co-ordinated interaction between CTCF and ER in breast cancer cells. *BMC Genomics*, **12**, 593.

489. Fiorito,E., Sharma,Y., Gilfillan,S., Wang,S., Singh,S.K., Satheesh,S.V., Katika,M.R., Urbanucci,A., Thiede,B., Mills,I.G., *et al.* (2016) CTCF modulates Estrogen Receptor function through specific chromatin and nuclear matrix interactions. *Nucleic Acids Res.*, **44**, 10588–10602.

490. Diao,Y., Azatyan,A., Rahman,M.F.-U., Zhao,C., Zhu,J., Dahlman-Wright,K. and Zaphiropoulos,P.G. (2016) Blockade of the Hedgehog pathway downregulates estrogen receptor alpha signaling in breast cancer cells. *Oncotarget*, **7**, 71580–71593.

491. Khaleque,M.A., Bharti,A., Gong,J., Gray,P.J., Sachdev,V., Ciocca,D.R., Stati,A., Fanelli,M. and Calderwood,S.K. (2008) Heat shock factor 1 represses estrogen-dependent transcription through association with MTA1. *Oncogene*, **27**, 1886–1893.

492. Rigaud,G., Roux,J., Pictet,R. and Grange,T. (1991) In vivo footprinting of rat TAT gene: dynamic interplay between the glucocorticoid receptor and a liver-specific factor. *Cell*, **67**, 977–986.

493. Li,G., Margueron,R., Hu,G., Stokes,D., Wang,Y.-H. and Reinberg,D. (2010) Highly compacted chromatin formed in vitro reflects the dynamics of transcription activation in vivo. *Mol. Cell*, **38**, 41–53.

494. Swinstead,E.E., Miranda,T.B., Paakinaho,V., Baek,S., Goldstein,I., Hawkins,M., Karpova,T.S., Ball,D., Mazza,D., Lavis,L.D., *et al.* (2016) Steroid Receptors Reprogram FoxA1 Occupancy through Dynamic Chromatin Transitions. *Cell*, **165**, 593–605.

495. King,H.W. and Klose,R.J. (2017) The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. *eLife*, **6**.

496. Singhal,N., Graumann,J., Wu,G., Araúzo-Bravo,M.J., Han,D.W., Greber,B., Gentile,L., Mann,M. and Schöler,H.R. (2010) Chromatin-Remodeling Components of the BAF Complex Facilitate Reprogramming. *Cell*, **141**, 943–955.

497. Wang,L., Du,Y., Ward,J.M., Shimbo,T., Lackford,B., Zheng,X., Miao,Y., Zhou,B., Han,L., Fargo,D.C., *et al.* (2014) INO80 facilitates pluripotency gene activation in embryonic stem cell self-renewal, reprogramming, and blastocyst development. *Cell Stem Cell*, **14**, 575–591.

498. Goardon,N., Lambert,J.A., Rodriguez,P., Nissaire,P., Herblot,S., Thibault,P., Dumenil,D., Strouboulis,J., Romeo,P.-H. and Hoang,T. (2006) ETO2 coordinates cellular proliferation and differentiation during erythropoiesis. *EMBO J.*, **25**, 357–366.

499. Guseva,N.V., Rokhlin,O.W., Bair,T.B., Glover,R.B. and Cohen,M.B. (2012) Inhibition of p53 expression modifies the specificity of chromatin binding by the androgen receptor. *Oncotarget*, **3**, 183–194.

500. Bochkis,I.M., Schug,J., Ye,D.Z., Kurinna,S., Stratton,S.A., Barton,M.C. and Kaestner,K.H. (2012) Genome-wide location analysis reveals distinct transcriptional circuitry by paralogous regulators Foxa1 and Foxa2. *PLoS Genet.*, **8**, e1002770.

501. Minoo,P., Hu,L., Zhu,N., Borok,Z., Bellusci,S., Groffen,J., Kardassis,D. and Li,C. (2008) SMAD3 prevents binding of NKX2.1 and FOXA1 to the SpB promoter through its MH1 and MH2 domains. *Nucleic Acids Res.*, **36**, 179–188.

502. Ji,L., Xu,J., Liu,J., Amjad,A., Zhang,K., Liu,Q., Zhou,L., Xiao,J. and Li,X. (2015) Mutant p53 promotes tumor cell malignancy by both positive and negative regulation of the transforming growth factor β (TGF-β) pathway. *J. Biol. Chem.*, **290**, 11729–11740.

503. Harvey,M., Vogel,H., Morris,D., Bradley,A., Bernstein,A. and Donehower,L.A. (1995) A mutant p53 transgene accelerates tumour development in heterozygous but not nullizygous p53-deficient mice. *Nat. Genet.*, **9**, 305–311.

504. Liu,G., McDonnell,T.J., Montes de Oca Luna,R., Kapoor,M., Mims,B., El-Naggar,A.K. and Lozano,G. (2000) High metastatic potential in mice inheriting a targeted p53 missense mutation. *Proc. Natl. Acad. Sci. U. S. A.*, **97**, 4174–4179.

505. Willis,A., Jung,E.J., Wakefield,T. and Chen,X. (2004) Mutant p53 exerts a dominant negative effect by preventing wild-type p53 from binding to the promoter of its target genes. *Oncogene*, **23**, 2330–2338.

506. Oren,M. and Rotter,V. (2010) Mutant p53 gain-of-function in cancer. *Cold Spring Harb. Perspect. Biol.*, **2**, a001107.

507. Wiech,M., Olszewski,M.B., Tracz-Gaszewska,Z., Wawrzynow,B., Zylicz,M. and Zylicz,A. (2012) Molecular mechanism of mutant p53 stabilization: the role of HSP70 and MDM2. *PloS One*, **7**, e51426.

508. Di Agostino,S., Strano,S., Emiliozzi,V., Zerbini,V., Mottolese,M., Sacchi,A., Blandino,G. and Piaggio,G. (2006) Gain of function of mutant p53: the mutant p53/NF-Y protein complex reveals an aberrant transcriptional mechanism of cell cycle regulation. *Cancer Cell*, **10**, 191–202.

509. Stambolsky,P., Tabach,Y., Fontemaggi,G., Weisz,L., Maor-Aloni,R., Siegfried,Z., Sigfried,Z., Shiff,I., Kogan,I., Shay,M., *et al.* (2010) Modulation of the vitamin D3 response by cancer-associated mutant p53. *Cancer Cell*, **17**, 273–285.

510. Barretina,J., Caponigro,G., Stransky,N., Venkatesan,K., Margolin,A.A., Kim,S., Wilson,C.J., Lehár,J., Kryukov,G.V., Sonkin,D., *et al.* (2012) The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, **483**, 603–607.

511. Rasti,M., Arabsolghar,R., Khatooni,Z. and Mostafavi-Pour,Z. (2012) p53 Binds to estrogen receptor 1 promoter in human breast cancer cells. *Pathol. Oncol. Res. POR*, **18**, 169–175.

512. Arabsolghar,R., Azimi,T. and Rasti,M. (2013) Mutant p53 binds to estrogen receptor negative promoter via DNMT1 and HDAC1 in MDA-MB-468 breast cancer cells. *Mol. Biol. Rep.*, **40**, 2617–2625.

513. Yu,X., Vazquez,A., Levine,A.J. and Carpizo,D.R. (2012) Allele-specific p53 mutant reactivation. *Cancer Cell*, **21**, 614–625.

514. Wang,Z.-T., Chen,Z.-J., Jiang,G.-M., Wu,Y.-M., Liu,T., Yi,Y.-M., Zeng,J., Du,J. and Wang,H.-S. (2016) Histone deacetylase inhibitors suppress mutant p53 transcription via HDAC8/YY1 signals in triple negative breast cancer cells. *Cell. Signal.*, **28**, 506–515.

515. Reed,S.M. and Quelle,D.E. (2014) p53 Acetylation: Regulation and Consequences. *Cancers*, **7**, 30–69.

516. Eckert,R.L., Mullick,A., Rorke,E.A. and Katzenellenbogen,B.S. (1984) Estrogen receptor synthesis and turnover in MCF-7 breast cancer cells measured by a density shift technique. *Endocrinology*, **114**, 629–637.

517. Welsbie,D.S., Xu,J., Chen,Y., Borsu,L., Scher,H.I., Rosen,N. and Sawyers,C.L. (2009) Histone deacetylases are required for androgen receptor function in hormone-sensitive and castrate-resistant prostate cancer. *Cancer Res.*, **69**, 958–966.

518. Wang,C., Fu,M., Angeletti,R.H., Siconolfi-Baez,L., Reutens,A.T., Albanese,C., Lisanti,M.P., Katzenellenbogen,B.S., Kato,S., Hopp,T., *et al.* (2001) Direct acetylation of the estrogen receptor alpha hinge region by p300 regulates transactivation and hormone sensitivity. *J. Biol. Chem.*, **276**, 18375–18383.

519. Lerin,C., Rodgers,J.T., Kalume,D.E., Kim,S., Pandey,A. and Puigserver,P. (2006) GCN5 acetyltransferase complex controls glucose metabolism through transcriptional repression of PGC-1alpha. *Cell Metab.*, **3**, 429–438.

520. Chen,H., Lin,R.J., Xie,W., Wilpitz,D. and Evans,R.M. (1999) Regulation of hormone-induced histone hyperacetylation and gene activation via acetylation of an acetylase. *Cell*, **98**, 675–686.

521. Yamaguchi,N., Ito,E., Azuma,S., Honma,R., Yanagisawa,Y., Nishikawa,A., Kawamura,M., Imai,J.-I., Tatsuta,K., Inoue,J.-I., *et al.* (2008) FoxA1 as a lineage-specific oncogene in luminal type breast cancer. *Biochem. Biophys. Res. Commun.*, **365**, 711–717.

522. Johnston,S.J., Holmes,K.A. and Carroll,J.S. (2016) Abstract 2906: Targeting FoxA1 in estrogen receptor-positive breast cancer: biological characterization of kinase regulators. *Cancer Res.*, **76**, 2906–2906.

523. Liu,Y.-N., Lee,W.-W., Wang,C.-Y., Chao,T.-H., Chen,Y. and Chen,J.H. (2005) Regulatory mechanisms controlling human E-cadherin gene expression. *Oncogene*, **24**, 8277–8290.