IL6/STAT3 signaling hijacks ER enhancers to drive breast cancer metastasis

Rasmus Siersbæk[1,*,#], Valentina Scabia[2], Sankari Nagarajan[1], Igor Chernukhin[1], Evangelia K. Papachristou[1], Rebecca Broome[1], Simon J. Johnston[3,§], Stacey E. P. Joosten[4], Andrew R. Green[3], Sanjeev Kumar[1,5], Julia Jones[1], Soleilmane Omarjee[1], Ruben Alvarez-Fernandez[1], Silvia Glont[1], Sarah J. Aitken[1,6,7], Kamal Kishore[1], Danya Cheeseman[1], Emad A. Rakha[3], Clive D'Santos[1], Wilbert Zwart[4,8], Alasdair Russell[1], Cathrin Brisken[2], Jason S. Carroll[1,9*]

[1] Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge CB2 0RE, United Kingdom.

[2] ISREC – Swiss Institute for Experimental Cancer Research, School of Life Sciences, Ecole Polytechnique fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

[3] Nottingham Breast Cancer Research Centre, Division of Cancer and Stem Cells, School of Medicine, University of Nottingham Biodiscovery Institute, University Park, Nottingham NG7 2RD, United Kingdom.

[4] Division of Oncogenomics, Oncode Institute, Netherlands Cancer Institute, Amsterdam, The Netherlands.

[5] Addenbrookes Hospital, Cambridge CB2 0QQ, United Kingdom.

[6] Department of Histopathology, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Cambridge, CB2 0QQ, United Kingdom.

[7] Department of Pathology, University of Cambridge, Cambridge CB2 1QP, United Kingdom.

[8] Laboratory of Chemical Biology and Institute for Complex Molecular Systems, Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands.

[9] Lead Contact

#Current address: University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark.

§Current address: Translational Medicine, AstraZeneca, Cambridge, CB2 0AA, UK.

*correspondence: siersbaek@bmb.sdu.dk, Jason.carroll@cruk.cam.ac.uk

Summary

The cytokine interleukin 6 (IL6) and its downstream effector STAT3 constitute a key oncogenic pathway, which has been thought to be functionally connected to estrogen receptor α (ER) in breast cancer. We demonstrate that IL6/STAT3 signaling drives metastasis in ER^+ breast cancer independent of ER. STAT3 hijacks a subset of ER enhancers to drive a distinct transcriptional program. Although these enhancers are shared by both STAT3 and ER, IL6/STAT3 activity is refractory to standard ER targeted therapies. Instead, inhibition of STAT3 activity using JAK inhibitor ruxolitinib decreases breast cancer invasion *in vivo*. Therefore, IL6/STAT3 and ER oncogenic pathways are functionally decoupled highlighting the potential of IL6/STAT3-targeted therapies in ER^+ breast cancer.

Significance

The ubiquitous oncogenic IL6/STAT3 pathway has been linked to cancer progression in many solid tumors, and it is therefore a prime candidate for the development of anti-cancer therapies. In breast cancer, IL6/STAT3 signaling has previously been linked ER function. Here, we demonstrate that IL6/STAT3 signaling induces a more invasive and metastatic phenotype in breast cancer independent of ER and its pioneer factor FOXA1 by hijacking shared ER-FOXA1-STAT3 enhancers. This effectively decouples these two oncogenic pathways and highlights the potential for targeting the IL6/STAT3 pathway in ER^+/$IL6^+$/$pSTAT3^+$ breast cancer to inhibit this microenvironment-driven pathway that is not targeted by current therapies.

Introduction

Interleukin-6 (IL6) is a pleiotropic cytokine that plays a central role in both normal human physiology and disease (Hunter and Jones, 2015; Mauer et al., 2015). IL6 binding to its receptor complex IL6R/gp130, activates downstream Janus kinases (JAKs), which subsequently activate Signal transducer and activator of transcription 3 (STAT3) through phosphorylation of Tyrosine 705 (Mauer et al., 2015). IL6/STAT3 signaling has been shown to play an important role in tumor progression in many solid tumor types (e.g. breast cancer and lung cancer) by inducing epithelial-to-mesenchymal transition (EMT) and angiogenesis (Johnson et al., 2018). Thus, the IL6/STAT3 pathway is an attractive pharmacological target in oncology, and multiple different approaches to target this pathway are being pursued pre-clinically as well as clinically including targeting of upstream JAKs (e.g. ruxolitinib), direct targeting of STAT3 phosphorylation and activation (e.g. OPB compounds) and downregulating STAT3 expression (e.g. AZD9150) (Huynh et al., 2018; Wong et al., 2017).

The estrogen receptor α (ER) is a key driver of tumor growth and metastasis in ER^+ breast cancer, and anti-estrogen treatments (e.g. aromatase inhibitors, tamoxifen and fulvestrant) are the standard of care for patients with this disease (Matutino et al., 2018). ER cooperates with several other transcription factors to control gene expression and ultimately tumor growth (Siersbaek et al., 2018), most notably the forkhead transcription factor FOXA1, which is a pioneer factor important for facilitating ER-chromatin association (Hurtado et al., 2011). IL6 has been implicated as a driver in pre-clinical models of ER^+ breast cancer (Casneuf et al., 2016; De Luca et al., 2012; Ibrahim et al., 2016), and high IL6 serum and tumor levels have been associated with aggressiveness and poor outcome in patients (Bachelot et al., 2003; Jiang et al., 2017; Salgado et al., 2003). It has been reported that ER and IL6/STAT3 can modulate the transcriptional activity of each other in luciferase assays (Speirs et al., 2000; Yamamoto et al., 2000) presumably directly by forming a

complex (Binai et al., 2010; Yamamoto et al., 2000). IL6/STAT3 signaling and ER function are therefore thought to be functionally connected. Here, we investigate the transcriptional mechanisms through which IL6/STAT3 signaling controls breast cancer metastasis and its functional connection to ER and FOXA1.

Results

IL6/STAT3 signaling drives tumor growth and metastasis in an intraductal breast cancer xenograft model

To first establish the functional importance of IL6 signaling for tumor progression in ER^+ breast cancer, we took advantage of a recently developed mouse xenograft model using intraductal injection of human cancer cells into the mammary ducts (Fiche et al., 2019; Sflomos et al., 2016). This mouse intraductal (MIND) model recapitulates the different stages of breast cancer development seen in the clinic, including metastasis to clinically relevant distant organs such as liver, lung, bone and brain. Since mouse IL6 does not bind to the human receptor (Hammacher et al., 1994), we generated luciferase-expressing T47D cells that also express human IL6 (Figure S1A), which induces constitutive phosphorylation and activation of STAT3 (Figure S1B), and injected these cells intraductally into the mouse mammary glands of NSG females. IL6 increased phosphorylation of STAT3 in the primary tumors as expected (Figure 1A and Figure S1C) to a level comparable to that seen in $ER^+/pSTAT3^+$ patients (Figure S1D). Interestingly, IL6 expression significantly increased the growth rate of the primary tumor, relative to non-IL6 expressing control MIND tumors (Figure 1B and S1E). Importantly, activation of the IL6/STAT3 pathway significantly increased metastasis to the lungs by ~5-fold, and produced a trend towards increased metastasis in the other organs (Figure 1C,D and S1F). This demonstrates that in a clinically

relevant xenograft model, activation of the IL6/STAT3 pathway in ER^+ tumors promotes metastatic spread. Consistently, activation of the IL6-STAT3 pathway, as indicated by positivity for IL6 and $pSTAT3^{Y705}$ by immunohistochemistry (IHC) (Ahmad et al., 2018; Aleskandarany et al., 2016), is associated with lower breast cancer-specific survival (BCSS, 54.2 ($IL6^+/pSTAT3^+$) versus 85.3 (all other cases) months; Log-rank p = 0.022) and overall survival (OS, 47.2 ($IL6^+/pSTAT3^+$) versus 73.6 (all other cases) months; Log-rank p = 0.003) in patients (All patients: Figure 1E,F; ER^+ patients: Figure S1G,H).

Activated STAT3 associates with the ER/FOXA1 complex on chromatin

To understand the mechanisms underlying IL6-induced activation of tumor progression, we performed quantitative multiplexed rapid immunoprecipitation mass spectrometry of endogenous proteins (qPLEX-RIME), an unbiased protein interactome method we recently developed (Papachristou et al., 2018). Strikingly, ER qPLEX-RIME under full estrogenic conditions demonstrated that increased association between ER and STAT3 was the only significant change in the ER complex following IL6-induced phosphorylation of STAT3 (Figure 2A and Figure S2A,B). In contrast to ovarian cancer (Wang et al., 2015), Ser167 phosphorylation of ER was not induced by IL6 (Figure S2C). qPLEX-RIME of STAT3 showed that IL6 also increased the association between STAT3 and several well-known ER-associated factors such as FOXA1 and nuclear receptor coactivator 3 (NCOA3) (Figure 2A and Figure S2B). We could validate an increased association between ER and (p)STAT3 in response to IL6 using proximity ligation assay (PLA) (Figure S2D). Importantly, ER qPLEX-RIME in nine patient tumor samples (five primary ER^+ tumors and four ER^+ pleural effusions – see Table S1) demonstrated that STAT3 was a robust ER-associated protein also in clinical breast cancer samples (Figure 2B and Figure S2B). There

was no difference in this association between primary solid tumors and pleural effusions (Figure S2E), indicating that ER and STAT3 associate both in primary and metastatic cancer cells.

To understand the association between ER/FOXA1 and STAT3 at the genomic level, we performed chromatin immunoprecipitation coupled with deep sequencing (ChIP-seq) of ER, pSTAT3, total STAT3 and FOXA1 in response to short-term IL6 stimulation under full estrogenic conditions in T47D cells (Figure 3A and Figure S3A). IL6 induced extensive binding of pSTAT3 and total STAT3 to chromatin, whereas there was very little binding in the absence of IL6, demonstrating the necessity for exogenous IL6 for recruitment of STAT3 to chromatin (Figure 3A and Figure S3B). Consistent with the qPLEX-RIME data, pSTAT3 and total STAT3 chromatin binding in response to IL6 overlapped extensively with both ER and FOXA1 binding (Figure 3A and Figure S3C). Interestingly, binding of ER and FOXA1 was induced at a subset of sites shared between pSTAT3 and ER in response to IL6 (Figure 3B). These putative *cis*-regulatory elements were enriched for the STAT3 motif (Figure S3D), indicating that STAT3 likely mediates recruitment of ER/FOXA1 to these sites. Importantly, the extensive overlap in genomic binding of STAT3, ER and FOXA1 was validated in MCF7 cells (Figure S4A and S4B) and two $ER^+/pSTAT3^+$ primary breast tumors (Figure 3C).

IL6/STAT3 signaling controls a distinct gene program through shared STAT3-ER-FOXA1 binding sites

To link these binding events to transcriptional regulation, we profiled differential gene expression events in response to IL6 under full estrogenic conditions using RNA-seq in both T47D and MCF7 cells (Figure 4A and Figure S5A). Most of the genes in both cell lines changing significantly in response to one-hour IL6 stimulation are induced (T47D: 90%, 328 genes; MCF7: 96%, 67 genes;

Figure 4A and Figure S5A), consistent with the global recruitment of the transcriptional activator STAT3 to chromatin in response to IL6 (Figure 3A and S3B). Gene Set Enrichment Analysis (GSEA) showed that although some of the induced genes are known estrogen responsive genes, including *MYC* and X-box binding protein 1 (*XBP1*), most of the IL6 induced genes belong to different gene pathways (Figure 4A, right). This includes genes linked to an aggressive and more metastatic breast cancer phenotype, including B-cell lymphoma 3-encoded protein (*BCL3*) (Wakefield et al., 2013), intracellular adhesion molecule 1 (*ICAM1*) (Di et al., 2016) and vascular endothelial growth factor A (*VEGFA*) (Kim et al., 2017). Importantly, binding sites shared between pSTAT3, ER and FOXA1 were the most enriched sites adjacent to the IL6 target genes (Figure 4B and Figure S5B), suggesting that *cis*-regulatory elements near IL6 target genes are more likely to be co-bound by all three transcription factors.

Correlation of RNA-seq and Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq) signal within 500 kb across multiple TCGA patient tumors has recently allowed Corces et al. to infer direct target genes controlled by regulatory regions in patients without relying solely on gene proximity (Corces et al., 2018). We took advantage of these established promoter-enhancer connections to infer direct target genes controlled by pSTAT3-ER binding sites. Importantly, ATAC-seq peaks in patient breast tumors that are occupied by both pSTAT3 and ER in T47D cells or patients are functionally connected to the control of IL6 target gene expression through promoter-enhancer connections more than expected by chance (Figure 4C,D and Figure S5C, approximate random permutation p value <0.0001). In contrast, sites only bound by pSTAT3 do not show this enrichment (Figure 4C). Furthermore, shared pSTAT3-ER sites are more highly connected to regulation of IL6 target genes through expression quantitative trait loci (eQTL) than regions bound by either factor alone (Figure S5D). Taken together, this indicates that IL6 largely

signals through shared pSTAT3-ER sites to activate its target gene program, which is distinct from the classical estrogen responsive gene program.

IL6/STAT3 signaling induces the establishment of shared ER-FOXA1-STAT3 enhancers

The key role of shared ER-pSTAT3 binding sites in driving the IL6 gene program led us to investigate the role of IL6/STAT3 signaling in establishing active regulatory regions at these shared sites. First, we assessed changes in chromatin accessibility in response to IL6 using ATAC-seq and compared this to the effect of depleting FOXA1, which is known to be involved in chromatin remodeling at ER binding sites (Hurtado et al., 2011). In response to short-term IL6 treatment, we identify a remarkable dynamic response with 8,723 and 3,159 sites showing increased and decreased accessibility, respectively (Figure S6A,B). This is comparable to the extent of chromatin remodeling observed in response to two days of FOXA1 knock down (7,718 and 8,529 sites with increased and decreased accessibility, respectively, in the IL6 condition (Figure S6A,C)). Interestingly, IL6-induced chromatin accessibility is associated with increased binding of ER and FOXA1 in addition to pSTAT3 at shared sites, but chromatin remodeling is largely independent on the presence of FOXA1 (Figure 5A,B and S6D). Strikingly, ER binding to these IL6-dependent sites is also not dependent on FOXA1 (Figure 5B, top), which has previously been shown to be important for most classical estrogen-dependent ER binding events (Hurtado et al., 2011) (Figure 5B, bottom). This demonstrates that IL6/STAT3 signaling induces ER and FOXA1 binding at a subset of binding sites, where FOXA1 is not a pioneer factor. In contrast, chromatin accessibility and binding of ER and pSTAT3 at a distinct subset of shared sites is highly dependent on FOXA1, but shows limited response to IL6 treatment, despite robust IL6-induced pSTAT3 binding (Figure 5B and S6D). Interestingly, IL6-dependent and FOXA1-dependent

binding sites are highly enriched for the STAT3 and FOXA1 motifs, respectively (Figure S6E), indicating that the selective initiating function of these factors is at least partly dictated by the DNA sequence. Importantly, profiling of the well-established active chromatin mark H3K27ac (Creyghton et al., 2010) together with the mostly distal location of the shared binding sites (Figure S6F) shows that IL6-induced chromatin remodeling is associated with robust enhancer activation independent of FOXA1 (Figure 5B, top). Similarly, IL6 has no effect on H3K27ac at sites dependent on the pioneering function of FOXA1 (Figure 5B, bottom). This further demonstrates that distinct initiating functions of IL6/STAT3 and FOXA1 induce activation of distinct enhancers.

IL6/STAT3 hijacks shared ER-FOXA1-pSTAT3 enhancers to drive target genes independent of ER/FOXA1 function

The IL6-dependent shared ER-FOXA1-pSTAT3 binding sites defined above are more highly enriched near IL6-activated genes than FOXA1-dependent sites (Figure 5C), indicating that they directly control the IL6 target gene program. To determine the functional roles of STAT3, ER and FOXA1 for the transcriptional response to IL6, we therefore performed RNA-seq in response to IL6 stimulation following siRNA-mediated depletion of these factors in T47D cells (Figure S6A). As expected, STAT3 is critical for IL6-induced transcriptional activation (Figure 5D and S6G). This could be validated by RNA-seq in two $STAT3^{-/-}$ MCF7 clones generated using a CRISPR/Cas9 deletion approach (Figure 5E and S6H). FOXA1-dependent genes are largely distinct from IL6-dependent genes indicating that these factors drive different gene programs (Figure S6I). Consistently, FOXA1 is not required for the transcriptional response to IL6 (Figure 5D and S6G), which is consistent with the finding that FOXA1 is not required for activation of IL6-dependent shared binding sites (Figure 5B). Given the robust binding of ER to IL6-dependent

binding sites (Figure 5B), we hypothesized that ER rather than FOXA1 binding in response to IL6/STAT3-induced chromatin remodeling was important for transcriptional activation. Surprisingly however, ER is not required for the transcriptional activation of genes near IL6-dependent enhancers (Figure 5D) or the IL6 gene program as a whole (Figure S6G). To validate this, we used the potent and rapid ER-degrader fulvestrant in MCF7 cells. Consistent with our findings using siRNA-mediated knock down of ER in T47D cells, four-hour treatment with fulvestrant under full estrogenic conditions (Figure S6H) did not affect transcriptional activation of the IL6-induced genes (Figure 5E), although it robustly decreased the classical ER target gene program (Figure S6J). Thus, FOXA1 and ER are not required for the transcriptional response to IL6/STAT3 signaling. Similarly, FOXA1-induced genes near FOXA1-dependent shared binding sites do not respond to IL6, but depend on FOXA1 and ER (Figure 5F). Taken together, these results demonstrate that IL6/STAT3 and ER/FOXA1 drive distinct and functionally independent gene programs from shared regulatory regions.

Ruxolitinib inhibits the IL6/STAT3 target gene program and invasion in patient-derived tumors

Consistent with the functionally independent roles of ER and STAT3 at shared regulatory regions described above, pSTAT3 and ER are independent prognostic factors for BCSS and OS in $IL6^+$ patients as determined by multivariate analyses (Figure S7A,B). To investigate whether the IL6/STAT3 target gene program is also refractory to ER-targeted therapies in patients, we took advantage of matched pre- and post-treatment gene expression data from ER^+ breast cancer patients treated with the aromatase inhibitors anastrozole (Dunbier et al., 2013) and letrozole (Selli et al., 2019) in the neoadjuvant setting. Importantly, while around one third of Hallmark (Liberzon et al., 2015) estrogen responsive genes representing classical ER target genes (e.g. *PGR*, retinoic

acid receptor alpha (*RARA*), *TFF1*, growth regulation by estrogen in breast cancer 1 (*GREB1*), cyclin D1 (*CCND1*)) were repressed by aromatase inhibitor treatment in both data sets, IL6 target genes were largely unaltered (Figure 6A and S7C). To investigate whether pharmacological targeting of IL6/STAT3 signaling instead could inhibit this pathway to potentially benefit patients, we took advantage of a recently developed *ex vivo* explant system (Figure S7D) (Centenera et al., 2018) to investigate the response of ER^+/PR^- PDX tumors to the JAK inhibitor ruxolitinib. Ruxolitinib inhibited both basal and IL6-induced STAT3 phosphorylation in these explant models (Figure 6B,C and S7E,F). Importantly, ruxolitinib treatment also decreased expression of the proliferation maker Ki67 (Figure 6B,C) and expression of the IL6 target genes *MYC*, ETS-related transcription factor Elf3 (*ELF3*) and *BCL3* (Figure 6D).

To investigate how inhibition of IL6/STAT3 affects tumor progression in a clinically relevant setting of endocrine resistance *in vivo*, we took advantage of a highly resistant ER^+/PR^+ PDX, where the patient progressed in chemotherapy, aromatase inhibitors, fulvestrant and palbociclib (CDK4/6 inhibitor). Ruxolitinib treatment in this MIND PDX model significantly, but very modestly, decreases primary tumor growth (~20%, Figure 6E), despite very efficient on-target efficacy as shown by the absence of pSTAT3 upon ruxolitinib treatment (Figure 6F,G). This is similar to the modest response to fulvestrant (which the patient progressed on) or the combination treatment (~30% reduction in tumor growth), which are also associated with decreased Ki67 levels (Figure 6E-H). Interestingly however, ruxolitinib or the combination treatment dramatically decreased the ability of the tumor cells to invade through the myoepithelial layer and into the surrounding tissue (Figure 6I). This was not observed in mice treated with only fulvestrant despite efficient inhibition of ER function as indicated by reduced expression of the classical ER target gene and breast cancer biomarker PR (Figure 6G and S7G). This indicates that ruxolitinib can

efficiently inhibit progression of a highly endocrine-resistant tumor, particularly by specifically inhibiting the invasive potential of the cancer cells.

Discussion

We demonstrate that IL6/STAT3 signaling drives metastasis in MIND xenograft models of ER^+ breast cancer. Previous reports have shown that pSTAT3 is associated with good outcome in breast cancer (Aleskandarany et al., 2016; Sonnenblick et al., 2018), but here we show in the Nottingham cohort (Ahmad et al., 2018; Aleskandarany et al., 2016) that pSTAT3 is associated with poor outcome, when specifically focusing on patients with an $IL6^+$ tumor. This is consistent with our *in vivo* data and highlights the clinical importance of the IL6/STAT3 pathway.

Interestingly, IL6/STAT3 induces chromatin remodeling to establish shared ER-FOXA1-pSTAT3 enhancers independent of the pioneering activity of FOXA1. This demonstrates that pioneering activity is context-dependent and different factors can have pioneering or initiating function at different sites consistent with previous findings for steroid receptors (Paakinaho et al., 2019). Surprisingly, in contrast to previous reports suggesting functional crosstalk between IL6/STAT3 and ER (Speirs et al., 2000; Yamamoto et al., 2000), IL6/STAT3 drives a distinct oncogenic gene program independent of ER/FOXA1 function by hijacking these shared ER-FOXA1-pSTAT3 enhancers. Thus, ER and FOXA1 have limited functional importance and are largely silent passengers at these IL6/STAT3-driven enhancers. Consistent with independent roles of ER/FOXA1 and IL6/STAT3, we show that pSTAT3 and ER are independent prognostic factors in breast cancer. Importantly, IL6/STAT3 is therefore also resistant to endocrine therapy, yet responsive to IL6/STAT3-targeted therapy like ruxolitinib. This decouples these pathways and suggests that, although STAT3 and ER/FOXA1 share regulatory regions, they constitute

functionally independent oncogenic pathways (Figure 7). We therefore propose that treatment with IL6/STAT3-targeting agents will benefit $ER^+/IL6^+/pSTAT3^+$ breast cancer patients in the adjuvant or metastatic setting by inhibiting a tumor microenvironment-driven metastatic pathway that is not targeted by current endocrine therapies. In this regard, it is interesting to note the recent structure-based design of SD36 as a small molecule that specifically and efficiently induces proteosomal degradation of STAT3 protein to inhibit growth of lymphoma and leukemia xenografts (Bai et al., 2019). Furthermore, a STAT3 inhibitor, TTI-101, is being investigated in a recently established phase I clinical trial for safety and efficacy in multiple cancers including breast cancer (NCT03195699). While many of the IL6/STAT3 target genes described here have previously been linked to a more aggressive and metastatic phenotype (e.g. *BCL3* (Wakefield et al., 2013), *ICAM1* (Di et al., 2016) and *VEGFA* (Kim et al., 2017)), it will be important to determine how such IL6/STAT3-targeting drugs affect the IL6/STAT3 target gene program to control the breast cancer cell phenotype in the future.

Interestingly, resistance to CDK4/6 inhibitors, which are now standard of care for metastatic ER^+ patients, has been linked to repression of ER and activation of IL6/STAT3 signaling (Kettner et al., 2019). This suggests that the ER-independent IL6/STAT3 mechanism described here is active in CDK4/6 resistant patients supporting the rationale for exploring IL6/STAT3 inhibitors in these patients. Furthermore, given that IL6 is an adipokine that is increased in obesity (Bastard et al., 2000) and both IL6 and obesity are linked to metastasis of ER^+ breast cancers (Madeddu et al., 2014), the IL6/STAT3 pathway is a potential driver of obesity-driven ER^+ breast cancer. Our conclusions suggest that this is independent of the ER pathway and unaffected by ER targeted treatments, but sensitive to IL6/STAT3-targeted therapies.

Previous work has shown that tumor necrosis factor α (TNFα) and IL1β cytokine signaling can utilize the transcriptional activity of ER to activate target gene expression that ultimately regulates growth and metastasis (Franco et al., 2015; Stender et al., 2017). The mechanism through which IL6/STAT3 signaling controls breast cancer progression is therefore different from the mechanisms that have been described for these cytokines, since this pathway does not depend on ER function. Thus, we propose a model of paracrine signaling by cytokines in the tumor microenvironment, whereby IL6/STAT3 induces a pro-metastatic phenotype without functional convergence with the key oncogenic driver, a potential mechanism applicable for multiple cancer types associated with IL6/STAT3 signaling.

Acknowledgments

We thank Carlos Caldas and Alejandra Bruna for providing PDX material for the explant experiments and Kelly Holmes for generating the parental luciferase-mStrawberry expressing MCF7 cells used for generating $STAT3^{-/-}$ clones. We also thank the core facilities at Cancer Research UK (genomics, proteomics, histopathology, preclinical genome editing, bioinformatics, biorepository, flow cytometry, research instrumentation), in particular Cara Brodie from the histopathology core, for technical support. We thank the NKI Core Facility Molecular Pathology and Karianne Schuurman for technical support and Nottingham Health Science Biobank and Breast Cancer Now Tissue Bank for the provision of tissue samples. We would like to acknowledge the support of the University of Cambridge, Cancer Research UK and Hutchison Whampoa Limited. J.S.C. is funded by Cancer Research UK, an ERC Consolidator Award and a Komen Scholarship. R.S. is funded by the Novo Nordisk Foundation (NNF15OC0014136). C.B. received funding from the Swiss Cancer Ligue KFS-3701-08-2015 and SNF310030_179163. W.Z.

is supported by a KWF Dutch Cancer Society/ Alpe d'HuZes Bas Mulder Award (NKI 2014-6711), a KWF Dutch Cancer Society research grant (NKI 2015-7733) and a VIDI grant (016.156.401) from The Netherlands Organisation for Scientific Research (NWO). The Fusion Lumos Orbitrap mass spectrometer was purchased with the support from a Wellcome Trust Multi-user Equipment Grant (Grant # 108467/Z/15/Z).

Author contributions

Conceptualisation: R.S. and J.S.C. Investigation: R.S., V.S., S.N., E.K.P., R.B., S.J.J., A.R.G., S.E.P.J., S.K., J.J., S.O., R.A., S.G., S.J.A., D.C. Formal analysis: R.S., I.C., K.K., E.K.P., S.J.J., J.J., S.J.A. Visualization: R.S., V.S., I.C. Funding acquisition and supervision: R.S., E.A.R., C.S., A.R., W.Z., C.B., J.S.C. Writing – original draft: R.S. and J.S.C. Writing – review & editing: All authors.

Declaration of Interests

J.S.C. is the Founder and CSO of Azeria Therapeutics and S.J.J. has moved to AstraZeneca during the revision of this manuscript. The other authors declare no competing interests.

Main Figure Titles and Legends

Figure 1. IL6/STAT3 signaling increases primary tumor growth and metastasis.

(A) Representative H&E staining and immunohistochemistry (IHC) for phosphorylated STAT3 ($pSTAT3^{Y705}$) and ER in luciferase-expressing T47D primary MIND tumors expressing human IL6 or empty vector (EV). Scale bar = 50 μm. **(B)** Growth of the primary T47D MIND tumors as determined by changes in radiance relative to day 1 after cell injection. 5 mice per condition, 17 glands for EV and 15 glands for IL6. Mean of the fold change in radiance per gland is shown with error bars representing SEM. ****padj<0.0001 (two way ANOVA, Sidak's multiple comparisons test). **(C)** Relative metastatic burden in mice having primary tumors expressing IL6 or EV (n=5) as measured by radiance. Error bars represent SEM. ***padj<0.001 (two-way ANOVA, Sidak's multiple comparisons test). **(D)** Representative images showing radiance from micrometastases in bones and lungs used in panel C. **(E,F)** Kaplan-Meier (KM) plots showing the association between BCSS (E) or OS (F) and IHC staining of $pSTAT3^{Y705}$ and IL6 in breast tumors from the Nottingham cohort (Ahmad et al., 2018; Aleskandarany et al., 2016). Log-rank p values are indicated for $IL6^{+}/pSTAT3^{+}$ vs all other cases. See also Figure S1.

Figure 2. STAT3 associates with the ER/FOXA1 complex in response to activation by phosphorylation in cell lines and patients.

(A) ER and STAT3 qPLEX-RIME in MCF7 cells in response to 30 min treatment with recombinant human IL6. Four replicates of the ER or STAT3 RIME and one pooled IgG control RIME were included in each 9plex tandem mass tags experiment. Red points, padj<0.05, qPLEXanalyzer (Papachristou et al., 2018). **(B)** ER qPLEX-RIME in nine patient samples, i.e. five primary ER^{+} breast tumors and four ER^{+} pleural effusions. Non-specific proteins previously

identified in clinical breast tumors (Papachristou et al., 2018) were removed. In addition to ER (red) and STAT3 (blue), known ER-associated proteins are highlighted in light blue. The peptide coverage of STAT3 is shown in the insert. The schematic illustration was created with BioRender.com. See also Figure S2 and Table S1.

Figure 3. STAT3 chromatin binding overlaps extensively with binding of ER and FOXA1 in cell lines and patients.

(A) ChIP-seq in T47D cells in response to 30 min IL6 treatment (n=4). Groups of binding sites are defined based on the positional overlap between pSTAT3 and ER binding sites in the IL6 condition. Color bars indicate scale for normalized tag densities per 20 M subsampled reads. **(B)** Shared pSTAT3-ER binding sites from panel A are divided into constitutive or increased (padj<0.05, DiffBind (Stark and Brown, 2011)) ER binding sites. UCSC genome browser (http://genome.ucsc.edu, Dec. 2013 GRCh38/hg38) (Kent et al., 2002) screen shots for a representative gained and common ER binding site are shown to the right. **(C)** Positional overlap between ER and pSTAT3 chromatin binding based on ChIP-seq in ER^+/$pSTAT3^+$ primary breast tumors. The heatmaps show binding in the shared pSTAT3-ER sites for each patient. Color bars indicate scale for normalized tag densities per 20 M subsampled reads. The schematic illustration was created with BioRender.com. See also Figures S3 and S4.

Figure 4. Shared pSTAT3-ER-FOXA1 sites control the IL6-induced gene program.

(**A**) Left, differentially expressed genes (padj<0.05, DESeq2 (Love et al., 2014)) in response to one hour IL6 treatment of T47D cells as determined by RNA-seq (n=6). Right, significantly enriched Hallmark terms from MSigDB (Liberzon et al., 2015) in IL6-induced genes. **(B)** Enrichment of binding sites from T47D cells and patients in the vicinity of IL6-induced genes

compared to an equal number of constitutively expressed genes in T47D cells. Lines illustrate the cumulative percentage of sites within a given distance from the TSS. **(C)** Percentage of breast cancer ATAC-seq peaks occupied by both pSTAT3 and ER or only pSTAT3 in T47D cells that are connected to IL6 target gene promoters through promoter-enhancer connections (Corces et al., 2018) (Red line). The distribution of connections between the same number of random ATAC-seq peaks and IL6 target gene promoters is shown for comparison in blue (10,000 iterations to estimate the distribution of the test statistic under the null hypothesis). *Approximate random permutation p value <0.0001. **(D)** Identified functional promoter-enhancer links between IL6-induced genes and pSTAT3-ER occupied breast cancer ATAC-seq peaks. See also Figure S5.

Figure 5. IL6/STAT3 signaling through pSTAT3-ER-FOXA1 enhancers is functionally independent of ER/FOXA1.

(A) Venn diagram showing the overlap between IL6-dependent (1,282 sites) and FOXA1-dependent (2,303 sites) ATAC-seq peaks (n=3). IL6-dependent sites were defined as those with padj<0.05 (DiffBind (Stark and Brown, 2011)) comparing siControl-IL6 to siControl-veh. FOXA1-dependent sites were defined as those with padj<0.05 comparing siFOXA1-IL6 to siControl-IL6. **(B)** ATAC-seq and ChIP-seq for pSTAT3, ER, FOXA1 and H3K27ac in IL6- and FOXA1-dependent sites from panel A (n=3-4). FOXA1 ChIP-seq is from Figure 3A. **(C)** Enrichment of IL6- and FOXA1-dependent sites from panel A in the vicinity of IL6-induced genes compared to constitutively expressed genes. Lines illustrate the cumulative percentage of sites within a given distance from the TSS. **(D)** Gene expression as reads per kb per million (RPKM) for IL6-induced genes with TSS within 50 kb of IL6-dependent ATAC-seq peaks (113 genes) in response to IL6 and STAT3, ER or FOXA1 knock down (n=4). *$p<10^{-6}$, **$p<10^{-9}$ (two-sided

Wilcoxon test comparing veh and IL6 treatment for the different knock downs). Boxplot definitions: Center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; any outliers have been hidden. Notch shows the approximate 95% confidence interval for the median. **(E)** Regulation of IL6-induced genes (1 h treatment, padj<0.1, DESeq2 (Love et al., 2014)) in STAT3$^{-/-}$ MCF7 cells and in response to fulvestrant treatment (100 nM for 3 h prior to the start of the IL6 treatment) in WT cells (n=4). **(F)** Expression of FOXA1-induced genes with TSS within 50 kb of FOXA1-dependent ATAC-seq peaks (369 genes) as in panel D. *p<10^{-4}, **p=$2.2x10^{-308}$ (two-sided Wilcoxon test comparing veh and IL6 in siControl to veh and IL6 in the different knock downs, respectively). See also Figure S6.

Figure 6. IL6/STAT3 signaling is resistant to endocrine therapy, but the JAK inhibitor ruxolitinib inhibits activation of STAT3 and invasion in resistant PDX.

(A) Regulation of Hallmark early estrogen responsive genes (n=200) (Liberzon et al., 2015) and IL6-induced genes (n=328 for T47D and n=67 for MCF7) in matched samples from 42 patients treated with letrozole for at least four months in the neoadjuvant setting (Selli et al., 2019). The mean regulation of 200 random genes is included for comparison with the error bar representing the standard deviation of 1,000 iterations. *p=1.8e-07 and **p< 2.2e-16 (Fisher's exact test, two-sided). The schematic illustration was created with BioRender.com. **(B)** Representative H&E, pSTAT3^{Y705} and Ki67 IHC staining for a WT ER^{+}/PR^{-} PDX (AB555P) (Bruna et al., 2016) explant treated with ruxolitinib (500 nM) for two days *ex vivo*. Scale bars = 50 μm. **(C)** Quantification of pSTAT3^{Y705} and Ki67 IHC staining in the PDX explant from panel B using the Halo TM (Indica labs) Image analysis platform. Error bars represent SEM between different tumor pieces (n=9-10) from different parts of the PDX tumor cultured on the same sponge. *p=0.099, **p=1.1e-06 (Student's t-test, two-sided). **(D)** Relative mRNA expression of *MYC*, *ELF3* and *BCL3* in the PDX

explant from panel B. Average dots per cell for these three genes and the housekeeping gene TATA-binding protein (*TBP*) were determined by RNAscope *in situ* hybridization. Error bars represent SEM between different tumor pieces (n=7-8) from different parts of the PDX tumor cultured on the same sponge. *p= 0.098, **p=0.0074, ***p=0.00013 (Student's t-test, two-sided). **(E)** Primary tumor growth of endocrine-resistant intraductal PDX in response to fulvestrant (Fulv) or ruxolitinib (Ruxo) treatment (initiated one week after injection). Points represent average growth of tumors and error bars represent SEM. n=19, 22, 22, 16 tumors from 8 mice per treatment arm. *p<0.05, ****p<0.0001 (Two way ANOVA, followed by Tukey's multiple comparison test). Insert shows the relative fold change in radiance at endpoint. **(F)** Quantification of expression of protein markers in the PDX from panel E as determined by IHC. Error bars represent SEM, n=4-5. **(G)** H&E staining and representative images of protein marker expression in adjacent sections in the PDX from panel E. **(H)** Quantification of Ki67 expression from panel G. Error bars represent SEM, n=4-5. **(I)** Quantification of tumor invasion in the mammary gland in the PDX from panel E. The invasive areas of the tumor, characterized by loss of the myoepithelial layer of the duct and tumor cell invasion into the surrounding tissue, were quantified by automated cell detection, manually curated and annotated. The percentage area of invasion relative to total tumor area (invasive and *in situ*) was determined. Error bars represent SEM, n=4-5. The presence of one outlier in the fulvestrant arm was identified and removed using Graphpad Prism outlier function with Q=2%. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 (One way ANOVA, followed by Tukey's multiple comparison test). See also Figure S7.

Figure 7. Schematic model of IL6/STAT3 and ER/FOXA1 function in breast cancer.

The two potent oncogenic pathways, ER/FOXA1 and IL6/STAT3, drive breast tumor growth and metastasis through functionally independent mechanisms using shared enhancers. Created partly

with BioRender.com. FHRE, Forkhead response element; ERE, Estrogen response element; STAT-RE, STAT response element.

STAR Methods

Resource Availability

Lead Contact

Requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Jason S. Carroll (jason.carroll@cruk.cam.ac.uk).

Materials Availability

Cell lines generated in this study are available upon request.

Data and Code Availability

The ChIP-seq and RNA-seq datasets generated during this study are available at Gene Expression Omnibus (GEO) with accession GSE126006. The proteomics datasets generated during this study are available at Proteomics Identification Database (PRIDE) with accession PXD012980. The code supporting the current study is available upon request.

Experimental Model and Subject Details

Animals

Animal experiments were performed in accordance with protocol VD1865.4 approved by the Service de la Consommation et des Affaires Vétérinaires of Canton de Vaud. NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ mice (NSG) were purchased from Charles River.

Animals were housed in IVC cages – Green line – from Tecniplast®, type II long in polysulfone, with bottles in polysulfone as well (5 mice per cage). Cages were in over-pressure compared to the pressure in the housing room. Housing rooms were in over-pressure compared to the pressure

outside of the barrier unit. Bedding was provided by Aspen Tapvei® (little squares about 4 x 4 x 1 mm). Water was acidified (pH between 2.5 to 3) on resin column (Prominent® CH system) and diet was from Provimi-Kliba® (cat# 3242, irradiated). The cages were enriched with nesting material (tissues) and cardboard and wood tunnels. Housing room temperature was at 22 °C +/- 2 °C, humidity 55% +/- 10%. Twelve hours light cycle, 7 am to 7 pm.

Human tumors

ChIP-seq

Matched fresh-frozen and FFPE samples from early stage, primary, ER^+ invasive breast cancers from female patients undergoing surgical resection at the Nottingham City Hospital between 1992 and 1997 with high $pSTAT3^{Y705}$ expression (Aleskandarany et al., 2016) were sectioned according to institutional protocols. None of these patients received neoadjuvant therapy.

This study was approved by the North West – Greater Manchester Central Research Ethics Committee under the title; Nottingham Health Science Biobank (15/NW/0685). All patients provided written informed consent. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Release of data was also pseudo-anonymised as per the UK Human Tissue Act regulations.

qPLEX-RIME

Primary tumor material from female patients was obtained from the biobank at the Netherlands Cancer Institute and cryosectioned in 30x30 μm. Pleural effusions were collected from female patients treated at the same hospital and processed as described previously (Schrijver et al., 2018).

Pellets were stored at -80 °C without DMSO. In brief, the cells in the pleural effusions were isolated by centrifugation, exposed to an erythrocyte lysis buffer and remaining cells were then split for formalin-fixed and paraffin-embedded for histological assessment/staining as well as cryo-storage as a cell pellet in eppendorf tubes kept at -80 °C until further use. H&E's were sectioned from the FFPE blocks to assess tumor cell percentage. Pleural samples used for qPLEX-RIME contained 70-90% tumor cells, as assessed by a pathologist. Available patient information is shown in Table S1.

This study was performed in accordance with the institutional medical ethical guidelines. The use of anonymous or coded leftover material for scientific purposes is part of the standard treatment agreement with patients, and therefore, informed consent was not required according to Dutch law for patients not actively opting out.

PDX models for explants and MIND

PDX models for tumor explants were previously published with all required ethics approvals (Bruna et al., 2016). For the MIND PDX, the study of primary tumors was approved by the Commission cantonale d'éthique de la recherche sur l'être humain (CER-VD 38/15), patients signed an informed consent. Clinical information on patient age, menopausal status, primary tumor size, nodal status, histological type, ER, PR, Ki67, HER2 status, was collected from clinical charts and anonymized.

Cell lines

MCF7 cells (female, ATCC) were grown in DMEM (Gibco, 41966-029) supplemented with 10% fetal bovine serum (FBS – Gibco, 10500-064), 50 U/ml penicillin and 50 μg/ml streptomycin (Gibco, 15070-063) and 2 mM L-glutamine (Gibco, 25030). T47D cells (female, ATCC) were

grown in RPMI (Gibco, 21875-034) supplemented with 10% FBS, penicilin/streptomycin and L-glutamine as for MCF7 cells. Cells were routinely genotyped by short-tandem repeat (STR) genetic profiling using the PowerPlex 16HS Cell Line panel and analyzed using the Applied Biosystems Gene Mapper ID v3.2.1 software by the external provider Genetica DNA Laboratories (LabCorp Specialty Testing Group). Cells were mycoplasma tested (MycoProbe Mycoplasma detection kit, R&D) around every major experiment and at least every six months. All cell lines were grown at 37 ºC.

Method Details

Mouse intraductal xenografts experiments

T47D cells were subjected to lentiviral transduction with pLEX_TRC211/L2N-TurboRFP-c (Stock et al., 2016) driving both turboRFP and luciferase2 expression followed by lentiviral transduction with either empty vector or with IL6 expression vector as described below. The PDX was established from ascites effusion of a patient diagnosed with an ER^+/PR^+ NST stage IV cancer, that was treated with multiple lines of both chemotherapy and endocrine therapy during a period of 10 years.

The cells were injected intraductally in 3 glands per mouse on average, as previously described (Sflomos et al., 2016). Briefly, 10-week-old NSG females were anesthetized by intraperitoneal injection with 10 mg/kg xylazine and 90 mg/kg ketamine (Graeub) and injected into the cleaved teat with a blunt end Hamilton syringe (cat. No. HAMI80508), specifications: 50 µl 705N, gauge 30/13 mm/pst3) with 500,000 cells.

Tumor growth was measured weekly on the lateral sides of the mice by IVIS Spectrum (Perkin Elmer). To assess metastatic burden, organs were carefully collected; bones were deprived of muscle to avoid any background and the *ex vivo* bioluminescence was measured by IVIS.

Ruxolitinib (Adooq, cat# INCB018424) was prepared for dosing in 0.5% methylcellulose (400 cps) (Sigma-Aldrich, cat# M0262) in water. Formulations were prepared approximately biweekly as suspensions, even distribution was achieved by sonication at 4 °C for 30 min. Vehicle and test article suspensions were administered once daily by oral gavage at a dose volume of 10 ml/kg (Shuey et al., 2016). The last dose was administered 5 h before sacrifice.

Fulvestrant (cat# S1191, Selleckchem) was purchased already dissolved in 4% DMSO-castor oil solution at a concentration of 20 mg/ml. It was administered weekly by subcutaneous injection at a dose of 160 mg/kg. The last dose was administered 3 days before sacrifice.

Treatment of cells

Cells were treated in full media described above, rather than in media supplemented with charcoal stripped serum, to maintain full estrogenic conditions and normal growth factor and hormone signaling. This approach most accurately recapitulates the conditions in treatment naive patients. An estrogenic background is also necessary for studying ER function on chromatin in response to treatment.

Unless otherwise noted, cells were treated with 20 ng/ml human recombinant IL6 (R&D systems, 206-IL-200) dissolved in PBS+0.1% bovine serum albumin (BSA) by directly adding the required volume of IL6 stock solution to the media. Cells were treated for 30 min (ChIP-seq and ATAC-seq) or 1 h (RNA-seq) before cell harvest. Fulvestrant (SelleckChem Cat# S1191) treatment (100 nM in ethanol) was done for three hours prior to initiation of IL6 treatment.

Generation of MCF7 STAT3$^{-/-}$ by CRISPR

A clone from MCF7 cells stably transfected with a luciferase/mStrawberry construct with high expression of luciferase was identified. The growth, responsiveness to anti-estrogens and STR genotype of this clone was validated to confirm that the phenotype was similar to parental MCF7 cells. This provided a clonal MCF7 cell line (MCF7-luc) as a parental clone for the CRISPR-mediated knock out of STAT3, which circumvented confounding factors associated with the comparison of clonal knock out cell lines and multiclonal parental cells.

Four CRISPR guides were designed against Exons 3 and 4 of STAT3 (NM_139276). Sequences are available upon request. Oligos (Sigma Aldrich) were cloned into pSpCas9(BB)-2A-GFP (PX458, Addgene#48138), and constructs were transfected into the clonal MCF7-luc cell line. Three days post-transfection, GFP^{+} cells were FACS sorted into 6 well plates. After another eight days, mStrawberry^{+} cells were single cell sorted into 96 well plates by FACS and allowed to form colonies for ~3 weeks. STAT3 expression in these colonies was determined by In-Cell Western using a STAT3-specific antibody (Cell Signaling, 9139) and the CellTag 700 cellular stain (Li-Cor, 926-41090) to normalize cell number. Eight putative STAT3$^{-/-}$ clones were expanded and the loss of STAT3 was confirmed by western blotting using a different STAT3 antibody (Cell Signaling, 4904). The frameshift mutations induced by CRISPR were identified by targeted next generation sequencing (NGS) of an amplicon containing the guide RNA target site. The two most likely off-target sites for all but one guide, which had no predicted off-target sites (Hsu et al., 2013), were tested for off-target editing using the same NGS approach as described above. Two clones with heterozygotic frameshift mutations and no off-target mutations were used for subsequent experiments. The target frameshift mutations in these two clones were validated by

Sanger sequencing. The STAT3$^{-/-}$ clones were STR genotyped and mycoplasma tested as described under “Cell lines”.

siRNA-mediated knock down

T47D cells were transfected with ON-TARGETplus SMARTPools (Dharmacon) targeting STAT3 (L-003544-00-0020), ER (L-003401-00-0020) or FOXA1 (L-010319-00-0020) using Lipofectamine RNAiMAX (Invitrogen, 13778-150). A non-targeting pool (D-001810-10-20) was used as a control. The media was changed 24 h after transfection, and cells were treated for one hour with IL6 48 h after transfection.

Ectopic expression of luciferase and IL6

T47D cells were infected with lentivirus driving expression of luciferase2 and turboRFP (pLEX_TRC211/L2N-TurboRFP-c) (Stock et al., 2016) followed by FACS sorting to obtain a population of cells expressing high levels of turboRFP. These cells were then infected with lentivirus encoding human IL6 (EX-A0301-Lv197) or empty vector (EX-NEG-Lv197) followed by selection with blasticidin (20 μg/ml, Gibco, A11139-03) until non-infected controls cells had all died. Expression and release of IL6 into the media was validated by ELISA (see Figure S1A) and activation of STAT3 by phosphorylation downstream of IL6 signaling was validated by western blot (see Figure S1B). The T47D-luc cells expressing IL6 or empty vector were genotyped and mycoplasma tested as described under “Cell lines”.

Explant of PDX tumors

Fresh WT ER^+/PR^- PDX tumors (STG335 and AB555P) (Bruna et al., 2016) were cut into small (~2 mm^3) pieces and placed on Spongostan gelatin dental sponges (Ferrosan Medical Devices, MS0005) soaked in explant media (RPMI 1640 supplemented with 10% FBS, 2 mM L-glutamine (Gibco, 25030), 10 μg/ml hydrocortisone (Sigma, H0888), 1xinsulin-transferrin-selenium supplement (Sigma, I3146) and 1xAntibiotic Antimycotic Solution (Sigma, A5955)). For STG335 (Figure S7E,F), media was replaced after 24 h with media containing 50 ng/ml IL6 (R&D systems, 206-IL-200) and 500 nM ruxolitinib (Selleckchem, S1378) as indicated for two days. Treated media was refreshed every day and two hours before fixation in 10% neutral-buffered formalin for 24 h. For AB555P (Figure 6B-D), tumor pieces were placed on sponges soaked in media with or without ruxolitinib (500 nM) and incubated for two days without media change before fixation in 10% neutral-buffered formalin. Tumor sections were subjected to staining with hematoxylin and eosin (H&E), $pSTAT3^{Y705}$ IHC staining (Cell Signaling, 9145), Ki67 IHC staining (Dako, M7240) or RNAscope *in situ* hybridization.

IL6 ELISA

10^6 T47D-luc cells expressing IL6 or empty vector were seeded in a 10 cm dish. After three days of culture, the media was harvested, and the IL6 levels were determined by ELISA (EH2IL6, Thermo Fisher Scientific) according to the manufacturer's instructions.

qPLEX-RIME

Cell lines

qPLEX-RIME on cell lines was performed essentially as described previously (Papachristou et al., 2018). Briefly, cells were crosslinked at room temperature by incubating with 2 mM

disuccinimidyl glutarate (DSG) for 20 min followed by 1% formaldehyde for 10 min before quenching in 0.1 M glycine for 10 min. Cells were washed twice in cold PBS and harvested in cold PBS containing protease (Roche, Complete, 05056489001) and Halt phosphatase inhibitors (Thermo Scientific, 78427). Crosslinked cells were incubated at 4 °C with LB1 (50 mM Hepes–KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.5% NP-40/Igepal CA-630, 0.25% Triton X-100) for 10 min followed by 5 min in LB2 (10 mM Tris–HCL pH8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) before resuspending in LB3 (10 mM Tris–HCl pH 8, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na–Deoxycholate, 0.5% N-lauroylsarcosine). Chromatin was sonicated using the Bioruptor Plus (Diagenode) for 15 min (30 sec on/30 sec off) to generate DNA fragments of around 100-800 bp. Chromatin was immunoprecipitated overnight using protein A (Invitrogen, used for rabbit antibodies) or protein G (Invitrogen, used for goat antibodies) dynabeads with specific antibodies against ER (Santa-Cruz, sc-543 or mix of abcam, ab3575 and Millipore, 06-935), STAT3 (Santa Cruz, sc-482), pSTAT3^{Y705} (Cell Signaling, 9145), FOXA1 (abcam, ab5089) or H3K27ac (abcam, ab4729). Beads were washed 10 times in RIPA buffer (50 mM HEPES pH 7.6, 1 mM EDTA, 0.7% Na-deoxycholate, 1% NP-40, 0.5 M LiCl) followed by two washes in AMBIC (100 mM ammonium hydrogen carbonate). Washed beads were frozen at -20 °C.

LC-MS analysis and data processing

For the proteomic analysis, samples were prepared as previously described (Papachristou et al., 2018). Briefly, the bound proteins were digested on the beads overnight with the addition of trypsin at final concentration 15 ng/ul (Pierce) followed by a second digestion step the next day for 4 h. Peptide samples were cleaned and labelled with the TMT-10plex reagents (Thermo Fisher

Scientific) followed by fractionation using Reversed-Phase spin columns at high pH (Pierce #84868). LC-MS analysis was performed on the Dionex Ultimate 3000 UHPLC system coupled with the nano-ESI Fusion Lumos (Thermo Scientific) mass spectrometer. The Lumos was operated in a data-dependent mode for both the MS2 and the SPS-MS3 methods. Precursors were selected at 120k resolution and the collision energy was set at 35%. Peptides were isolated in the quadrupole with isolation window 0.7 Th. The top 10 most intense fragments were selected for Synchronous Precursor Selection (SPS) HCD-MS3 analysis with MS2 isolation window 2.0 Th. The HCD collision energy was set at 55% and the detection was performed with Orbitrap resolution 50k or 60k. The raw data were processed with the SequestHT search engine in Proteome Discoverer 2.1. The SequestHT node included the following parameters: Precursor Mass Tolerance 20ppm, Fragment Mass Tolerance 0.5 Da, Dynamic Modifications were Oxidation of M (+15.995 Da), Deamidation of N/Q (+0.984 Da) and Static Modifications were TMT6plex at any N-Terminus/K (+229.163 Da). Only unique peptides were used for quantification and the consensus workflow included calculation of TMT signal-to-noise. The percolator node with decoy database search was used to estimate the confidence level for peptide identifications. The peptide intensities were normalized using median scaling and aggregated (by summing) to protein intensities. qPLEX-RIME data was analyzed by the qPLEXanalyzer (10.18129/B9.bioc.qPLEXanalyzer) (Papachristou et al., 2018) to identify significant changes in protein interactions using limma based analysis (Ritchie et al., 2015). Data was visualized using ggplot2 (Wickham, 2016) in R.

Clinical samples

qPLEX-RIME on tumor material was performed essentially as described previously (Papachristou et al., 2018). Briefly, flash frozen breast cancer material from female patients was cryosectioned

into 30 μm sections, crosslinked by 2 mM DSG for 45 min total together with 1% formaldehyde for the last 20 min and quenched by 0.1 M glycine for 10 min at room temperature. After two washes in cold PBS, crosslinked sections were resuspended in LB3 and sonicated using a probe sonicator (Fisher Scientific) until most DNA fragments were 100-800 bp. Chromatin was immunoprecipitated, washed and processed for mass spectrometry analysis as described for cell lines. Specific ER-associated proteins were identified by removing non-specific proteins previously identified by ER/IgG qPLEX-RIME in human breast tumors (Papachristou et al., 2018). Non-specific proteins were defined as those with padj>0.05 or Log2 fold change ≤1 comparing ER and IgG samples. However, even with this filtering step a few unspecific proteins with high intensity remain. These include keratins, fibrinogen, ribonucleoproteins and a couple of abundant metabolic enzymes.

ChIP-seq

Cell lines and tumors

ChIP-seq was performed on three to four independent biological replicates. Chromatin was prepared from both cell lines/tumors and immunoprecipitated as described for qPLEX-RIME above. After washing of the beads, chromatin was eluted and decrosslinked by incubating over night at 65 °C in elution buffer (50 mM TrisHCl pH8, 10 mM EDTA, 1% SDS). Samples were treated with RNase A (20 ng/ml) for 30 min-1 h followed by proteinase K (200 ng/ml) for 1-2 h before DNA was purified by phenol-chloroform extraction. Purified DNA was subjected to library preparation using the SMARTer ThruPLEX® DNA-Seq Kit (TaKaRa, R400676) and DNA HT Dual Index Kit – 96N Set A (TaKaRa, R400660) followed by Illumina next generation sequencing to reach approximately 30 M reads per sample.

Analysis

50 bp single-end reads were mapped to hg38 genome using bowtie2 2.2.6 (Langmead and Salzberg, 2012). Aligned reads with mapping quality less than 5 were filtered out. The read alignments from all replicates were combined into single library and peaks were called with MACS2 version 2.1.1.2016 (Zhang et al., 2008) using input sequences as a background control. The peaks that absorbed statistically significant tag density from all replicates were selected for downstream analysis. MEME version 4.9.1 (Bailey et al., 2009) was used to detect known and discover novel binding motifs amongst tag-enriched sequences. For visualizing tag density and signal distribution heatmap the read coverage in a window of +/- 2.5 or 5 kb region flanking the peak midpoint was generated using a bin size of 1/100 of the window length. Differential binding analysis (Diffbind) was performed as described previously (Stark and Brown, 2011). Motif matrices for ER, FOXA1 and STAT3 from the JASPAR database (Mathelier et al., 2016) were used for motif enrichment analyses in ChIP-seq peaks relative to open chromatin as previously determined by ATAC-seq (Toska et al., 2017).

RNA-seq

Total RNA was purified using RNeasy kit (Qiagen, 74106) and mRNA libraries were prepared using the TruSeq stranded mRNA library preparation kit (Illumina, cat#20020595) and IDT for Illumina – TruSeq RNA UD Indexes (cat# 20022371). Samples were subjected to Illumina next generation sequencing to reach around 30 M reads per sample.

Analysis

50 bp single-end reads were aligned to the human genome version GRCh37 using STAR version 2.6.1a (Dobin et al., 2013). Differentially expressed genes were identified based on the negative

binomial distribution using the DESeq2 package (Love et al., 2014) (https://bioconductor.org/packages/release/bioc/html/DESeq2.html) version 1.14.1 in R. Heatmap visualisation, clustering and various plotting functions were used from statistical and bioinfo modules in MATLAB framework (https://www.mathworks.com).

ATAC-sequencing

T47D cells were transfected with siRNA targeting FOXA1 or control siRNA as described above (see "siRNA-mediated knock down"). After two days of knock down, cells were treated with IL6 (20 ng/ml) for 30 min before processing cells for ATAC-seq. Omni-ATAC-sequencing was performed as previously described (Corces et al., 2017). ATAC-seq resuspension buffer (RSB) was prepared as follows: For 50 ml of the buffer, 500 µl 1M Tris-HCl pH 7.4, 100 µl 5M NaCl, 150 µl 1M $MgCl_2$, and 49.25 ml sterile water were added. 50 µl cold ATAC-RSB containing 0.1% NP40, 0.1% Tween-20 and 0.01% Digitonin (Promega cat# G9441) was added to 50,000 viable cells and pipetted up and down three times. The cells were incubated on ice for 3 minutes to lyse the cells and washed out with 1 ml of cold ATAC-RSB containing 0.1% Tween-20, but without NP40 and digitonin. The tube was inverted three times to mix. The nuclei were pelleted at 500xg for 10 min at 4 ºC in a fixed angle centrifuge. The pellet was resuspended in 50 µl transposition mix (25 µl 2xTagment buffer, 2.5 µl transposase (100 nM final, Illumina Tagment DNA Enzyme and Buffer Small Kit, Cat. No. 20034197), 16.5 µl PBS, 0.5 µl 1% digitonin, 0.5 µl 10% Tween-20, 5 µl H_2O) and the reaction was incubated at 37 ºC for 30 min in a thermomixer with 1,000 RPM mixing. Zymo DNA Clean and Concentrator-5 Kit (cat# D4014) was used to clean up the reaction and DNA was eluted in 21 µl elution buffer and stored at -20 °C until amplification. 20 µl of the product was used for the following PCR. Preamplification was performed for 5 cycles

using 1 µl of i5 primer, 1 µl of i7 primer, 25 µl 2xQ5 HotStart NEBNext master mix (M0494L), and 20 µl transposed/cleaned-up sample. Thermocycler conditions (72 ºC for 5 min, 98 ºC for 30 sec, followed by 5 cycles of [98 ºC for 10 sec, 63 ºC for 30 sec, 72 ºC for 1 min] then hold at 4 ºC. Using 5 µl (10%) of the pre-amplified mixture, 15 µl qPCR was run to determine the number of additional cycles needed as follows: 3.85 µl sterile water, 0.2 µl Nextera XT Index kit (24 indexes, FC-131-1001) i5 primer, 0.2 µl i7 primer, 0.75 µl 20x EvaGreen (in DMSO), 5 µl 2xNEBNext master mix, 5 µl pre-amplified sample. Cycling conditions: 98 ºC for 30 sec, followed by 20 cycles of [98 ºC for 10 sec, 63 ºC for 30 sec, 72 ºC for 1 min]. After qPCR amplification, the amplification profiles were assessed to determine the required number of additional cycles to amplify. Final amplification and clean-up was performed on the remaining 45 µl of the pre-amplified DNA. Samples were size selected and purified using Ampure XP beads. Next generation sequencing was performed using Illumina Novaseq 6000 with SP flow cell system with 50 bp paired end reads to get around 40 M reads per sample.

ChIP-seq enrichment near regulated genes

The cumulative numbers of ChIP-Seq binding sites located within the 10 kb linear scale from the TSS of significantly up-regulated genes were determined. As a background control, binding sites near an equal number of random constantly expressed genes revealed by the RNA-seq analyses were tested in 100 permutations. Fraction of binding sites near these groups of genes were then visualized as curves including the standard deviation for the random constitutively expressed control genes as error bars.

Promoter-enhancer links to IL6-induced genes

Table S2 and S7 from (Corces et al., 2018) were downloaded. They identify open chromatin regions, as determined by ATAC-seq, and functional links between distal regulatory open chromatin regions and their target genes in TCGA tumors, respectively. Briefly, these promoter-enhancer links were established by correlating changes in accessibility to changes in expression of genes within 500 kb using 373 tumor samples with matched ATAC-seq and RNA-seq data (Corces et al., 2018).

All the ATAC-seq peaks from TCGA breast cancer samples that had a positional overlap with a pSTAT3-ER binding site from T47D cells or patient samples were identified by GenomicRanges (Lawrence et al., 2013) in R (R Core Team, 2020). They were functionally linked to their target genes using the pan-cancer links table (Corces et al., 2018). The promoter-enhancer links between pSTAT3-ER occupied breast cancer ATAC-seq peaks and IL6-induced genes from T47D cells were compared to the links of random ATAC-seq peaks using 1,000 iterations to show that pSTAT3-ER sites are functionally linked to IL6-induced genes more than expected by chance. Circlize (Gu et al., 2014) was used to visualize these links in a circular plot.

eQTL analyses

eQTL (GTEx_Analysis_v8_eQTL.tar) were downloaded from the GTEx portal (11/14/2019). eQTL linked to expression of genes significantly (padj<0.05) induced by IL6 in T47D cells (4,756 eQTL linked to 78 out of the 328 IL6-induced genes) were identified, and the number of binding sites overlapping at least one eQTL was determined by the GenomicRanges package (Lawrence et al., 2013) in R (R Core Team, 2020). The overlap between the same set of binding sites and eQTL linked to the same number of random expressed genes with eQTL information (n=78) was

included as control (1,000 iterations to the estimate the distribution of the test statistic under the null hypothesis).

Proximity ligation assay

Cells were fixed and permeabilized by the addition of ice-cold methanol (−20 °C) for 3 min followed by three washing steps with cold PBS. PLA was carried out according to manufacturer's instructions (Sigma Aldrich, #DUO92007). The following primary antibodies were used for the PLA assay: ERα (sc-8002, 1:200), STAT3 (sc-482, 1:200), pSTAT3^{Y705} (Cell Signaling, 9131, 1:200) and the incubation on the coverslips was performed for 1 h at 37 °C. The secondary proximity probes (Sigma Aldrich, anti-Rabbit-PLUS, #DUO92002 and anti-Mouse-MINUS, #DUO92004) were incubated for 1 h at 37 °C. The Leica DFC340FX microscope was used and images were captured with Leica Imaging software. Raw TIFF images were opened in FIJI version 2.0.0-rc-69/1.52p. Channels were split into RGB using Image/Colour/Split Channels function. The green channel was discarded to remove background as it can cross bleed into the blue channel. Channels were remerged using only red and blue channels corresponding to DAPI and PLA staining. Brightness and contrast were adjusted identically for all images within each experiment. This adjustment was performed individually for the red and blue channels. DAPI and PLA fluorescence were captured at high resolution for a total of 8 separate observation fields. Cell numeration and PLA labelling were carried out using ImageJ software (https://imagej.net/Welcome). Cells and red PLA dots were counted using the 'Analyze Particles' function. For each condition at least 200 cells were imaged and analyzed. Then, the average number of spots per nucleus was calculated. All statistical analyses were carried out by performing Student's t-test.

Western blot

Cells were washed twice in cold PBS and harvested in RIPA buffer (Pierce, 89901). Whole cell extract was sonicated using the Bioruptor Plus (Diagenode) for 2 min (30 sec on/30 sec off) to degrade the DNA, quantified by Direct detect Spectrometer (Millipore) and run on NuPAGE 4-12% bis-Tris gels (Invitrogen). Proteins were transferred to nitrocellulose membranes by iBlot2 (Invitrogen), and the membrane was then blocked by Odyssey TBS blocking solution (Li-Cor, 927-50000) and incubated with antibody over night at 4 °C. After washing, the membrane was then incubated with appropriate secondary antibodies (Li-Cor) and developed using the Odyssey CLx Imaging System (Li-Cor). Primary antibodies: ER (Novocastra, NCL-L-ER-6F11), STAT3 (Santa Cruz, sc-482), pSTAT3^{Y705} (Cell Signaling, 9145), FOXA1 (Abcam, ab23738), beta-actin (Cell Signaling, 4970; Sigma, A5441). Secondary antibodies: Goat anti-rabbit (926-32211, 926-68071), goat anti-mouse (926-32210, 926-68070), donkey anti-goat (926-32214).

IHC on T47D-IL6 tumors and PDX explants

IHC was run on Leica's Polymer Refine Detection System (DS9800) using modified versions of their standard template on the automated Bond-III platform. De-waxing and re-hydration prior to IHC were done on the automated Leica ST5020, as was the post-IHC de-hydration and clearing. Antigen retrieval was performed in a sodium citrate buffer (pSTAT3^{Y705} and ER) and Tris-EDTA buffer (Ki67) at 100 °C for 20 min (pSTAT3) or 30 min (ER and Ki67). Primary antibodies against the following proteins were used: pSTAT3^{Y705} (Cell Signaling, 9145, 1:25 dilution), ER (Novocastra, NCL-ER-6F11/2, 1.071 μg/ml) and Ki67 (Dako, M7240, 1:400 dilution). Since the mouse Ki67 antibody was used on human xenografts in mice, sections were blocked with mouse

IgG blocking solution (Vector, MKB-2213) prior to incubation with the Ki67 antibody. Slides were incubated with primary antibody for 60 min (pSTAT3) or 15 min (ER and Ki67) at room temperature. DAB Enhancer was used for all antibodies (Leica, AR9432). Sections were mounted on Leica's coverslipper (CV5030), and the slides were imaged on the Aperio AT2 (Leica Biosystems) to create whole slide images. Images were captured at 20x magnification, with a resolution of 0.5 microns per pixel. Whole slide images were imported into the Halo TM (Indica labs) Image analysis platform. Annotations for tumor were selected and percentage of marker-positive cells was quantified using the Indica Labs Multiplex IHC v1.2 module.

IHC on PDX MIND model, Figure 6 and S7

IHC staining

Five tumors per treatment arm were fixed with PFA and paraffin embedded. Sections were cut, dewaxed, rehydrated and subsequently stained with antibodies on Ventana (Roche) using the fully automated Ventana Discovery ULTRA (Roche Diagnostics, Rotkreuz, Switzerland) with Ventana solutions. Paraffin sections were pretreated with heat using standard condition (40 min) CC1 solution. The primary antibody was applied and incubated for specific time (listed below). After incubation with a rabbit Impress HRP antibody (Vector laboratories), chromogenic revelation was performed with ChromoMap DAB kit (Roche Diagnostics, Rotkreuz, Switzerland). IHC staining for $pSTAT3^{Y705}$ was done as described above for the T47D-IL6 model.

Protein	REF	Dilution	Incubation time	Supplier
ER	BRB053	Ready to use	16 min	Zytomed systems
PR	790-2223	Ready to use	32 min	Ventana
Ki67	ab16667	1:400	60 min	Abcam

IHC quantifications

The slides were scanned with Olympus VS120-L100 slide scanner using a 20x/0.75 objective connected to a Pike F505 C Color camera. Images were loaded into QuPath (Bankhead et al., 2017) using the BioFormats Extension (https://github.com/qupath/qupath-bioformats-extension) and tumor cell regions were identified using a custom script. In short, the RGB images were converted to HSB (Hue Saturation Brightness), and the Brightness image was used to detect the tissue regions.

A median filter of radius 3 pixels was applied, followed by ImageJ's Triangle auto-threshold method [https://imagej.net/Auto_Threshold#Triangle]. Only regions larger than 1e4 square microns were kept for analysis.

The regions were then manually curated to ensure accurate detection of the tissue of interest and to define *in situ* versus invasive foci. Individual cells on each region were segmented using QuPath's built-in Watershed Cell Detection algorithm parameters as follows. Color deconvolution stains: hematoxylin (0.56273 0.71108 0.42155), DAB (0.32201 0.61392 0.7207). Brightfield image detection on optical density sum: pixel size 0.2 microns. Nucleus parameters in microns: background radius 8.0, median radius 0.0, sigma 1.5, minimum area 10.0, maximum area 400.0. Intensity and cell parameters: threshold 0.05, maximum background intensity 2.0, watershed process, include nuclei, smooth boundaries, make measurements, threshold compartment DAB OD mean, threshold positive 0.15. The number of positive cells was counted, and the average values were then plotted. The invasive versus *in situ* area was quantified for each tumor analyzed.

RNAscope *in situ* hybridization

Detection of Human *TBP*, *MYC*, *ELF3* and *BCL3* mRNA was performed on FFPE sections using Advanced Cell Diagnostics (ACD) RNAscope® 2.5 LS Reagent Kit-RED (Cat No. 322150), RNAscope® 2.5 LS Probe- Hs-TBP (Cat No. 314298), RNAscope® LS 2.5 Probe- Hs-MYC (Cat No. 311768), RNAscope® LS 2.5 Probe- Hs-ELF3 (Cat No. 416408), and RNAscope® LS 2.5 Probe- Hs-BCL3 (Cat No. 417608). Negative controls were performed using the RNAscope® 2.5 LS Negative Control Probe - DapB (Cat No. 312038). Briefly, sections were cut at 3 µm thick, baked for 1 h at 60 ºC before loading onto a Bond RX instrument (Leica Biosystems). Slides were deparaffinised and rehydrated on board before pre-treatments using Epitope Retrieval Solution 2 (Cat No. AR9640, Leica Biosystems) at 95 °C for 10 min, and ACD Enzyme from the LS Reagent kit at 40 ºC for 10 min. Probe hybridization and signal amplification was performed according to manufacturer's instructions. Fast red detection of human *TBP*, *MYC*, *ELF3*, *BCL3* mRNA was performed on the Bond Rx using the Bond Polymer Refine Red Detection Kit (Leica Biosystems, Cat No. DS9390) according to ACD protocol. Slides were then removed from the Bond Rx and were heated at 60 ºC for 1 h, dipped in Xylene and mounted using EcoMount Mounting Medium (Biocare Medical, CA, USA. Cat No. EM897L).

The slides were imaged on the Aperio AT2 (Leica Biosystems) to create whole slide images. Images were captured at 40x magnification, with a resolution of 0.25 microns per pixel. Whole slide images were imported into the Halo TM (Indica labs) Image analysis platform. Annotations for tumor were selected and the surface area and intensity of positive staining for each annotation was quantified using the Indica Labs – Area Quantification FL v 1.0 positive pixel counting algorithm. The average number of probe copies per cell were measured using the Indica Labs – ISH IHC v1.3 module.

IL6/pSTAT3^{Y705} and breast cancer outcome

Integrative IHC analysis of pSTAT3^{Y705} and IL6 expression was performed using a combined cohort of N=2,497 patients with primary, stage I-III breast cancer treated at a single institution in Nottingham. Within the combined cohort, pSTAT3 and IL6 expression was assessed by IHC using tissue microarrays of n=1,270 and n=1,380 cases respectively, as previously described (Ahmad et al., 2018; Aleskandarany et al., 2016). From the combined cohort, Kaplan-Meier (KM) plots of OS and BCSS were stratified by ER, IL6 and pSTAT3 status and statistical comparisons were performed using log-rank tests as indicated in the relevant figure legends using SPSS Statistics (ver26). Multivariate analysis of pSTAT3 status with clinical biomarkers (ER, PR, HER2 and Ki67 status) was performed using Cox regression in SPSS Statistics. Plots of univariate and multivariate analyses were prepared using GraphPad Prism (ver7).

Quantification and Statistical Analyses

Two way ANOVA tests were used to determine significant differences in the size of the primary and metastatic tumors in the MIND experiments, expression of protein markers as determined by IHC and invasion (Figure 1B,C and S1E; 6E,F,H,I; S7G). The p values and n are shown in the legends. p values for KM plots showing the association between pSTAT3^{Y705} and IL6 IHC staining and BCSS/OS was determined by log-rank test using SPSS Statistics, ver26 (Figure 1E,F and Figure S1G,H). Univariate and multivariate analyses were performed using Cox regression in SPSS Statistics (Figure S7A,B). Significant differences in RNA-seq, ChIP-seq and ATAC-seq data (corrected for multiple testing) were determined using DEseq2 (RNA-seq) (Love et al., 2014) and DiffBind (ChIP-seq and ATAC-seq) (Stark and Brown, 2011), respectively. padj and n values for these experiments are indicated in the relevant figure legends. Significant differences in qPLEX-

RIME were determined by qPLEXanalyzer (10.18129/B9.bioc.qPLEXanalyzer) (Papachristou et al., 2018). padj and n values are indicated in the legend for Figure 2A. In two analyses, an approximate random permutation p value was determined by comparing to an estimation of the distribution of the test statistic under the null hypothesis: 1) To determine whether the percentage of binding sites connected to IL6 target gene promoters was significantly different from random sites (Fig 4C and Figure S5C), and 2) to determine whether the overlap between pSTAT3 and ER binding and eQTLs linked to IL6 target genes was significantly different from the overlap with eQTLs linked to the same number of random expressed genes (Figure S5D). p values are indicated in the relevant figure legends. p values for the differences in the proportions of ER and IL6/STAT3 target genes repressed by aromatase inhibitor treatment in patients (Figure 6A and S7C) were determined by two-sided Fisher's exact tests and indicated in the relevant figure legends. p values for quantitative difference in PLA in response to IL6/fulvestrant and staining by IHC and RNAscope *in situ* hybridization in tumor explants were determined by two-sided Student's t-test. p and n values are indicated in the relevant figure legends. Student's t-test and Wilcoxon test were used for specific comparison throughout the manuscript with p values indicated in the relevant legends.

References

Ahmad, N., Ammar, A., Storr, S.J., Green, A.R., Rakha, E., Ellis, I.O., and Martin, S.G. (2018). IL-6 and IL-10 are associated with good prognosis in early stage invasive breast cancer patients. Cancer Immunol Immunother *67*, 537-549.

Aleskandarany, M.A., Agarwal, D., Negm, O.H., Ball, G., Elmouna, A., Ashankyty, I., Nuglozeh, E., Fazaludeen, M.F., Diez-Rodriguez, M., Nolan, C.C., *et al.* (2016). The prognostic significance of STAT3 in invasive breast cancer: analysis of protein and mRNA expressions in large cohorts. Breast Cancer Res Treat *156*, 9-20.

Bachelot, T., Ray-Coquard, I., Menetrier-Caux, C., Rastkha, M., Duc, A., and Blay, J.Y. (2003). Prognostic value of serum levels of interleukin 6 and of serum and plasma levels of vascular endothelial growth factor in hormone-refractory metastatic breast cancer patients. Br J Cancer *88*, 1721-1726.

Bai, L., Zhou, H., Xu, R., Zhao, Y., Chinnaswamy, K., McEachern, D., Chen, J., Yang, C.Y., Liu, Z., Wang, M., *et al.* (2019). A Potent and Selective Small-Molecule Degrader of STAT3 Achieves Complete Tumor Regression In Vivo. Cancer Cell *36*, 498-511 e417.

Bailey, T., Boden, M., Buske, F., Frith, M., Grant, C., Clementi, L., Ren, J., Li, W., and Noble, W. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic acids research *37*, 8.

Bankhead, P., Loughrey, M.B., Fernandez, J.A., Dombrowski, Y., McArt, D.G., Dunne, P.D., McQuaid, S., Gray, R.T., Murray, L.J., Coleman, H.G., *et al.* (2017). QuPath: Open source software for digital pathology image analysis. Scientific reports *7*, 16878.

Bastard, J.P., Jardel, C., Bruckert, E., Blondy, P., Capeau, J., Laville, M., Vidal, H., and Hainque, B. (2000). Elevated levels of interleukin 6 are reduced in serum and subcutaneous adipose tissue of obese women after weight loss. J Clin Endocrinol Metab *85*, 3338-3342.

Binai, N.A., Damert, A., Carra, G., Steckelbroeck, S., Lower, J., Lower, R., and Wessler, S. (2010). Expression of estrogen receptor alpha increases leptin-induced STAT3 activity in breast cancer cells. Int J Cancer *127*, 55-66.

Bruna, A., Rueda, O.M., Greenwood, W., Batra, A.S., Callari, M., Batra, R.N., Pogrebniak, K., Sandoval, J., Cassidy, J.W., Tufegdzic-Vidakovic, A., *et al.* (2016). A Biobank of Breast Cancer Explants with Preserved Intra-tumor Heterogeneity to Screen Anticancer Compounds. Cell *167*, 260-274 e222.

Casneuf, T., Axel, A.E., King, P., Alvarez, J.D., Werbeck, J.L., Verhulst, T., Verstraeten, K., Hall, B.M., and Sasser, A.K. (2016). Interleukin-6 is a potential therapeutic target in interleukin-6 dependent, estrogen receptor-alpha-positive breast cancer. Breast Cancer (Dove Med Press) *8*, 13-27.

Centenera, M.M., Hickey, T.E., Jindal, S., Ryan, N.K., Ravindranathan, P., Mohammed, H., Robinson, J.L., Schiewer, M.J., Ma, S., Kapur, P., *et al.* (2018). A patient-derived explant (PDE) model of hormone-dependent cancer. Mol Oncol *12*, 1608-1622.

Corces, M.R., Granja, J.M., Shams, S., Louie, B.H., Seoane, J.A., Zhou, W., Silva, T.C., Groeneveld, C., Wong, C.K., Cho, S.W., *et al.* (2018). The chromatin accessibility landscape of primary human cancers. Science *362*.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., *et al.* (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat Methods *14*, 959-962.

Creyghton, M., Cheng, A., Welstead, G., Kooistra, T., Carey, B., Steine, E., Hanna, J., Lodato, M., Frampton, G., Sharp, P., *et al.* (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America *107*, 21931-21936.
De Luca, A., Lamura, L., Gallo, M., Maffia, V., and Normanno, N. (2012). Mesenchymal stem cell-derived interleukin-6 and vascular endothelial growth factor promote breast cancer cell migration. J Cell Biochem *113*, 3363-3370.
Di, D., Chen, L., Wang, L., Sun, P., Liu, Y., Xu, Z., and Ju, J. (2016). Downregulation of human intercellular adhesion molecule-1 attenuates the metastatic ability in human breast cancer cell lines. Oncology reports *35*, 1541-1548.
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. Bioinformatics *29*, 15-21.
Dunbier, A.K., Ghazoui, Z., Anderson, H., Salter, J., Nerurkar, A., Osin, P., A'Hern, R., Miller, W.R., Smith, I.E., and Dowsett, M. (2013). Molecular profiling of aromatase inhibitor-treated postmenopausal breast tumors identifies immune-related correlates of resistance. Clin Cancer Res *19*, 2775-2786.
Fiche, M., Scabia, V., Aouad, P., Battista, L., Treboux, A., Stravodimou, A., Zaman, K., Rls, Dormoy, V., Ayyanan, A., *et al.* (2019). Intraductal patient-derived xenografts of estrogen receptor alpha-positive breast cancer recapitulate the histopathological spectrum and metastatic potential of human lesions. J Pathol *247*, 287-292.
Franco, H.L., Nagari, A., and Kraus, W.L. (2015). TNFalpha signaling exposes latent estrogen receptor binding sites to alter the breast cancer cell transcriptome. Mol Cell *58*, 21-34.
Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize Implements and enhances circular visualization in R. Bioinformatics *30*, 2811-2812.
Hammacher, A., Ward, L.D., Weinstock, J., Treutlein, H., Yasukawa, K., and Simpson, R.J. (1994). Structure-function analysis of human IL-6: identification of two distinct regions that are important for receptor binding. Protein Sci *3*, 2280-2293.
Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., *et al.* (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol *31*, 827-832.
Hunter, C.A., and Jones, S.A. (2015). IL-6 as a keystone cytokine in health and disease. Nat Immunol *16*, 448-457.
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.
Huynh, J., Chand, A., Gough, D., and Ernst, M. (2018). Therapeutically exploiting STAT3 activity in cancer - using tissue repair as a road map. Nat Rev Cancer.
Ibrahim, S.A., El-Ghonaimy, E.A., Hassan, H., Mahana, N., Mahmoud, M.A., El-Mamlouk, T., El-Shinawi, M., and Mohamed, M.M. (2016). Hormonal-receptor positive breast cancer: IL-6 augments invasion and lymph node metastasis via stimulating cathepsin B expression. J Adv Res *7*, 661-670.
Jiang, M., Chen, J., Zhang, W., Zhang, R., Ye, Y., Liu, P., Yu, W., Wei, F., Ren, X., and Yu, J. (2017). Interleukin-6 Trans-Signaling Pathway Promotes Immunosuppressive Myeloid-Derived Suppressor Cells via Suppression of Suppressor of Cytokine Signaling 3 in Breast Cancer. Front Immunol *8*, 1840.
Johnson, D.E., O'Keefe, R.A., and Grandis, J.R. (2018). Targeting the IL-6/JAK/STAT3 signalling axis in cancer. Nat Rev Clin Oncol *15*, 234-248.

Kent, W., Sugnet, C., Furey, T., Roskin, K., Pringle, T., Zahler, A., and Haussler, D. (2002). The human genome browser at UCSC. Genome research *12*, 996-1006.
Kettner, N.M., Vijayaraghavan, S., Durak, M.G., Bui, T., Kohansal, M., Ha, M.J., Liu, B., Rao, X., Wang, J., Yi, M., *et al.* (2019). Combined Inhibition of STAT3 and DNA Repair in Palbociclib-Resistant ER-Positive Breast Cancer. Clin Cancer Res.
Kim, M., Jang, K., Miller, P., Picon-Ruiz, M., Yeasky, T.M., El-Ashry, D., and Slingerland, J.M. (2017). VEGFA links self-renewal and metastasis by inducing Sox2 to repress miR-452, driving Slug. Oncogene *36*, 5199-5211.
Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.
Lawrence, M., Huber, W., Pages, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput Biol *9*, e1003118.
Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst *1*, 417-425.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.
Madeddu, C., Gramignano, G., Floris, C., Murenu, G., Sollai, G., and Maccio, A. (2014). Role of inflammation and oxidative stress in post-menopausal oestrogen-dependent breast cancer. J Cell Mol Med *18*, 2519-2529.
Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.Y., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R., *et al.* (2016). JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. Nucleic Acids Res *44*, D110-115.
Matutino, A., Joy, A.A., Brezden-Masley, C., Chia, S., and Verma, S. (2018). Hormone receptor-positive, HER2-negative metastatic breast cancer: redrawing the lines. Curr Oncol *25*, S131-S141.
Mauer, J., Denson, J.L., and Bruning, J.C. (2015). Versatile functions for IL-6 in metabolism and cancer. Trends Immunol *36*, 92-101.
Papachristou, E.K., Kishore, K., Holding, A.N., Harvey, K., Roumeliotis, T.I., Chilamakuri, C.S.R., Omarjee, S., Chia, K.M., Swarbrick, A., Lim, E., *et al.* (2018). A quantitative mass spectrometry-based approach to monitor the dynamics of endogenous chromatin-associated protein complexes. Nat Commun *9*, 2311.
Paakinaho, V., Swinstead, E.E., Presman, D.M., Grontved, L., and Hager, G.L. (2019). Meta-analysis of Chromatin Programming by Steroid Receptors. Cell reports *28*, 3523-3534 e3522.
R Core Team (2020). R: A language and environment for statistical computing.
Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res *43*, e47.
Salgado, R., Junius, S., Benoy, I., Van Dam, P., Vermeulen, P., Van Marck, E., Huget, P., and Dirix, L.Y. (2003). Circulating interleukin-6 predicts survival in patients with metastatic breast cancer. Int J Cancer *103*, 642-646.
Schrijver, W., Schuurman, K., van Rossum, A., Droog, M., Jeronimo, C., Salta, S., Henrique, R., Wesseling, J., Moelans, C., Linn, S.C., *et al.* (2018). FOXA1 levels are decreased in pleural breast cancer metastases after adjuvant endocrine therapy, and this is associated with poor outcome. Mol Oncol *12*, 1884-1894.
Selli, C., Turnbull, A.K., Pearce, D.A., Li, A., Fernando, A., Wills, J., Renshaw, L., Thomas, J.S., Dixon, J.M., and Sims, A.H. (2019). Molecular changes during extended neoadjuvant letrozole

treatment of breast cancer: distinguishing acquired resistance from dormant tumours. Breast Cancer Res *21*, 2.
Sflomos, G., Dormoy, V., Metsalu, T., Jeitziner, R., Battista, L., Scabia, V., Raffoul, W., Delaloye, J.F., Treboux, A., Fiche, M., *et al.* (2016). A Preclinical Model for ERalpha-Positive Breast Cancer Points to the Epithelial Microenvironment as Determinant of Luminal Phenotype and Hormone Response. Cancer Cell *29*, 407-422.
Shuey, D.L., Oliver, J., Zhou, G., and Roberts, A. (2016). Results from oral gavage carcinogenicity studies of ruxolitinib in Tg.rasH2 mice and Sprague-Dawley (Crl:CD) rats. Regul Toxicol Pharmacol *81*, 305-315.
Siersbaek, R., Kumar, S., and Carroll, J.S. (2018). Signaling pathways and steroid receptors modulating estrogen receptor alpha function in breast cancer. Genes Dev *32*, 1141-1154.
Sonnenblick, A., Salgado, R., Brohee, S., Zahavi, T., Peretz, T., Van den Eynden, G., Rouas, G., Salmon, A., Francis, P.A., Di Leo, A., *et al.* (2018). p-STAT3 in luminal breast cancer: Integrated RNA-protein pooled analysis and results from the BIG 2-98 phase III trial. Int J Oncol *52*, 424-432.
Speirs, V., Kerin, M.J., Walton, D.S., Newton, C.J., Desai, S.B., and Atkin, S.L. (2000). Direct activation of oestrogen receptor-alpha by interleukin-6 in primary cultures of breast cancer epithelial cells. Br J Cancer *82*, 1312-1316.
Stark, R., and Brown, G. (2011). DiffBind : differential binding analysis of ChIP-Seq peak data.
Stender, J.D., Nwachukwu, J.C., Kastrati, I., Kim, Y., Strid, T., Yakir, M., Srinivasan, S., Nowak, J., Izard, T., Rangarajan, E.S., *et al.* (2017). Structural and Molecular Mechanisms of Cytokine-Mediated Endocrine Resistance in Human Breast Cancer Cells. Mol Cell *65*, 1122-1135 e1125.
Stock, K., Estrada, M.F., Vidic, S., Gjerde, K., Rudisch, A., Santo, V.E., Barbier, M., Blom, S., Arundkar, S.C., Selvam, I., *et al.* (2016). Capturing tumor complexity in vitro: Comparative analysis of 2D and 3D tumor models for drug discovery. Scientific reports *6*, 28951.
Toska, E., Osmanbeyoglu, H.U., Castel, P., Chan, C., Hendrickson, R.C., Elkabets, M., Dickler, M.N., Scaltriti, M., Leslie, C.S., Armstrong, S.A., *et al.* (2017). PI3K pathway regulates ER-dependent transcription in breast cancer through the epigenetic regulator KMT2D. Science *355*, 1324-1330.
Wakefield, A., Soukupova, J., Montagne, A., Ranger, J., French, R., Muller, W.J., and Clarkson, R.W. (2013). Bcl3 selectively promotes metastasis of ERBB2-driven mammary tumors. Cancer research *73*, 745-755.
Wang, Y., Niu, X.L., Guo, X.Q., Yang, J., Li, L., Qu, Y., Xiu Hu, C., Mao, L.Q., and Wang, D. (2015). IL6 induces TAM resistance via kinase-specific phosphorylation of ERalpha in OVCA cells. J Mol Endocrinol *54*, 351-361.
Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New York).
Wong, A.L.A., Hirpara, J.L., Pervaiz, S., Eu, J.Q., Sethi, G., and Goh, B.C. (2017). Do STAT3 inhibitors have potential in the future for cancer therapy? Expert Opin Investig Drugs *26*, 883-887.
Yamamoto, T., Matsuda, T., Junicho, A., Kishi, H., Saatcioglu, F., and Muraguchi, A. (2000). Cross-talk between signal transducer and activator of transcription 3 and estrogen receptor signaling. FEBS Lett *486*, 143-148.
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biology *9*, R137.

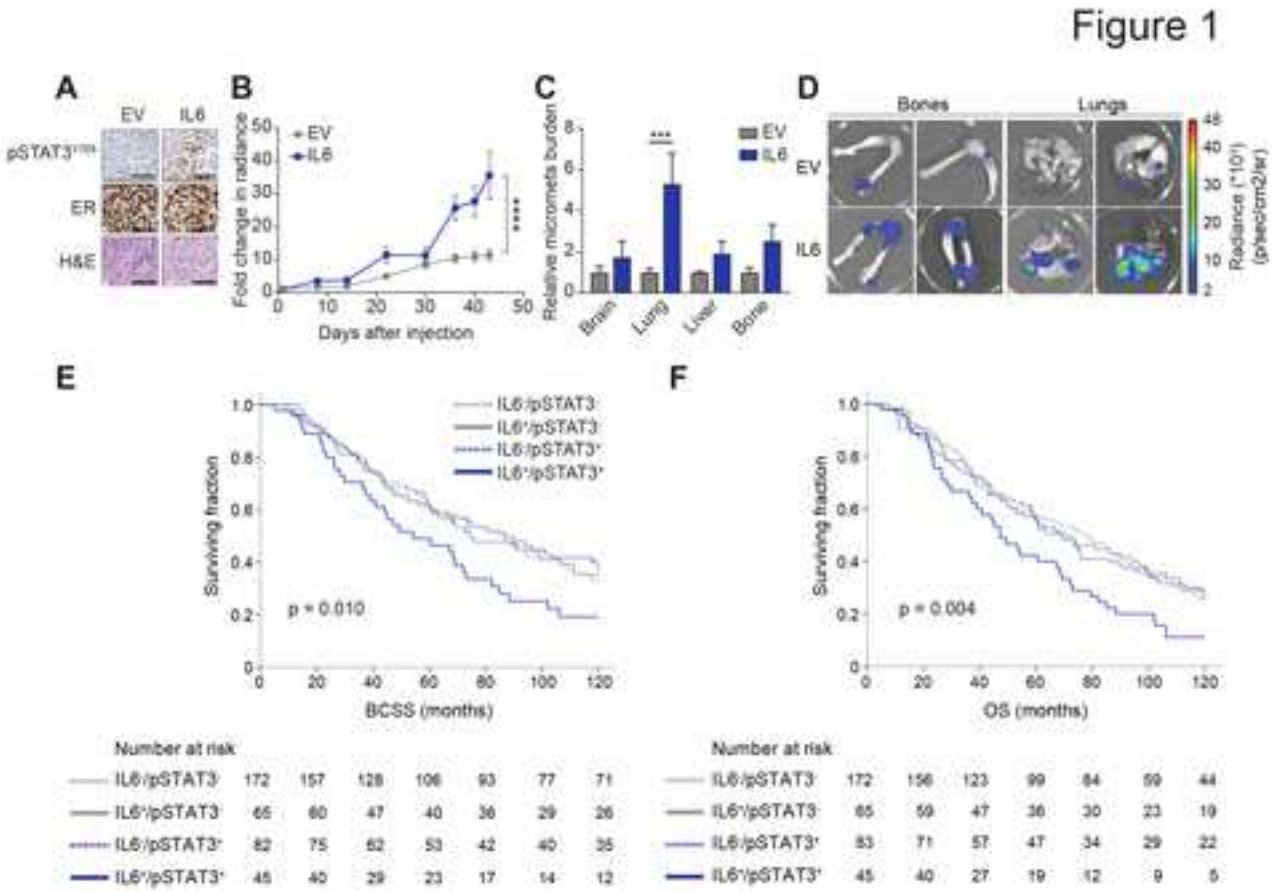


Figure 2

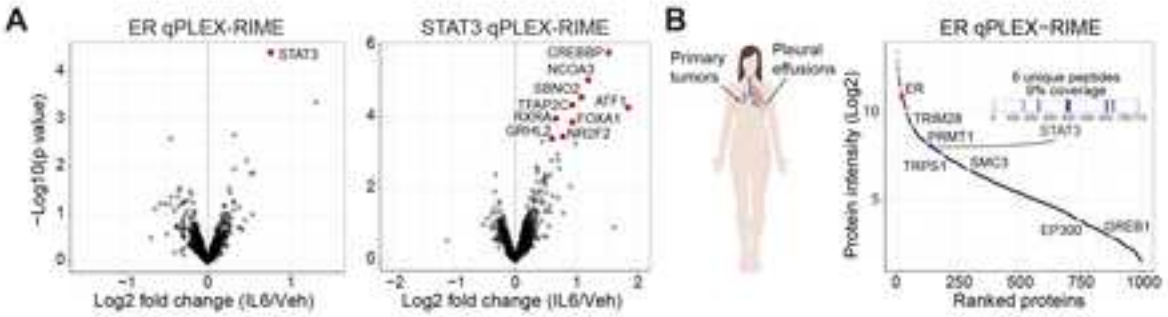


Figure 3

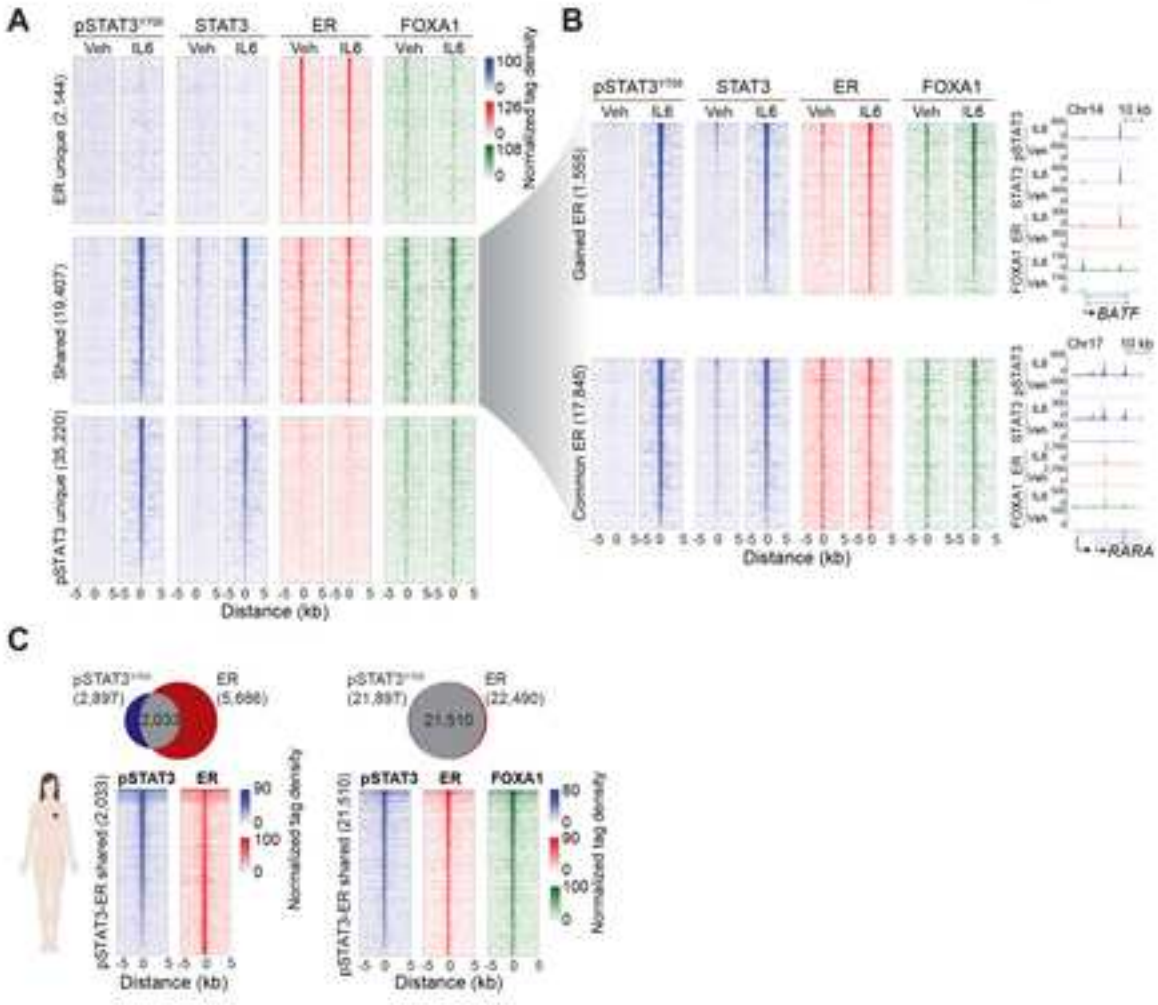
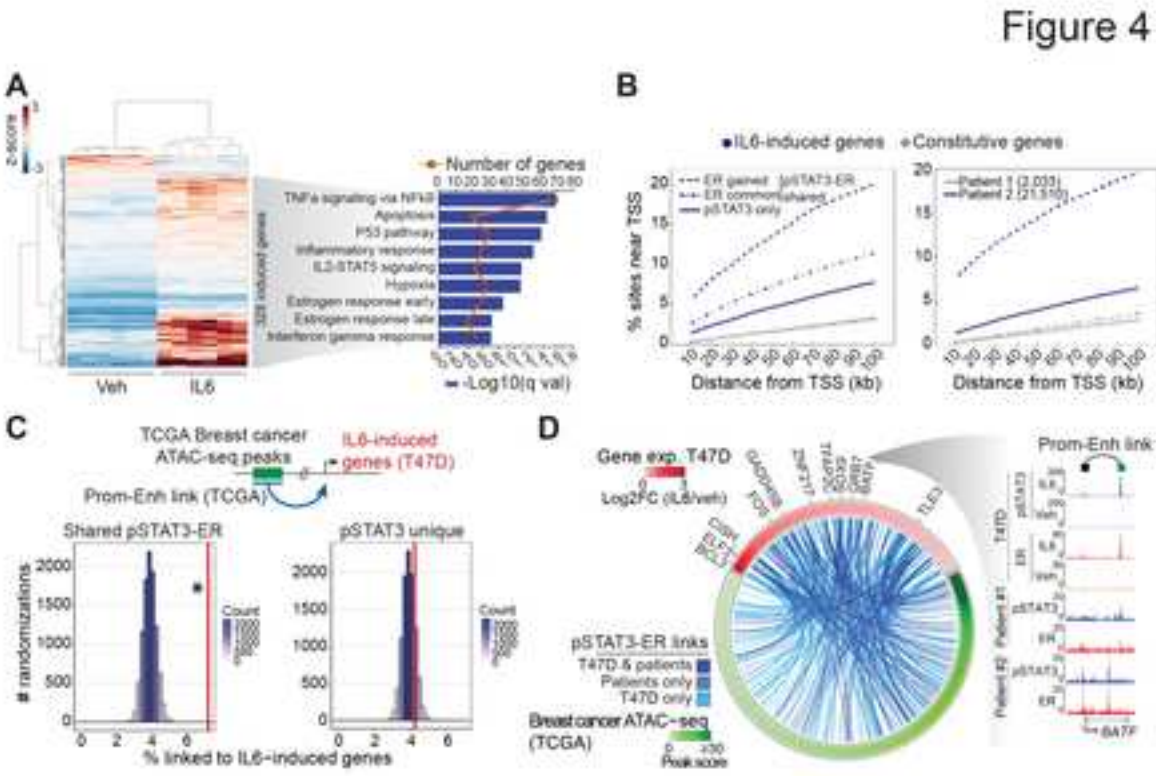


Figure 4

Click here to access/download;Figure;Main_figures_CancerCell_Final_revisio



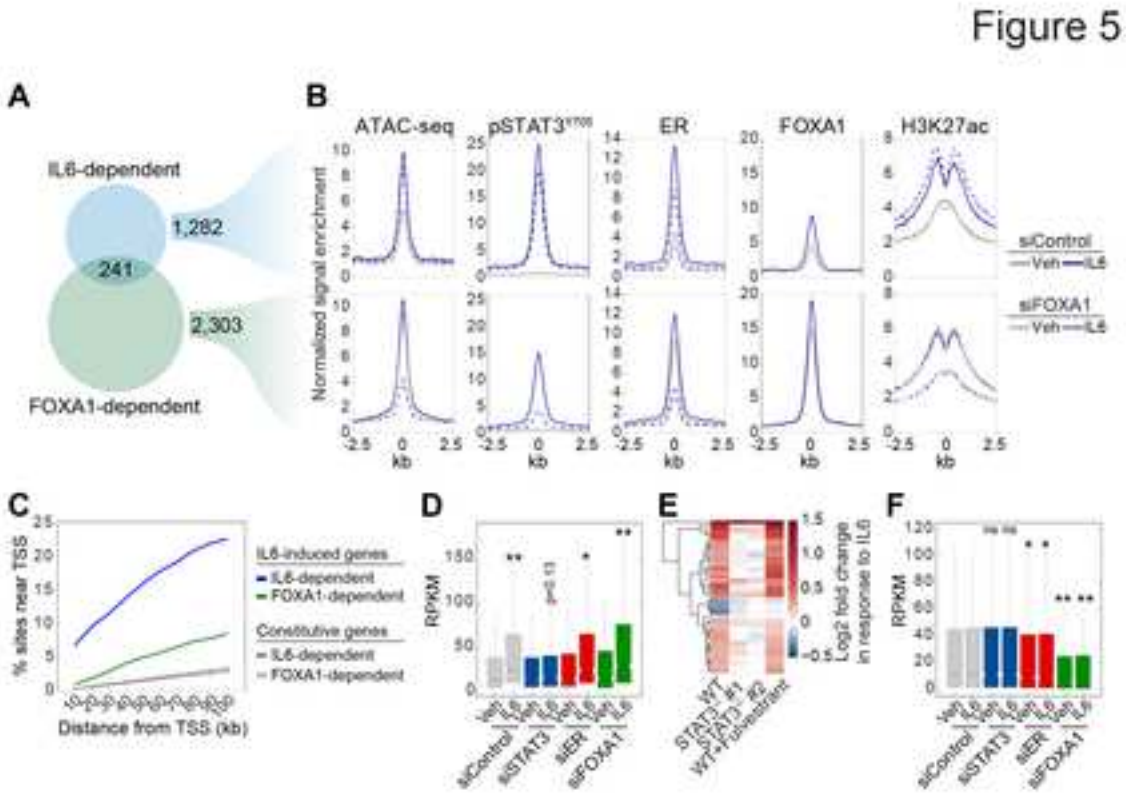


Figure 6

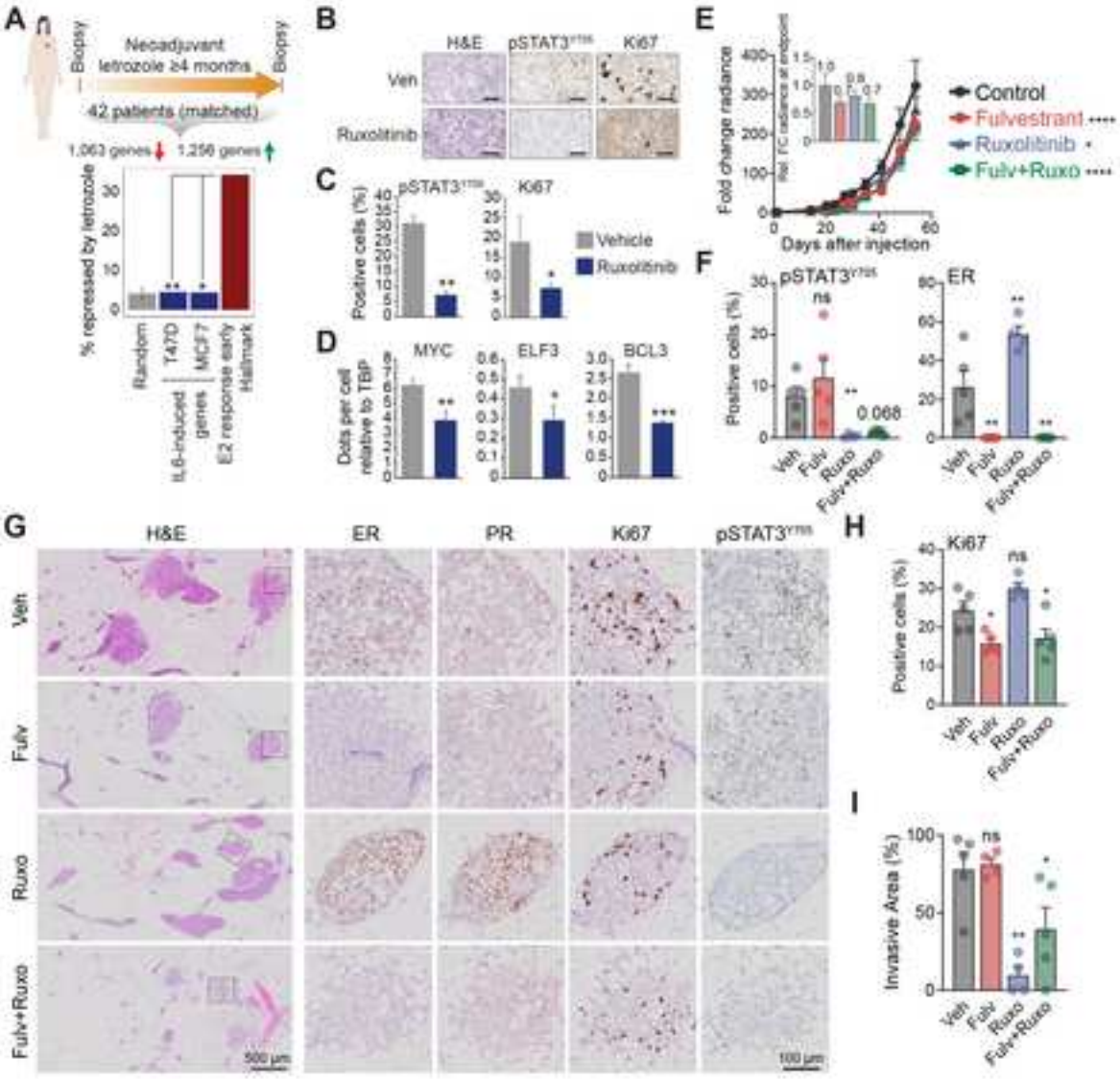
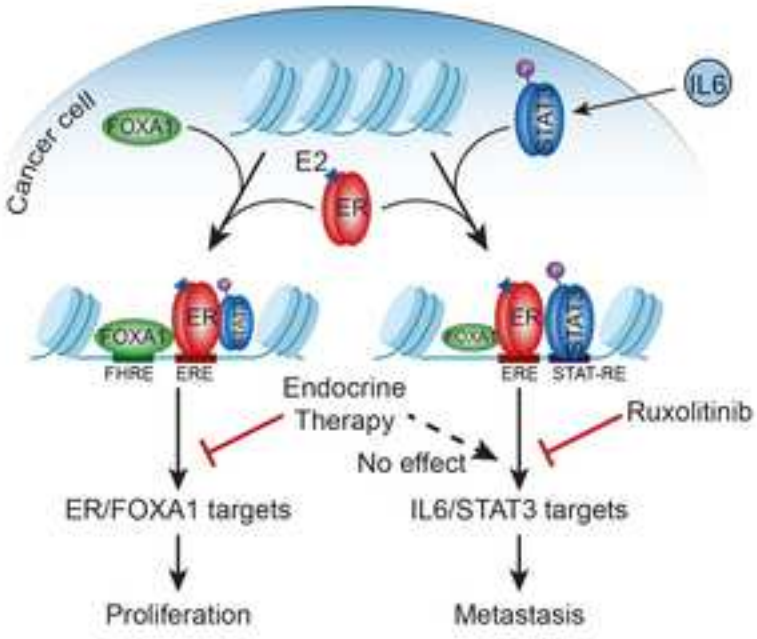


Figure 7



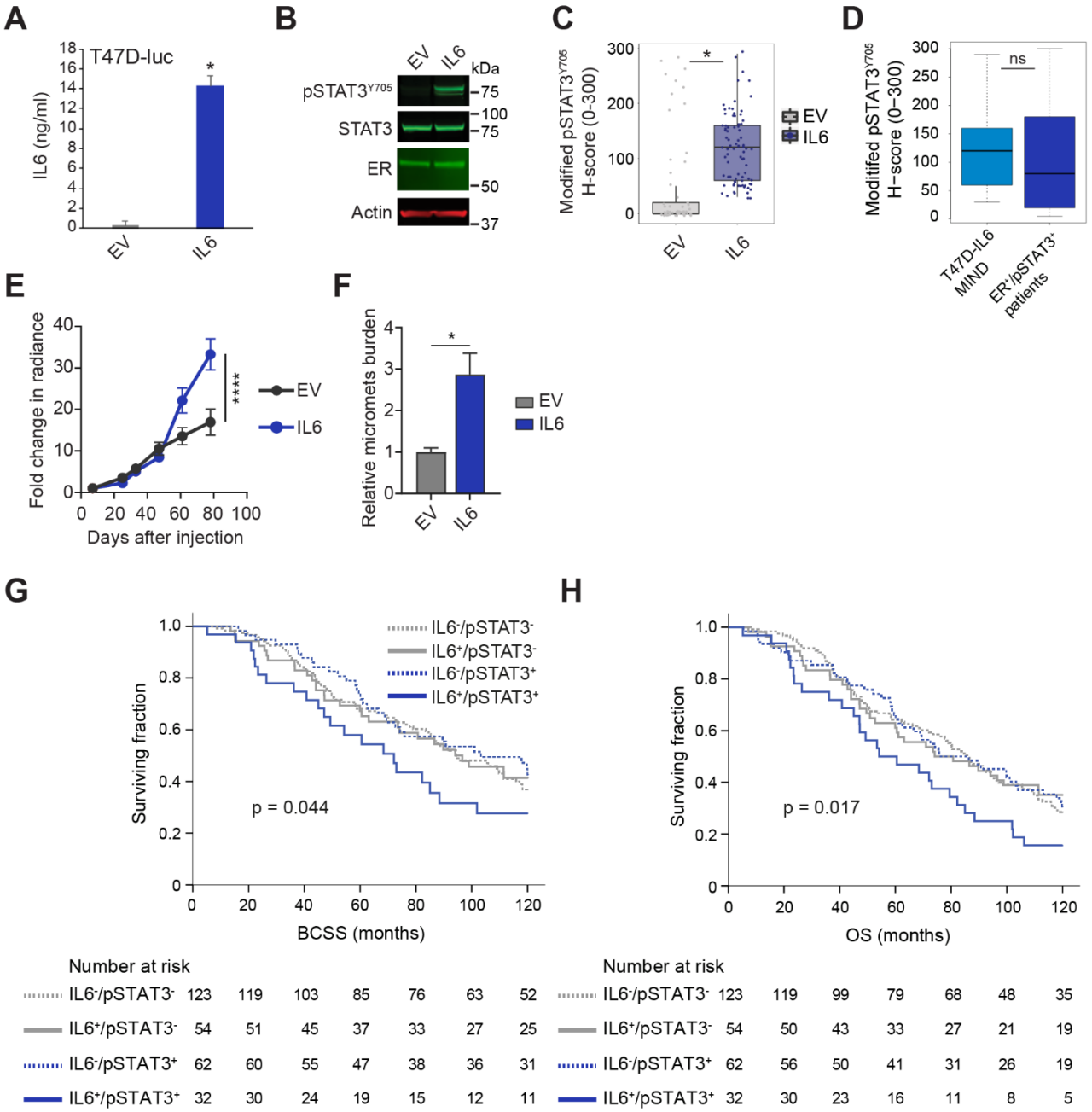


Figure S1. Ectopic expression of human IL6 in T47D cells activates STAT3 by Y705 phosphorylation and induces *in vivo* growth and metastasis, Related to Figure 1.

(A) Human IL6 levels in the media used for growing luciferase-expressing T47D cells expressing human IL6 or empty vector (EV) were determined by ELISA. Results are representative of two independent experiments, and error bars show standard deviations based on three technical replicates of the ELISA. *p=0.00052 (Student's t-test, two-sided). **(B)** Western blot showing phosphorylation of STAT3 at Y705 in luciferase-expressing T47D cells expressing IL6, but not in cells infected with empty vector (EV). Actin is used as a loading control. Representative of two independent biological replicates. **(C)** Modified H-score for pSTAT3^{Y705} in MIND samples from luciferase-expressing T47D cells expressing empty vector (EV) or IL6 (four mice from each group). Due to the fact that

phosphorylation of STAT3 is dictated by the tumor microenvironment, each tumor area was quantified separately by a histopathologist and visualized as a point in the boxplot. *p=1.6e-12 (two-sided Wilcoxon test). Boxplot definitions: Center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points, separate tumor areas. **(D)** Modified H-scores from $pSTAT3^{Y705}$ IHC (Cell Signaling, 9145) of tumor foci (n=85 from four mice, see panel C) from the T47D-IL6 MIND tumors were compared to modified $pSTAT3^{Y705}$ H-scores (Ab76315) from ER^+ patients (n=103) (Aleskandarany et al., 2016). Non-significant (ns) p=0.058 was determined by two-sided Wilcoxon test. Boxplot definitions: Center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; any outliers have been hidden. **(E)** Growth of primary T47D MIND tumors expressing empty vector (EV) or IL6 as determined by changes in radiance relative to day 7 after cell injection. 10 mice per condition, 23 glands for EV and 24 glands for IL6. Mean of the fold change in radiance per gland is shown with error bars representing SEM. ****p<0.0001 (two way ANOVA, Sidak's multiple comparisons test). **(F)** Relative combined metastatic burden for all metastatic sites, i.e. bone, lung, liver and brain, in the T47D-IL6 MIND model (n=60 organs from 15 mice per condition). Metastatic burden in individual organs is shown in Figure 1C. *p=0.016 (Mann-Whitney test). **(G,H)** Kaplan-Meier (KM) plots showing the association between BCSS (G) or OS (H) and IHC staining of $pSTAT3^{Y705}$ and IL6 in ER^+ breast tumors from the Nottingham cohort (Ahmad et al., 2018; Aleskandarany et al., 2016). Log-rank p values are indicated for $ER^+/IL6^+/pSTAT3^+$ vs all other ER^+ cases.

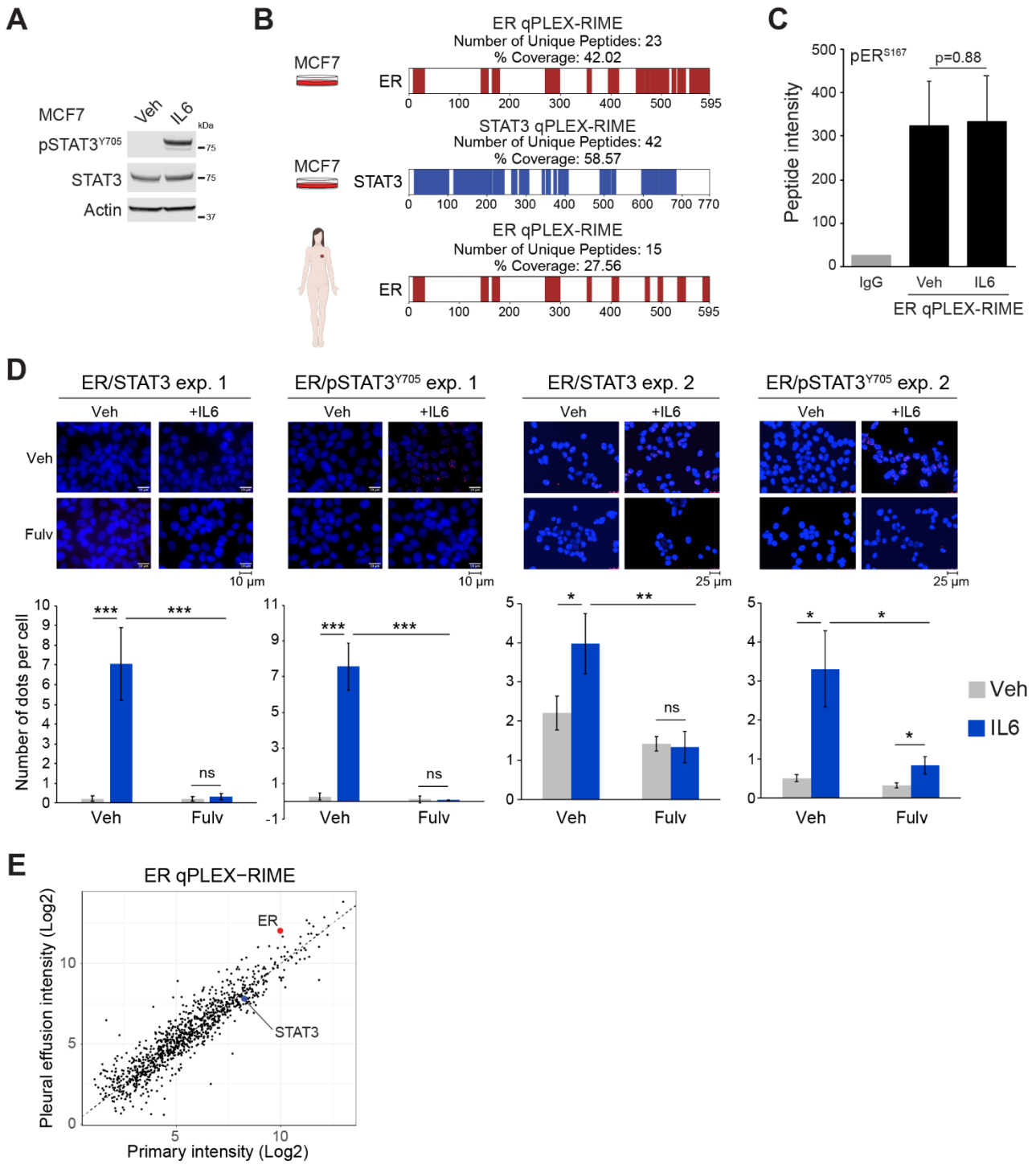


Figure S2. Western blot, PLA and qPLEX-RIME in cell lines and patient material, Related to Figure 2.

(A) Western blot showing phosphorylation of STAT3 at Y705 in response to 30 min IL6 treatment in MCF7 cells. Representative of more than four independent biological replicates. **(B)** Peptide coverage of ER and STAT3 in the qPLEX-RIME experiments in MCF7 cells (top and middle) and clinical samples (bottom). See Figure 2 for more details. The schematic illustration was created with BioRender.com. **(C)** Intensity of the ER peptide phosphorylated on serine 167 in ER qPLEX-RIME

(see Figure 2A for more details). Errors bars represent standard deviation of four independent biological replicates. p value was determined by Student's t-test, two-sided. **(D)** PLA of ER/STAT3 and ER/pSTAT3^{Y705} in MCF7 cells treated with vehicle (veh) or IL6 (30 min) following three hours of veh or fulvestrant treatment (100 nM). Top, representative images from two independent PLA experiments. Brightness and contrast have been adjusted identically for all images within each experiment. This adjustment was performed individually for the red and blue channels. Bottom, bar plots showing quantification of the PLA signal. Error bars represent standard deviations. *$p<0.01$, **$p<0.001$, ***$p<0.0001$ (Student's t-test, two-sided). **(E)** Comparison of protein intensity in ER^{+} primary tumors and pleural effusions from ER qPLEX-RIME data. ER and STAT3 are highlighted in red and blue, respectively. See Figure 2B for more details.

Table S1. Clinical information for the patients whose tumors were used for ER qPLEX-RIME, Related to Figure 2.

Patient	*ER*	*PR*	*Type*	*Hormonal Treatment*
1	+	+	Ductal/grade 2	None
2	+	+	Ductal /grade 2 with small lobular component	Treatment info is missing
3	+	+	Ductal/grade 2	None
4	+	+	Ductal/grade2	None
5	+	+	Ductal/grade 1	None
6	+	Not stained for PR	Pleural effusion	Received adjuvant hormonal therapy
7	+	-	Pleural effusion	Received adjuvant hormonal therapy
8	+	Not stained for PR	Pleural effusion	Received adjuvant hormonal therapy
9	+	-	Pleural effusion	Received adjuvant hormonal therapy

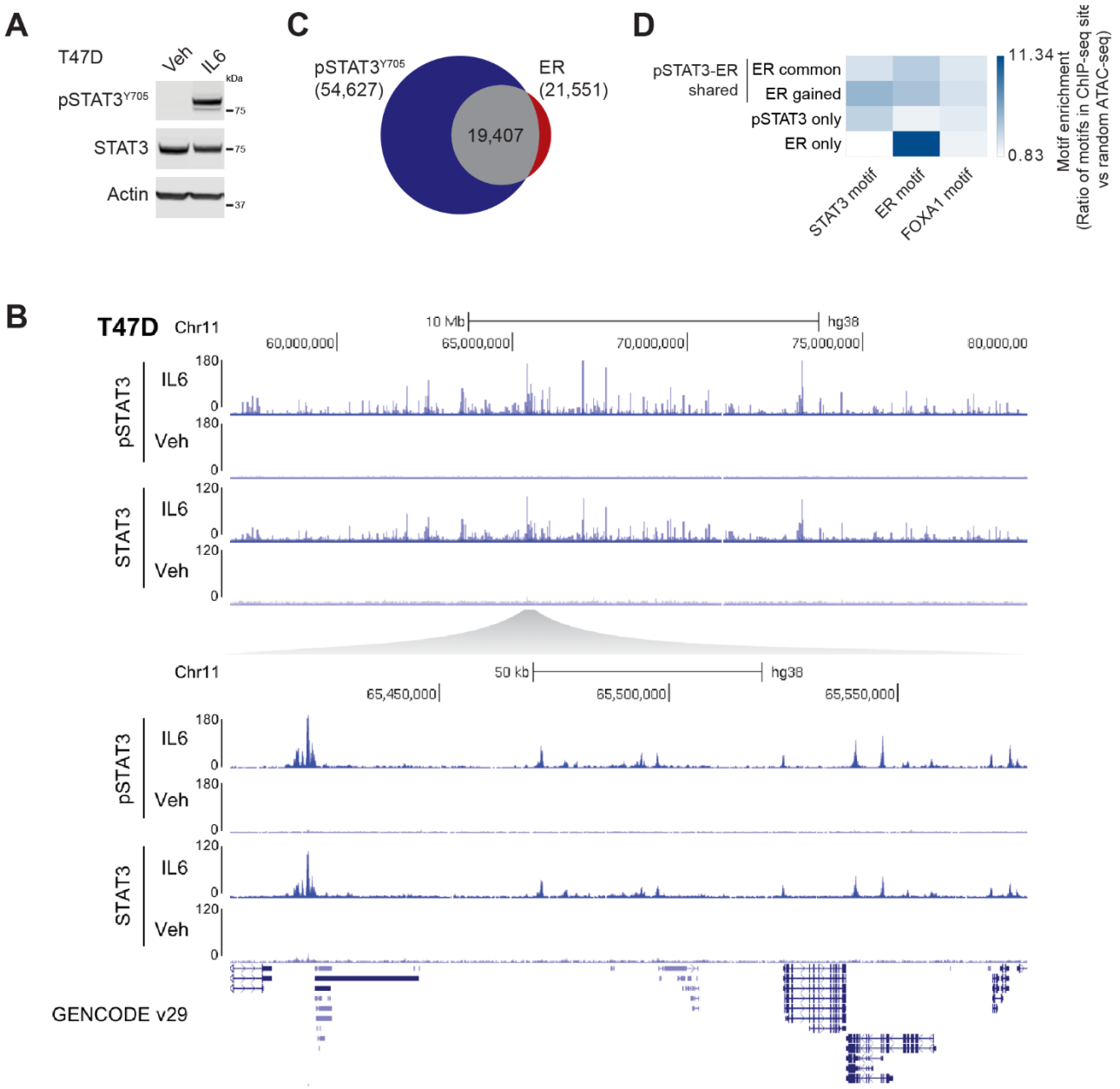


Figure S3. pSTAT3, STAT3, ER, and FOXA1 ChIP-seq in T47D cells, Related to Figure 3.

(A) Western blot showing phosphorylation of STAT3 at Y705 in response to 30 min IL6 treatment in T47D cells. Representative of four independent biological replicates. **(B)** Screen shots from the UCSC genome browser (http://genome.ucsc.edu, Dec. 2013 GRCh38/hg38) (Kent et al., 2002) showing binding of pSTAT3^{Y705} and total STAT3 as determined by ChIP-seq in response to 30 min IL6 treatment in T47D cells. **(C)** Venn diagram showing the positional overlap of pSTAT3^{Y705} and ER binding in T47D cells in response to 30 min IL6 treatment as determined by ChIP-seq (n=4). **(D)** Analysis of motif enrichment in ER-pSTAT3 binding sites from Figure 3A,B was performed using ER, FOXA1 and STAT3 motif matrices from the Jaspar database (Mathelier et al., 2016). The enrichment was calculated as the number of motif instances in each group of sites relative to the number of motif instances in the same number of random T47D ATAC-seq peaks (Toska et al., 2017).

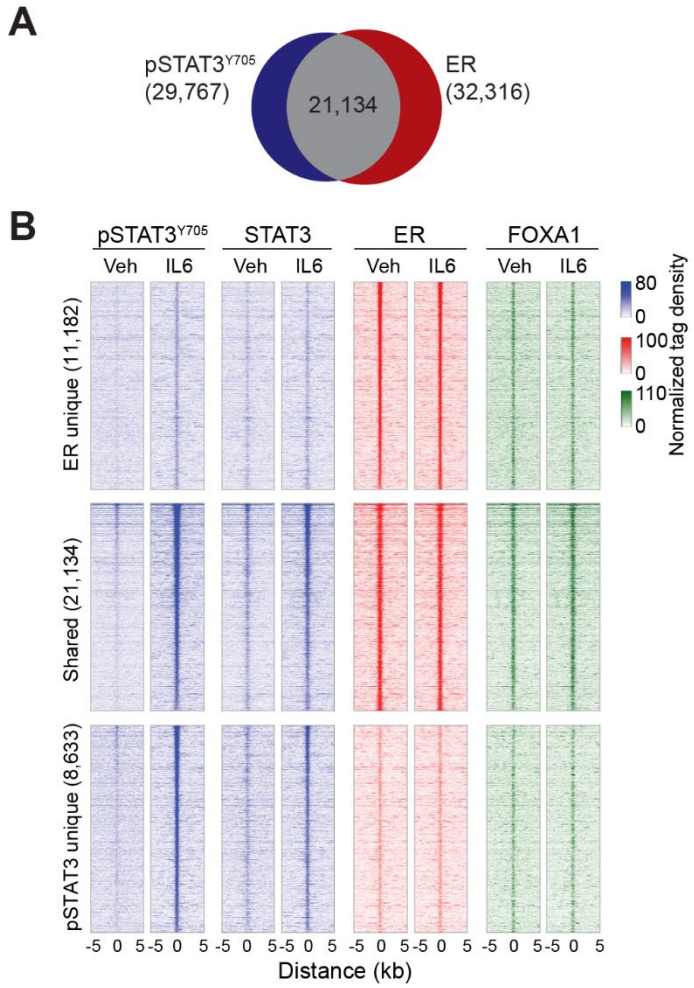


Figure S4. pSTAT3, STAT3, ER and FOXA1 ChIP-seq in MCF7 cells, Related to Figure 3.
(A) Venn diagram showing positional overlap of pSTAT3^{Y705} and ER binding in MCF7 cells upon 30 min stimulation with IL6 as determined by ChIP-seq (n=4). **(B)** Heatmap showing the ChIP-seq intensity for pSTAT3^{Y705}, total STAT3, ER and FOXA1 at the binding sites defined in panel A in vehicle (Veh) control conditions and upon treatment with IL6. Color bars indicate scale for normalized tag densities per 20 M subsampled reads.

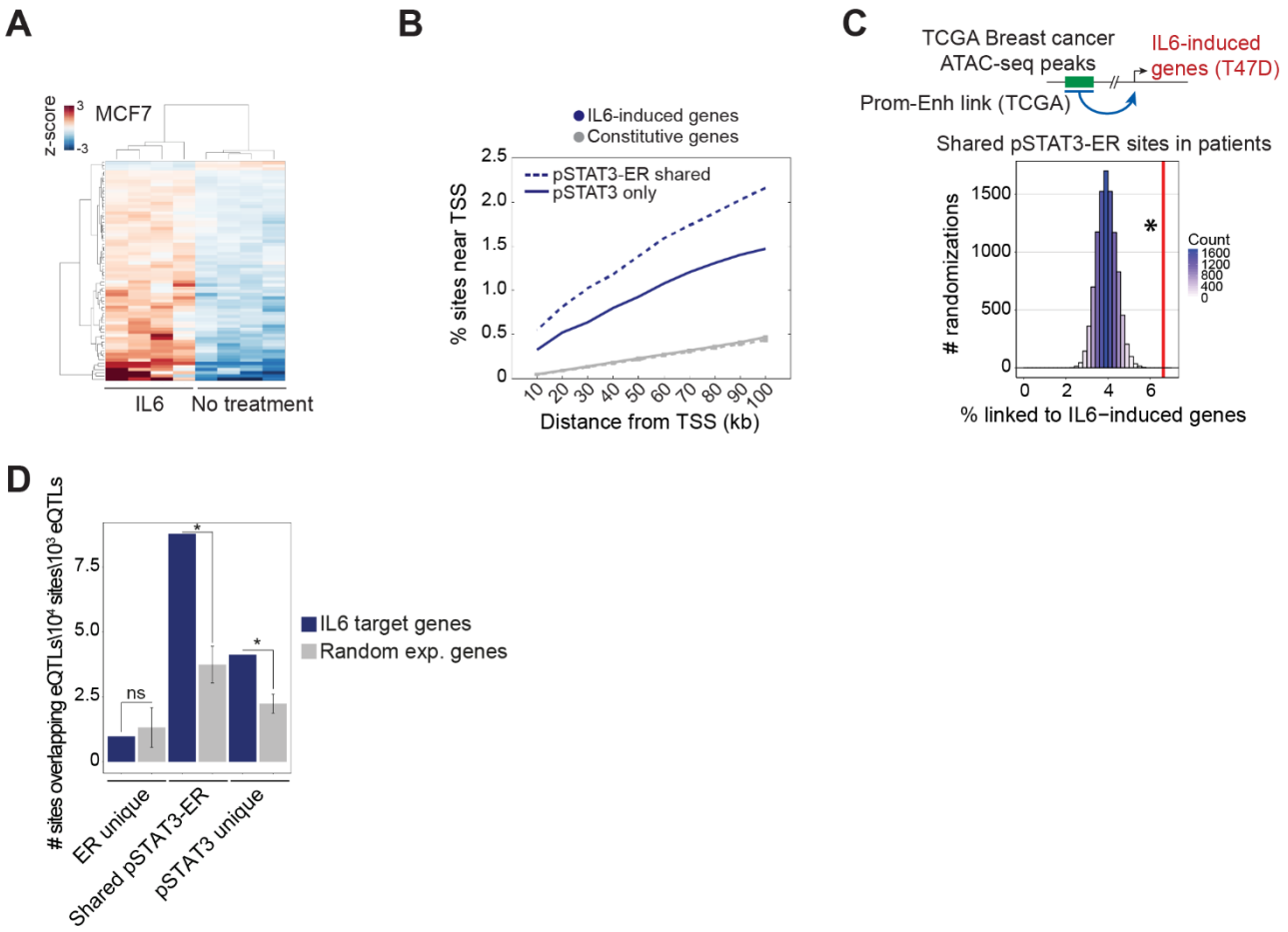


Figure S5. pSTAT3-ER binding sites are functionally connected to IL6-activated genes, Related to Figure 4.

(A) Differentially expressed genes (padj<0.1, DESeq2 (Love et al., 2014)) in MCF7 cells in response to one-hour treatment with recombinant human IL6 as determined by RNA-seq (n=4). **(B)** Enrichment of shared pSTAT3-ER binding sites (dashed line) and sites only bound by pSTAT3 (solid line) in the vicinity of IL6-induced genes (blue) compared to an equal number of constitutively expressed genes (grey) in MCF7 cells. Lines illustrate the cumulative percentage of sites within a given distance from the TSS. **(C)** Percentage of breast cancer ATAC-seq peaks occupied by pSTAT3 and ER in patients (Figure 3C) that are connected to IL6 target gene promoters from T47D cells through promoter-enhancer connections (Corces et al., 2018) (Red line). The distribution of connections between random ATAC-seq peaks and IL6 target gene promoters is shown for comparison in blue (10,000 iterations to estimate the distribution of the test statistic under the null hypothesis). *Approximate random permutation p value <0.0001. **(D)** Number of binding sites occupied by pSTAT3 and/or ER that contain expression quantitative trait loci (eQTL) linked to expression of IL6 target genes in T47D cells (Figure 4A) (blue) or the same number of random expressed genes (grey) were determined using eQTL data from the GTEx Portal (GTEx_Analysis_v8_eQTL.tar was downloaded on 11/14/2019). Error bars indicate standard deviations from 1,000 random permutations. *p<0.001 (1,000 iterations to estimate the distribution of the test statistic under the null hypothesis).

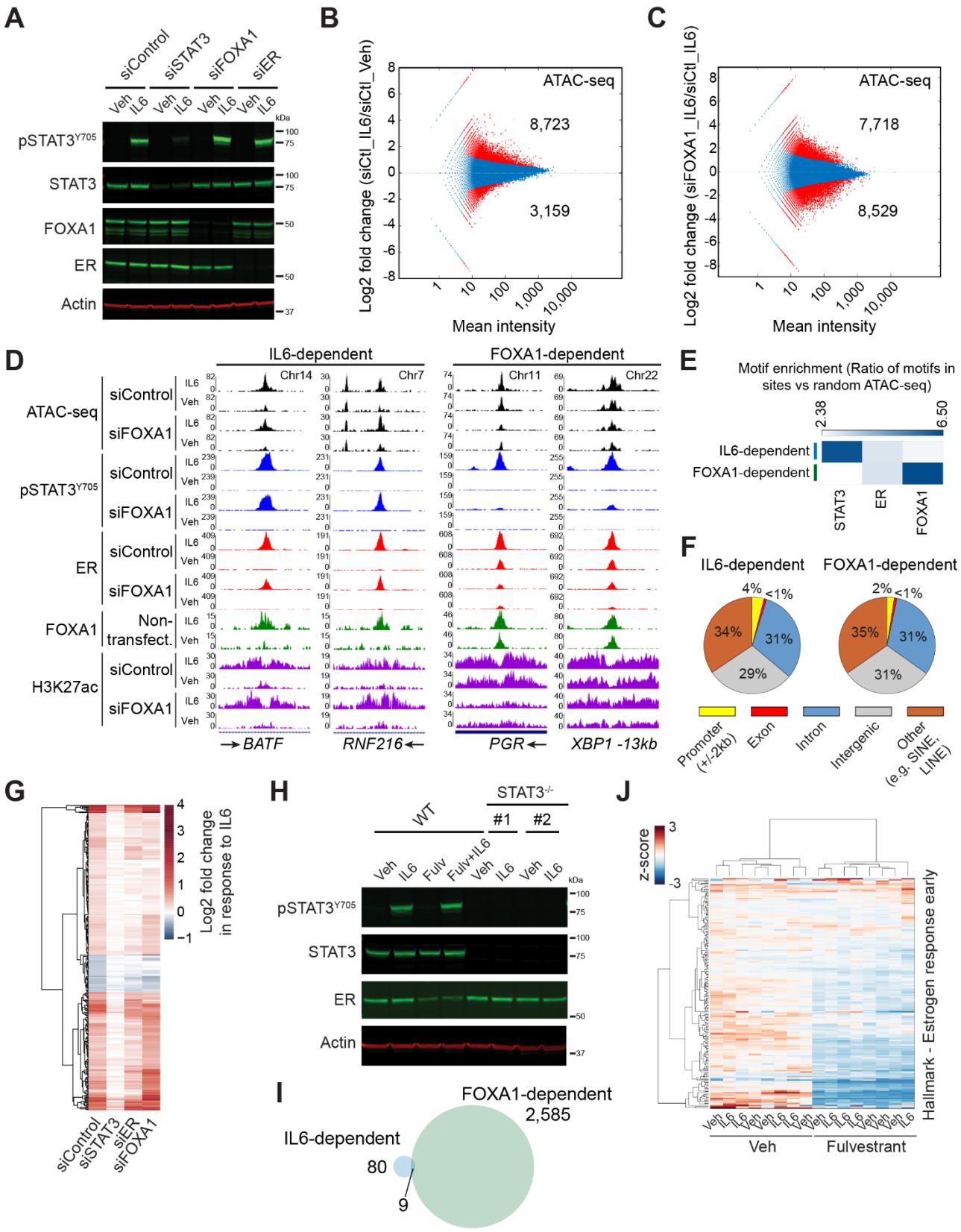


Figure S6. Western blot, ATAC-seq and RNA-seq after transcription factor perturbations in MCF7 and T47D cells, Related to Figure 5.

(A) Western blot showing siRNA-mediated knock down of STAT3, ER and FOXA1 in T47D cells after 48 h. Representative of two independent biological replicates. **(B,C)** MA plot showing changes in ATAC-seq signal in response to IL6 under siControl (ciCtl) conditions (B) or in response to FOXA1 knock down under IL6-stimulated conditions (C) (n=3). Red points indicate significant changes (padj<0.05, DiffBind (Stark and Brown, 2011)). **(D)** UCSC genome browser screen shots for representative IL6- and FOXA1-dependent ATAC-seq peaks from Figure 5A,B (http://genome.ucsc.edu, Dec. 2013 GRCh38/hg38) (Kent et al., 2002). **(E)** Analysis of motif enrichment in IL6- and FOXA1-dependent ATAC-seq peaks (see Figure 5A,B) was performed using ER, FOXA1 and STAT3 motif matrices from the Jaspar database (Mathelier et al., 2016). The enrichment was calculated as the number of motif instances in each group of sites relative to the number of motif instances in the same number of random T47D ATAC-seq peaks from the IL6 condition. **(F)** Genomic distribution of IL6- and FOXA1-dependent ATAC-seq peaks (see Figure 5A,B). **(G)** Regulation of the transcriptional response to one hour of IL6 treatment (padj<0.05, DESeq2 (Love et al., 2014)) upon depletion of STAT3, ER, or FOXA1 as determined by RNA-seq in T47D cells (n=4). **(H)** Western blot showing knock out of STAT3 and response to fulvestrant in MCF7 cells. WT is the parental wild-type MCF7 clone expressing luciferase and mStrawberry used to make the two $STAT3^{-/-}$ clones. Representative of two independent biological replicates. **(I)** Overlap between IL6- and FOXA1-dependent genes as determined by RNA-seq in T47D cells (n=4). FOXA1-dependent genes were defined as those with padj<0.05 and Log2 fold change<-0.5 (DESeq2 (Love et al., 2014)) comparing siFOXA1-IL6 to siControl-IL6. IL6-dependent genes were defined as those with padj<0.05 and Log2 fold change>0.5 comparing siControl-IL6 to siControl-Veh. **(J)** Heatmap showing a general repression of the Hallmark Estrogen response early gene set (Liberzon et al., 2015) in response to four hours of fulvestrant treatment of the MCF7 cells from panel H as determined by RNA-seq (n=4).

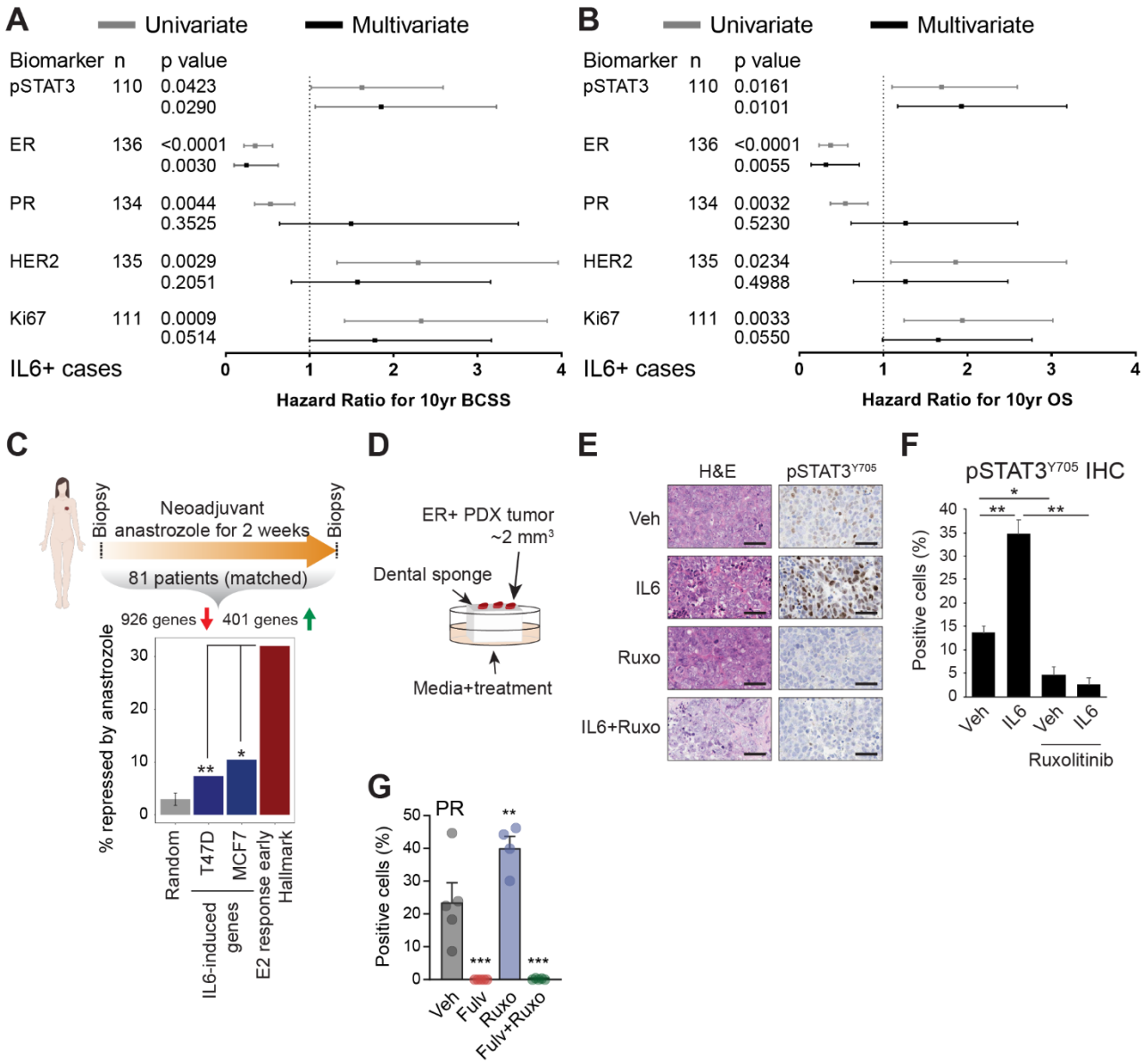


Figure S7. Univariate and multivariate analyses and response of the IL6 gene program to treatment, Related to Figure 6.

(A,B) Univariate and multivariate analyses of the prognostic significance of well-established breast cancer marker proteins (i.e. ER, HER2, PR, Ki67) and pSTAT3^{Y705} for 10 year BCSS (A) and OS (B) in the Nottingham cohort (Ahmad et al., 2018; Aleskandarany et al., 2016). p values were determined by Cox regression analyses. **(C)** Regulation of Hallmark early estrogen responsive genes (n=200) (Liberzon et al., 2015) and IL6-induced genes (n=328 for T47D and n=67 for MCF7) in matched samples from 81 patients treated with anastrozole for two weeks prior to surgery (Dunbier et al., 2013). The mean regulation of 200 random genes is included for comparison with the error bar representing the standard deviation of 1,000 iterations. *p=0.00038 and **p=4.0e-13 (Fisher's exact test, two-sided). The schematic illustration was created with BioRender.com. **(D)** Schematic overview of the explant approach used to culture ER$^+$ PDX tumors *ex vivo*. **(E)** Representative images of H&E and IHC for pSTAT3^{Y705} in a WT ER$^+$/PR$^-$ PDX (STG335) (Bruna et al., 2016) explant

treated with vehicle (Veh), IL6 (50 ng/ml) or IL6+Ruxolitinib (500 nM) for two days *ex vivo*. Scale bar = 50 μm. **(F)** Quantification of pSTAT3^{Y705} IHC staining in the PDX explant from panel E using the Halo TM (Indica labs) Image analysis platform. Error bars represent SEM between different tumor pieces (n=3-4) from different parts of the PDX tumor cultured on the same sponge. *p<0.05, **p<0.01 (Student's t-test, two-sided). **(G)** Quantification of PR expression in the PDX from Figure 6E as determined by IHC. Error bars represent SEM, n=4-5. **p<0.01, ***p<0.001 (One way ANOVA, followed by Tukey's multiple comparison test).