VGLL1-directed TEAD activation drives endocrine therapy resistance in estrogen receptor positive breast cancer

Carolina Gemma1,*, Anup K Singh1, Antonino Belfiore2, Chun-Fui Lai1, Manikandan Periyasamy1, Sara Abuelmaaty3,4, Claire M Davies5, Xiangyi Wang1, Georgia M. Simmons1, Van T M Nguyen1, Luca Magnani1, Emad Rakha3,4, Lesley-Ann Martin6, R. Charles Coombes1, Giancarlo Pruneri2, Lakjaya Buluwela1, Simak Ali1,*

1Department of Surgery and Cancer, Imperial College London, Hammersmith Campus, London, W12 0NN, UK.
2Pathology Department, Fondazione IRCCS Istituto Nazionale Tumori and University of Milan, School of Medicine.
3Division of Cancer and Stem Cells, Nottingham Breast Cancer Research Centre, University of Nottingham Biodiscovery Institute, University Park, Nottingham, NG7 2RD, UK.
4Department of Histopathology, Nottingham University Hospital NHS Trust, City Hospital Campus, Hucknall Road, Nottingham, NG5 1PB, UK.
5ECMC Imperial College. Department of Surgery and Cancer, Imperial College London, Hammersmith Campus, London, W12 0NN, UK.

* To whom correspondence should be addressed: c.gemma@imperial.ac.uk; simak.ali@imperial.ac.uk
Abstract

Endocrine therapies are standard-of-care treatments for estrogen receptor (ER) positive breast cancer. However, patients with ER+ breast cancer develop resistance to these therapies and most relapsed patients die with endocrine-resistant metastatic disease. Here we show that resistance to the ER degrader, fulvestrant, is accompanied by epigenetic activation of the transcriptional co-activator VGLL1. Rewiring of the epigenome in therapy resistant cells also results in increased binding of the transcription factor TEAD4. Through interaction with TEAD4, VGLL1 induces the expression of genes implicated in cell proliferation in the resistant cells. We demonstrate that VGLL1 is necessary for the growth of fulvestrant-resistant breast cancer cells. Pharmacological disruption of VGLL1/TEAD4 interaction blocked growth of fulvestrant-resistant cells, accompanied by inhibition of VGLL1/TEAD transcriptional programmes. Furthermore, we identify EGFR as an important downstream VGLL1 target, whereby VGLL1-directed EGFR upregulation sensitises fulvestrant-resistant breast cancer cells to EGFR inhibitors. Taken together, our findings identify VGLL1 as a key transcriptional driver in endocrine-resistant breast cancer and identify new therapeutic approaches for advanced breast cancer patients.
Introduction

ER is the key transcriptional driver of tumour growth in three-quarters of breast cancers\(^1\). As such, endocrine therapies are the main therapeutic options for ER positive (ER+) breast cancer. These drugs inhibit ER signalling through distinct mechanisms. Aromatase inhibitors (AI) block estrogen biosynthesis to prevent ER activation\(^2\). Selective estrogen receptor modulators (SERMs), typified by tamoxifen\(^3\), bind to ER to inhibit its activity, while fulvestrant represents a class of anti-estrogens whose binding to ER promotes its degradation\(^4\). Despite the effectiveness of these treatments, many women will present with intrinsic or acquired resistance to these drugs\(^1-3,5\), necessitating determination of mechanisms of resistance and the development of new therapies.

Cell identity is established through epigenetic activation of distal regulatory elements that drive cell-type-specific gene expression programs in development and differentiation\(^6,7\). Cancer cells are also characterised by specific transcriptional programmes, frequently defined by cancer-type-specific epigenetic states\(^8,9\), while altered epigenetic landscapes are likely to signpost therapy resistance pathways\(^10\).

Indeed, profiling isogenic ER+ breast cancer cell models of resistance to different endocrine therapies, for acetylation of the active transcription histone mark, histone 3-lysine 27 (H3K27ac)\(^11-15\), has demonstrated extensive and drug-specific epigenetic reprogramming\(^16\), resulting in sweeping enhancer landscape reorganization\(^17\).

The TEA domain (TEAD) family of transcription factors have been implicated in cancer, by virtue of the fact that they are DNA binding partners for Yes-associated protein (YAP), and its paralog, transcriptional co-activator with PDZ-binding motif (TAZ)\(^18-21\). YAP/TAZ are downstream effectors of the Hippo pathway, required for organogenesis, tissue homeostasis and as important players in cancer initiation and progression, including resistance to cancer therapies\(^22,23\). Very little is known about the second family of TEAD co-activators, named vestigial-like (VGLL) due to their homology to the Drosophila co-activator vestigial (vg). Vestigial (vg) is a master regulator of wing development in Drosophila, regulating gene expression upon dimerization with scalloped (sd), the Drosophila homologue of mammalian TEAD\(^24-27\).

There are four mammalian vg-like genes (VGLL1-4), all of which function through...
interaction with TEADs\textsuperscript{28,29}. Human VGLL1 can substitute for \textit{Drosophila} vg in wing formation\textsuperscript{30}, underlining the conserved role of VGLL1 in driving TEAD-directed gene expression. Indeed, VGLL1 competes with YAP for binding to TEAD4 \textit{in vitro}, and the structural similarity between VGLL1-TEAD and YAP-TEAD complexes has suggested that VGLL1 may have a function similar to that of YAP\textsuperscript{28,31}.

Fulvestrant is a key treatment for advanced ER+ breast cancer following relapse on prior endocrine therapies\textsuperscript{32,33}, with particular utility in combination with drugs targeting other pathways, e.g. CDK4/6 inhibitors\textsuperscript{34} and for treating AI-resistance in metastatic breast cancer due to activating ESR1 mutations\textsuperscript{35-37}. Here, we show that VGLL1 is greatly upregulated in fulvestrant-resistant breast cancer and reveal a dependency of fulvestrant-resistant breast cancer cells on VGLL1. Indeed, overexpression of VGLL1 in endocrine-sensitive breast cancer cells was sufficient to promote resistance to fulvestrant. With ChIP-seq for VGLL1 and TEAD4, together with RNA-seq following VGLL1 knockdown, we demonstrate that VGLL1 drives resistance to fulvestrant through induction of TEAD target genes and identify potential therapeutic approaches for fulvestrant-resistant breast cancer.

\textbf{Results}

\textbf{Epigenetic plasticity results in the induction of VGLL1 expression in endocrine therapy resistant breast cancer.} To discover genes that are activated following epigenetic reprogramming with endocrine therapies, we analysed H3K27ac ChIP-seq data from well-established endocrine therapy resistant breast cancer cell lines\textsuperscript{16}. We used H3K27ac ChIP-seq signal at gene promoters as a mark of active promoters, as previously reported\textsuperscript{38,39}, to rank gene promoters according to the ratio of H3K27ac signal in fulvestrant-resistant (FULVR) MCF7 breast cancer cells relative to the isogenic fulvestrant-sensitive MCF7 cells. The VGLL1 and VGLL3 genes were ranked third and fourth highest for H3K27ac gain in FULVR cells (Fig. 1a, b). By contrast, H3K27ac was reduced at promoters of ER target genes, including \textit{GREB1} and \textit{PGR} in FULVR cells. Epigenetic activation of \textit{VGLL1} and \textit{VGLL3} was accompanied by strong increases in their expression in FULVR cells and reductions in ER target gene expression in RNA-seq analysis (Supplementary Fig. 1a). Interestingly, \textit{VGLL1} and
VGLL3 expression was also higher in tamoxifen resistant (TAMR) MCF7 cells, but not in long-term estrogen-deprived (LTED) cells (mimicking AI resistance) (Supplementary Fig. 1b), suggesting a mechanism for VGLL1/3 induction that requires inhibition of ER activity. Indeed, our previous work revealed that in LTED cells activation of cholesterol biosynthesis pathways, generate the ER ligand 27-hydroxycholesterol\(^\text{16}\). Hence, ER is not inhibited in LTED cells. RT-qPCR and immunoblotting confirmed the results of the H3K27ac ChIP-seq and RNA-seq analyses (Fig. 1c, d). Expression of several TEADs was also increased in MCF7-FULVR cells. YAP and TAZ, the best-known cofactors for TEADs, were expressed in MCF7 cells, but levels of YAP were reduced in FULVR cells. Other genes implicated in the HIPPO pathway, were unchanged in MCF7-FULVR cells (Fig. 1d and Supplementary Fig. 1c). VGLL1 and VGLL3 were similarly over-expressed in independently derived MCF7-FULVR cells (Supplementary Fig. 1d, e). Furthermore, VGLL1, but not VGLL3, was induced in additional FULVR cells generated in other ER+ breast cancer cell lines. In these additional FULVR cell lines, expression of YAP, TAZ and other HIPPO components was unchanged, or was reduced (Supplementary Fig. 1d, e).

We also measured VGLL1 expression in a cohort of patients with tumour biopsies collected prior to and following progression on fulvestrant, using RNA in situ hybridisation and immunohistochemical detection. In agreement with the results in FULVR cell lines, patients with breast cancer relapse post-fulvestrant treatment had substantially higher VGLL1 mRNA and protein (Fig. 1e and Supplementary Fig.1f). Collectively, these data clearly show that upregulation of the TEAD coactivator VGLL1 is a frequent feature of fulvestrant-resistant breast cancer.

**VGLL1 expression is induced upon ER downregulation.** Since VGLL1 expression emerges in fulvestrant resistance, we determined if anti-estrogen-directed ER inhibition can induce VGLL1 expression. Indeed, 24 or 48 hours of fulvestrant treatment stimulated VGLL1 expression in MCF7 cells (Fig. 1f and Supplementary Fig. 2a) and in other ER+ breast cancer cell lines (Fig. 1f and Supplementary Fig. 2c). RNAi-mediated ESR1 knockdown similarly promoted VGLL1 expression in all cell lines (Fig. 1g and Supplementary Fig. 2d, e). However, change in VGLL3 expression by fulvestrant or ESR1 siRNA was less consistently observed, being stimulated in...
MCF7 and BT474, but not in the other cell lines (Supplementary Fig. 2c, d, f). Other HIPPO pathway genes were not affected by fulvestrant or ER knockdown, although small increases in the expression of some TEADs and the LATS1/2 HIPPO pathway kinases (which promote YAP/TAZ degradation), were observed (Supplementary Fig. 2b, c, d, f). The important clinical activity of fulvestrant, has led to the development of next generation SERDs, some of which have progressed through clinical testing. As seen for fulvestrant, VGLL1 expression was increased by other SERDs (GDC-0810, AZD9496, RAD1901), with accompanying inhibition of ER activity (Supplementary Fig. 2g). Together, these data show clearly that downregulation of ER is sufficient to induce VGLL1 expression in ER+ breast cancer cells.

VGLL1 is recruited to TEAD binding sites at active regulatory regions in fulvestrant-resistant breast cancer cells. To investigate the mechanistic roles of VGLL1 and TEADs in fulvestrant resistance, we performed ChIP-seq for VGLL1 in MCF7-FULVR cells and for TEAD4, as it is the best characterized TEAD in cancers (Supplementary Fig. 3a). Remarkably, three times more TEAD4 binding events were identified in FULVR cells than in the isogenic MCF7 cells (20,894 vs 6,968 peaks, respectively) (Fig. 2a). Most (75%) of the TEAD4 binding events in MCF7 cells were also present in FULVR cells. TEAD binding sites were by far the most strongly enriched motifs identified with de novo motif analysis in MCF7 and FULVR cells (Fig. 2a and Supplementary Fig. 3b). These findings suggest that TEAD4 activity is substantially elevated in FULVR cells. Also enriched at TEAD4 binding regions in FULVR cells were activating protein-1 (AP-1) binding sequences (Supplementary Fig. 3b), consistent with a previous report showing that TEADs and AP-1 are part of a multiprotein complex that regulates genes involved in proliferation and migration in cancer cells.

The transcriptional activity of TEADs requires binding of co-activators, including YAP/TAZ18-20 and VGLLs28,29. ChIP-seq for VGLL1 in FULVR cells, identified 29,930 VGLL1 peaks (q-value < 0.01; Supplementary Table 1 and Supplementary Fig. 3c). There was a strong positive correlation between VGLL1 and TEAD4 binding events (p<2.2x10^-16; Fig. 2b, c). Moreover, the TEAD4 binding sequence was the most highly enriched motif at VGLL1 binding regions (Supplementary Fig. 3d), indicative of VGLL1 recruitment to the TEAD DNA binding proteins. Indeed, VGLL1 peaks were highly
enriched at the TEAD4 binding sites in FULVR cells ($p<0.0001$, Fig. 2d), indicating that VGLL1 recruitment to chromatin is mainly through interaction with TEADs. Regions co-bound by VGLL1 and TEAD4 were present at active regulatory regions in FULVR cells and were strongly associated with active chromatin, as revealed by correlation with H3K27ac signal and reduced nucleosome occupancy at the peak centre (Fig. 2e, f). Moreover, in MCF7 cells these regions were associated with an inactive chromatin conformation (Supplementary Fig. 3f), indicating that the higher TEAD4 occupancy observed in FULVR cells is associated with increased chromatin accessibility at the TEAD4 binding regions. The vast majority of the VGLL1 and TEAD4 peaks were present at intergenic and intronic regions (Supplementary Fig. 3e), suggesting that they are mainly present at distal enhancers, as also shown previously for YAP/TEAD binding in other cell types\textsuperscript{18,43,47}. In MCF7-FULVR cells, VGLL1 and TEAD4 were commonly co-bound at known YAP and TEAD target genes, including CTGF, AMOTL2 and ANKRD1\textsuperscript{47,48}, and was accompanied by increased H3K27ac at these genes (Fig. 2g). Consistent with this, expression of these genes was greatly elevated in MCF7-FULVR cells, compared to MCF7 cells. ChIP-qPCR confirmed that VGLL1 was recruited to these binding sites in MCF7-FULVR cells, but not in MCF7 cells (Supplementary Fig. 3g). YAP and TAZ were largely absent at these genes in FULVR cells, suggesting that VGLL1 is the main TEAD co-activator in driving expression of these genes in MCF7-FULVR cells. Collectively, our data identify VGLL1 and TEAD4 as previously unknown core components of the regulatory landscape in breast cancer cells resistant to fulvestrant.

**VGLL1 promotes resistance of breast cancer cells to fulvestrant.** To gain further insights into the transcriptional program regulated by VGLL1, we performed functional annotation of the VGLL1 binding regions in FULVR cells. Notably, genes associated with VGLL1 peaks were highly enriched in gene sets linked to acquired resistance to endocrine therapies in breast cancer cells (Supplementary Fig. 4a and Supplementary Table 2). Moreover, siRNA-directed VGLL1 knockdown reduced growth of fulvestrant-resistant MCF7 and T47D cells (Fig. 3a and Supplementary Fig. 4b), demonstrating its functional importance in resistant cells.

We next asked whether activation of the endogenous VGLL1 gene is sufficient for development of resistance to fulvestrant. Applying the CRISPR/Cas9 Synergistic
Activation Mediator (SAM) method\textsuperscript{49}, using a specific sgRNA targeted to the VGLL1 gene promoter (Fig. 3b), we successfully established VGLL1 expression in MCF7 cells (Fig. 3c and Supplementary Fig. 4c). As reported previously, treatment of ER+ breast cancer cells with fulvestrant results in rapid and near complete cell loss and resistant cells only emerge after 3-6 months\textsuperscript{50-52}. Culturing MCF7 cells with transcriptional activation of endogenous VGLL1 (MCF7 ActCas9-VGLL1) in the presence of 100 nM fulvestrant showed an initial period of growth inhibition, as did the control (MCF7 ActCas9-Vector) cells (Fig. 3d). Remarkably, growth recovery of MCF7 ActCas9-VGLL1 cells was observed after ~20 days, with no cell recovery evident in the control cells even after 100 days (Fig. 3d and Supplementary Fig. 4d). We further evaluated two of the independently generated ActCas9-VGLL1 fulvestrant-resistant lines, named MCF7 ActCas9-VGLL1-FULVR-1 and -2. Interestingly, development of fulvestrant resistance in these cells was accompanied by further upregulation of VGLL1, with a concomitant downregulation in the mRNA and protein levels of ER and PR (Fig 3e, f and Supplementary Fig. 4e), as noted for the acquired fulvestrant-resistant models (Fig. 1c and Supplementary Fig. 1d). Moreover, comparison of global gene expression profiles generated by RNA-seq between MCF7 ActCas9-VGLL1-FULVR cells and other fulvestrant-resistant cell lines, revealed that similar transcriptional changes were observed in the ActCas9-VGLL1-FULVR cells, as in the MCF7-FULVR cells (Supplementary Fig. 4f). The growth dependence of MCF7 ActCas9-VGLL1-FULVR cells on VGLL1 could be confirmed by siRNA-mediated VGLL1 knockdown (Fig. 3g), indicating that MCF7 ActCas9-VGLL1-FULVR cells become dependent on VGLL1 for survival, as we observed in the other fulvestrant resistance models. Moreover, not only were MCF7 ActCas9-VGLL1-FULVR cells resistant to fulvestrant (Fig. 3h) but were also insensitive to next generation SERDs (Supplementary Fig. 4h).

The above results suggest a switch in dependency from ER in the ER+ treatment-naïve, breast cancer cells, to VGLL1 in the fulvestrant-resistant cells. Indeed, induction of VGLL1 expression in the fulvestrant-naïve MCF7 cells was insufficient to induce a shift in the transcriptional program (Supplementary Fig. 4g), which continued to be dependent on ER, as evidenced by the maintained sensitivity to fulvestrant (Fig. 3h). These data suggest that downregulation of ER and ER transcriptional activity following fulvestrant treatment is a necessary first step for adaptation to VGLL1 dependence. Taken together, our data show that fulvestrant-resistant breast cancer cells are
dependent on VGLL1 for their growth and that resistance to fulvestrant is rapidly
installed in a VGLL1-dependent manner.

**VGLL1 transcriptional dependency in fulvestrant-resistant breast cancer.** To
explore the role of VGLL1 as a coactivator in fulvestrant resistance, we next asked
whether FULVR cells rely on VGLL1 to activate their transcriptional programs. Indeed,
analysis of our VGLL1 ChIP-seq data revealed that genes predicted as VGLL1 targets
and associated with two or more VGLL1 peaks were significantly more highly
expressed in the fulvestrant-resistant than in the fulvestrant-sensitive MCF7 cells
compared to genes not associated with VGLL1 peaks (Fig. 4a), a finding confirmed in
the MCF7 ActCas9-VGLL1-FULVR cells (Supplementary Fig. 5a).

To confirm these findings using high confidence VGLL1 direct targets, we performed
RNA-seq following siRNA-mediated VGLL1 downregulation in FULVR cells. We
identified a set of direct VGLL1 targets as those genes that were downregulated upon
VGLL1 downregulation in FULVR cells, which we further filtered based on the
presence of at least one associated VGLL1 binding site, which generated a list of 762
genes (VGLL1-activated genes). Genes whose expression was not altered by siRNA-
downregulation that were also further filtered based on the absence
of VGLL1 associated peaks were considered as not-VGLL1 targets (n=8,932 genes).
(Supplementary Fig. 5b, c). We found that VGLL1 activated genes were significantly
more highly expressed in MCF7-FULVR cells than the isogenic fulvestrant-sensitive
MCF7 cells, compared to not-VGLL1 targets (Fig. 4b). Moreover, VGLL1 activated
genes were highly enriched in VGLL1 and TEAD4 peaks (Supplementary Fig. 5d, e)
and were associated with higher levels of H3K27ac at gene promoters in FULVR cells
compared to MCF7 cells (Fig. 4c and Supplementary Fig. 5f), further supporting the
notion that VGLL1 regulates genes that become more active in the resistant cells. In
line with this, TEAD4 peaks associated with the VGLL1-activated genes were present
at regions with higher H3K27ac signal and TEAD4 recruitment in FULVR cells
compared to MCF7 cells (Supplementary Fig. 5g). Taken together, these data indicate
that VGLL1 drives expression of TEAD-target genes in FULVR cells. Interestingly, we
also found that genes whose expression depends on VGLL1 in FULVR cells were
significantly more expressed than not-VGLL1 targets (Fig. 4d). Functional annotation
of the VGLL1-activated genes revealed enrichment for pathways associated with
growth factor signalling, including growth factor binding and transmembrane receptor protein kinase activity (Fig. 4e). Indeed, genes included in these functional categories, e.g. IGFBP3, ITGB6, EGFR and TGFB2, were among the VGLL1 target genes most highly upregulated in fulvestrant-resistant compared to fulvestrant-sensitive cells (Fig. 4f, g and Supplementary Fig. 5h). Collectively, these data indicate that VGLL1 acts as an essential transcriptional co-factor in FULVR cells, driving the gene expression programs that facilitate survival and growth of fulvestrant-resistant breast cancer cells.

**VGLL1 drives EGFR expression in breast cancer cells resistant to fulvestrant.**

Our findings link VGLL1 activity with genes associated with receptor tyrosine kinase (RTK) signalling. High expression of the RTKs, EGFR and HER2, has been linked to reduced therapeutic efficacy of anti-estrogens, including tamoxifen and fulvestrant and ectopic expression of EGFR promotes fulvestrant resistance in breast cancer cells. The TEAD4 ChIP-seq showed TEAD4 recruitment to the EGFR enhancer and promoter region in MCF7-FULVR, but not in MCF7 cells (Fig. 5a). VGLL1 co-localised with TEAD4 in MCF7-FULVR cells at the EGFR enhancer. Interestingly, EGFR, together with VGLL1, were among the top genes with higher TEAD4 occupancy in FULVR cells compared to MCF7 cells (Supplementary Fig. 6a). VGLL1 and TEAD4 co-binding at the EGFR enhancer, together with acquisition of H3K27ac, was also evident in MCF7-ActCas9-VGLL1-FULVR cells (Fig. 5b), confirming the binding of VGLL1 and TEAD4 to the regulatory regions of the EGFR gene in fulvestrant-resistant cells. VGLL1 and TEAD4 co-binding at the EGFR enhancer was accompanied by a large increase in EGFR expression in MCF7-ActCas9-VGLL1-FULVR cells (Fig. 5c, d). Importantly, siRNA-mediated downregulation of Cas9 in these cells led to a decrease in VGLL1 and EGFR expression (Fig. 5e). Similarly, EGFR expression was significantly increased in several fulvestrant-resistant cells in which VGLL1 expression was induced, compared to their isogenic fulvestrant-sensitive breast cancer cell lines (Supplementary Fig. 6b). In agreement with these findings, VGLL1 knockdown was sufficient to reduce the levels of EGFR in fulvestrant-resistant breast cancer cell lines (Supplementary Fig. 6c, d, e).

VGLL1-induced EGFR expression in MCF7ActCas9-FULVR cells was accompanied by a strong increase in AKT and ERK1/2 MAPK phosphorylation (Fig. 5d), while VGLL1 knockdown also reduced AKT and ERK1/2 phosphorylation (Supplementary
Fig. 6d, e), indicating that VGLL1-directed EGFR over-expression in FULVR cells facilitates activation of downstream signalling pathways that promote cell survival and proliferation\textsuperscript{55,56}. MCF7 cells are characterised by low EGFR and HER2 expression and insensitivity to EGFR inhibitors\textsuperscript{57}. Consistent with high EGFR expression and activity, treatment of MCF7-ActCas9-VGLL1-FULVR cells with the EGFR inhibitor erlotinib\textsuperscript{58}, potently inhibited cell growth, whereas the parental fulvestrant sensitive cell lines (MCF7-ActCas9-Vector and MCF7-ActCas9-VGLL1) were largely insensitive to erlotinib (Fig. 5f). In line with this, erlotinib treatment led to reductions in EGFR phosphorylation and inactivation of the downstream signalling pathways in MCF7-ActCas9-VGLL1-FULVR cells (Fig. 5g). A similar result was obtained in additional fulvestrant resistant cell lines (Supplementary Fig.6f, g).

Our results provide a novel mechanism leading to increased EGFR signalling in breast cancer cells, in which VGLL1-TEAD4 co-binding at the EGFR enhancer induces EGFR expression. In agreement with this, EGFR expression was highly positively correlated with VGLL1 expression in ER+ breast cancer patients from TCGA\textsuperscript{59} and METABRIC\textsuperscript{60} (Fig. 5h and Supplementary Fig. 6i). Remarkably, EGFR was the highest-ranked gene co-expressed with VGLL1 in breast cancer patients from METABRIC and the highest ranked protein significantly correlated with VGLL1 in breast cancer from TCGA (Supplementary Fig. 6h).

Collectively, these data identify EGFR as an important VGLL1 target gene in fulvestrant-resistant breast cancer cells, in which rewiring of the epigenome results in VGLL1 and TEAD4 binding to the EGFR enhancer, leading to induction in EGFR expression and activation of downstream EGFR signalling pathways (Fig. 6i).

**Targeting VGLL1 transcriptional activity with verteporfin.** Our findings identify VGLL1 as a new therapeutic target in metastatic, endocrine-resistant breast cancer. Verteporfin (VP), a drug used for photodynamic treatment of macular degeneration, has been identified as an inhibitor of YAP/TEAD. VP disrupts the physical interaction between YAP and TEADs, resulting in downregulation of YAP/TEAD target genes\textsuperscript{61,62}. Given that the VGLL1-TEAD interaction is structurally similar to that of YAP-TEAD\textsuperscript{28,31}, we reasoned that VP would disrupt VGLL1-TEAD4 interaction and so inhibit expression of VGLL1-regulated genes. Indeed, VP treatment significantly reduced the
growth of fulvestrant-resistant breast cancer cells (Fig. 6a and Supplementary Fig. 7a, b). Moreover, ChIP-qPCR for VGLL1 and TEAD4 in MCF7-FULVR cells treated with VP for 24h revealed that VP reduced the binding of VGLL1 to the target genes, including EGFR (Fig. 6b), whereas TEAD4 binding at target genes was either unaffected or only moderately reduced by VP, indicating that VP prevents binding of VGLL1 to TEAD4. To assess whether the VGLL1 co-transcriptional activity is inhibited by VP, we performed RNA-seq in FULVR cells following treatment with VP. Remarkably, we found that VP strongly and very significantly downregulated the expression of VGLL1 activated genes including EGFR, and several other VGLL1 target genes associated with VGLL1 binding at the promoter or enhancer region, including classical TEAD targets such as CTGF, as well as the previously described VGLL1 target genes IGFBP5 and MMP9, which we also found to be VGLL1 activated genes in FULVR cells (Fig. 6c, d and Supplementary Fig. 7c, d). Importantly, we found that the most significantly downregulated genes after treatment with VP were highly enriched for VGLL1 peaks (Fig. 6e). Indeed, genes with 2 or more associated VGLL1 peaks were preferentially downregulated by VP, compared to genes without VGLL1 peaks (Supplementary Fig. 7e). The selectivity of VP in downregulating VGLL1 target genes was further confirmed by analysing all the high confidence VGLL1 targets and not-VGLL1 targets, revealing that VGLL1 activated genes were significantly downregulated by VP compared to not-VGLL1 targets (Fig. 6f, g). Indeed, the majority of VGLL1 activated genes were downregulated by VP and displayed preferential sensitivity to VP (Supplementary Fig. 7f). In line with this, we also found that genes downregulated by VP were expressed at higher levels, compared to genes not inhibited by VP, in FULVR cells (Supplementary Fig. 7g), consistent with VGLL1 activated genes being more highly expressed than not-VGLL1 targets in FULVR cells (Fig. 4d). In agreement with the above results, we found that genes downregulated by VP were significantly enriched in functional categories attributed to VGLL1 activated genes such as growth factor binding, extracellular matrix binding and transmembrane receptor protein kinase activity (Supplementary Fig. 7h and Supplementary Table 3).

Interestingly, VGLL1 expression levels were associated with expression of VGLL1 activated genes also in breast cancer patients (METABRIC) (Fig 6h). Moreover, when we stratified patients according to the expression of the combined VGLL1 activated genes, we found that patients with higher expression levels of this signature had a
worse prognosis (Fig. 6i). A similar result was obtained when considering only ER+ breast cancer cases (Supplementary Fig. 7i), indicating that our VGLL1 transcriptional signature of VGLL1 activated genes has useful prognostic value for breast cancer patients treated with endocrine therapies. In summary, our results show that VGLL1 transcriptional activity is sensitive to VP and suggest that inhibition of VGLL1 recruitment to TEADs, as with VP, could provide an approach to treating VGLL1-dependent, endocrine-resistant breast cancer.

Discussion

Fulvestrant is the main endocrine agent for advanced ER+ breast cancer following relapse on prior endocrine therapies32,33 and newer generation SERDs with improved pharmacological properties are in advanced clinical development41,42. Understanding resistance to SERDs is urgently needed for aiding patient selection and in identifying therapeutic interventions for patients who progress on SERDs. Our work has identified VGLL1/TEAD4, as previously unknown key transcriptional regulators and drivers of resistance to fulvestrant and suggests new therapeutic opportunities for endocrine-resistant breast cancer. We show that SERD inhibition of estrogen-regulated gene expression is accompanied by induction of VGLL1 expression. The induction of VGLL1 expression upon ER inhibition drives fulvestrant resistance through the establishment of a VGLL1/TEAD directed transcriptional program that supports cell survival and proliferation. Our work strongly supports the idea that VGLL1/TEAD4 act as transcriptional master regulators in FULVR cells and that resistance to fulvestrant emerges as a consequence of a switch in transcriptional dependencies from ER in pre-treatment ER+ breast cancer, to VGLL1/TEAD in fulvestrant-resistant breast cancer. The clinical significance of our findings is further emphasized by the validity of the signature of VGLL1 activated genes as a prognostic tool in a large set of breast cancer patients treated with endocrine therapies.

The well-established oncogenic role of TEADs in different cancers has been closely linked to the activity of YAP/TAZ as TEAD coactivators in the majority of studies43,48. However, it is increasingly clear that the canonical functions of TEADs as downstream effectors of YAP and TAZ in the Hippo pathway are not the only mechanism by which
TEADs contribute to cancer progression in different cancers\textsuperscript{44,64,65}. For instance, in ER+ breast cancer, a non-canonical function of YAP and TEAD4, has recently been reported, where they act as ER cofactors, thereby regulating ER target genes rather than canonical TEAD target genes\textsuperscript{44}. Our results are in agreement with these findings, as we found that classical TEAD target genes were expressed at low levels in ER+ breast cancer cells and we did not observe enrichment of YAP/TAZ binding at these target genes. Our work also revealed that many of the established YAP/TAZ/TEAD target genes\textsuperscript{47,66,67}, are upregulated in FULVR cells, but YAP/TAZ are largely absent at these genes. Instead, VGLL1 was co-bound with TEAD at these genes in FULVR cells and the expression of these genes was regulated by VGLL1.

Several lines of evidence indicate that endocrine therapies are drivers of epigenetic reprogramming\textsuperscript{16,68}. Our results further show that reprogramming of the epigenetic landscape in breast cancer cells resistant to fulvestrant is associated with increased TEAD4 recruitment to the chromatin and association of VGLL1/TEAD4 with active enhancers in the resistant cells, and uncover a previously unrecognized role for VGLL1/TEAD4 as transcriptional drivers in advanced ER+ breast cancer.

Additionally, our work identified EGFR as a VGLL1/TEAD activated gene in FULVR cells. We show that instalment of fulvestrant resistance following VGLL1 overexpression was accompanied by VGLL1 and TEAD4 recruitment to the EGFR enhancer, upregulation of EGFR, associated activation of EGFR downstream effectors and sensitivity to EGFR inhibition. Recent studies show that EGFR and HER2 amplification, as well as mutations in HER2 and genes in the MAPK pathway are increased in metastatic, endocrine-resistant ER+ breast cancer, together accounting for 10-15\% of cases\textsuperscript{54,69} and demonstrating the importance of elevated EGFR and HER2 signalling in acquired resistance to endocrine treatments. Our results evidence an alternate mechanism by which EGFR expression is induced in endocrine-resistant breast cancer and thus advance a rationale for expanded clinical utility for EGFR inhibitors.

Targeting transcription factors in cancer has proved to be challenging, yet it offers great promise for cancer treatment as many cancers are driven by aberrant transcription factor expression and/or activities\textsuperscript{70}. As the importance of VGLL1 in
cancer has largely been unexplored, there are at present no reported VGLL1 inhibitors. Given the well-established functions of YAP/TAZ/TEAD as tumor promoters in many cancers, there have been considerable efforts to identify inhibitors of their activity. Elucidation of the structure of the VGLL1/TEAD complex has revealed that VGLL1-TEAD interaction is structurally similar to that of YAP-TEAD. Indeed, we show that verteporfin, a FDA approved photodynamic therapy drug for macular degeneration, which disrupts YAP/TEAD interaction, also disrupts the binding of VGLL1 to TEAD4 at the target genes, and selectively inhibits VGLL1 transcriptional activity. Thus, our findings show that VGLL1 transcriptional and growth dependencies could be exploited as a therapeutic vulnerability in advanced ER+ breast cancer and so inhibiting VGLL1 interaction with TEAD, as well as inhibition of downstream VGLL1-activated genes such as EGFR, could be a viable therapeutic approach for patients who progress on endocrine therapies.

Collectively, our work uncovers multiple new insights into the biological function of the TEAD coactivator VGLL1 in advanced therapy-resistant breast cancer. The biological function and importance of VGLL1 in other cancers remains poorly understood. However, because of the tissue-restricted expression of VGLL1 and its specific upregulation in several cancers, it is tempting to speculate that VGLL1 could emerge as a promising new therapeutic target in several cancers. Thus, our findings and the VGLL1-EGFR connection may be relevant in other cancers where VGLL1 upregulation is observed.

**Methods**

Methods, including statements of data availability and any associated accession codes and references, are available online at XXX

**Methods**

**Cell lines and treatments.** Cells lines were originally purchased from ATCC, were authenticated at the time of the studies by LGC Standards (Bury, UK) and were regularly screened for mycoplasma infection. BT474, ZR-75-1 and T47D cells were cultured in Dulbecco's modified Eagle's medium (DMEM) with 10% of fetal bovine serum (FBS) and 100U penicillin/0.1 mg ml⁻¹ streptomycin. MCF7 cells were cultured
in DMEM with 10% of fetal bovine serum (FBS) and 100U penicillin/0.1 mg ml\(^{-1}\) streptomycin plus 10nM estradiol (SIGMA E8875). MCF7(2) cells were cultured in RPMI 1640 with 10% of fetal bovine serum (FBS) and 100U penicillin/0.1 mg ml\(^{-1}\) streptomycin plus 1nM estradiol. All fulvestrant resistant isogenic derivatives of ER+ breast cancer cell lines have been previously described \(^{16,50,73,74}\), and were cultured as the corresponding parental cells with the addition of 100nM fulvestrant (Sigma I4409).

For VGLL1 overexpression, MCF7-ActCas9-VGLL1 cells were generated by transducing MCF7 cells with lentiviral MS2-P65-HSF1_Hygro (Addgene plasmid #61426), lentiviral dCAS-VP64_Blast (Addgene plasmid #61425) and sgRNA(MS2) cloning backbone (Addgene plasmid #61424) using the following VGLL1 sgRNA: ATTCCTGCAGGTGCCCAACCAGG, that was cloned into the sgRNA(MS2) cloning backbone plasmid. The control MCF7-ActCas9-Vector cells were generated simultaneously using the above mentioned plasmids, except that sgRNA(MS2) cloning backbone empty vector was used instead. MCF7-ActCas9-Vector and MCF7-ActCas9-VGLL1 cells were maintained in media containing 0.2 mg/ml hygromycin (Corning 30-240-CR), 10 \(\mu\)g/ml blasticidin (Corning 30-100-RB) and 1\(\mu\)g/ml puromycin (gibco A1113803).

MCF7-ActCas9-VGLL1-FULVR cells were generated by culturing MCF7-ActCas9-VGLL1 cells in the appropriate media with the addition of 100nM or 1000nM fulvestrant.

All drugs were solubilised in DMSO. Verteporfin (Sigma SML0534), was used at a final concentration of 2\(\mu\)M for 24 h, unless otherwise specified. SERMs: tamoxifen (Sigma H7904) and raloxifene (Tocris 2280), as well as SERDs: fulvestrant, GDC-0810 (Genentech), AZD9496 (Astra Zeneca) and RAD1901(MedChem Express HY-19822A) were used at a final concentration of 100nM for 48 h, unless otherwise specified. Erlotinib (VWR CAYM10483) was used at a final concentration of 3.1 uM for 24 h, unless otherwise specified.

Clinical samples. From the institutional database of the Fondazione IRCCS Istituto Nazionale Tumori of Milan, we retrospectively identified 72 metastatic ER+, HER2 negative breast cancer patients diagnosed between 1997 and 2018, treated with endocrine therapy, including at least one line of fulvestrant, and standard
chemotherapy. For the present study, we selected 15 patients for which breast cancer biopsies were available both prior to fulvestrant (including primary tumor samples and metastatic localizations) and post fulvestrant administration. All selected cases were carefully revised by an expert breast pathologist, including the evaluation of ER and PR receptor status, proliferative index and tumor cellularity. All patients included in this study gave their written consent to donate the tissue remaining after their diagnostic procedures to Istituto Nazionale Tumori of Milan.

siRNA transfections. siRNAs (25nM) were transfected with Lipofectamine RNAi-MAX (Life Technologies) in antibiotics-free medium according to the manufacturer’s instructions. Cells were harvested 72 h after transfection. The following individual ON TARGETplus siRNAs (Horizon Discovery) were used for VGLL1: J-017939-06 (siVGLL1 #1), J-017939-08 (siVGLL1 #2) and J-017939-07 (siVGLL1 #3). For dCAS9 the following siRNAs were used: CCGAAGAGGUCGUGAAGAA (Cas9 siRNA#1) and GGGAAAGAUCGAGAAGAU (Cas9 siRNA#2). ESR1 siRNA was purchased from Qiagen (SI03114979, FlexiTube siRNA). The negative control siRNA (siControl) was purchased from Horizon Discovery (ON TARGETplus Non-targeting control siRNA #2, D-001810-02)

RNA in situ hybridization. Formalin fixed, paraffin embedded breast cancer tissue sections (5 μm) from matched pre- and post-fulvestrant patient biopsies were processed for RNA in situ detection using the RNAscope Multiplex Fluorescent Reagent Kit v2 according to the manufacturer’s instructions (Advanced Cell Diagnostics), using RNAscope probe Hs-VGLL1 446731. Nuclei were stained with DAPI and the RNAscope probe was detected using Cy5 fluorescent dye (PerkinElmer NEL745E001KT TSA Plus Cyanine 5). Data was analysed using Fiji and CellProfiler using custom macros, provided upon request.

Immunohistochemistry. Immunohistochemical staining was performed on formalin-fixed, paraffin embedded 4 μm tissue sections using VGLL1 polyclonal antibody (Proteintech, 10124-2-AP). Antigen retrieval was performed based on the manufacturer’s recommendations (citrate buffer pH 6.0 at 1000 W for 20 min using microwave). Visualization was performed using the Novocastra Novolink Polymer Detection Systems kit (RE7280-K, Leica, Biosystems, UK).
Western blot. Cells were washed twice in ice-cold PBS and harvested in RIPA Buffer (10 mM Tris-HCl (pH 7.6), 1mM EDTA, 0.1% SDS, 0.1% sodium deoxycholate, 1% Triton X-100 and phosphatase and protease inhibitors) and lysed by sonication. Samples were centrifuged at 4 °C, maximum speed, for 10 min, then the supernatant was transferred to a clean Eppendorf tube. Extracts were quantified using the Pierce BCA Protein Assay Kit, according to the manufacturer's instructions (ThermoFisher Scientific, 23227). Samples were run on 10% polyacrylamide gels and transferred onto polyvinylidene difluoride (PVDF) membranes by wet electrophoretic transfer. Blots were blocked with either 5% non-fat dry milk (non-phosphoprotein) or 5% BSA (phosphoprotein) and incubated with primary antibodies overnight at 4°C. Secondary antibodies were incubated for 1 h at room temperature and then blots were developed with goat anti-mouse (Biorad, 1706516) and anti-rabbit (Biorad, 1706515) horseradish peroxidase (HRP) conjugated secondary antibodies. SuperSignal West Pico PLUS Chemiluminescent Substrate (ThermoFisher Scientific, 34577) or SuperSignal West Femto Maximum Sensitivity Substrate (ThermoFisher Scientific, 34095) was used for chemiluminescent imaging using a Fusion solo (Vilber) imager.

Antibodies. The VGLL1 antibody was purchased from Proteintech (10124-2-AP). The ER antibody was from Leica Biosystem (NCL-ER-6F11). VGLL3 (ab68262), VGLL4 (ab140290), YAP1 (ab56701), TEAD4 (ab58310) and GAPDH (ab9484) antibodies were supplied by Abcam. Antibodies for PR (8757S), TEAD3 (13224S), EGFR (2232S), phospho-Tyr1068 EGFR (2234S), phospho-Tyr1148 EGFR (4404S), AKT (9272S), phospho-Ser473-AKT (9271S), p44/42 MAPK (4695S) and phospho-p44/p42 MAPK (Thr202/Tyr204) (9101S) were from Cell Signaling Technology. TEAD1 (sc-376113) and TEAD2 (sc-67115) antibodies were purchased from Santa Cruz Biotechnology and the TAZ antibody was obtained from BD (clone M2-616, 560235),

RT-qPCR. Total RNA was extracted from cells using RNeasy Mini Kit (Qiagen 74106). For cDNA synthesis 500 ng of total RNA was reverse transcribed using RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific, K1622). RT-qPCR reactions were carried out using Fast SYBR Green Master Mix (ThermoFisher Scientific,
4385616) in a StepOnePlus Real-Time PCR thermal cycler (ThermoFisher Scientific).

Gene expression levels were calculated relative to GAPDH. The following primers were used:

**VGLL1:** CCCCTCGAGTCAGAGTGAAG, CAGGGACGGTGAGAACTGAT;

**VGLL3:** GCTGGTAAGAGCTGGTCCAA, CCATCCAGAATCTGCCATTT;

**ESR1:** CAGGTGCCCTACTACCTGGA, TCCTTGCCAGATTTCCATAGC;

**PGR:** GATGCTTCATCCCCCCACAGAT, AGGTCTACCCGCCCTATCTC;

**GREB1:** AAGGAGGCTGGAAACAAAT, CGTGGAAATGGAGACAAGG;

**GAPDH:** TGGAAATCCCATCACCATCT, TTCACACCCATGACGAACAT;

**YAP1:** CAGCAACTGCAGATGGAGAA, TGGATTTTGAGTCCCACCAT;

**TAZ:** TCATCACCCTGCTCAATCAC, GTGGGAGTGTAGCTCCTTGG;

**TEAD1:** CAAGCCTTTTGTGCACAGCAG, AAAATCCACCAGGCGAAG;

**TEAD2:** CTTGGACTGGATTTCCCTTG, CCGCTACATCAAGCTGAGAA;

**TEAD3:** GAGGCAATGGTACGGTCCT, CTTTGGCAGATTCCATAGC;

**TEAD4:** GAGGCAATGGTACGGTCCT, CTTTGGCAGATTCCATAGC;

**LATS1:** ACCTTTCCAGCTCTGTTTGC, AGATCCTCGACGAGAGCAGA;

**LATS2:** GTGGTAGGACGCAAACGAAT, CCGAGGAATGAGCAGATTGT;

**BTRC:** ACAGGATCATCGGATTCCAC, TTGAAACGCAAGTGCAGAAC;

**SAV1:** TGGCTGGTATGTGACAGGAG, ACTTCCTCCTGGATGGGAAC;

**MOB1a:** ACAGCTTTGCTTCAGTGCAGA, TATGTGGCTTGAGGGAGAGG;

**EGFR:** GAGGGCAAAATACAGCTTTTG, GCCCTTCGACATTCTTACC;

**dCAS9:** GAACCGGATCTGCTATCTGC, CGCTCGTGCTTCTTATCCTC;

**CTGF:** TGGAGATTTTGGGAGTACGG, CCTGGTCCAGACCACAGAGT;

**TGFB2:** ACAAGAGCAGGAAGGCCAGATG, TGCAGCAGGAGCAGATGTAAG;

**IGFBP5:** AGGTGTGGCAGCTCAAGCTCC, ATTTGTGACCACCAAGGATTCC;

**IGFBP3:** AGGCTGCCCATCTTATCCA, GGGGTGACACATTTCACAC;

**AMOTL2:** ACCACTGCCAGTCTACCAC, AGCAGGCTATGTGGAGAAA;

**APEX1:** CCCCCAGATCGAAAAACCTCA, TTTGGTCTCTTGAAGGCGACA;

**ANKRD1:** CCAAATGCTCCTCCAAGCAGAT, TGAAGGCTGCTCTGAGAAGAT.

**ChIP-seq and ChIP-qPCR.** ChIP was performed as previously described, with few modifications. Briefly, cells were crosslinked with 1% formaldehyde in culture medium without antibiotics for 10 min at room temperature. Chromatin extracts were sonicated using a Bioruptor Pico sonication device (Diagenode, B01060001) using 15 cycles (30
s on and 30 s off) at maximum intensity. Chromatin was immunoprecipitated using 10
µg of antibodies for TEAD4 (Abcam, ab58310), VGLL1 (Proteintech, 10124-2-AP),
YAP1 (Abcam, ab56701), TAZ (BD, clone M2-616, 560235) and H3K27ac (Abcam,
ab4729). Non-immunoprecipitated chromatin was used as input control. After
chromatin de-crosslinking DNA was purified using the QIAquick PCR Purification Kit
(Qiagen, 28106). For all ChIP experiments the enrichment was determined by qPCR
relative to the input sample using positive and negative control primers. For TEAD4
and VGLL1 ChIP-seq two and three highly reproducible biological replicates were
generated, with Spearman’s correlation values of 0.94-0.99 in all biological replicates
(Supplementary Fig. 2a). ChIP enrichment was assessed by ChIP-qPCR before library
construction using primers previously described for YAP/TEAD ChIP in other cell
types47. Libraries for ChIP-seq were prepared using the NEBNext Ultra DNA Library
Prep Kit for Illumina (New England Biolabs, E7370S) according to the manufacturer’s
instructions. Sequencing was performed on an Illumina HiSeq 2500 platform using
50bp single end reads. The following primers were used:

ANKRD1 Promoter: GAGGGGAGGACAAGCTAACCC,
AGCTGTCCCCTGACTCTTGA;
TGFβ2 Enhancer: AGCTTTGATCACACTGATTCCA,
TGCCTCTTCACATCTGTTCATT;
AMOTL2 3’ Enhancer: ACAGCCCTCCAACTATGCTAAG,
CCACAGCACATTTCAGGATA;
CTGF Promoter: GAGCTGAATGGAGTCCTACACA,
GGAGGAATGCTGAGTGTCAAG;
LSAMP (negative): CTGAATAATGAGTGAGGGACAA,
GATTTGGGCTTGGCAGGTGT;
CTGF -8.3kb (negative): TTGCTGGTGGTAGGGAAATACT,
TCACTGCACCTTTGCTTTTCTA;
TUBB Promoter (negative): TCCCTGACCACAAGACTG,
ATTGTTGTCATCTGGCAGG;
EGFR (enh): CACACCTGAGCATGTCCTTG, GCAATGGGATCGAGTTGT;
GAPDH (negative): TCGACAGTCAGCCGCATCT, CTAGCCTCCCCGGGTCTTCT.
RNA-seq. Total RNA was extracted from cells using RNeasy Mini Kit (Qiagen 74106). RNA-seq libraries were prepared using NEBNext Ultra II RNA Library Prep Kit for Illumina (New England Biolabs, E7775) and sequencing was performed on an Illumina HiSeq 2500 platform. Raw reads were aligned to the human reference genome hg19 version using TopHat. Redundant reads were removed using SAMtools. HTSeq was used to count reads at UCSC annotated genes. Normalization and differential expression analysis were carried out using DESeq2 package. Integrated Genomics Viewer (IGV, Broad Institute) was used to visualize the Z-score normalized coverage tracks generated with R packages Rsamtools (http://bioconductor.org/packages/Rsamtools) and rtracklayer. Fold changes in gene expression between two cell lines or two conditions were calculated as the ratio of normalized read counts obtained with DESeq2. The 5th percentile, first quartile, median, third quartile and 95th percentile are plotted in box-and-whiskers graphs.

GO analysis. GO analysis on the VGLL1 activated genes and genes downregulated by VP was performed using g:Profiler (https://biit.cs.ut.ee/gprofiler/gost).

Growth assays. Cell growth following siRNAs or drug treatments was assessed using the sulphorhodamine B (SRB) assay. Briefly, 2000-5000 cells were seeded per well in 96-well plates. Cells were allowed to attach overnight before treatment with drugs or transfection with siRNAs. Medium was changed every 3 days and cells were fixed by adding 100 μl of cold 40% (wt/vol) trichloroacetic acid (TCA) to each well for at least 60 min. The plates were washed five times with distilled water, 100 μl of SRB reagent (0.4% wt/vol SRB in 1% wt/vol acetic acid) was added to each well, and the plates were allowed to incubate for 1h. The plates were then washed five times with 1% (wt/vol) acetic acid and allowed to dry overnight. SRB solubilization was performed by adding 100 μl of 10 mM Tris HCl per well to the plates, followed by shaking for 10 min. Optical density (OD) was then measured using a Sunrise microplate reader (Tecan) at λ=492 nm. Six wells were analysed for each experimental condition. Data are presented as mean ± standard error of the mean (s.e.m.). Data are presented as growth relative to control cells (treated with vehicle or transfected with control siRNA), using the first day of treatment as baseline. For time-course experiments, multiple plates were seeded and drugged in identical fashion, and at the indicated time points.
Each growth experiment was independently confirmed using three biological replicates.

**Live cell imaging.** To follow the growth of cells over the course of three months in the presence of drug containing media, we used the live cell imaging system Incucyte Zoom (EssenBioscience). One million cells were seeded in a T75 flask and after 24h the corresponding media was added to the cells with the addition of fulvestrant at the indicated concentrations. Medium and drugs were replaced twice per week and cells were imaged every 2 days. Analysis was performed using the Incucyte Zoom software with parameters optimized for each cell line and using confluency (µm²) as a measure of cell growth. Raw data was extracted and plotted in GraphPad Prism. Each experiment was independently confirmed using three biological replicates.

**ChIP-seq analysis.**

Raw reads were aligned to the human reference genome hg19 version using Bowtie. Redundant reads were removed using SAMtools. VGLL1 and TEAD4 peaks were called using MACS2 using the input sample as a control, with default parameters and q value less than 0.01. For each ChIP-seq target, peaks common to all the biological replicates were kept for further analysis. Integrated Genomics Viewer (IGV, Broad Institute) was used to visualize the Z-score normalized coverage tracks generated with R packages Rsamtools (http://bioconductor.org/packages/Rsamtools) and rtracklayer. The overlap of peaks from different ChIP–seq experiments was determined using the BEDTools2 suite.

**Ratio of H3K27ac between FULVR vs MCF7 cells:** ChIP-seq data for H3K27ac in FULVR and MCF7 cells was obtained from. A list of gene promoters was downloaded from the Table Browser of UCSC genome browser by selecting 3kb-wide regions centred on each TSS annotated in the human genome version hg19. The BEDTools2 suite was used to calculate the coverage of H3K37ac reads at gene promoters. The H3K27ac coverage at promoters was normalized to the total number of million mapped reads to generate normalized read counts in units of reads per million mapped reads (RPM). For each gene promoter we calculated the ratio of normalized H3K27ac coverage between FULVR and MCF7 cells to generate the plot in Fig.1a.
ChIP-seq heatmaps and average profiles. Normalized coverage tracks were plotted as heatmaps and average signal profiles with SeqPlots\textsuperscript{87} using a window of 2-3kb centred around the peak centre. Heatmap rows were sorted from high to low signal. Average profiles are presented as the mean of the normalized coverage ± s.e.m. and 95% confidence interval.

Peak annotation and Motif analysis. The HOMER suite\textsuperscript{88} was used for de novo motif discovery analysis of the VGLL1 and TEAD4 peaks using the findMotifsGenome.pl script with default parameters. The annotatePeaks.pl script in the HOMER suite was used for peak annotation analysis to associate peaks with gene targets, and to annotate the location of peaks to different genomic features (see Supplementary Fig. 2f).

VGLL1-TEAD4 co-binding enrichment analysis. VGLL1/TEAD4 co-bound regions were defined with the bedtools intersect script from the BEDTools2 suite. To calculate VGLL1/TEAD4 co-binding enrichment, the TEAD4 binding regions were randomized in the hg19 version of the human reference genome using the bedtools shuffle script from the BEDTools2 suite. Chi-square test was applied to assess the co-binding enrichment based on the observed co-binding relative to the expected (random) co-binding, after performing ten permutations of the TEAD4 peaks.

Functional annotations. Functional annotation of the VGLL1 ChIP-seq peaks was performed using the MSigDB Perturbation annotation from GREAT version 3.0.0 \textsuperscript{89}. Supplementary Table 2 shows the complete annotation. To visualize the top nine most significant terms from the annotation (according to binomial P value), we used ggplot2 (https://ggplot2.tidyverse.org) with R scripts adapted from REVIGO\textsuperscript{90}, scripts provided upon request (see Supplementary Fig. 3a).

Generation of the signature of VGLL1 activated genes and not VGLL1 targets. To identify the VGLL1 direct targets we performed RNA-seq in FULVR cells transfected with three different VGLL1 siRNAs (siVGLL1#1-3) and control siRNA (siControl) (three biological replicates for each treatment). To generate the signature
of VGLL1 activated genes we selected all the genes significantly downregulated (adjusted $p < 0.05$, negative log2(fold change)) by at least two different VGLL1 siRNAs compared to siControl. We further filtered these genes by those associated with at least one VGLL1 binding site to define a more stringent set of direct VGLL1 targets which we called VGLL1 activated genes. To define the set of not VGLL1 targets we selected the genes whose expression was not significantly altered (adjusted $p > 0.05$), by any of the VGLL1 siRNAs, and we further filtered these genes by those without any VGLL1 associated binding site and by those with non-zero read counts among FULVR and MCF7 cells.

**Kaplan–Meier survival analysis.** To evaluate the prognostic value of the VGLL1 activated genes signature we analysed the survival of patients from the METABRIC breast cancer dataset\(^{60}\) treated with endocrine therapies. We calculated the median expression of the VGLL1 activated genes signature across all patients and used it as a cut off to select patients with high expression of the VGLL1 activated genes signature as those with median expression of the combined VGLL1 activated genes above the median cut off and patients with low expression of the VGLL1 activated genes signature as those with median expression of the combined VGLL1 activated genes below the median cut off. The Kaplan–Meier plot was generated to compare the survival curves of patients with high and low expression of the VGLL1 activated genes signature. Survival analysis was performed in GraphPad Prism and $p$ values were calculated using the log-rank (Mantel–Cox) test.

**Statistical analysis.** All statistical analyses were performed in R or GraphPad Prism. To perform pairwise comparisons between two groups we used the Mann-Whitney test, two tailed and the Student’s t-test, two tailed. One-tailed Wilcoxon signed rank test was used to compare gene expression values in the same patients at two different time points. Correlations were performed using the Spearman test. Chi-square test was applied to assess the co-binding enrichment from ChIP-seq over expected co-binding.

**Data availability**
The RNA-seq and ChIP-seq data generated in this study have been deposited in the GEO database under accession XXXX.
Acknowledgements

We are grateful to Dr Chris Morrow (Astra Zeneca) for providing AZD9496 and to Dr Jun Liang (Genentech) for GDC-0810. This work was funded by Cancer Research UK (C37/A18784). Imperial College Healthcare NHS Trust Tissue Bank provided tissue samples. Other investigators may have received samples from these same tissues. The research was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre, the Experimental Cancer Medicine Centre and the Cancer Research UK Centre at Imperial College Healthcare NHS Trust, Imperial College, London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Author contributions


References


63. Kim, B.K., *et al.* PI3K/AKT/beta-Catenin Signaling Regulates Vestigial-Like 1 Which Predicts Poor Prognosis and Enhances Malignant Phenotype in Gastric Cancer. *Cancers (Basel)* **11**(2019).


**Figure Legends**

**Fig. 1 VGLL1 expression is induced in fulvestrant-resistant breast cancer.**

A, Ratio of H3K27ac in FULVR/MCF7 cells at gene promoters in a window of ± 1.5 kb centred on the transcriptional start site. H3K27ac is highly enriched at the promoters of the VGLL1 and VGLL3 genes in fulvestrant cells relative to MCF7 cells, whereas ER target genes (including PGR, GREB1 and TFF1) show H3K27ac loss in FULVR cells.  

B, Genome browser view of H3K27ac ChIP-seq signal at the promoter region of the VGLL1 and VGLL3 genes in MCF7-FULVR and MCF7 cells.  

C, RT-qPCR showing VGLL1 and VGLL3 upregulation in FULVR cells and loss of ESR1 and PGR expression. Data are presented as mean ± s.e.m of n=3 experiments. *P<0.05 (Student’s t-test, two-tailed).  

D, Immunoblotting of VGLL1, VGLL3 and proteins in the YAP/TEAD pathway in MCF7 and FULVR cells.  

E, RNA-FISH showing increased VGLL1 mRNA levels in breast cancer patients after fulvestrant treatment and relapse compared to matched pre-treatment samples. Representative images show the results for matched samples from one patient. Nuclei are labelled with DAPI, and individual VGLL1 mRNA molecules detected with Cy5-labelled probes. P=0.0177 (one-tailed Wilcoxon signed rank test). Scale bar, 10µm.  

F, RT-qPCR for ESR1 and VGLL1 after treatment of ER+ breast cancer cell lines with fulvestrant (100nM, 24 h).  

G, RT-qPCR for ESR1 and VGLL1 after ESR1 siRNA. Data in f and g are presented as in c. *P<0.05 (Student’s t-test, two-tailed).

**Fig. 2 VGLL1 is recruited to TEAD4 binding regions in fulvestrant-resistant MCF7 cells.**

A, Overlap between TEAD4 ChIP-seq peaks in MCF7 and FULVR cells showing increased TEAD4 binding in FULVR cells. Top-most enriched motifs in each cell line are shown. More detailed motif analysis is shown in Supplementary Fig. 3b.  

B, Correlation between VGLL1 and TEAD4 occupancy at TEAD4 peaks in FULVR cells. P-value was calculated using the Spearman test; r_s, Spearman’s correlation coefficient.  

C, Heatmap showing TEAD4 and VGLL1 binding at shared and unique TEAD4 peaks in each cell line, in a window of ± 2kb centred on the peak centre.  

D, VGLL1 binding is enriched at TEAD4 peaks in FULVR cells. Binding enrichment was calculated as VGLL1-TEAD4 co-binding over the mean expected value after...
generating random permutations of the TEAD4 peaks (Chi-squared test p-value <0.0001). e, Average normalized ChIP-seq signal of TEAD4, VGLL1 and H3K27ac at TEAD4 peaks in FULVR cells, in a window of ± 3 kb from the peak centre. f, Average ChIP-seq signal of H3K27ac on VGLL1 peaks divided in quantiles based on the peak coverage in FULVR cells. g, Genome browser view of VGLL1, TEAD4 and H3K27ac ChIP-seq signal, together with the RNA-seq signal in MCF7 and FULVR cells at YAP/TEAD target genes described in other cell lines.

Fig. 3 VGLL1 mediates resistance to fulvestrant. a, VGLL1 is required for the growth of MCF7-FULVR cells. Cells were transfected with two independent siRNAs for VGLL1. Growth was determined with the SRB assay five days after transfection. Data are mean ± s.e.m of n=6 independent wells. One representative experiment is shown; similar results were obtained in two additional independent experiments. * P<0.05 (Mann-Whitney test, two tailed). Also shown are expression levels of VGLL1 following siVGLL1 transfection (n=3, *p<0.05 (Student’s t-test, two-tailed)). b, The synergistic activation mediators (SAM) technique using a catalytically dead Cas9 (dCas9) for transcriptional activation of endogenous genes⁴⁹, with a sgRNA for the region -156 to -134 of the VGLL1 gene, was used to activate VGLL1 expression in MCF7 cells. c, RT-qPCR of VGLL1 in MCF7 cells overexpressing VGLL1 (MCF7 ActCas9-VGLL1) generated as described in b and control cells (MCF7 ActCas9-Vector). d, Transcriptional activation of endogenous VGLL1 in MCF7 cells promotes fulvestrant resistance. Growth of MCF7 ActCas9-VGLL1 cells and MCF7 ActCas9-Vector cells was monitored using live cell imaging in the presence of 100 nM fulvestrant. Shown are the results of one of three independent experiments. Data are mean ± s.e.m of n=9 representative images from one experiment. e, Immunoblotting for VGLL1, ER and PR in the indicated cell lines. VGLL1-FULVR-1 and VGLL1-FULVR-2 correspond to two independent fulvestrant resistant cell lines derived from the parental MCF7 ActCas9-VGLL1 cell line after continuous culturing in the presence of 100 nM or 1 µM fulvestrant, respectively. f, Normalized RNA-seq read counts of VGLL1, ESR1 and PGR in the indicated cell lines showing strong increase in VGLL1 expression in the resistant cells, with a concurrent downregulation of ER and PGR. g, MCF7 ActCas9-VGLL1-FULVR cells were transfected with siVGLL1 and growth assessed as in a. Data are mean ± s.e.m of n=6. One representative experiment is
shown; similar results were obtained in two additional independent experiments. * P<0.05 (Mann-Whitney test, two tailed). RT-qPCR for VGLL1 is also shown (n=3, *p<0.05 (Student’s t-test, two-tailed)). h, Growth of the indicated cell lines treated with increasing concentrations of fulvestrant to a maximum of 1 µM, for five days. Cell growth was estimated using the SRB assay and is shown as percentage relative to vehicle (n=6). Half maximum inhibitory concentration (IC₅₀) is also shown for each line.

**Fig. 4 Genes upregulated in fulvestrant-resistant breast cancer cells depend on VGLL1 transcriptional activity.** a, Genes predicted as VGLL1 targets in FULVR cells are more highly expressed in FULVR cells than MCF7 cells. Genes were segregated into those with no VGLL1 peaks, and those with 2 or more and 5 or more VGLL1 peaks. The y-axis shows the log2 fold change in gene expression determined from RNA-seq in FULVR cells versus the parental MCF7 cells. Data are presented as box-and-whiskers plots (whiskers extend from the 5th to the 95th percentile; the box extends from the 25th to the 75th percentile; the line within the box represents the median). ****P<0.0001 (Mann-Whitney test, two tailed). b, Genes activated by VGLL1 (n = 762) are over-expressed in FULVR cells relative to MCF7 cells, compared to not-VGLL1 targets (n = 8,932). VGLL1-activated genes and not-VGLL1 targets were determined by RNA-seq in FULVR cells transfected with VGLL1 siRNAs. The VGLL1-activated genes were defined as genes downregulated by VGLL1 siRNAs (see Methods, and Supplementary Fig. 4b). Data are presented as in a. P<0.0001 (Mann-Whitney test, two tailed). c, Normalized average H3K27ac signal on the promoters of VGLL1 activated genes in FULVR cells and MCF7 cells. d, VGLL1-activated genes (n = 762) display higher expression levels in FULVR cells than not-VGLL1 targets (n = 8,932). The y-axis shows normalized gene expression values from RNA-seq. P<0.0001 (Mann-Whitney test, two tailed). e, GO molecular function sets enriched in VGLL1-activated genes. f, RNA-seq data represented as a volcano plot for the comparison between FULVR vs MCF7 cells showing the VGLL1-activated genes as red dots. Four VGLL1 targets highly upregulated in FULVR cells and associated with growth functional categories enriched in VGLL1-activated genes are highlighted. g, RT-qPCR in the indicated MCF7-ActCas9 cells represented as a heatmap. Gene expression values were normalized to GAPDH expression and are shown as log2 fold change relative to MCF7 ActCas9-Vector cells.
**Fig. 5** VGLL1 promotes EGFR expression in FULVR cells.  

**a,** Genome browser view of VGLL1 and TEAD4 ChIP-seq signal at the EGFR gene and EGFR enhancer (highlighted).  

**b,** ChIP-qPCR for TEAD4, VGLL1 and H3K27ac in MCF7 ActCas9-Vector, MCF7 ActCas9-VGLL1 and MCF7 ActCas9-VGLL1-FULVR cells showing VGLL1 and TEAD4 binding at the EGFR enhancer together with EGFR enhancer activation exclusively in FULVR cells. The CTGF -8.3kb region was used as a negative control for VGLL1/TEAD4 binding.  

**c,** RT-qPCR for EGFR in the indicated MCF7 ActCas9 cells (n=3, *p<0.05 (Student’s t-test, two-tailed)).  

**d,** Immunoblotting for EGFR and downstream EGFR signalling proteins showing EGFR up-regulation and activation of the EGFR pathway in MCF7 ActCas9-VGLL1-FULVR cells  

**e,** RT-qPCR for dCAS9, VGLL1 and EGFR in MCF7 ActCas9-VGLL1-FULVR cells transfected with dCAS9 siRNAs showing reduction in VGLL1 and EGFR expression.  

**f,** Growth of the indicated MCF7 ActCas9 cell lines treated with increasing concentrations of the EGFR inhibitor erlotinib (0.05 µM to 12.5 µM) for 5 days. Growth is shown as percentage relative to vehicle treatment for one representative experiment (n=6 replicates) of two independent experiments.  

**g,** Western blot of EGFR and the downstream EGFR signalling proteins in the indicated MCF7 ActCas9 cell lines treated with erlotinib at the indicated concentrations (24 h).  

**h,** Correlation between EGFR and VGLL1 gene expression in ER + breast cancers from TCGA (n = 609 samples).  

**i,** The mechanism by which VGLL1/TEAD4 activate the expression of EGFR in fulvestrant-resistant breast cancer cells.

**Fig. 6** Verteporfin inhibits VGLL1 transcriptional activity.  

**a,** VP impairs the growth of FULVR cells. Growth in the indicated FULVR cells treated with increasing concentrations of VP is shown as percentage of growth relative to vehicle.  

**b,** ChIP-qPCR for VGLL1 and TEAD4 in MCF7-FULVR cells showing reduced VGLL1 binding at the target genes in the presence of VP (2µM, 24 h). The y axis shows DNA enrichment calculated as the percentage of input.  

**c,** RNA-seq data represented as a volcano plot showing upregulated genes (red, adjusted P < 0.01, log2 (fold change) > 1) and downregulated genes (blue, adjusted P < 0.01, log2 (fold change) < -1) in MCF7-FULVR cells treated with VP, (2µM, 24 h) compared to vehicle (DMSO), n=4 biological replicates. VGLL1 target genes such as IGFBP5, EGFR and TGFB2 are strongly downregulated by VP.  

**d,** RT-qPCR in MCF7-FULVR cells treated with VP.
(2µM, 24 h) showing that VP selectively downregulates the expression VGLL1 targets, while the expression of not VGLL1 targets (APEX1 and URI1) is not affected by VP.

e, Genes downregulated by VP are highly enriched in VGLL1 peaks. The y axis shows the number of VGLL1 peaks over the expected value after generating random permutations of the VGLL1 peaks showing that the top most highly downregulated genes after VP treatment in FULVR cells ($n = 631$) have significantly higher number of VGLL1 peaks compared to the bottom genes not differentially expressed by VP ($n = 631$). Data are presented as box-and-whiskers plots (whiskers extend from the 5th to the 95th percentile; the box extends from the 25th to the 75th percentile; the line within the box represents the median). **** $P<0.0001$ (Mann-Whitney test, two tailed).

f, Genome browser view of VGLL1 and TEAD4 normalized ChIP-seq signal and normalized RNA-seq signal in FULVR cells showing representative examples of VGLL1 activated genes (direct VGLL1 targets) and not-VGLL1 targets. VGLL1 activated genes are downregulated by VGLL1 siRNA and are also selectively downregulated by VP (2µM, 24 h) compared to not-VGLL1 targets. g, VGLL1 activated genes ($n = 762$) are preferentially downregulated by VP compared to not-VGLL1 targets ($n = 8,932$). The y axis shows fold change in gene expression from RNA-seq between VP versus DMSO treatment in MCF7-FULVR cells. Data are presented as in e. ****$P<0.0001$ (Mann-Whitney test, two tailed).

h, Breast cancer patients with higher VGLL1 expression display increased levels of expression of the VGLL1 activated genes compared to patients with lower VGLL1 expression. Patients from the METABRIC breast cancer dataset were stratified according to high (top quantile, n=495) or low (bottom quantile, n=495) VGLL1 expression levels. Data are presented as in e. ****$P<0.0001$ (Mann-Whitney test, two tailed).

i, Kaplan–Meier plot representing the percentage of metastasis-free survival in patients with breast cancer treated with endocrine therapies showing that patients with higher expression of the VGLL1 activated genes signature display lower survival rates (log-rank Mantel-Cox test).

**Supplementary Figure Legends**

**Supplementary Figure 1.**

a, RNA-seq data represented as a volcano plot showing upregulated genes (red, adjusted $P < 0.01$, log2 (fold change) > 2.5) and downregulated genes (blue, adjusted...
P < 0.01, log2 (fold change) < -2.5) in FULVR cells compared to MCF7 cells. VGLL1 and VGLL3 expression is strongly increased in FULVR cells. Classical ER target genes (TFF1, PGR and GREB1) are shown as controls for ER signaling downregulation in FULVR cells. b, Normalized gene expression of TEAD coactivators determined by RNA-seq. VGLL1 and VGLL3 are the only TEAD coactivators that show the strongest upregulation in FULVR cells, as well as being up-regulated in tamoxifen-resistant MCF7 (TAMR) cells. Data are presented as box-and-whiskers plots showing the median of two independent biological RNA-seq replicates (whiskers extend to minimum and maximum values). LTED, long term estradiol deprived. RNA-seq data were obtained from ref16. c, RT-qPCR for genes in the YAP/TEAD pathway in MCF7 and FULVR cells. Data are presented as mean + s.e.m of three independent experiments. *P<0.05 (Student’s t-test, two-tailed). d, Heatmap showing the expression of genes in the YAP/TEAD pathway in FULVR cell lines derived from the ER positive breast cancer cell lines MCF7, T47D or ZR-75-1. Upregulation of VGLL1 expression is common to all the FULVR cell lines. The expression values are represented as log2 (fold change) relative to the corresponding parental cell lines and normalized to GAPDH. e, Immunoblotting showing VGLL1 expression in different FULVR cell lines, accompanied by reductions in ER and PR levels. ER, estrogen receptor; PR-A, progesterone receptor isoform A; PR-B, progesterone receptor isoform B. f, Immunohistochemistry of VGLL1 performed in sections from matched breast cancer samples prior to fulvestrant treatment and after fulvestrant treatment and relapse. VGLL1 staining is significantly higher after relapse on fulvestrant treatment, p=0.0098 (one-tailed Wilcoxon signed rank test). Representative images of matched breast cancer tissue pre and post fulvestrant, show increased VGLL1 staining after fulvestrant treatment and relapse. Scale bar, 100µm.

Supplementary Figure 2.

a, Immunoblotting of VGLL1, ER and PR after treatment with vehicle or fulvestrant at the indicated concentrations (48 h) in MCF7 cells. Increased levels of VGLL1 are observed after ER inhibition by fulvestrant. b, RT-qPCR for genes in the YAP/TEAD pathway in MCF7 cells after treatment with vehicle or fulvestrant (100nM, 24 h). Data are presented as mean + s.e.m of three independent experiments. *P<0.05 (Student’s t-test, two-tailed). c, Heatmap showing the expression of genes in the YAP/TEAD pathway in the ER positive breast cancer cell lines T47D, ZR-75-1 and BT474 after
treatment with vehicle or fulvestrant (100nM, 24 h). The expression values are represented as log2 (fold change) relative to the vehicle control and normalized to GAPDH. 

d, Heatmap showing the expression of genes in the YAP/TEAD pathway after ESR1 siRNA in the ER positive breast cancer cell lines T47D, ZR-75-1 and BT474. A strong upregulation of VGLL1 after ESR1 downregulation is common to all cell lines. The expression values are represented as log2 (fold change) relative to the control siRNA (siControl) and normalized to GAPDH. 
e, Western blot of VGLL1 and ER showing VGLL1 upregulation following ER downregulation by ESR1 siRNA in MCF7 cells. 
f, RT-qPCR for genes in the YAP/TEAD pathway transfected with ESR1 siRNA. Data are presented as mean ± s.e.m of two independent experiments. *P<0.05 (Student’s t-test, two-tailed)). 
g, RT-qPCR for VGLL1 and the ER target genes PGR and GREB1 in MCF7 cells showing downregulation of ER transcriptional activity and VGLL1 upregulation after treatment with different SERDs (GDC-0810, AZD9496 and RAD1901, 100nM, 48 h). Data are presented as mean ± s.e.m of two independent experiments.

Supplementary Figure 3. 

a, Correlation between two biological replicates of TEAD4 ChIP-seq generated in MCF7 cells and FULVR cells. P-value was calculated using Spearman test, rs, Spearman’s correlation coefficient. 
b, De novo motif analysis of TEAD4 peaks in MCF7 cells and FULVR cells. 
c, Correlation between three biological replicates of VGLL1 ChIP-seq in FULVR cells. P-value was calculated using Spearman test, rs, Spearman’s correlation coefficient. 
d, De novo motif analysis of VGLL1 peaks in FULVR cells. 
e, Fraction of VGLL1 and TEAD4 peaks present at the indicated genomic features showing that the majority of VGLL1 and TEAD4 peaks are present at intronic and intergenic regions. 
f, Normalized H3K27ac signal at TEAD4/VGLL1 co-bound regions from FULVR cells showing lower H3K27ac signal at these sites in MCF7 cells compared to FULVR cells. 
g, ChIP-qPCR showing preferential VGLL1 binding compared to YAP/TAZ binding on promoters and enhancers of known YAP/TEAD target genes from other cell lines47, in FULVR cells but not in MCF7 cells. Negative control regions are indicated which were previously reported as non-YAP/TAZ targets in other cell types47. Binding enrichment was calculated as percentage of input. 

*P<0.05 (Student’s t-test, two-tailed).
Supplementary Figure 4.

a, Functional annotation of the VGLL1 ChIP-seq peaks in FULVR cells was performed using GREAT\textsuperscript{89}. REVIGO\textsuperscript{90} was used to visualize the top nine most significant categories (see Supplementary Table 2 for the complete list). b, Left: Growth of T47D-FULVR cells and MCF7-FULVR(2) determined with the SRB assay five days after VGLL1 siRNA transfection. Growth was calculated relative to the siRNA control (siControl). Data are mean ± s.e.m of n=6. *P < 0.05 (Mann-Whitney test, two tailed).

Right: VGLL1 gene expression after VGLL1 siRNA transfection in the indicated cell lines (n=3, *p<0.05 (Student's t-test, two-tailed)). c, Western blot of VGLL1 in MCF7 cells with stable expression of dCAS9 fused to transcriptional activators\textsuperscript{49} targeted to the VGLL1 promoter (MCF7 ActCas9-VGLL1) and control cells (MCF7 ActCas9-Vector). d, Growth of MCF7 ActCas9-VGLL1 cells and control MCF7 ActCas9-Vector cells monitored using live cell imaging over a period of 100 days in the presence of fulvestrant (100nM, similar results were obtained with 1µM). e, RT-qPCR of VGLL1, ESR1 and PGR in the indicated MCF7 ActCas9 cell lines. f, Comparison of genes differentially expressed in MCF7 ActCas9-VGLL1-FULVR-1 cells across independently generated fulvestrant resistant breast cancer cell line models show similar transcriptional perturbations across the different models. The y axis shows the fold change in gene expression from RNA-seq data between each FULVR cell line and the corresponding parental cells for genes upregulated (adjusted p<0.05, fold change > 2) and downregulated (adjusted p<0.05, fold change < 0.5) in MCF7 ActCas9-VGLL1-FULVR-1 cells compared to the parental MCF7 ActCas9-VGLL1 cells. Data are presented as box-and-whiskers plots (whiskers extend from the 5th to the 95th percentile; the box extends from the 25th to the 75th percentile; the line within the box represents the median). ****P<0.0001 (Mann-Whitney test, two tailed). g, RNA-seq data represented as a volcano plot for the comparison between MCF7 ActCas9-VGLL1 vs MCF7 ActCas9-Vector cells, showing that MCF7 cells overexpressing VGLL1 which have not been exposed to fulvestrant show almost identical transcriptional profiles compared to the control cells. h, Growth of the indicated MCF7 ActCas9 cells following five days of treatment with increasing doses (0.025 nM to 1µM) of three different SERDs (AZD9496, RAD1901, GDC-0810). Data are plotted as the mean ± s.e.m of n=6.
Supplementary Figure 5.

a, MCF7 ActCas9-VGLL1-FULVR cells display higher levels of expression of predicted VGLL1 targets from the independent FULVR cells. Genes with 2 or more and 5 or more VGLL1 associated peaks were considered as VGLL1 targets for this analysis together with genes without VGLL1 associated peaks. The y axis shows the fold change in gene expression determined by RNA-seq between MCF7 ActCas9-VGLL1-FULVR-1 and the parental MCF7 ActCas9-VGLL1 cells. Data are presented as box-and-whiskers plots (whiskers extend from the 5th to the 95th percentile; the box extends from the 25th to the 75th percentile; the line within the box represents the median). ****P<0.0001 (Mann-Whitney test, two tailed); **P = 0.0056 (Mann-Whitney test, two tailed).

b, VGLL1 activated genes are defined as genes downregulated by VGLL1 siRNAs. Fold change in gene expression between FULVR cells transfected with VGLL1 siRNAs versus control siRNA for VGLL1 activated genes (n = 762) and not-VGLL1 targets (n = 8,932). Gene expression was determined by RNA-seq in FULVR cells transfected with three different VGLL1 siRNAs, n=3 biological replicates (see Methods). Data are presented as box-and-whiskers plots, as in a. ****P<0.0001 (Mann-Whitney test, two tailed).

c, VGLL1 gene expression determined by RNA-seq in FULVR cells transfected with three different VGLL1 siRNAs from the same RNA-seq experiment as in b. The y axis shows normalized read counts from RNA-seq data, n=3 biological replicates.

d, VGLL1 binding is enriched at VGLL1 activated genes in FULVR cells. The y axis shows the number of VGLL1 peaks associated with VGLL1 activated genes and not-VGLL1 targets over the mean expected value after generating random permutations of the VGLL1 peaks. P<0.0001 (Mann-Whitney test, two tailed).

e, TEAD4 binding is enriched at VGLL1 activated genes in FULVR cells. The y axis shows the number of TEAD4 peaks associated with VGLL1 activated genes and not-VGLL1 targets over the mean expected value after generating random permutations of the TEAD4 peaks. P<0.0001 (Mann-Whitney test, two tailed).

f, Fold change in H3K27ac signal between FULVR versus MCF7 cells at the promoters of VGLL1 activated genes (n = 762) and not-VGLL1 targets (n = 8,932). Data are presented as box-and-whiskers plots, as in a. P<0.0001 (Mann-Whitney test, two tailed).

g, Increased H3K27ac and TEAD4 recruitment in FULVR cells compared to MCF7 cells at regulatory elements associated with VGLL1 activated genes. Data are presented as average normalized ChIP-seq signal of VGLL1, TEAD4 and H3K27ac in FULVR cells together with TEAD4 and H3K27ac signal in MCF7 cells, in a window of...
± 2.5 kb centred on the VGLL1/TEAD4 peaks associated with VGLL1 activated genes in FULVR cells. **h**, Heatmap showing a strong upregulation of VGLL1 gene expression and upregulation of VGLL1 activated genes, with a concurrent downregulation of ER transcriptional activity as evidenced by a significant reduction in PGR expression, in MCF7 ActCas9-VGLL1-FULVR cells. The expression values are shown as normalized read counts from RNA-seq and are represented as log2 fold change relative to MCF7 ActCas9-Vector cells.

**Supplementary Figure 6.**

**a**, Ratio of TEAD4 ChIP-seq signal in FULVR versus MCF7 cells at TEAD4 peaks from FULVR and MCF7 cells. The position of the VGLL1 promoter and EGFR enhancer are shown in this plot and are among the top genes with high TEAD4 signal in FULVR cells and absence of TEAD4 binding in MCF7 cells. VGLL1 is recruited to the TEAD4 binding sites at the VGLL1 promoter and EGFR enhancer in FULVR cells (see also Fig. 5a). **b**, RT-qPCR for EGFR showing EGFR upregulation in different fulvestrant resistant cell lines which also upregulate VGLL1 (see Fig. 1c and Supplementary Fig. 1d, e for VGLL1 expression in these cell lines). **c**, RT-qPCR for EGFR and VGLL1 in MCF7 ActCas9-VGLL1-FULVR cells transfected with VGLL1 siRNA, showing reduced EGFR expression following VGLL1 downregulation. **d**, Western blot of VGLL1, EGFR and downstream EGFR targets in MCF7-FULVR cells transfected with VGLL1 siRNAs showing reduced EGFR and downstream EGFR signalling following VGLL1 downregulation. **e**, Western blot of VGLL1, EGFR and downstream EGFR targets in T47D-FULVR cells transfected with VGLL1 siRNAs showing reduced EGFR and downstream EGFR signalling following VGLL1 downregulation. **f**, Growth of the indicated FULVR cells and the corresponding isogenic fulvestrant-sensitive cells treated with increasing concentrations of the EGFR inhibitor erlotinib for 5 days. Growth is shown as percentage relative to vehicle (n=6 replicates). **g**, Western blot of EGFR and downstream EGFR targets in MCF7-FULVR cells and the isogenic fulvestrant-sensitive MCF7 cells, treated with erlotinib at the indicated concentrations (24 h). **h**, EGFR ranks first as the most significantly co-expressed gene with VGLL1 in breast cancer patients from METABRIC and is also the highest ranked protein significantly correlated with VGLL1 in breast cancer from TCGA. **i**, EGFR expression in ER+ breast cancer patients with high (top 10%) and low (bottom 10%) VGLL1 expression.
expression from TCGA and METABRIC datasets. ****P<2.2x10^{-16} (Mann-Whitney test, two tailed).

Supplementary Figure 7.

a, Growth of MCF7 ActCas9-VGLL1-FULVR cells is reduced following treatment with increasing concentrations of VP. Data is shown as percentage growth relative to vehicle (n=6 replicates). b, Fold growth relative to day 0 of MCF7-FULVR(2) cells treated with the indicated concentrations of VP over 9 days. c, Heatmap showing the expression of representative VGLL1 activated genes (direct VGLL1 targets) and not VGLL1 targets in FULVR cells after downregulation of VGLL1 with two different VGLL1 siRNAs. The expression values were determined by RNA-seq and are represented as fold change relative to siControl. d, RT-qPCR for VGLL1, representative VGLL1 activated genes, and not-VGLL1 targets (APEX1 and URI1) in FULVR cells after downregulation of VGLL1 with two different VGLL1 siRNAs. Expression values are represented as fold change relative to siControl and normalized to GAPDH. e, Genes associated with VGLL1 peaks are preferentially downregulated by VP compared to genes without VGLL1 associated peaks. Genes with 2 or more and 5 or more VGLL1 associated peaks were considered as VGLL1 targets for this analysis. The y axis shows the fold change in gene expression determined by RNA-seq between VP versus DMSO. Data are presented as box-and-whiskers plots (whiskers extend from the 5th to the 95th percentile; the box extends from the 25th to the 75th percentile; the line within the box represents the median). ****P<0.0001 (Mann-Whitney test, two tailed); ** P = 0.0056 (Mann-Whitney test, two tailed). f, VGLL1 activated genes display high sensitivity to VP downregulation. Pie charts showing that the proportion of VGLL1 activated genes downregulated by VP is larger than the proportion of not VGLL1 targets downregulated by the same treatment.

g, Box-and-whiskers plot showing that genes downregulated by VP in FULVR cells (adjusted P < 0.01, log2(fold change) <-1) display higher expression levels than genes not inhibited by VP (adjusted P > 0.05, log2(fold change) > -1). The y axis shows normalized gene expression values determined by RNA-seq. Data are presented as box-and-whiskers plots, as in e. ****P<0.0001 (Mann-Whitney test, two tailed). h, Genes downregulated by VP are associated with growth functional categories enriched in the VGLL1 activated genes. The y axis shows the fold change in gene expression determined by RNA-seq between VP versus DMSO for genes included in
the following GO terms: Growth factor binding (GO:0019838), extracellular matrix binding (GO:0050840) and transmembrane receptor protein kinase activity (GO:0019199), showing that genes included in these categories are significantly downregulated by VP. \( P_{adj} = 2.39 \times 10^{-25}, \) \# \( P_{adj} = 2.76 \times 10^{-9}, \) \( ^{\wedge} P_{adj} = 1.34 \times 10^{-22} \) (Adjusted \( p \)-values from the pathway enrichment analysis of VP downregulated genes, see also Supplementary Table 3). i, Kaplan–Meier plot representing the percentage of metastasis-free survival in ER+ breast cancer patients from METABRIC showing that patients with higher expression of the VGLL1 activated genes signature display lower survival rates in response to endocrine therapies treatment (log-rank Mantel-Cox test).
Figure 1
**Figure 2**

**a**

TEAD4 ChIP-seq

Motifs Enriched in MCF7

Motifs Enriched in FULVR Cells Only

**b**

FULVR ChIP-seq

Fold Enrichment over randomized TEAD4 peaks

p < 2.2 x 10^{-16}

r_s = 0.74

**c**

Shared TEAD4 (5,205)

FULVR Only (15,689)

MCF7 Only (1,763)

**d**

VGLL1-TEAD4 co-binding (FULVR)

Fold Enrichment over randomized TEAD4 peaks

Observed/Expected

Expected/Expected

p = 0.0001

**e**

Normalized Read Density at TEAD peaks in FULVR-Cells

Fold Enrichment over randomized TEAD4 peaks

**f**

VGLL1 peaks

Q1 (top 25%)

Q2

Q3

Q4 (bottom 25%)

**g**

ChIP-seq

RNA-seq

Input

CTGF

AMOTL2

ANKRD1

Fold Enrichment over randomized TEAD4 peaks
Figure 3
Figure 4

**VGLL1-activated genes - Pathway Enrichment**

- growth factor binding
- virus receptor activity
- extracellular matrix binding
- laminin binding
- receptor activity
- molecular transducer activity
- sulfur compound binding
- transporter activity
- collagen binding
- anion transmembrane transporter activity
- cytokine binding
- transmembrane receptor protein kinase activity

**RT-qPCR**

- Log2 FC
- VGLL1
- YAP
- TAZ
- AMOTL2
- ANKRD1
- CTGF
- EGFR
- ITGB6
- IGFBP3
- ESR1
- PGR

**VGLL1 Targets**

- Vector
- VGLL1
- VGLL1-FULVR-1
- VGLL1-FULVR-2

**Gene Expression (log FC)**

- FULVR
- MCF7

**Normalized gene expression (log)**

- FULVR
- MCF7

**-Log10 p-value**

- VGLL1-activated genes

**-Log10 padj**

- VGLL1-activated genes

**Log2 Fold Change**

- TGFB2
- ITGB6
- EGFR
- IGFBP3
Figure 5

**a**

![Diagram showing gene expression and regulation](image)

**b**

![Bar graph showing gene expression](image)

**c**

![Graph showing growth and Erlotinib concentration](image)

**d**

![Bar graph showing gene expression](image)

**e**

![Bar graph showing gene expression](image)

**f**

![Graph showing growth and Erlotinib concentration](image)

**g**

![Western blot images](image)

**h**

![Scatter plot showing gene expression](image)

**i**

![Diagram showing epigenetic reprogramming](image)
**Figure 6**

**a**
- MCF7-FULVR
- T47D-FULVR

**b**
- ChIP-qPCR
- VGLL1
- TEAD4
- Vehicle
- VP

**c**
- VP v DMSO (RNA-seq)
- Downregulated (1798 genes)
- Up-regulated (83 genes)

**d**
- RT-qPCR
- Gene Expression
- Vehicle
- VP

**e**
- Enrichment of VGLL1 peaks at genes downregulated by VP

**f**
- VGLL1 activated genes
- Not VGLL1 targets

**g**
- Fold change in gene expression (log)
  - VP:DMSO
  - Not VGLL1 targets
  - VGLL1 activated genes

**h**
- VGLL1 activated genes
  - Breast Cancer (METABRIC)
  - Normalized gene expression
  - Low VGLL1
  - High VGLL1

**i**
- VGLL1 activated genes
  - Breast Cancer (METABRIC)
  - Survival (%)
  - Low (n = 591)
  - High (n = 624)
  - p = 0.0008
**Supplementary Figure 1**

**Panel a:**
- DOWN in FULVR (427 genes)
- UP in FULVR (627 genes)

**Panel b:**
- Normalized Read Counts
  - VGLL1
  - VGLL3
  - VGLL4
  - YAP1
  - TAZ

**Panel c:**
- Gene Expression
  - MCF7
  - FULVR
  - Normalized Read Counts
  - VGLL1
  - VGLL3
  - VGLL4
  - YAP1
  - TAZ

**Panel d:**
- mRNA Expression (Log2 Fold Change)
  - MCF7 (2)
  - MCF7-FULVR (2)
  - T47D
  - T47D-FULVR
  - ZR751
  - ZR751-FULVR

**Panel e:**
- Breast cancer VGLL1 IHC
  - Pre-Fulvestrant treatment
  - Relapsed Post-fulvestrant

**Panel f:**
- Nuclear Hscore
  - p=0.0098; n=15 pairs
  - Pre vs Post Fulv
Supplementary Figure 2
### TEAD4 Binding Regions (MCF7)

<table>
<thead>
<tr>
<th>Motif</th>
<th>Rank</th>
<th>TF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEAD4</td>
<td>1</td>
<td>TEAD4</td>
<td>1e-1479</td>
</tr>
<tr>
<td>TEAD2</td>
<td>2</td>
<td>TEAD2</td>
<td>1e-1439</td>
</tr>
<tr>
<td>TEAD</td>
<td>3</td>
<td>TEAD</td>
<td>1e-1168</td>
</tr>
<tr>
<td>TEAD4</td>
<td>4</td>
<td>TEAD4</td>
<td>1e-107</td>
</tr>
<tr>
<td>FOSL2</td>
<td>5</td>
<td>FOXL2</td>
<td>1e-105</td>
</tr>
<tr>
<td>FOSL2</td>
<td>6</td>
<td>AP2</td>
<td>1e-104</td>
</tr>
<tr>
<td>BATF</td>
<td>7</td>
<td>BATF</td>
<td>1e-102</td>
</tr>
<tr>
<td>AP2</td>
<td>8</td>
<td>AP2</td>
<td>1e-101</td>
</tr>
<tr>
<td>AP2</td>
<td>9</td>
<td>AP1</td>
<td>1e-100</td>
</tr>
<tr>
<td>FOXE BOX</td>
<td>10</td>
<td>FOXE BOX</td>
<td>1e-65</td>
</tr>
</tbody>
</table>

### TEAD4 Binding Regions (FULVR)

<table>
<thead>
<tr>
<th>Motif</th>
<th>Rank</th>
<th>TF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEAD4</td>
<td>1</td>
<td>TEAD4</td>
<td>1e-4752</td>
</tr>
<tr>
<td>TEAD2</td>
<td>2</td>
<td>TEAD2</td>
<td>1e-4374</td>
</tr>
<tr>
<td>TEAD</td>
<td>3</td>
<td>TEAD</td>
<td>1e-3999</td>
</tr>
<tr>
<td>TEAD2</td>
<td>4</td>
<td>TEAD2</td>
<td>1e-314</td>
</tr>
<tr>
<td>JUN-AP1</td>
<td>5</td>
<td>JUN-AP1</td>
<td>1e-312</td>
</tr>
<tr>
<td>AP2</td>
<td>6</td>
<td>AP2</td>
<td>1e-298</td>
</tr>
<tr>
<td>AP1</td>
<td>7</td>
<td>AP1</td>
<td>1e-277</td>
</tr>
<tr>
<td>AP1</td>
<td>8</td>
<td>AP1</td>
<td>1e-261</td>
</tr>
<tr>
<td>AP1</td>
<td>9</td>
<td>AP1</td>
<td>1e-247</td>
</tr>
<tr>
<td>GRHL2</td>
<td>10</td>
<td>GRHL2</td>
<td>1e-241</td>
</tr>
</tbody>
</table>

### VGLL1 Binding Regions (FULVR)

<table>
<thead>
<tr>
<th>Motif</th>
<th>Rank</th>
<th>TF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEAD4</td>
<td>1</td>
<td>TEAD4</td>
<td>1e-1479</td>
</tr>
<tr>
<td>GRHL2</td>
<td>2</td>
<td>GRHL2</td>
<td>1e-1546</td>
</tr>
<tr>
<td>TEAD</td>
<td>3</td>
<td>TEAD</td>
<td>1e-1432</td>
</tr>
<tr>
<td>FOXL2</td>
<td>4</td>
<td>FOXL2</td>
<td>1e-1346</td>
</tr>
<tr>
<td>FOXA1</td>
<td>5</td>
<td>FOXA1</td>
<td>1e-1323</td>
</tr>
<tr>
<td>FOXL1</td>
<td>6</td>
<td>FOXL1</td>
<td>1e-1319</td>
</tr>
<tr>
<td>TEAD2</td>
<td>7</td>
<td>TEAD2</td>
<td>1e-1267</td>
</tr>
<tr>
<td>FOSL2</td>
<td>8</td>
<td>FOSL2</td>
<td>1e-1172</td>
</tr>
<tr>
<td>JUN-AP1</td>
<td>9</td>
<td>JUN-AP1</td>
<td>1e-1130</td>
</tr>
<tr>
<td>GRHL2</td>
<td>10</td>
<td>GRHL2</td>
<td>1e-1125</td>
</tr>
</tbody>
</table>

### Supplementary Figure 3

- **a** TEAD4 - MCF7
- **b** TEAD4 - FULVR
- **c** VGLL1 - FULVR
- **d** VGLL1 Binding Regions (FULVR)
- **e** TEAD4
  - MCF7: 11%
  - FULVR: 6%
  - VGLL1: 3%
- **f** H3K27ac signal at TEAD4/VGLL1 co-bound regions in FULVR cells
- **g** % of INPUT
  - CTGF Promoter
  - AMOTL2 3' Enhancer
  - ANKR21 Promoter
  - TGFBR2 Enhancer
  - LAMP Inton 1
  - CTGF -8:3kb
  - TUBB Promoter
  - YAP/TAZ Negative Regions

*VGLL1 (MCF7) VGLL1 (FULVR) YAP1 (MCF7) YAP1 (FULVR) TAZ (MCF7) TAZ (FULVR)
Supplementary Figure 4

a) Functional annotation of VGLL1 peaks

b) Relative Growth

MCF7 ActCas9-VGLL1:ActCas9-Vector

RNA-seq

ActCas9-VGLL1:ActCas9-Vector

RT-qPCR

Gene Expression

VGLL1-FULVR-1 and VGLL1-FULVR-2

f) RNA-seq

Growth (% relative to vehicle)

h) AZD9496 [log_{10} nM]
**Supplementary Figure 6**

**a**

![Graph showing TEAD4 peaks and VGLL1-fulvr comparison](image)

**b**

![Bar graph showing EGFR expression in MCF7, T47D, and ZR-75-1 cell lines](image)

**c**

![Gene expression of EGFR and VGLL1](image)

**d**

![Western blot images of MCF7-FULVR and T47D-FULVR](image)

**e**

![Western blot images of T47D-FULVR](image)

**f**

![Growth curve of MCF7 and ZR-75-1 with Erlotinib](image)

**g**

![Western blot images of MCF7 and MCF7-FULVR](image)

**h**

<table>
<thead>
<tr>
<th>METABRIC - 1904 samples</th>
<th>TCGA RPPA - 673 samples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rank</strong></td>
<td><strong>Gene</strong></td>
</tr>
<tr>
<td>1</td>
<td>EGFR</td>
</tr>
<tr>
<td>2</td>
<td>CDH3</td>
</tr>
<tr>
<td>3</td>
<td>PPP1R14C</td>
</tr>
<tr>
<td>1</td>
<td>EGFR</td>
</tr>
<tr>
<td>2</td>
<td>CDH3</td>
</tr>
<tr>
<td>3</td>
<td>CCNE1</td>
</tr>
</tbody>
</table>

**i**

![Graph showing EGFR expression in ER+ Breast Cancer METABRIC and TCGA](image)
**Figures and Tables**

**Figure a:** Growth % relative to vehicle vs Verteporfin [log(uM)].

**Figure b:** Fold growth vs Days for MCF7-FULVR(2) with different concentrations of VP.

**Figure c:** Gene expression (FC) for RNA-seq and RT-qPCR.

**Figure d:** Bar graph showing the relative expression of genes with different treatments.

**Figure e:** Expression (log FC) for different gene sets.

**Figure f:** Circle diagram showing the percentage of genes downregulated by VP.

**Figure g:** Box plot showing the expression levels from RNA-seq.

**Figure h:** Fold change in gene expression (log) with different treatments.

**Figure i:** Survival analysis for ER+ Breast Cancer (METABRIC) with different treatments.

**Supplementary Figure 7**

---

*Note: The figure includes various statistical analyses and visual representations to support the findings presented in the main text.*