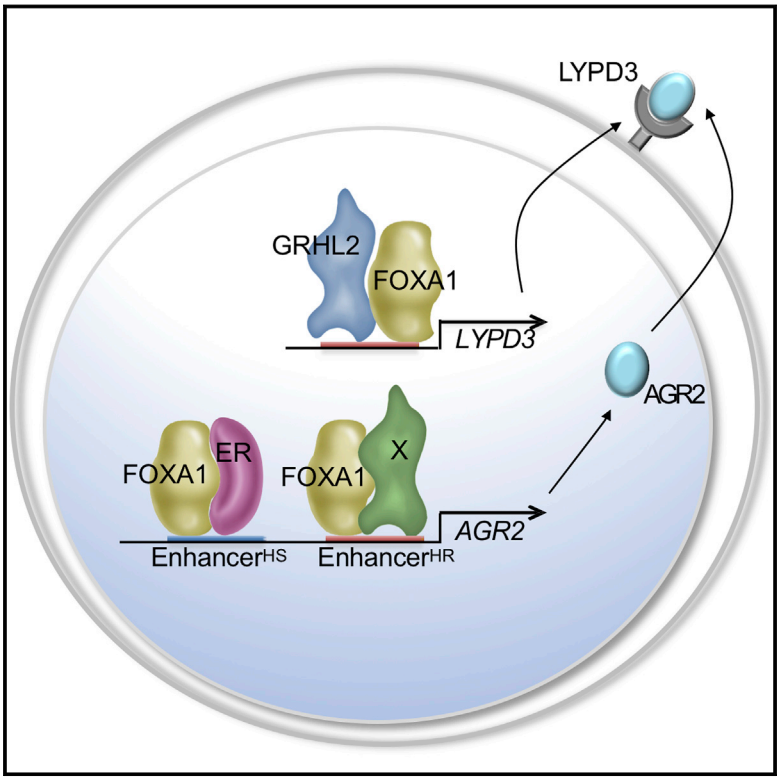
Article

Cell Reports

The Lineage Determining Factor GRHL2 Collaborates with FOXA1 to Establish a Targetable Pathway in Endocrine Therapy-Resistant Breast Cancer

Graphical Abstract



Authors

Kimberly J. Cocce, Jeff S. Jasper, Taylor K. Desautels, ..., John D. Norris, Ching-yi Chang, Donald P. McDonnell

Correspondence

donald.mcdonnell@duke.edu

In Brief

Cocce et al. show that FOXA1 contributes to disease pathogenesis by cooperating with GRHL2 in endocrine therapy-resistant breast cancer. LYPD3 is identified as an actionable downstream target of FOXA1/GRHL2, and humanized antibodies against LYDP3, or its ligand AGR2, demonstrate anti-tumor efficacy in animal models of endocrine therapy-resistant breast tumors.

Highlights

- FOXA1 cooperates with GRHL2 to drive resistance to endocrine therapy in breast cancer
- LYPD3 is a targetable node downstream of FOXA1/GRHL2
- Targeting LYPD3 and its ligand AGR2 inhibit the growth of therapy-resistant tumors

Cocce et al., 2019, Cell Reports *29*, 889–903
October 22, 2019
https://doi.org/10.1016/j.celrep.2019.09.032

CellPress

Cell Reports
Article

The Lineage Determining Factor GRHL2 Collaborates with FOXA1 to Establish a Targetable Pathway in Endocrine Therapy-Resistant Breast Cancer

Kimberly J. Cocce,[1,15] Jeff S. Jasper,[1,15] Taylor K. Desautels,[1] Logan Everett,[2] Suzanne Wardell,[1] Thomas Westerling,[3,4] Robert Baldi,[1] Tricia M. Wright,[1] Kendall Tavares,[1] Alex Yllanes,[1] Yeeun Bae,[1] Jeremy T. Blitzer,[5] Craig Logsdon,[6] Daniel P. Rakiec,[7] David A. Ruddy,[7] Tiancong Jiang,[8] Gloria Broadwater,[8] Terry Hyslop,[8] Allison Hall,[9] Muriel Laine,[13] Linda Phung,[13] Geoffrey L. Greene,[13] Lesley-Ann Martin,[14] Sunil Pancholi,[14] Mitch Dowsett,[10] Simone Detre,[10] Jeffrey R. Marks,[11] Gregory E. Crawford,[12] Myles Brown,[3,4] John D. Norris,[1] Ching-yi Chang,[1] and Donald P. McDonnell[1,16,*]

[1]Department of Pharmacology and Cancer Biology, Duke University School of Medicine, Durham, NC 27710, USA
[2]Department of Biological Sciences, North Carolina State University, Raleigh, NC 27695, USA
[3]Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[4]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA
[5]Viba Therapeutics, San Francisco, CA 94158, USA
[6]Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77054, USA
[7]Novartis Institutes for Biomedical Research, Oncology Disease Area, Cambridge, MA 02139, USA
[8]Department of Biostatistics, Duke Cancer Institute, Duke University School of Medicine, Durham, NC 27710, USA
[9]Department of Pathology, Duke University School of Medicine, Durham, NC 27710, USA
[10]Ralph Lauren Centre for Breast Cancer Research, Royal Marsden Hospital NHS Trust, London, SW3 6JJ, UK
[11]Department of Surgery, Duke University School of Medicine, Durham, NC 27710, USA
[12]Department of Pediatrics, Duke University School of Medicine, Durham, NC 27710, USA
[13]The Ben May Department for Cancer Research, The University of Chicago, Chicago, IL 60637, USA
[14]Breast Cancer Now, Toby Robins Research Centre, Institute of Cancer Research, London, SW3 6JB, UK
[15]These authors contributed equally
[16]Lead Contact
*Correspondence: donald.mcdonnell@duke.edu
https://doi.org/10.1016/j.celrep.2019.09.032

SUMMARY

Notwithstanding the positive clinical impact of endocrine therapies in estrogen receptor-alpha (ERα)-positive breast cancer, *de novo* and acquired resistance limits the therapeutic lifespan of existing drugs. Taking the position that resistance is nearly inevitable, we undertook a study to identify and exploit targetable vulnerabilities that were manifest in endocrine therapy-resistant disease. Using cellular and mouse models of endocrine therapy-sensitive and endocrine therapy-resistant breast cancer, together with contemporary discovery platforms, we identified a targetable pathway that is composed of the transcription factors FOXA1 and GRHL2, a coregulated target gene, the membrane receptor LYPD3, and the LYPD3 ligand, AGR2. Inhibition of the activity of this pathway using blocking antibodies directed against LYPD3 or AGR2 inhibits the growth of endocrine therapy-resistant tumors in mice, providing the rationale for near-term clinical development of humanized antibodies directed against these proteins.

INTRODUCTION

The majority of breast cancers express estrogen receptor-alpha (ERα), and drugs that target the production of estrogens or which directly interfere with the transcriptional activity of ERα have become frontline interventions in the treatment and prevention of this disease (Brodie, 2002; Fisher et al., 1998, 2001; Perou et al., 2000; McDonnell et al., 2015). Although these treatments have been effective, clinical experience with currently available ERα modulators and the results of preclinical studies of drugs currently under development indicate that resistance is a seemingly inevitable adaptive event that will limit the efficacy of any endocrine therapy in breast cancer (Jeselsohn et al., 2014, 2018; Toy et al., 2013).

Whereas aromatase inhibitors (AIs) have largely replaced tamoxifen as first-line endocrine therapy in post-menopausal women with ER+ breast cancer, it is now apparent that there is considerable overlap in the mechanisms that underlie resistance to both drugs, a finding that may explain the high level of cross-resistance between these types of interventions (Brodie, 2002; Dowsett and Howell, 2002; Lønning, 2002; Mokbel, 2002; Palmieri et al., 2014). Of particular relevance is the observation that long-term estrogen deprivation facilitates adaptive events that permit ERα and its co-regulators to activate transcription in a ligand-independent manner (Britton et al., 2006;

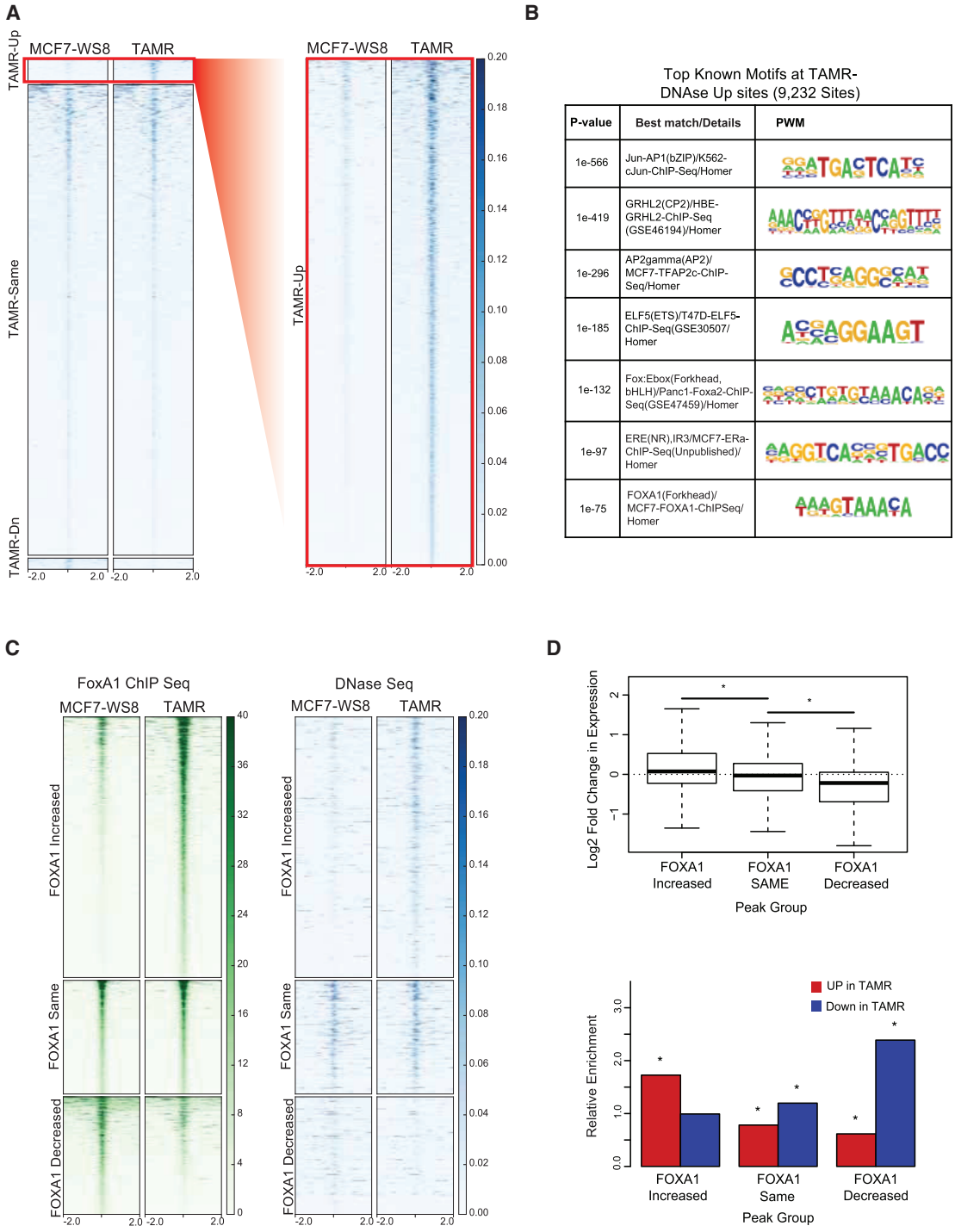


Figure 1. FOXA1 as a Key Mediator of Acquired Alterations in the Cistrome in Setting of Tamoxifen Resistance

(A) Heatmap of DNase signal in a 4 kb window of (left) all DHSs identified in MCF7-WS8 and TAMR, subdivided on the basis of whether they are significantly increased in TAMR (TAMR-Up), significantly decreased in TAMR (TAMR-Dn), or not significantly different between cell lines (TAMR-Same) and (right) zoomed in view of only those TAMR-Up DHSs.

(B) Position weight matrices (PWMs) indicating known motifs enriched in TAMR-Up DHSs.

(C) Heatmaps showing signal in a 4 kb window of (left) FOXA1 binding events as determined by ChIP-seq in MCF7-WS8 relative to TAMR and (right) DNase-seq, ordered on the basis of FOXA1 binding profiles. Subgroup naming is determined on the basis of FOXA1 binding profile in TAMR relative to MCF7-WS8: sites where FOXA1 binding is significantly increased (FOXA1 increased), where there is no statistically significant difference (FOXA1 same), and sites where FOXA1 binding is decreased (FOXA1 decreased).

(legend continued on next page)

Knowlden et al., 2005; Lupien et al., 2010; Massarweh et al., 2008; Santen et al., 2005; Smith et al., 1993). Ligand-independent activation of ERα can also occur in cells in which the expression and/or activity of receptor-interacting co-regulators are elevated or in which direct phosphorylation of the receptor stabilizes its interaction with co-regulators. In either case, it is assumed that existing ERα modulators enable the outgrowth of a subpopulation of cells that express the appropriate co-regulator repertoire and/or signaling kinases needed to support ligand-independent activity of the receptor (Osborne et al., 2003; Smith et al., 1997). Such activities are associated with resistance to endocrine therapies.

In this study, we used pharmacological and biochemical approaches to identify targets whose expression and activity accompanies the development of resistance to endocrine therapies through interaction with FOXA1, a key lineage-selective transcription factor whose overexpression and/or increased activity has been shown to be associated with the development of endocrine therapy resistance (Carroll et al., 2005; Fu et al., 2016; Hurtado et al., 2011; Kong et al., 2011; Ross-Innes et al., 2012; Sérandour et al., 2011; Wright et al., 2014). Specifically, we determined that FOXA1 collaborates with GRHL2 to increase the expression and activity of LYPD3/AGR2, a receptor ligand complex that regulates processes of pathological importance in cancer. This work culminated in the validation of inactivating antibodies directed against LYPD3, and its extracellular protein ligand AGR2, as therapeutic approaches in advanced endocrine therapy-resistant breast cancer.

RESULTS

Alterations in the FOXA1 Cistrome Accompany the Development of Endocrine Therapy Resistance in Preclinical Models of Luminal Breast Cancer

We developed a model of tamoxifen-resistant breast cancer by serially passaging an aggressive subline of ERα-positive MCF7 breast cancer cells (MCF7-WS8) as a xenograft in the presence of tamoxifen (Connor et al., 2001; Gottardis and Jordan, 1988; Pink et al., 1995). The resulting tumors, whose characteristics reflect that seen in patients with endocrine therapy-resistant disease, grow in an estrogen-independent manner and recognize tamoxifen as an agonist (Figure S1A) (Martz et al., 2014; Nelson et al., 2013; Wardell et al., 2015; Wright et al., 2014). In this study, we used the parental tamoxifen-sensitive (MCF7-WS8) cell line and a cell line derived from a tamoxifen-resistant variant of MCF7-WS8-derived tumors (TAMR) to identify molecular events that are associated with estrogen-independent growth and tamoxifen resistance. To this end, DNase sequencing (DNase-seq) analysis was used to perform an unbiased genome-wide survey of changes in chromatin architecture with the goal of identifying potential *cis*-acting elements enriched in TAMR versus MCF7-WS8 cells. In this manner, 205,924 DNase-hypersensitive sites (DHSs) were identified in the two genomes, the majority of which (192,136) were equally enriched in MCF7-WS8 and TAMR cells. Notably, 9,232 sites with significantly increased hypersensitivity and an additional 4,556 sites that show decreased hypersensitivity in TAMR cells were identified, when compared with the parental MCF7-WS8 line. Our continued analysis focused on those sites that demonstrated increased hypersensitivity in TAMR, as we considered that gain-of-function processes that enabled or occurred as a consequence of these changes were the most likely to be informative with respect to the identification of new therapeutic targets (Figure 1A).

Motif enrichment analysis of sequences within the "gained" DHSs in TAMR indicated that the transcription factors most likely to be interacting at these genomic loci include those that bind bZip motifs, GRHL2, AP2 factors, ER, and FOXA family members (Figure 1B). It has been observed previously that tamoxifen-resistant cells derived from MCF7 cells acquire enhanced FOXA1 activity at functional enhancers, some of which overlap with ER binding sites (Fu et al., 2016; Hurtado et al., 2011). Thus, to examine the potential role(s) of FOXA1 in our system, we overlaid our DNase hypersensitivity data with previously published FOXA1 and ER chromatin immunoprecipitation sequencing (ChIP-seq) data from MCF7 cells (Hurtado et al., 2011) (Figure S1B). This analysis indicated that the differences observed in the genome-wide DHSs between MCF7-WS8 and TAMR correspond to those sites bound by FOXA1 (either alone or at sites shared with ER). Surprisingly, no significant increase in hypersensitivity was noted in TAMR cells at sites previously described as FOXA1-independent ER binding events. To confirm that the sites identified using DNase hypersensitivity analysis were indeed bona fide FOXA1 binding events, we performed FOXA1 ChIP-seq in both MCF7-WS8 and TAMR cell lines (Figure 1C, left panel). In this study, 47,337 high-confidence FOXA1 binding sites were identified in the TAMR and MCF7-WS8 genomes, with 25,093 sites scoring as novel or enhanced in TAMR relative to MCF7-WS8. A further 10,939 sites were found to be represented equivalently in both cell lines, and an additional 11,315 sites showed decreased binding in TAMR relative to MCF7-WS8 cells. An overlap analysis of the FOXA1 ChIP-seq and DNase-seq data revealed a strong degree of concordance between FOXA1 binding and DNase hypersensitivity (Figure 1C, right panel). Thus, whereas increased FOXA1 binding at ERα containing enhancers has been observed by others in cellular models of tamoxifen resistance, our analysis highlights additional roles for FOXA1-dependent/ERα-independent enhancers in the pathobiology of breast cancer (Fu et al., 2016; Hurtado et al., 2011).

(D) FOXA1 binding events as defined in (C) were compared with gene expression in TAMR and MCF7-WS8 cells as determined by RNA-seq. Genes with transcription start sites within ±10 kb of any FOXA1 binding event were assigned to at most one set of peaks, with the leftmost group having highest assignment priority. Analysis is based on (top) $\log_2$ fold change in gene expression; *$p < 0.01$ by Mann-Whitney test comparing the pair of boxplots marked by the horizontal line, or (bottom) relative enrichment of significantly differentially expressed genes relative to a control set of genes with similar average expression level but minimal fold change. Red bars denote genes significantly upregulated in TAMR versus MCF7-WS8. Blue bars denote genes significantly downregulated in TAMR versus MCF7-WS8. Asterisk denotes significant difference from 1, with p value cutoff of 0.01 using Fisher's exact test.
See also Figure S1.

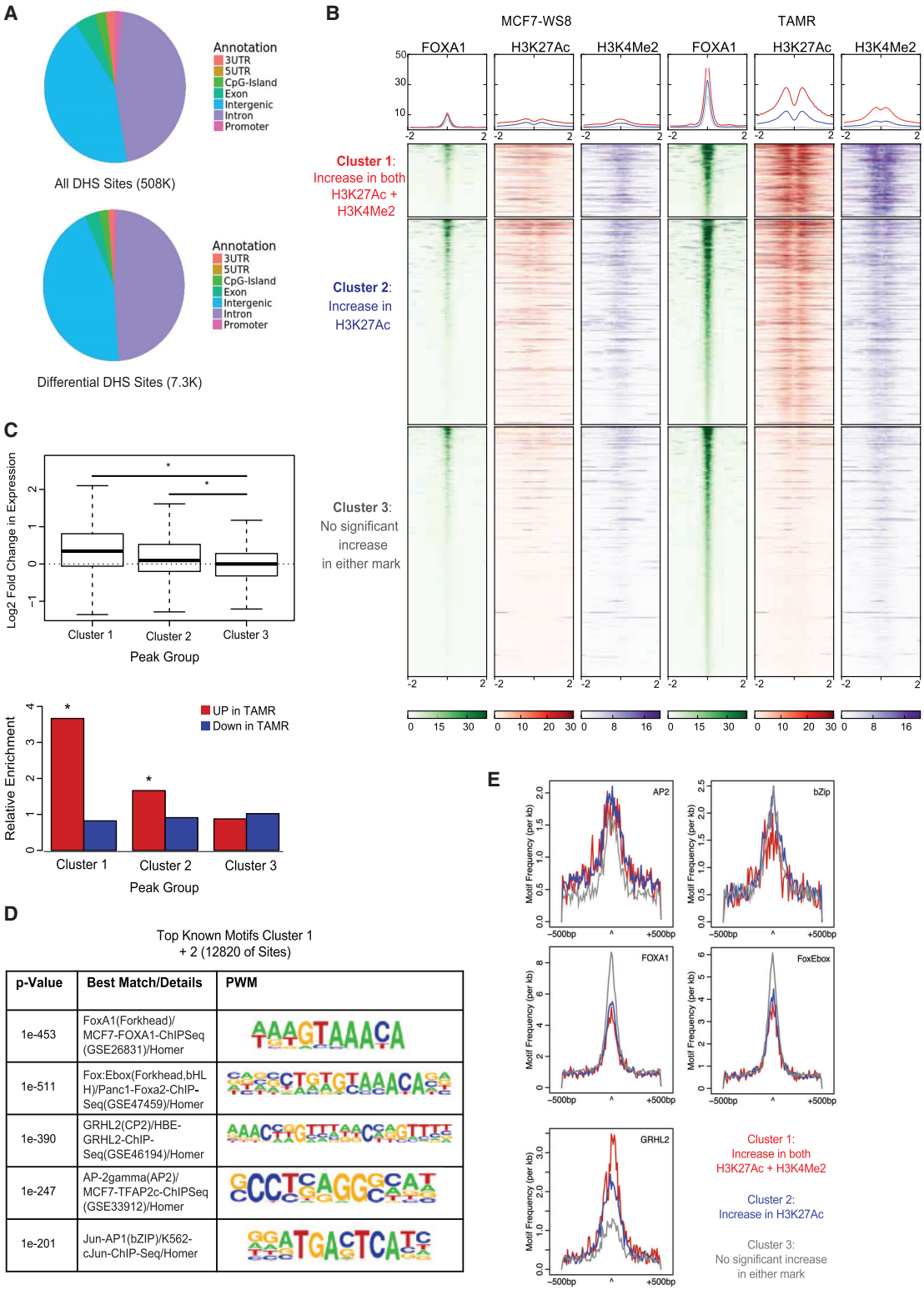


Figure 2. Epigenetic Signatures of Histones Flanking FOXA1 Binding Events Increased in TAMR Define Different Subsets of Enhancers

(A) Pie charts indicating genomic distribution of DHSs across background sites (top) and at sites determined to be significantly different between TAMR and MCF7-WS8 (bottom).

(legend continued on next page)

We next performed RNA sequencing (RNA-seq) on MCF7-WS8 and TAMR cells and used this information to evaluate potential changes in gene transcription that were associated with the FOXA1 binding sites identified (Figure 1D). Genes that were located within 10 kb of the FOXA1 sites identified in the TAMR cells were divided into three categories (increased, equivalent, or decreased binding of FOXA1). It was determined that genes within 10 kb of FOXA1 sites that demonstrate increased binding in TAMR are more likely to be increased in expression in the TAMR cells relative to the MCF7-WS8 cells. Lesser differences in the expression of genes within 10 kb of FOXA1 binding events that were equivalent in the two cell lines were noted, whereas decreased expression was found for those genes located at the same distance from FOXA1 sites that have decreased binding of TAMR relative to MCF7-WS8. Collectively, these data suggest that independent of ERα, FOXA1 is involved in the regulation of the expression of genes that distinguish tamoxifen-sensitive from tamoxifen-resistant cells. These findings provided the impetus to identify the factor(s) that cooperate with FOXA1, with the goal of identifying new exploitable therapeutic targets.

FOXA1-Dependent *cis*-Regulatory Elements Demonstrating Enhanced Activation Status in TAMR Cells Are Co-occupied by GRHL2

The majority of DHSs observed in our cell line models occur within intronic and intergenic areas of the genome (Figure 2A). We next probed whether the gained FOXA1 binding events in TAMR were associated with increased enhancer function (Lupien et al., 2008; Sérandour et al., 2011). Thus, we performed H3K4Me2 and H3K27Ac ChIP-seq and compared these data with our FOXA1 ChIP-seq data to define those binding events most likely to be associated with active enhancers. The dynamics of these histone marks were evaluated at those FOXA1 binding sites that were gained in TAMR cells, and this information was used to subdivide these sites into three major classes (Figure 2B). Cluster 1 contains gained FOXA1 binding sites that were associated with increases in both H3K4Me2 and H3K27Ac marks (active enhancers). Cluster 2 contains sites that demonstrate an enrichment of H3K27Ac marks alone (poised enhancers). Cluster 3 is composed of gained FOXA1 binding sites for which no significant changes in either H3K4Me2 or H3K27Ac marks were apparent.

The expression of transcripts associated with genes in each of the three clusters was next evaluated. Both the $\log_2$ fold change distribution and the relative enrichment of genes that are differentially expressed in TAMR relative to MCF7-WS8 in each cluster were analyzed (Figure 2C). This analysis indicated that cluster 1 binding sites are most associated with genes exhibiting increased expression in TAMR relative to MCF7-WS8 cells. Less robust, though significant, increases in transcription of genes associated with cluster 2 binding sites were observed. No association with expression changes in either direction were noted in genes associated with the cluster 3 binding sites. Thus, a significant number of the FOXA1 binding events that are increased in TAMR are located within active enhancers and are associated with increased gene transcription.

A motif enrichment analysis of the activated, FOXA1-dependent enhancers present in both clusters 1 and 2 revealed a significant enrichment of binding sites for Fox.Ebox, GRHL2, AP1/bZIP, and AP2 (Figure 2D). Positional enrichment analysis of the motifs around FOXA1 peaks identified by ChIP-seq was used to assess the relative enrichment of the binding sites for these transcription factors across clusters 1–3 (Figure 2E). It was determined that AP2 and AP1/bZIP binding motifs are equally distributed across all three clusters. FOXA1 and Fox.Ebox binding sites were enriched in cluster 3. Interestingly, the centrality and enrichment of GRHL2 motifs correlated specifically with H3K27Ac as it relates to the different clusters (cluster 1 > cluster 2 > cluster 3) (Figure 2F). This observation, indicating a likely functional association between FOXA1 and GRHL2, is further supported by the results of a liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis of FOXA1 interaction proteins in TAMR that revealed that GRHL2 interacts with FOXA1 in the setting of tamoxifen resistance (Table S1). Such interactions have been reported previously to occur in endocrine therapy-sensitive breast cancer cells (Jozwik et al., 2016).

A GRHL2 ChIP-seq analysis was performed in MCF7-WS8 and TAMR cells to probe more directly its potential roles in FOXA1 activity. Using this approach, 20,283 (MCF7-WS8) and 35,406 (TAMR) high-confidence GRHL2 sites were identified. The FOXA1 binding events that were significantly increased in TAMR cells relative to MCF7-WS8 cells were then evaluated as a function of the degree of GRHL2 binding. This resulted in the identification of some gained FOXA1 binding sites that were associated with increased GRHL2 binding in TAMR relative to MCF-WS8 (GRHL2 increased). The GRHL2 binding activity associated with a second group of FOXA1 binding sites was unchanged (GRHL2 same). The remainder of the gained FOXA1

(B) Heatmaps centered on 4 kb window indicating comparison of ChIP-seq of all FOXA1 binding events that were significantly increased in TAMR relative to MCF7-WS8, across histone mark signature: H3K27Ac and H3K4Me2. This comparison indicates three patterns, those sites where FOXA1 binding is significantly increased in TAMR and both histone marks are significantly increased in TAMR relative to MCF7-WS8 (cluster 1, red), those where only H3K27Ac is significantly increased in TAMR relative to MCF7-WS8 (cluster 2, blue), and those that do not have statistically significant difference in either mark (cluster 3, gray). Line graphs above each heatmap indicate average signal intensity for each cluster within heatmap below.

(C) FOXA1 binding events within these categories as defined in (B) were compared with gene expression in TAMR and MCF7-WS8 cells as determined by RNA-seq. Genes with transcription start sites within ±10 kb of any FOXA1 binding event were assigned to at most one set of peaks, with the leftmost group having highest assignment priority. Analysis is based on (top) $\log_2$ fold change in gene expression; *$p < 0.01$ by Mann-Whitney test comparing the pair of boxplots marked by the horizontal line, or (bottom) relative enrichment of significantly differentially expressed genes relative to a control set of genes with similar average expression level but minimal fold change. Red bars denote genes significantly upregulated in TAMR versus MCF7-WS8. Blue bars denote genes significantly downregulated in TAMR versus MCF7-WS8. Asterisk denotes significant difference from 1, with p value cutoff of 0.01 using Fisher's exact test.

(D) PWM for motifs enriched within cluster 1 and cluster 2. All analysis is done on ±500 bp of sequence around FOXA1 peak call center.

(E) The top five distinct motifs as determined in (D) are presented and scanned against three different clusters of sites as defined in (B).

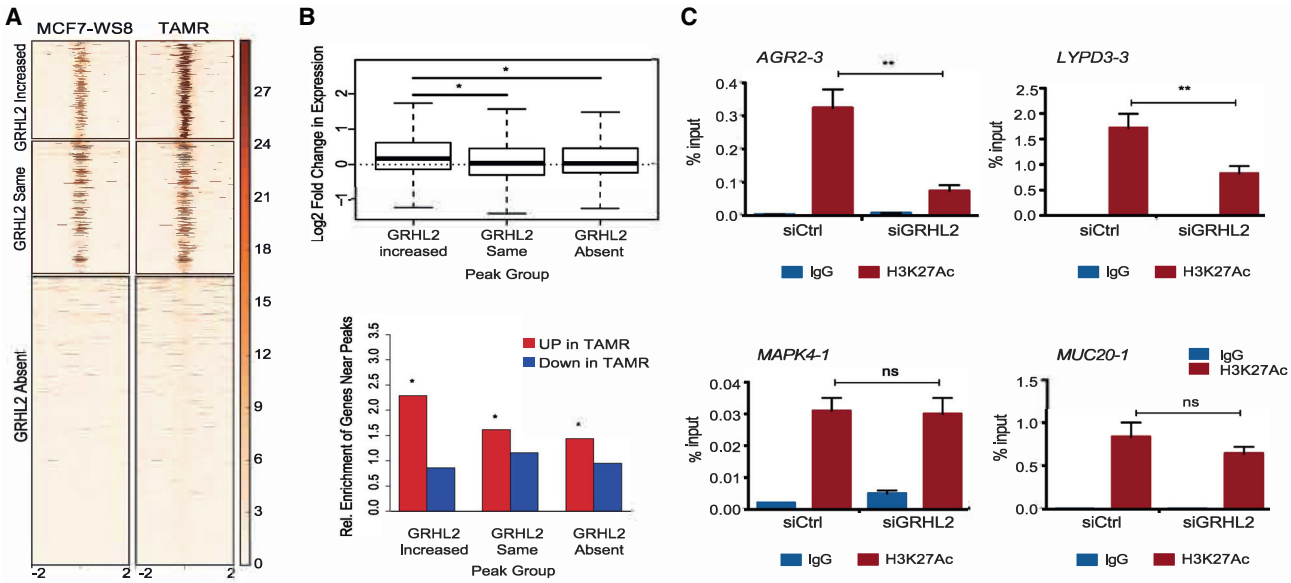


Figure 3. GRHL2 Interacts with FOXA1 at Subset of Active *cis*-Regulatory Elements

(A) Heatmap of GRHL2 ChIP-seq in a 4 kb window at sites in TAMR and MCF7-WS8 with increased FOXA1 binding intensity in TAMR relative to MCF7-WS8 as determined by ChIP-seq.

(B) FOXA1 binding events were categorized on the basis of whether the peak call determined by ChIP-seq was associated with a GRHL2 binding event that was significantly increased (GRHL2 increased), unchanged (GRHL2 same), or not detected (GRHL2 absent) in TAMR relative to MCF7-WS8. FOXA1 binding events within these categories were compared with gene expression in TAMR and MCF7-WS8 cells as determined by RNA-seq. Genes with transcription start sites within ±10 kb of any FOXA1 binding event were then assigned to at most one set of peaks, with the leftmost group having highest assignment priority. Analysis is based on (top) $\log_2$ fold change in gene expression; *$p < 0.01$ by Mann-Whitney test comparing the pair of boxplots marked by the horizontal line, or (bottom) relative enrichment of significantly differentially expressed genes relative to a control set of genes with similar average expression level but minimal fold change. Red bars denote genes significantly upregulated in TAMR versus MCF7-WS8. Blue bars denote genes significantly downregulated in TAMR versus MCF7-WS8. Asterisk denotes significant difference from 1, with p value cutoff of 0.01 using Fisher's exact test.

(C) The impact of GRHL2 knockdown on the status of H3K27 acetylation at FOXA1 enhancers within candidate GRHL2 target genes was assessed using ChIP-qPCR in TAMR cells. The bars represent the mean percentage input ± SD (three technical replicates). The experiment was repeated four times with similar results, and representative data are shown. Significance was determined using t test between siCtrl and siGRHL2. **$p < 0.05$; ns, not significant.

binding sites were not associated with GRHL2 binding events in either cell (GRHL2 absent) (Figure 3A). It was determined that the expression of genes within 10 kb of a FOXA1 gained site and that exhibited increased binding of GRHL2 in TAMR (GRHL2 increased) was more likely to be associated with increased gene transcription in TAMR relative to the MCF7-WS8 (Figure 3B). Together, these data suggest that GRHL2 may collaborate with FOXA1 in establishing new enhancers or lead to enhanced transcriptional activity at established enhancers in TAMR.

To explore the mechanisms underlying FOXA1/GRHL2 cooperativity, we evaluated the impact of GRHL2 knockdown on the status of H3K27 acetylation at FOXA1 enhancers within candidate GRHL2 target genes. For some enhancer sites, such as those associated with AGR2 or LYPD3, the absence of GRHL2 resulted in a significant decrease in H3K27Ac (Figure 3C); in contrast, H3K27 acetylation at enhancers associated with genes such as MAPK4 and MUC20 was unchanged following GRHL2 knockdown (Figure 3D). These data, while confirming a direct role for GRHL2 in regulating the deposition of H3K27Ac marks at a subset of candidate genes, indicate that cells may have redundant mechanisms to maintain acetylation in the absence of GRHL2.

GRHL2 Protein Expression Levels Are Associated with Decreased Responsiveness to Tamoxifen

Previously, it has been reported that elevated expression of FOXA1 mRNA is associated with decreased relapse-free survival (RFS) in patients with ERα-positive tumors receiving tamoxifen, but not in patients who are not on endocrine therapy (Fu et al., 2016). This association, however, was limited to the highest quartile of FOXA1 mRNA expression levels. Others have reported a significant association between elevated expression of GRHL2 mRNA and decreased time to recurrence and increased risk for metastasis in breast cancer patients (Xiang et al., 2012). Looking at gene expression data from 4,885 breast cancers, we were unable to identify a statistically significant difference in recurrence-free survival interval or distant metastasis survival interval for all comers or in the luminal breast cancer subgroups (Figure S2A). GRHL2 protein expression, but not mRNA levels, were increased in TAMR relative to MCF7-WS8 cells, and this was not influenced by ER modulation (Figures 4A and 4B). This encouraged us to examine the extent to which GRHL2 protein expression associates with clinical outcome in human breast cancers (a summary of patient characteristics appears in Table S2). A board-certified pathologist blinded to sample identification scored GRHL2 staining intensity on a

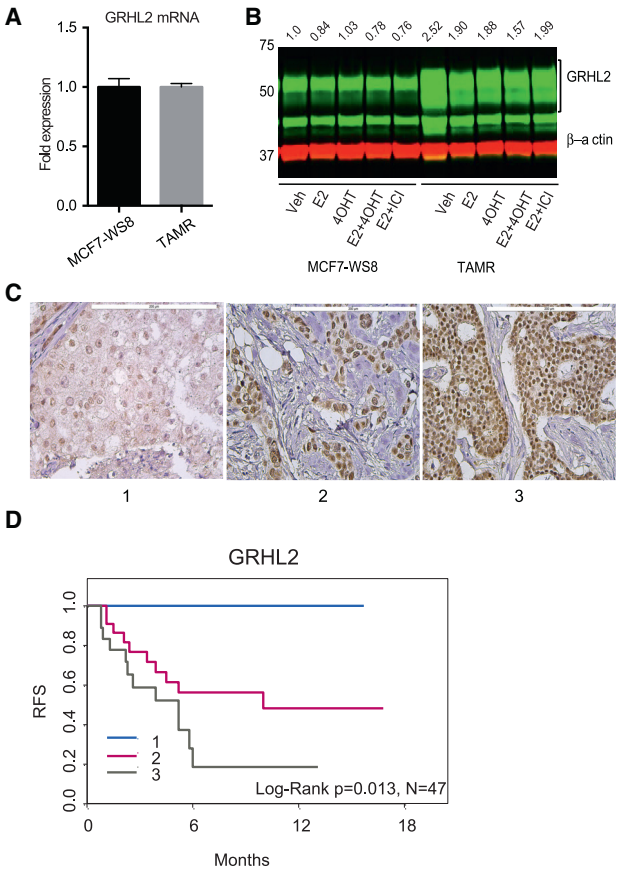


Figure 4. Increased GRHL2 Protein Expression Is Associated with Tamoxifen Resistance and Decreased Time to Recurrence

(A) Assessment of GRHL2 mRNA expression in MCF7-WS8 and TAMR cells using qPCR. The bars represent the fold change in CT values from three triplicate wells per condition, with error bars representing SEM. The experiment was repeated at least three times with similar results, and representative data are shown.

(B) Protein expression was assessed in MCF7-WS8 and TAMR cells treated with 10 nM E2, 100 nM 4OHT, or 100 nM fulvestrant as indicated using the indicated antibodies. Relative GRHL2 protein expression is indicated on top (normalized to β-actin, then to WS8 vehicle control). The expression of GRHL2 mRNA and protein was assessed at least three times with similar results, and representative data are shown.

(C) Representative immunohistochemistry examples (with scores 1, 2, and 3; 1 = low, 3 = high) from breast tumor tissue microarray stained with GRHL2 antibody. Scale bar, 200 μm.

(D) Kaplan-Meier estimator of time to recurrence (RFS, months) of tumors derived from patients with ER-positive disease, stratified on the basis of GRHL2 protein expression. Statistical significance was determined using log rank test, with p = 0.013. n = 47. Because of the small sample size and no events at level 1, hazard ratios are not estimable.

See also Figure S2.

scale ranging from 0 to 3; this score was then used to probe associations with T and N stage and time to recurrence (see representative staining in Figure 4C). GRHL2 protein expression did not significantly associate with T or N stage in this sample set. However, when assessing all comers (independent of hormone receptor status), there was a strong trend toward a shorter time to recurrence observed with increasing expression; however, this trend did not reach significance (p = 0.08) (Figure S2B). In those patients with ER-positive disease, it was determined that patients with the highest GRHL2 staining intensity (3) have decreased time to recurrence relative to those with lower GRHL2 staining intensity (1 or 2) (log rank p = 0.013) (Figure 4D). Together, these data implicate GRHL2 as an important regulator of tumor progression in ER-positive luminal breast cancers.

Cells Derived from Tamoxifen-Resistant Tumors Demonstrate Enhanced Activation of Pathways Associated with Aggressive Cancer Phenotypes

The impact of GRHL2 expression on processes of pathological importance in TAMR cells was next examined. Knockdown of GRHL2 expression using three independent small interfering RNAs (siRNAs) inhibited cell proliferation and decreased the migratory activity of the TAMR cells, and these effects correlate with the degree of GRHL2 knockdown (Figures 5A, 5B, and S3A). Pathway analysis of RNA-seq data generated from TAMR cells treated with control siRNA or siRNAs directed against GRHL2 was next undertaken. Using the EnrichR algorithm to mine the ChEA, ENCODE, and ChEA consensus TFs from ChIP X datasets, several interesting findings emerged (Kuleshov et al., 2016). First, the family of genes whose expression was attenuated by GRHL2 knockdown in TAMR (sig_genes_dn) is enriched for genes that are associated with forkhead binding events (Figure 5C). Second, both sets of genes (those that are increased and those that are decreased following GRHL2 knockdown) are enriched for genes associated with ER binding events, suggesting that GRHL2 may also play a role in both activating and repressing ER activity. Third, there is also a striking inverse correlation between genes which decrease following GRHL2 knockdown and genes that are associated with lysine demethylase 2B (KDM2B) binding events. Finally, analysis of the ChEA dataset reveals that there is an overlap between the GRHL2 target genes we identified in TAMR cells and genes described previously to be regulated by TCF3 (E47) and SALL4, the expression of which are associated with early-stage breast cancer, SOX2, and NFE2L2, both of which have been strongly linked to tamoxifen resistance, and the androgen receptor (AR), the activity of which is closely associated with FOXA1 in the setting of prostate cancer (Arif et al., 2015; Jeter et al., 2016; Kim et al., 2008; Kobayashi et al., 2011; Slyper et al., 2012). Thus, in addition to FOXA1, GRHL2 may also collaborate with additional factors in TAMR.

In parallel to the *in vitro* studies outlined above we assessed the expression of genes that were differentially regulated in TAMR tumors relative to MCF7-WS8 tumors using the gene classification schemes established in the DHS and ChIP-seq studies (FOXA1 increased, FOXA1 same, and FOXA1 decreased) (Figure 1). When comparing the relative expression of genes stratified on the basis of FOXA1 binding differences alone (as in Figure 1D), we observe very little enrichment across the three groups (Figure S3B). A clearer picture emerged by focusing on those genes associated with increased FOXA1 binding and increased marks of histone activation (see Figures 2B and 2C; Figure S3C), and even more significant biology was revealed when we examined gene stratification on the basis of the

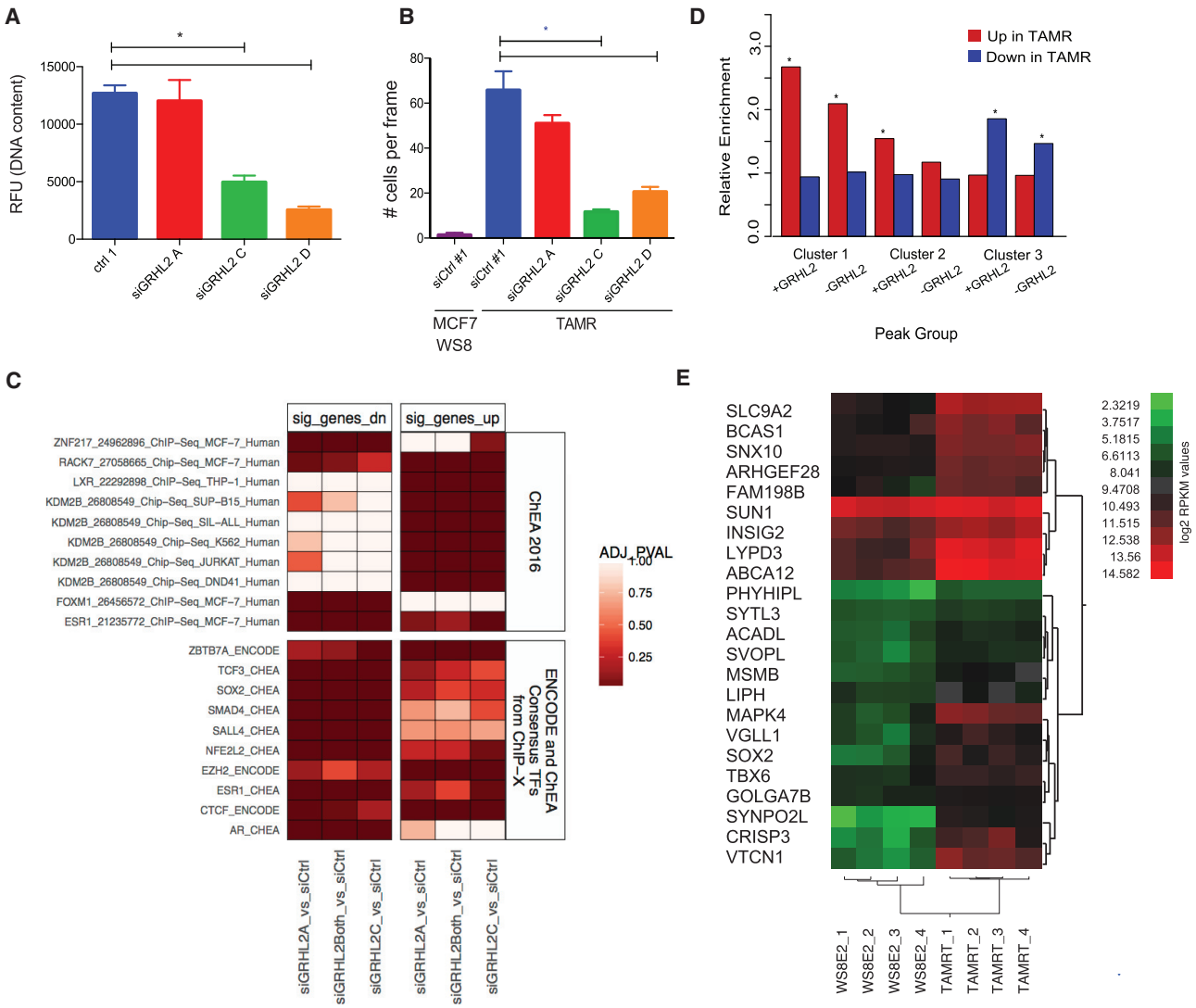


Figure 5. GRHL2 Regulates Proliferation and Migration, via Interaction with Several Candidate Transcription Factors

(A and B) TAMR cells were transfected with siCtrl or three unique siRNAs targeting GRHL2 and monitored for cell proliferation (A) or migration (B). MCF7-WS8 treated with siCtrl was included for comparison. The bars in (A) represent the mean relative fluorescence intensity of triplicate wells per condition. The bars in (B) represent the average migrated cells per field of view counting six fields of view per transwell and two transwells per condition. Error bars are SEM. The experiments were repeated three times with similar results, and representative data are shown.

(C) EnrichR analysis of RNA-seq with two different siRNAs to GRHL2 compared to siCtrl in TAMR cells.

(D) Relative enrichment of genes differentially expressed in TAMR relative to MCF7-WS8 within 10 kb of FOXA1 increased binding events subdivided on the basis of histone marks (as in Figure 2B) and further subdivided on the basis of presence (+GRHL2) or absence (−GRHL2) of GRHL2 binding event. Asterisk denotes significant difference from 1, with p value cutoff of 0.01 using Fisher's exact test.

(E) Heatmap indicating relative mRNA expression of genes in TAMR xenograft tumors treated with tamoxifen relative to MCF7-WS8 xenograft tumors treated with estradiol, which meet the following criteria: (1) within a 10 kb window of a FOXA1 binding event that is increased in TAMR cells associated with a significant increase in H3K27Ac relative to MCF7-WS8 and overlaps with a GRHL2 binding event; (2) the expression of which is increased in TAMR versus MCF7-WS8; and (3) the expression is decreased with siGRHL2 relative to control siRNA in TAMR cells on the basis of RNA-seq.

See also Figure S3.

presence or absence of a GRHL2 binding event in TAMR cells at the FOXA1 gained sites (Figure 5D). Taken together, the results of these studies indicate that the transcriptional activity of GRHL2 increases as the cells/tumors develop resistance to tamoxifen and highlight the potential utility of inhibiting its activity or a downstream target(s) in late-stage disease.

LYPD3 Is Regulated Downstream of GRHL2

GRHL2 does not exhibit any features that would suggest that it could be easily targeted with small molecules, and thus we explored the utility of exploiting proteins/processes downstream of GRHL2 for new drug development. To identify such targets, we mined the datasets generated in this study for genes (1)

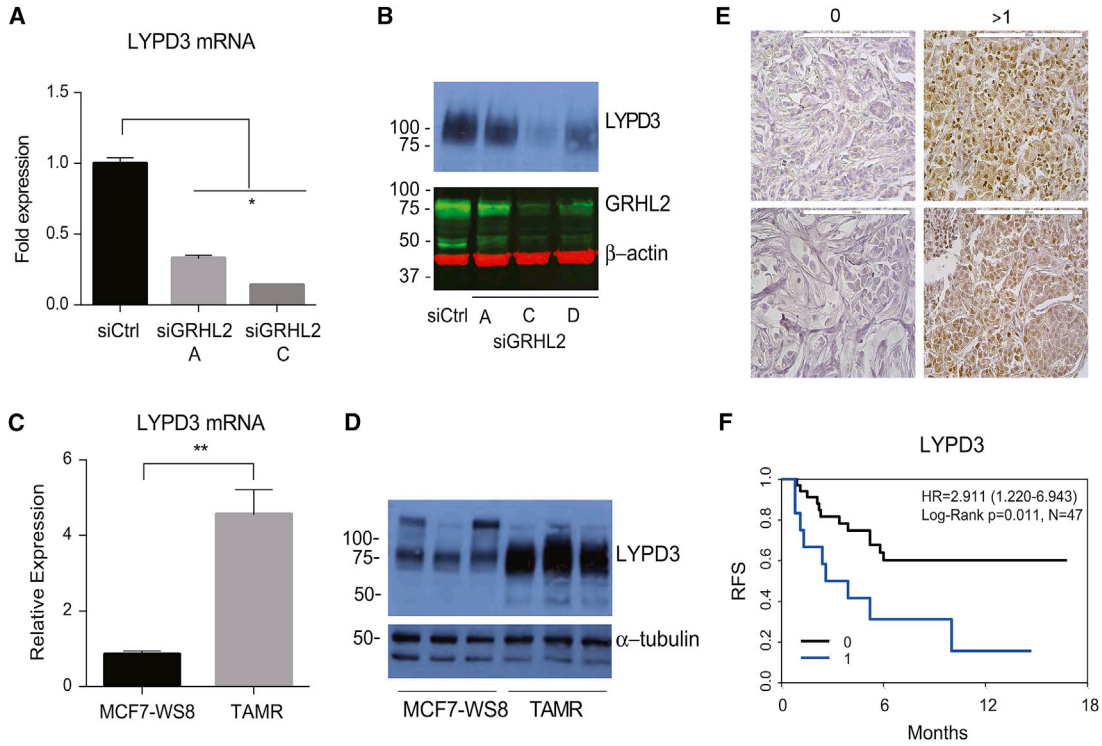


Figure 6. LYPD3 is Regulated by GRHL2

(A and B) LYPD3 (A) mRNA and (B) protein expression in TAMR cells following GRHL2 knockdown. RNA expression was assessed using qPCR, with the bars representing the fold change in CT values from three triplicate wells per condition, with error bars representing SEM. Protein was assessed using western blot using the indicated antibodies. Significance was determined using one-way ANOVA with Bonferroni's test. *$p < 0.05$. This experiment was repeated three times with similar results, and representative data are shown.

(C and D) LYPD3 mRNA (C) and protein expression (D) was assessed in MCF7-WS8 and TAMR tumors; n = 3 xenograft tumors per group. Error bars are SEM.

(E) Representative immunohistochemistry examples from breast tumor tissue microarray stained with LYPD3 antibody. Scale bar, 200 μm.

(F) Kaplan Meier-estimator of time to recurrence (RFS, months) of tumors derived from patients with ER-positive disease, stratified on the basis of LYPD3 protein expression (0 = no staining, 1 = positive staining). Statistical significance was determined using log rank test, with p = 0.011, n = 47. Hazard ratio (HR) was determined using univariate Cox proportional-hazards model (p = 0.016).

See also Figures S4 and S5.

that were associated with a FOXA1 binding event (within 10 kb) that is increased in TAMR relative to MCF7-WS8, exhibits increased H3K27Ac, and overlaps with a GRHL2 binding event; (2) whose mRNA expression in TAMR cells was increased relative to MCF7-WS8 cells; and (3) whose expression was decreased upon siRNA-mediated knockdown of GRHL2 relative to siCtrl in TAMR cells. The expression of genes which met these criteria were further evaluated for their expression in MCF7-WS8 and TAMR tumor models (Figure 5E). This informative analysis led to the identification of several genes, the products of which are likely to be involved in breast cancer cell biology. Among the potential candidate genes identified, one candidate, LYPD3 (C4.4a), was of particular interest because (1) it was expressed at the cell surface, thus making it a desirable pharmacological target, and (2) we had previously shown that expression of its ligand, Anterior Grade 2 (AGR2), was highly induced by both estradiol and tamoxifen in breast cancer cells and in tumor xenografts (Wright et al., 2014).

LYPD3 is a GPI-anchored membrane protein that belongs to the Ly6 family of receptors and has been shown to be overexpressed in several human malignancies (Hansen et al., 2007; Seiter et al., 2001; Smith et al., 2001). Functionally, LYPD3 has been linked to increased invasion and metastasis, mediated by regulation of the focal adhesion pathway via interactions with cell surface integrins (Ngora et al., 2012). Little is known about the signaling events up or downstream of LYPD3 activation, although it has recently been shown to serve as a functional cell surface receptor for AGR2, a protein whose expression we have shown to be increased in the setting of tamoxifen resistance (Arumugam et al., 2015; Wright et al., 2014). Cumulatively these findings suggest that GRHL2-dependent upregulation of LYPD3 expression, and signaling events downstream of LYDP3/AGR2, may be critical to maintaining the growth of TAMR cells and tumors. In support of this hypothesis, we demonstrated that (1) the expression of LYPD3 in TAMR cells absolutely requires GRHL2 (Figures 6A and 6B). The same dependency was confirmed in additional breast cancer cell lines (Figure S4); (2) there are several robust GRHL2 chromatin binding sites downstream, within, and around the LYPD3 promoter in both TAMR and MCF7-WS8 cells (Figure S5A); and (3) LYPD3 protein and mRNA expression are elevated in TAMR tumors (Figures 6C and 6D). Of note, although the mRNA and

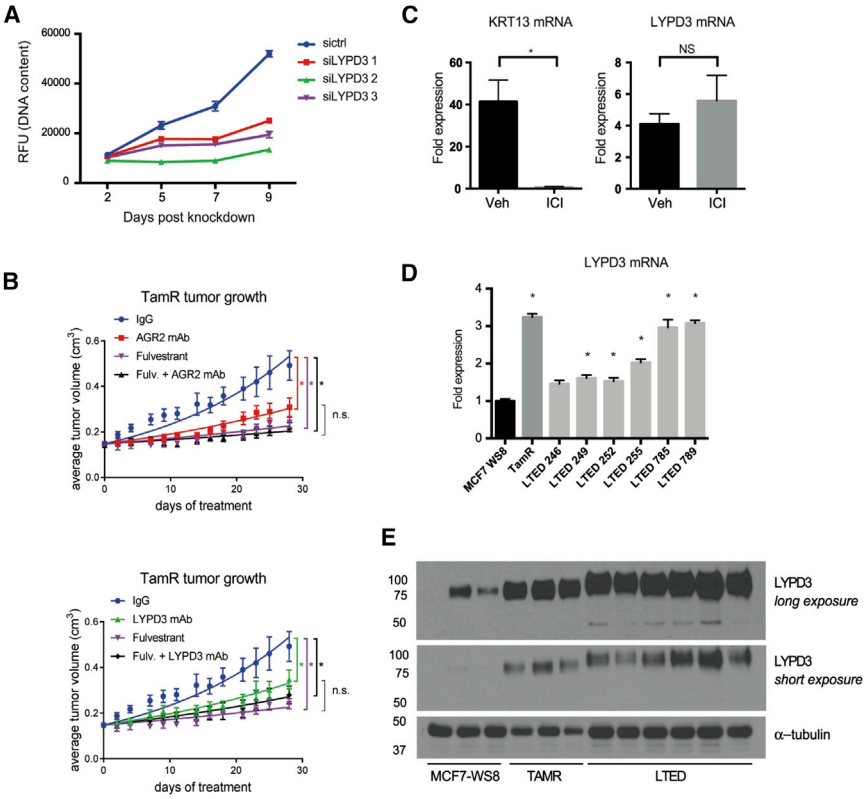


Figure 7. LYPD3 as a Candidate Drug Target for the Treatment of Aggressive Luminal Cancer

(A) TAMR cells were transfected with siCtrl or three unique siRNA sequences targeting LYPD3 and monitored for 9 days. Individual points on the curve represent the mean relative fluorescent intensity of triplicate wells per condition on that day. Error bars calculated as SEM. The experiment was repeated three times with similar results, and representative data are shown.

(B) Tamoxifen-treated J:nu mice bearing TamR xenograft tumors were randomized to treatment with 45 mg/kg IgG, 15 mg/kg anti-AGR2 (top), or 45 mg/kg anti-LYPD3 (bottom) antibodies intraperitoneally (i.p.) twice weekly, with groups further subdivided to receive subcutaneous (s.c.) injection of corn oil or 25 mg/kg fulvestrant. To facilitate interpretation, data for anti-AGR2 and anti-LYPD3 are presented in separate graphs, with controls (IgG and fulvestrant administered alone) included in both graphs. Data presented indicate the average tumor volume for each group (mean ± SEM) at each time point of tumor measurement. Two-way ANOVA analysis followed by Bonferroni multiple-comparison test detected significant differences between the IgG control and all treatment groups between days 14 and 28 (*$p < 0.05$).

(C) TAMR tumors from mice treated with corn oil (vehicle) or fulvestrant were assessed for mRNA expression of KRT13 and LYPD3; n = 9 xenograft tumors per group. Data plotted are mean fold change ± SEM.

(D and E) LYPD3 mRNA (D) and protein expression (E) was assessed in LTED tumors and compared with representative samples of MCF7-WS8 and TAMR tumors. Each bar indicates an independent biological replicate and plotted as mean fold change ± SD (three technical replicates). Asterisk indicates samples with mRNA expression significantly different ($p < 0.05$) than a representative MCF7-WS8 control tumor sample.

See also Figures S6 and S7.

protein levels are robustly increased in TAMR tumors, cell line data highlight a more profound increase in protein levels, while mRNA differences are not as significant (Figures S5B and S5C). We also determined that LYPD3 is heavily glycosylated in TAMR cells (and tumors) and are exploring the possibility that this event contributes to its pathobiological actions. Collectively, these results indicate that LYPD3 is a candidate target protein whose function is defined by a variety of processes enhanced in TAMR and whose expression is directly regulated by GRHL2.

LYPD3 Is a Functional Component of an ER-Independent, Collateral Pathway, the Targeting of Which Inhibits the Growth of Endocrine Therapy Resistant Tumors

To investigate the potential clinical relevance of LYPD3, we first explored whether the expression of LYPD3 mRNA was associated with outcomes in patients with breast tumors. Using the TCGA database, we found that recurrence-free survival did not differ between patients with high or low LYPD3 mRNA expression (Figure S5D). This was true for all breast cancers and for ER+ tumors. LYPD3 protein expression in tumors was more informative (Figure 6E) where higher LYPD3 protein levels were found to associate with decreased time to recurrence in ER+ breast cancers (log rank $p = 0.011$) (Figure 6F). When all comers are considered, higher LYPD3 expression also trends toward decreased time to recurrence (Figure S5E).

Our data suggest GRHL2/LYPD3 and AGR2 are components of a signaling pathway whose activity increases as tumors develop resistance to endocrine therapy. Furthermore, our previously published data demonstrated that AGR2 is essential for the viability of TAMR cells (Wright et al., 2014). Consistent with these results, we observed that siRNA-mediated knockdown of LYPD3 compromises the growth of TAMR cells (Figure 7A; Figure S6A). Similarly, knockdown of LYPD3 also reduced the growth of another model of tamoxifen resistance, HCC1428-TamR breast cancer cells (Figure S6B) (Guest et al., 2016). Thus, we evaluated the impact of monoclonal antibodies targeting either AGR2 or LYPD3 on TAMR xenograft tumor growth. Treatment with either anti-AGR2 or anti-LYPD3 resulted in a significant decrease in the growth of tamoxifen-resistant tumors (Figure 7B). Of note, the treatment of tumors with either AGR2 or LYPD3 targeted agents resulted in similar growth inhibition. Furthermore, treating tumors with anti-AGR2 and anti-LYPD3 antibodies in combination was equivalent to either alone, a likely reflection of the epistatic nature of the two targets (Figure S6C). Of

clinical relevance is the observation that these antibody treatments are as effective as the selective estrogen receptor downregulator (SERD) fulvestrant (ICI) (modeled in mice to human exposure levels) although there was no benefit in this model to combining any two of the three agents (Figure 7B). Importantly, the expression of LYPD3 is not influenced by treatment of TAMR tumors with the SERD fulvestrant (Figure 7C).

We demonstrated that AGR2 and LYPD3 expression were both dramatically upregulated in a xenograft model of long-term estrogen deprivation (a surrogate for AI activity) (Figures 7D, 7E, and S6D). Treatment of these tumors with anti-LYPD3 antibody led to a significant reduction in the time to tumor doubling (growth velocity) (Figure S6E).

Finally, to reinforce the potential clinical significance of these tumor data we demonstrated that GRHL2 and LYPD3 expression was maintained in tumors from patients who had progressed while on endocrine therapy (Figure S7).

DISCUSSION

Forkhead family members, including FOXA1, have previously been described as critical regulators of cell fate and identity where they are engaged in determining the transcriptional landscape in cells. However, dysregulated FOXA1 activity also contributes to the development of tamoxifen resistance in luminal breast cancer (Hurtado et al., 2011; Ross-Innes et al., 2012; Sérandour et al., 2011). Specifically, it had been shown that FOXA1 binding profiles were altered in cell models of tamoxifen resistance relative to tamoxifen-sensitive models. What is not apparent from previous work is (1) how FOXA1 attains this altered activity in the setting of resistance and (2) how these findings can be translated into new treatment or intervention strategies.

In this study, we focused on those FOXA1 binding events that were increased in TAMR relative to MCF7-WS8 cells and used markers of enhancer activity to assess their likely functionality. This allowed us to define enhancers that were contained within latent chromatin in MCF7-WS8 cells but within active chromatin in TAMR. This pattern is consistent with the classic view of FOXA1 function as a pioneer factor (Sérandour et al., 2011). However, we also observed that some of the gained FOXA1 enhancers were in a "poised" state in MCF7-WS8 cells before adopting the characteristics of an activated state in TAMR. This transition is more indicative of a "signal-dependent" change, whereby an extracellular signaling event activates a transcription factor facilitating its recruitment to a genomic site where a lineage-dependent transcription factor/pioneer transcription factor is already present (Heinz et al., 2010). We also identified a third set of gained FOXA1 binding sites, wherein no changes in histone marks were identified. These binding events may be the result of nonspecific sampling of the genome as proposed previously for other transcription factors (Coons et al., 2017). Examination of the architecture of sites exhibiting enhanced FOXA1 binding and an "active histone signature" in TAMR versus MCF7-WS8 cells led to the discovery that GRHL2 was a likely collaborator of FOXA1, a result confirmed using biochemical studies.

GRHL2 as a Key Determinant of the Functional Cistrome in Luminal Breast Cancer

GRHL2 has been shown to serve as a key determinant of keratinocyte differentiation and lung epithelial morphogenesis and suggested also to be a lineage-determining factor in breast epithelial cells (Xiang et al., 2012). A role for GRHL2 in determining epithelial identity in the 4T1 mouse model of mammary carcinoma and in human breast cancer cell line models has also been suggested (Cieply et al., 2012; Lønning, 2002; Werner et al., 2013; Xiang et al., 2012) and reviewed (Frisch et al., 2017). These published studies indicate that the expression of GRHL2 suppresses epithelial-to-mesenchymal transition (EMT) in breast cancer cells and that GRHL2 directly or indirectly regulates a broad range of epithelial genes. As such, not only does GRHL2 participate with and potentially regulate FOXA1 binding events, but it is also required to maintain luminal breast cell identity.

Given that we have shown that GRHL2 interacts at sites within chromatin adjacent to FOXA1, the question remains as to how GRHL2 binding is regulated. We demonstrated that GRHL2 levels, although not differing significantly at the transcript level, are increased at the level of protein expression in the setting of tamoxifen resistance. Several putative post-translational modifications in GRHL2 have been identified by high-throughput mass spectroscopy, whether one or more of these modifications contribute to GRHL2 stability remains to be determined. It was also of significance that we were able to show that elevated GRHL2 protein expression is strongly associated with a decreased time to recurrence in patients with ERα-positive breast cancer. Although striking, it is likely that in addition to increased GRHL2 expression that increased activity of this protein may also contribute to disease pathobiology.

The FOXA1 gained binding events that we observe in TAMR can be subcategorized in three different ways: those in which FOXA1 and GRHL2 are increased, those in which FOXA1 is increased and GRHL2 is unchanged, and those in which FOXA1 is increased but GRHL2 is not present. Examining these subsets separately suggests several different mechanisms which define GRHL2 activity. At sites where both FOXA1 and GRHL2 are low in the setting of tamoxifen sensitivity but both are increased in the setting of resistance, it is possible that the two factors (FOXA1 and GRHL2) respond to a similar differentiation cue and result in collaborative binding. At sites where FOXA1 increases in TAMR relative to MCF7-WS8 and GRHL2 is present, but does not increase further in the setting of resistance, suggests a potential pioneer-like role of GRHL2, whereby GRHL2 pre-marks a site and stabilizes nucleosomes, facilitating FOXA1 binding in the setting of resistance. Indeed, there is a subset of DHSs bound by GRHL2 and not by FOXA1. The presence of these types of sites suggests that GRHL2 could be the initiating transcription factor. Along these lines, *Grh*, the *Drosophila* homolog of GRHL2, has been suggested to have intrinsic nucleosome binding and displacement ability (Nevil et al., 2017). As such, GRHL2 may have the same ability as FOXA1 and other pioneer factors to non-specifically bind nucleosomes and scan chromatin. Interestingly, our bioinformatic studies suggest that in addition to FOXA1, GRHL2 may interact functionally with TCF3 (E47), SALL4 SOX2,

NFE2L2, and the nuclear receptors ER and AR (Arif et al., 2015; Jeter et al., 2016; Kim et al., 2008; Kobayashi et al., 2011; Slyper et al., 2012).

Analysis of the FOXA1 GRHL2 Cistrome Reveals Targetable Pathways in Breast Cancer

Previously we reported that AGR2 expression is increased by either estradiol or tamoxifen and tamoxifen-resistant tumors, even in the absence of treatment, are found to express higher levels of AGR2 than endocrine therapy-sensitive tumors. This was an important finding, as high AGR2 expression is a predictor of poor prognosis and decreased response to endocrine therapy in patients with luminal breast cancer (Wright et al., 2014). Until recently, however, it was unclear as to how AGR2 influences tumor biology. Of relevance to our present work, however, Arumugam et al. (2015), working in models of pancreatic cancer, determined that LYPD3 is the putative receptor for AGR2.

Our studies led to the observation that breast tumors that are resistant to tamoxifen (or to estrogen deprivation) have elevated levels of LYPD3 protein. Expression of this protein was also predictive of a poor response to endocrine therapy in patients. We determined that LYPD3 is a direct transcriptional target of GRHL2. LYPD3 has been shown to interact with extracellular matrix components such as lamininins and galectins and play a role in cell-cell attachment. It has also been shown to be specifically expressed on the leading edge of invasive tumors and to be present in exosomes released from metastatic cells (Ngora et al., 2012). As such, it stands to reason that part of the mechanism by which mammary tumor cells demonstrate decreased migratory ability following GRHL2 knockdown is via decreased expression of LYPD3, which functionally results in disruption of cell contacts.

Herein we have demonstrated that targeting LYPD3 directly as well as its ligand AGR2 with specific inactivating antibodies can effectively decrease tumor growth in two different models of endocrine therapy-resistant breast cancer. Given that LYPD3 expression (1) is maintained in the presence of anti-estrogen therapy, (2) is increased in models of tamoxifen and aromatase resistance, and (3) serves as a marker of poor prognosis in breast cancer patients, we believe that it will be a useful target in advanced disease. Humanized antibodies directed against AGR2 and LYPD3 are now in late-stage preclinical development and are expected to be in clinical trials in the near future. Furthermore, an anti-LYPD3 antibody-auristatin conjugate is currently in clinical trials for squamous cell carcinoma of the lung, although considering these data it may be a useful intervention also in breast cancer (Willuda et al., 2017). Continued exploitation of additional GRHL2 targets is the subject of ongoing studies in our laboratory.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.09.032.

ACKNOWLEDGMENTS

This study was supported by NIH grant DK048807 to D.P.M.; Susan G. Komen grants SAC180085 to D.P.M. and SAB100008 to M.B.; NRSA 5F30CA183395-03 to K.J.C.; and Viba Therapeutics, Novartis, and funds from the Royal Marsden – Institute of Cancer Research NIHR BRC to M.D. We thank Alexias Safi for her help with processing DNaseq, Will Thompson and the Duke Proteomics Core for performing the mass spectrometry analysis, and Andrew Dodson for assistance with TMA.

AUTHOR CONTRIBUTIONS

J.D.N., C.Y.C., and D.P.M. conceived and supervised the project. K.J.C., C.Y.C., J.S.J., L.E., G.E.C., T.M.W., D.A.R., D.P.R., T.J., G.B., M.B., T.H., A.H., G.L.G., and D.P.M. contributed to data analysis. J.T.B. and C.L. developed the inactivating antibodies to AGR2 and LYPD3. K.J.C., J.S.J., S.W., G.E.C., T.M.W., D.A.R., D.P.R., T.K.D., C.Y.C., R.B., T.W., A.Y., K.T., Y.B., M.L., and L.P. contributed to the experiments. J.S.J. implemented gene analytics web portal. J.R.M., M.D., and S.D. provided tissue microarray. L.-A.M. and S.P. provided the HCC1428-TamR cell line. K.J.C., J.S.J., C.Y.C., and D.P.M. wrote the manuscript with input from other authors.

DECLARATION OF INTERESTS

J.T.B. and C.L. are cofounders of Viba Therapeutics, a company that is developing strategies to inhibit LYPD3/AGR2 signaling for a variety of cancers. Subsequent to the work defining the roles of GRHL2/LYPD3 in breast cancer, D.P.M. became a member of the scientific advisory board of Viba Therapeutics. D.P.M., J.T.B., and C.L. have equity in Viba Therapeutics. M.B. receives sponsored research support from Novartis and serves on the scientific advisory board of Kronos Bio.

Received: December 13, 2017
Revised: July 2, 2019
Accepted: September 12, 2019
Published: October 22, 2019

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Arif, K., Hussain, I., Rea, C., and El-Sheemy, M. (2015). The role of Nanog expression in tamoxifen-resistant breast cancer cells. OncoTargets Ther. *8*, 1327–1334.

Arumugam, T., Deng, D., Bover, L., Wang, H., Logsdon, C.D., and Ramachandran, V. (2015). New blocking antibodies against novel AGR2-C4.4A pathway reduce growth and metastasis of pancreatic tumors and increase survival in mice. Mol. Cancer Ther. *14*, 941–951.

Boyle, A.P., Guinney, J., Crawford, G.E., and Furey, T.S. (2008). F-Seq: a feature density estimator for high-throughput sequence tags. Bioinformatics *24*, 2537–2538.

Britton, D.J., Hutcheson, I.R., Knowlden, J.M., Barrow, D., Giles, M., McClelland, R.A., Gee, J.M.W., and Nicholson, R.I. (2006). Bidirectional cross talk between ERalpha and EGFR signalling pathways regulates tamoxifen-resistant growth. Breast Cancer Res. Treat. *96*, 131–146.

Brodie, A. (2002). Aromatase inhibitors in breast cancer. Trends Endocrinol. Metab. *13*, 61–65.

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

Cieply, B., Riley, P., 4th, Pifer, P.M., Widmeyer, J., Addison, J.B., Ivanov, A.V., Denvir, J., and Frisch, S.M. (2012). Suppression of the epithelial-mesenchymal transition by Grainyhead-like-2. Cancer Res. *72*, 2440–2453.

Connor, C.E., Norris, J.D., Broadwater, G., Willson, T.M., Gottardis, M.M., Dewhirst, M.W., and McDonnell, D.P. (2001). Circumventing tamoxifen resistance in breast cancers using antiestrogens that induce unique conformational changes in the estrogen receptor. Cancer Res. *61*, 2917–2922.

Coons, L.A., Hewitt, S.C., Burkholder, A.B., McDonnell, D.P., and Korach, K.S. (2017). DNA sequence constraints define functionally active steroid nuclear receptor binding sites in chromatin. Endocrinology *158*, 3212–3234.

Delhomme, N., Padioleau, I., Furlong, E.E., and Steinmetz, L.M. (2012). easyRNASeq: a bioconductor package for processing RNA-seq data. Bioinformatics *28*, 2532–2533.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dowsett, M., and Howell, A. (2002). Breast cancer: aromatase inhibitors take on tamoxifen. Nat. Med. *8*, 1341–1344.

Drury, S.C., Detre, S., Leary, A., Salter, J., Reis-Filho, J., Barbashina, V., Marchio, C., Lopez-Knowles, E., Ghazoui, Z., Habben, K., et al. (2011). Changes in breast cancer biomarkers in the IGF1R/PI3K pathway in recurrent breast cancer after tamoxifen treatment. Endocr. Relat. Cancer *18*, 565–577.

Fisher, B., Costantino, J.P., Wickerham, D.L., Redmond, C.K., Kavanah, M., Cronin, W.M., Vogel, V., Robidoux, A., Dimitrov, N., Atkins, J., et al. (1998). Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. J. Natl. Cancer Inst. *90*, 1371–1388.

Fisher, B., Dignam, J., Bryant, J., and Wolmark, N. (2001). Five versus more than five years of tamoxifen for lymph node-negative breast cancer: updated findings from the National Surgical Adjuvant Breast and Bowel Project B-14 randomized trial. J. Natl. Cancer Inst. *93*, 684–690.

Frisch, S.M., Farris, J.C., and Pifer, P.M. (2017). Roles of Grainyhead-like transcription factors in cancer. Oncogene *36*, 6067–6073.

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

Gottardis, M.M., and Jordan, V.C. (1988). Development of tamoxifen-stimulated growth of MCF-7 tumors in athymic mice after long-term antiestrogen administration. Cancer Res. *48*, 5183–5187.

Guest, S.K., Ribas, R., Pancholi, S., Nikitorowicz-Buniak, J., Simigdala, N., Dowsett, M., Johnston, S.R., and Martin, L.A. (2016). Src is a potential therapeutic target in endocrine-resistant breast cancer exhibiting low estrogen receptor-mediated transactivation. PLoS ONE *11*, e0157397.

Haibe-Kains, B., Desmedt, C., Loi, S., Culhane, A.C., Bontempi, G., Quackenbush, J., and Sotiriou, C. (2012). A three-gene model to robustly identify breast cancer molecular subtypes. J. Natl. Cancer Inst. *104*, 311–325.

Hansen, L.V., Skov, B.G., Ploug, M., and Pappot, H. (2007). Tumour cell expression of C4.4A, a structural homologue of the urokinase receptor, correlates with poor prognosis in non-small cell lung cancer. Lung Cancer *58*, 260–266.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

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

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

Jeselsohn, R., Bergholz, J.S., Pun, M., Cornwell, M., Liu, W., Nardone, A., Xiao, T., Li, W., Qiu, X., Buchwalter, G., et al. (2018). Allele-specific chromatin recruitment and therapeutic vulnerabilities of ESR1 activating mutations. Cancer Cell *33*, 173–186.e5.

Jeter, C.R., Liu, B., Lu, Y., Chao, H.-P., Zhang, D., Liu, X., Chen, X., Li, Q., Rycaj, K., Calhoun-Davis, T., et al. (2016). NANOG reprograms prostate cancer cells to castration resistance via dynamically repressing and engaging the AR/FOXA1 signaling axis. Cell Discov. *2*, 16041.

Jiang, H., Lei, R., Ding, S.-W., and Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics *15*, 182.

Jozwik, K.M., Chernukhin, I., Serandour, A.A., Nagarajan, S., and Carroll, J.S. (2016). FOXA1 directs H3K4 monomethylation at enhancers via recruitment of the methyltransferase MLL3 (Cold Spring Harbor Laboratory).

Kim, S.K., Yang, J.W., Kim, M.R., Roh, S.H., Kim, H.G., Lee, K.Y., Jeong, H.G., and Kang, K.W. (2008). Increased expression of Nrf2/ARE-dependent antioxidant proteins in tamoxifen-resistant breast cancer cells. Free Radic. Biol. Med. *45*, 537–546.

Knowlden, J.M., Hutcheson, I.R., Barrow, D., Gee, J.M.W., and Nicholson, R.I. (2005). Insulin-like growth factor-I receptor signaling in tamoxifen-resistant breast cancer: a supporting role to the epidermal growth factor receptor. Endocrinology *146*, 4609–4618.

Kobayashi, D., Kuribayshi, K., Tanaka, M., and Watanabe, N. (2011). SALL4 is essential for cancer cell proliferation and is overexpressed at early clinical stages in breast cancer. Int. J. Oncol. *38*, 933–939.

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

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44* (*W1*), W90-7.

Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E., and Storey, J.D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics *28*, 882–883.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv, arXiv:1303.3997(q-bio). https://arxiv.org/abs/1303.3997.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Lin, A., Hu, Q., Li, C., Xing, Z., Ma, G., Wang, C., Li, J., Ye, Y., Yao, J., Liang, K., et al. (2017). The LINK-A lncRNA interacts with PtdIns(3,4,5)P_3 to hyperactivate AKT and confer resistance to AKT inhibitors. Nat. Cell Biol. *19*, 238–251.

Lønning, P.E. (2002). Aromatase inhibitors and inactivators for breast cancer therapy. Drugs Aging *19*, 277–298.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

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

Lupien, M., Meyer, C.A., Bailey, S.T., Eeckhoute, J., Cook, J., Westerling, T., Zhang, X., Carroll, J.S., Rhodes, D.R., Liu, X.S., and Brown, M. (2010). Growth factor stimulation induces a distinct ER(alpha) cistrome underlying breast cancer endocrine resistance. Genes Dev. *24*, 2219–2227.

Martz, C.A., Ottina, K.A., Singleton, K.R., Jasper, J.S., Wardell, S.E., Peraza-Penton, A., Anderson, G.R., Winter, P.S., Wang, T., Alley, H.M., et al. (2014). Systematic identification of signaling pathways with potential to confer anticancer drug resistance. Sci. Signal. *7*, ra121.

Massarweh, S., Osborne, C.K., Creighton, C.J., Qin, L., Tsimelzon, A., Huang, S., Weiss, H., Rimawi, M., and Schiff, R. (2008). Tamoxifen resistance in breast tumors is driven by growth factor receptor signaling with repression of classic estrogen receptor genomic function. Cancer Res. *68*, 826–833.

McDonnell, D.P., Wardell, S.E., and Norris, J.D. (2015). Oral selective estrogen receptor downregulators (SERDs), a breakthrough endocrine therapy for breast cancer. J. Med. Chem. *58*, 4883–4887.

Mokbel, K. (2002). The evolving role of aromatase inhibitors in breast cancer. Int. J. Clin. Oncol. *7*, 279–283.

Nelson, E.R., Wardell, S.E., Jasper, J.S., Park, S., Suchindran, S., Howe, M.K., Carver, N.J., Pillai, R.V., Sullivan, P.M., Sondhi, V., et al. (2013). 27-Hydroxycholesterol links hypercholesterolemia and breast cancer pathophysiology. Science *342*, 1094–1098.

Nevil, M., Bondra, E.R., Schulz, K.N., Kaplan, T., and Harrison, M.M. (2017). Stable binding of the conserved transcription factor Grainy Head to its target genes throughout Drosophila melanogaster development. Genetics *205*, 605–620.

Ngora, H., Galli, U.M., Miyazaki, K., and Zöller, M. (2012). Membrane-bound and exosomal metastasis-associated C4.4A promotes migration by associating with the α(6)β(4) integrin and MT1-MMP. Neoplasia *14*, 95–107.

Osborne, C.K., Bardou, V., Hopp, T.A., Chamness, G.C., Hilsenbeck, S.G., Fuqua, S.A.W., Wong, J., Allred, D.C., Clark, G.M., and Schiff, R. (2003). Role of the estrogen receptor coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer. J. Natl. Cancer Inst. *95*, 353–361.

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

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods *14*, 417–419.

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

Pink, J.J., Jiang, S.Y., Fritsch, M., and Jordan, V.C. (1995). An estrogen-independent MCF-7 breast cancer cell line which contains a novel 80-kilodalton estrogen receptor-related protein. Cancer Res. *55*, 2583–2590.

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

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44* (*W1*), W160-5.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

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

Santen, R.J., Song, R.X., Zhang, Z., Kumar, R., Jeng, M.H., Masamura, A., Lawrence, J., Jr., Berstein, L., and Yue, W. (2005). Long-term estradiol deprivation in breast cancer cells up-regulates growth factor signaling and enhances estrogen sensitivity. Endocr. Relat. Cancer *12* (*Suppl 1*), S61–S73.

Seiter, S., Stassar, M., Rappl, G., Reinhold, U., Tilgen, W., and Zöller, M. (2001). Upregulation of C4.4A expression during progression of melanoma. J. Invest. Dermatol. *116*, 344–347.

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

Slyper, M., Shahar, A., Bar-Ziv, A., Granit, R.Z., Hamburger, T., Maly, B., Peretz, T., and Ben-Porath, I. (2012). Control of breast cancer growth and initiation by the stem cell-associated transcription factor TCF3. Cancer Res. *72*, 5613–5624.

Smith, C.L., Conneely, O.M., and O'Malley, B.W. (1993). Modulation of the ligand-independent activation of the human estrogen receptor by hormone and antihormone. Proc. Natl. Acad. Sci. U S A *90*, 6120–6124.

Smith, C.L., Nawaz, Z., and O'Malley, B.W. (1997). Coactivator and corepressor regulation of the agonist/antagonist activity of the mixed antiestrogen, 4-hydroxytamoxifen. Mol. Endocrinol. *11*, 657–666.

Smith, B.A., Kennedy, W.J., Harnden, P., Selby, P.J., Trejdosiewicz, L.K., and Southgate, J. (2001). Identification of genes involved in human urothelial cell-matrix interactions: implications for the progression pathways of malignant urothelium. Cancer Res. *61*, 1678–1685.

Song, L., and Crawford, G.E. (2010). DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. Cold Spring Harb. Protoc. *2010*, pdb.prot5384.

Song, L., Zhang, Z., Grasfeder, L.L., Boyle, A.P., Giresi, P.G., Lee, B.K., Sheffield, N.C., Gräf, S., Huss, M., Keefe, D., et al. (2011). Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. Genome Res. *21*, 1757–1767.

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

Wardell, S.E., Ellis, M.J., Alley, H.M., Eisele, K., VanArsdale, T., Dann, S.G., Arndt, K.T., Primeau, T., Griffin, E., Shao, J., et al. (2015). Efficacy of SERD/SERM Hybrid-CDK4/6 inhibitor combinations in models of endocrine therapy-resistant breast cancer. Clin. Cancer Res. *21*, 5121–5130.

Werner, S., Frey, S., Riethdorf, S., Schulze, C., Alawi, M., Kling, L., Vafaizadeh, V., Sauter, G., Terracciano, L., Schumacher, U., et al. (2013). Dual roles of the transcription factor grainyhead-like 2 (GRHL2) in breast cancer. J. Biol. Chem. *288*, 22993–23008.

Willuda, J., Linden, L., Lerchen, H.G., Kopitz, C., Stelte-Ludwig, B., Pena, C., Lange, C., Golfier, S., Kneip, C., Carrigan, P.E., et al. (2017). Preclinical antitumor efficacy of BAY 1129980—a novel auristatin-based anti-C4.4A (LYPD3) antibody-drug conjugate for the treatment of non-small cell lung cancer. Mol. Cancer Ther. *16*, 893–904.

Wright, T.M., Wardell, S.E., Jasper, J.S., Stice, J.P., Safi, R., Nelson, E.R., and McDonnell, D.P. (2014). Delineation of a FOXA1/ERα/AGR2 regulatory loop

that is dysregulated in endocrine therapy-resistant breast cancer. Mol. Cancer Res. *12*, 1829–1839.

Xiang, X., Deng, Z., Zhuang, X., Ju, S., Mu, J., Jiang, H., Zhang, L., Yan, J., Miller, D., and Zhang, H.-G. (2012). Grhl2 determines the epithelial phenotype of breast cancers and promotes tumor progression. PLoS ONE *7*, e50781.

Xiao, T., Li, W., Wang, X., Xu, H., Yang, J., Wu, Q., Huang, Y., Geradts, J., Jiang, P., Fei, T., et al. (2018). Estrogen-regulated feedback loop limits the efficacy of estrogen receptor-targeted breast cancer therapy. Proc. Natl. Acad. Sci. U S A *115*, 7869–7878.

Zhang, Y., Lin, Y.-H., Johnson, T.D., Rozek, L.S., and Sartor, M.A. (2014). PePr: a peak-calling prioritization pipeline to identify consistent or differential peaks from replicated ChIP-Seq data. Bioinformatics *30*, 2568–2575.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit Polyclonal Anti-FOXA1	Abcam	Cat#ab23738; RRID:AB_04842
Rabbit Polyclonal Anti-Alpha-Tubulin (E-19)	Santa Cruz	Cat#Sc-12462-R; RRID:AB_2241125
Rabbit Polyclonal Anti-H3K4Me2	Sigma	Cat#07-030; RRID:AB_310342
Rabbit Polyclonal Anti-H3K27Ac	Diagenode	Cat#C15410196; RRID:AB_2637079
Rabbit Polyclonal Anti-H3K27Ac	Abcam	Cat#ab4729; RRID:AB_2118291
Sheep Polyclonal Anti-C4.4a/LYPD3	R & D Systems	Cat#AF5428; RRID:AB_2234844
Rabbit Monoclonal Anti-LYPD3	Abcam	Cat#ab151709
Rabbit Polyclonal Anti-GRHL2	Sigma	Cat#HPA004820; RRID: AB_1857928
Mouse monoclonal Anti-Beta actin (AC15)	Sigma	Cat#A5441; RRID:AB_476744
Mouse monoclonal anti-LYPD3	Arumugam et al., 2015	N/A
Rabbit Polyclonal Anti-LYPD3	Sigma	Cat#HPA041797; RRID:AB_2677679
Rabbit Monoclonal Anti-AGR2	Novus Bio	Cat#NBP1-40630; RRID:AB_2305344
Rabbit Polyclonal Anti-Lamin A	Santa Cruz	Cat#sc-20680; RRID:AB_648148
Mouse monoclonal anti-AGR2	Arumugam et al., 2015	N/A
Biological Samples		
Breast Tumor Tissue Microarray	Lin et al., 2017; Jeffrey Marks, PhD; Duke IRB approved protocol Pro00012025	N/A
Breast Tumor Tissue Microarray	Xiao et al., 2018; Drury et al., 2011; Mitchell Dowsett, PhD	N/A
Chemicals, Peptides, and Recombinant Proteins		
17-Beta-estradiol time-released sc pellet (0.72mg / 60 days)	Innovative Research of America	Cat#SE-121 Cas#50-28-2
Tamoxifen time-released sc pellet (5mg/ 60 days)	Innovative Research of America	Cat#E-361 Cas#10540-29-1
Fulvestrant (for animal studies)	MedChem express	Cat#HY-13636 Cas#129453-61-8
17-Beta-estradiol [50-28-2]	Sigma	Cat#E8875 Cas#50-28-2
Fulvestrant (ICI) [129453-61-8]	Sigma	Cat#I4409 Cas#129453-61-8
4-hydroxytamoxifen [68047-06-3]	Sigma	Cat#H7904 Cas# 68047-06-3
Critical Commercial Assays		
Aurum Total RNA Mini Kit	Bio-Rad	Cat#7326820
iScript cDNA synthesis Kit	Bio-Rad	Cat#1708890
iQ SYBR Green supermix	Bio-Rad	Cat#1708880
Fluoreporter Assay	Invitrogen	Cat#F-2962
Agilent RNA 6000 Nano Kit	Agilent	Cat#5067-1511
Ilumina TruSeq RNA Sample Prep Kit – Sets A/B	Illumina	Cat#FC-122-1001 Cat#FC-122-1002
QIAGEN Elution Buffer	QIAGEN	Cat#1014609
TruSeq PE Cluster Kit v3- cBot – HS	Illumina	Cat#FC-401-3001
KAPA HTP library preparation kit	Kapa Biosystems	Cat#KR0426
Background Terminator	Biocare	Cat#BT967L
4plus Biotinylated Universal Goat Link	Biocare	Cat#GU600H
4plus Streptavidin HRP Label	Biocare	Cat#HP604H
Da Vinci Green Diluent	Biocare	Cat#PD900L
Dako liquid DAB+ substrate chromogen system	Abcam	Cat#Ab64238

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
MCF7, MCF7-WS8, TAMR cell line RNASeq	This paper	GSE106695
MCF7-WS8, TAMR xenograft tumor RNASeq	This paper	GSE106695
MCF7-WS8, TAMR FOXA1, H3K27Ac, H3K4Me2, GRHL2 ChIP Seq	This paper	GSE106995
TAMR siCtrl, siGRHL2 A, siGRHL2 C RNASeq	This paper	GSE106995
Mendeley Data	This paper	https://data.mendeley.com/datasets/s6y9mzbhx7/draft?a=b22d3e7e-5cd0-4b5f-834a-59a36523a531
Experimental Models: Cell Lines		
MCF7	ATCC	N/A
MCF7-WS8	Gottardis and Jordan, 1988; Connor et al., 2001	N/A
TAMR	Connor et al., 2001; Wright et al., 2014	N/A
CAMA-1	ATCC	N/A
MDA-MB-361	ATCC	N/A
HCC1428-TamR	Guest et al., 2016	N/A
Experimental Models: Organisms/Strains		
MCF7-WS8 xenograft	Gottardis and Jordan, 1988; Connor et al., 2001	N/A
TAMR xenograft	Connor et al., 2001; Wright et al., 2014	N/A
J:nu nude mice	Duke Breeding Core	JAX stock #007850
Oligonucleotides		
Primers: LYPD3 Forward 5′ GTCACCTTGACGGCAGCTAA 3′	This paper	LYPD3
Primers: LYPD3 Reverse 5′ GTCTTGTTGCGGAGGTCAGA 3′	This paper	LYPD3
Primers: KRT13 Forward 5′ CGAGGGCCAGGACGCCAAGATGAT 3′	This paper	KRT13
Primers: KRT13 Reverse 5′ ACGGACATCAGAAGTGCGGCG 3′	This paper	KRT13
Primers: RPLP0 Forward 5′ GGACATGTTGCTGGCCAATAA 3′	This paper	36B4
Primers: RPLP0 Reverse 5′ GGGCCCGAGACCAGTGTT 3′	This paper	36B4
Primers: GRHL2 Forward 5′ AACAGGAAGAAAGGGAAAGGCCAGG 3′	This paper	GRHL2
Primers: GRHL2 Reverse 5′ TAGATTTCCATGAGCGTGACCTTG 3′	This paper	GRHL2
Primers: LYPD3-3 (ChIP-qPCR) Forward 5′ TCTCTCTCTCTCTTGCTGTCTCT 3′	This paper	LYPD3-3
Primers: LYPD3-3 (ChIP-qPCR) Reverse 5′ AACGAAGGGCTTGTTTAATTTTAATT 3′	This paper	LYPD3-3
Primers: AGR2-3 (ChIP-qPCR) Forward 5′ TCTGATGTGGTCCCATGAGG 3′	This paper	AGR2-3
Primers: AGR2-3 (ChIP-qPCR) Reverse 5′ TCTGATGTTTCTTGGTTCTTGCT 3′	This paper	AGR2-3
Primers: MUC20-1 (ChIP-qPCR) Forward 5′ TGACGCTGCCATCATAAGGG 3′	This paper	MUC20-1

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Primers: MUC20-1 (ChIP-qPCR)	This paper	MUC20-1
Reverse 5′ CCCACTTACTGTCCCACGTT 3′		
Primers: MAPK4-1 (ChIP-qPCR)	This paper	MAPK4-1
Forward 5′ TGTAGGGCTAGCGACTGAGA 3′		
Primers: MAPK4-1 (ChIP-qPCR)	This paper	MAPK4-1
Reverse 5′ TGGGTAAGATCTACATGTAGACAGG 3′		
Silencer Negative Control No. 1	Ambion/ Thermo Fisher	Cat#AM4611 "siCtrl "
siRNA to GRHL2 – 109594	Ambion/ Thermo Fisher	Cat#AM16708 "siGRHL2 A"
siRNA to GRHL2 – 109596	Ambion/ Thermo Fisher	Cat#AM16708 "siGRHL2 C"
siRNA to GRHL2 – 116387	Ambion/ Thermo Fisher	Cat#AM16708 "siGRHL2 D"
Negative Control siRNA	QIAGEN	Cat#1027310 "siCtrl"
siRNA to LYPD3 – Hs_LYPD3_1	QIAGEN	Hs_LYPD3_1, Cat#SI03082072"siLYPD3 1"
siRNA to LYPD3 – Hs_LYPD3_2	QIAGEN	Hs_LYPD3_2, Cat#SI03084291"siLYPD3 2"
siRNA to LYPD3 – Hs_C4.4A_2	QIAGEN	Hs_C4.4A_2, Cat#SI00105707"siLYPD3 3"
DNASeq Oligos: Linker 1 Oligo 1a: 5 –Bio-ACAGGTTCAGAGTTCTA CAGTCCGAC-3 Oligo 1b: 5 –P-GTCGGA CTGTAGAACTCTGAAC-Amm-3	Song and Crawford, 2010	N/A
DNASeq Oligos Linker 2 Oligo 2a: 5 –P-TCGTATGCCGTCTTCTGCTTG-3, Oligo 2b: 5 –CAAGCAGAAGACGGCATACGANN-3 (N represents any of A, T, G, or C)	Song and Crawford, 2010	N/A
DNaseq Oligs: Library amplification primer 1 – 5′ –CAAGCAGAA GACGGCATACGA- 3′ primer 2 – 5′ –AATGATACGGCGACCACCG ACAGGTTCAGAGTTCTACAGTCCGA- 3′	Song and Crawford, 2010	N/A
Software and Algorithms		
CASAVA 1.8.2 software suite	Illumina	http://support.illumina.com/sequencing/sequencing_software/casava.html http://gensoft.pasteur.fr/docs/casava/1.8.2/
easyRNASeq	Delhomme et al., 2012	http://master.bioconductor.org/packages/release/bioc/
edgeR	Robinson et al., 2010	http://master.bioconductor.org/packages/release/bioc/
DESeq2	Love et al., 2014	http://master.bioconductor.org/packages/release/bioc/
Skewer v2.2	Jiang et al., 2014	https://github.com/relipmoc/skewer/releases
Samtools v1.3.1	Li and Durbin, 2009	http://samtools.sourceforge.net/
STAR v2.5.2b	Dobin et al., 2013	https://github.com/alexdobin/STAR/releases
Salmon v0.8.0	Patro et al., 2017	https://github.com/COMBINE-lab/salmon/releases
Sciclone NGS Workstation	Sciclone	P/N SG3-31020-0300
Peak Prioritization Pipeline (Pepr) v.1.1.18	Zhang et al., 2014	https://github.com/shawnzhangyx/PePr/releases
F-Seq v1.8.4	Boyle et al., 2008	http://fureylab.web.unc.edu/software/fseq/
Bedtools v2.25	Quinlan and Hall, 2010	https://github.com/arq5x/bedtools2/releases
DeepTools v 2.4.2	Ramírez et al., 2016	https://github.com/fidelram/deepTools/releases

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
HOMER v4.8.2	Heinz et al., 2010	http://homer.ucsd.edu/homer/ngs/
Geneanalytics v1.0	Martz et al., 2014	http://geneanalytics.duhs.duke.edu
Other		
Mouse IgG	Southern Biotech	Cat#0107-01
Lipofectamine RNAi Max	Thermo Fisher	Cat#13778150
Protein A/G beads	Pierce	Cat#20421
Falcon cell culture inserts, transparent PET membrane 8.0 micron pore	Corning	Cat#353097
Protein A magnetic beads	Invitrogen	Cat#10001D
DNA Purification Beads	MagBio	Cat#AC-60050
Complete Protease Inhibitor Cocktail	Roche/Sigma	Cat#11697498001

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Donald McDonnell (Donald.mcdonnell@duke.edu).

This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Culture

Female MCF7-WS8 and its derivative, TAMR, cell lines were obtained as previously described (Gottardis and Jordan 1988; Connor et al., 2001) and validated (Wright et al., 2014), and maintained in Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM/F12). Female MCF7 cells were obtained from American Type Culture Collection (ATCC) and were used for comparison as in Figure S5, maintained in the same media. TAMR cells were kept under constant selection with 100nM 4-OHT. All cell lines were supplemented with 8% fetal bovine serum (FBS) or twice charcoal-stripped FBS (CFS) (Hyclone Laboratories, Logan, UT), 0.1 mM non-essential amino acids (NEAA) and 1 mM sodium pyruvate (NaPyr). ER ligands used for cell culture treatment were obtained from Sigma (St. Louis, MO) include: 17β-Estradiol [50-28-2] (E8875), Fulvestrant (ICI) [129453-61-8] (I4409), and 4-hydroxytamoxifen [68047-06-3] (H7904).

Xenograft Studies

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. All procedures performed in studies involving animals were in accordance with the ethical standards of the Duke University Institutional Animal Care and Use Committee.

For confirmation of TAMR model using cell lines, 2 days prior to cell injection, female J:nu mice, JAX stock #007850 (~6 weeks of age) were ovariectomized under isoflurane anesthesia and tamoxifen pellets (5 mg/60 days from Innovative Research of America) were implanted. Log phase TAMR cells were injected as a 1:1 mixture of serum-free media and matrigel (242 and 248: 5×10^7 cells; 699 and 700: 8×10^7 cells, as indicated in Figure S1A) orthotopically subcutaneously under the second nipple. Tumors were measured 3X weekly (volume = l x w^2 x 0.5) with concurrent weight monitoring. Following the final measurement, animals were euthanized and tumors were sterilely excised prior to being subdivided into ~8 mm^3 sections. These sections were then serially implanted under anesthesia into ovariectomized mice (prepared as above) receiving tamoxifen or no treatment and monitored for growth for the time indicated.

MCF7-WS8 and TAMR tumor samples used for RNA-seq analysis and immunoblotting were prepared as follows: female J:nu mice (~6 weeks in age) were ovariectomized under isoflurane anesthesia; a slow release estradiol pellet (0.72mg/60 days from Innovative Research of America) or tamoxifen treatment pellet (5mg/60 days from Innovative Research of America) was implanted in the scapular region during the same procedure. The next day an approximate 8 mm^3 of tumor tissue (derived from sectioning a freshly harvested tumor of 0.8 – 1 cm^3 volume) was engrafted orthotopically (right axial mammary fatpad) under anesthesia using a 10 g trocar. Tumors were measured 3 times weekly, concurrent with weight and behavior monitoring. For MCF7-WS8 tumors, tumors were grown for 30 days when they reached a size of ~0.15-0.2 cm^3 volume. Animals were then randomized to treatment with vehicle (0.1cc corn oil sc 3 times weekly) or tamoxifen (40mg/kg 3 times weekly). Animals were euthanized and tissues preserved following a total of 6 weeks of tumor growth. For TAMR tumors, tamoxifen treated animals were randomized to vehicle (0.2cc corn oil sc weekly), or fulvestrant (200mg/kg sc weekly). Animals were euthanized and tissues preserved following a total of 8 weeks of tumor growth.

For anti-AGR2 and anti-LYPD3 antibody treatment studies, tamoxifen stimulated TAMR tumors were initiated orthotopically by serial tumor transfer into female J:nu mice (~6 weeks age), as indicated above

Briefly, ovariectomized recipient mice received no treatment or tamoxifen treatment via a timed-release pellet (5 mg/60 days from Innovative Research of America) implanted subcutaneously. Two days later, TAMR tumors (~0.8cm^3 volume) were sterilely excised from euthanized tamoxifen treated donor mice, diced to ~2mm^3 sections and implanted into the axial mammary gland of recipient mice under anesthesia using a 10 g trocar. Tumor growth was measured 3 times weekly by caliper. When tumor volume reached ~0.15cm^3 (~20 days), mice were randomized to receive 45 mg/kg IgG, 15 mg/kg anti-AGR2 or 45 mg/kg anti-LYPD3 antibodies ip twice weekly, with groups further subdivided to receive sc injection of corn oil or 25 mg/kg fulvestrant.

Long-term estrogen-deprived (LTED) xenograft tumor model was derived by withdrawal of estradiol treatment of a growing (~0.4 cm^3 volume) MCF7-WS8 xenograft tumor engrafted into the axial mammary fat pad of an ovariectomized J:nu mouse. After tumor regression and stasis (18 weeks), this initial (parent) tumor re-entered exponential growth. When tumor volume reached ~2 cm^3 (37 weeks after estradiol withdrawal), the donor mouse was euthanized, and the tumor was resected, sectioned, and implanted (~8 mm^3 initial volume) into the mammary fat pad of ovariectomized (10 days prior) 6-week old female J:nu mice. This xenograft model of estrogen withdrawal is continuously maintained via serial passage of tumor tissue as described above. Tumors were measured 3 times weekly, concurrent with weight and behavior monitoring. When tumors volume reached ~0.1cm^3, animals were euthanized and tumors were harvested, as described above.

METHOD DETAILS

Antibodies

The following antibodies were purchased from Santa Cruz Biotechnology: α-Tubulin (E-19, R) (sc-12462-R), Abcam: FOXA1 (ab23738), H3K27Ac (ab4729, for ChIP-qPCR), LYPD3 (Ab151709, for immunoblot/LiCOR), Sigma: H3K4Me2 (07-030), GRHL2 (HPA004820), LYPD3 (HPA041797, for IHC), and β-actin (AC15) (a5441), R&D Biosystems: Human C4.4a/LYPD3 (AF5428, immunoblot/ECL), Diagenode: H3K27Ac (C15410196, for ChIP-seq), Novus: AGR2 (NBP1-40630), Santa Cruz: Lamin A (sc-20680).

AGR2 and LYPD3 monoclonal antibodies as produced and validated previously (Arumugam et al., 2015) for treatment of xenograft tumors were provided by Craig Logsdon at MD Anderson. Mouse IgG Isotype control (0107-01) was obtained from Southern Biotech.

siRNA transfection

For experiments involving transient transfection of small interfering RNA (siRNA), validated siRNA or siRNA control were used as indicated and listed below. Cells were plated in phenol red-free DMEM/12 containing 8% charcoal-stripped serum (CFS), 0.1 mM NEAA and 1 mM NaPyr in the presence of 60 nM siRNA or associated siRNA control for siGRHL2 or 30nM siRNA or associated siRNA control for siLYPD3 using Lipofectamine RNAi MAX as the transfection agent, according to the manufacturer's recommendations. For qPCR and immunoblot analysis, cells were harvested following 3 days of transfection. For proliferation assays, cells were allowed to grow for 9 days following transfection, as detailed in "proliferation assay" below. For migration assay, cells were allowed to grow for 18 h following transfection, as detailed in "migration assay" below.

siRNA sequences used, include:

siCtrl No. 1 (Ambion/ThermoFisher) (AM4611)
siGRHL2 A (Ambion/ThermoFisher) (109594)
siGRHL2 C (Ambion/ThermoFisher) (109596)
siGRHL2 D (Ambion/ThermoFisher) (116387)
siCtrl (QIAGEN, 1027310):
siLYPD3 #1 (QIAGEN, Hs_LYPD3_1, SI03082072)
siLYPD3 #2 (QIAGEN, Hs_LYPD3_2, SI03084291)
siLYPD3 #3 (QIAGEN, Hs_C4.4A_2, SI00105707)

RNA Isolation and qPCR

For cell line studies, cells were seeded in 12-well plates in phenol red-free media containing 8% CFS for 2 days and treated with ligands as indicated. After the indicated time period, cells were harvested and total RNA was isolated using the Aurum^{TM} Total RNA Mini Kit (Bio-Rad, Hercules, CA). For tumor tissue, tumors were dissected, and then snap frozen in liquid nitrogen. Tissue was then pulverized using mortar and pestle and then isolated. 500ng to 1ug of purified RNA was reverse transcribed using the iScript cDNA synthesis kit (Bio-Rad). Reactions for qPCR were performed with diluted cDNA, specified primers, and iQ SYBRGreen supermix (Bio-Rad). Data are normalized to RPLP0 (36B4) housekeeping gene and presented as fold expression relative to controls, as previously described (Wright et al., 2014).

Primer sequences used for qPCR, include:

LYPD3:
(forward): 5′ GTCACCTTGACGGCAGCTAA 3′

(reverse) 5′ GTCTTGTTGCGGAGGTCAGA 3′
GRHL2:
(forward): 5′ AACAGGAAGAAAGGGAAAGGCCAGG 3′
(reverse): 5′ TAGATTTCCATGAGCGTGACCTTG 3′
KRT13
(forward) 5′ CGAGGGCCAGGACGCCAAGATGAT 3′
(reverse) 5′ ACGGACATCAGAAGTGCGGCG 3′
36 B4 (RPLP0):
(forward): 5′ GGACATGTTGCTGGCCAATAA 3′
(reverse): 5′ GGGCCCGAGACCAGTGTT 3′

Immunoblotting

For cell line studies, cells were seeded in 6-well plates in phenol red-free DMEM containing 8% CFS, 0.1 mM NEAA and 1 mM NaPyr for 2 days and treated as indicated. Following treatment for the indicated time periods, cells were harvested in ice-cold PBS and lysed in RIPA Buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 1.5 mM $MgCl_2$, 1% Triton X-100, 1 mM EDTA, 10% glycerol, 50 mM NaF, 2 mM Na_3VO_4, and protease inhibitors (Sigma #8340-ML) while rotating at 4°C for 30 min. For tumor tissue, tumors were dissected, and then snap frozen in liquid nitrogen. Tissue was then pulverized using mortar and pestle and lysed in RIPA lysis buffer as above. 20-25μg of whole-cell extract was resolved by SDS-PAGE, transferred to a PVDF membrane (Bio-Rad) and probed with the appropriate antibodies.

Proliferation assays

TAMR cells were reverse transfected in triplicate with siRNA as indicated above at the time of plating into 96-well plate. For siGRHL2 and respective siCtrl, transfection was repeated on Day 3 by aspirating media suspension and re-transfecting with respective siRNA. For siLYPD3 and respective controls, transfection was carried out only on day 0. On collection day, media was decanted and plates were frozen at −80°C. Plates were thawed completely at room temperature after which 100ul of H_2O was added to each well and incubated at 37°C for 1 hr. Plates were refrozen at −80°C, thawed at room temperature, and DNA content was detected using a Fluoreporter assay (Invitrogen) per manufacturer's instructions.

Migration assays

Cells were serum starved for 24 h with phenol red-free DMEM/F12, 0.1 mM NEAA and 1 mM NaPyr. Cells were subsequently plated on Falcon cell culture inserts (transparent PET membrane, 8.0 micron pore) (Corning) in duplicate and migrated toward 8% FBS in DMEM/F12, 0.1 mM NEAA and 1 mM NaPyr for 18 h. Migrating cells were fixed with 10% formalin, stained with 1% crystal violet in PBS, and counted.

Preparation of Nuclear Extracts for LC/MS-MS analysis

FOXA1 Antibody (60ug) was incubated with 90ul Protein A/G beads (Pierce #20421) in PBS overnight. The next morning beads were washed three times in 0.2M sodium borate, pH 9.0. Complexed beads were conjugated with dimethylpimelimidate (DMP) with 0.0259 g DMP and 5ml of 0.2 M sodium borate to make a 20mM solution. Beads were incubated for 40 min with end over end rocking at room temperature. After, beads were washed in 0.2 M ethanolamine (pH 8.0) to quench residual DMP and then suspended in 0.2 M ethanolamine for additional 1 h incubation. Uncoupled antibody was then washed three times using 0.58% acetic acid with 150 mM NaCl. Beads were stored in PBS with sodium azide at 4°C.

Nuclear extracts of TAMR cells were then prepared for Mass-Spec analysis. Briefly, 10 cm culture plates were washed with PBS and harvested with 0.25% trypsin. Cells were scraped into a conical tube and spun for 5 min at 1500 g in a pre-cooled centrifuge. Cell pellets were suspended in 5 times the cell pellet volume in hypotonic buffer for 5 min and thereafter checked every minute with trypan blue until greater than 90% of cells stained positive. NP40 was added to 0.1% and vortexed on mid setting for 10 s and immediately centrifuged at 3000 g for 1 min. Resulting supernatant, predominately consisting of cytosolic extract was set aside and nuclear pellet was suspended in a half a cell pellet volume of low salt buffer [20 mM HEPES pH 7.9, 0.02 M KCl, 0.2 mM EDTA, 25% Glycerol 1 mM DTT, beta-glycerophosphate, protease inhibitors, NaF, Na3V4, and sodium butyrate (NaB)] being careful not to break nuclei. A half bed volume of high salt buffer [20mM HEPES pH 7.9, 1 M KCl, 0.2 mM EDTA, 25% Glycerol 1 mM DTT, beta-glycerophosphate, protease inhibitors, NaF, NaV, and NaB] was added gently and tube was rocked for 1 h in cold room. Nuclear debris was pelleted at 14000 g for 15 min and supernatant was dialyzed against dialysis buffer [20 mM HEPES pH 7.9, 100 mM KCl, 0.2 mM EDTA, 20% Glycerol] to normalize salts for two changes at 1 h each.

Nuclear extracts were normalized at 10 mg each and pre-cleared with A/G beads for 1 h and then incubated overnight with the prepared conjugated beads. The next day beads were washed 5 times with wash buffer [50 mM HEPES pH 7.9, 0.2% NP40, 150 mM KCl, 1 mM EGTA, 1 mM DTT, beta-glycerophosphate, protease inhibitors, NaF, NaV, and NaB]. Final washes were performed using PBS three times and 50 mM Ammonium Bicarbonate twice before submitting to the Duke Proteomics Core for LC-MS/MS analysis.

DNase-Seq

DNase-Seq was performed as previously described (Song and Crawford, 2010; Song et al., 2011). Briefly, cells were plated in phenol red-free DMEM/12 containing 8% charcoal-stripped serum (CFS), 0.1 mM NEAA and 1 mM NaPyr in 15cm culture plates and treated 2 days later with either vehicle or 4-OHT for 24 h before harvesting. Two independent biological replicates of each condition were prepared. Nuclei were extracted and digested with DNaseI enzyme. After confirmation of adequate digestion, DNaseI-digested ends were blunt ended, and a biotinylated linker was ligated to these ends. Linkers were generated using the following oligos: Linker 1: annealed oligonucleotides 1a and 1b (HPLC-purified; Integrated DNA Technologies)

Oligo 1a: 5 -Bio-ACAGGTTCAGAGTTCTACAGTCCGAC-3 Oligo 1b: 5 -P-GTCGGACTGTAGAACTCTGAAC-Amm-3; Linker 2: annealed oligonucleotides 2a and 2b (HPLC-purified; Integrated DNA Technologies) Oligo 2a: 5 -P-TCGTATGCCGTCTTCTGCTTG-3, Oligo 2b: 5 -CAAGCAGAAGACGGCATACGANN-3 (N represents any of A, T, G, or C). Fragments with linker attached were isolated, digested with MmeI, and captured using streptavidin-conjugated magnetic beads. A second linker was ligated to the MmeI-digested end, and then the fragments were amplified and subsequently purified via gel electrophoresis. Primers used for library amplification include primer 1 – 5′ -CAAGCAGAAGACGGCATACGA- 3′ and primer 2 – 5′ -AATGATACGGCGACCACCGACAGGTTCAGAGTTC TACAGTCCGA- 3′. The libraries were sequenced using 50bp SR on Illumina HiSeq, using the custom sequencing primer 5 -CCACC GACAGGTTCAGAGTTCTACAGTCCGAC-3

FASTQ files were aligned to hg19 human genome reference from UCSC using BWA aln (Li and Durbin, 2009). Reads were filtered from the SAM file that do not align based on the 0X004 flag, align to multiple locations, align to more than two ambiguous locations, align to the Y chromosome, and fall of chromosome boundaries using the chrom.sizes file from UCSC. Additional alignments were also filtered to remove problematic repetitive regions such as alpha satellites and sequence artifacts as defined by the ENCODE blacklist. Biological replicates were compared for reproducibility and correlation. Final base-pair resolution signal as a Wig file was generated using F-Seq at 300bp signal bandwidth (Boyle et al., 2008) and converted to bigwig using the UCSC utility, WigtobigWig. Peaks were called by F-Seq and significance of the peaks were determined by fitting DNase-Seq signal data to a gamma distribution and then determining the signal value that corresponded to a p value < 0.05. Sequencing tags were quantified in each DNase peak for each condition using multicov from the bedtools suite (Quinlan and Hall, 2010). Reads overlapping within 500bp each direction from the TSS were also subset to discriminate against sites likely unrelated to TF binding using the Refseq annotated genes. To identify regions of significant change across cell lines and treatments in DNase-Seq data, we used the DESeq (Anders and Huber, 2010) package from bioconductor.

RNA-Seq

RNA was harvested after appropriate treatments as indicated above in ‘‘RNA isolation.’’ For MCF7-WS8 and TAMR cell lines, individual biological duplicates were collected and analyzed. For MCF7-WS8 and TAMR xenograft tumors, biological quadruplicates were collected and analyzed. For both cell lines and tumors, total RNA was quantified using the Agilent RNA 6000 Nano kit (#5067-1511) on the Agilent 2100 BioAnalyzer. 1ug of high purity total RNA (defined as greater than 7.0 RNA Integrity Number (RIN)) was used as input to the Illumina TruSeq RNA Sample Prep Kit – Sets A/B (48Rxn) (#FC-122-1001 and FC-122-1002). The gel-free protocol was employed for the TruSeq RNA Sample Prep Kit per manufacturer’s specifications, and performed on the Biomek Fxp robotics platform. The PCR amplified RNA-seq library products were then quantified using the Advance Analytical Fragment Analyzer Standard Sensitivity NGS Fragment Analysis Kit (#DNF-479). The samples were diluted 10 nM in QIAGEN Elution Buffer (#1014609), and denatured and loaded at 3 pM on an Illumina cBOT using the TruSeq PE Cluster Kit v3 – cBot – HS (#PE-401-3001). The resultant flow cells were loaded on a HiSeq2500 with the TruSeq SBS Kit v3 – HS (200-cycles) reagents (#FC-401-3001). The RNA-seq libraries were sequenced at 100 bp paired end with 7 bp index using the standard Illumina primers. The sequence intensity files were generated on instrument using the Illumina Real Time Analysis software. The intensity files were demultiplexed and FASTQ files created using the CASAVA 1.8.2 software suite.

For RNA-seq analysis of TAMR cell lines treated with control siRNA or two independent siRNA sequences to GRHL2, RNA was harvested as above from independent biological triplicates. Samples were assessed using stranded mRNaseq on Illumina Hi-Seq with 50bp Paired End Rapid Run Sequencing.

RNA-seq samples (TAMR and MCF7-WS8 cell lines and TAMR cell lines treated with siCtrl or siGRHL2) were clipped using Skewer (Jiang et al., 2014) and aligned using STAR (Dobin et al., 2013) to the GRCh37 genome with Gencode v23Lift37 transcripts defined in the index. Default parameters were used with the exception that only 5 multi-mapping reads were allowed with outFilterMultimapNmax = 5. Following alignment, transcript quantification was performed using Salmon (Patro et al., 2017) with the transcriptome BAM files from STAR and the Gencode transcript reference. Differential expression analysis was performed using DESeq2 (Love et al., 2014) with tximport.

Xenograft samples were separately processed to account for reads that align to both human and mouse. FASTQ files were preprocessed to remove adapters and low-quality 3′ reads then aligned independently to the hg19 and mm10 genomes. Resulting BAM files were subsequently filtered to remove multi-mapping across species and rRNA reads. Reads were quantified using easyRNASeq (Delhomme et al., 2012) over ensemble transcripts in R and genes with more than 2 CPM in at least two conditions were brought forward to edgeR (Robinson et al., 2010) for differential expression. A multi-factorial design was incorporated to account for resistance and sensitive cells in each treatment state. TMM normalization was used and genes with $p < 0.01$ were considered differentially expressed for subsequent analyses.

ChIP-Seq

Cells were seeded in a 15cm dishes with appropriate media described above. Cells were grown to 90% confluence in phenol red-free DMEM/F12 media supplemented with 8% CFS, 0.1 mM NEAA and 1 mM NaPyr for 3 days and subsequently treated for 45 min with ETOH (1:10,000 dilution) to serve as vehicle control. Cells were then subjected to ChIP analysis. Each plate of cells was cross-linked with 1% formaldehyde PBS solution for a maximum of 10 min at room temperature and quenched with ice-cold, 125 mM glycine solution containing 5mg/ml bovine serum albumin (BSA) for 5 min. Cells were then rinsed once and harvested with ice cold PBS, pelleted at 8000 rpm for 30 s at room temperature and snap frozen in liquid nitrogen for storage at −80 degrees. All solutions were supplemented with 10 mM Na(C_3H_7COO). Cell pellets were thawed on ice, and then resuspended in Lysis Buffer containing 1% SDS, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0, protease inhibitor (Roche, Complete protease inhibitor tablets 11697498001). Cell lysates were sonicated using the Covaris E210 (for FOXA1 H3K4Me2, and H3K27Ac ChIP-seq samples) or E220 (for GRHL2 ChIP-seq samples) instrument according to the manufacturers' instructions. Sheared chromatin was diluted using Dilution Buffer (20 mM Tris [pH 8.0], 150 mM NaCl, 2 mM EDTA, 1% Triton X-100). Sheared, diluted chromatin was incubated with respective antibodies overnight in a deep 96-well plate at 4°C and then captured on protein A magnetic beads (Invitrogen, Dynal). After 45 min of incubation with the beads the immunoprecipitates were washed on a 96-well microplate in 150 uL volumes of the following solutions. For FoxA1, H3K4me2, and H3K27Ac ChIP, beads were washed a total of 6 times: 4 times in RIPA buffer containing 500 mM LiCl (50 mM HEPES, 1 mM EDTA, 0.7% Na Deoxycholate, 1% NP-40, 500 mM LiCl), and twice with TE buffer (20 mM Tris pH 8.0, 2mM EDTA). For GRHL2 ChIP, beads were washed twice with Buffer A (50 mM HEPES, 500 mM NaCl, 1mM EDTA, 0.1% SDS, 1% Triton X, 0.1% Na Deoxycholate), twice with Buffer B (20 mM Tris pH 8.0, 1 mM EDTA, 0.5% NP40, 0.5% Na Deoxycholate, 250 mM LiCl), twice with TE Buffer. Following washes, precipitates were re-suspended in a Reverse Crosslinking Buffer containing 100 mM NaHCO3 and 1% (w/v) SDS. Crosslink reversal was done at 65°C for 6 h. ChIP DNA were isolated using DNA purification beads (MagBio).

ChIP library construction was done by an automated protocol using the Kapa HTP library preparation kit (KR0426 Kapa Biosystems). All automation was performed using the Sciclone NGS Workstation (P/N SG3-31020-0300, PerkinElmer, Waltham, MA). Independent biological triplicate samples were prepared for all ChIP experiments and were sequenced using Illumina NS500 Single-End 75bp (SE75).

FASTQ files were clipped using Skewer and aligned with BWA mem (Li, 2013) to hg19 human genome reference. Differential peak calling was performed using the peak calling peak prioritization pipeline (PePr) (Zhang et al., 2014). In short, PePr considers all sample variance estimates in a sliding window approach to obtain differential binding sites across contrasts. Each contrast and corresponding set of ChIP replicates were therefore run with PePr to identify differential peaks. Normalized BigWig files were created using deeptools (previously cited).

ChIP-qPCR was performed essentially as described for ChIP-seq with the following exception: Cells seeded in 15cm dishes were transfected with siCtrl or siGRHL2-C for 72hrs before performing formaldehyde crosslink. Cell lysates were sonicated using Misonix ultrasonic cell disruptor for 13 s x 13 times, with output power of 10-11 W. Sheared, diluted chromatin was incubated with respective antibodies overnight in 1.5 mL microcentrifuge tubes at 4°C and then captured on protein A magnetic beads (Invitrogen, Dynal). After 45 min of incubation with the beads the immunoprecipitates were washed with 1 mL of each wash buffers as indicated in the ChIP-seq procedure above. ChIP DNA were isolated using QIAGEN PCR clean up kit. Purified chromatin was diluted 1:5 with H2O and qPCR performed using the following primers and iQ SYBRGreen supermix (Bio-Rad):

LYPD3-3_F TCTCTCTCTCTCTTGCTGTCTCT
LYPD3-3_R AACGAAGGGCTTGTTTAATTTTAATT
AGR2-3_F TCTGATGTGGTCCCATGAGG
AGR2-3_R TCTGATGTTTCTTGGTTCTTGCT
MUC20-1_F TGACGCTGCCATCATAAGGG
MUC20-1_R CCCACTTACTGTCCCACGTT
MAPK4-1_F TGTAGGGCTAGCGACTGAGA
MAPK4-1_R TGGGTAAGATCTACATGTAGACAGG

Integrative Analysis

Overlapping sets of ChIP-seq and DNase-seq peaks were determined using bedtools v2.25 (Quinlan and Hall, 2010). Heatmaps and profile plots of ChIP-seq and DNase-Seq data were drawn using deepTools v2.4.2 (Ramírez et al., 2016). Motif enrichment analysis was performed using HOMER v4.8.2 (Heinz et al., 2010) with random matched genomic control regions used as background sequence. Integration of ChIP-seq and differential expression data was performed by comparing the minimum distance between each ChIP-seq peak center and any known TSS for each expressed gene. Distributions of log2 fold-changes for genes associated with each peak were determined by assigning each gene to the first group of ChIP-seq peaks (starting from left to right in each figure) within the specified distance cut-off. For each differential expression experiment, a set of non-differential control genes was determined as those with average fold-change < 1.3x in either direction, and an uncorrected p value > 0.5. The enrichment of up and down-regulated genes around each set of sites of interest was compared to these control gene sets for each differential expression experiment.

Tissue Microarray

Tissue microarray was obtained from and prepared by the laboratory of JRM. Samples were acquired in compliance with the informed consent policy approved by the Duke University Institutional Review Board under the protocol Pro00012025, as previously published (Lin et al., 2017). A total of 100 patient samples were assessed on the TMA. All samples were de-identified prior to receipt of materials. Samples were previously evaluated and classified based on tumor and nodal stage, and estrogen receptor (ER), progesterone receptor (PR) status. An additional TMA from seventy-seven breast cancer patients diagnosed between 1981 and 2004 at the Royal Marsden Hospital, who had tissue available from both primary invasive tumor and subsequent invasive recurrence following adjuvant tamoxifen treatment were included to assess the expression levels of GHRL2 and LYPD3 in primary versus recurrent tumors (Drury et al., 2011; Xiao et al., 2018). Immunohistochemistry (IHC) was performed on the prepared microarray slides using Biocare Medical Supply IHC staining kit, including Background terminator (BT967L), 4+ Biotinylated Universal Goat Link (GU600H), 4+ Streptavidin HRP Label (HP604H), and Da Vinci Green Diluent (PD900L). Slides heated to 60°C for 1 h to melt paraffin, and then were deparaffinized in Clearify (MasterTech) and hydrated in a graded series of ethanol. Heat retrieval was performed using a sodium citrate buffer at a pH of 6 in a pressure cooker for 25 min and subsequently allowed to cool in sodium citrate to room temperature. Suppression of endogenous peroxidase was achieved using 3% hydrogen peroxide for 15 min. Background terminator was applied for 10 min. Slides were then rinsed for 15 min with tap water to remove residual hydrogen peroxide. Primary antibody staining was performed using GRHL2 (HPA0004820) at 0.05 mg/ml and LYPD3 (HPA041797) at 0.4 mg/ml, diluted in Da Vinci Green antibody dilution solution with 1% goat serum, at 4°C overnight. Tissue sections were washed with 3 times with Tris buffered saline with 0.1% polysorbate 20 (TBS-T), incubated with 4+ Biotinylated Universal Goat link at room temperature for 10 min, and then washed 3 more times with TBS-T. Immunoreactivity was detected using the Dako liquid DAB+ substrate chromogen system, as follows: incubation with 4+ Streptavidin HRP for 10 min, rinse 3 times with TBS-T, and incubation with DAB chromogen for 5 min. After washing, hematoxylin staining and Blue Nuclei staining were performed, followed by further washing, and stepwise dehydration with ethanol washes. Final steps include mounting and coverslip application.

The degree of GRHL2 staining was scored by board certified pathologist (AH) in the Duke Department of Pathology, who was provided only the de-identified patient ID number and sample grid. Hemotoxylin & Eosin staining was used to confirm presence of carcinoma within the samples. Samples were then scored on a 0 (absent) to 3 (high) scale to reflect degree of staining.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical Analyses

All statistical analyses for cell proliferation, cell migration, mRNA expression and tumor growth were performed using GraphPad Prism 6 (GraphPad Software, La Jolla, CA, USA). p values are indicated in figure legends. Cell proliferation and migration and cell line mRNA expression was compared by one-way ANOVA followed by Bonferroni's multiple comparisons test. For data shown of cell line experiments, n = 3 samples (technical replicates) per experiment for qPCR, immunoblotting and proliferations data; n = 2 samples per experiment (technical replicates) for migration; results are representative of experiments with coincident results, performed at least in triplicate, independently. SEM are plotted as error bars. For tumor mRNA experiments, unpaired two tailed t test was performed to compare expression differences between two groups. Specific n number of biologic replicates is indicated in the figure legend. Tumor growth was analyzed by exponential growth curve analysis and by 2-way ANOVA of matched values followed by Bonferroni's multiple comparisons test to establish significance between groups at each day of treatment. SEM are plotted as error bars.

For tissue microarray, statistical analyses were performed using SAS v9.4. Assessment of GRHL2 protein was merged with de-identified clinical variables, including pathologic T and N stage, ER, PR, and time to recurrence. Associations of GRHL2 with T stage, N stage, ER and PR status were completed with chi-square tests. Association of GRHL2 with time to recurrence was completed with the Kaplan-Meier estimator and the Log-Rank test.

For integrative ChIP-sequencing and RNA-sequencing analysis, all statistical analysis, differential gene calls and Fisher's exact test was performed using R.

DATA AND CODE AVAILABILITY

Deposited Data

Raw data files for the RNA, DNase and ChIP-sequencing analysis have been deposited in the NCBI Gene Expression Omnibus (GEO) under accession number: GSE106695

Mendeley Dataset including original figures is available at https://data.mendeley.com/datasets/s6y9mzbhx7/draft?a=b22d3e7e-5cd0-4b5f-834a-59a36523a531

Software

Gene Analytics is publicly available via http://geneanalytics.duhs.duke.edu. In brief, gene expression omnibus (GEO) was queried for breast cancer datasets that were performed on HGU133A or HGU133Plus2 Affymetrix platforms. In total, 25 non-redundant datasets were identified comprising 4885 patients. Datasets used: GSE10780, GSE11121, GSE12093, GSE12276, GSE1456, GSE16391,

GSE16446, GSE17705, GSE17907, GSE19615, GSE20194, GSE2034, GSE20685, GSE20711, GSE2109, GSE21653, GSE22093, GSE24185, GSE25066, GSE3494, GSE5460, GSE6532, GSE6532, GSE7390, and GSE9195. The raw data were downloaded from GEO, and each dataset was normalized with fRMA to remove platform-specific batch effects. The data were then combined using the COMBAT algorithm implemented in the sva package within R (Leek et al., 2012) with a design matrix to account for known covariates including data source and platform. Each tumor was then classified into PAM50 molecular subtypes using genefu (Haibe-Kains et al., 2012). To confirm normalization, a Multi-dimensional Scaling (MDS) plot was used to visually inspect the data in relation to platform and tumor subtype.

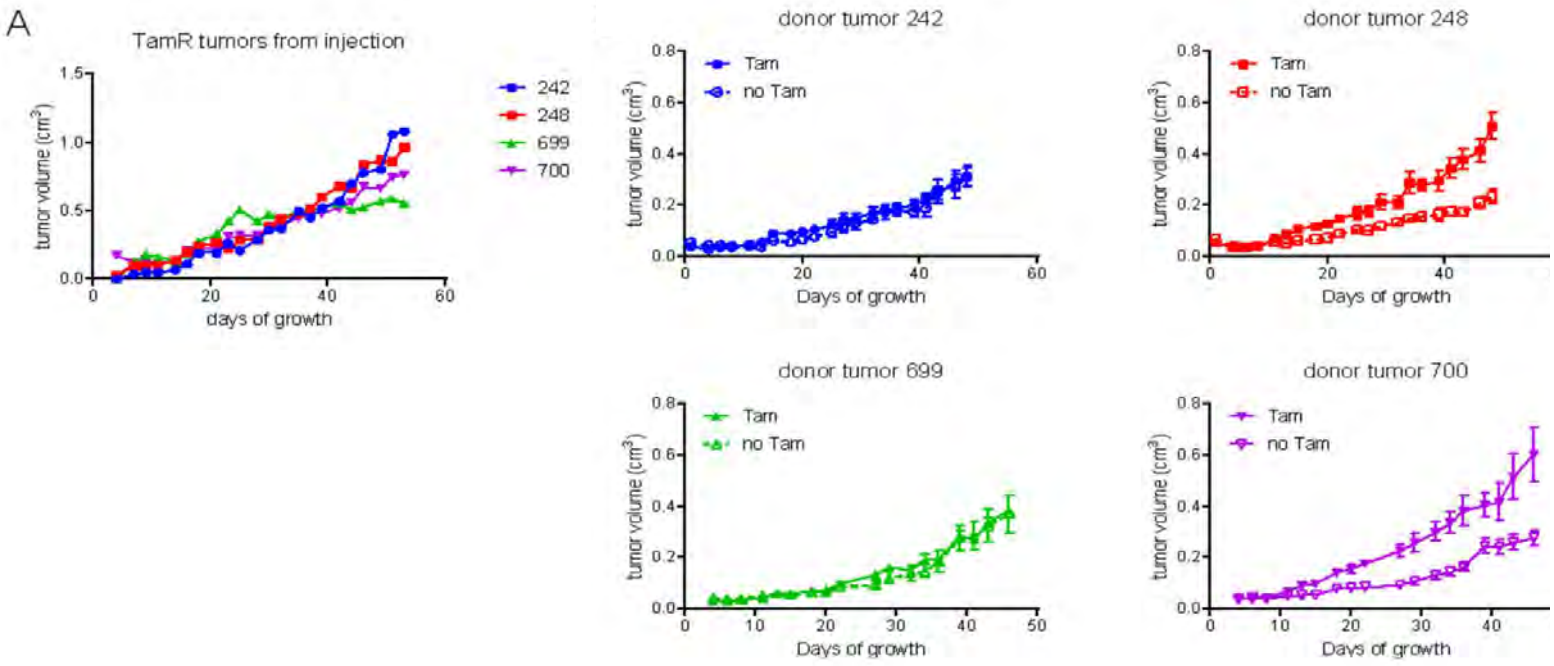
Cell Reports, Volume *29*

Supplemental Information

The Lineage Determining Factor GRHL2 Collaborates with FOXA1 to Establish a Targetable Pathway in Endocrine Therapy-Resistant Breast Cancer

Kimberly J. Cocce, Jeff S. Jasper, Taylor K. Desautels, Logan Everett, Suzanne Wardell, Thomas Westerling, Robert Baldi, Tricia M. Wright, Kendall Tavares, Alex Yllanes, Yeeun Bae, Jeremy T. Blitzer, Craig Logsdon, Daniel P. Rakiec, David A. Ruddy, Tiancong Jiang, Gloria Broadwater, Terry Hyslop, Allison Hall, Muriel Laine, Linda Phung, Geoffrey L. Greene, Lesley-Ann Martin, Sunil Pancholi, Mitch Dowsett, Simone Detre, Jeffrey R. Marks, Gregory E. Crawford, Myles Brown, John D. Norris, Ching-yi Chang, and Donald P. McDonnell

Supplementary Information



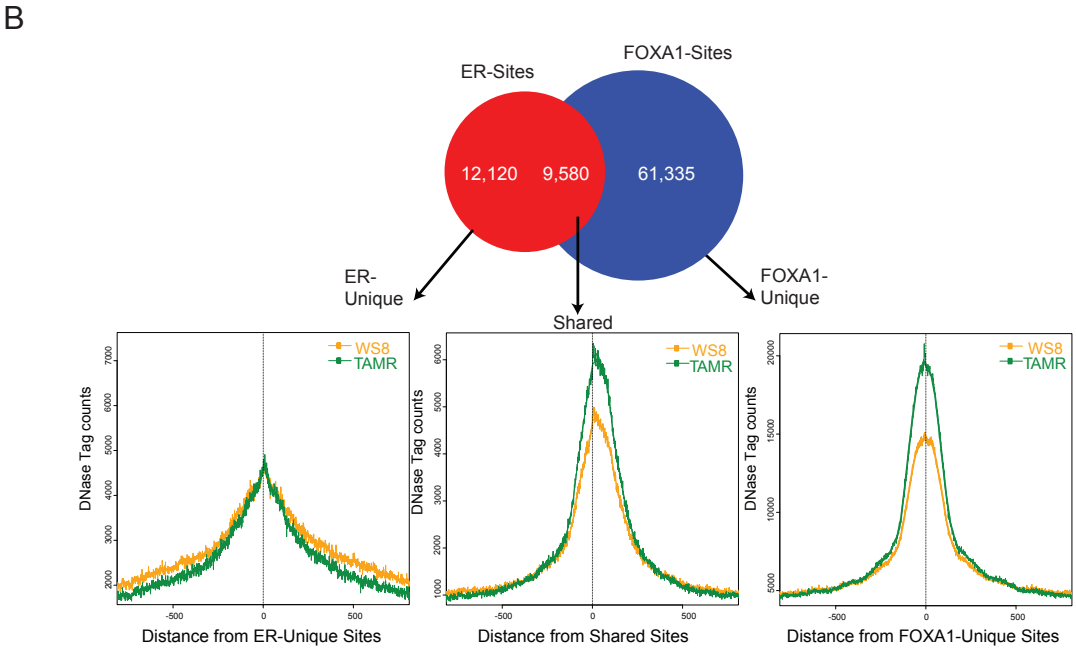


Figure S1. Related to Figure 1

A) Validation of TAMR tumor growth. Left is the growth curves of 4 individual founder tumors (formed by injection of the TAMR cells), and right four panels show the growth patterns of the resulting TAMR tumor models with tamoxifen/no tamoxifen (n = 4-5 for each group) treatment. B) Schematic representing comparison of DNase Seq aligned with published ER and FOXA1 ChIP Seq (Hurtado, Holmes et al. 2011); A venn diagram was first created for ER and FOXA1 known binding sites which define subsets for 9580 coincident sites, 12,120 sites private to ER, and 61,335 sites private to FOXA1. The aggregate DNase signal profile was then plotted for each subset with a 500bp window to each side of the features.

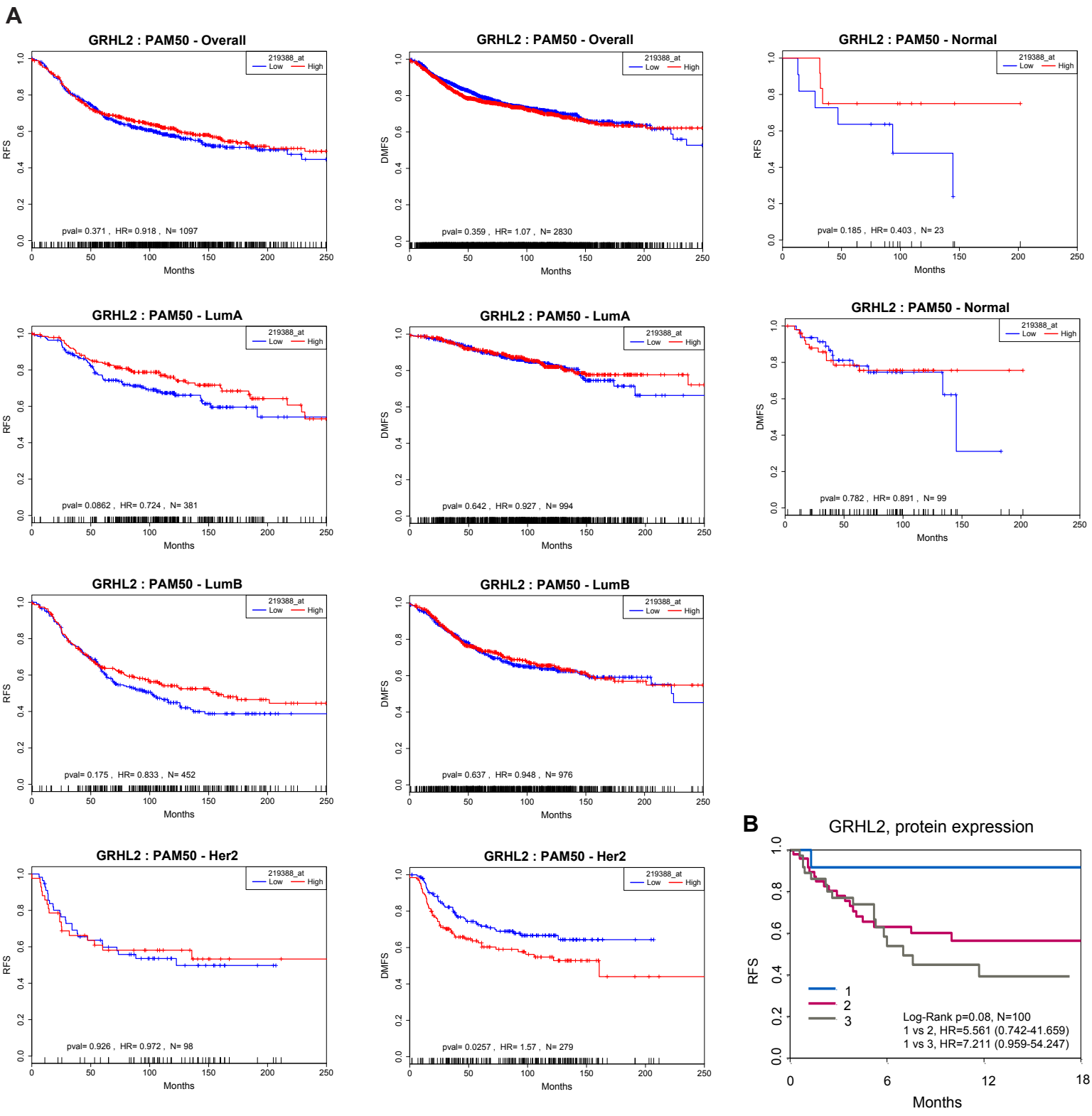
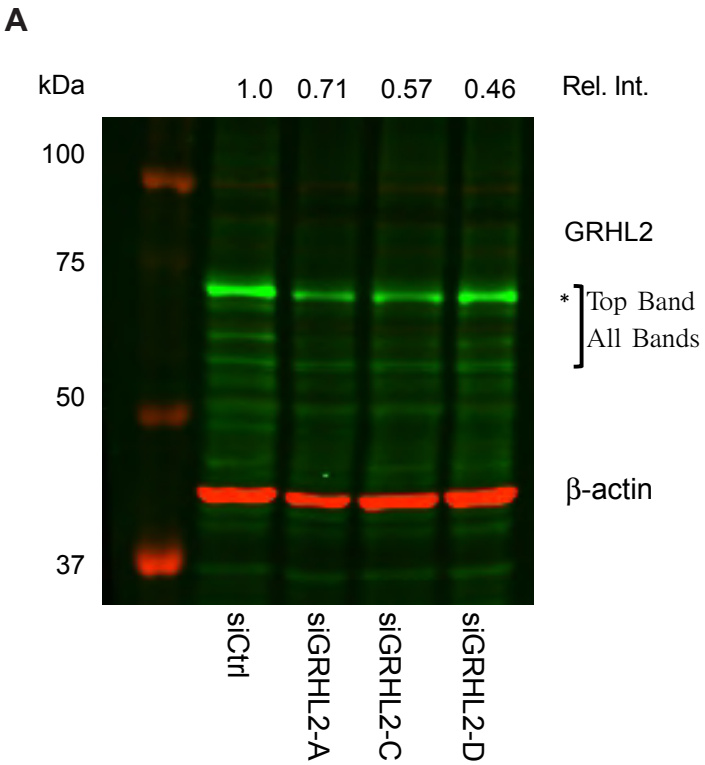
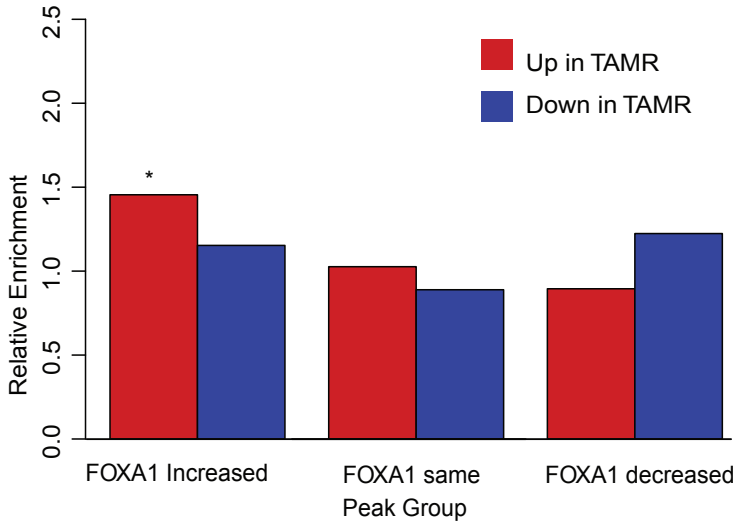
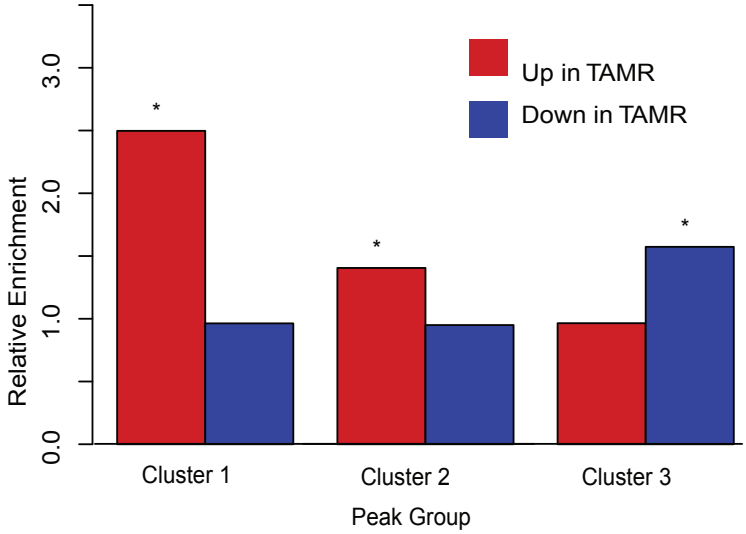


Figure S2, related to Figure 4

A) Assessment of mRNA expression of GRHL2 as predictor of outcome (Recurrence Free Survival (RFS) and Distant Metastasis Free Survival (DMFS)) using the Gene Analytics Tool in overall patient populations, or by PAM50 subgroup. B) Kaplan Meier- estimator of time to recurrence (RFS, months) of breast tumors derived from all patients (patient characteristics see **Table S2**), stratified based on GRHL2 protein expression (1 = low, 3 = high). Statistical significance to determine differences between groups based on scoring was determined using Log-Rank test, p = 0.08; Hazard ratios were determined using univariate Cox proportional hazards model, p=0.14.







		siCtrl	siA	siC	siD
GRHL2 Ab	All Bands	34.22	13.45	17.81	23.69
	Top Band	21.60	7.89	9.87	14.41
Actin Ab		72.73	37.20	58.03	106.61

Normalized to Actin

GRHL2 Ab	All Bands	0.471	0.362	0.307	0.222
	Top Band	0.300	0.212	0.170	0.135

Normalized to siCtrl

GRHL2 Ab	All Bands	1	0.768	0.652	0.472
	Top Band	1	0.714	0.572	0.455

Figure S3, related to Figure 5

A)Western blot showing efficacy of GRHL2 knockdown using three different siRNA targeting GRHL2. TamR cells were transfected with siRNA, final concentration 60nM, for 72hrs in phenol red-free DMEM:F12 containing 8% 2x charcoal-stripped FBS. After which, cells were scraped off plate and lysed with RIPA buffer. Cell lysates were resolved on SDS-PAGE and GRHL2 expression normalized to β- actin and then to the control siRNA sample. Knockdown of GRHL2 protein was achieved whether the top band or all bands was assessed. This experiment was repeated with biological replicates n=3 and this figure is a representative image of such replicates. Comparison of differential gene expression in TAMR and MCF7-WS8 xenografts data with subcategories of regulatory elements indicates subset of genes increased in vivo in tamoxifen resistance, downstream of FOXA1/GRHL2 collaboration. Relative enrichment of genes differentially expressed in TAMR relative to MCF7-WS8 within 10kb of B) FOXA1 binding events alone (as in Figure 1), C) FOXA1 increased binding events subdivided based on histone marks (as in Figure 2B)

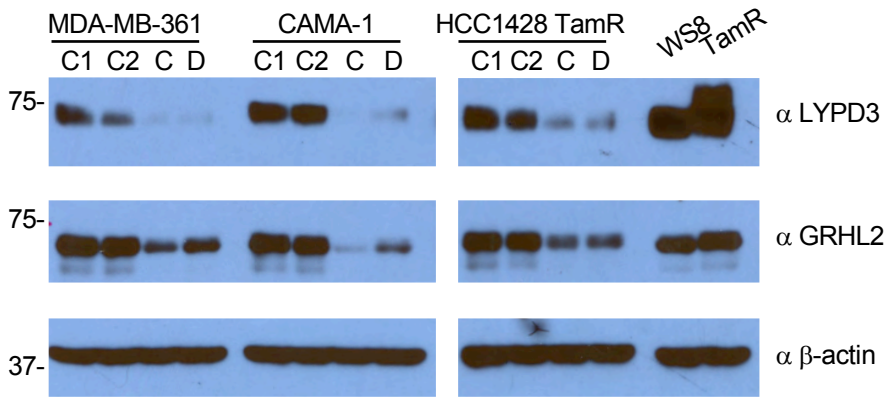


Figure S4, related to Figure 6

LYPD3 protein expression was assessed in MDA-MB-361, CAMA-1 and HCC1428-TamR cells following siRNA knockdown of GRHL2 (C and D) or control siRNAs (C1 and C2) as indicated. Cells were collected following 3 days of knockdown. Whole cell extracts were analyzed via Western Blot using the indicated antibodies. All 3 cell lines are ER+; MDA-MB-361 and CAMA-1 were obtained from ATCC and cultured in DMEM (MDA-MB-361) and Eagle Minimum Essential Media (CAMA-1) supplemented with 8% FBS, 0.1 mM non-essential amino acids and 1 mM sodium pyruvate. HCC1428-TamR is an *in vitro*-derived tamoxifen resistant cell line and was cultured in phenol-red free RPMI with double charcoal-stripped fetal bovine serum, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate and supplemented with 100 nM 4-hydroxytamoxifen and 0.1 nM 17β-estradiol.

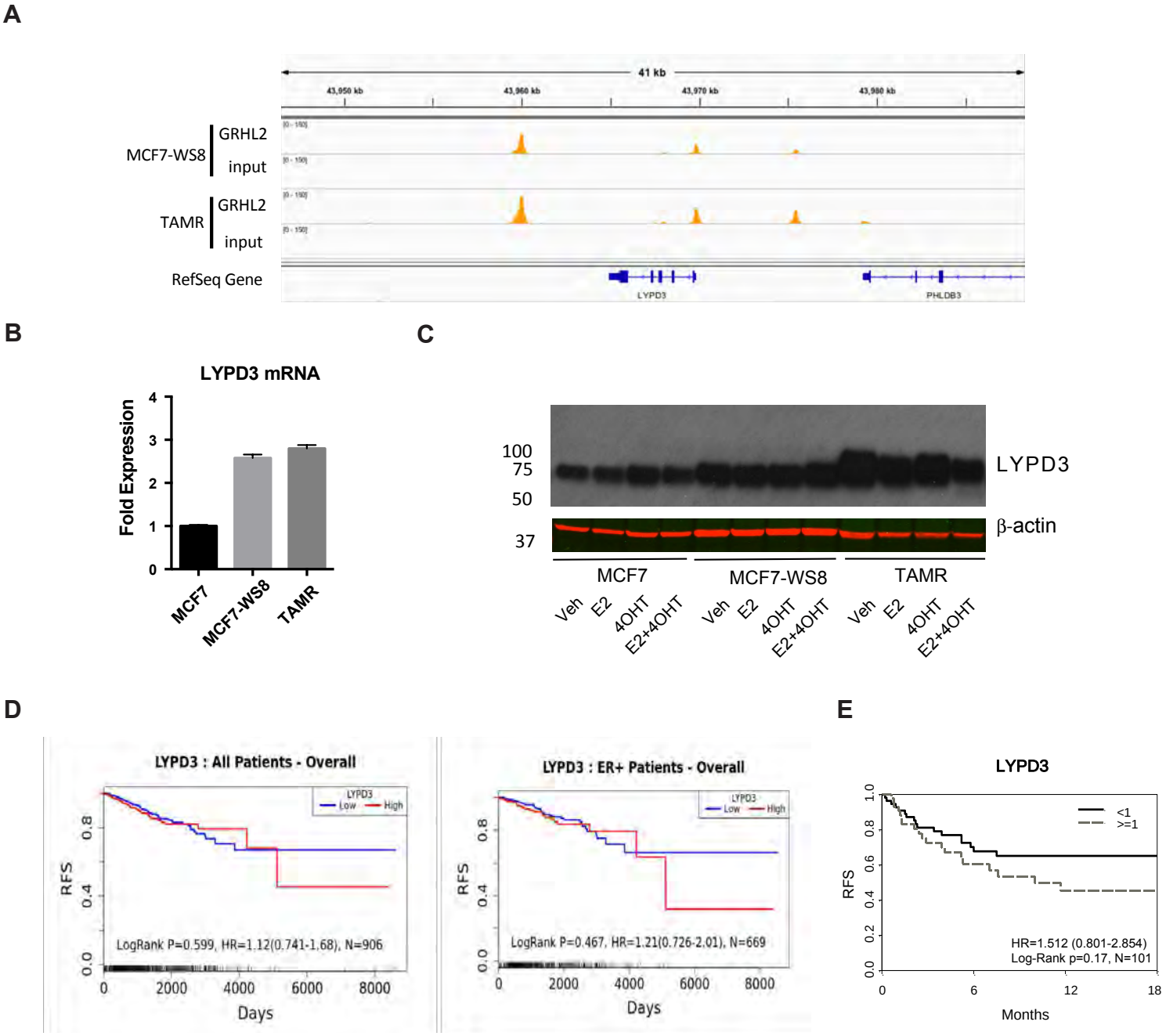


Figure S5, related to Figure 6

A) Screenshot of Integrative Genomic Viewer (IGV) ((Robinson, Thorvaldsdottir et al. 2011, Thorvaldsdottir, Robinson et al. 2013) window demonstrating GRHL2 binding events around LYPD3 genomic locus in MCF7-WS8 and TAMR. LYPD3 mRNA (B) and protein expression (C) in MCF7 (ATCC), MCF7-WS8, and TAMR cell lines. For RNA, Cells were grown in charcoal stripped serum for 72 hrs and RNA expression was assessed by qPCR, normalized to 36B4. For protein, cells were grown for 48 hrs in charcoal stripped serum and treated with 10nM E2, 100nM 4OHT, or 100nM ICI as indicated for 24 hrs. Whole cell extracts were analyzed via Western Blot using the indicated antibodies. (D) Correlation between LYPD3 mRNA expression and relapse-free survival (RFS) was assess in TCGA breast tumor dataset(TCGA-BRCA(June 2018)). RPKM values of TCGA data(RNA-SeqV2) was downloaded from the Broad GDAC Firehose using the R package "TCGA2STAT". Matching clinical annotations for these patients were sourced from UCSC (https://tcga.xenahubs.net/download/TCGA.BRCA.sampleMap/BRCA_clinicalMatrix). Only tumor samples with sufficient clinical annotations were used in this analysis. Total samples was 1093 with sufficient clinical evidence. PAM50 calls are limited on this clinical set to 840 samples so the R package genefu was used to call molecular subtypes using the pam50.robust centroids taking care to rename KNTC2 and CDCA1 to the equivalent gene names used in TCGA data. Using this approach there was a high-degree of correlation to existing PAM50 calls. Survival analysis was performed using R with the "survival" package. All p-values shown were derived using the log-rank approach. Neither all-comers or ER+ patients show significant differences on RFS. Several cut points were explored with no major changes. This is at 60%.(60% of cohort low, 40% high). No PAM50 subtype showed significant differences in survival. (E) Kaplan Meier-estimator of time to recurrence (RFS, months) of tumors derived from all patients, stratified based on LYPD3 protein expression (<1 no staining, >=1 positive staining). Statistical significance to determine differences between groups based on scoring was determined using Log-Rank test, with p=0.17; Hazard ratio was determined using univariate Cox proportional hazards regression, p=0.2025.

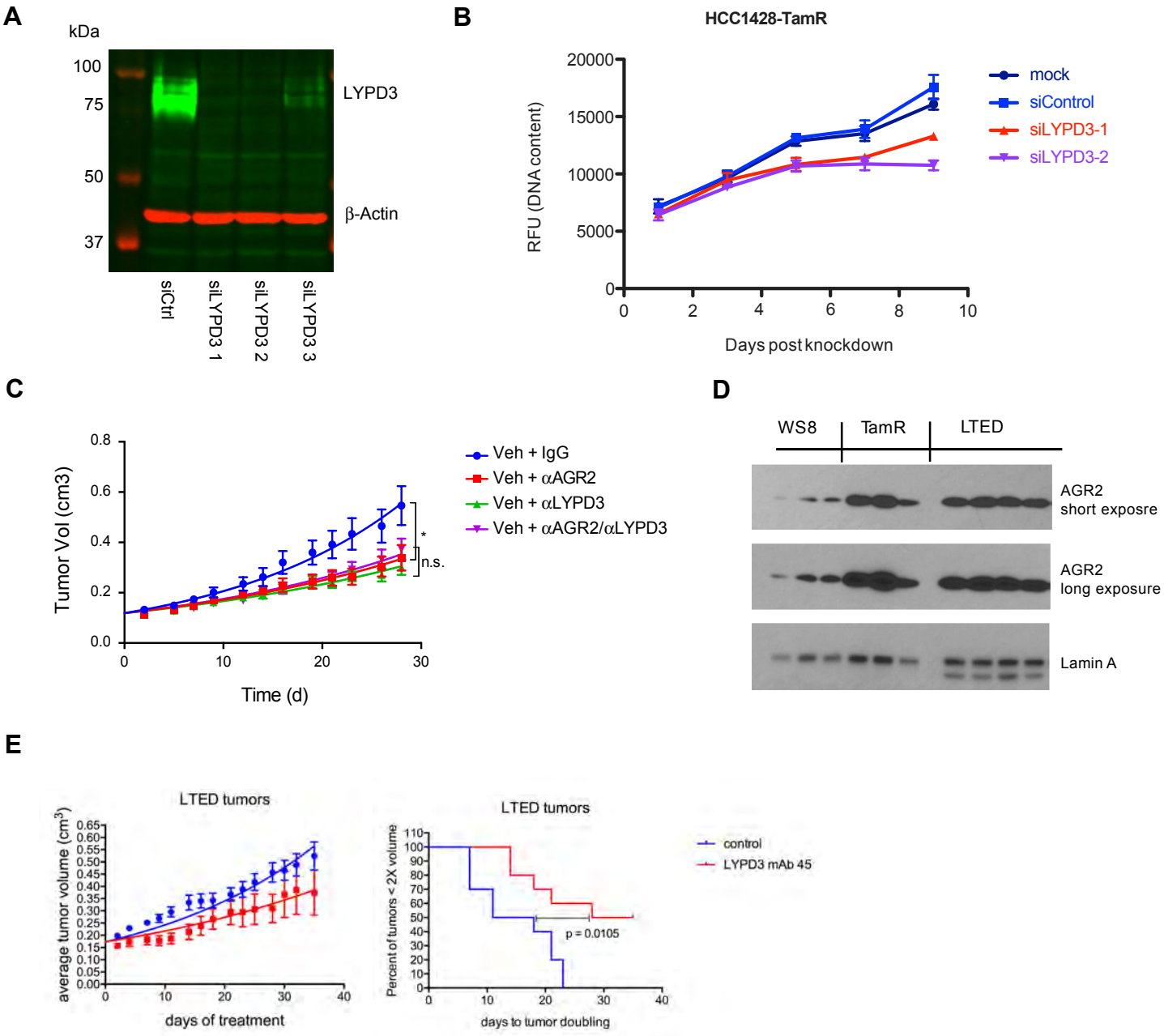


Figure S6, related to Figure 7

A) Knockdown efficiency of siRNAs against LYPD3 was assessed in TAMR cells transfected with control siRNA or 3 different siRNA against LYPD3. Cell lysates were resolved on SDS-PAGE and the LYPD3 protein levels were determined by immunoblots using antibody against LYPD3. β- actin antibody was used as normalization control. This experiment was repeated for a total of 3 times and shown is representative data. B) HCC1428-TAMR cells were transfected with RNAiMAX only (mock), siCtrl or 2 unique siRNA sequences targeting LYPD3 and monitored for cell growth for 9 days. The experiment was repeated 4 times with similar results and representative data was shown. C) TAMR tumors were established in tamoxifen-treated mice and then received further treatment with anti-LYPD3, anti-AGR2 or IgG control antibodies (15mg/kg, 2X weekly) or both anti-LYPD3 + anti-AGR2 (7.5mg/kg/antibody, 2X weekly) and were monitored over time. Average tumor volume -/+ SEM is plotted (n = 8-9 per group) and significance was determined by two-way ANOVA followed by Bonferroni's multiple comparison test. $p < 0.05$ as indicated by *. D) The expression of AGR2 in xenograft tumor models (MCF7-WS8, TamR and LTED) was assessed using immunoblot with anti-AGR2 antibody. Antibody against Lamin A was used as a normalization control. E) Ovariectomized J/nu mice bearing LTED (resistant to long term estrogen withdrawal) xenograft tumors were randomized to treatment with 45 mg/kg IgG or LYPD3 antibody. Data presented indicate the average tumor volume for each group (mean +/- SEM) at each time point of tumor measurement (left). Time to progression analysis (Kaplan-Meier) analysis was conducted using 2X tumor volume (twice the tumor volume at time of randomization) as an endpoint(right). High variability confounded the results of 2-way ANOVA analysis, and therefore no significant differences were detected between treatment groups. Time to progression analysis revealed a significant ($p = 0.0105$) delay in tumor progression.

In all samples by ATR

Classif=Primary

Variable	N	Median	Minimum	Lower Quartile	Upper Quartile	Maximum
MeanATR_LYPD	41	1.0	0.0	0.8	1.3	2.0
MeanATR_GRHL2	42	2.0	0.0	1.0	2.3	3.0

Classif=Recurrence

Variable	N	Median	Minimum	Lower Quartile	Upper Quartile	Maximum
MeanATR_LYPD	38	1.0	0.0	0.8	1.5	3.0
MeanATR_GRHL2	39	1.7	0.0	1.0	2.5	3.0

Wilcoxon Rank Sum Test p-values	
MeanATR_LYPD3	0.7544
MeanATR_GRHL2	0.7823

In all samples by TRA

Classif=Primary

Variable	N	Median	Minimum	Lower Quartile	Upper Quartile	Maximum
MeanTRA_LYPD	30	1.0	0.0	0.5	1.5	2.0
MeanTRA_GRHL2	30	2.0	0.0	1.0	2.5	3.0

Classif=Recurrence

Variable	N	Median	Minimum	Lower Quartile	Upper Quartile	Maximum
MeanTRA_LYPD	28	1.0	0.0	0.5	1.3	2.0
MeanTRA_GRHL2	28	2.0	0.0	1.4	2.7	3.0

Wilcoxon Rank Sum p-values	
MeanATR_LYPD3	0.7155
MeanATR_GRHL2	0.4616

Figure S7, related to Figure 7
The intensity of LYPD3 and GRHL2 staining in primary vs recurrent tumors was assessed in two different cohorts of patient TMAs (ATR and TRA, Drury et al, 2011 and Xiao et al, 2018). No statistical difference in expression was found.

Table S1, related to Figure 2: Endogenous FOXA1 interactors in TAMR cells

Swissprot_Id	Ensembl_Id	Symbol	Description
ACINU_HUMAN	ENSG00000100813	ACIN1	apoptotic chromatin condensation inducer 1 [Source:HGNC Symbol;Acc:17066]
AKP13_HUMAN	ENSG00000170776	AKAP13	A kinase (PRKA) anchor protein 13 [Source:HGNC Symbol;Acc:371]
P5CS_HUMAN	ENSG00000059573	ALDH18A1	aldehyde dehydrogenase 18 family, member A1 [Source:HGNC Symbol;Acc:9722]
COPD_HUMAN	ENSG00000095139	ARCN1	archain 1 [Source:HGNC Symbol;Acc:649]
COPD_HUMAN	ENSG00000269382	ARCN1	archain 1 [Source:HGNC Symbol;Acc:649]
AT1A1_HUMAN	ENSG00000163399	ATP1A1	ATPase, Na+/K+ transporting, alpha 1 polypeptide [Source:HGNC Symbol;Acc:799]
AT2A2_HUMAN	ENSG00000174437	ATP2A2	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2 [Source:HGNC Symbol;Acc:812]
ATPA_HUMAN	ENSG00000152234	ATP5A1	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit 1, cardiac muscle [Source:HGNC Symbol;Acc:823]
CN166_HUMAN	ENSG00000087302	C14orf166	chromosome 14 open reading frame 166 [Source:HGNC Symbol;Acc:23169]
CAZA2_HUMAN	ENSG00000198898	CAPZA2	capping protein (actin filament) muscle Z-line, alpha 2 [Source:HGNC Symbol;Acc:1490]
COPG1_HUMAN	ENSG00000181789	COPG1	coatomer protein complex, subunit gamma 1 [Source:HGNC Symbol;Acc:2236]
CPSF2_HUMAN	ENSG00000165934	CPSF2	cleavage and polyadenylation specific factor 2, 100kDa [Source:HGNC Symbol;Acc:2325]
CTNB1_HUMAN	ENSG00000168036	CTNNB1	catenin (cadherin-associated protein), beta 1, 88kDa [Source:HGNC Symbol;Acc:2514]
RT29_HUMAN	ENSG00000132676	DAP3	death associated protein 3 [Source:HGNC Symbol;Acc:2673]
DDX1_HUMAN	ENSG00000079785	DDX1	DEAD (Asp-Glu-Ala-Asp) box helicase 1 [Source:HGNC Symbol;Acc:2734]
DDX18_HUMAN	ENSG00000088205	DDX18	DEAD (Asp-Glu-Ala-Asp) box polypeptide 18 [Source:HGNC Symbol;Acc:2741]
DECR_HUMAN	ENSG00000104325	DECR1	2,4-dienoyl CoA reductase 1, mitochondrial [Source:HGNC Symbol;Acc:2753]
DEK_HUMAN	ENSG00000124795	DEK	DEK oncogene [Source:HGNC Symbol;Acc:2768]
DHX15_HUMAN	ENSG00000109606	DHX15	DEAH (Asp-Glu-Ala-His) box helicase 15 [Source:HGNC Symbol;Acc:2738]
DHX30_HUMAN	ENSG00000132153	DHX30	DEAH (Asp-Glu-Ala-His) box helicase 30 [Source:HGNC Symbol;Acc:16716]
DIAP3_HUMAN	ENSG00000139734	DIAPH3	diaphanous-related formin 3 [Source:HGNC Symbol;Acc:15480]
DNJA2_HUMAN	ENSG00000069345	DNAJA2	DnaJ (Hsp40) homolog, subfamily A, member 2 [Source:HGNC Symbol;Acc:14884]
DYN2_HUMAN	ENSG00000079805	DNM2	dynamin 2 [Source:HGNC Symbol;Acc:2974]
ECH1_HUMAN	ENSG00000104823	ECH1	enoyl CoA hydratase 1, peroxisomal [Source:HGNC Symbol;Acc:3149]
EF1G_HUMAN	ENSG00000254772	EEF1G	eukaryotic translation elongation factor 1 gamma [Source:HGNC Symbol;Acc:3213]
IF4A3_HUMAN	ENSG00000141543	EIF4A3	eukaryotic translation initiation factor 4A3 [Source:HGNC Symbol;Acc:18683]
EPIPL_HUMAN	ENSG00000261150	EPPK1	epiplakin 1 [Source:HGNC Symbol;Acc:15577]
ERF1_HUMAN	ENSG00000120705	ETF1	eukaryotic translation termination factor 1 [Source:HGNC Symbol;Acc:3477]
EWS_HUMAN	ENSG00000182944	EWSR1	EWS RNA-binding protein 1 [Source:HGNC Symbol;Acc:3508]
FANCI_HUMAN	ENSG00000140525	FANCI	Fanconi anemia, complementation group I [Source:HGNC Symbol;Acc:25568]
FBLL1_HUMAN	ENSG00000188573	FBLL1	fibrillarin-like 1 [Source:HGNC Symbol;Acc:35458]
FEN1_HUMAN	ENSG00000168496	FEN1	flap structure-specific endonuclease 1 [Source:HGNC Symbol;Acc:3650]
FLOT2_HUMAN	ENSG00000132589	FLOT2	flotillin 2 [Source:HGNC Symbol;Acc:3758]
GBB1_HUMAN	ENSG00000078369	GNB1	guanine nucleotide binding protein (G protein), beta polypeptide 1 [Source:HGNC Symbol;Acc:4396]
GRHL2_HUMAN	ENSG00000083307	GRHL2	grainyhead-like 2 (Drosophila) [Source:HGNC Symbol;Acc:2799]

TF3C3_HUMAN	ENSG00000119041	GTF3C3	general transcription factor IIIC, polypeptide 3, 102kDa [Source:HGNC Symbol;Acc:4666]
H1X_HUMAN	ENSG00000184897	H1FX	H1 histone family, member X [Source:HGNC Symbol;Acc:4722]
HDAC1_HUMAN	ENSG00000116478	HDAC1	histone deacetylase 1 [Source:HGNC Symbol;Acc:4852]
H2B1B_HUMAN	ENSG00000196226	HIST1H2BB	histone cluster 1, H2bb [Source:HGNC Symbol;Acc:4751]
HMGB2_HUMAN	ENSG00000164104	HMGB2	high mobility group box 2 [Source:HGNC Symbol;Acc:5000]
ROAA_HUMAN	ENSG00000197451	HNRNPAB	heterogeneous nuclear ribonucleoprotein A/B [Source:HGNC Symbol;Acc:5034]
HNRPF_HUMAN	ENSG00000169813	HNRNPF	heterogeneous nuclear ribonucleoprotein F [Source:HGNC Symbol;Acc:5039]
HNRH3_HUMAN	ENSG00000096746	HNRNPH3	heterogeneous nuclear ribonucleoprotein H3 (2H9) [Source:HGNC Symbol;Acc:5043]
HNRPR_HUMAN	ENSG00000125944	HNRNPR	heterogeneous nuclear ribonucleoprotein R [Source:HGNC Symbol;Acc:5047]
HORN_HUMAN	ENSG00000197915	HRNR	hornerin [Source:HGNC Symbol;Acc:20846]
DHB4_HUMAN	ENSG00000133835	HSD17B4	hydroxysteroid (17-beta) dehydrogenase 4 [Source:HGNC Symbol;Acc:5213]
GRP78_HUMAN	ENSG00000044574	HSPA5	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa) [Source:HGNC Symbol;Acc:5238]
GRP75_HUMAN	ENSG00000113013	HSPA9	heat shock 70kDa protein 9 (mortalin) [Source:HGNC Symbol;Acc:5244]
ILF3_HUMAN	ENSG00000129351	ILF3	interleukin enhancer binding factor 3, 90kDa [Source:HGNC Symbol;Acc:6038]
IPO4_HUMAN	ENSG00000196497	IPO4	importin 4 [Source:HGNC Symbol;Acc:19426]
SYK_HUMAN	ENSG00000065427	KARS	lysyl-tRNA synthetase [Source:HGNC Symbol;Acc:6215]
KHDR1_HUMAN	ENSG00000121774	KHDRBS1	KH domain containing, RNA binding, signal transduction associated 1 [Source:HGNC Symbol;Acc:18116]
FUBP2_HUMAN	ENSG00000088247	KHSRP	KH-type splicing regulatory protein [Source:HGNC Symbol;Acc:6316]
K1C16_HUMAN	ENSG00000186832	KRT16	keratin 16 [Source:HGNC Symbol;Acc:6423]
MDHM_HUMAN	ENSG00000146701	MDH2	malate dehydrogenase 2, NAD (mitochondrial) [Source:HGNC Symbol;Acc:6971]
MDHM_HUMAN	ENSG00000262847	MDH2	malate dehydrogenase 2, NAD (mitochondrial) [Source:HGNC Symbol;Acc:6971]
MLH1_HUMAN	ENSG00000076242	MLH1	mutL homolog 1 [Source:HGNC Symbol;Acc:7127]
RT31_HUMAN	ENSG00000102738	MRPS31	mitochondrial ribosomal protein S31 [Source:HGNC Symbol;Acc:16632]
MSH6_HUMAN	ENSG00000116062	MSH6	mutS homolog 6 [Source:HGNC Symbol;Acc:7329]
MBB1A_HUMAN	ENSG00000132382	MYBBP1A	MYB binding protein (P160) 1a [Source:HGNC Symbol;Acc:7546]
MYO5B_HUMAN	ENSG00000167306	MYO5B	myosin VB [Source:HGNC Symbol;Acc:7603]
MYO5C_HUMAN	ENSG00000128833	MYO5C	myosin VC [Source:HGNC Symbol;Acc:7604]
NELFB_HUMAN	ENSG00000188986	NELFB	negative elongation factor complex member B [Source:HGNC Symbol;Acc:24324]
NOP56_HUMAN	ENSG00000101361	NOP56	NOP56 ribonucleoprotein [Source:HGNC Symbol;Acc:15911]
NOP58_HUMAN	ENSG00000055044	NOP58	NOP58 ribonucleoprotein [Source:HGNC Symbol;Acc:29926]
NUMA1_HUMAN	ENSG00000137497	NUMA1	nuclear mitotic apparatus protein 1 [Source:HGNC Symbol;Acc:8059]
RRP5_HUMAN	ENSG00000148843	PDCD11	programmed cell death 11 [Source:HGNC Symbol;Acc:13408]
K6PP_HUMAN	ENSG00000067057	PFKP	phosphofructokinase, platelet [Source:HGNC Symbol;Acc:8878]
PHB_HUMAN	ENSG00000167085	PHB	prohibitin [Source:HGNC Symbol;Acc:8912]
PICAL_HUMAN	ENSG00000073921	PICALM	phosphatidylinositol binding clathrin assembly protein [Source:HGNC Symbol;Acc:15514]
PKP3_HUMAN	ENSG00000184363	PKP3	plakophilin 3 [Source:HGNC Symbol;Acc:9025]
PLEC_HUMAN	ENSG00000178209	PLEC	plectin [Source:HGNC Symbol;Acc:9069]

PLEC_HUMAN	ENSG00000261109	PLEC	plectin [Source:HGNC Symbol;Acc:9069]
PLST_HUMAN	ENSG00000268767	PLS3	plastin 3 [Source:HGNC Symbol;Acc:9091]
PLST_HUMAN	ENSG00000102024	PLS3	plastin 3 [Source:HGNC Symbol;Acc:9091]
PTBP3_HUMAN	ENSG00000119314	PTBP3	polypyrimidine tract binding protein 3 [Source:HGNC Symbol;Acc:10253]
HACD3_HUMAN	ENSG00000074696	PTPLAD1	protein tyrosine phosphatase-like A domain containing 1 [Source:HGNC Symbol;Acc:24175]
RBM14_HUMAN	ENSG00000239306	RBM14	RNA binding motif protein 14 [Source:HGNC Symbol;Acc:14219]
RBM14_HUMAN	ENSG00000248643	RBM14-RBM4	RBM14-RBM4 readthrough [Source:HGNC Symbol;Acc:38840]
RBM28_HUMAN	ENSG00000106344	RBM28	RNA binding motif protein 28 [Source:HGNC Symbol;Acc:21863]
RFC2_HUMAN	ENSG00000049541	RFC2	replication factor C (activator 1) 2, 40kDa [Source:HGNC Symbol;Acc:9970]
RFC2_HUMAN	ENSG00000261911	RFC2	replication factor C (activator 1) 2, 40kDa [Source:HGNC Symbol;Acc:9970]
RFC5_HUMAN	ENSG00000111445	RFC5	replication factor C (activator 1) 5, 36.5kDa [Source:HGNC Symbol;Acc:9973]
RGS20_HUMAN	ENSG00000147509	RGS20	regulator of G-protein signaling 20 [Source:HGNC Symbol;Acc:14600]
RL3_HUMAN	ENSG00000100316	RPL3	ribosomal protein L3 [Source:HGNC Symbol;Acc:10332]
RL5_HUMAN	ENSG00000122406	RPL5	ribosomal protein L5 [Source:HGNC Symbol;Acc:10360]
RL8_HUMAN	ENSG00000161016	RPL8	ribosomal protein L8 [Source:HGNC Symbol;Acc:10368]
RRP12_HUMAN	ENSG00000052749	RRP12	ribosomal RNA processing 12 homolog (S. cerevisiae) [Source:HGNC Symbol;Acc:29100]
RL1D1_HUMAN	ENSG00000171490	RSL1D1	ribosomal L1 domain containing 1 [Source:HGNC Symbol;Acc:24534]
RUVB2_HUMAN	ENSG00000183207	RUVBL2	RuvB-like AAA ATPase 2 [Source:HGNC Symbol;Acc:10475]
SAFB1_HUMAN	ENSG00000160633	SAFB	scaffold attachment factor B [Source:HGNC Symbol;Acc:10520]
SF3B3_HUMAN	ENSG00000189091	SF3B3	splicing factor 3b, subunit 3, 130kDa [Source:HGNC Symbol;Acc:10770]
SF3B4_HUMAN	ENSG00000143368	SF3B4	splicing factor 3b, subunit 4, 49kDa [Source:HGNC Symbol;Acc:10771]
SF3B4_HUMAN	ENSG00000263977	SF3B4	splicing factor 3b, subunit 4, 49kDa [Source:HGNC Symbol;Acc:10771]
SIN3A_HUMAN	ENSG00000169375	SIN3A	SIN3 transcription regulator family member A [Source:HGNC Symbol;Acc:19353]
AAAT_HUMAN	ENSG00000105281	SLC1A5	solute carrier family 1 (neutral amino acid transporter), member 5 [Source:HGNC Symbol;Acc:10943]
LAT1_HUMAN	ENSG00000103257	SLC7A5	solute carrier family 7 (amino acid transporter light chain, L system), member 5 [Source:HGNC Symbol;Acc:11063]
CYTSB_HUMAN	ENSG00000128487	SPECC1	sperm antigen with calponin homology and coiled-coil domains 1 [Source:HGNC Symbol;Acc:30615]
SPTN2_HUMAN	ENSG00000173898	SPTBN2	spectrin, beta, non-erythrocytic 2 [Source:HGNC Symbol;Acc:11276]
SRSF7_HUMAN	ENSG00000115875	SRSF7	serine/arginine-rich splicing factor 7 [Source:HGNC Symbol;Acc:10789]
SVIL_HUMAN	ENSG00000197321	SVIL	supervillin [Source:HGNC Symbol;Acc:11480]
TADBP_HUMAN	ENSG00000120948	TARDBP	TAR DNA binding protein [Source:HGNC Symbol;Acc:11571]
AP2C_HUMAN	ENSG00000087510	TFAP2C	transcription factor AP-2 gamma (activating enhancer binding protein 2 gamma) [Source:HGNC Symbol;Acc:11744]
TBB4A_HUMAN	ENSG00000104833	TUBB4A	tubulin, beta 4A class IVa [Source:HGNC Symbol;Acc:20774]
TXLNA_HUMAN	ENSG00000084652	TXLNA	taxilin alpha [Source:HGNC Symbol;Acc:30685]
UBF1_HUMAN	ENSG00000108312	UBTF	upstream binding transcription factor, RNA polymerase I [Source:HGNC Symbol;Acc:12511]
UN45A_HUMAN	ENSG00000140553	UNC45A	unc-45 homolog A (C. elegans) [Source:HGNC Symbol;Acc:30594]
VDAC1_HUMAN	ENSG00000213585	VDAC1	voltage-dependent anion channel 1 [Source:HGNC Symbol;Acc:12669]
YES_HUMAN	ENSG00000176105	YES1	v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1 [Source:HGNC Symbol;Acc:12841]

Table S2, related to Figure 4: Summary of patient characteristics

	Average	**SD**
Age at Diagnosis	52.05	13.15
Time to Recurrence	3.26	2.86
Receptor status		
ERA	positive	71
	negative	55
	unknown	34
PRA	positive	57
	negative	60
	borderline	6
	unknown	34
Histological Diagnosis		
Ductal	120	
Infiltrating duct and lobular	12	
Inflammatory carcinoma	8	
Adenocarcinoma	1	
Carcinoma, NOS	1	
Comedocarcinoma	1	
Lobular carcinoma	11	
Mucinous adenocarcinoma	3	
Paget disease and intraductal	1	
Phyllodes tumor	1	
Tubular	1	