Published in final edited form as:

Sci Transl Med. 2017 August 09; 9(402): . doi:10.1126/scitranslmed.aai7993.

Genomic profiling of ER+ breast cancers after short-term estrogen suppression reveals alterations associated with endocrine resistance

J.M. Giltane, MD, PhD^{1,2,†}, K.E. Hutchinson, PhD^{3,†}, T.P. Stricker, MD, PhD^{1,2,†}, L. Formisano, MD³, C.D. Young, PhD³, M.V. Estrada, MD², M.J. Nixon, PhD⁴, L. Du, PhD⁵, V. Sanchez², P. Gonzalez Ericsson, MD², M.G. Kuba, MD¹, M.E. Sanders, MD^{1,2}, X.J. Mu, PhD^{6,7}, E.M. Van Allen, MD⁶, N. Wagle, MD^{6,7}, I. Mayer, MD, MSCI^{2,3}, V. Abramson, MD^{2,3}, H. Gómez, MD⁸, M. Rizzo, MD⁹, W. Toy, PhD¹⁰, S. Chandarlapaty, MD, PhD¹⁰, E.L. Mayer, MD, MPH⁷, J. Christiansen, PhD¹¹, D. Murphy, PhD¹¹, K. Fitzgerald, PhD¹², K. Wang, PhD¹¹, J.S. Ross, MD¹², V.A. Miller, MD¹², P.J. Stephens, PhD¹², R. Yelensky, PhD¹², L. Garraway, MD, PhD^{6,7,13}, Y. Shyr, PhD⁵, I. Meszoely, MD^{2,14}, J.M. Balko, PharmD, PhD^{2,3,4}, and C.L. Arteaga, MD^{2,3,4,*}

¹Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN

²Breast Cancer Research Program, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN

³Department of Medicine, Vanderbilt University Medical Center, Nashville, TN

⁴Department of Cancer Biology, Vanderbilt University School of Medicine, Nashville, TN

⁵Vanderbilt Center for Quantitative Sciences, Vanderbilt University School of Medicine, Nashville, TN

⁶Broad Institute of MIT and Harvard, Cambridge, MA

⁷Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA

⁸Instituto Nacional de Enfermedades Neoplásias, Lima, Perú

⁹Department of Surgery, Emory University, Atlanta, GA

¹⁰Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY

*To whom correspondence should be addressed: Carlos L. Arteaga, MD, Div. of Oncology, Vanderbilt University Medical Center, 2220 Pierce Avenue, 777 Preston Research Building, Nashville, TN 37232-6307, carlos.artea@vanderbilt.edu, Tel. 615.343.6653.

†These authors contributed equally to this work.

Author Contributions: Trial design/conception: IM, CLA

Clinical research procedures/regulatory approvals: IM, CLA, MR, HG, JMG

Data acquisition / analysis: All authors

Writing manuscript: JMG, KEH, TPS, CLA

Competing Interests: The authors declare no competing interests.

Data And Material Access. RNA sequencing data generated in this study can be accessed through dbGaP number phs000857. Whole exome sequencing data have been deposited and the dbGaP number is pending assignment.

¹¹Genoptix Medical Laboratories, Carlsbad, CA

¹²Foundation Medicine, Incorporated, Cambridge, MA

¹³Joint Center for Cancer Precision Medicine, Dana-Farber Cancer Institute, Boston, MA

¹⁴Department of Surgery, Vanderbilt University Medical Center, Nashville, TN

Abstract

Proliferative inhibition of estrogen-receptor positive (ER+) breast cancers after short-term antiestrogen therapy correlates with long-term patient outcome. We profiled 155 ER+/HER2– early breast cancers from 143 patients treated with the aromatase inhibitor letrozole for 10–21 days before surgery. Twenty-one percent of tumors remained highly proliferative suggesting these tumors harbor alterations associated with intrinsic endocrine therapy resistance. Whole-exome sequencing revealed a correlation between 8p11–12 and 11q13 gene amplifications, including *FGFR1* and *CCND1*, respectively, and high Ki67. We corroborated these findings in a separate cohort of serial pre-treatment, post-neoadjuvant chemotherapy, and recurrent ER+ tumors. Combined inhibition of FGFR1 and CDK4/6 reversed antiestrogen resistance in ER+ *FGFR1*/*CCND1* co-amplified CAMA1 breast cancer cells. RNA sequencing of letrozole-treated tumors revealed intrachromosomal *ESR1* fusion transcripts and gene expression signatures in cancers with high Ki67, indicative of enhanced E2F-mediated transcription and cell cycle processes. These data suggest short-term pre-operative estrogen deprivation followed by genomic profiling can be used to identify druggable alterations potentially causal to intrinsic endocrine therapy resistance.

Introduction

Estrogen receptor positive (ER+) breast cancer is the most common clinical subtype of breast cancer, comprising approximately 80% of patients (1). Adjuvant endocrine therapies, such as selective estrogen receptor modulators (SERMs, *i.e.*, tamoxifen), ER downregulators (SERDs, *i.e.*, fulvestant), and aromatase inhibitors (AIs, *i.e.*, letrozole) are approved for adjuvant treatment of women with ER+ breast cancer. Randomized clinical trials have proven the effectiveness of these anti-estrogens in preventing disease recurrence (2). However, approximately 20% of patients diagnosed with operable ER+ tumors will recur during or following adjuvant endocrine therapy. Of note, mortality from these endocrine resistant tumors accounts for the majority of breast cancer deaths in the United States each year (3). An increasing number of mechanisms of endocrine resistance have been reported (4, 5). To date, the only mechanisms of endocrine therapy resistance that have been observed in the clinic are *ERBB2* amplification (6) and mutations in the ligand binding domain (LBD) of *ESR1* (7–10).

Breast cancer cell proliferation measured by Ki67 immunohistochemistry (IHC) after short-term anti-estrogen therapy was first shown to correlate with recurrence-free survival (RFS) in the Immediate Preoperative Anastrozole, Tamoxifen, or Combined with Tamoxifen (IMPACT) neoadjuvant trial (11). In this trial, tertiles of the post-treatment 2-week Ki67 labeling index showed a strong inverse association with RFS and identified a group of patients with a high 5-year RFS rate of ~40%. Further in this trial, the aromatase inhibitor,

anastrozole, induced a stronger suppression of Ki67 expression compared to tamoxifen or the combination of these drugs. This difference translated into improved long-term outcome in patients treated with aromatase inhibitors in the large adjuvant ATAC (Arimidex, Tamoxifen Alone or in Combination) and BIG (Breast International Group) 1-98 trials (12, 13). These data suggest that tumors exhibiting profound inhibition of cellular proliferation by Ki67 are highly hormone dependent and may identify patients with an excellent prognosis after adjuvant endocrine therapy alone. Conversely, high tumor cell proliferation upon short-term estrogen deprivation can serve as a biomarker to identify patients with antiestrogen-resistant cancers potentially destined to recur early. Thus, we hypothesized that profiling operable ER+ tumors after short-term estrogen suppression with an aromatase inhibitor would identify actionable molecular alterations associated with endocrine resistance. These alterations may ultimately serve as therapeutic targets to combat resistance to anti-estrogen therapy in ER+ breast cancer.

Results

A subset of ER+ breast cancers remains highly proliferative despite letrozole-induced estrogen deprivation

One hundred fifty-five tumor biopsies were obtained from a population-representative set of 143 patients with stage I-III operable ER+/HER2- breast cancer enrolled in a clinical trial of the aromatase inhibitor letrozole administered for 10-21 days prior to surgery (Vanderbilt University NCT00651976). Patients provided written informed consent according to a protocol approved by the Vanderbilt-Ingram Cancer Center Institutional Review Board. Intra-operative biopsies or surgical specimens, snap-frozen in liquid nitrogen or formalin-fixed paraffin-embedded (FFPE), were obtained from each patient's tumor(s). A diagnostic (pre-treatment) FFPE tumor biopsy was obtained for assessment of baseline Ki67, ER, PR and HER2 status at diagnosis were confirmed by IHC or FISH as per clinical guidelines (Fig. 1A). Mean patient age was 64 (range 45-87) with tumors distributed among stage I (54%), II (38%) and III (7.25%). A majority of cancers were of low (35%) and intermediate histological grade (54%); 92% had an ER Allred score of 5 (67% ER+ cells) and 46.5% had a PR Allred score of 5 (67% PR+ cells). Detailed clinical characteristics are shown in Table 1. A detailed illustration of the number of patients enrolled in the trial, the number of evaluable tissue samples, and the number available for molecular analysis can be found in Suppl. Fig. S1.

We measured ER, PR, and Ki67 expression in pre- and post-letrozole tumor samples by automated quantitative immunofluorescence (AQUA) (14, 15). This method utilizes multiplexed immunofluorescence to delineate invasive cancer cells with tagged antibodies against pan-cytokeratin and nuclei with DAPI staining in 10-25 regions of interest on whole tumor sections identified by a breast pathologist, scoring approximately 10,000 tumor nuclei (7). As all tumors were clinically HER2 negative (not amplified), this allowed for allocation of molecular subtype (luminal A or B) by the AQUA-IHC4 algorithm (14) (Fig. 1B).

Tumor response to letrozole was categorized by tertiles of the calculated Ki67 labeling index in the post-treatment (mastectomy) sample according to those used in the IMPACT study (11). The tertiles were: sensitive [Ki67 ln (natural log) $\leq$ 1.0; 0-2.7% Ki67+ tumor cells or

labeling index], intermediate (Ki67 $\ln=1.1-1.9$; 2.8-7.3%), or resistant (Ki67 $\ln \geq 2.0$; 7.4%). Ki67 was evaluable in 140 FFPE tumors. A representative tumor in which the Ki67 was reduced from $\ln \geq 2.0$ (37%) to $\ln \leq 1.0$ (0%) is shown in Suppl. Fig. S2A. Twenty-one percent (21%) of tumors were resistant and retained high Ki67 despite estrogen deprivation (mean Ki67 index 24.5%), 56% were categorized as sensitive (mean Ki67 index 2.7%) and considered to have achieved complete cell cycle arrest (CCCA), and the remaining tumors were deemed intermediate (mean Ki67 index 4.7%) (Fig. 1C, Table 2). Several tumors with high proliferation at baseline ($\ln \geq 2.0$ or Ki67 index $\geq 7.4\%$) exhibited dramatic reductions in Ki67 upon estrogen deprivation with letrozole. The mean number of days sensitive, intermediate, and resistant patients underwent letrozole therapy was not significantly different: sensitive = 15 ± 4 days; intermediate = 14 ± 4 days; resistant = 15 ± 4 days on letrozole.

Consistent with inhibition of ER signaling by letrozole, there was a significant overall reduction in post-treatment levels of progesterone receptor (PR), a well-established target of ER α -mediated transcription (16), compared to pre-treatment levels as measured by AQUA ($p < 0.0001$, paired t-test; Suppl. Fig. S2B). Endocrine resistance, defined by the post-treatment Ki67 tertiles, was independent of the percentage of ER+ cells as scored by a breast pathologist (M.G.K.) (Suppl. Fig. S2C). With AQUA-based scoring, however, ER α -positivity was significantly lower in the resistant tumors (Suppl. Fig. S2D). Tumors with available baseline biopsies were categorized as luminal A or B based on pretreatment AQUA-IHC4 (Fig. 1B); 70 of 134 (52%) were luminal A and 64 of 134 (48%) were luminal B. Of note, 23 of 64 (36%) luminal B tumors exhibited CCCA upon treatment, suggesting they were highly hormone dependent despite unfavorable characteristics by AQUA-IHC4 profiling.

Whole exome sequencing identifies copy number alterations associated with resistance to estrogen deprivation

To uncover single nucleotide variations (SNVs), insertions/deletions (indels) and copy number alterations, we performed successful whole exome sequencing (WES) on DNA extracted from 54 post-treatment tumors and matched normal/blood and 5 unpaired tumor samples (see Suppl. Fig. S1). Somatic mutations were detected in 54 tumor-normal pairs and copy number calls were made in 59 tumors. WES mean target coverage was 90 $\times$. Figure 2 shows an overview of the recurrent somatic mutations (Fig. 2A) and copy number changes (Fig. 2B) detected by CoMut and GISTIC analysis, respectively. A summary of mutations per megabase (mean mutation rate = 1.503×10^{-6} ; standard deviation, SD = 1.400×10^{-6}) showed one outlier patient with a mutation rate higher than two standard deviations from the mean (mean + 2SD = 4.303×10^{-6}). This tumor genome contained a loss-of-function mutation in *RAD52*, which is required for the repair of double-strand DNA breaks (17). Consistent with other breast cancer genomic studies, *PIK3CA* mutations were the most common, with 27 mutations present in 24/54 tumors (44%, n=19 exon 20, n=4 exon 9, n=4 other). *TP53*, *CDH1* (E-cadherin), and *TBX3* mutations were also common, present in 20% or more of samples. Similarly, we found alterations in other genes known to be recurrent in ER+ breast cancer, such as *GATA3* (5/54, 9%) and *SF3B1* (6/54, 11%) (18) and frequent copy number alterations (CNAs/CNVs) at the 1q32 (19 of 59, 32%), 8p11-12 (11 of 59, 19%), and 11q13

amplicons (9 of 59, 15%) (19, 20) (Fig. 2B, Suppl. Fig. S3). Other notable alterations detected were one *ESR1* amplification and one novel ER α mutation (L429V), present in the ligand binding domain (LBD), and four *AKT1* missense mutations (4/54, 7%). Overall, the genomic profile of the letrozole-treated tumors in our cohort was similar to the genomics of treatment-naïve breast tumors in TCGA (1).

Suppl. Table S1 includes an analysis of all recurrent SNVs and CNVs with respect to categorical response to letrozole across all tumors. None of the recurrent somatic mutations (>5%), including *PIK3CA* and *TP53*, correlated with response or resistance to letrozole. Recurrent amplifications with significant correlation to response occurred included 17q21-23, 11q13.3, and 8p11.23. The 8p11-12 and 11q13 amplicons have been previously associated with endocrine resistance (19) and include the histone methyltransferase *WHSC1L1*, the receptor tyrosine kinase (RTK) *FGFR1*, the cell cycle regulator *CCND1* (encodes cyclin D1), and the FGFR ligands *FGF3*, 4, and 19, among others (Suppl. Fig. S3). The 17q21-23 amplicon has not been previously associated with endocrine therapy resistance; it includes *S6K* and *BRCA1*. Neither of these two genes was amplified in letrozole-resistant tumors.

To confirm the association of 8p11 and 11q13 amplification with lack of response to estrogen deprivation, focusing on actionable somatic alterations in such loci, we performed fluorescence *in situ* hybridization (FISH) using clinically-validated probes for *FGFR1* and *CCND1*, respectively. FFPE tumors from the initial WES study and additional tumors from the clinical trial were interrogated (Fig. 3A). *FGFR1* was amplified in 12/72 (17%) *FGFR1*-evaluable tumors and *CCND1* was amplified in 18/68 (26%) *CCND1*-evaluable tumors (Suppl. Figs. S4A & S4B, respectively). *FGFR1* and *CCND1* amplification were both evaluable in a total of 67 tumors (Fig. 3B, Table 3). We observed a statistically significant correlation between *CCND1* amplification by FISH, copy number by WES (Suppl. Fig. S4C), and transcript expression by RNA sequencing (Suppl. Fig. S4E, RNAseq is described in further detail below and in Materials & Methods). *FGFR1* amplification by FISH statistically correlated with WES-derived log₂ copy number values (Suppl. Fig. S4D) but did not correlate with transcript expression (Suppl. Fig. S4F). Next, we compared FISH and tumor response data and found that 7 out of 20 (35%) endocrine therapy resistant cancers were *FGFR1* amplified, whereas only 3 out of 35 (9%) of sensitive tumors harbored *FGFR1* amplification (Bonferroni adjusted $p = 0.05$, Fisher's exact test). Nine out of 20 (45%) letrozole-resistant tumors were *CCND1* amplified, while only 7 out of the 35 sensitive tumors harbored *CCND1* amplification (20%) (Bonferroni adjusted $p = 0.13$, Fisher's exact test). Interestingly, seven out of 67 tumors (10%) harbored co-amplification of *FGFR1* and *CCND1*; 6 of those 7 (86%) were endocrine therapy resistant (Fig. 3B, Table 3). Overall, a high proportion (40%) of ER+ tumors exhibiting *FGFR1* and/or *CCND1* amplification maintained elevated proliferation despite estrogen deprivation with letrozole (Pearson $\chi^2 = 12.67$, Bonferroni adjusted $p = 0.06$). In addition, we observed a trend toward higher baseline Ki67 in *FGFR1*/*CCND1*-co-amplified tumors compared to non-amplified or singly-amplified tumors (Suppl. Fig. S4G). A larger sample size with adequate power would be needed to confirm this association.

Gene expression analysis reveals that E2F, cell cycle and T-cell activation correlate with resistance to estrogen deprivation

To identify transcriptomic alterations, we performed RNA sequencing (RNAseq) on 56 post-treatment tumors (47 fresh frozen, 9 FFPE). RNA sequencing libraries were prepared in two batches: the first containing 41 fresh frozen tumors with RNA prepared using standard, unstranded protocols following polyA purification. The second batch contained a mix of frozen (6) and FFPE (9) tumors, and were prepared using RNA ACCESS (see Materials & Methods). Each batch was analyzed separately to avoid batch effects, with the first cohort serving as a discovery cohort and the second batch serving as a validation cohort. This approach allowed us to observe that the first two principle components separated tumors by response to estrogen deprivation (Figs. 4A & 4B). We performed differential gene expression analysis to identify profiles associated with endocrine resistance. Previous work in our laboratory showed that E2F activation is associated with acquired estrogen independence, and identified a gene expression signature of E2F activation that correlates with poor tumor response to aromatase inhibitors in patients (21). In the current study, this signature clearly segregated tumors by response to aromatase inhibition with the majority of sensitive tumors exhibiting a low E2F activation gene expression signature (Suppl. Fig. 5A). Because it may be expected an E2F/cell cycle gene expression signature would be upregulated in tumors that have high Ki67, we performed the differential expression analysis again after removal of genes associated with proliferation [list of removed genes can be found in Suppl. Table S2, compiled from (22)]. Upon doing so, endocrine therapy resistant tumors stratified based upon enrichment of genes that harbor SH2 and SH3 adaptor protein domains (Fig. 4C). Next, pathway analyses on differentially expressed genes were identified using a gene-by-gene linear model with post-treatment Ki67 categorization (sensitive, intermediate, resistant). Following correction for multiple hypothesis testing, 2328 genes (false discovery rate, FDR <0.025) were deemed differentially expressed in letrozole-resistant tumors. Specifically, genes involved in the cell cycle and mitosis (Fig. 4D, enrichment score 6.08), as well as genes involved in extracellular matrix (ECM) remodeling and migration (Fig. 4E, enrichment score 7.56), were differentially enriched in resistant versus sensitive tumors. Finally, we observed an enrichment of genes involved in secretome-related signaling in letrozole-resistant tumors (Fig. 4F, enrichment score 7.91). Together, these data support the idea that signaling pathways other than cell proliferation are ‘programmed’ differently between endocrine therapy resistant and sensitive tumors.

Because *FGFR1* and/or *CCND1* amplification correlated with a high Ki67 index in letrozole-treated tumors, we used RNAseq data from TCGA ER+/HER2- breast cohort samples to identify differentially expressed genes in *FGFR1*-amplified and *CCND1*-amplified cancers. Pathway analysis of genes differentially expressed in *CCND1*-amplified breast tumors showed robust upregulation of genes involved in the cell cycle and mitosis (Fig. 4F, enrichment score 31.10), whereas in *FGFR1*-amplified tumors, upregulation of cell cycle and mitotic genes was modest (Fig. 4G, enrichment score 6.10). Additionally, genes involved in T-cell activation and NFκB1 pathway signaling were enriched in *FGFR1*-amplified tumors (Suppl. Figs. S5B & S5C). Tumors with co-amplification of *FGFR1* and *CCND1* showed additional enrichment of cell cycle genes greater than the enrichment

explained by single amplification, consistent with an interaction between FGFR1 and cyclin D1 to drive estrogen independent proliferation (Fig. 4H, enrichment score 8.12).

Gene Set Enrichment Analysis (GSEA) also demonstrated a strong direct association between resistant tumors and gene signatures upregulated in high-grade breast cancers, E2F target genes, and genes involved in cell proliferation (Suppl. Figs. S5D, S5E, S5F, respectively). Interestingly, genes that were previously identified as part of the estradiol response (23) were also elevated, suggesting that the ER was still transcriptionally active despite letrozole-induced estrogen deprivation in these tumors (Suppl. Fig. S5G). Furthermore, and consistent with the differential gene expression analyses above, genes involved in the early activation of T-cells were also upregulated (Suppl. Fig. S5H). Finally, transcription factor and target analysis by iRegulon (24) of the differentially expressed genes showed strong enrichment for genes co-regulated by the transcription factors E2F4, FOXM1, and CBFB, all of which were upregulated in the resistant tumors (Suppl. Fig. S5I). Cumulatively, these data suggest that canonical cell cycle signaling pathways are still active in estrogen-deprived tumors, consistent with their high Ki67 scores despite letrozole therapy.

RNAseq Identifies Fusions Associated with Lack of Response to Letrozole

From the RNAseq data, we sought to identify fusion transcripts that may correlate with response to anti-estrogen therapy (see **Methods**). Fusion transcript identification from frozen samples resulted in 346 putative gene fusions from 50 tumors (~7 fusions per tumor); Suppl. Table S3). Primers with universal sequencing tags were designed against 3' and 5' sites of each fusion with a breakpoint mapping to RefSeq exon coding regions (n=187) (Suppl. Table S4). Quantitative PCR was used to amplify breakpoint flanking sequences from tumor cDNA and positive qPCR products were visualized by gel electrophoresis (not shown). Sanger sequencing results from these products were mapped to the human RNA reference transcriptome using the National Center for Biotechnology Information's (NCBI) Basic Local Alignment Search Tool (BLAST) (Suppl. Table S5). Overall, 26 putative fusion transcripts, from 15 unique tumors, were validated by mapping to open reading frames of the predicted 3' and 5' genes (Suppl. Table S5, Suppl. Fig. S6A). Validated transcripts were more likely to be called by more than one program [13 of 46 duplicate programmatic calls (28%) vs. 13 of 300 unique programmatic calls (5%), $p < 0.0001$ Fisher's exact test] (Suppl. Table S3).

Four of these 26 fusions mapped to chromosome 6q25.1, each involving the 5' end of *ESR1* (Fig. 5A). All four *ESR1* fusions were present in tumors that did not respond to estrogen deprivation (Suppl. Table S5). We also observed a statistically significant increase in *ESR1* coding transcript levels in tumors harboring an ER α fusion over those that did not (Suppl. Fig. S6B). Although these genes are in the same chromosome, the fusions are not the result of run-on transcription, as the 5' end of *ESR1* is fused to the 3' ends of *CCDC170*, *AKAP12*, or *c6orf211*, which are upstream of *ESR1* (Fig. 5A). All of these rearrangements fuse the two non-coding 5' exons of *ESR1* to the C-terminal portions of *CCDC170*, *AKAP12*, or *c6orf211*. The *ESR1-CCDC170* fusion variants are thought to result in the expression of a truncated form of CCDC170 which was previously shown to reduce sensitivity to tamoxifen (25).

Consistent with other recent reports (25, 26), variants of the *ESR1-CCDC170* fusion were also identified in MCF7 and other ER+ breast cancer cell lines by RNA sequencing (Fig. 5A). Interestingly, we observed putative *ESR1* fusion transcripts in 12 breast cancers from TCGA (Suppl. Table S6). To validate the *ESR1* fusions identified by RNAseq in our trial set and to uncover additional *ESR1* fusions, we designed NanoString® nCounter probes against each exon of *ESR1*, *CCDC170*, *c6orf211*, and *AKAP12* on chromosome 6q25.1. For this analysis, we included 40 samples from our RNAseq discovery cohort and an additional 23 tumors from this trial (total 63 tumor samples) (Fig. 5B). We also screened several ER+ breast cancer cell lines, and select long-term estrogen deprived (LTED), HER2+, and triple-negative breast cancer (TNBC) cell lines (Fig. 5B). As positive controls, MCF-7 and T47D cells were transduced with the *ESR1-CCDC170* fusion (specifically, the variant fusing *ESR1* exon 2 to *CCDC170* exon 2, E2/C2) (Fig. 5B). Reverse transcript PCR was used to confirm the presence of any putative *ESR1* fusions. The results of the nCounter analysis are summarized in Fig. 5B and the results of nCounter-identified *ESR1* fusion transcript validation are summarized in Suppl. Tables S7 & S8. Of 11 putative *ESR1* fusion transcripts identified in 6 patient samples, four fusions were confirmed by PCR (36%). These fusions were the same ones validated after identification by RNAseq (Suppl. Table S6), and were only validated in endocrine-resistant or –intermediate tumors. Retrospective nCounter analysis revealed that these fusions were detectable above background levels in 4 of 5 of matched pretreatment FFPE tumor biopsies, suggesting that these fusions occur *de novo* (Suppl. Fig. S6C). Of 22 putative *ESR1* fusion transcripts identified by nCounter analysis in 11 cell lines (including LTED, *ESR1* fusion-transduced, and GFP-transduced cell lines), 16 fusion transcripts were verified by fusion-directed PCR (72%) (Suppl. Tables S7 & S8).

***FGFR1* and *CCND1* Amplification-Mediated Resistance to Estrogen Deprivation is Therapeutically Actionable**

The overall purpose of this pre-surgical study was to identify therapeutically actionable somatic alterations associated with resistance to estrogen deprivation in ER+ breast cancers. WES analysis identified amplification of *WHSC1L1*, *FGFR1*, *CCND1* and *FGF3/4/19* as the top amplifications associated with resistance to letrozole (Suppl. Table S1). *WHSC1L1* is a histone methyltransferase that, when amplified, has been proposed as a driver of estrogen-independent ERα activity (27). At this time, however, *WHSC1L1* may not be therapeutically actionable. On the other hand, aberrant signaling triggered by *FGFR1* and its ligands or the interaction between cyclin D1 (*CCND1*) and CDK4/6 can be targeted with clinically-available small molecule inhibitors.

Therefore, to document a causal association of *FGFR1* and *CCND1* amplification/overexpression with therapeutic resistance, we modeled the effects of estrogen deprivation in CAMA1 ER+ breast cancer cells, which harbor co-amplification of *FGFR1* and *CCND1* [cBio Cancer Genomics Portal, Cancer Cell Line Encyclopedia dataset (28, 29)]. Upon stimulation with FGF3/19, CAMA1 cells proliferate in the absence of estradiol (Fig. 6A). These data suggest that activated *FGFR1* and cyclin D1 can mediate persistent growth of ER+ breast cancer cells in the absence of estradiol, a scenario akin to that of the ER+ tumors treated with letrozole in our clinical trial. Phosphorylation of the tumor suppressor Rb by the cyclin D1-CDK4/6 complex uncouples Rb from E2F transcription factors. As a result, E2Fs

induce transcription of genes necessary for the G1-to-S transition. To inhibit FGFR1 and/or cyclin D1 overexpression, we used the FGFR tyrosine kinase inhibitor (TKI) lucitanib (30) and the CDK4/6 inhibitor palbociclib, respectively. Combined inhibition of FGFR1 and CDK4/6 in estrogen-deprived CAMA1 cells with lucitanib and palbociclib potently suppressed estrogen-independent cell growth more effectively than either single agent (Fig. 6A). To further demonstrate the ability of combined blockade of FGFR1 and CDK4/6 activity to overcome resistance to estrogen withdrawal, we calculated the combination index (CI) measuring drug synergy based on the Chou-Talalay method (31) where a $CI < 1$ reflects drug synergism, $CI = 1$ reflects an additive drug effect, and $CI > 1$ represents drug antagonism. The table of CI values (Fig. 6B) for CAMA-1 cells treated with increasing doses of palbociclib $\pm$ lucitanib clearly demonstrated the synergistic effect of this combination on *FGFR1/CCND1* co-amplified cells. Lucitanib and/or palbociclib elicited a similar impact on the percent of CAMA1 cells in S-phase as measured by flow cytometry (Fig. 6C & Suppl. Fig. S7A). Furthermore, only the combination simultaneously reduced FGFR1-mediated ERK1/2 activation and pRB levels in CAMA1 cells, thus restoring Rb activity and cell cycle arrest (Fig. 6D). To complement the studies in CAMA1 cells, we stably transduced ER+ MCF-7 breast cancer cells with FGFR1 or cyclin D1 expression vectors (Suppl. Fig. S7B immunoblot). 2D growth and proliferation assays showed that MCF-7 cells overexpressing cyclin D1 were markedly resistant to estrogen deprivation, while MCF-7 cells transduced with FGFR1 were modestly resistant (Suppl. Figs. S7C & S7D). Collectively, these data support that amplification/overexpression of FGFR1 and cyclin D1 at amplicons 8p11-12 and 11q13 promote resistance to estrogen deprivation.

Profiling of serial breast tumor biopsies identifies 8p11-12 and 11q13 amplifications and ESR1 mutations as markers of endocrine therapy resistance

Finally, we performed targeted next-generation sequencing (FoundationOne™ assay) on tumors from a small cohort of seven patients with poor prognosis ER+ breast cancer. These patients were treated with neoadjuvant chemotherapy followed by adjuvant endocrine therapy, and all recurred with metastatic disease within 5 years (Fig. 7A). For all seven patients, DNA was extracted from a diagnostic tumor biopsy, from the chemotherapy-resistant mastectomy specimen, and from the metastatic biopsy (referred to as pre, post, and met in Fig. 7B).

Consistent with the above WES analysis, amplification of 8p11-12 and/or 11q13 genes was frequent in this cohort, present at diagnosis in 5 of 7 (71%) tumor sets. Among these, 3 of 7 (43%) patients harbored *FGFR1* amplification in all three metachronous biopsies and 2 of 7 (29%) exhibited *CCND1* amplification in all three biopsies. Additionally, amplification of *ZNF217* at 20q13.2, an amplicon known to be associated with poor prognosis, was observed *de novo* in three patients who displayed rapid recurrence on adjuvant therapy. GATA3 truncations were detected in 4 of 7 (57%) tumor sets across all biopsies. In general, most of the identified alterations were present in all three consecutive biopsies. Three tumor sets contained ER α ligand binding domain (LBD) mutations (p.V422del, p.Y537S, p.D538G), present only in the metastatic recurrence. Consistent with these data, the p.Y537S and p.D538G ER α mutants are rare in early breast cancers but detected in about 30% of ER+ metastatic breast cancers that progress on aromatase inhibitors (7-10). Importantly, these

mutants have been shown to be transcriptionally active in the absence of estrogen. To date, ER α p.V422del has only been reported in a single case of endometrial cancer in the Catalogue of Somatic Mutations in Cancer (COSMIC) database (32) and most recently, by Toy, *et al* (33). Secondary analysis of FoundationOne™-profiled breast tumors revealed 10 additional tumors that harbor ER α V422del (Suppl. Table S9). All 10 tumors were noted as metastatic and the 6 tumors with available clinical data were all ER/PR+ and HER2 negative, similar to the tumor with ER α V422del in this study. To better understand the potential effects of the V422del mutant on ER α transcriptional activity, we transfected ER-negative MDA-MB-231 and SKBR3 breast cancer cells with expression constructs encoding HA-tagged wild-type, V422del, or Y537S ER α ; ER α transcriptional activity was then measured with an ER α luciferase reporter. Cells expressing V422del exhibited two-fold higher ER α transcriptional activity in estrogen-free conditions compared to cells transfected with WT ER α (Suppl. Fig. S8A & S8B), as recently reported by Toy, *et al*. (33). Further investigation is necessary to confirm a causal association of ER α V422del with resistance to anti-estrogen therapy.

Discussion

The purpose of this study was to identify actionable alterations potentially causal to *de novo* endocrine resistance in ER+ breast cancers. To stratify patient tumor responses to estrogen deprivation with letrozole, we quantified pre- and post-treatment Ki67 and divided the post-Ki67 into tertiles based on the criteria delineated in the IMPACT trial (11) (Fig. 1C, Table 2). A similar approach has been reported by Ellis *et al.* who labeled aromatase inhibitor sensitive tumors as those with a Ki67 index at or below 10% at surgery following 4 months of neoadjuvant antiestrogen therapy (18). Resistant tumors, with a Ki67 index >10% at surgery, were associated with luminal B status and presence of patterns of somatic mutations mapping to cellular signaling pathways. Conversely, mutations in *GATA3* and *MAP3K1/* *MAP2K4* were associated with significant suppression of proliferation upon treatment with letrozole. Another approach has been to treat patients with post-menopausal ER+ breast cancer for 2-4 weeks with an aromatase inhibitor and rebiopsy the tumor at that time. Patients with tumors with a residual Ki67 $\geq$ 10% were categorized as endocrine resistant and switched to neoadjuvant chemotherapy. The rate of pathological complete responses among these patients was very low (34), suggesting these tumors were also resistant to chemotherapy and underscoring the need to identify molecular drivers of estrogen-independent progression in ER+ tumors.

For patient tumors with a low baseline (pre-treatment) Ki67 index ($\leq$ 2.7% or $\ln \leq$ 1.0), the post-letrozole 2-week Ki67 value may not be sensitive enough to provide a true reflection of drug action and hormone dependence. In the trial reported herein, 6/134 (4.4%) tumors had a baseline Ki67 $\leq$ 2.7% ($\ln \leq$ 1.0). Three of those 6 were PR+ by AQUA and in 2/3, there was a clear reduction in PR upon treatment with letrozole (Suppl. Fig. S2B). In addition, we observed an overall significant decrease in PR levels post-letrozole. These results suggest the presence of active ER α signaling in tumors with a low pre-treatment Ki67 can still be downregulated by estrogen deprivation (Suppl. Fig. S2). These data also suggest that in low-proliferative, ER α -driven tumors, the baseline Ki67 index and/or the difference between baseline and post-treatment Ki67 index are not optimal biomarkers for the detection of a

pharmacodynamic drug effect. Other surrogates of antiestrogen action are needed for these hormone-dependent cancers.

We did not find an association between *PIK3CA* activating mutations and endocrine resistance in this cohort of patients with newly-diagnosed ER+ breast cancer (Fig. 2A, Suppl. Table S1). The clinical activity reported in trials of PI3K inhibitors in combination with antiestrogens in patients with advanced ER+ breast cancer progressing on aromatase inhibitors (35, 36) may suggest these mutations are a marker of acquired, but not *de novo* resistance as has been posited by previous studies [discussed in (37)]. We did, however, identify significant, recurrent copy number amplifications in chromosome amplicons 8p11-12 and 11q13 (Fig. 2B, Suppl. Table S1, Fig. 3, Table 3) (19, 20). We also observed that a large proportion of the resistant tumors harbored co-amplification of *CCND1* (11q13) and *FGFR1* (8p11-12) (Fig. 3B, Table 3). *FGFR1* amplification occurs in ~15% of patients with advanced ER+/HER2- breast cancer in which it is associated with early relapse following adjuvant tamoxifen therapy (38). Simultaneous amplification of *CCND1*, *FGFR1*, and *FGF3/4/19* are reported to occur in 30-40% of breast tumors and are associated with reduced patient survival (19). While *CCND1* and *FGFR1* singly amplified tumors exhibited increased expression of genes involved in cell cycle progression and mitosis, *CCND1*/*FGFR1* co-amplified tumors showed additional enrichment of cell cycle genes above that explained by amplification of either gene alone. This finding suggests a potential interaction between FGFR1 and cyclin D1 to drive estrogen-independent proliferation in co-amplified tumors (Fig. 4, Suppl. Fig. S5). Further strengthening the link between endocrine therapy resistance and *FGFR1*/*CCND1* co-amplification, a previously-published gene signature associated with *RBI* loss (39) (i.e. enhanced proliferation) scored significantly higher in letrozole-resistant tumors and in co-amplified tumors (Fig. 4G & 4H). We also showed that co-targeting of FGFR1 and cyclin D1 signaling with the kinase inhibitors lincitinib and palbociclib, respectively, can synergistically inhibit proliferation of *FGFR1* and *CCND1* co-amplified ER+ breast cancer cells in estrogen-deprived conditions (Fig. 6, Suppl. Fig. S7). Thus, co-amplification of *FGFR1*/*CCND1* in ER+/HER2- tumors may help identify candidate patients for enrollment into investigational adjuvant combination therapy studies designed to trump or significantly delay cancer recurrence.

In line with previously-published reports, we identified a potential association of ER α fusions with endocrine resistant and intermediately-resistant tumors (Fig. 5, Suppl. Table S5) (25). We also observed an increase in total *ESR1* coding transcript levels in samples with ER α fusions versus those without (Suppl. Fig. S6B). Veeraraghavan, *et al.* reported that *ESR1* transcript fusions result in expression of truncated variants of the *ESR1* 5' fusion partner (i.e. *CCDC170*). Overexpression of these *CCDC170* truncated variants in non-tumorigenic MCF10A mammary epithelial cells resulted in increased colony formation, migration, and invasion (25). We posit instead that by relocating the *ESR1* promoter upstream of genes not thought to be relevant toward ER α -dependent signaling, this may release possible negative transcriptional regulation of the *ESR1* coding exons. Testing of this hypothesis requires additional research beyond the scope of this report.

Finally, ER α ligand binding domain mutations have been found predominantly in patients with advanced ER+ breast cancer who have progressed on aromatase inhibitors (7, 9, 10). It

is less clear whether they are associated with resistance to SERDs like fulvestrant or SERMs like tamoxifen (40). We found only one case with an ER α mutation, L429V, in the LBD (data not shown) among newly diagnosed tumors in our trial. Retrospective profiling of metastatic recurrences in patients with ER+/HER2- breast cancer revealed three ER α LBD alterations (Y537S, D538G, V422del), consistent with other reports that ER α mutations arise in late-stage disease (Fig. 7B). The Y537S and D538G mutants are known to induced ligand-independent ER α transcriptional activity (7-10), but less is known about the V422del variant. From FoundationOne™ profiling data, ER α V422del is clearly recurrent in metastatic ER+ disease (Suppl. Table S9). In addition to and consistent with a recent report (33), we showed that V422del activates estrogen-independent ER α transcription approximately two-fold (Suppl. Fig. S8). Further experimental studies are required to fully understand the significance of V422del toward ER α signaling, disease biology, and response to different endocrine therapies.

In sum, we propose that ER+ tumors that remain highly proliferative upon short-term estrogen deprivation harbor molecular lesions causally associated with antiestrogen resistance. Some of these alterations are retained throughout the natural and treated history of these cancers and may mediate early recurrence. As such, this presurgical clinical platform can be used for the unbiased discovery of mechanisms of endocrine resistance that, in turn, can be tested in patients with advanced ER+ cancers of the same genotype. Conversely, tumors that exhibit exquisite inhibition of proliferation under this approach are highly hormone dependent and may be candidates for adjuvant endocrine therapy alone. Finally, the presurgical approach reported herein could be of use in hormone-dependent luminal B tumors that, despite high-risk clinical features at diagnosis, respond well to antiestrogens and may also be candidates for endocrine therapy alone.

Materials and Methods

Clinical trial and tumor biopsies

Tumor samples were obtained from post-menopausal patients with stage I-III operable ER+/HER2- breast cancer enrolled in a clinical trial with the aromatase inhibitor letrozole administered for 10-21 days prior to surgery (Vanderbilt University NCT00651976) (41). Please refer to Supplementary Figure S1 for a detailed illustration of the number of patients enrolled on this trial, the number of evaluable tissue samples, and of those, the number available for the various molecular analyses performed in this study. At the time of this study, 157 patients were enrolled: 14 from Allegheny Cancer Center, 31 from Emory University's Winship Cancer Institute, 5 from Surgical Associates, 98 from Vanderbilt-Ingram Cancer Center, and 9 from Instituto Nacional de Enfermedades Neoplásicas in Lima, Perú. Each patient provided written informed consent for tumor biopsies according to a protocol approved by the Vanderbilt-Ingram Cancer Center Institutional Review Board. Patients were required to fill a pill diary. Intraoperative or surgical specimen biopsies, both fresh-frozen and formalin-fixed paraffin-embedded (FFPE) were obtained from each patient's tumor(s). A diagnostic (pre-treatment) FFPE tumor was obtained for assessment of baseline ER, PR, HER2, and Ki67 by AQUA/IHC4.

Automated quantitative analysis (AQUA) and IHC4 score

Fluorescent multiplexed IHC was conducted at Genoptix, Inc. Medical Laboratories according to clinical standards on both the diagnostic biopsy and post-treatment surgical biopsy for ER, PR, HER2, and Ki67 (antibody clone SP6, Ventana) (14, 15). This method utilizes multiplexed immunofluorescence to delineate invasive cancer cells with tagged pan-cytokeratin and nuclei with DAPI staining in 10-25 regions of interest on whole tumor sections identified by a breast pathologist, measuring approximately 10,000 tumor nuclei (7). AQUA scores were expressed as percent of tumor nuclei expressing Ki67; the Ki67 labeling index was calculated by natural log (ln) transformation of the sum of the AQUA score and a small factor to account for samples with a score of zero [Ki67 Labeling Index = $\ln(\text{AQUAScore} + 0.01)$]. This addition did not change the categorization of any sample. To directly compare IHC and AQUA, tumor sections assayed by both methods using the SP6 antibody were reviewed by experts at Genoptix, Inc. (n=128). A strong correlation in Ki67 expression between IHC and AQUA was observed (Spearman Rank RHO = 0.82; $p < 0.0001$). To divide tumors into luminal A vs. B, luminal A tumors were defined as PR positive (AQUA >36) and Ki67 low (<14%) as described (42).

Fluorescent in situ hybridization (FISH)

For *FGFR1* and *CCND1* FISH analysis, four- μm tissue sections were mounted on charged slides and hybridized overnight with the SPEC *FGFR1/CEN 8* Dual Color Probe (ZytoVision, catalog# ZTV-Z-2072) and *CCND1/CEN11* Dual Color Probe (ZytoVision, catalog# ZTV-Z-2071). Briefly, deparaffinization, protease treatment and washes were performed. After this pretreatment, slides were denatured in the presence of 10 μl of the probe for 6 min at 72°C and hybridized at 37°C overnight in StatSpin (Thermobrite, Abbott Molecular, Inc.). Post-hybridization saline sodium citrate (SSC) washes were performed at 72°C and slides were stained with DAPI before analysis. Endothelial cells, fibroblasts and/or other non-tumor cells in the same section served as internal positive controls. Cases were only further evaluated if control cells displayed one or two clearly distinct hybridization signals of each color. Tumor tissue was scanned using 10 $\times$ magnification (Olympus BX60 Fluorescent microscope). If the *FGFR1* or *CCND1* signals were homogeneously distributed, then random areas were used for counting the signals. Twenty to sixty tumor cell nuclei from random areas were individually evaluated with the 100 $\times$ oil immersion objective by counting green *FGFR1* or *CCND1* and orange centromere 8 (CEN8) or centromere 11 (CEN11) signals. The *FGFR1/CEN8* or *CCND1/CEN11* ratio and the average *FGFR1* or *CCND1* copy number per cell were calculated. Cases were considered as *FGFR1* or *CCND1* amplified under one of the following conditions: a) the *FGFR1/CEN8* or *CCND1/CEN11* ratio is ≥ 2.0 ; b) the average number of *FGFR1* or *CCND1* signals per tumor cell nucleus is ≥ 6 . Pearson correlation analysis was used to determine the relationship of amplification with response; a p value of <0.05 was considered statistically significant.

Whole Exome Sequencing (WES)

DNA extraction—DNA extraction was performed as previously described (43) using frozen or FFPE sections after pathologist review for tumor cellularity. Macrodissection was performed in order to achieve at least 50% tumor cellularity. For normal DNA controls, a

blood sample was obtained during the course of treatment. Peripheral blood mononuclear cells were stored at -80°C until DNA extraction was performed.

Library construction—DNA libraries for massively parallel sequencing were generated as previously described (43) with the following modifications: the initial genomic DNA input into the shearing step was reduced from 3 μg to 10–100 ng in 50 μL of solution. For adapter ligation, Illumina paired-end adapters were replaced with palindromic forked adapters (Integrated DNA Technologies) with unique 8-base index molecular barcode sequences included in the adapter sequence to facilitate downstream pooling. With the exception of the palindromic forked adapters, all reagents used for end repair, A-base addition, adapter ligation, and library enrichment PCR were purchased from KAPA Biosciences in 96-reaction kits. In addition, during the post-enrichment solid phase reversible immobilization (SPRI) bead cleanup, elution volume was reduced to 20 μL to maximize library concentration, and a vortexing step was added to maximize the amount of template eluted from the beads. Libraries with concentrations above 40 ng/ μL , as measured by a PicoGreen assay automated on an Agilent Bravo instrument, were considered acceptable for hybrid selection and sequencing.

Solution-phase hybrid selection—The exon capture procedure was performed as previously described (43) with the following modifications: prior to hybridization, any libraries with concentrations >60 ng/ μL (as determined by PicoGreen) were brought to 60 ng/ μL , and 8.3 μL of library was combined with blocking agent, bait, and hybridization buffer. Libraries with concentrations between 50 and 60 ng/ μL were normalized to 50 ng/ μL , and 10.3 μL of library was combined with blocking agent, bait, and hybridization buffer. Libraries with concentrations between 40 and 50 ng/ μL were normalized to 40 ng/ μL ; 12.3 μL of library were combined with blocking agent, bait, and hybridization buffer. Finally, the hybridization reaction was reduced to 17 hours with no changes to the subsequent capture protocol.

Preparation of libraries for cluster amplification and sequencing—After post-capture enrichment, libraries were quantified using PicoGreen, normalized to an equal concentration using a Perkin Elmer MiniJanus instrument, and pooled by equal volume on the Agilent Bravo platform. Library pools were then quantified using quantitative PCR (KAPA Biosystems) with probes specific to the ends of the adapters; this assay was automated using Agilent's Bravo liquid handling platform. Based on qPCR quantification, libraries were brought to 2 nM and denatured using 0.2 N NaOH on the Perkin-Elmer MiniJanus. After denaturation, libraries were diluted to 20 pM using hybridization buffer purchased from Illumina.

Cluster amplification and sequencing—Cluster amplification of denatured templates was performed according to the manufacturer's protocol (Illumina). HiSeq v3 cluster chemistry and flowcells, as well as Illumina's Multiplexing Sequencing Primer Kit. DNAs were added to flowcells and sequenced using the HiSeq 2000 v3 Sequencing-by-Synthesis method, and then analyzed using Real-Time Analysis (RTA) v.1.12.4.2 or later. Each pool of whole exome libraries was subjected to paired 76-bp runs. An 8-base index sequencing read

was performed to read molecular indices across the number of lanes needed to meet coverage for all libraries in the pool.

Sequence data processing—Exome sequence data processing was performed using established analytical pipelines at the Broad Institute. A BAM file was produced with the Picard pipeline (<http://picard.sourceforge.net/>), which aligns the tumor and normal sequences to the hg19 human genome build using Illumina sequencing reads. The BAM was uploaded into the Firehose pipeline (<http://www.broadinstitute.org/cancer/cga/Firehose>), which manages input and output files to be executed by GenePattern (44).

Sequencing quality control—Quality control modules within Firehose were applied to all sequencing data for comparison of the origin for tumor and normal genotypes and to assess fingerprinting concordance. Cross-contamination of samples was estimated using ContEst (45); 10% of samples were estimated to have contamination between 0.1-0.5% while the rest had no estimated cross-contamination.

Somatic alteration assessment—MuTect (46) was applied to identify somatic single-nucleotide variants. Indelocator (<http://www.broadinstitute.org/cancer/cga/indelocator>) was applied to identify small insertions or deletions. Artifacts introduced by DNA oxidation during sequencing were computationally removed using a filter-based method (47). Annotation of identified variants was done using Oncotator (<http://www.broadinstitute.org/cancer/cga/oncotator>). Copy number ratios were calculated for each captured target by dividing the tumor coverage by the median coverage obtained in a set of reference normal samples. The resulting copy ratios were segmented using the circular binary segmentation algorithm (48). Genes in regions with copy number ratio of greater than 4 were evaluated for focal amplifications, and genes in regions with copy number ratio of less than 0.5 were evaluated for deletions.

Cancer cell fraction analysis—Genome wide copy-ratios were estimated from WES data by comparison of the observed depth of coverage at each exon to that achieved in normal samples. Allelic copy-ratios were then estimated by analysis of allelic fractions for all heterozygous SNPs identified in the paired normal sample. These allelic copy-ratios were then analyzed using ABSOLUTE to generate cancer cell fractions as previously described (49). Pearson correlation analysis was used to determine the relationship of copy number changes with response; a *p* value of <0.05 was considered statistically significant.

Targeted Next Generation Sequencing (FoundationOne™)—Prior to targeted profiling, histopathological diagnoses was independently confirmed on hematoxylin and eosin (H&E)-stained slides. Genomic profiling was performed in a CLIA-certified, CAP-accredited reference laboratory (Foundation Medicine). At least 50 ng of DNA per sample were extracted from 40 μ s of FFPE tumor sections and analyzed by hybridization capture of all coding exons from 236 cancer-related genes and selected introns of 19 genes commonly rearranged in cancer. Specimens were sequenced to high, uniform coverage on Illumina HiSeq instruments, as previously described (50). Genomic base substitutions and indels (short insertions and deletions) were detected using custom tools optimized for mutation calling based on statistical modeling of sequence quality scores and local sequence

assembly. Variations were filtered using dbSNP (<http://www.ncbi.nlm.nih.gov/projects/SNP/>) and a custom artifact database, then annotated for known and likely somatic mutations using COSMIC (32). Copy number alterations were detected by comparing targeted genomic DNA sequence coverage with a process-matched normal control sample. Genomic rearrangements were detected by clustering chimeric reads mapping to targeted introns.

RNA extraction—H&E-stained frozen sections were reviewed by two expert breast pathologists for assessment of adequate tumor cellularity (>20%) with macrodissection if required, for RNA extraction. Fifty-seven fresh-frozen biopsies were deemed adequate for RNA extraction and stored at -80°C. Total RNA extraction was performed using standard TRIzol® techniques (Life Technologies, Grand Island, NY) with Phase Lock gel centrifugation tubes (5Prime, Gaithersburg, MD) to aid in phase separation. Following collection of the aqueous phase, 0.5× isopropanol was added and the resulting slurry was transferred to an RNeasy silica membrane (Qiagen, Valencia, CA) for column purification following the manufacturer's protocol, including the optional on-column DNase digest. All samples were quantified by Nanodrop and quality was evaluated using Agilent's Bioanalyzer 2100. In 10 cases where the frozen biopsy was inadequate for high quality RNA extraction, extraction was attempted from a paired FFPE sample. To control for tissue fixation, 8 tumors were sequenced using tissue from both frozen and FFPE sources. RNA was extracted using the Promega Maxwell® 16 LEV RNA FFPE Purification Kit and instrument, according to the manufacturer's instructions, starting with 4-8 10-µm sections. Total RNA was quantified using the Quant-iT™ RiboGreen® RNA Assay Kit and normalized to 4 ng/µL.

cDNA library construction for RNA sequencing

Frozen: Total RNA was quantified using the Quant-iT™ RiboGreen® RNA Assay Kit (Invitrogen) and normalized to 4 ng/µL; 200 ng of each sample were used for library preparation on an automated variant of the Illumina TruSeq™ RNA Sample Preparation protocol (Revision A, 2010). This method uses oligo(dT) beads to select mRNA from the total RNA sample and is followed by heat fragmentation and cDNA synthesis from the RNA template. The resultant cDNA then goes through library preparation (end repair, base 'A' addition, adapter ligation and enrichment) using Broad Institute-designed indexed adapters for multiplexing. After enrichment, the libraries were quantified with qPCR using the KAPA Library Quantification Kit for Illumina Sequencing Platforms and then pooled equimolarly. The entire process is performed in a 96-well format and all pipetting is done by either Agilent Bravo or PerkinElmer JANUS Mini liquid handlers.

FFPE: Total RNA was quantified using a Qubit (Life Technologies) and quality was assessed using an Agilent Bioanalyzer. 100ng of each sample was used for library preparation following the manufacturer's protocol for RNA ACCESS (Illumina). Briefly, first and second strand cDNA synthesis was performed, universal adapters were ligated, and coding regions were selected by two consecutive hybrid captures followed by PCR enrichment. After enrichment, the libraries were quantified with qPCR using the KAPA Library Quantification Kit for Illumina Sequencing platforms and then pooled equimolarly.

Illumina RNA sequencing—Pooled libraries were normalized to 2 nM and denatured using 0.2 N NaOH prior to sequencing. Flowcell cluster amplification and sequencing were performed according to the manufacturer's protocols using either the HiSeq 2000 v3, HiSeq 2500, HiSeq3000. Each run was a 76-bp paired-end run with an eight-base index barcode. Sequencing was successful in 50 frozen tumors and 19 FFPE tumor sections. Data were analyzed using the Broad Picard Pipeline which includes de-multiplexing and data aggregation. TopHat (v2.0.9) spliced aligner software was used to align reads to hg19, using refseq transcripts as a guide (47). RNAseq GCT files were generated from BAM files using RNA-SeQC (51). Transcripts were assembled and quantified using refseq transcripts as a guide with cufflinks, and normalized FPKMs generated using cuffnorm, following standard protocols.

Fusion transcript discovery pipeline—Fusion transcript analysis was performed using four programmatic algorithms (dRanger, TopHat-fusion, DeFuse, and Chimera Scan), using default parameters. Breakpoint spanning reads and discordant reads were counted for each fusion; a potential fusion list was generated containing fusions with at least 2 discordant/ spanning reads in at least one program. Fusions in genes of interest or detected in two or more programs were prioritized for validation. For fusions selected for validation, breakpoint sequence coordinates were extracted from the output from the respective programs. FASTA sequence corresponding to these coordinates was extracted from hg19 using samtools view. Custom Perl scripts were used to submit breakpoint FASTA sequences to Primer3, with default conditions except a length of at least 20 and a GC-clamp of 2. The resulting primers were checked against the genome for multiple alignments with blat, and primers with more than 2 hits were disqualified. T7 and M13 sequences were added to the 3' end of the primers to facilitate Sanger sequencing.

NanoString nCounter® analysis of putative ESR1 fusion transcripts—

NanoString® analysis was utilized to detect possible ESR1 fusions in RNA samples isolated from patients' tumors and cell lines. The NanoString® probe 'codeset' was first designed using the standard chemistry, which uses probes fluorescently barcoded at the 3' end for identification of the 3' fusion partner. These probes covered the 5' end of each exon in *ESR1* (exons 1-10). The 5' probes were designed against the 3' end of each exon of the upstream genes *AKAP12* (exons 1-5), *c6orf211* (exons 1-5), and *CCDC170* (exons 1-11). Fusions were called by a positive signal from a 5' and 3' pair; using this strategy, only the barcoded 3' probe could be quantitated. Using this strategy, samples with a positive signal at any *ESR1* exon were subsequently screened in the opposite direction (i.e. using reverse chemistry where the 5' exons were barcoded in the pair). Coupling both strategies allowed identification of both the 5' and 3' partner. A total of 144 samples were screened using standard NanoString chemistry, which will detect fusions with resolution at the 3' end. These included RNA samples from 16 cell lines, 48 FFPE tumors sections and 80 frozen tumors. Of these 144, 94 samples were screened in the opposite direction (i.e. for detection of a 5' fusion exon). Raw counts were first subtracted from background and then normalized. Background subtraction was calculated as the number of counts for a series of probes recognizing bacterial genes as negative controls. The mean of these probe counts plus three standard deviations was called as 'background signal'. Any fusion probe pairs giving

less than background signal were zeroed out. Technical normalization was then carried out by dividing all counts within a sample by the geometric mean of a dilution series of spike-in positive controls for the same sample. Finally, sample input was normalized by dividing each count within a sample by the geometric mean of the counts for a series of 5 housekeeper control genes (*GAPDH*, *POLR1B*, *TUBB*, *NPAS1*, *NAGA*). Samples with poor housekeeper control signal (arbitrarily defined as <16 counts of the geometric mean across all 5 housekeeper controls) were removed. Positive calls for a fusion were defined as the any sample with a signal over the mean of the probe pair across all samples, plus two SDs. Positive calls for the 3' chemistry were subsequently selected for 5' complementary analysis.

Cell lines, chemicals and drugs—MCF-7, CAMA-1, ZR75-1, HCC1428, T47D, HCC1500 (ER+/HER2-, cultured in IMEM), BT474 (ER+/HER2+, cultured in IMEM), SKBR3 (ER-/HER2+, cultured in McCoy's 5A), MDA-MB-231 and BT549 (triple negative, cultured in IMEM and RPMI, respectively) human breast cancer cell lines were obtained from ATCC. All lines were cultured in their respective media (Gibco) supplemented with penicillin (100 U/mL)-streptomycin (100 µg/mL) solution (Mediatech) and 10% heat-inactivated fetal bovine serum (FBS, Denville Scientific). For experiments conducted in the absence of estrogen, cells were cultured in phenol-red free IMEM supplemented with the penicillin-streptomycin solution and 10% charcoal/dextran treated FBS (DCC-FBS, HyClone). 293FT and HEKgpIRES cells were cultured in DMEM supplemented with 10% heat-inactivated FBS and the penicillin-streptomycin solution. Long-term estrogen deprived (LTED) MCF-7 and HCC1428 cell lines were generated and maintained as described previously (21). MCF-7 and T47D cell lines were made to stably express V5-tagged eGFP or the *ESR1*(ex1-2):*CCDC170*(ex2-11) fusion via constructs described below. Fulvestrant was provided by AstraZeneca and lucitanib and palbociclib were purchased from Selleck Chemicals and resuspended in DMSO. 17β-Estradiol (E2) was purchased from Sigma and resuspended in 100% ethanol. FGF3 & FGF19 ligands were also purchased from Sigma and resuspended in PBS containing 0.1% human albumin.

cDNA constructs—The coding sequences for *eGFP* and the *ESR1* (ex1-2):*CCDC170* (ex2-11) fusion (i.e. E2/C2 fusion) were developed by Genewiz™ and shuttled into the Gateway® lentiviral expression vector pLX302, resulting in C-terminally V5-tagged constructs. Using the Q5® High-Fidelity DNA Polymerase (New England Biolabs), full-length wildtype (WT) *CCND1* and *FGFR1* were cloned from other plasmids in the laboratory. C-terminal FLAG tags were added to the WT *CCND1* and *FGFR1* sequences by PCR and the constructs were then delivered into the pMXs-puro retroviral vector (Cell Biolabs, Inc.). Direct sequencing of the *eGFP-V5_pLX302*, *E2/C2-V5_pLX302*, *CCND1-FLAG_pMXs-puro* and *FGFR1-FLAG_pMXs-puro* constructs was performed to confirm the sequences and ensure no other mutations were introduced during the cloning process. Primers for *CCND1* and *FGFR1* cloning are listed in Suppl. Table S12. pcDNA3.1+ constructs that express N-terminally HA-tagged ERα wild-type, p.V422del, or p.Y537S were synthesized by inserting the HA-ERα fragment into pcDNA3.1+ using ApaI and BamHI restriction sites. Point mutations were then introduced into pcDNA3.1+-HA-ERα

using the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene) as per the manufacturer's instructions.

Lentiviral and retroviral transductions—eGFP-V5- or E2/C2-V5-encoding pLX302 vectors were transfected along with the pMD2.G VSVG-encoding (vesicular stomatitis virus surface protein envelope plasmid) and psPAX2 lentiviral packaging plasmids into 293FT cells. Media was changed at 24 h. Lentivirus containing media was harvested and filtered at 48 h. Viral supernatant was supplemented with 10 µg/mL polybrene and added to target cells. Puromycin selection (2 µg/mL) began at 48 h following infection for 1-2 weeks, changing media and puromycin every other day. The empty pMXs-puro retroviral plasmid or pMXs-puro vector encoding cyclinD1-FLAG or FGFR1-FLAG were transfected along with pCMV-VSVG into HEKgpIRES cells stably harboring a gag-pol internal ribosome entry site). Retroviral supernatant was harvested filtered and used for infection as above for the lentiviruses.

Immunoblot analysis—Cells were lysed using standard RIPA buffer (50 mM Tris-HCl, pH 7.5; 150 mM NaCl; 1% IGEPAL/NP-40 substitute; 0.1% SDS) and supplemented with protease and phosphatase inhibitors (Roche Complete Mini Protease Inhibitor cocktail tablet, EDTA-free, used as per manufacturer's instructions; 40 mM sodium fluoride; 1 mM sodium orthovanadate; 1 µM okadaic acid). Lysates were quantified by BCA assay (Pierce) and subjected to SDS-PAGE followed by immunoblot analyses using primary antibodies against the following targets: Actin (Cell Signaling, #4970), FGFR1 (Abcam, ab76464), Cyclin D1 (Santa Cruz, sc-718), ERα (Santa Cruz, sc-8002), FLAG (Sigma, F3165), pERK1/2 (T202/Y204, Cell Signaling #9101), pRb (S780, Cell Signaling #9307), and Rb (Santa Cruz, sc-7905). HRP-conjugated anti-rabbit and anti-mouse were used as secondary antibodies (Santa Cruz Biotechnology). Immunoreactive proteins were visualized by enhanced chemiluminescence (Pierce, Rockford, IL, USA).

Cell Proliferation Assays—MCF-7 cells were plated in duplicate at low density (500-1000 cells per well) in 12-well plates in IMEM + 10% DCC-FBS (for estradiol deprivation studies) (see “Cell Lines” above). The following day, baseline cell number per well was assessed using a Coulter Counter (Beckman Coulter). The remaining plates of cells were treated with vehicle or 1 nM 17β-estradiol. Drugs and media were replenished every 72 h for 14 days; cell counts were recorded on days 7 and 14. Each well was counted twice to ensure instrument accuracy and statistical significance was determined with Student's two-tailed *t*-test.

2D Clonogenic/Growth Assays—CAMA-1 cells were seeded in triplicate in 6-well plates (5000 cells/well) in IMEM + 10% DCC-FBS ± 100 ng/mL FGF3/19 ± 1 µM lucitanib ± 500 nM palbociclib. Media, FGF3/19, and drugs were replenished every 3 days until 50-70% confluency was achieved in control wells (~day 12). Cells were then fixed and stained with a solution containing 20% methanol + 80% water + 0.5% crystal violet for 20 min, washed with water, and then dried. After photographs of the wells were obtained, the crystal violet was solubilized in a 20% acetic acid solution and image intensity of the cell monolayers quantified by spectrophotometric detection at 490 nm (GloMax®-Multi

Detection System, Promega). Combination Index (CI) values were calculated as per the Chou-Talalay method (31) based on the average of the fold change in 2D growth relative to controls from three independent experiments. $CI < 1$ represents synergism, $CI = 1$ represents an additive effect, and $CI > 1$ represents antagonism.

Cell Cycle Analysis—After 6 hours of stimulation with 100 ng/mL FGF3/19, CAMA1 cells were harvested, washed in PBS and fixed in cold 70% ethanol. The cells were incubated in 50 µg/mL of propidium iodide and 1 mg/mL of RNase A for 30 min. Cell cycle analyses were performed on a BD flow cytometer and the data were analyzed using Cell Quest software. The experiments were carried out in triplicate and repeated two times.

Gene Expression Analyses—Normalized FPKM expression levels were analyzed in R/Bioconductor. For each gene, surrogate variable analysis (SVA) was performed on the matrix of expression measurements after controlling for the effects of extraction batch and histology. For each gene, we then constructed a linear model $y = batch + histology + KI67 + SV_{i..n} + e$, where y is the $\log_2$ transformed FPKM, *batch* is the extraction batch, *histology* is the diagnosed histology, *KI67* is a factor for response to letrozole (resistant, intermediate, sensitive), $SV_{i..n}$ represents the effects of a matrix of 5 surrogate variables, and e is the residual error. The effects of tissue fixation were analyzed similarly, with the *batch* term representing frozen vs. FFPE. The significance of differential gene expression was assessed by estimating false discovery rates using Storey's q-value method. Pathway enrichment for differentially expressed genes was assessed using DAVID. Additionally, a ranked list of differentially expressed genes was used for Gene Set Enrichment Analysis (GSEA). Principle Component Analysis (PCA) was performed using *pcaMethods*. Heatmaps were generated using *heatmap.2* in the *gplots* package. For the *RB1* loss signature analysis (39), simple gene set variation analysis (GSVA) scores were calculated for each sample in R, using package *GSVA*, method *plage*, using the genes in the *RB1* loss signature.

Statistical Analysis—Appropriate statistical analyses were applied to all comparisons herein under the consultation of biostatistical experts (T.P.S., Y.S., L.D.). Specific tests used are indicated within the manuscript text and/or the figure legend(s) and table(s) they support.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: This study was funded by NIH Breast SPORE grant P50 CA098131, Vanderbilt-Ingram Cancer Center Support grant P30 CA68485, Susan G. Komen for the Cure Foundation grant SAC100013 (CLA), Vanderbilt Institute for Clinical and Translational Research (VICTR) Intramural Clinical and Translational Science Award (CTSA) VR2998, and a grant from the Breast Cancer Research Foundation (CLA). TPS is supported by NIH grant K08 CA148912. JMB is supported by NIH/NCI R00 CA181491-01A1 grant, Department of Defense (DOD) grant BC131494, and Susan G. Komen Career Catalyst Research award CCR 299052. SC is supported by a DOD Breast Cancer Research Program award W81XWH-14-1-0359 and NCI Cancer Center Support Grant P30 CA08748.

References

1. N. Cancer Genome Atlas. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012; 490:490–70. [PubMed: 23099400]
2. Howell A, Cuzick J, Baum M, Buzdar A, Dowsett M, Forbes JF, Hocht-Boes G, Houghton J, Locker GY, Tobias JS. A. T. Group. Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet*. 2005; 365:365–62. [PubMed: 15680439]
3. Cuzick J, Sestak I, Baum M, Buzdar A, Howell A, Dowsett M, Forbes JF. A. L. investigators. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. *The Lancet Oncology*. 2010; 11:11–1141. [PubMed: 20129125]
4. Ma CX, Reinert T, Chmielewska I, Ellis MJ. Mechanisms of aromatase inhibitor resistance. *Nature reviews Cancer*. 2015; 15:15–275.
5. Osborne CK, Schiff R. Mechanisms of endocrine resistance in breast cancer. *Annual review of medicine*. 2011; 62:62–247.
6. Arpino G, Green SJ, Allred DC, Lew D, Martino S, Osborne CK, Elledge RM. HER-2 amplification, HER-1 expression, and tamoxifen response in estrogen receptor-positive metastatic breast cancer: a southwest oncology group study. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2004; 10:10–5676.
7. Jeselsohn R, Yelensky R, Buchwalter G, Frampton G, Meric-Bernstam F, Gonzalez-Angulo AM, Ferrer-Lozano J, Perez-Fidalgo JA, Cristofanilli M, Gomez H, Arteaga CL, Giltane J, Balko JM, Cronin MT, Jarosz M, Sun J, Hawryluk M, Lipson D, Otto G, Ross JS, Dvir A, Soussan-Gutman L, Wolf I, Rubinek T, Gilmore L, Schnitt S, Come SE, Pusztai L, Stephens P, Brown M, Miller VA. Emergence of constitutively active estrogen receptor-alpha mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2014; 20:20–1767.
8. Merenbakh-Lamin K, Ben-Baruch N, Yeheskel A, Dvir A, Soussan-Gutman L, Jeselsohn R, Yelensky R, Brown M, Miller VA, Sarid D, Rizel S, Klein B, Rubinek T, Wolf I. D538G mutation in estrogen receptor-alpha: A novel mechanism for acquired endocrine resistance in breast cancer. *Cancer research*. 2013; 73:73–6864.
9. Robinson DR, Wu YM, Vats P, Su F, Lonigro RJ, Cao X, Kalyana-Sundaram S, Wang R, Ning Y, Hodges L, Gursky A, Siddiqui J, Tomlins SA, Roychowdhury S, Pienta KJ, Kim SY, Roberts JS, Rae JM, Van Poznak CH, Hayes DF, Chugh R, Kunju LP, Talpaz M, Schott AF, Chinnaiyan AM. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nature genetics*. 2013; 45:45–1451.
10. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, Li Z, Gala K, Fanning S, King TA, Hudis C, Chen D, Taran T, Hortobagyi G, Greene G, Berger M, Baselga J, Chandralapaty S. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nature genetics*. 2013; 45:45–1445.
11. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, A'Hern R, Salter J, Detre S, Hills M, Walsh G. I. T. Group. Prognostic value of Ki67 expression after short-term presurgical endocrine therapy for primary breast cancer. *Journal of the National Cancer Institute*. 2007; 99:99–170. [PubMed: 17227988]
12. Baum M, Budzar AU, Cuzick J, Forbes J, Houghton JH, Klijn JG, Sahmoud T. A. T. Group. Anastrozole alone or in combination with tamoxifen versus tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. *Lancet*. 2002; 359:359–2139.
13. Thurlimann B, Keshaviah A, Coates AS, Mouridsen H, Mauriac L, Forbes JF, Paridaens R, Castiglione-Gertsch M, Gelber RD, Rabaglio M, Smith I, Wardley A, Price KN, Goldhirsch A. G. Breast International Group 1-98 Collaborative. A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. *The New England journal of medicine*. 2005; 353:353–2757.
14. Bartlett JM, Christiansen J, Gustavson M, Rimm DL, Piper T, van de Velde CJ, Hasenburger A, Kieback DG, Putter H, Markopoulos CJ, Dirix LY, Seynaeve C, Rea DW. Validation of the IHC4 Breast Cancer Prognostic Algorithm Using Multiple Approaches on the Multinational TEAM

- Clinical Trial. Archives of pathology & laboratory medicine. 2016; 140:140–74. [PubMed: 26910218]
15. Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nature medicine*. 2002; 8:8–1327.
 16. Lin CY, Strom A, Vega VB, Kong SL, Yeo AL, Thomsen JS, Chan WC, Doray B, Bangarusamy DK, Ramasamy A, Vergara LA, Tang S, Chong A, Bajic VB, Miller LD, Gustafsson JA, Liu ET. Discovery of estrogen receptor alpha target genes and response elements in breast tumor cells. *Genome biology*. 2004; 5:R66. [PubMed: 15345050]
 17. Lok BH, Powell SN. Molecular pathways: understanding the role of Rad52 in homologous recombination for therapeutic advancement. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2012; 18:18–6406.
 18. Ellis MJ, Ding L, Shen D, Luo J, Suman VJ, Wallis JW, Van Tine BA, Hoog J, Goiffon RJ, Goldstein TC, Ng S, Lin L, Crowder R, Snider J, Ballman K, Weber J, Chen K, Koboldt DC, Kandoth C, Schierding WS, McMichael JF, Miller CA, Lu C, Harris CC, McLellan MD, Wendl MC, DeSchryver K, Allred DC, Esserman L, Unzeitig G, Margenthaler J, Babiera GV, Marcom PK, Guenther JM, Leitch M, Hunt K, Olson J, Tao Y, Maher CA, Fulton LL, Fulton RS, Harrison M, Oberkfell B, Du F, Demeter R, Vickery TL, Elhammali A, Piwnica-Worms H, McDonald S, Watson M, Dooling DJ, Ota D, Chang LW, Bose R, Ley TJ, Piwnica-Worms D, Stuart JM, Wilson RK, Mardis ER. Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature*. 2012; 486:486–360.
 19. Kwek SS, Roy R, Zhou H, Climent J, Martinez-Climent JA, Fridlyand J, Albertson DG. Co-amplified genes at 8p12 and 11q13 in breast tumors cooperate with two major pathways in oncogenesis. *Oncogene*. 2009; 28:28–1903.
 20. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, Zhang H, McLellan M, Yau C, Kandoth C, Bowlby R, Shen H, Hayat S, Fieldhouse R, Lester SC, Tse GM, Factor RE, Collins LC, Allison KH, Chen YY, Jensen K, Johnson NB, Oesterreich S, Mills GB, Cherniack AD, Robertson G, Benz C, Sander C, Laird PW, Hoadley KA, King TA, T. R. Network. Perou CM. Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell*. 2015; 163:163–519.
 21. Miller TW, Balko JM, Fox EM, Ghazoui Z, Dunbier A, Anderson H, Dowsett M, Jiang A, Smith RA, Maira SM, Manning HC, Gonzalez-Angulo AM, Mills GB, Higham C, Chanthaphychith S, Kuba MG, Miller WR, Shyr Y, Arteaga CL. ERalpha-dependent E2F transcription can mediate resistance to estrogen deprivation in human breast cancer. *Cancer discovery*. 2011; 1:1–351. [PubMed: 22586302]
 22. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, Weinberg RA. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nature genetics*. 2008; 40:40–507.
 23. Dutertre M, Gratadou L, Dardenne E, Germann S, Samaan S, Lidereau R, Driouch K, de la Grange P, Auboeuf D. Estrogen regulation and physiopathologic significance of alternative promoters in breast cancer. *Cancer research*. 2010; 70:70–3770.
 24. Janky R, Verfaillie A, Imrichova H, Van de Sande B, Standaert L, Christiaens V, Hulselmans G, Hertens K, Naval Sanchez M, Potier D, Svetlichnyy D, Kalender Atak Z, Fiers M, Marine JC, Aerts S. iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS computational biology*. 2014; 10:e1003731. [PubMed: 25058159]
 25. Veeraraghavan J, Tan Y, Cao XX, Kim JA, Wang X, Chamness GC, Maiti SN, Cooper LJ, Edwards DP, Contreras A, Hilsenbeck SG, Chang EC, Schiff R, Wang XS. Recurrent ESR1-CCDC170 rearrangements in an aggressive subset of oestrogen receptor-positive breast cancers. *Nature communications*. 2014; 5:4577.
 26. Sakarya O, Breu H, Radovich M, Chen Y, Wang YN, Barbacioru C, Utiramerur S, Whitley PP, Brockman JP, Vatta P, Zhang Z, Popescu L, Muller MW, Kudlingar V, Garg N, Li CY, Kong BS, Bodeau JP, Nutter RC, Gu J, Bramlett KS, Ichikawa JK, Hyland FC, Siddiqui AS. RNA-Seq mapping and detection of gene fusions with a suffix array algorithm. *PLoS computational biology*. 2012; 8:e1002464. [PubMed: 22496636]
 27. Irish JC, Mills JN, Turner-Ivey B, Wilson RC, Guest ST, Rutkovsky A, Dombkowski A, Kappler CS, Hardiman G, Ethier SP. Amplification of WHSC1L1 regulates expression and estrogen-

independent activation of ERalpha in SUM-44 breast cancer cells and is associated with ERalpha over-expression in breast cancer. *Molecular oncology*. 2016

28. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery*. 2012; 2:2–404.
29. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science signaling*. 2013; 6:pl1. [PubMed: 23550210]
30. Bello E, Colella G, Scarlato V, Oliva P, Berndt A, Valbusa G, Serra SC, D'Incalci M, Cavalletti E, Giavazzi R, Damia G, Camboni G. E-3810 is a potent dual inhibitor of VEGFR and FGFR that exerts antitumor activity in multiple preclinical models. *Cancer research*. 2011; 71:71–1405.
31. Chou TC. Drug combination studies and their synergy quantification using the Chou-Talalay method. *Cancer research*. 2010; 70:70–446.
32. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, Ding M, Bamford S, Cole C, Ward S, Kok CY, Jia M, De T, Teague JW, Stratton MR, McDermott U, Campbell PJ. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic acids research*. 2015; 43:D805–811. [PubMed: 25355519]
33. Toy W, Weir H, Razavi P, Lawson M, Goepfert AU, Mazzola AM, Smith A, Wilson J, Morrow C, Wong WL, De Stanchina E, Carlson KE, Martin TS, Uddin S, Li Z, Fanning S, Katzenellenbogen JA, Greene G, Baselga J, Chandralapaty S. Activating ESR1 Mutations Differentially Impact the Efficacy of ER Antagonists. *Cancer discovery*. 2016
34. Ellis MJ, Suman V, McCall L, Luo R, Hoog J, Brink A, Watson M, Ma C, Unzeitig G, Pluard T, Whitworth P, Babiera G, Guenther M, Dayao Z, Leitch M, Ota D, Olson J, Hunt K, Allred C. Z1031B Neoadjuvant Aromatase Inhibitor Trial: A Phase 2 study of Triage to Chemotherapy Based on 2 to 4 week Ki67 level > 10%. *Cancer research*. 2012; 72
35. Baselga J, Im SA, Iwata H, Clemon M, Ito Y, Awada A, Chia S, Jagiello-Gruszfeld A, Pistilli B, Tseng LM, Hurvitz S, Masuda N, Cortes J, De Laurentiis M, Arteaga CL, Jiang Z, Jonat W, Hachemi S, Le Mouhaer S, Di Tomaso E, Urban P, Mssacasi C, Campone M. PIK3CA status in circulating tumor DNA (ctDNA) predicts efficacy of buparlisib (BUP) plus fulvestrant (FULV) in postmenopausal women with endocrine-resistant HR+/HER2- advanced breast cancer (BC): First results from the randomized, phase III BELLE-2 trial. [abstract]. *Proceedings of the Thirty-Eighth Annual CTRC-AACR San Antonio Breast Cancer Symposium*. 2015
36. Mayer IA, Abramson VG, Isakoff SJ, Forero A, Balko JM, Kuba MG, Sanders ME, Yap JT, Van den Abbeele AD, Li Y, Cantley LC, Winer E, Arteaga CL. Stand up to cancer phase Ib study of pan-phosphoinositide-3-kinase inhibitor buparlisib with letrozole in estrogen receptor-positive/human epidermal growth factor receptor 2-negative metastatic breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2014; 32:32–1209.
37. Mayer IA, Arteaga CL. PIK3CA activating mutations: a discordant role in early versus advanced hormone-dependent estrogen receptor-positive breast cancer? *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2014; 32:32–2934.
38. Turner N, Pearson A, Sharpe R, Lambros M, Geyer F, Lopez-Garcia MA, Natrajan R, Marchio C, Iorns E, Mackay A, Gillett C, Grigoriadis A, Tutt A, Reis-Filho JS, Ashworth A. FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. *Cancer research*. 2010; 70:70–2094.
39. Herschkowitz JI, He X, Fan C, Perou CM. The functional loss of the retinoblastoma tumour suppressor is a common event in basal-like and luminal B breast carcinomas. *Breast cancer research : BCR*. 2008; 10:R75. [PubMed: 18782450]
40. Spoerke JM, Gendreau S, Walter K, Qiu J, Wilson TR, Savage H, Aimi J, Derynck MK, Chen M, Chan IT, Amler LC, Hampton GM, Johnston S, Krop I, Schmid P, Lackner MR. Heterogeneity and clinical significance of ESR1 mutations in ER-positive metastatic breast cancer patients receiving fulvestrant. *Nature communications*. 2016; 7:11579.
41. Balko JM, Mayer IA, Sanders ME, Miller TW, Kuba MG, Meszoely IM, Wagle N, Garraway LA, Arteaga CL. Discordant cellular response to presurgical letrozole in bilateral synchronous ER+

breast cancers with a KRAS mutation or FGFR1 gene amplification. *Molecular cancer therapeutics*. 2012; 11:11–2305.

42. Inic Z, Zegarac M, Inic M, Markovic I, Kozomara Z, Djuricic I, Inic I, Pupic G, Jancic S. Difference between Luminal A and Luminal B Subtypes According to Ki-67, Tumor Size, and Progesterone Receptor Negativity Providing Prognostic Information. *Clinical Medicine Insights. Oncology*. 2014; 8:8–111.
43. Fisher S, Barry A, Abreu J, Minie B, Nolan J, Delorey TM, Young G, Fennell TJ, Allen A, Ambrogio L, Berlin AM, Blumenstiel B, Cibulskis K, Friedrich D, Johnson R, Juhn F, Reilly B, Shamas R, Stalker J, Sykes SM, Thompson J, Walsh J, Zimmer A, Zwirko Z, Gabriel S, Nicol R, Nusbaum C. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome biology*. 2011; 12:R1. [PubMed: 21205303]
44. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. *Nature genetics*. 2006; 38:38–501. [PubMed: 16369534]
45. Cibulskis K, McKenna A, Fennell T, Banks E, DePristo M, Getz G. ContEst: estimating cross-contamination of human samples in next-generation sequencing data. *Bioinformatics*. 2011; 27:27–2602.
46. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature biotechnology*. 2013; 31:31–219.
47. Costello M, Pugh TJ, Fennell TJ, Stewart C, Lichtenstein L, Meldrim JC, Fostel JL, Friedrich DC, Perrin D, Dionne D, Kim S, Gabriel SB, Lander ES, Fisher S, Getz G. Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic acids research*. 2013; 41:e67. [PubMed: 23303777]
48. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*. 2004; 5:5–572.
49. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, Laird PW, Onofrio RC, Winckler W, Weir BA, Beroukhir R, Pellman D, Levine DA, Lander ES, Meyerson M, Getz G. Absolute quantification of somatic DNA alterations in human cancer. *Nature biotechnology*. 2012; 30:30–421.
50. Frampton GM, Fichtenholtz A, Otto GA, Wang K, Downing SR, He J, Schnall-Levin M, White J, Sanford EM, An P, Sun J, Juhn F, Brennan K, Iwanik K, Maillet A, Buell J, White E, Zhao M, Balasubramanian S, Terzic S, Richards T, Banning V, Garcia L, Mahoney K, Zwirko Z, Donahue A, Beltran H, Mosquera JM, Rubin MA, Dogan S, Hedvat CV, Berger MF, Puztai L, Lechner M, Boshoff C, Jarosz M, Vietz C, Parker A, Miller VA, Ross JS, Curran J, Cronin MT, Stephens PJ, Lipson D, Yelensky R. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nature biotechnology*. 2013; 31:31–1031.
51. DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C, Reich M, Winckler W, Getz G. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics*. 2012; 28:28–1532.

Fig. 1. A subset of ER+ breast cancers remains highly proliferative despite letrozole-mediated estrogen deprivation

(A) Schema of clinical trial of 143 patients with ER+/HER2- breast cancer treated for 10-21 days with letrozole. Arrows indicate general time points at which a biopsy was taken or surgery was performed. (B) Heat map displaying pre- and post-letrozole treatment immunohistochemistry (by AQUA) scores for Ki67, estrogen receptor (ER), and progesterone receptor (PR) in tumor specimens stratified by Ki67 response to letrozole. Molecular subtype, recurrence score by IHC4, and histologic type (invasive ductal, invasive lobular) are also noted. (C) Paired pre- and post-letrozole treatment tumor specimens from the trial were stratified into sensitive, intermediate or resistant response categories based on post-treatment Ki67 scores. BrCa, breast cancer; pts, patients; ER+, estrogen receptor positive; HER2-, HER2 negative; QD, once daily; hr, hour; FFPE, formalin-fixed, paraffin-embedded; IHC, immunohistochemistry; AQUA-IHC4, Automated Quantitative Analysis-IHC for ER, PR, HER2, Ki67; NST, invasive carcinoma of no special type; ILC, invasive lobular carcinoma.

Fig. 2. Whole exome sequencing identifies copy number alterations associated with endocrine therapy resistance

(A) Tile plot of variants identified in significantly mutated genes detected by whole-exome sequencing (WES) in 55 tumor samples. Samples are listed by response category (13 resistant, 8 intermediate, 30 sensitive, 3 unknown). Genes were considered significantly mutated if their associated q-value was ≤ 0.1 (*i.e.* $-\log_{10}(\text{q-value}) \geq 1.0$), delineated by the solid red line in the histogram at right). (B) Heat map showing $\log_2$ copy number ratios for genomic regions with recurrent gains (red) or losses (green) by GISTIC. Available for CNV analysis were 12 resistant, 8 intermediate, 35 sensitive, and 4 unknown tumors.

Fig. 3. *CCND1* and *FGFR1* amplification by FISH are associated with resistance to estrogen deprivation

(A) Representative FISH images from our cohort displaying samples with *CCND1* amplification (Patient 7629), *FGFR1* amplification (Patient 7670), *CCND1* and *FGFR1* co-amplification (Patient 7657), and a patient negative for *CCND1* and *FGFR1* amplification (Patient 1213). Magnification is 100X for each image. (B) Graphical summary of *CCND1* amplification, *FGFR1* amplification, and *FGFR1-CCND1* co-amplification across sensitive, intermediate, and letrozole-resistant patients as per their post-treatment Ki67 categorization. Numbers for this analysis are shown in Table 3.

Figure 4. Gene expression analysis reveals multiple pathways that strongly correlate with resistance to estrogen deprivation

(A) Principal component analysis of gene expression shows that tumors are separated by response to estrogen in both frozen- (A) and FFPE- (B) derived tumors. (C) Following removal of proliferation-associated genes, differential gene expression analysis using RNA-seq data for response to letrozole shows enrichment for SH2 and SH3 domain containing genes. (D, E) Pathway analyses performed using RNA-seq data from letrozole-treated breast cancers reveals an upregulation of genes involved in the cell cycle, particularly in tumors with *FGFR1* and *CCND1* co-amplification, as well as extracellular matrix proteins. (F) Pathway analysis after removal of proliferation genes shows upregulation of secretome-related signaling. (G, H) Analysis of the *RB1* loss gene expression signature shows increased expression in the resistant tumors (G) and tumors co-amplified for *CCND1* and *FGFR1* (H). Bars, mean $\pm$ SD. *p*-values represent results of a one-way ANOVA followed by Tukey's post hoc test (* $p < 0.05$, *** $p < 0.001$).

Fig. 5. RNA sequencing identifies *ESR1* fusion transcripts associated with resistance to antiestrogen therapy

(A) Diagram of chromosome 6 highlighting the 6q25 locus and showing *ESR1* is recurrently rearranged with other genes in 6q25, including *AKAP12*, *c6orf211*, and *CCDC170*. (B) Heat map of *ESR1* fusion results from NanoString nCounter® analysis, where high counts are represented in pink, low counts are represented in aqua (light blue/green), and absent counts are represented in blue. Left panels represent counts resulting from analysis of probes covering each exon of 5' 6q25 genes; right panels represent counts resulting from probes covering each exon of *ESR1* (3' on 6q25). See Supplemental Tables S6 and S7 for validation results. E2/C2, cells transduced with a construct encoding the *ESR1* ex2-*CCDC170* ex2 fusion; GFP, cells transduced with a construct encoding green fluorescent protein; LTED, long-term estrogen deprivation.

Figure 6. *FGFR1* and *CCND1* Amplification-Mediated Resistance to Estrogen Deprivation is Therapeutically Actionable

(A) *FGFR1* and *CCND1* co-amplified CAMA-1 breast cancer cells were seeded in estrogen-free media supplemented with 100 ng/mL FGF3/19 and treated with increasing doses (0 to 1000 nM) of the CDK4/6 inhibitor, palbociclib, and/or the *FGFR1* inhibitor, lucitanib. When the untreated control cells reached 50-70% confluence (day 12), the cells were fixed and stained with crystal violet and 2D growth was quantified (see Materials and Methods). The combination of palbociclib + lucitanib was sufficient to overcome resistance to estrogen withdrawal better than either single agent. Further highlighting this effect, the table presented in (B) represents the Combination Index (CI) values calculated based on the average of the fold change in 2D growth relative to untreated controls from three independent experiments. CI<1 represents synergism, CI=1 represents an additive effect, and CI>1 represents antagonism. (C) S-phase analysis of CAMA1 cells treated with 100ng/mL FGF3/19, ± 1 μM lucitanib, ± 500 nM palbociclib reveals that treatment with palbociclib or the combination of lucitanib + palbociclib can overcome FGF-induced persistence of CAMA1 cells in estrogen-deprived conditions. Individual plots for each treatment condition can be found in Supplementary Figure S7A. Finally, immunoblot analysis (D) of CAMA1 whole-cell lysates shows that only combined inhibition of *FGFR1* and CDK4/6 can simultaneously decrease ERK1/2 activation and increase Rb activity. Student's t-test, **p<0.05, ***p<0.005. Lucitanib (Luc) = 1 μM and Palbociclib (Palbo) = 500 nM.

Fig. 7. Profiling of serial breast tumor biopsies suggests 8p11-12 and 11q13 amplifications and *ESR1* mutations as markers of endocrine therapy resistance

(A) Schema of the treatment course of seven patients with ER+ breast cancer from which pre-treatment, post-neoadjuvant chemotherapy, and metastatic recurrence biopsies were collected. Dotted lines represent time points at which biopsies were taken. (B) Diagram of the landscape of genomic alterations for the seven patients as per targeted next-generation profiling by FoundationOne™ including time to recurrence and prior treatment. The *ESR1* mutations that were identified are specifically named.

Table 1

Baseline Clinical Characteristics of the 143 Study Patients (No. tumors, 155)

Age (* n = 140 patients)		Years	
Range		45 - 87	
Mean		64	
Tumor Type (n = 155)		No. (%)	
NST (no special type)		84 (54)	
Lobular		20 (13)	
NST w/ special features (lobular, mucinous, tubular, cribriform)		44 (28)	
Special Type		7 (5)	
Tumor Histological Grade (n = 155)		No. (%)	
Low		54 (35)	
Intermediate		83 (54)	
High		17 (11)	
Pathological Stage ** (n = 149)		No. (%)	
IA		74 (50)	
IB		6 (4)	
IIA		39 (26)	
IIB		18 (12)	
IIIA		4 (2.5)	
IIIB		1 (0.75)	
IIIC		6 (4)	
Not available		1 (0.75)	
Tumor Size (n = 154)		No. (%)	
T1a: 5.0 mm		7 (5)	
T1b: 5.1 – 10.0 mm		30 (19)	
T1c: 10.1 – 20.0 mm		70 (45)	
T2: 20.1 – 50.0 mm		42 (27)	
T3: 50.1 mm		5 (3)	
Nodal Status (n = 146)		No. (%)	
Negative		98 (67)	
1 – 3		37 (25)	
4 – 9		5 (3)	
10		6 (4)	
Tumor laterality (n = 143)		No. (%)	
Unilateral		Unifocal	133 (93)
		Multifocal	4 (3)
Bilateral		6 (4)	
Hormone Receptor Positivity		Percent	

Age (* n = 140 patients)	Years	
Estrogen Receptor		
Range	14-100	
Mean	89.5	
Progesterone Receptor		
Range	0-100	
Mean	52.5	
Allred Score (% positive range, n = 155)	ER+ (%)	PR+ (%)
0 (0)	0 (0)	17 (11)
1 (<1)	0 (0)	0 (0)
2 (1-10)	0 (0)	25 (16)
3 (11-33)	5 (3)	17 (11)
4 (34-66)	6 (4)	21 (13.5)
5 (67)	142 (92)	72 (46.5)
Unknown ***	2 (1)	3 (2)

* n represents the number of patients or number of tumors for which data were available

** 143 unique tumors + 6 contralateral tumors

*** positive but unknown proportion

Table 2
Categorization of Study Samples by Post-Letrozole Ki67 Proliferative Index

Category	%Ki67	IMPACT tertile	Samples (%)
Sensitive	2.7	1.0	78 (56%)
Intermediate	2.8 - 7.3	1.1 - 1.9	32 (23%)
Resistant	7.4	2.0	30 (21%)
		Total:	140

Table 3
Summary of *FGFR1* and *CCND1* FISH Analysis in Trial Tumors

	Sens (%)	Int (%)	Res (%)	Total
Neg/Equiv	26 (74)	10 (83)	10 (50)	46
FGFR1 Amp	2 (6)	1 (8.5)		4
CCND1 Amp	6 (17)	1 (8.5)		10
Co-Amp	1 (3)	0 (0)		7
No. Tumors	35	12	20	67
Total Amp	9/35 (26)	2/12 (17)		
Pearson χ^2	$\chi^2 = 12.6664$		p = 0.0487	