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

Carolina Gemma[1,*], Anup K Singh[1], Antonino Belfiore[2], Chun-Fui Lai[1], Manikandan Periyasamy[1], Sara Abuelmaaty[3, 4], Claire M Davies[5], Xiangyi Wang[1], Georgia M. Simmons[1], Van T M Nguyen[1], Luca Magnani[1], Emad Rakha[3, 4], Lesley-Ann Martin[6], R. Charles Coombes[1], Giancarlo Pruneri[2], Lakjaya Buluwela[1], Simak Ali[1,*]

[1]Department of Surgery and Cancer, Imperial College London, Hammersmith Campus, London, W12 0NN, UK.

[2]Pathology Department, Fondazione IRCCS Istituto Nazionale Tumori and University of Milan, School of Medicine.

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

[4]Department of Histopathology, Nottingham University Hospital NHS Trust, City Hospital Campus, Hucknall Road, Nottingham, NG5 1PB, UK.

[5]ECMC Imperial College. Department of Surgery and Cancer, Imperial College London, Hammersmith Campus, London, W12 0NN, UK.

[6]Breast Cancer Now Toby Robins Research Centre, Institute of Cancer Research, London, SW7 3RP, 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[28,29]. Human VGLL1 can substitute for *Drosophila* vg in wing formation[30], underlining the conserved role of VGLL1 in driving TEAD-directed gene expression. Indeed, VGLL1 competes with YAP for binding to TEAD4 *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[28,31].

Fulvestrant is a key treatment for advanced ER+ breast cancer following relapse on prior endocrine therapies[32,33], with particular utility in combination with drugs targeting other pathways, e.g. CDK4/6 inhibitors[34] and for treating AI-resistance in metastatic breast cancer due to activating ESR1 mutations[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.

Results

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[16]. We used H3K27ac ChIP-seq signal at gene promoters as a mark of active promoters, as previously reported[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 *GREB1* and *PGR* in FULVR cells. Epigenetic activation of *VGLL1* and *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, *VGLL1* and

138 *VGLL3* expression was also higher in tamoxifen resistant (TAMR) MCF7 cells, but not
139 in long-term estrogen-deprived (LTED) cells (mimicking AI resistance)
140 (Supplementary Fig. 1b), suggesting a mechanism for VGLL1/3 induction that requires
141 inhibition of ER activity. Indeed, our previous work revealed that in LTED cells
142 activation of cholesterol biosynthesis pathways, generate the ER ligand 27-
143 hydroxycholesterol[16]. Hence, ER is not inhibited in LTED cells. RT-qPCR and
144 immunoblotting confirmed the results of the H3K27ac ChIP-seq and RNA-seq
145 analyses (Fig. 1c, d). Expression of several TEADs was also increased in MCF7-
146 FULVR cells. YAP and TAZ, the best-known cofactors for TEADs, were expressed in
147 MCF7 cells, but levels of YAP were reduced in FULVR cells. Other genes implicated
148 in the HIPPO pathway, were unchanged in MCF7-FULVR cells (Fig. 1d and
149 Supplementary Fig. 1c). VGLL1 and VGLL3 were similarly over-expressed in
150 independently derived MCF7-FULVR cells (Supplementary Fig. 1d, e). Furthermore,
151 VGLL1, but not VGLL3, was induced in additional FULVR cells generated in other ER+
152 breast cancer cell lines. In these additional FULVR cell lines, expression of YAP, TAZ
153 and other HIPPO components was unchanged, or was reduced (Supplementary Fig.
154 1d, e).
155
156 We also measured VGLL1 expression in a cohort of patients with tumour biopsies
157 collected prior to and following progression on fulvestrant, using RNA *in situ*
158 hybridisation and immunohistochemical detection. In agreement with the results in
159 FULVR cell lines, patients with breast cancer relapse post-fulvestrant treatment had
160 substantially higher VGLL1 mRNA and protein (Fig. 1e and Supplementary Fig.1f).
161 Collectively, these data clearly show that upregulation of the TEAD coactivator VGLL1
162 is a frequent feature of fulvestrant-resistant breast cancer.
163
164 **VGLL1 expression is induced upon ER downregulation.** Since VGLL1 expression
165 emerges in fulvestrant resistance, we determined if anti-estrogen-directed ER
166 inhibition can induce VGLL1 expression. Indeed, 24 or 48 hours of fulvestrant
167 treatment stimulated VGLL1 expression in MCF7 cells (Fig. 1f and Supplementary Fig.
168 2a) and in other ER+ breast cancer cell lines (Fig. 1f and Supplementary Fig. 2c).
169 RNAi-mediated *ESR1* knockdown similarly promoted *VGLL1* expression in all cell
170 lines (Fig. 1g and Supplementary Fig. 2d, e). However, change in VGLL3 expression
171 by fulvestrant or ESR1 siRNA was less consistently observed, being stimulated in

172 MCF7 and BT474, but not in the other cell lines (Supplementary Fig. 2c, d, f). Other
173 HIPPO pathway genes were not affected by fulvestrant or ER knockdown, although
174 small increases in the expression of some TEADs and the LATS1/2 HIPPO pathway
175 kinases (which promote YAP/TAZ degradation[22]), were observed (Supplementary Fig.
176 2b, c, d, f). The important clinical activity of fulvestrant[32,33], has led to the development
177 of next generation SERDs, some of which have progressed through clinical testing[40-
178 42]. As seen for fulvestrant, *VGLL1* expression was increased by other SERDs (GDC-
179 0810, AZD9496, RAD1901), with accompanying inhibition of ER activity
180 (Supplementary Fig. 2g). Together, these data show clearly that downregulation of ER
181 is sufficient to induce VGLL1 expression in ER+ breast cancer cells.
182
183 **VGLL1 is recruited to TEAD binding sites at active regulatory regions in**
184 **fulvestrant-resistant breast cancer cells.** To investigate the mechanistic roles of
185 VGLL1 and TEADs in fulvestrant resistance, we performed ChIP-seq for VGLL1 in
186 MCF7-FULVR cells and for TEAD4, as it is the best characterized TEAD in cancers[43-
187 45] (Supplementary Fig. 3a). Remarkably, three times more TEAD4 binding events
188 were identified in FULVR cells than in the isogenic MCF7 cells (20,894 vs 6,968 peaks,
189 respectively) (Fig. 2a). Most (75%) of the TEAD4 binding events in MCF7 cells were
190 also present in FULVR cells. TEAD binding sites were by far the most strongly
191 enriched motifs identified with *de novo* motif analysis in MCF7 and FULVR cells (Fig.
192 2a and Supplementary Fig. 3b). These findings suggest that TEAD4 activity is
193 substantially elevated in FULVR cells. Also enriched at TEAD4 binding regions in
194 FULVR cells were activating protein-1 (AP-1) binding sequences (Supplementary Fig.
195 3b), consistent with a previous report showing that TEADs and AP-1 are part of a
196 multiprotein complex that regulates genes involved in proliferation and migration in
197 cancer cells[43].
198
199 The transcriptional activity of TEADs requires binding of co-activators[46], including
200 YAP/TAZ[18-20] and VGLLs[28,29]. ChIP-seq for VGLL1 in FULVR cells, identified 29,930
201 VGLL1 peaks (q-value < 0.01; Supplementary Table 1 and Supplementary Fig. 3c).
202 There was a strong positive correlation between VGLL1 and TEAD4 binding events
203 (p<2.2x10^{-16}; Fig. 2b, c). Moreover, the TEAD4 binding sequence was the most highly
204 enriched motif at VGLL1 binding regions (Supplementary Fig. 3d), indicative of VGLL1
205 recruitment to the TEAD DNA binding proteins. Indeed, VGLL1 peaks were highly

206 enriched at the TEAD4 binding sites in FULVR cells (p<0.0001, Fig. 2d), indicating
207 that VGLL1 recruitment to chromatin is mainly through interaction with TEADs.
208 Regions co-bound by VGLL1 and TEAD4 were present at active regulatory regions in
209 FULVR cells and were strongly associated with active chromatin, as revealed by
210 correlation with H3K27ac signal and reduced nucleosome occupancy at the peak
211 centre (Fig. 2e, f). Moreover, in MCF7 cells these regions were associated with an
212 inactive chromatin conformation (Supplementary Fig. 3f), indicating that the higher
213 TEAD4 occupancy observed in FULVR cells is associated with increased chromatin
214 accessibility at the TEAD4 binding regions. The vast majority of the VGLL1 and TEAD4
215 peaks were present at intergenic and intronic regions (Supplementary Fig. 3e),
216 suggesting that they are mainly present at distal enhancers, as also shown previously
217 for YAP/TEAD binding in other cell types[18,43,47]. In MCF7-FULVR cells, VGLL1 and
218 TEAD4 were commonly co-bound at known YAP and TEAD target genes, including
219 CTGF, AMOTL2 and ANKRD1[47,48], and was accompanied by increased H3K27ac at
220 these genes (Fig. 2g). Consistent with this, expression of these genes was greatly
221 elevated in MCF7-FULVR cells, compared to MCF7 cells. ChIP-qPCR confirmed that
222 VGLL1 was recruited to these binding sites in MCF7-FULVR cells, but not in MCF7
223 cells (Supplementary Fig. 3g). YAP and TAZ were largely absent at these genes in
224 FULVR cells, suggesting that VGLL1 is the main TEAD co-activator in driving
225 expression of these genes in MCF7-FULVR cells. Collectively, our data identify VGLL1
226 and TEAD4 as previously unknown core components of the regulatory landscape in
227 breast cancer cells resistant to fulvestrant.
228
229 **VGLL1 promotes resistance of breast cancer cells to fulvestrant.** To gain further
230 insights into the transcriptional program regulated by VGLL1, we performed functional
231 annotation of the VGLL1 binding regions in FULVR cells. Notably, genes associated
232 with VGLL1 peaks were highly enriched in gene sets linked to acquired resistance to
233 endocrine therapies in breast cancer cells (Supplementary Fig. 4a and Supplementary
234 Table 2). Moreover, siRNA-directed VGLL1 knockdown reduced growth of fulvestrant-
235 resistant MCF7 and T47D cells (Fig. 3a and Supplementary Fig. 4b), demonstrating
236 its functional importance in resistant cells.
237
238 We next asked whether activation of the endogenous VGLL1 gene is sufficient for
239 development of resistance to fulvestrant. Applying the CRISPR/Cas9 Synergistic

240 Activation Mediator (SAM) method[49], using a specific sgRNA targeted to the *VGLL1*
241 gene promoter (Fig. 3b), we successfully established VGLL1 expression in MCF7 cells
242 (Fig. 3c and Supplementary Fig. 4c). As reported previously, treatment of ER+ breast
243 cancer cells with fulvestrant results in rapid and near complete cell loss and resistant
244 cells only emerge after 3-6 months[50-52]. Culturing MCF7 cells with transcriptional
245 activation of endogenous *VGLL1* (MCF7 ActCas9-VGLL1) in the presence of 100 nM
246 fulvestrant showed an initial period of growth inhibition, as did the control (MCF7
247 ActCas9-Vector) cells (Fig. 3d). Remarkably, growth recovery of MCF7 ActCas9-
248 VGLL1 cells was observed after ~20 days, with no cell recovery evident in the control
249 cells even after 100 days (Fig. 3d and Supplementary Fig. 4d). We further evaluated
250 two of the independently generated ActCas9-VGLL1 fulvestrant-resistant lines, named
251 MCF7 ActCas9-VGLL1-FULVR-1 and -2. Interestingly, development of fulvestrant
252 resistance in these cells was accompanied by further upregulation of VGLL1, with a
253 concomitant downregulation in the mRNA and protein levels of ER and PR (Fig 3e, f
254 and Supplementary Fig. 4e), as noted for the acquired fulvestrant-resistant models
255 (Fig. 1c and Supplementary Fig. 1d). Moreover, comparison of global gene expression
256 profiles generated by RNA-seq between MCF7 ActCas9-VGLL1-FULVR cells and
257 other fulvestrant-resistant cell lines, revealed that similar transcriptional changes were
258 observed in the ActCas9-VGLL1-FULVR cells, as in the MCF7-FULVR cells
259 (Supplementary Fig. 4f). The growth dependence of MCF7 ActCas9-VGLL1-FULVR
260 cells on VGLL1 could be confirmed by siRNA-mediated VGLL1 knockdown (Fig. 3g),
261 indicating that MCF7 ActCas9-VGLL1-FULVR cells become dependent on VGLL1 for
262 survival, as we observed in the other fulvestrant resistance models. Moreover, not only
263 were MCF7 ActCas9-VGLL1-FULVR cells resistant to fulvestrant (Fig. 3h) but were
264 also insensitive to next generation SERDs (Supplementary Fig. 4h).
265
266 The above results suggest a switch in dependency from ER in the ER+, treatment-
267 naïve, breast cancer cells, to VGLL1 in the fulvestrant-resistant cells. Indeed, induction
268 of VGLL1 expression in the fulvestrant-naïve MCF7 cells was insufficient to induce a
269 shift in the transcriptional program (Supplementary Fig. 4g), which continued to be
270 dependent on ER, as evidenced by the maintained sensitivity to fulvestrant (Fig. 3h).
271 These data suggest that downregulation of ER and ER transcriptional activity following
272 fulvestrant treatment is a necessary first step for adaptation to VGLL1 dependence.
273 Taken together, our data show that fulvestrant-resistant breast cancer cells are

274 dependent on VGLL1 for their growth and that resistance to fulvestrant is rapidly
275 installed in a VGLL1-dependent manner.
276
277 **VGLL1 transcriptional dependency in fulvestrant-resistant breast cancer.** To
278 explore the role of VGLL1 as a coactivator in fulvestrant resistance, we next asked
279 whether FULVR cells rely on VGLL1 to activate their transcriptional programs. Indeed,
280 analysis of our VGLL1 ChIP-seq data revealed that genes predicted as VGLL1 targets
281 and associated with two or more VGLL1 peaks were significantly more highly
282 expressed in the fulvestrant-resistant than in the fulvestrant-sensitive MCF7 cells
283 compared to genes not associated with VGLL1 peaks (Fig. 4a), a finding confirmed in
284 the MCF7 ActCas9-VGLL1-FULVR cells (Supplementary Fig. 5a).
285
286 To confirm these findings using high confidence VGLL1 direct targets, we performed
287 RNA-seq following siRNA-mediated VGLL1 downregulation in FULVR cells. We
288 identified a set of direct VGLL1 targets as those genes that were downregulated upon
289 VGLL1 downregulation in FULVR cells, which we further filtered based on the
290 presence of at least one associated VGLL1 binding site, which generated a list of 762
291 genes (VGLL1-activated genes). Genes whose expression was not altered by siRNA-
292 mediated VGLL1 downregulation that were also further filtered based on the absence
293 of VGLL1 associated peaks were considered as not-VGLL1 targets (n=8,932 genes).
294 (Supplementary Fig. 5b, c). We found that VGLL1 activated genes were significantly
295 more highly expressed in MCF7-FULVR cells than the isogenic fulvestrant-sensitive
296 MCF7 cells, compared to not-VGLL1 targets (Fig. 4b). Moreover, VGLL1 activated
297 genes were highly enriched in VGLL1 and TEAD4 peaks (Supplementary Fig. 5d, e)
298 and were associated with higher levels of H3K27ac at gene promoters in FULVR cells
299 compared to MCF7 cells (Fig. 4c and Supplementary Fig. 5f), further supporting the
300 notion that VGLL1 regulates genes that become more active in the resistant cells. In
301 line with this, TEAD4 peaks associated with the VGLL1-activated genes were present
302 at regions with higher H3K27ac signal and TEAD4 recruitment in FULVR cells
303 compared to MCF7 cells (Supplementary Fig. 5g). Taken together, these data indicate
304 that VGLL1 drives expression of TEAD-target genes in FULVR cells. Interestingly, we
305 also found that genes whose expression depends on VGLL1 in FULVR cells were
306 significantly more expressed than not-VGLL1 targets (Fig. 4d). Functional annotation
307 of the VGLL1-activated genes revealed enrichment for pathways associated with

308 growth factor signalling, including growth factor binding and transmembrane receptor
309 protein kinase activity (Fig. 4e). Indeed, genes included in these functional categories,
310 e.g. IGFBP3, ITGB6, EGFR and TGFB2, were among the VGLL1 target genes most
311 highly upregulated in fulvestrant-resistant compared to fulvestrant-sensitive cells (Fig.
312 4f, g and Supplementary Fig. 5h). Collectively, these data indicate that VGLL1 acts as
313 an essential transcriptional co-factor in FULVR cells, driving the gene expression
314 programs that facilitate survival and growth of fulvestrant-resistant breast cancer cells.
315
316 **VGLL1 drives EGFR expression in breast cancer cells resistant to fulvestrant.**
317 Our findings link VGLL1 activity with genes associated with receptor tyrosine kinase
318 (RTK) signalling. High expression of the RTKs, EGFR and HER2, has been linked to
319 reduced therapeutic efficacy of anti-estrogens, including tamoxifen and fulvestrant[53]
320 and ectopic expression of EGFR promotes fulvestrant resistance in breast cancer
321 cells[54]. The TEAD4 ChIP-seq showed TEAD4 recruitment to the EGFR enhancer and
322 promoter region in MCF7-FULVR, but not in MCF7 cells (Fig. 5a). VGLL1 co-localised
323 with TEAD4 in MCF7-FULVR cells at the EGFR enhancer. Interestingly, *EGFR*,
324 together with *VGLL1*, were among the top genes with higher TEAD4 occupancy in
325 FULVR cells compared to MCF7 cells (Supplementary Fig. 6a). VGLL1 and TEAD4
326 co-binding at the *EGFR* enhancer, together with acquisition of H3K27ac, was also
327 evident in MCF7-ActCas9-VGLL1-FULVR cells (Fig. 5b), confirming the binding of
328 VGLL1 and TEAD4 to the regulatory regions of the *EGFR* gene in fulvestrant-resistant
329 cells. VGLL1 and TEAD4 co-binding at the *EGFR* enhancer was accompanied by a
330 large increase in EGFR expression in MCF7-ActCas9-VGLL1-FULVR cells (Fig. 5c,
331 d). Importantly, siRNA-mediated downregulation of Cas9 in these cells led to a
332 decrease in VGLL1 and EGFR expression (Fig. 5e). Similarly, EGFR expression was
333 significantly increased in several fulvestrant-resistant cells in which VGLL1 expression
334 was induced, compared to their isogenic fulvestrant-sensitive breast cancer cell lines
335 (Supplementary Fig. 6b). In agreement with these findings, VGLL1 knockdown was
336 sufficient to reduce the levels of EGFR in fulvestrant-resistant breast cancer cell lines
337 (Supplementary Fig. 6c, d, e).
338
339 VGLL1-induced EGFR expression in MCF7ActCas9-FULVR cells was accompanied
340 by a strong increase in AKT and ERK1/2 MAPK phosphorylation (Fig. 5d), while
341 VGLL1 knockdown also reduced AKT and ERK1/2 phosphorylation (Supplementary

342 Fig. 6d, e), indicating that VGLL1-directed EGFR over-expression in FULVR cells
343 facilitates activation of downstream signalling pathways that promote cell survival and
344 proliferation[55,56]. MCF7 cells are characterised by low EGFR and HER2 expression
345 and insensitivity to EGFR inhibitors[57]. Consistent with high EGFR expression and
346 activity, treatment of MCF7-ActCas9-VGLL1-FULVR cells with the EGFR inhibitor
347 erlotinib[58], potently inhibited cell growth, whereas the parental fulvestrant sensitive cell
348 lines (MCF7-ActCas9-Vector and MCF7-ActCas9-VGLL1) were largely insensitive to
349 erlotinib (Fig. 5f). In line with this, erlotinib treatment led to reductions in EGFR
350 phosphorylation and inactivation of the downstream signalling pathways in MCF7-
351 ActCas9-VGLL1-FULVR cells (Fig. 5g). A similar result was obtained in additional
352 fulvestrant resistant cell lines (Supplementary Fig.6f, g).
353
354 Our results provide a novel mechanism leading to increased EGFR signalling in breast
355 cancer cells, in which VGLL1-TEAD4 co-binding at the EGFR enhancer induces
356 EGFR expression. In agreement with this, *EGFR* expression was highly positively
357 correlated with *VGLL1* expression in ER+ breast cancer patients from TCGA[59] and
358 METABRIC[60] (Fig. 5h and Supplementary Fig. 6i). Remarkably, EGFR was the
359 highest-ranked gene co-expressed with VGLL1 in breast cancer patients from
360 METABRIC and the highest ranked protein significantly correlated with VGLL1 in
361 breast cancer from TCGA (Supplementary Fig. 6h).
362
363 Collectively, these data identify EGFR as an important VGLL1 target gene in
364 fulvestrant-resistant breast cancer cells, in which rewiring of the epigenome results in
365 VGLL1 and TEAD4 binding to the EGFR enhancer, leading to induction in EGFR
366 expression and activation of downstream EGFR signalling pathways (Fig. 6i).
367
368 **Targeting VGLL1 transcriptional activity with verteporfin.** Our findings identify
369 VGLL1 as a new therapeutic target in metastatic, endocrine-resistant breast cancer.
370 Verteporfin (VP), a drug used for photodynamic treatment of macular degeneration,
371 has been identified as an inhibitor of YAP/TEAD. VP disrupts the physical interaction
372 between YAP and TEADs, resulting in downregulation of YAP/TEAD target genes[61,62].
373 Given that the VGLL1-TEAD interaction is structurally similar to that of YAP-TEAD[28,31],
374 we reasoned that VP would disrupt VGLL1-TEAD4 interaction and so inhibit
375 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[28] and MMP9[63], 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

410 worse prognosis (Fig. 6i). A similar result was obtained when considering only ER+
411 breast cancer cases (Supplementary Fig. 7i), indicating that our VGLL1 transcriptional
412 signature of VGLL1 activated genes has useful prognostic value for breast cancer
413 patients treated with endocrine therapies. In summary, our results show that VGLL1
414 transcriptional activity is sensitive to VP and suggest that inhibition of VGLL1
415 recruitment to TEADs, as with VP, could provide an approach to treating VGLL1-
416 dependent, endocrine-resistant breast cancer.
417
418
419 **Discussion**
420
421 Fulvestrant is the main endocrine agent for advanced ER+ breast cancer following
422 relapse on prior endocrine therapies[32,33] and newer generation SERDs with improved
423 pharmacological properties are in advanced clinical development[41,42]. Understanding
424 resistance to SERDs is urgently needed for aiding patient selection and in identifying
425 therapeutic interventions for patients who progress on SERDs. Our work has identified
426 VGLL1/TEAD4, as previously unknown key transcriptional regulators and drivers of
427 resistance to fulvestrant and suggests new therapeutic opportunities for endocrine-
428 resistant breast cancer. We show that SERD inhibition of estrogen-regulated gene
429 expression is accompanied by induction of VGLL1 expression. The induction of
430 VGLL1 expression upon ER inhibition drives fulvestrant resistance through the
431 establishment of a VGLL1/TEAD directed transcriptional program that supports cell
432 survival and proliferation. Our work strongly supports the idea that VGLL1/TEAD4 act
433 as transcriptional master regulators in FULVR cells and that resistance to fulvestrant
434 emerges as a consequence of a switch in transcriptional dependencies from ER in
435 pre-treatment ER+ breast cancer, to VGLL1/TEAD in fulvestrant-resistant breast
436 cancer. The clinical significance of our findings is further emphasized by the validity of
437 the signature of VGLL1 activated genes as a prognostic tool in a large set of breast
438 cancer patients treated with endocrine therapies
439
440 The well-established oncogenic role of TEADs in different cancers has been closely
441 linked to the activity of YAP/TAZ as TEAD coactivators in the majority of studies[43,48].
442 However, it is increasingly clear that the canonical functions of TEADs as downstream
443 effectors of YAP and TAZ in the Hippo pathway are not the only mechanism by which

444 TEADs contribute to cancer progression in different cancers[44,64,65]. For instance, in
445 ER+ breast cancer, a non-canonical function of YAP and TEAD4, has recently been
446 reported, where they act as ER cofactors, thereby regulating ER target genes rather
447 than canonical TEAD target genes[44]. Our results are in agreement with these findings,
448 as we found that classical TEAD target genes were expressed at low levels in ER+
449 breast cancer cells and we did not observe enrichment of YAP/TAZ binding at these
450 target genes. Our work also revealed that many of the established YAP/TAZ/TEAD
451 target genes[47,66,67], are upregulated in FULVR cells, but YAP/TAZ are largely absent
452 at these genes. Instead, VGLL1 was co-bound with TEAD at these genes in FULVR
453 cells and the expression of these genes was regulated by VGLL1.
454
455 Several lines of evidence indicate that endocrine therapies are drivers of epigenetic
456 reprogramming[16,68]. Our results further show that reprograming of the epigenetic
457 landscape in breast cancer cells resistant to fulvestrant is associated with increased
458 TEAD4 recruitment to the chromatin and association of VGLL1/TEAD4 with active
459 enhancers in the resistant cells, and uncover a previously unrecognized role for
460 VGLL1/TEAD4 as transcriptional drivers in advanced ER+ breast cancer.
461
462 Additionally, our work identified EGFR as a VGLL1/TEAD activated gene in FULVR
463 cells. We show that instalment of fulvestrant resistance following VGLL1
464 overexpression was accompanied by VGLL1 and TEAD4 recruitment to the EGFR
465 enhancer, upregulation of EGFR, associated activation of EGFR downstream
466 effectors and sensitivity to EGFR inhibition. Recent studies show that EGFR and
467 HER2 amplification, as well as mutations in HER2 and genes in the MAPK pathway
468 are increased in metastatic, endocrine-resistant ER+ breast cancer, together
469 accounting for 10-15% of cases[54,69] and demonstrating the importance of elevated
470 EGFR and HER2 signalling in acquired resistance to endocrine treatments. Our
471 results evidence an alternate mechanism by which EGFR expression is induced in
472 endocrine-resistant breast cancer and thus advance a rationale for expanded clinical
473 utility for EGFR inhibitors.
474
475 Targeting transcription factors in cancer has proved to be challenging, yet it offers
476 great promise for cancer treatment as many cancers are driven by aberrant
477 transcription factor expression and/or activities[70]. As the importance of VGLL1 in

478 cancer has largely been unexplored, there are at present no reported VGLL1
479 inhibitors. Given the well-established functions of YAP/TAZ/TEAD as tumor promoters
480 in many cancers, there have been considerable efforts to identify inhibitors of their
481 activity[71]. Elucidation of the structure of the VGLL1/TEAD complex has revealed that
482 VGLL1-TEAD interaction is structurally similar to that of YAP-TEAD[28,31]. Indeed, we
483 show that verteporfin, a FDA approved photodynamic therapy drug for macular
484 degeneration, which disrupts YAP/TEAD interaction[61,62], also disrupts the binding of
485 VGLL1 to TEAD4 at the target genes, and selectively inhibits VGLL1 transcriptional
486 activity. Thus, our findings show that VGLL1 transcriptional and growth dependencies
487 could be exploited as a therapeutic vulnerability in advanced ER+ breast cancer and
488 so inhibiting VGLL1 interaction with TEAD, as well as inhibition of downstream VGLL1-
489 activated genes such as EGFR, could be a viable therapeutic approach for patients
490 who progress on endocrine therapies.
491 Collectively, our work uncovers multiple new insights into the biological function of the
492 TEAD coactivator VGLL1 in advanced therapy-resistant breast cancer. The biological
493 function and importance of VGLL1 in other cancers remains poorly understood.
494 However, because of the tissue-restricted expression of VGLL1 and its specific
495 upregulation in several cancers, it is tempting to speculate that VGLL1 could emerge
496 as a promising new therapeutic target in several cancers[72]. Thus, our findings and the
497 VGLL1-EGFR connection may be relevant in other cancers where VGLL1
498 upregulation is observed.
499
500
501 **Methods**
502 Methods, including statements of data availability and any associated accession
503 codes and references, are available online at XXX
504
505
506 **Methods**
507 **Cell lines and treatments.** Cells lines were originally purchased from ATCC, were
508 authenticated at the time of the studies by LGC Standards (Bury, UK) and were
509 regularly screened for mycoplasma infection. BT474, ZR-75-1 and T47D cells were
510 cultured in Dulbecco's modified Eagle's medium (DMEM) with 10% of fetal bovine
511 serum (FBS) and 100U penicillin/0.1 mg ml^{-1} streptomycin. MCF7 cells were cultured

512 in DMEM with 10% of fetal bovine serum (FBS) and 100U penicillin/0.1 mg ml^{-1}
513 streptomycin plus 10nM estradiol (SIGMA E8875). MCF7(2) cells were cultured in
514 RPMI 1640 with 10% of fetal bovine serum (FBS) and 100U penicillin/0.1 mg ml^{-1}
515 streptomycin plus 1nM estradiol. All fulvestrant resistant isogenic derivatives of ER+
516 breast cancer cell lines have been previously described [16,50,73,74], and were cultured
517 as the corresponding parental cells with the addition of 100nM fulvestrant (Sigma
518 I4409).
519 For VGLL1 overexpression, MCF7-ActCas9-VGLL1 cells were generated by
520 transducing MCF7 cells with lentiviral MS2-P65-HSF1_Hygro (Addgene plasmid
521 #61426), lentiviral dCAS-VP64_Blast (Addgene plasmid #61425) and sgRNA(MS2)
522 cloning backbone (Addgene plasmid #61424) using the following VGLL1 sgRNA:
523 ATTCCTGCAGGTGCCCAACCAGG, that was cloned into the sgRNA(MS2) cloning
524 backbone plasmid. The control MCF7-ActCas9-Vector cells were generated
525 simultaneously using the above mentioned plasmids, except that sgRNA(MS2) cloning
526 backbone empty vector was used instead. MCF7-ActCas9-Vector and MCF7-
527 ActCas9-VGLL1 cells were maintained in media containing 0.2 mg/ml hygromycin
528 (Corning 30-240-CR), 10 μg/ml blasticidin (Corning 30-100-RB) and 1μg/ml puromycin
529 (gibco A1113803).
530 MCF7-ActCas9-VGLL1-FULVR cells were generated by culturing MCF7-ActCas9-
531 VGLL1 cells in the appropriate media with the addition of 100nM or 1000nM
532 fulvestrant.
533 All drugs were solubilised in DMSO. Verteporfin (Sigma SML0534), was used at a final
534 concentration of 2μM for 24 h, unless otherwise specified. SERMs: tamoxifen (Sigma
535 H7904) and raloxifene (Tocris 2280), as well as SERDs: fulvestrant, GDC-0810
536 (Genentech), AZD9496 (Astra Zeneca) and RAD1901(MedChem Express HY-
537 19822A) were used at a final concentration of 100nM for 48 h, unless otherwise
538 specified. Erlotinib (VWR CAYM10483) was used at a final concentration of 3.1 uM for
539 24 h, unless otherwise specified.
540
541 **Clinical samples.** From the institutional database of the Fondazione IRCCS Istituto
542 Nazionale Tumori of Milan, we retrospectively identified 72 metastatic ER+, HER2
543 negative breast cancer patients diagnosed between 1997 and 2018, treated with
544 endocrine therapy, including at least one line of fulvestrant, and standard

545 chemotherapy. For the present study, we selected 15 patients for which breast cancer
546 biopsies were available both prior to fulvestrant (including primary tumor samples and
547 metastatic localizations) and post fulvestrant administration. All selected cases were
548 carefully revised by an expert breast pathologist, including the evaluation of ER and
549 PR receptor status, proliferative index and tumor cellularity. All patients included in
550 this study gave their written consent to donate the tissue remaining after their
551 diagnostic procedures to Istituto Nazionale Tumori of Milan.
552
553 **siRNA transfections.** siRNAs (25nM) were transfected with Lipofectamine RNAi-
554 MAX (Life Technologies) in antibiotics-free medium according to the manufacturer's
555 instructions. Cells were harvested 72 h after transfection. The following individual ON
556 TARGETplus siRNAs (Horizon Discovery) were used for VGLL1: J-017939-06
557 (siVGLL1 #1), J-017939-08 (siVGLL1 #2) and J-017939-07 (siVGLL1 #3). For dCAS9
558 the following siRNAs were used: CCGAAGAGGUCGUGAAGAA (Cas9 siRNA#1) and
559 GGGAAAAGAUCGAGAAGAU (Cas9 siRNA#2). ESR1 siRNA was purchased from
560 Qiagen (SI03114979, FlexiTube siRNA). The negative control siRNA (siControl) was
561 purchased from Horizon Discovery (ON TARGETplus Non-targeting control siRNA #2,
562 D-001810-02)
563
564 **RNA in situ hybridization.** Formalin fixed, paraffin embedded breast cancer tissue
565 sections (5 µm) from matched pre- and post-fulvestrant patient biopsies were
566 processed for RNA in situ detection using the RNAscope Multiplex Fluorescent
567 Reagent Kit v2 according to the manufacturer's instructions (Advanced Cell
568 Diagnostics), using RNAscope probe Hs-VGLL1 446731. Nuclei were stained with
569 DAPI and the RNAscope probe was detected using Cy5 fluorescent dye (PerkinElmer
570 NEL745E001KT TSA Plus Cyanine 5). Data was analysed using Fiji and CellProfiler
571 using custom macros, provided upon request.
572
573 **Immunohistochemistry.** Immunohistochemical staining was performed on formalin-
574 fixed, paraffin embedded 4 µm tissue sections using VGLL1 polyclonal antibody
575 (Proteintech, 10124-2-AP). Antigen retrieval was performed based on the
576 manufacturer's recommendations (citrate buffer pH 6.0 at 1000 W for 20 min using
577 microwave). Visualization was performed using the Novocastra Novolink Polymer
578 Detection Systems kit (RE7280-K, Leica, Biosystems, UK).

579

580 **Western blot.** Cells were washed twice in ice-cold PBS and harvested in RIPA Buffer
581 (10 mM Tris-HCl (pH 7.6), 1mM EDTA, 0.1% SDS, 0.1% sodium deoxycholate, 1%
582 Triton X-100 and phosphatase and protease inhibitors) and lysed by sonication.
583 Samples were centrifuged at 4 °C, maximum speed, for 10 min, then the supernatant
584 was transferred to a clean Eppendorf tube. Extracts were quantified using the Pierce
585 BCA Protein Assay Kit, according to the manufacturer's instructions (ThermoFisher
586 Scientific, 23227). Samples were run on 10% polyacrylamide gels and transferred onto
587 polyvinylidene difluoride (PVDF) membranes by wet electrophoretic transfer. Blots
588 were blocked with either 5% non-fat dry milk (non-phosphoprotein) or 5% BSA
589 (phosphoprotein) and incubated with primary antibodies overnight at 4°C. Secondary
590 antibodies were incubated for 1 h at room temperature and then blots were developed
591 with goat anti-mouse (Biorad, 1706516) and anti-rabbit (Biorad, 1706515) horseradish
592 peroxidase (HRP) conjugated secondary antibodies. SuperSignal West Pico PLUS
593 Chemiluminescent Substrate (ThermoFisher Scientific, 34577) or SuperSignal West
594 Femto Maximum Sensitivity Substrate (ThermoFisher Scientific, 34095) was used for
595 chemiluminescent imaging using a Fusion solo (Vilber) imager.
596

597 *Antibodies.* The VGLL1 antibody was purchased from Proteintech (10124-2-AP). The
598 ER antibody was from Leica Biosystem (NCL-L-ER-6F11). VGLL3 (ab68262), VGLL4
599 (ab140290), YAP1 (ab56701), TEAD4 (ab58310) and GAPDH (ab9484) antibodies
600 were supplied by Abcam. Antibodies for PR (8757S), TEAD3 (13224S), EGFR
601 (2232S), phospho-Tyr1068 EGFR (2234S), phospho-Tyr1148 EGFR (4404S), AKT
602 (9272S), phospho-Ser473-AKT (9271S), p44/42 MAPK (4695S) and phospho-
603 p44/p42 MAPK (Thr202/Tyr204) (9101S) were from Cell Signaling Technology.
604 TEAD1 (sc-376113) and TEAD2 (sc-67115) antibodies were purchased from Santa
605 Cruz Biotechnology and the TAZ antibody was obtained from BD (clone M2-616,
606 560235),
607

608 **RT-qPCR.** Total RNA was extracted from cells using RNeasy Mini Kit (Qiagen 74106).
609 For cDNA synthesis 500 ng of total RNA was reverse transcribed using RevertAid First
610 Strand cDNA Synthesis Kit (ThermoFisher Scientific, K1622). RT-qPCR reactions
611 were carried out using Fast SYBR Green Master Mix (ThermoFisher Scientific,

612 4385616) in a StepOnePlus Real-Time PCR thermal cycler (ThermoFisher Scientific).
613 Gene expression levels were calculated relative to GAPDH. The following primers
614 were used:
615 VGLL1: CCCCTCGAGTCAGAGTGAAG, CAGGGACGGTGAGAACTGAT;
616 VGLL3: GCTGGTAAGAGCTGGTCCAA, CCATCCAGAATCTGCCATTT;
617 ESR1: CAGGTGCCCTACTACCTGGA, TCCTTGGCAGATTCCATAGC;
618 PGR: GATGCTTCATCCCCACAGAT, AGGTCTACCCGCCCTATCTC;
619 GREB1: AAGGAGGGCTGGAAACAAAT, CGTTGGAAATGGAGACAAGG;
620 GAPDH: TGGAAATCCCATCACCATCT, TTCACACCCATGACGAACAT;
621 YAP1: CAGCAACTGCAGATGGAGAA, TGGATTTTGAGTCCCACCAT;
622 TAZ: TCATCACCGTGTCCAATCAC, GTGGGAGTGTAGCTCCTTGG;
623 TEAD1: CAAGCCTTTTGTGCAGCAG, AAAATTCCACCAGGCGAAG;
624 TEAD2: CTTGGACTGGATTTCCCTTG, CCGCTACATCAAGCTGAGAA;
625 TEAD3: GAGGCAATGGTACGGTCCT, CTTTGCACAGCCAGCCTAC;
626 TEAD4: TCCACGAAGGTCTGCTCTTT, GTGCTTGAGCTTGTGGATGA;
627 LATS1: ACCTTTCCAGCTCTGTTTGC, AGATCCTCGACGAGAGCAGA;
628 LATS2: GTGGTAGGACGCAAACGAAT, CCGAGGAATGAGCAGATTGT;
629 BTRC: ACAGGATCATCGGATTCCAC, TTGAAACGCAAGTGCAGAAC;
630 SAV1: TGGCTGGTATGTGACAGGAG, ACTTCCTCCTGGATGGGAAC;
631 MOB1a: ACAGCTTGCTTCAGTGCAGA, TATGTTGCCTGAGGGAGAGG;
632 EGFR: GAGGGCAAATACAGCTTTGG, GCCCTTCGCACTTCTTACAC;
633 dCAS9: GAACCGGATCTGCTATCTGC, CGCTCGTGCTTCTTATCCTC;
634 CTGF: TGGAGATTTTGGGAGTACGG, CCTGGTCCAGACCACAGAGT;
635 TGFB2: ACAAGAGCAGAAGGCGAATG, TGCAGCAGGGACAGTGTAAG;
636 IGFBP5: AGGTGTGGCACTGAAAGTCC, ATTGTGACCGCAAAGGATTC;
637 IGFBP3: AGGCTGCCCATACTTATCCA, GGGGTGTACACATTCCCAAC;
638 AMOTL2: ACCACTGCCACTGCTACCAC, AGCAGGCCTATGTGGAGAAA;
639 APEX1: CCCCAGATCAGAAAACCTCA, TTTGGTCTCTTGAAGGCACA;
640 ANKRD1: CCAAATGTCCTTCCAAGCAT, TGAAGGCTGCTCTGGAGAAT.
641
642 **ChIP-seq and ChIP-qPCR.** ChIP was performed as previously described[75], with few
643 modifications. Briefly, cells were crosslinked with 1% formaldehyde in culture medium
644 without antibiotics for 10 min at room temperature. Chromatin extracts were sonicated
645 using a Bioruptor Pico sonication device (Diagenode, B01060001) using 15 cycles (30

646 s on and 30 s off) at maximum intensity. Chromatin was immunoprecipitated using 10
647 µg of antibodies for TEAD4 (Abcam, ab58310), VGLL1 (Proteintech, 10124-2-AP),
648 YAP1 (Abcam, ab56701), TAZ (BD, clone M2-616, 560235) and H3K27ac (Abcam,
649 ab4729). Non-immunoprecipitated chromatin was used as input control. After
650 chromatin de-crosslinking DNA was purified using the QIAquick PCR Purification Kit
651 (Qiagen, 28106). For all ChIP experiments the enrichment was determined by qPCR
652 relative to the input sample using positive and negative control primers. For TEAD4
653 and VGLL1 ChIP-seq two and three highly reproducible biological replicates were
654 generated, with Spearman's correlation values of 0.94-0.99 in all biological replicates
655 (Supplementary Fig. 2a). ChIP enrichment was assessed by ChIP-qPCR before library
656 construction using primers previously described for YAP/TEAD ChIP in other cell
657 types[47]. Libraries for ChIP-seq were prepared using the NEBNext Ultra DNA Library
658 Prep Kit for Illumina (New England Biolabs, E7370S) according to the manufacturer's
659 instructions. Sequencing was performed on an Illumina HiSeq 2500 platform using
660 50bp single end reads. The following primers were used:
661 ANKRD1 Promoter: GAGGGGAGGACAAGCTAACC,
662 AGCTGTCCCCTGACTCTTGA;
663 TGFβ2 Enhancer: AGCTTTGATCACACTGATTCCA,
664 TGCCTCTTCACATCTGTTCATT;
665 AMOTL2 3' Enhancer: ACAGCCCTCCAACTATGCTAAG,
666 CCAACAGCACATTTTCAGGATA;
667 CTGF Promoter: GAGCTGAATGGAGTCCTACACA,
668 GGAGGAATGCTGAGTGTCAAG;
669 LSAMP (negative): CTGAAATGAAAGTGGGGACA,
670 GATTTGGTCTTGGCAGGTGT;
671 CTGF -8.3kb (negative): TTGCTGGTGGTAGGGAAATACT,
672 TCACTGCACCTTTGCTTTTCTA;
673 TUBB Promoter (negative): TCCTGTACCCCCAAGAACTG,
674 ATTGTTGTCCATGCTGCAAA;
675 EGFR (enh): CACACCTGAGCATGTCCTTG, GCAATGGGATCGAGTTGTTT;
676 GAPDH (negative): TCGACAGTCAGCCGCATCT, CTAGCCTCCCGGGTTTCTCT.
677

678 **RNA-seq.** Total RNA was extracted from cells using RNeasy Mini Kit (Qiagen 74106).
679 RNA-seq libraries were prepared using NEBNext Ultra II RNA Library Prep Kit for
680 Illumina (New England Biolabs, E7775) and sequencing was performed on an Illumina
681 HiSeq 2500 platform. Raw reads were aligned to the human reference genome hg19
682 version using TopHat[76]. Redundant reads were removed using SAMtools[77]. HTSeq[78]
683 was used to count reads at UCSC annotated genes. Normalization and differential
684 expression analysis were carried out using DESeq2 package[79]. Integrated Genomics
685 Viewer (IGV, Broad Institute) was used to visualize the Z-score normalized coverage
686 tracks generated with R packages Rsamtools
687 (http://bioconductor.org/packages/Rsamtools) and rtracklayer[80]. Fold changes in gene
688 expression between two cell lines or two conditions were calculated as the ratio of
689 normalized read counts obtained with DESeq2. The 5th percentile, first quartile,
690 median, third quartile and 95th percentile are plotted in box-and-whiskers graphs.
691
692 **GO analysis.** GO analysis on the VGLL1 activated genes and genes downregulated
693 by VP was performed using g:Profiler (https://biit.cs.ut.ee/gprofiler/gost)[81]
694
695 **Growth assays.** Cell growth following siRNAs or drug treatments was assessed using
696 the sulphorhodamine B (SRB) assay[82]. Briefly, 2000-5000 cells were seeded per well
697 in 96-well plates. Cells were allowed to attach overnight before treatment with drugs
698 or transfection with siRNAs. Medium was changed every 3 days and cells were fixed
699 by adding 100 µl of cold 40% (wt/vol) trichloroacetic acid (TCA) to each well for at least
700 60 min. The plates were washed five times with distilled water, 100 µl of SRB reagent
701 (0.4% wt/vol SRB in 1% wt/vol acetic acid) was added to each well, and the plates
702 were allowed to incubate for 1h. The plates were then washed five times with 1%
703 (wt/vol) acetic acid and allowed to dry overnight. SRB solubilization was performed by
704 adding 100 µl of 10 mM Tris HCl per well to the plates, followed by shaking for 10 min.
705 Optical density (OD) was then measured using a Sunrise microplate reader (Tecan)
706 at λ=492 nm. Six wells were analysed for each experimental condition. Data are
707 presented as mean ± standard error of the mean (s.e.m.). Data are presented as
708 growth relative to control cells (treated with vehicle or transfected with control siRNA),
709 using the first day of treatment as baseline. For time-course experiments, multiple
710 plates were seeded and drugged in identical fashion, and at the indicated time points.

711 Each growth experiment was independently confirmed using three biological
712 replicates.
713 **Live cell imaging.** To follow the growth of cells over the course of three months in the
714 presence of drug containing media, we used the live cell imaging system Incucyte
715 Zoom (EssenBioscience). One million cells were seeded in a T75 flask and after 24h
716 the corresponding media was added to the cells with the addition of fulvestrant at the
717 indicated concentrations. Medium and drugs were replaced twice per week and cells
718 were imaged every 2 days. Analysis was performed using the Incucyte Zoom software
719 with parameters optimized for each cell line and using confluency (μm^2) as a measure
720 of cell growth. Raw data was extracted and plotted in GraphPad Prism. Each
721 experiment was independently confirmed using three biological replicates.
722
723 **ChIP-seq analysis.**
724 Raw reads were aligned to the human reference genome hg19 version using Bowtie[83].
725 Redundant reads were removed using SAMtools[77]. VGLL1 and TEAD4 peaks were
726 called using MACS2[84] using the input sample as a control, with default parameters
727 and q value less than 0.01. For each ChIP-seq target, peaks common to all the
728 biological replicates were kept for further analysis. Integrated Genomics Viewer (IGV,
729 Broad Institute) was used to visualize the Z-score normalized coverage tracks
730 generated with R packages Rsamtools (http://bioconductor.org/packages/Rsamtools)
731 and rtracklayer[80]. The overlap of peaks from different ChIP–seq experiments was
732 determined using the BEDTools2 suite[85]
733
734 *Ratio of H3K27ac between FULVR vs MCF7 cells*: ChIP-seq data for H3K27ac in
735 FULVR and MCF7 cells was obtained from[16]. A list of gene promoters was
736 downloaded from the Table Browser of UCSC genome browser[86] by selecting 3kb-
737 wide regions centred on each TSS annotated in the human genome version hg19. The
738 BEDTools2 suite was used to calculate the coverage of H3K37ac reads at gene
739 promoters. The H3K27ac coverage at promoters was normalized to the total number
740 of million mapped reads to generate normalized read counts in units of reads per
741 million mapped reads (RPM). For each gene promoter we calculated the ratio of
742 normalized H3K27ac coverage between FULVR and MCF7 cells to generate the plot
743 in Fig.1a.

744

745 *ChIP-seq heatmaps and average profiles.* Normalized coverage tracks were plotted
746 as heatmaps and average signal profiles with SeqPlots[87] using a window of 2-3kb
747 centred around the peak centre. Heatmap rows were sorted from high to low signal.
748 Average profiles are presented as the mean of the normalized coverage ± s.e.m. and
749 95% confidence interval.
750

751 *Peak annotation and Motif analysis.* The HOMER suite[88] was used for *de novo* motif
752 discovery analysis of the VGLL1 and TEAD4 peaks using the findMotifsGenome.pl
753 script with default parameters. The annotatePeaks.pl script in the HOMER suite was
754 used for peak annotation analysis to associate peaks with gene targets, and to
755 annotate the location of peaks to different genomic features (see Supplementary Fig.
756 2f).
757

758 *VGLL1-TEAD4 co-binding enrichment analysis.* VGLL1/TEAD4 co-bound regions
759 were defined with the bedtools intersect script from the BEDTools2 suite. To calculate
760 VGLL1/TEAD4 co-binding enrichment, the TEAD4 binding regions were randomized
761 in the hg19 version of the human reference genome using the bedtools shuffle script
762 from the BEDTools2 suite. Chi-square test was applied to assess the co-binding
763 enrichment based on the observed co-binding relative to the expected (random) co-
764 binding, after performing ten permutations of the TEAD4 peaks.
765

766 *Functional annotations.* Functional annotation of the VGLL1 ChIP-seq peaks was
767 performed using the MSigDB Perturbation annotation from GREAT version 3.0.0 [89].
768 Supplementary Table 2 shows the complete annotation. To visualize the top nine most
769 significant terms from the annotation (according to binomial P value), we used ggplot2
770 (https://ggplot2.tidyverse.org) with R scripts adapted from REVIGO[90], scripts provided
771 upon request (see Supplementary Fig. 3a).
772

773 **Generation of the signature of VGLL1 activated genes and not VGLL1 targets.**
774 To identify the VGLL1 direct targets we performed RNA-seq in FULVR cells
775 transfected with three different VGLL1 siRNAs (siVGLL1#1-3) and control siRNA
776 (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

C.G. and S.A. conceived the study. C.G. performed the RNA-seq, ChIP-seq and RNA-FISH experiments and the bioinformatics analyses. C.G., A.K.S., C.F.L., M.P., X.W., G.M.S., L.B., and V.T.M.N performed all other experiments. A.B., G.P. and R.C.C. recruited patients and supplied tissues. S. Abuelmaaty, C.M.D. and E.R. performed IHC staining and scoring. L.A.M. and L.M. provided cell lines. C.G. and S.A. wrote the manuscript.

References

1. Ali, S. & Coombes, R.C. Endocrine-responsive breast cancer and strategies for combating resistance. *Nat Rev Cancer* **2**, 101-112 (2002).
2. Johnston, S.R. & Dowsett, M. Aromatase inhibitors for breast cancer: lessons from the laboratory. *Nat Rev Cancer* **3**, 821-831 (2003).
3. Ali, S., Buluwela, L. & Coombes, R.C. Antiestrogens and their therapeutic applications in breast cancer and other diseases. *Annu Rev Med* **62**, 217-232 (2011).
4. Long, X. & Nephew, K.P. Fulvestrant (ICI 182,780)-dependent interacting proteins mediate immobilization and degradation of estrogen receptor-alpha. *J Biol Chem* **281**, 9607-9615 (2006).
5. Ma, C.X., Reinert, T., Chmielewska, I. & Ellis, M.J. Mechanisms of aromatase inhibitor resistance. *Nature reviews. Cancer* **15**, 261-275 (2015).
6. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
7. Xie, W.*, et al.* Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* **153**, 1134-1148 (2013).

848 8. Hnisz, D., *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**,
849 934-947 (2013).
850 9. Bradner, J.E., Hnisz, D. & Young, R.A. Transcriptional Addiction in Cancer. *Cell* **168**,
851 629-643 (2017).
852 10. Sharma, S.V., *et al.* A chromatin-mediated reversible drug-tolerant state in cancer
853 cell subpopulations. *Cell* **141**, 69-80 (2010).
854 11. Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for
855 systematic annotation of the human genome. *Nat Biotechnol* **28**, 817-825 (2010).
856 12. Whyte, W.A., *et al.* Master transcription factors and mediator establish super-
857 enhancers at key cell identity genes. *Cell* **153**, 307-319 (2013).
858 13. Heintzman, N.D., *et al.* Histone modifications at human enhancers reflect global cell-
859 type-specific gene expression. *Nature* **459**, 108-112 (2009).
860 14. Malik, A.N., *et al.* Genome-wide identification and characterization of functional
861 neuronal activity-dependent enhancers. *Nat Neurosci* **17**, 1330-1339 (2014).
862 15. Stasevich, T.J., *et al.* Regulation of RNA polymerase II activation by histone
863 acetylation in single living cells. *Nature* **516**, 272-275 (2014).
864 16. Nguyen, V.T., *et al.* Differential epigenetic reprogramming in response to specific
865 endocrine therapies promotes cholesterol biosynthesis and cellular invasion. *Nat*
866 *Commun* **6**, 10044 (2015).
867 17. Bi, M., *et al.* Enhancer reprogramming driven by high-order assemblies of
868 transcription factors promotes phenotypic plasticity and breast cancer endocrine
869 resistance. *Nat Cell Biol* **22**, 701-715 (2020).
870 18. Stein, C., *et al.* YAP1 Exerts Its Transcriptional Control via TEAD-Mediated Activation
871 of Enhancers. *PLoS Genet* **11**, e1005465 (2015).
872 19. Zhang, H., *et al.* TEAD transcription factors mediate the function of TAZ in cell growth
873 and epithelial-mesenchymal transition. *J Biol Chem* **284**, 13355-13362 (2009).
874 20. Lamar, J.M., *et al.* The Hippo pathway target, YAP, promotes metastasis through its
875 TEAD-interaction domain. *Proc Natl Acad Sci U S A* **109**, E2441-2450 (2012).
876 21. Ota, M. & Sasaki, H. Mammalian Tead proteins regulate cell proliferation and contact
877 inhibition as transcriptional mediators of Hippo signaling. *Development* **135**, 4059-
878 4069 (2008).
879 22. Zanconato, F., Cordenonsi, M. & Piccolo, S. YAP/TAZ at the Roots of Cancer. *Cancer*
880 *Cell* **29**, 783-803 (2016).
881 23. Nguyen, C.D.K. & Yi, C. YAP/TAZ Signaling and Resistance to Cancer Therapy. *Trends*
882 *Cancer* **5**, 283-296 (2019).
883 24. Davidson, I., Xiao, J.H., Rosales, R., Staub, A. & Chambon, P. The HeLa cell protein
884 TEF-1 binds specifically and cooperatively to two SV40 enhancer motifs of unrelated
885 sequence. *Cell* **54**, 931-942 (1988).
886 25. Jacquemin, P. & Davidson, I. The Role of the TEF Transcription Factors in
887 Cardiogenesis and Other Developmental Processes. *Trends Cardiovasc Med* **7**, 192-
888 197 (1997).
889 26. Halder, G., *et al.* The Vestigial and Scalloped proteins act together to directly
890 regulate wing-specific gene expression in Drosophila. *Genes & development* **12**,
891 3900-3909 (1998).
892 27. Simmonds, A.J., *et al.* Molecular interactions between Vestigial and Scalloped
893 promote wing formation in Drosophila. *Genes & development* **12**, 3815-3820 (1998).

894 28. Pobbati, A.V., Chan, S.W., Lee, I., Song, H. & Hong, W. Structural and functional
895 similarity between the Vgll1-TEAD and the YAP-TEAD complexes. *Structure* **20**, 1135-
896 1140 (2012).
897 29. Pobbati, A.V. & Hong, W. Emerging roles of TEAD transcription factors and its
898 coactivators in cancers. *Cancer Biol Ther* **14**, 390-398 (2013).
899 30. Vaudin, P., Delanoue, R., Davidson, I., Silber, J. & Zider, A. TONDU (TDU), a novel
900 human protein related to the product of vestigial (vg) gene of Drosophila
901 melanogaster interacts with vertebrate TEF factors and substitutes for Vg function in
902 wing formation. *Development* **126**, 4807-4816 (1999).
903 31. Mesrouze, Y.*, et al.* The surprising features of the TEAD4-Vgll1 protein-protein
904 interaction. *Chembiochem* **15**, 537-542 (2014).
905 32. Buzdar, A.U. Fulvestrant--a novel estrogen receptor antagonist for the treatment of
906 advanced breast cancer. *Drugs Today (Barc)* **44**, 679-692 (2008).
907 33. Howell, A.*, et al.* Comparison of fulvestrant versus tamoxifen for the treatment of
908 advanced breast cancer in postmenopausal women previously untreated with
909 endocrine therapy: a multinational, double-blind, randomized trial. *J Clin Oncol* **22**,
910 1605-1613 (2004).
911 34. Turner, N.C.*, et al.* Overall Survival with Palbociclib and Fulvestrant in Advanced
912 Breast Cancer. *N Engl J Med* **379**, 1926-1936 (2018).
913 35. Jeselsohn, R., Buchwalter, G., De Angelis, C., Brown, M. & Schiff, R. ESR1 mutations-a
914 mechanism for acquired endocrine resistance in breast cancer. *Nat Rev Clin Oncol*
915 **12**, 573-583 (2015).
916 36. Katzenellenbogen, J.A., Mayne, C.G., Katzenellenbogen, B.S., Greene, G.L. &
917 Chandarlapaty, S. Structural underpinnings of oestrogen receptor mutations in
918 endocrine therapy resistance. *Nat Rev Cancer* **18**, 377-388 (2018).
919 37. Pejerrey, S.M.*, et al.* The Impact of ESR1 Mutations on the Treatment of Metastatic
920 Breast Cancer. *Horm Cancer* **9**, 215-228 (2018).
921 38. Heintzman, N.D.*, et al.* Distinct and predictive chromatin signatures of transcriptional
922 promoters and enhancers in the human genome. *Nat Genet* **39**, 311-318 (2007).
923 39. Ernst, J.*, et al.* Mapping and analysis of chromatin state dynamics in nine human cell
924 types. *Nature* **473**, 43-49 (2011).
925 40. Bardia, A.*, et al.* EMERALD: Phase III trial of elacestrant (RAD1901) vs endocrine
926 therapy for previously treated ER+ advanced breast cancer. *Future Oncol* **15**, 3209-
927 3218 (2019).
928 41. Cheung, K.W.K.*, et al.* GDC-0810 Pharmacokinetics and Transporter-Mediated Drug
929 Interaction Evaluation with an Endogenous Biomarker in the First-in-Human, Dose
930 Escalation Study. *Drug Metab Dispos* **47**, 966-973 (2019).
931 42. Hamilton, E.P.*, et al.* A First-in-Human Study of the New Oral Selective Estrogen
932 Receptor Degrader AZD9496 for ER(+)/HER2(-) Advanced Breast Cancer. *Clin Cancer*
933 *Res* **24**, 3510-3518 (2018).
934 43. Zanconato, F.*, et al.* Genome-wide association between YAP/TAZ/TEAD and AP-1 at
935 enhancers drives oncogenic growth. *Nat Cell Biol* **17**, 1218-1227 (2015).
936 44. Zhu, C.*, et al.* A Non-canonical Role of YAP/TEAD Is Required for Activation of
937 Estrogen-Regulated Enhancers in Breast Cancer. *Mol Cell* **75**, 791-806 e798 (2019).
938 45. Liu, X.*, et al.* Tead and AP1 Coordinate Transcription and Motility. *Cell Rep* **14**, 1169-
939 1180 (2016).

940 46. Xiao, J.H., Davidson, I., Matthes, H., Garnier, J.M. & Chambon, P. Cloning, expression,
941 and transcriptional properties of the human enhancer factor TEF-1. *Cell* **65**, 551-568
942 (1991).
943 47. Galli, G.G*., et al.* YAP Drives Growth by Controlling Transcriptional Pause Release
944 from Dynamic Enhancers. *Mol Cell* **60**, 328-337 (2015).
945 48. Zhao, B*., et al.* TEAD mediates YAP-dependent gene induction and growth control.
946 *Genes Dev* **22**, 1962-1971 (2008).
947 49. Konermann, S*., et al.* Genome-scale transcriptional activation by an engineered
948 CRISPR-Cas9 complex. *Nature* **517**, 583-588 (2015).
949 50. Shaw, L.E., Sadler, A.J., Pugazhendhi, D. & Darbre, P.D. Changes in oestrogen
950 receptor-alpha and -beta during progression to acquired resistance to tamoxifen and
951 fulvestrant (Faslodex, ICI 182,780) in MCF7 human breast cancer cells. *J Steroid*
952 *Biochem Mol Biol* **99**, 19-32 (2006).
953 51. Larsen, S.S., Egeblad, M., Jaattela, M. & Lykkesfeldt, A.E. Acquired antiestrogen
954 resistance in MCF-7 human breast cancer sublines is not accomplished by altered
955 expression of receptors in the ErbB-family. *Breast Cancer Res Treat* **58**, 41-56 (1999).
956 52. Kirkegaard, T*., et al.* T47D breast cancer cells switch from ER/HER to HER/c-Src
957 signaling upon acquiring resistance to the antiestrogen fulvestrant. *Cancer Lett* **344**,
958 90-100 (2014).
959 53. Hanker, A.B., Sudhan, D.R. & Arteaga, C.L. Overcoming Endocrine Resistance in
960 Breast Cancer. *Cancer Cell* **37**, 496-513 (2020).
961 54. Razavi, P*., et al.* The Genomic Landscape of Endocrine-Resistant Advanced Breast
962 Cancers. *Cancer Cell* **34**, 427-438 e426 (2018).
963 55. Engelman, J.A. Targeting PI3K signalling in cancer: opportunities, challenges and
964 limitations. *Nat Rev Cancer* **9**, 550-562 (2009).
965 56. Downward, J. Targeting RAS signalling pathways in cancer therapy. *Nat Rev Cancer* **3**,
966 11-22 (2003).
967 57. Stanley, A., Ashrafi, G.H., Seddon, A.M. & Modjtahedi, H. Synergistic effects of
968 various Her inhibitors in combination with IGF-1R, C-MET and Src targeting agents in
969 breast cancer cell lines. *Sci Rep* **7**, 3964 (2017).
970 58. Dowell, J., Minna, J.D. & Kirkpatrick, P. Erlotinib hydrochloride. *Nat Rev Drug Discov*
971 **4**, 13-14 (2005).
972 59. Ciriello, G*., et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast
973 Cancer. *Cell* **163**, 506-519 (2015).
974 60. Curtis, C*., et al.* The genomic and transcriptomic architecture of 2,000 breast
975 tumours reveals novel subgroups. *Nature* **486**, 346-352 (2012).
976 61. Liu-Chittenden, Y*., et al.* Genetic and pharmacological disruption of the TEAD-YAP
977 complex suppresses the oncogenic activity of YAP. *Genes Dev* **26**, 1300-1305 (2012).
978 62. Gibault, F., Sturbaut, M., Bailly, F., Melnyk, P. & Cotelle, P. Targeting Transcriptional
979 Enhanced Associate Domains (TEADs). *J Med Chem* **61**, 5057-5072 (2018).
980 63. Kim, B.K*., et al.* PI3K/AKT/beta-Catenin Signaling Regulates Vestigial-Like 1 Which
981 Predicts Poor Prognosis and Enhances Malignant Phenotype in Gastric Cancer.
982 *Cancers (Basel)* **11**(2019).
983 64. Liu, Y*., et al.* Increased TEAD4 expression and nuclear localization in colorectal cancer
984 promote epithelial-mesenchymal transition and metastasis in a YAP-independent
985 manner. *Oncogene* **35**, 2789-2800 (2016).

986 65. Lin, K.C., *et al.* Regulation of Hippo pathway transcription factor TEAD by p38 MAPK-
987 induced cytoplasmic translocation. *Nat Cell Biol* **19**, 996-1002 (2017).
988 66. Song, S., *et al.* The Hippo Coactivator YAP1 Mediates EGFR Overexpression and
989 Confers Chemoresistance in Esophageal Cancer. *Clin Cancer Res* **21**, 2580-2590
990 (2015).
991 67. Lee, D.H., *et al.* LATS-YAP/TAZ controls lineage specification by regulating TGFbeta
992 signaling and Hnf4alpha expression during liver development. *Nat Commun* **7**, 11961
993 (2016).
994 68. Patten, D.K., *et al.* Enhancer mapping uncovers phenotypic heterogeneity and
995 evolution in patients with luminal breast cancer. *Nat Med* **24**, 1469-1480 (2018).
996 69. Nayar, U., *et al.* Acquired HER2 mutations in ER(+) metastatic breast cancer confer
997 resistance to estrogen receptor-directed therapies. *Nat Genet* **51**, 207-216 (2019).
998 70. Bushweller, J.H. Targeting transcription factors in cancer - from undruggable to
999 reality. *Nat Rev Cancer* **19**, 611-624 (2019).
1000 71. Elisi, G.M., *et al.* Repurposing of Drugs Targeting YAP-TEAD Functions. *Cancers*
1001 *(Basel)* **10**(2018).
1002 72. Bradley, S.D., *et al.* Vestigial-like 1 is a shared targetable cancer-placenta antigen
1003 expressed by pancreatic and basal-like breast cancers. *Nat Commun* **11**, 5332 (2020).
1004 73. Martin, L.A., *et al.* Discovery of naturally occurring ESR1 mutations in breast cancer
1005 cell lines modelling endocrine resistance. *Nat Commun* **8**, 1865 (2017).
1006 74. Pancholi, S., *et al.* Combination of mTORC1/2 inhibitor vistusertib plus fulvestrant in
1007 vitro and in vivo targets oestrogen receptor-positive endocrine-resistant breast
1008 cancer. *Breast Cancer Res* **21**, 135 (2019).
1009 75. Schmidt, D., *et al.* ChIP-seq: using high-throughput sequencing to discover protein-
1010 DNA interactions. *Methods* **48**, 240-248 (2009).
1011 76. Kim, D., *et al.* TopHat2: accurate alignment of transcriptomes in the presence of
1012 insertions, deletions and gene fusions. *Genome Biol* **14**, R36 (2013).
1013 77. Li, H., *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
1014 2078-2079 (2009).
1015 78. Anders, S., Pyl, P.T. & Huber, W. HTSeq--a Python framework to work with high-
1016 throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).
1017 79. Anders, S. & Huber, W. Differential expression analysis for sequence count data.
1018 *Genome Biol* **11**, R106 (2010).
1019 80. Lawrence, M., Gentleman, R. & Carey, V. rtracklayer: an R package for interfacing
1020 with genome browsers. *Bioinformatics* **25**, 1841-1842 (2009).
1021 81. Raudvere, U., *et al.* g:Profiler: a web server for functional enrichment analysis and
1022 conversions of gene lists (2019 update). *Nucleic Acids Res* **47**, W191-W198 (2019).
1023 82. Skehan, P., *et al.* New colorimetric cytotoxicity assay for anticancer-drug screening. *J*
1024 *Natl Cancer Inst* **82**, 1107-1112 (1990).
1025 83. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient
1026 alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25
1027 (2009).
1028 84. Zhang, Y., *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137
1029 (2008).
1030 85. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic
1031 features. *Bioinformatics* **26**, 841-842 (2010).

86. Karolchik, D.*, et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res* **32**, D493-496 (2004).
87. Stempor, P. & Ahringer, J. SeqPlots - Interactive software for exploratory data analyses, pattern discovery and visualization in genomics. *Wellcome Open Res* **1**, 14 (2016).
88. Heinz, S.*, et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589 (2010).
89. McLean, C.Y.*, et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**, 495-501 (2010).
90. Supek, F., Bosnjak, M., Skunca, N. & Smuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, e21800 (2011).

1068

1069 **Figure Legends**

1070

1071 **Fig. 1 VGLL1 expression is induced in fulvestrant-resistant breast cancer. a**,
1072 Ratio of H3K27ac in FULVR/MCF7 cells at gene promoters in a window of ± 1.5 kb
1073 centred on the transcriptional start site. H3K27ac is highly enriched at the promoters
1074 of the *VGLL1* and *VGLL3* genes in FULVR cells relative to MCF7 cells, whereas ER
1075 target genes (including *PGR*, *GREB1* and *TFF1*) show H3K27ac loss in FULVR cells.
1076 **b**, Genome browser view of H3K27ac ChIP-seq signal at the promoter region of the
1077 VGLL1 and VGLL3 genes in MCF7-FULVR and MCF7 cells. **c**, RT-qPCR showing
1078 *VGLL1* and *VGLL3* upregulation in FULVR cells and loss of ESR1 and PGR
1079 expression. Data are presented as mean + s.e.m of n=3 experiments. *P<0.05
1080 (Student's t-test, two-tailed). **d**, Immunoblotting of VGLL1, VGLL3 and proteins in the
1081 YAP/TEAD pathway in MCF7 and FULVR cells **e**, RNA-FISH showing increased
1082 *VGLL1* mRNA levels in breast cancer patients after fulvestrant treatment and relapse
1083 compared to matched pre-treatment samples. Representative images show the
1084 results for matched samples from one patient. Nuclei are labelled with DAPI, and
1085 individual *VGLL1* mRNA molecules detected with Cy5-labelled probes. P=0.0177
1086 (one-tailed Wilcoxon signed rank test). Scale bar, 10μm. **f**, RT-qPCR for *ESR1* and
1087 *VGLL1* after treatment of ER+ breast cancer cell lines with fulvestrant (100nM, 24 h).
1088 **g**, RT-qPCR for *ESR1* and *VGLL1* after ESR1 siRNA. Data in **f** and **g** are presented
1089 as in **c**. *P<0.05 (Student's t-test, two-tailed).

1090

1091 **Fig. 2 VGLL1 is recruited to TEAD4 binding regions in fulvestrant-resistant**
1092 **MCF7 cells. a**, Overlap between TEAD4 ChIP-seq peaks in MCF7 and FULVR cells
1093 showing increased TEAD4 binding in FULVR cells. Top-most enriched motifs in each
1094 cell line are shown. More detailed motif analysis is shown in Supplementary Fig. 3b.
1095 **b**, Correlation between VGLL1 and TEAD4 occupancy at TEAD4 peaks in FULVR
1096 cells. *P-value* was calculated using the Spearman test; r_s, Spearman's correlation
1097 coefficient. **c**, Heatmap showing TEAD4 and VGLL1 binding at shared and unique
1098 TEAD4 peaks in each cell line, in a window of ± 2kb centred on the peak centre. **d**,
1099 VGLL1 binding is enriched at TEAD4 peaks in FULVR cells. Binding enrichment was
1100 calculated as VGLL1-TEAD4 co-binding over the mean expected value after

1101 generating random permutations of the TEAD4 peaks (Chi-squared test p-value
1102 <0.0001). **e**, Average normalized ChIP-seq signal of TEAD4, VGLL1 and H3K27ac at
1103 TEAD4 peaks in FULVR cells, in a window of ± 3 kb from the peak centre. **f**, Average
1104 ChIP-seq signal of H3K27ac on VGLL1 peaks divided in quantiles based on the peak
1105 coverage in FULVR cells. **g**, Genome browser view of VGLL1, TEAD4 and H3K27ac
1106 ChIP-seq signal, together with the RNA-seq signal in MCF7 and FULVR cells at
1107 YAP/TEAD target genes described in other cell lines.
1108
1109 **Fig. 3 VGLL1 mediates resistance to fulvestrant. a,** VGLL1 is required for the
1110 growth of MCF7-FULVR cells. Cells were transfected with two independent siRNAs
1111 for VGLL1. Growth was determined with the SRB assay five days after transfection.
1112 Data are mean ± s.e.m of n=6 independent wells. One representative experiment is
1113 shown; similar results were obtained in two additional independent experiments. *
1114 *P*<0.05 (Mann-Whitney test, two tailed). Also shown are expression levels of VGLL1
1115 following siVGLL1 transfection (*n*=3, **p*<0.05 (Student's t-test, two-tailed)). **b**, The
1116 synergistic activation mediators (SAM) technique using a catalytically dead Cas9
1117 (dCas9) for transcriptional activation of endogenous genes[49], with a sgRNA for the
1118 region -156 to -134 of the VGLL1 gene, was used to activate VGLL1 expression in
1119 MCF7 cells. **c,** RT-qPCR of VGLL1 in MCF7 cells overexpressing VGLL1 (MCF7
1120 ActCas9-VGLL1) generated as described in **b** and control cells (MCF7 ActCas9-
1121 Vector). **d,** Transcriptional activation of endogenous VGLL1 in MCF7 cells promotes
1122 fulvestrant resistance. Growth of MCF7 ActCas9-VGLL1 cells and MCF7 ActCas9-
1123 Vector cells was monitored using live cell imaging in the presence of 100 nM
1124 fulvestrant. Shown are the results of one of three independent experiments. Data are
1125 mean ± s.e.m of n=9 representative images from one experiment. **e,** Immunoblotting
1126 for VGLL1, ER and PR in the indicated cell lines. VGLL1-FULVR-1 and VGLL1-
1127 FULVR-2 correspond to two independent fulvestrant resistant cell lines derived from
1128 the parental MCF7 ActCas9-VGLL1 cell line after continuous culturing in the presence
1129 of 100 nM or 1μM fulvestrant, respectively. **f,** Normalized RNA-seq read counts of
1130 *VGLL1*, *ESR1* and *PGR* in the indicated cell lines showing strong increase in VGLL1
1131 expression in the resistant cells, with a concurrent downregulation of ER and PGR. **g,**
1132 MCF7 ActCas9-VGLL1-FULVR cells were transfected with siVGLL1 and growth
1133 assessed as in **a.** Data are mean ± s.e.m of n=6. One representative experiment is

1134 shown; similar results were obtained in two additional independent experiments. *
1135 *P*<0.05 (Mann-Whitney test, two tailed). RT-qPCR for VGLL1 is also shown (*n*=3,
1136 **p*<0.05 (Student's t-test, two-tailed)). **h,** Growth of the indicated cell lines treated with
1137 increasing concentrations of fulvestrant to a maximum of 1 μM, for five days. Cell
1138 growth was estimated using the SRB assay and is shown as percentage relative to
1139 vehicle (n=6). Half maximum inhibitory concentration (IC_{50}) is also shown for each line.
1140

1141 **Fig. 4 Genes upregulated in fulvestrant-resistant breast cancer cells depend on**
1142 **VGLL1 transcriptional activity. a,** Genes predicted as VGLL1 targets in FULVR cells
1143 are more highly expressed in FULVR cells than MCF7 cells. Genes were segregated
1144 into those with no VGLL1 peaks, and those with 2 or more and 5 or more VGLL1
1145 peaks. The y-axis shows the log2 fold change in gene expression determined from
1146 RNA-seq in FULVR cells versus the parental MCF7 cells. Data are presented as box-
1147 and-whiskers plots (whiskers extend from the 5th to the 95th percentile; the box
1148 extends from the 25th to the 75th percentile; the line within the box represents the
1149 median). ****P<0.0001 (Mann-Whitney test, two tailed). **b,** Genes activated by VGLL1
1150 (n = 762) are over-expressed in FULVR cells relative to MCF7 cells, compared to not-
1151 VGLL1 targets (n = 8,932). VGLL1-activated genes and not-VGLL1 targets were
1152 determined by RNA-seq in FULVR cells transfected with VGLL1 siRNAs. The VGLL1-
1153 activated genes were defined as genes downregulated by VGLL1 siRNAs (see
1154 Methods, and Supplementary Fig. 4b). Data are presented as in **a**. P<0.0001 (Mann-
1155 Whitney test, two tailed). **c,** Normalized average H3K27ac signal on the promoters of
1156 VGLL1 activated genes in FULVR cells and MCF7 cells. **d,** VGLL1-activated genes (n
1157 = 762) display higher expression levels in FULVR cells than not-VGLL1 targets (n =
1158 8,932). The y-axis shows normalized gene expression values from RNA-seq.
1159 P<0.0001 (Mann-Whitney test, two tailed). **e,** GO molecular function sets enriched in
1160 VGLL1-activated genes. **f,** RNA-seq data represented as a volcano plot for the
1161 comparison between FULVR vs MCF7 cells showing the VGLL1-activated genes as
1162 red dots. Four VGLL1 targets highly upregulated in FULVR cells and associated with
1163 growth functional categories enriched in VGLL1-activated genes are highlighted. **g,**
1164 RT-qPCR in the indicated MCF7-ActCas9 cells represented as a heatmap. Gene
1165 expression values were normalized to *GAPDH* expression and are shown as log2 fold
1166 change relative to MCF7 ActCas9-Vector cells.
1167

1168 **Fig. 5 VGLL1 promotes EGFR expression in FULVR cells. a,** Genome browser
1169 view of VGLL1 and TEAD4 ChIP-seq signal at the EGFR gene and EGFR enhancer
1170 (highlighted). **b,** ChIP-qPCR for TEAD4, VGLL1 and H3K27ac in MCF7 ActCas9-
1171 Vector, MCF7 ActCas9-VGLL1 and MCF7 ActCas9-VGLL1-FULVR cells showing
1172 VGLL1 and TEAD4 binding at the EGFR enhancer together with EGFR enhancer
1173 activation exclusively in FULVR cells. The CTGF -8.3kb region was used as a negative
1174 control for VGLL1/TEAD4 binding. **c,** RT-qPCR for EGFR in the indicated MCF7-
1175 ActCas9 cells (n=3, *p<0.05 (Student's t-test, two-tailed)). **d,** Immunoblotting for
1176 EGFR and downstream EGFR signalling proteins showing EGFR up-regulation and
1177 activation of the EGFR pathway in MCF7 ActCas9-VGLL1-FULVR cells **e,** RT-qPCR
1178 for dCAS9, *VGLL1* and *EGFR* in MCF7 ActCas9-VGLL1-FULVR cells transfected with
1179 dCAS9 siRNAs showing reduction in *VGLL1* and *EGFR* expression. **f,** Growth of the
1180 indicated MCF7 ActCas9 cell lines treated with increasing concentrations of the EGFR
1181 inhibitor erlotinib (0.05 μM to 12.5 μM) for 5 days. Growth is shown as percentage
1182 relative to vehicle treatment for one representative experiment (n=6 replicates) of two
1183 independent experiments. **g,** Western blot of EGFR and the downstream EGFR
1184 signalling proteins in the indicated MCF7 ActCas9 cell lines treated with erlotinib at
1185 the indicated concentrations (24 h). **h,** Correlation between *EGFR* and *VGLL1* gene
1186 expression in ER + breast cancers from TCGA (n = 609 samples). **i**, The mechanism
1187 by which VGLL1/TEAD4 activate the expression of EGFR in fulvestrant-resistant
1188 breast cancer cells.
1189
1190 **Fig. 6 Verteporfin inhibits VGLL1 transcriptional activity. a,** VP impairs the growth
1191 of FULVR cells. Growth in the indicated FULVR cells treated with increasing
1192 concentrations of VP is shown as percentage of growth relative to vehicle. **b,** ChIP-
1193 qPCR for VGLL1 and TEAD4 in MCF7-FULVR cells showing reduced VGLL1 binding
1194 at the target genes in the presence of VP (2μM, 24 h). The y axis shows DNA
1195 enrichment calculated as the percentage of input. **c,** RNA-seq data represented as a
1196 volcano plot showing upregulated genes (red, adjusted P < 0.01, log2 (fold change) >
1197 1) and downregulated genes (blue, adjusted P < 0.01, log2 (fold change) < -1) in
1198 MCF7-FULVR cells treated with VP, (2μM, 24 h) compared to vehicle (DMSO), n=4
1199 biological replicates. VGLL1 target genes such as IGFBP5, EGFR and TGFB2 are
1200 strongly downregulated by VP. **d,** RT-qPCR in MCF7-FULVR cells treated with VP

1201 (2μM, 24 h) showing that VP selectively downregulates the expression VGLL1 targets,
1202 while the expression of not VGLL1 targets (APEX1 and URI1) is not affected by VP.
1203 **e,** Genes downregulated by VP are highly enriched in VGLL1 peaks. The y axis shows
1204 the number of VGLL1 peaks over the expected value after generating random
1205 permutations of the VGLL1 peaks showing that the top most highly downregulated
1206 genes after VP treatment in FULVR cells (*n* = 631) have significantly higher number
1207 of VGLL1 peaks compared to the bottom genes not differentially expressed by VP (*n*
1208 = 631). Data are presented as box-and-whiskers plots (whiskers extend from the 5th
1209 to the 95th percentile; the box extends from the 25th to the 75th percentile; the line
1210 within the box represents the median). **** *P*<0.0001 (Mann-Whitney test, two tailed).
1211 **f,** Genome browser view of VGLL1 and TEAD4 normalized ChIP-seq signal and
1212 normalized RNA-seq signal in FULVR cells showing representative examples of
1213 VGLL1 activated genes (direct VGLL1 targets) and not-VGLL1 targets. VGLL1
1214 activated genes are downregulated by VGLL1 siRNA and are also selectively
1215 downregulated by VP (2μM, 24 h) compared to not-VGLL1 targets. **g,** VGLL1 activated
1216 genes (*n* = 762) are preferentially downregulated by VP compared to not-VGLL1
1217 targets (*n* = 8,932). The y axis shows fold change in gene expression from RNA-seq
1218 between VP versus DMSO treatment in MCF7-FULVR cells. Data are presented as in
1219 **e**. ****P<0.0001 (Mann-Whitney test, two tailed). **h,** Breast cancer patients with higher
1220 *VGLL1* expression display increased levels of expression of the VGLL1 activated
1221 genes compared to patients with lower VGLL1 expression. Patients from the
1222 METABRIC breast cancer dataset were stratified according to high (top quantile,
1223 n=495) or low (bottom quantile, n=495) *VGLL1* expression levels. Data are presented
1224 as in **e**. ****P<0.0001 (Mann-Whitney test, two tailed). **i,** Kaplan–Meier plot
1225 representing the percentage of metastasis-free survival in patients with breast cancer
1226 treated with endocrine therapies showing that patients with higher expression of the
1227 VGLL1 activated genes signature display lower survival rates (log-rank Mantel-Cox
1228 test).
1229

1230 **Supplementary Figure Legends**

1231 **Supplementary Figure 1.**

1232 **a,** RNA-seq data represented as a volcano plot showing upregulated genes (red,
1233 adjusted P < 0.01, log2 (fold change) > 2.5) and downregulated genes (blue, adjusted

1234 *P* < 0.01, log2 (fold change) < -2.5) in FULVR cells compared to MCF7 cells. VGLL1
1235 and VGLL3 expression is strongly increased in FULVR cells. Classical ER target
1236 genes (TFF1, PGR and GREB1) are shown as controls for ER signaling
1237 downregulation in FULVR cells. **b,** Normalized gene expression of TEAD coactivators
1238 determined by RNA-seq. VGLL1 and VGLL3 are the only TEAD coactivators that show
1239 the strongest upregulation in FULVR cells, as well as being up-regulated in tamoxifen-
1240 resistant MCF7 (TAMR) cells. Data are presented as box-and-whiskers plots showing
1241 the median of two independent biological RNA-seq replicates (whiskers extend to
1242 minimum and maximum values). LTED, long term estradiol deprived. RNA-seq data
1243 were obtained from ref[16]. **c,** RT-qPCR for genes in the YAP/TEAD pathway in MCF7
1244 and FULVR cells. Data are presented as mean + s.e.m of three independent
1245 experiments. **P*<0.05 (Student's t-test, two-tailed)). **d,** Heatmap showing the
1246 expression of genes in the YAP/TEAD pathway in FULVR cell lines derived from the
1247 ER positive breast cancer cell lines MCF7, T47D or ZR-75-1. Upregulation of VGLL1
1248 expression is common to all the FULVR cell lines. The expression values are
1249 represented as log2 (fold change) relative to the corresponding parental cell lines and
1250 normalized to GAPDH. **e,** Immunoblotting showing VGLL1 expression in different
1251 FULVR cell lines, accompanied by reductions in ER and PR levels. ER, estrogen
1252 receptor; PR-A, progesterone receptor isoform A; PR-B, progesterone receptor
1253 isoform B. **f,** Immunohistochemistry of VGLL1 performed in sections from matched
1254 breast cancer samples prior to fulvestrant treatment and after fulvestrant treatment
1255 and relapse. VGLL1 staining is significantly higher after relapse on fulvestrant
1256 treatment, *p*=0.0098 (one-tailed Wilcoxon signed rank test). Representative images of
1257 matched breast cancer tissue pre and post fulvestrant, show increased VGLL1
1258 staining after fulvestrant treatment and relapse. Scale bar, 100μm.
1259
1260 **Supplementary Figure 2.**
1261 **a,** Immunoblotting of VGLL1, ER and PR after treatment with vehicle or fulvestrant at
1262 the indicated concentrations (48 h) in MCF7 cells. Increased levels of VGLL1 are
1263 observed after ER inhibition by fulvestrant. **b,** RT-qPCR for genes in the YAP/TEAD
1264 pathway in MCF7 cells after treatment with vehicle or fulvestrant (100nM, 24 h). Data
1265 are presented as mean + s.e.m of three independent experiments. **P*<0.05 (Student's
1266 t-test, two-tailed)). **c,** Heatmap showing the expression of genes in the YAP/TEAD
1267 pathway in the ER positive breast cancer cell lines T47D, ZR-75-1 and BT474 after

1268 treatment with vehicle or fulvestrant (100nM, 24 h). The expression values are
1269 represented as log2 (fold change) relative to the vehicle control and normalized to
1270 GAPDH. **d,** Heatmap showing the expression of genes in the YAP/TEAD pathway
1271 after *ESR1* siRNA in the ER positive breast cancer cell lines T47D, ZR-75-1 and
1272 BT474. A strong upregulation of *VGLL1* after *ESR1* downregulation is common to all
1273 cell lines. The expression values are represented as log2 (fold change) relative to the
1274 control siRNA (siControl) and normalized to *GAPDH*. **e,** Western blot of VGLL1 and
1275 ER showing VGLL1 upregulation following ER downregulation by ESR1 siRNA in
1276 MCF7 cells. **f**, RT-qPCR for genes in the YAP/TEAD pathway in MCF7 cells
1277 transfected with ESR1 siRNA. Data are presented as mean + s.e.m of two
1278 independent experiments. *P<0.05 (Student's t-test, two-tailed)). **g,** RT-qPCR for
1279 *VGLL1* and the ER target genes *PGR* and *GREB1* in MCF7 cells showing
1280 downregulation of ER transcriptional activity and *VGLL1* upregulation after treatment
1281 with different SERDs (GDC-0810, AZD9496 and RAD1901, 100nM, 48 h). Data are
1282 presented as mean + s.e.m of two independent experiments.
1283
1284 **Supplementary Figure 3.**
1285 **a,** Correlation between two biological replicates of TEAD4 ChIP-seq generated in
1286 MCF7 cells and FULVR cells. *P-value* was calculated using Spearman test, r_s,
1287 Spearman's correlation coefficient. **b,** *De novo* motif analysis of TEAD4 peaks in
1288 MCF7 cells and FULVR cells. **c,** Correlation between three biological replicates of
1289 VGLL1 ChIP-seq in FULVR cells. *P-value* was calculated using Spearman test, r_s,
1290 Spearman's correlation coefficient. **d,** *De novo* motif analysis of VGLL1 peaks in
1291 FULVR cells. **e,** Fraction of VGLL1 and TEAD4 peaks present at the indicated genomic
1292 features showing that the majority of VGLL1 and TEAD4 peaks are present at intronic
1293 and intergenic regions. **f,** Normalized H3K27ac signal at TEAD4/VGLL1 co-bound
1294 regions from FULVR cells showing lower H3K27ac signal at these sites in MCF7 cells
1295 compared to FULVR cells. **g,** ChIP-qPCR showing preferential VGLL1 binding
1296 compared to YAP/TAZ binding on promoters and enhancers of known YAP/TEAD
1297 target genes from other cell lines[47], in FULVR cells but not in MCF7 cells. Negative
1298 control regions are indicated which were previously reported as non-YAP/TAZ targets
1299 in other cell types[47]. Binding enrichment was calculated as percentage of input.
1300 *P<0.05 (Student's t-test, two-tailed).
1301

1302 **Supplementary Figure 4.**

1303 **a,** Functional annotation of the VGLL1 ChIP-seq peaks in FULVR cells was performed
1304 using GREAT[89]. REVIGO[90] was used to visualize the top nine most significant
1305 categories (see Supplementary Table 2 for the complete list). **b,** Left: Growth of T47D-
1306 FULVR cells and MCF7-FULVR(2) determined with the SRB assay five days after
1307 VGLL1 siRNA transfection. Growth was calculated relative to the siRNA control
1308 (siControl). Data are mean + s.e.m of n=6. *P < 0.05 (Mann-Whitney test, two tailed).
1309 Right: VGLL1 gene expression after VGLL1 siRNA transfection in the indicated cell
1310 lines (n=3, *p<0.05 (Student's t-test, two-tailed)). **c,** Western blot of VGLL1 in MCF7
1311 cells with stable expression of dCAS9 fused to transcriptional activators[49] targeted to
1312 the VGLL1 promoter (MCF7 ActCas9-VGLL1) and control cells (MCF7 ActCas9-
1313 Vector). **d,** Growth of MCF7 ActCas9-VGLL1 cells and control MCF7 ActCas9-Vector
1314 cells monitored using live cell imaging over a period of 100 days in the presence of
1315 fulvestrant (100nM, similar results were obtained with 1μM). **e,** RT-qPCR of *VGLL1*,
1316 *ESR1* and *PGR* in the indicated MCF7 ActCas9 cell lines. **f,** Comparison of genes
1317 differentially expressed in MCF7 ActCas9-VGLL1-FULVR-1 cells across
1318 independently generated fulvestrant resistant breast cancer cell line models show
1319 similar transcriptional perturbations across the different models. The y axis shows the
1320 fold change in gene expression from RNA-seq data between each FULVR cell line
1321 and the corresponding parental cells for genes upregulated (adjusted p<0.05, fold
1322 change > 2) and downregulated (adjusted p<0.05, fold change < 0.5) in MCF7
1323 ActCas9-VGLL1-FULVR-1 cells compared to the parental MCF7 ActCas9-VGLL1
1324 cells. Data are presented as box-and-whiskers plots (whiskers extend from the 5th to
1325 the 95th percentile; the box extends from the 25th to the 75th percentile; the line within
1326 the box represents the median). ****P<0.0001 (Mann-Whitney test, two tailed). **g,**
1327 RNA-seq data represented as a volcano plot for the comparison between MCF7
1328 ActCas9-VGLL1 vs MCF7 ActCas9-Vector cells, showing that MCF7 cells
1329 overexpressing VGLL1 which have not been exposed to fulvestrant show almost
1330 identical transcriptional profiles compared to the control cells. **h,** Growth of the
1331 indicated MCF7 ActCas9 cells following five days of treatment with increasing doses
1332 (0.025 nM to 1μM) of three different SERDs (AZD9496, RAD1901, GDC-0810). Data
1333 are plotted as the mean ± s.e.m of n=6.

1334

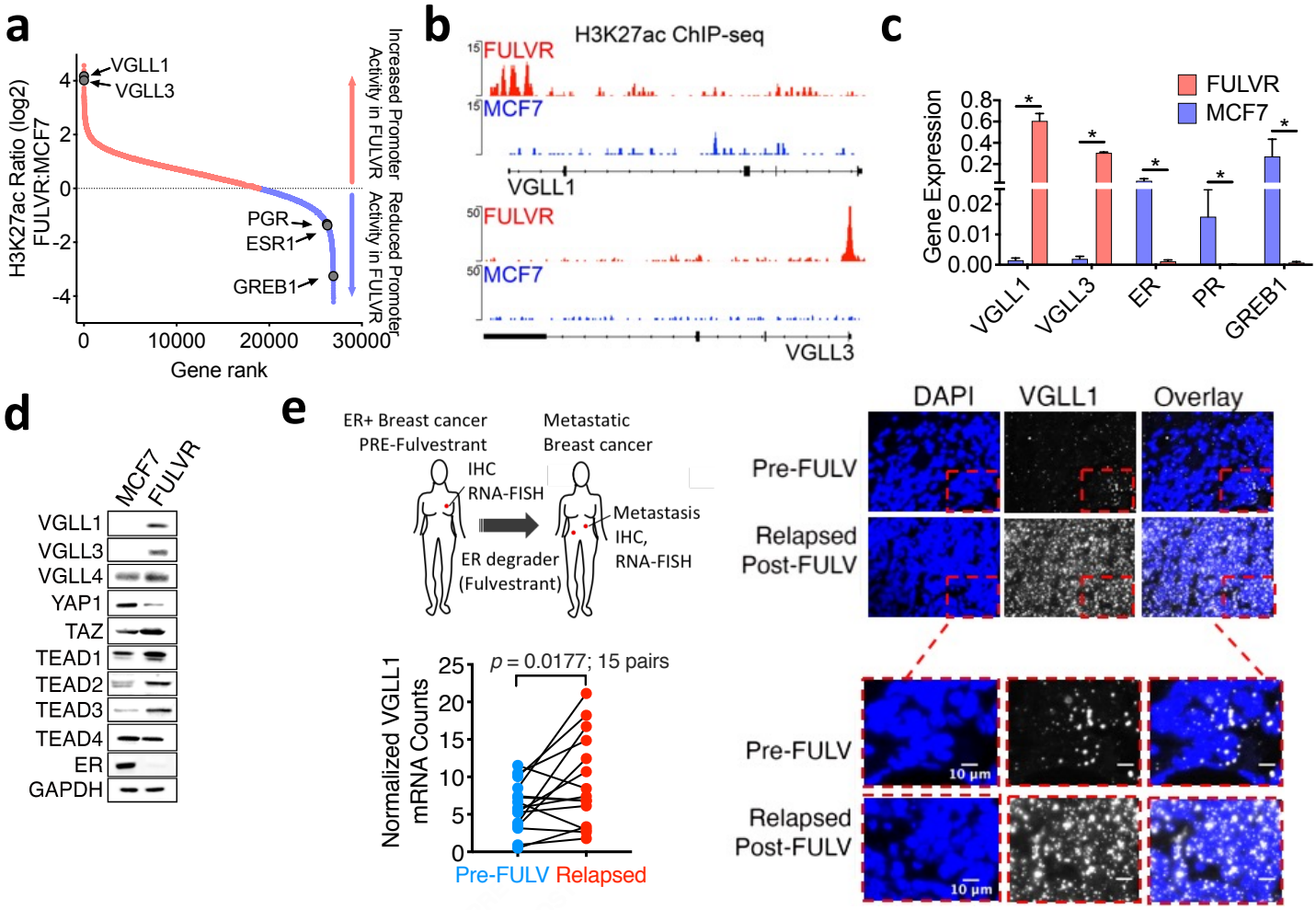
1335 **Supplementary Figure 5.**

1336 **a,** MCF7 ActCas9-VGLL1-FULVR cells display higher levels of expression of predicted
1337 VGLL1 targets from the independent FULVR cells. Genes with 2 or more and 5 or
1338 more VGLL1 associated peaks were considered as VGLL1 targets for this analysis
1339 together with genes without VGLL1 associated peaks. The y axis shows the fold
1340 change in gene expression determined by RNA-seq between MCF7 ActCas9-VGLL1-
1341 FULVR-1 and the parental MCF7 ActCas9-VGLL1 cells. Data are presented as box-
1342 and-whiskers plots (whiskers extend from the 5th to the 95th percentile; the box
1343 extends from the 25th to the 75th percentile; the line within the box represents the
1344 median). ****P<0.0001 (Mann-Whitney test, two tailed); **P = 0.0056 (Mann-Whitney
1345 test, two tailed). **b,** VGLL1 activated genes are defined as genes downregulated by
1346 VGLL1 siRNAs. Fold change in gene expression between FULVR cells transfected
1347 with VGLL1 siRNAs versus control siRNA for VGLL1 activated genes (n = 762) and
1348 not-VGLL1 targets (n = 8,932). Gene expression was determined by RNA-seq in
1349 FULVR cells transfected with three different VGLL1 siRNAs, n=3 biological replicates
1350 (see Methods). Data are presented as box-and-whiskers plots, as in **a**. ****P<0.0001
1351 (Mann-Whitney test, two tailed). **c,** *VGLL1* gene expression determined by RNA-seq
1352 in FULVR cells transfected with three different VGLL1 siRNAs from the same RNA-
1353 seq experiment as in **b**. The y axis shows normalized read counts from RNA-seq data,
1354 n=3 biological replicates. **d,** VGLL1 binding is enriched at VGLL1 activated genes in
1355 FULVR cells. The y axis shows the number of VGLL1 peaks associated with VGLL1
1356 activated genes and not-VGLL1 targets over the mean expected value after
1357 generating random permutations of the VGLL1 peaks. P<0.0001 (Mann-Whitney test,
1358 two tailed). **e,** TEAD4 binding is enriched at VGLL1 activated genes in FULVR cells.
1359 The y axis shows the number of TEAD4 peaks associated with VGLL1 activated genes
1360 and not-VGLL1 targets over the mean expected value after generating random
1361 permutations of the TEAD4 peaks. P<0.0001 (Mann-Whitney test, two tailed). **f,** Fold
1362 change in H3K27ac signal between FULVR versus MCF7 cells at the promoters of
1363 VGLL1 activated genes (n = 762) and not-VGLL1 targets (n = 8,932). Data are
1364 presented as box-and-whiskers plots, as in **a**. P<0.0001 (Mann-Whitney test, two
1365 tailed). **g,** Increased H3K27ac and TEAD4 recruitment in FULVR cells compared to
1366 MCF7 cells at regulatory elements associated with VGLL1 activated genes. Data are
1367 presented as average normalized ChIP-seq signal of VGLL1, TEAD4 and H3K27ac in
1368 FULVR cells together with TEAD4 and H3K27ac signal in MCF7 cells, in a window of

1369 ± 2.5 kb centred on the VGLL1/TEAD4 peaks associated with VGLL1 activated genes
1370 in FULVR cells. **h,** Heatmap showing a strong upregulation of *VGLL1* gene expression
1371 and upregulation of VGLL1 activated genes, with a concurrent downregulation of ER
1372 transcriptional activity as evidenced by a significant reduction in *PGR* expression, in
1373 MCF7 ActCas9-VGLL1-FULVR cells. The expression values are shown as normalized
1374 read counts from RNA-seq and are represented as log2 fold change relative to MCF7
1375 ActCas9-Vector cells.
1376
1377 **Supplementary Figure 6.**
1378 **a,** Ratio of TEAD4 ChIP-seq signal in FULVR versus MCF7 cells at TEAD4 peaks from
1379 FULVR and MCF7 cells. The position of the VGLL1 promoter and EGFR enhancer are
1380 shown in this plot and are among the top genes with high TEAD4 signal in FULVR
1381 cells and absence of TEAD4 binding in MCF7 cells. VGLL1 is recruited to the TEAD4
1382 binding sites at the VGLL1 promoter and EGFR enhancer in FULVR cells (see also
1383 Fig. 5a). **b,** RT-qPCR for EGFR showing EGFR upregulation in different fulvestrant
1384 resistant cell lines which also upregulate VGLL1 (see Fig. 1c and Supplementary Fig.
1385 1d, e for VGLL1 expression in these cell lines). **c,** RT-qPCR for *EGFR* and *VGLL1* in
1386 MCF7 ActCas9-VGLL1-FULVR cells transfected with VGLL1 siRNA, showing reduced
1387 *EGFR* expression following VGLL1 downregulation. **d,** Western blot of VGLL1, EGFR
1388 and downstream EGFR targets in MCF7-FULVR cells transfected with VGLL1 siRNAs
1389 showing reduced EGFR and downstream EGFR signalling following VGLL1
1390 downregulation. **e,** Western blot of VGLL1, EGFR and downstream EGFR targets in
1391 T47D-FULVR cells transfected with VGLL1 siRNAs showing reduced EGFR and
1392 downstream EGFR signalling following VGLL1 downregulation. **f,** Growth of the
1393 indicated FULVR cells and the corresponding isogenic fulvestrant-sensitive cells
1394 treated with increasing concentrations of the EGFR inhibitor erlotinib for 5 days.
1395 Growth is shown as percentage relative to vehicle (n=6 replicates). **g,** Western blot of
1396 EGFR and downstream EGFR targets in MCF7-FULVR cells and the isogenic
1397 fulvestrant-sensitive MCF7 cells, treated with erlotinib at the indicated concentrations
1398 (24 h). **h,** EGFR ranks first as the most significantly co-expressed gene with VGLL1 in
1399 breast cancer patients from METABRIC and is also the highest ranked protein
1400 significantly correlated with VGLL1 in breast cancer from TCGA. **i,** *EGFR* expression
1401 in ER+ breast cancer patients with high (top 10%) and low (bottom 10%) *VGLL1*

1402 expression from TCGA and METABRIC datasets. ****P<2.2x10^{-16} (Mann-Whitney test,
1403 two tailed).
1404
1405 **Supplementary Figure 7.**
1406 **a,** Growth of MCF7 ActCas9-VGLL1-FULVR cells is reduced following treatment with
1407 increasing concentrations of VP. Data is shown as percentage growth relative to
1408 vehicle (n=6 replicates). **b,** Fold growth relative to day 0 of MCF7-FULVR(2) cells
1409 treated with the indicated concentrations of VP over 9 days. **c,** Heatmap showing the
1410 expression of representative VGLL1 activated genes (direct VGLL1 targets) and not
1411 VGLL1 targets in FULVR cells after downregulation of VGLL1 with two different VGLL1
1412 siRNAs. The expression values were determined by RNA-seq and are represented as
1413 fold change relative to siControl. **d,** RT-qPCR for *VGLL1*, representative VGLL1
1414 activated genes, and not-VGLL1 targets (APEX1 and URI1) in FULVR cells after
1415 downregulation of VGLL1 with two different VGLL1 siRNAs. Expression values are
1416 represented as fold change relative to siControl and normalized to *GAPDH*. **e,** Genes
1417 associated with VGLL1 peaks are preferentially downregulated by VP compared to
1418 genes without VGLL1 associated peaks. Genes with 2 or more and 5 or more VGLL1
1419 associated peaks were considered as VGLL1 targets for this analysis. The y axis
1420 shows the fold change in gene expression determined by RNA-seq between VP
1421 versus DMSO. Data are presented as box-and-whiskers plots (whiskers extend from
1422 the 5th to the 95th percentile; the box extends from the 25th to the 75th percentile; the
1423 line within the box represents the median). ****P<0.0001 (Mann-Whitney test, two
1424 tailed); ** P = 0.0056 (Mann-Whitney test, two tailed). **f,** VGLL1 activated genes
1425 display high sensitivity to VP downregulation. Pie charts showing that the proportion
1426 of VGLL1 activated genes downregulated by VP is larger than the proportion of not
1427 VGLL1 targets downregulated by the same treatment.
1428 **g,** Box-and-whiskers plot showing that genes downregulated by VP in FULVR cells
1429 (*adjusted P* < 0.01, log2(fold change) < -1) display higher expression levels than genes
1430 not inhibited by VP (*adjusted P* > 0.05, log2(fold change) > -1). The y axis shows
1431 normalized gene expression values determined by RNA-seq. Data are presented as
1432 box-and-whiskers plots, as in **e**. ****P<0.0001 (Mann-Whitney test, two tailed). **h,**
1433 Genes downregulated by VP are associated with growth functional categories
1434 enriched in the VGLL1 activated genes. The y axis shows the fold change in gene
1435 expression determined by RNA-seq between VP versus DMSO for genes included in

1436 the following GO terms: Growth factor binding (GO:0019838), extracellular matrix
1437 binding (GO:0050840) and transmembrane receptor protein kinase activity
1438 (GO:0019199), showing that genes included in these categories are significantly
1439 downregulated by VP. + P_{adj} = 2.39×10^{-25}, # P_{adj} = 2.76×10^{-9}, ^ P_{adj} = 1.34×10^{-22}
1440 (Adjusted *p-values* from the pathway enrichment analysis of VP downregulated genes,
1441 see also Supplementary Table 3). **i,** Kaplan–Meier plot representing the percentage
1442 of metastasis-free survival in ER+ breast cancer patients from METABRIC showing
1443 that patients with higher expression of the VGLL1 activated genes signature display
1444 lower survival rates in response to endocrine therapies treatment (log-rank Mantel-
1445 Cox test).


a
VGLL1
VGLL3
PGR
ESR1
GREB1
H3K27ac Ratio (log2) FULVR:MCF7
Gene rank
Increased Promoter Activity in FULVR
Reduced Promoter Activity in FULVR
b
H3K27ac ChIP-seq
FULVR
MCF7
VGLL1
VGLL3
c
FULVR
MCF7
Gene Expression
VGLL1
VGLL3
ER
PR
GREB1
d
MCF7
FULVR
VGLL1
VGLL3
VGLL4
YAP1
TAZ
TEAD1
TEAD2
TEAD3
TEAD4
ER
GAPDH
e
ER+ Breast cancer
PRE-Fulvestrant
Metastatic
Breast cancer
IHC
RNA-FISH
ER degrader
(Fulvestrant)
Metastasis
IHC,
RNA-FISH
p = 0.0177; 15 pairs
Normalized VGLL1 mRNA Counts
Pre-FULV
Relapsed
DAPI
VGLL1
Overlay
Pre-FULV
Relapsed
Post-FULV
10 µm

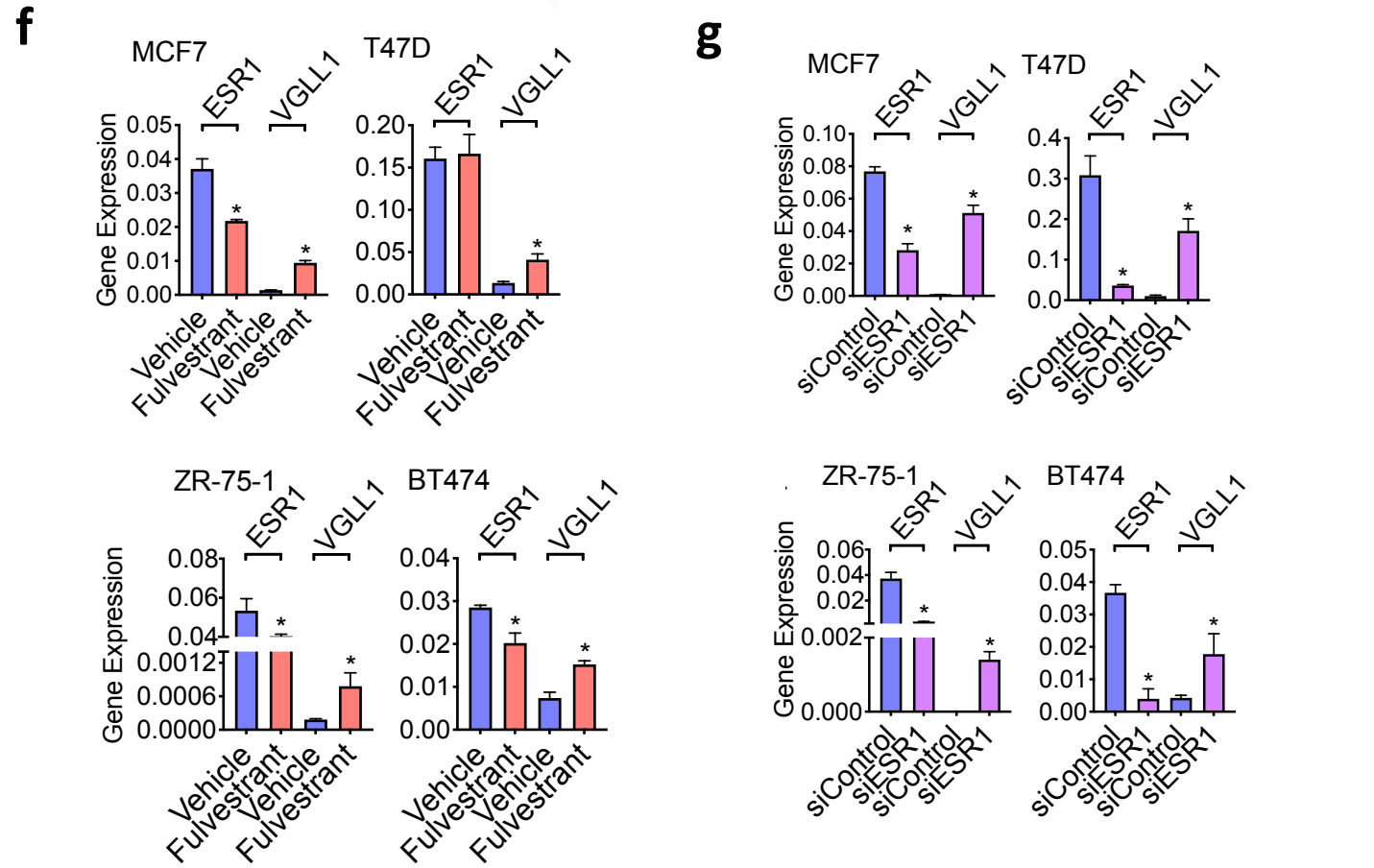

f
MCF7
T47D
ZR-75-1
BT474
ESR1
VGLL1
Gene Expression
Vehicle
Fulvestrant
g
siControl
siESR1

Figure 1

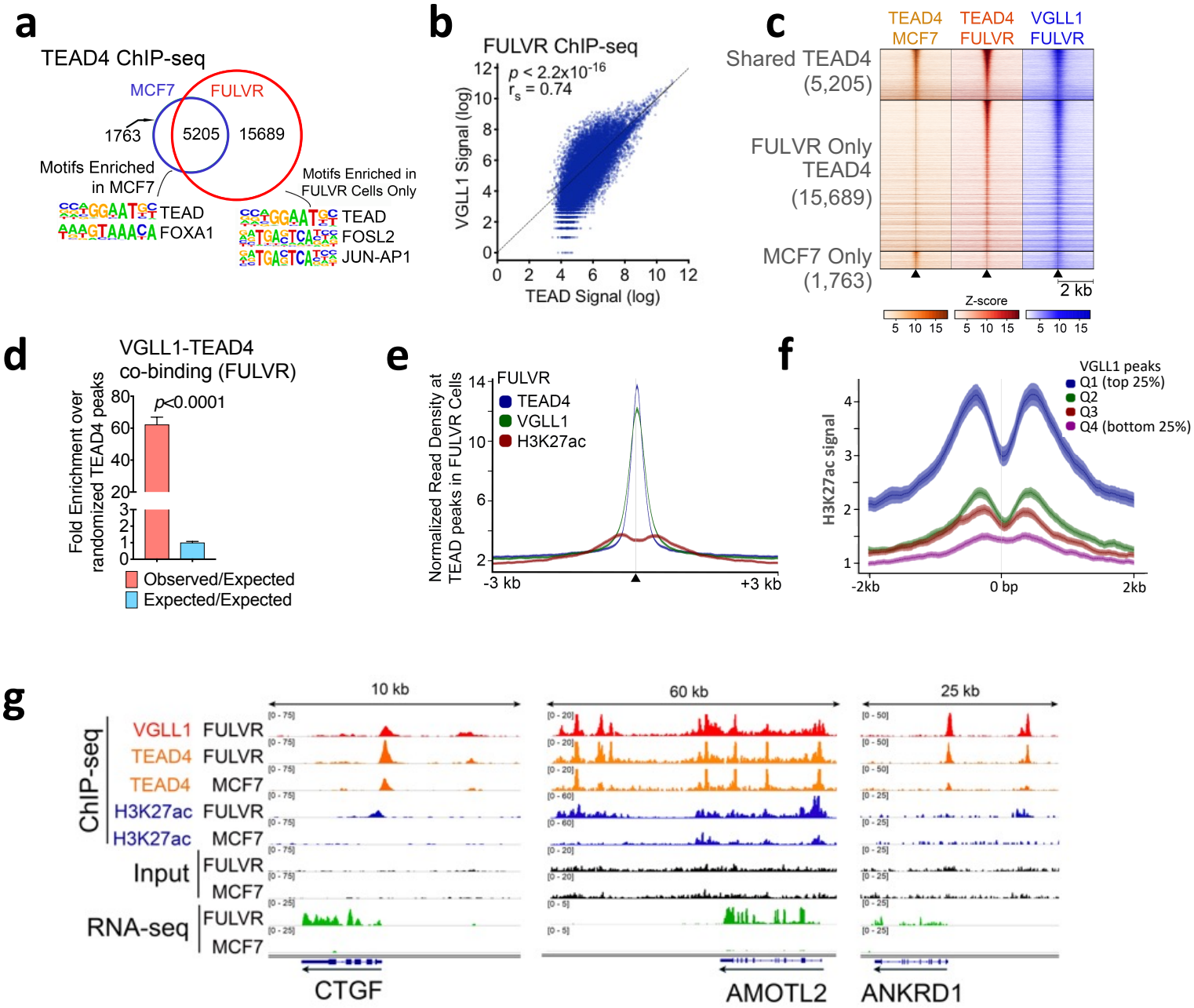
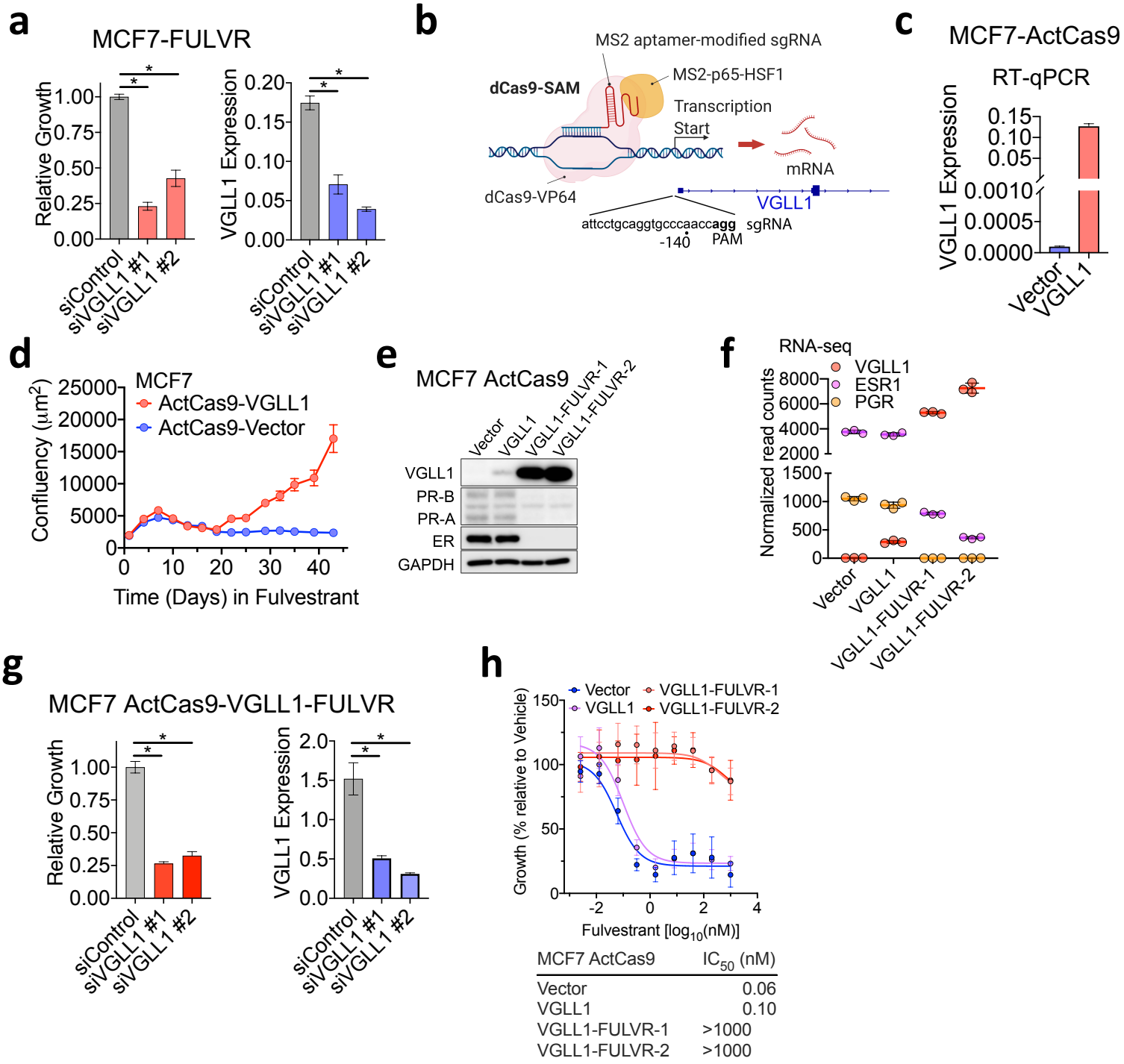

a
TEAD4 ChIP-seq
MCF7
FULVR
1763
5205
15689
Motifs Enriched in MCF7
Motifs Enriched in FULVR Cells Only
TEAD
FOXA1
TEAD
FOSL2
JUN-AP1
b
FULVR ChIP-seq
p < 2.2x10-16
rs = 0.74
VGLL1 Signal (log)
TEAD Signal (log)
c
TEAD4 MCF7
TEAD4 FULVR
VGLL1 FULVR
Shared TEAD4 (5,205)
FULVR Only TEAD4 (15,689)
MCF7 Only (1,763)
2 kb
Z-score
d
VGLL1-TEAD4 co-binding (FULVR)
Fold Enrichment over randomized TEAD4 peaks
p<0.0001
Observed/Expected
Expected/Expected
e
FULVR
TEAD4
VGLL1
H3K27ac
Normalized Read Density at TEAD peaks in FULVR Cells
-3 kb
+3 kb
f
VGLL1 peaks
Q1 (top 25%)
Q2
Q3
Q4 (bottom 25%)
H3K27ac signal
-2kb
0 bp
2kb
g
10 kb
60 kb
25 kb
ChIP-seq
VGLL1 FULVR
TEAD4 FULVR
TEAD4 MCF7
H3K27ac FULVR
H3K27ac MCF7
Input FULVR
Input MCF7
RNA-seq FULVR
RNA-seq MCF7
CTGF
AMOTL2
ANKRD1

Figure 2



MCF7 ActCas9	IC_{50} (nM)
Vector	0.06
VGLL1	0.10
VGLL1-FULVR-1	>1000
VGLL1-FULVR-2	>1000

Figure 3

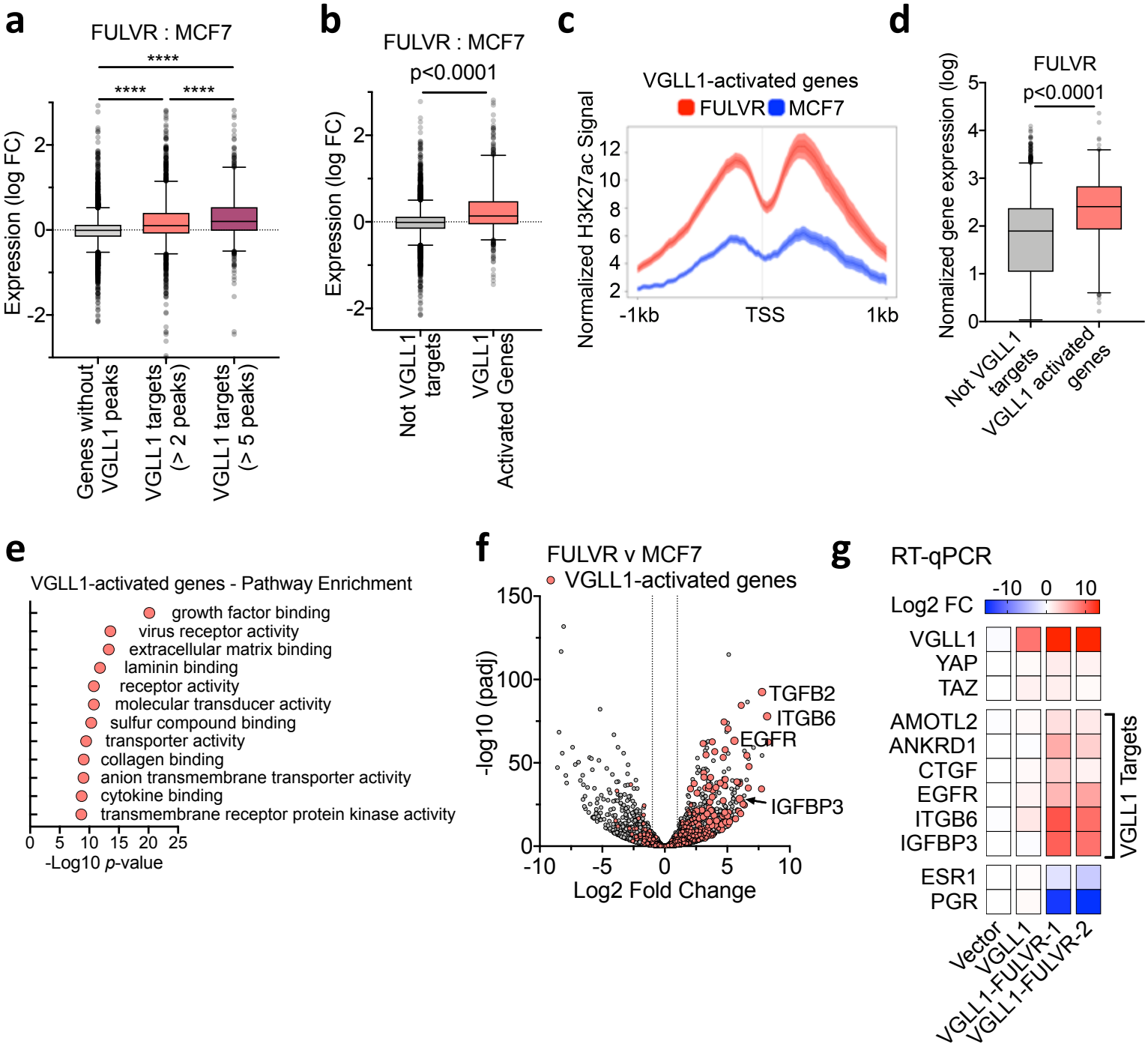

a
FULVR : MCF7
****
****
****
Expression (log FC)
2
0
-2
Genes without VGLL1 peaks
VGLL1 targets (> 2 peaks)
VGLL1 targets (> 5 peaks)
b
FULVR : MCF7
p<0.0001
Expression (log FC)
2
0
-2
Not VGLL1 targets
VGLL1 Activated Genes
c
VGLL1-activated genes
FULVR
MCF7
Normalized H3K27ac Signal
12
10
8
6
4
2
-1kb
TSS
1kb
d
FULVR
p<0.0001
Nomalized gene expression (log)
5
4
3
2
1
0
Not VGLL1 targets
VGLL1 activated genes
e
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
0
5
10
15
20
25
-Log10 p-value
f
FULVR v MCF7
VGLL1-activated genes
-log10 (padj)
150
100
50
0
TGFB2
ITGB6
EGFR
IGFBP3
-10
-5
0
5
10
Log2 Fold Change
g
RT-qPCR
Log2 FC
-10
0
10
VGLL1
YAP
TAZ
AMOTL2
ANKRD1
CTGF
EGFR
ITGB6
IGFBP3
VGLL1 Targets
ESR1
PGR
Vector
VGLL1
VGLL1-FULVR-1
VGLL1-FULVR-2

Figure 4

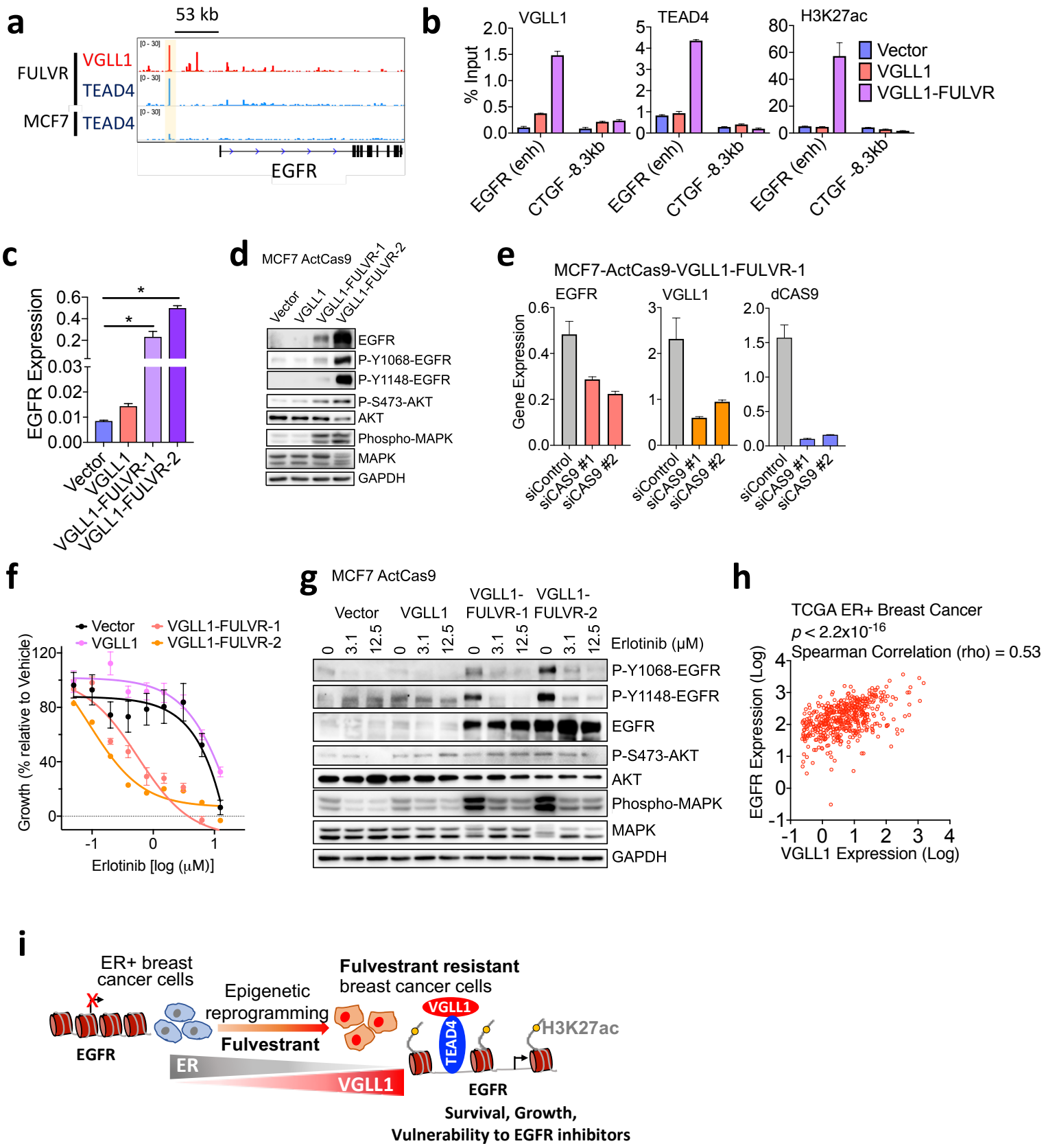

a
53 kb
FULVR
VGLL1
TEAD4
MCF7
TEAD4
EGFR
b
VGLL1
TEAD4
H3K27ac
% Input
Vector
VGLL1
VGLL1-FULVR
EGFR (enh)
CTGF -8.3kb
c
EGFR Expression
Vector
VGLL1
VGLL1-FULVR-1
VGLL1-FULVR-2
d
MCF7 ActCas9
Vector
VGLL1
VGLL1-FULVR-1
VGLL1-FULVR-2
EGFR
P-Y1068-EGFR
P-Y1148-EGFR
P-S473-AKT
AKT
Phospho-MAPK
MAPK
GAPDH
e
MCF7-ActCas9-VGLL1-FULVR-1
EGFR
VGLL1
dCAS9
Gene Expression
siControl
siCAS9 #1
siCAS9 #2
f
Vector
VGLL1
VGLL1-FULVR-1
VGLL1-FULVR-2
Growth (% relative to Vehicle)
Erlotinib [log (μM)]
g
MCF7 ActCas9
Vector
VGLL1
VGLL1-FULVR-1
VGLL1-FULVR-2
Erlotinib (μM)
P-Y1068-EGFR
P-Y1148-EGFR
EGFR
P-S473-AKT
AKT
Phospho-MAPK
MAPK
GAPDH
h
TCGA ER+ Breast Cancer
p < 2.2x10-16
Spearman Correlation (rho) = 0.53
EGFR Expression (Log)
VGLL1 Expression (Log)
i
ER+ breast cancer cells
EGFR
Epigenetic reprogramming
Fulvestrant
ER
VGLL1
Fulvestrant resistant breast cancer cells
VGLL1
TEAD4
H3K27ac
EGFR
Survival, Growth,
Vulnerability to EGFR inhibitors

Figure 5

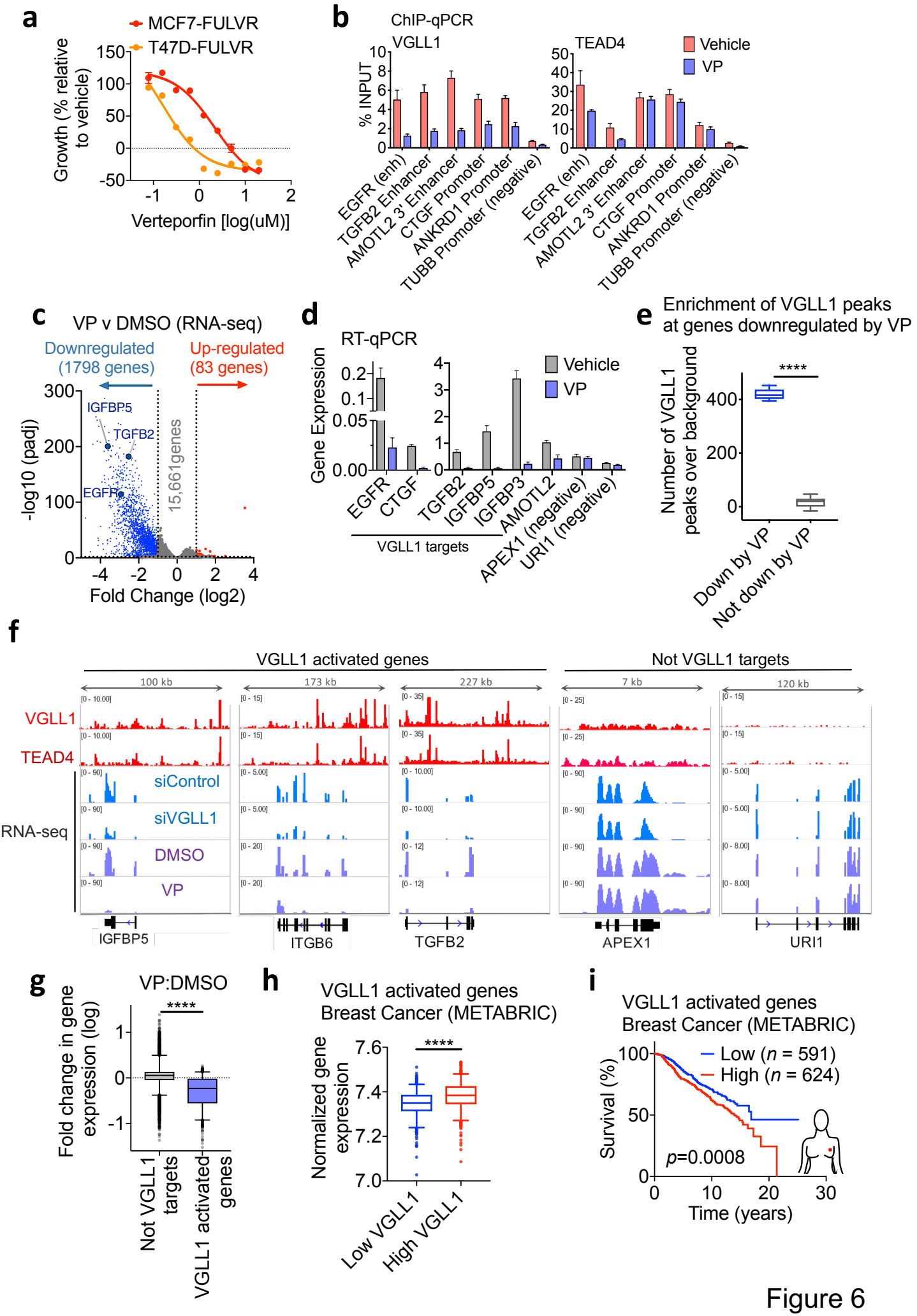
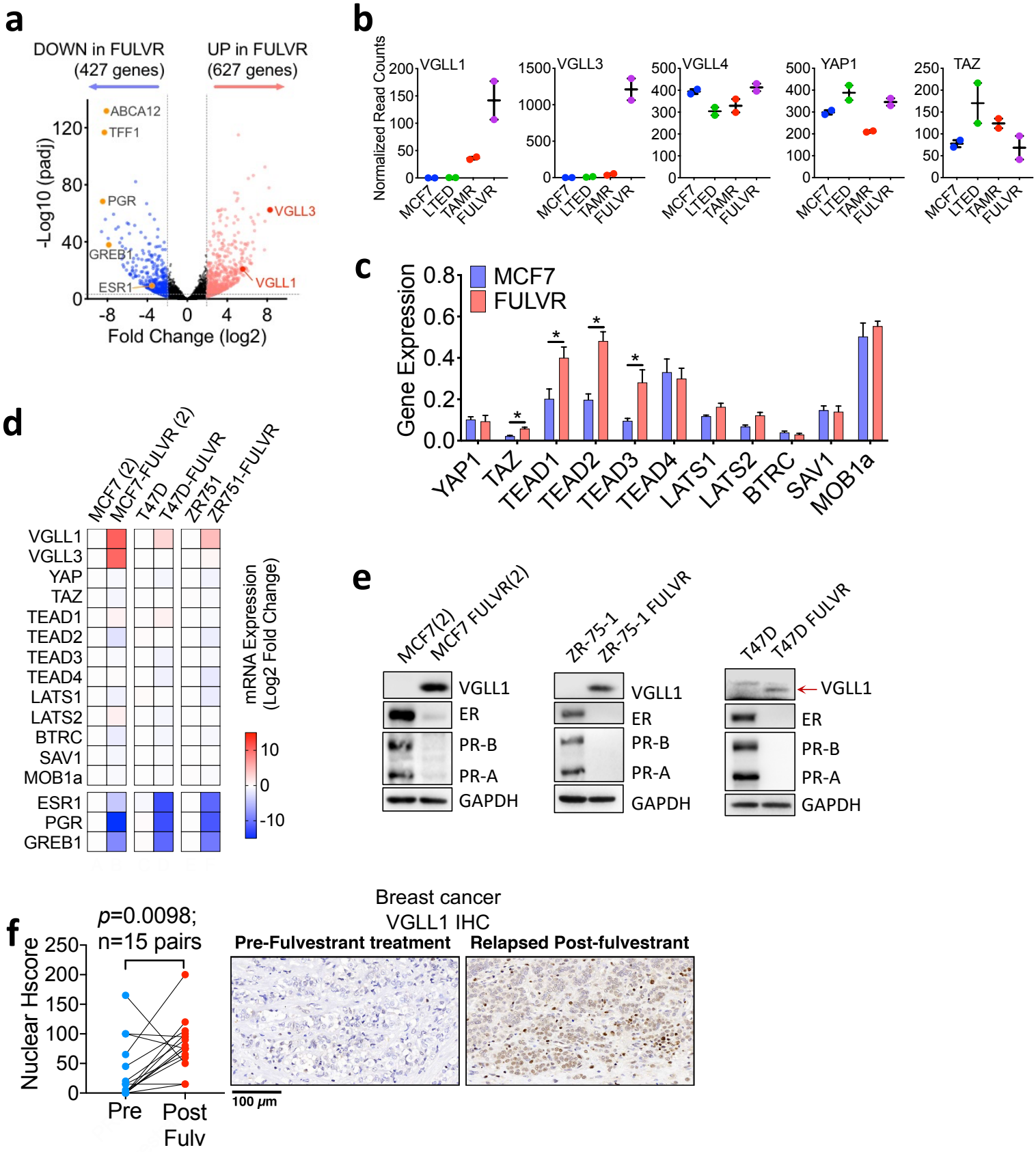
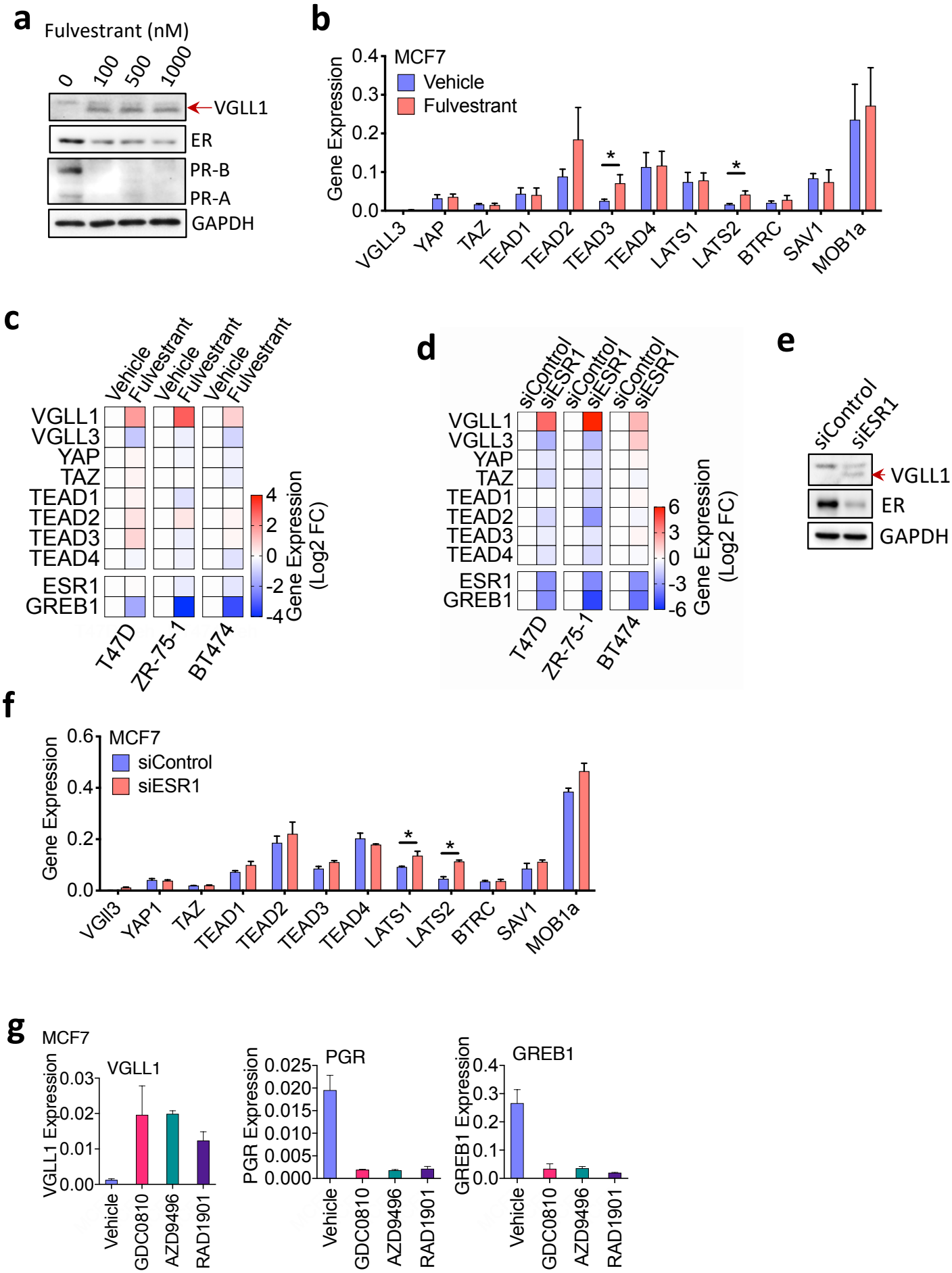

a
MCF7-FULVR
T47D-FULVR
Growth (% relative to vehicle)
Verteporfin [log(uM)]
b
ChIP-qPCR
VGLL1
TEAD4
% INPUT
Vehicle
VP
EGFR (enh)
TGFB2 Enhancer
AMOTL2 3' Enhancer
CTGF Promoter
ANKRD1 Promoter
TUBB Promoter (negative)
c
VP v DMSO (RNA-seq)
Downregulated (1798 genes)
Up-regulated (83 genes)
IGFBP5
TGFB2
EGFR
15,661genes
-log10 (padj)
Fold Change (log2)
d
RT-qPCR
Gene Expression
Vehicle
VP
EGFR
CTGF
TGFB2
IGFBP5
IGFBP3
AMOTL2
APEX1 (negative)
URI1 (negative)
VGLL1 targets
e
Enrichment of VGLL1 peaks at genes downregulated by VP
Number of VGLL1 peaks over background
****
Down by VP
Not down by VP
f
VGLL1 activated genes
Not VGLL1 targets
100 kb
173 kb
227 kb
7 kb
120 kb
VGLL1
TEAD4
RNA-seq
siControl
siVGLL1
DMSO
VP
IGFBP5
ITGB6
TGFB2
APEX1
URI1
g
VP:DMSO
Fold change in gene expression (log)
****
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
Time (years)

Figure 6

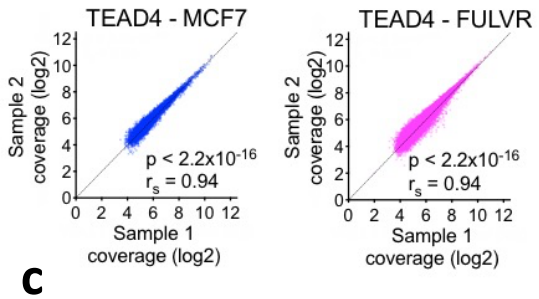


Supplementary Figure 1



Supplementary Figure 2

a



b

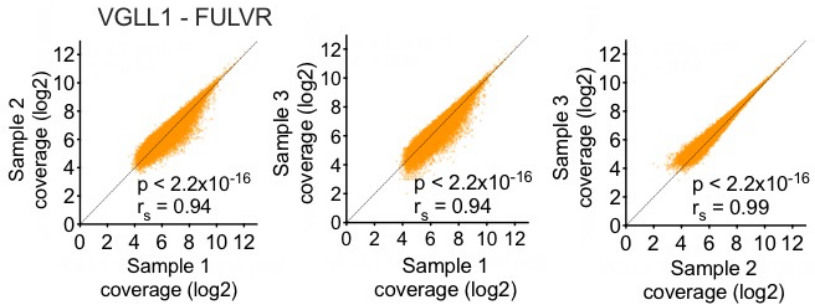
TEAD4 Binding Regions (MCF7)

Motif	Rank	TF	p-value
	1	TEAD4	1e-1479
	2	TEAD2	1e-1439
	3	TEAD	1e-1188
	4	FOXA1	1e-107
	5	FOXA1	1e-105
	6	AP2γ	1e-104
	7	FOXM1	1e-102
	8	AP2α	1e-101
	9	FOXA2	1e-88
	10	FOX:E BOX	1e-65

TEAD4 Binding Regions (FULVR)

Motif	Rank	TF	p-value
	1	TEAD4	1e-4752
	2	TEAD2	1e-4374
	3	TEAD	1e-3699
	4	FOSL2	1e-314
	5	JUN-AP1	1e-312
	6	FRA1	1e-296
	7	BATF	1e-277
	8	ATF3	1e-276
	9	AP1	1e-247
	10	GRHL2	1e-241

c

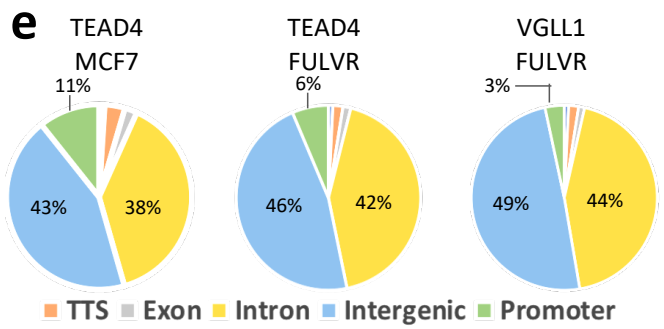


d

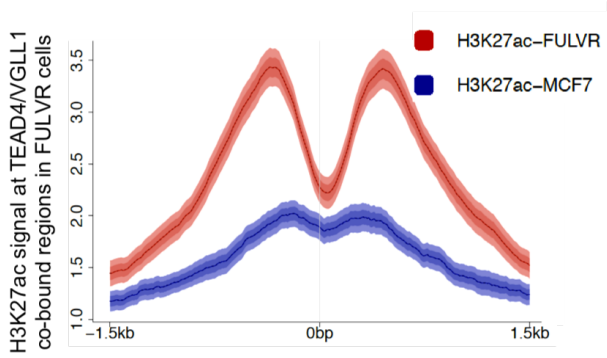
VGLL1 Binding Regions (FULVR)

Motif	Rank	TF	p-value
	1	TEAD4	1e-1581
	2	GRHL2	1e-1546
	3	TEAD	1e-1432
	4	FOXA1	1e-1346
	5	FOXA1	1e-1323
	6	FOXM1	1e-1319
	7	TEAD2	1e-1267
	8	FRA1	1e-1172
	9	JUN-AP1	1e-1130
	10	FOSL2	1e-1125

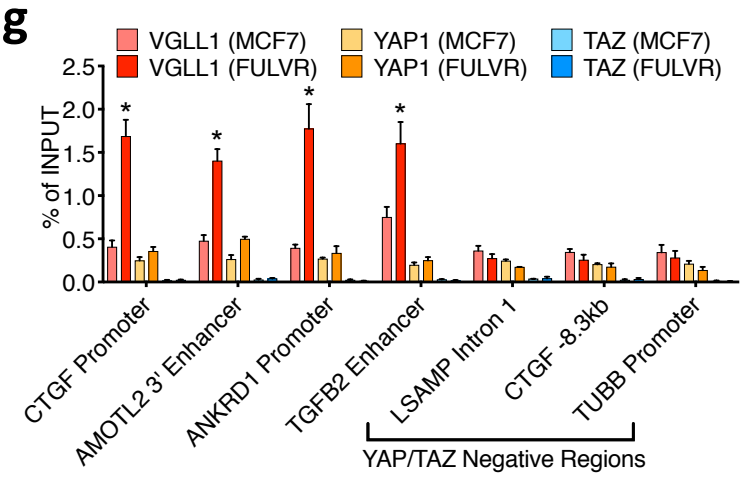
e



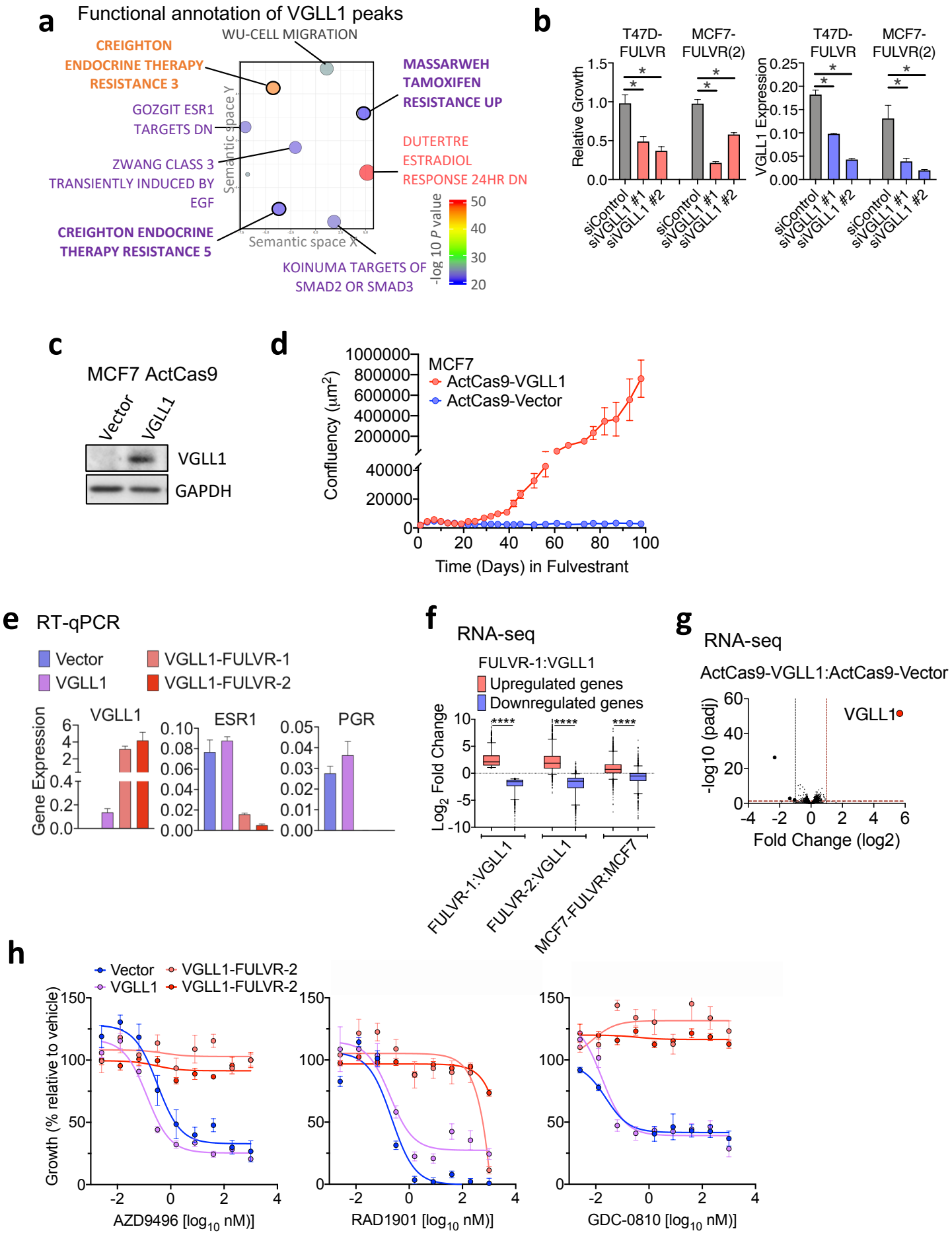
f



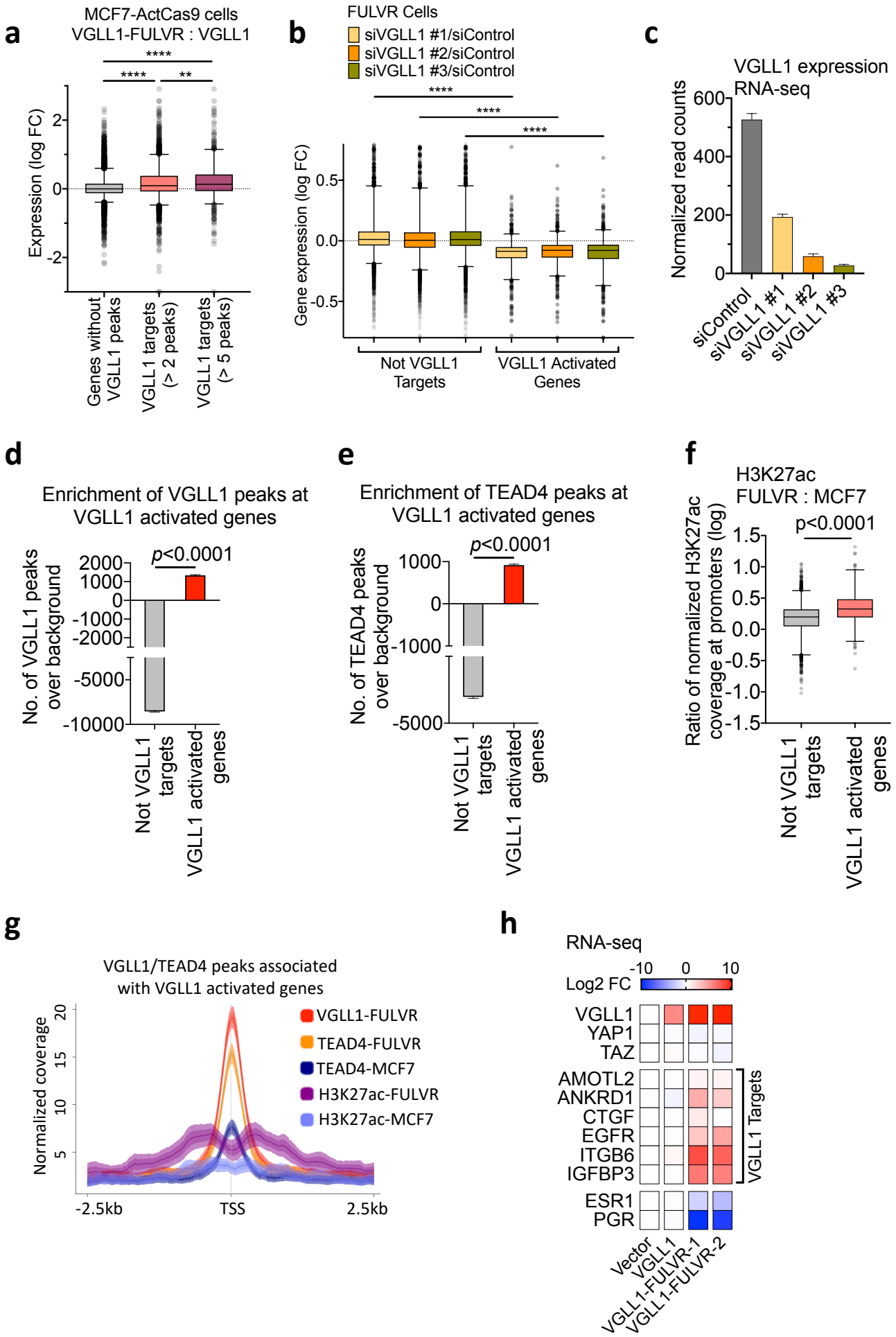
g



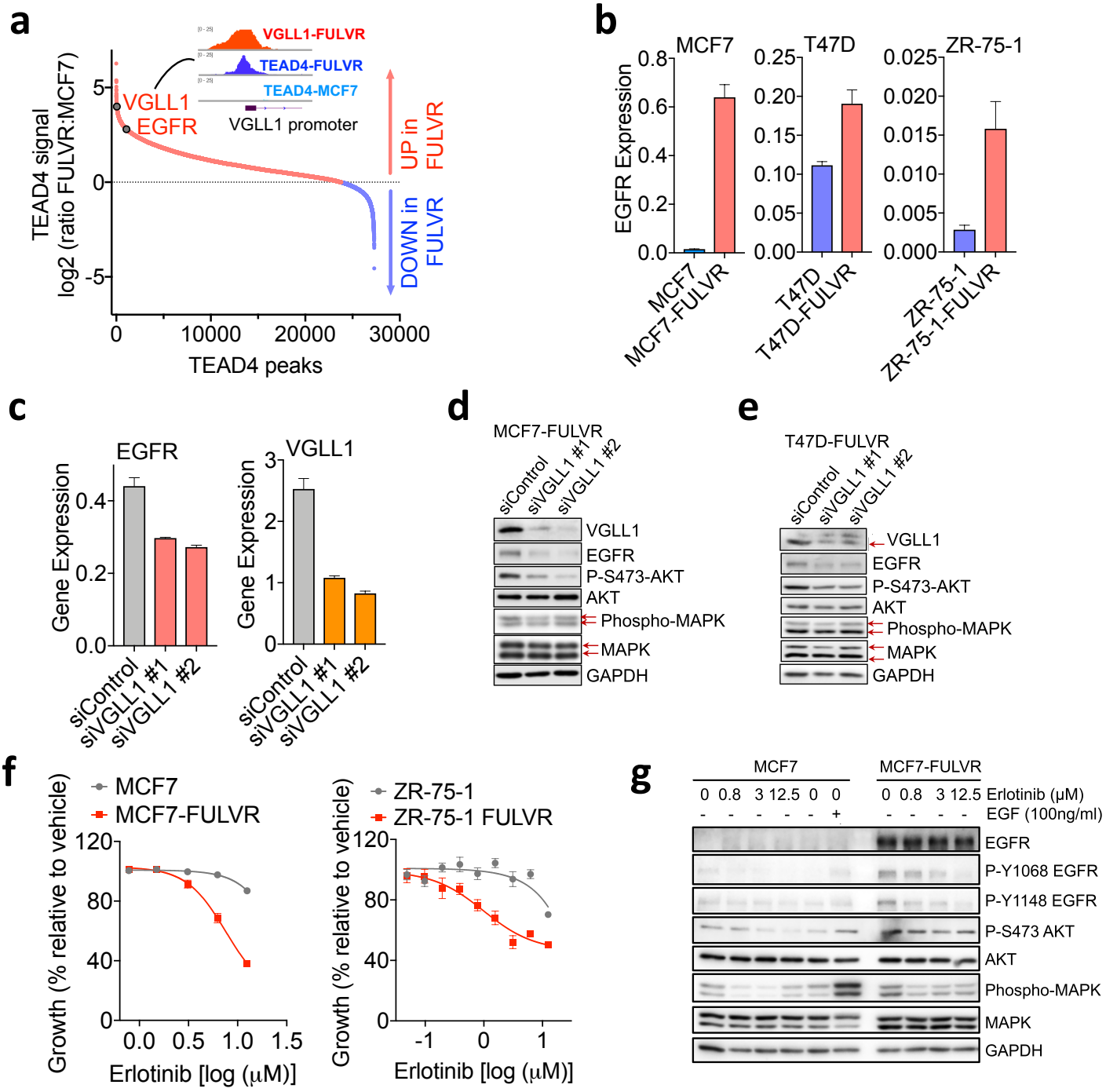
Supplementary Figure 3



Supplementary Figure 4

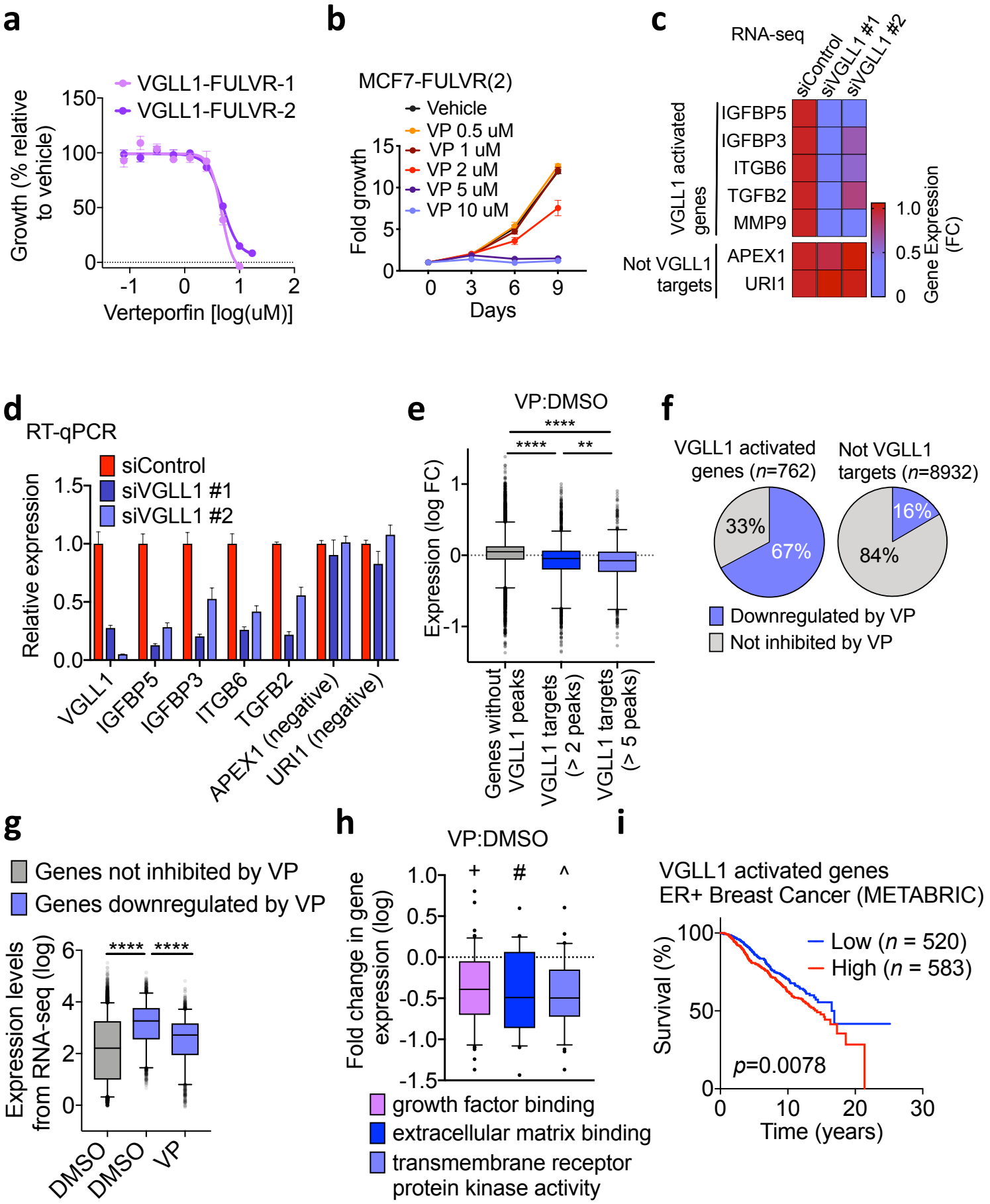


Supplementary Figure 5



	Rank	Gene	Spearman's Correlation	p-value
METABRIC - 1904 samples	1	EGFR	0.531	5.22e-139
	2	CDH3	0.524	5.22e-135
	3	PPP1R14C	0.519	5.22e-131
TCGA RPPA - 673 samples	1	EGFR	0.392	3.79e-26
	2	CDH3	0.382	8.92e-25
	3	CCNE1	0.357	1.14e-21

Supplementary Figure 6



Supplementary Figure 7