



Published in final edited form as:

Nat Cancer. 2021 June ; 2(6): 658–671. doi:10.1038/s43018-021-00215-7.

Serial single-cell genomics reveals convergent subclonal evolution of resistance as early-stage breast cancer patients progress on endocrine plus CDK4/6 therapy

Jason I. Griffiths^{1,2,^}, Jinfeng Chen^{1,^}, Patrick A. Cosgrove¹, Anne O’Dea³, Priyanka Sharma³, Cynthia Ma⁴, Meghna Trivedi⁵, Kevin Kalinsky⁵, Kari B. Wisinski⁶, Ruth O’Regan⁶, Issam Makhoul⁷, Laura M. Spring⁸, Aditya Bardia⁸, Frederick R. Adler^{2,9}, Adam L. Cohen¹⁰, Jeffrey T. Chang¹¹, Qamar J. Khan^{3,*,^}, Andrea H. Bild^{1,*,^}

¹Department of Medical Oncology & Therapeutics Research, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, CA, 91010, USA.

²Department of Mathematics, University of Utah 155 South 1400 East, Salt Lake City, UT, 84112, USA.

³Division of Medical Oncology, University of Kansas Medical Center, Westwood, KS, 66160, USA.

⁴Division of Oncology, Washington University School of Medicine, St. Louis, MO, 63130, USA.

⁵Department of Medicine, Columbia University Irving Medical Center, NY, 10032, USA.

⁶Department of Medicine, University of Wisconsin School of Medicine and Public Health, Carbone Cancer Center, WI, 53726, USA.

⁷Division of Internal Medical Oncology, University of Arkansas for Medical Sciences, AR, 72205, USA.

⁸Department of Medicine, Massachusetts General Hospital Cancer Center and Harvard Medical School, MA, 02114, USA.

⁹School of Biological Sciences, University of Utah 257 South 1400 East, Salt Lake City, UT, 84112, USA.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

*To whom correspondence should be addressed: Andrea Bild (abild@coh.org) and Qamar Khan (qkhan@kumc.edu).

Author Contributions Statement

J.I.G. contributed study design and coordination, evaluated patient response to therapies, analyzed tumor heterogeneity, identified response related phenotypes using scRNA-Seq ssGSEA pathway analysis, quantified resistance phenotypes using mathematical models, linked genetic copy number alterations to phenotypes, reconstructed cancer cell cycle transition and gene expression, and wrote manuscript. J.C. conducted the bioinformatics pipelines to process DNA and scRNAseq data, performed normalization and cell type classification, conducted structural variation analysis of WGS, determined subclonal tumor structure, analysed WES data, and wrote manuscript. P.C. performed scRNA and WGS experiments. A.O.D, P.S, C.M., M.T., K.K, K.B.W, R.O.R, I.M., L.M.S, A.B. contributed patient samples and contributed to writing the manuscript. F.R.A. developed analyses and models and contributed to writing the manuscript. J.T.C. developed bioinformatics pipelines, performed data management and curation, conducted data analysis and wrote the manuscript. A.L.C. contributed data analysis and study design, provided clinical insight, and contributed to writing the manuscript. Q.K., conceived and coordinated the clinical trial, contributed clinical support and infrastructure and provided clinical data and patient samples as well as contributed to writing the manuscript. A.H.B. designed the research project and analyses, performed scRNA experiments and data analysis, coordinated genomic and mathematical/statistical analyses, and wrote the manuscript.

[^]These authors contributed equally

Code availability

Custom code used in analyses are available on GitHub at https://github.com/U54Bioinformatics/FELINE_project.

¹⁰Department of Internal Medicine, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT 84112, USA

¹¹Department of Integrative Biology and Pharmacology, School of Medicine, School of Biomedical Informatics, UT Health Sciences Center at Houston, Houston, TX, 77030, USA.

Abstract

Combining cyclin-dependent kinase (CDK) inhibitors with endocrine therapy improves outcomes for metastatic estrogen receptor positive (ER+) breast cancer patients but its value in earlier stage patients is unclear. We examined evolutionary trajectories of early-stage breast cancer tumors, using single cell RNA sequencing (scRNAseq) of serial biopsies from the FELINE clinical trial (#NCT02712723) of endocrine therapy (letrozole) alone or combined with the CDK inhibitor ribociclib. Despite differences in subclonal diversity evolution across patients and treatments, common resistance phenotypes emerged. Resistant tumors treated with combination therapy showed accelerated loss of estrogen signaling with convergent up-regulation of JNK signaling through growth factor receptors. In contrast, cancer cells maintaining estrogen signaling during mono- or combination therapy showed potentiation of CDK4/6 activation and ERK upregulation through ERBB4 signaling. These results indicate that combination therapy in early-stage ER+ breast cancer leads to emergence of resistance through a shift from estrogen to alternative growth signal-mediated proliferation.

Introduction

Hormone receptor positive (estrogen receptor positive (ER+) and/or progesterone receptor positive (PR+)) breast cancer comprises 70–80% of all breast cancers (1). In ER+ breast cancer, estrogen receptors are activated by estrogen and transduction of this signal to the nucleus promotes cell proliferation and tumor growth. The primary treatment for ER+ breast cancer is endocrine therapy, which either depletes endogenous estrogen and estrogen made by breast cancer cells using aromatase-inhibition (AI) or blocks ER activity through direct modulation or degradation (2, 3). Approximately 90% of all patients with metastatic breast cancer eventually develop resistance to endocrine therapy and at least 33% of patients with early-stage disease will develop endocrine resistance (4, 5). Combination of endocrine therapy with cyclin-dependent kinase (CDK) 4/6 inhibitors has improved disease control in metastatic ER+ breast cancer and adjuvant trials to study the efficacy of this combination in early stage, non-metastatic breast cancer are ongoing (NATALEE) or have completed accrual and are in the follow-up phase (monarchE, PENELOPE-B and PALLAS) (6–9). Preliminary results after short term follow-up from adjuvant trials have shown contradictory results, with the MonarchE trial showing an improvement in invasive disease-free survival with two years of adjuvant abemaciclib added to endocrine therapy, while the PALLAS trial failed to show an improvement in the same endpoint with two years of adjuvant palbociclib. It is not known to what extent such differences are due to biologic effects of the drugs versus biologic differences in different populations of early breast cancer patients; therefore, additional research is needed to characterize the effects of CDK4/6 inhibitor and mechanisms of resistance in surviving cells.

CDK4/6 proteins form a complex with cyclin D that phosphorylates and deactivates the key cell cycle checkpoint regulator retinoblastoma protein (RB1), leading to E2F transcription factor activation, production of cell cycle promoting genes, and progression from G1 to S phase (10). Binding of estrogen to ER and of growth factors binding to growth factor receptors (GFR) drive proliferation through cyclin D/CDK4/6 activation (11, 12). ER can also activate extracellular signal-regulated kinase (ERK) mitogen-activated protein kinase (MAPK) signaling and drive transcription of cyclin D genes and cell cycle progression (13). Some known mechanisms of endocrine therapy resistance include loss or modification of the key estrogen receptor-alpha (ESR1) and activation of the phosphoinositide 3-kinases (PI3K) or epidermal growth factor receptor (EGFR) pathways (14). Prior studies have also revealed mechanisms underlying resistance to CDK4/6 inhibitor treatment in the metastatic setting (15), including disruption or loss of CDK6 and cyclin E2 (CCNE2), as well as activation of AKT1, RAS, ERBB2, and FGFR genes. Resistance mechanisms in early, non-metastatic breast tumors remain unknown.

Phosphorylation of ER serine residue at position 118 is required for its full activity. Ser118 is phosphorylated by MAPK, specifically MAPK1/3 (ERK1/2), initiated by GFR activation (16). Likewise, estrogens also activate ERK1/2 through multiple signaling pathways, highlighting the crosstalk among these pathways (17, 18). Additional MAPK pathways, including MAPK8–10 (Jun amino-terminal kinases; JNK1–3) and MAPK11–14 (p38 α), have also been shown to interact with ER signaling, but their role in response to therapy or evolution to resistance is unknown.

To address the questions detailed above including the impact of therapy on signaling and response in early-stage ER+ breast cancer, we detail how resistance evolves in response to endocrine and cell cycle inhibitor therapies in early stage ER+ breast cancers. These analyses demonstrate multiple convergent phenotypes conveying resistance to combination therapy.

Results

Clinical trial: Patient treatment and sample collection

We studied the evolution of endocrine and CDK inhibitor resistant cancer cell genotypes and phenotypes in post-menopausal women with node positive or >2 cm ER and or PR+, HER2 negative breast cancer enrolled in the FELINE trial ([clinicaltrials.gov # NCT02712723](https://clinicaltrials.gov/ct2/show/study/NCT02712723)) (19). This trial evaluated of the addition of CDK inhibition to endocrine therapy in the neo-adjuvant setting Patients (n=120) were randomized equally into three arms: A) endocrine therapy alone (letrozole plus placebo), B) intermittent high dose combination therapy (letrozole plus ribociclib (600 mg/d, three on/one week off)) or C) continuous lower dose combination therapy (letrozole plus ribociclib (400 mg/d)) (Figure 1a). Patients were treated for six cycles and biopsies were collected at baseline (day 0), following treatment initiation (day 14), and end of treatment (surgery around day 180).

We used tumor growth measurements over time to define resistance and sensitivity during the six months of therapy. Specifically, we mathematically reconstructed tumor size continuously over time during treatment using data from magnetic resonance imaging

(MRI), ultrasound (US), mammogram (MG), clinical physical assessment (CA) taken throughout therapy, and at time of surgical pathology (SP) (20). This estimation was determined using a Gaussian Process Latent Variable Model to account for known biases and differences in accuracy between measurement modalities (21, 22). Groups of similar tumor trajectories were then identified using a Gaussian mixture model. Patients exhibiting either a sustained shrinkage in tumor size during treatment or an initial shrinkage followed by a plateau during treatment were classified as sensitive to therapy. Alternatively, tumors classified as resistant show either 1) no change in size during treatment, 2) a rebound indicated by an initial tumor shrinkage followed by growth during treatment, or 3) continual increased growth (Extended Data Fig. 1a–b). Tumors defined as resistant had a significantly higher proportion of tumor remaining after therapy ($>2/3$ initial size) compared to sensitive tumors ($t=4.45$, $p<0.001$), and have a significant correlation between clinical and modeling classifications (Extended Data Fig. 1a–c). All patient modeling classifications match RECIST 1.1 classification except for two patients. One tumor classified as resistant by the model but PR by RECIST 1.1 shrunk on imaging and clinical exam at day 90 but then both pathology and ultrasound show a rebound by day 180. The second tumor classified as sensitive by the model but marginally SD by RECIST 1.1 exhibited a steady 30% decrease in size at surgery.

Biopsies from the first 60 patients were processed and analyzed (20 patients from each treatment arm). From 60 total patients, 45 had sufficient high-quality tissue in optimal cutting temperature compound (OCT) at day 0 and a follow-up time point. Serial single-cell RNA-sequencing profiles (scRNAseq) and whole exome sequencing (WES) pre- and post-treatment was performed as detailed in the methods (Figure 1a). From the 45 patients sampled, 34 had enough cancer cells and high-quality sequencing data for scRNAseq analysis of the progression of tumor RNA phenotypes and 24 patients yielded WES data (Figure 1e shows patient samples processed on the 10x and Supplementary Table 1 shows patient samples processed using the ICELL8). For samples with sufficient DNA, WES (mean depth 234x) was performed on pre- and day 180 post-treatment tumor biopsies with matched blood sequenced in parallel (mean depth 230x) to identify somatic mutations (Supplementary Table 2).

Genomic analysis of patient tumor samples

We obtained scRNAseq transcriptional profiles for 176,644 cells after filtering out low-quality cells and doublets (Supplementary Table 3). Counts were normalized using a zero-inflated negative binomial model (zinbwave) (23). Cells across patients were integrated using the Seurat normalization package and the reciprocal PCA method (24). Cancer cells were distinguished from normal cells by performing gene copy number analysis of the scRNAseq data using inferCNV for each single cell (25) (Supplementary Information Dataset 1/2). As shown in Figure 1c and Extended Data Fig. 2b, some cells have frequent and pronounced changes in copy numbers and were therefore classified as cancer cells. Using t-SNE, we confirmed that cancer cells with copy number alterations clustered together; and normal cells that show no copy number alterations also cluster with themselves (26). Finally, marker gene expression shown in Figure 1d was used to verify the cell type annotations. A total of 32,781 (18.56%) stromal cells and 16,672 (9.44%) immune cells

were identified, using singleR cell type annotation and verified by cell type specific marker gene expression (Figures 1b and 1c, Extended Data Fig. 2, and Supplementary Table 3) (27). Immune and stromal cells were clearly identifiable by expression of PTPRC and FAP/HTRA1, while cancer and normal epithelial cells expressed KRT19 and/or E-cadherin (CDH1) (Figure 1d).

Tumors undergo subclonal evolution during treatment

To understand how selective pressures of endocrine and CDK4/6 inhibitors drove evolution of cancer cells, WES data was analyzed as detailed in the methods. On average, 99 non-synonymous mutations (range 7–916) and 89 indels (range 24–380) were detected in each sample (Supplementary Table 4). Two patients had substantially higher mutation burden compared to average (Supplementary Table 4) and shared a distinct mutational signature enriched in APOBEC signatures 2 and 13 (Extended Data Fig. 3). Gene mutations known to be frequent in ER+ breast cancer were seen, including PIK3CA (46%), TP53 (29%), and MAP3K1 (21%) (Figure 2a). Gene copy number alterations (CNA) were also frequently identified in the WES data, including gains in AKT3, CCND1, CCNE2, CDK6, FGFR1 and losses in ESR1, RB1 and TP53 (Figure 2b). In general, copy number alterations were more frequent in resistant than sensitive tumors, with most present prior to therapy (Figure 2b, Supplementary Table 5). When summarized at the pathway level, cell cycle, TGF-beta, and TP53 pathway related genes were frequently mutated in this cohort (Extended Data Fig. 4 and Supplementary Table 6). Over time, allele frequency of variants in 19 genes (PTEN, GATA3, and others) increase in 10 patients, suggesting enrichment of clones carrying those mutations (Supplementary Table 7).

Subclonal cancer cell populations were identified using PyClone (28). All patients' tumors show polyclonal populations, with a range of 2–7 subclones present over the course of therapy (Figure 2c). Unlike later stage ER+ breast cancer (29) or triple negative breast cancer (TNBC) (30), few patients show bottleneck events, in which a single dominant subclone emerges during treatment, with the majority of patient tumors maintaining persistent polyclonal populations (Figure 2c) (23).

Evolution of tumor heterogeneity between treatment arms was examined by assessing the frequency of mono- or poly-clonal populations over time (Figure 2d). Diversity includes two components: richness (the number of subclones) and dominance (the uneven fraction of subclones) (Figure 2d). For each biopsy, overall tumor diversity was measured using Shannon's index and changes in dominance by differences in Simpson's index (34). Tumor heterogeneity was found to decrease during endocrine treatment, due to an increase in the dominance of resistant subclones ($t=2.33$, $p<0.05$). In contrast, overall heterogeneity increased during combination treatment due to decreasing dominance ($t=-5.06$, $p<0.001$) (Supplementary Table 8). The increase in tumor genetic diversity under combined therapy suggests that multiple genetic mechanisms of resistance to ribociclib can lead to a resistance phenotype with similar fitness.

Identifying resistant phenotypes during treatment

To determine how cancer cell phenotypes evolved during treatment, we analyzed single cell pathway activity across all three time points (day 0, 14 and 180), using single sample Gene Set Enrichment Analysis (ssGSEA) scores (31). We identified pathways from the Molecular Signatures Database (MSigDB; c2 and hallmark pathways) significantly dysregulated during each treatment (Supplementary Table 9) (32). A hierarchical regression model determined overall and patient specific trajectories of pathway expression within treatment arms, with a Holm's conservative correction for statistical significance (p-values) of pathway activation (33). The most significantly altered pathways (n=87) were categorized by major biological processes, including estrogen receptor activity, signal transduction and proliferation (cell cycle activity) (Supplementary Table 10). The phenotypic heterogeneity in these three processes were assessed in detail through the extraction, dimension reduction, and model-based analysis of the individual genes constituting the detected pathways. Of note, following 6 months of treatment, remaining tumor cells may include those with a resistance phenotype (29, 30).

Accelerated evolution of diminished estrogen signaling

Analysis of the pathway trajectories shows that persistent tumors treated with endocrine therapy alone maintained estrogen signaling following treatment (measured by Hallmark estrogen response early signature), indicating little or no sensitivity to therapy ($t=0.77$, $p=0.45$), while shrinking tumors showed a significant but modest decrease in estrogen signaling (Figure 3a, top left panel) ($t=-2.138$, $p<0.05$). This result indicates that effective response occurs with effective block of estrogen signaling. In contrast, combination therapy caused a stark decrease in estrogen signaling during treatment in persistent tumors ($t=-2.721$, $p=0.05$), demonstrating acquisition of a low estrogen dependent state with diminished endocrine signaling activity (Figure 3a, top middle and right panels). Tumors shrinking during combination therapy show similar diminished levels of endocrine signaling in cancer cells across all treatment arms.

During combination therapy, diminished ER expression and pathway activity was accompanied by a transition from luminal-like to basal-like characteristics (Figure 3a, middle and bottom panels) (transition under high dose combination therapy: $t=2.85$, $p<0.05$; low dose: $t=2.97$, $p<0.05$). The transition occurred repeatedly in different subclones from patient's tumors (Extended Data Fig. 5 and Supplementary Table 11). Cells with increased basal-like phenotype score had lower estrogen receptor (ESR1) expression ($t=-5.77$, $p<0.001$), and this negative association was strengthened under both ribociclib treatment arms ($t=-3.15$, $p<0.005$) (Supplementary Table 12). Estrogen signaling loss also correlated with an increased signature of endocrine therapy resistance (Supplementary Table 12). The switch to an estrogen independent state of proliferation was only weakly observable during endocrine therapy alone, indicating that the selective pressure of combination therapy, especially at high doses, promoted accelerated evolution of the phenotypic switch away from estrogen signaling activity in resistant tumors. Sensitive tumors that shrink during treatment show a more moderate decrease in estrogen signaling.

We also analyzed gene level changes in ESR1 during treatment. We quantified the trajectories of ESR1 expression using a hierarchical generalized additive model while accounting for differences in initial tumor expression. In line with other studies, we found from the WES data that loss of ESR1 heterozygosity (LOH) also significantly diminished expression of ESR1 mRNA (Figure 3b, $W=20$, $p<0.05$), including some pre-treatment samples, indicating innate resistance to endocrine therapy (34, 35). Combination therapy accelerated the reduction of ESR1 mRNA levels over time (ESR1 reduction: $t=-31.54$, $p<0.001$) (Figure 3c), with persistent tumors showing the most dramatic reductions in expression ($t=-80.28$, $p<0.001$). Shrinking tumors also show reductions in receptor mRNA expression, but more comparable to letrozole alone levels ($t=-79.50$, $p<0.001$).

Convergence on JNK MAPK signaling in growing tumors

To identify the proliferative stimuli in cancer cells with diminished estrogen signaling, we analyzed alternative growth pathways that were dysregulated during combination therapy (Supplementary Table 10). In addition to a loss of ER signaling, the pathway analysis detected several signatures of MAPK signaling network alterations. The JNK and ERK pathways are major component of the MAPK network, along with the p38 MAPK pathway; both showed divergent patterns of activity following combination therapy but not endocrine therapy alone (Figure 4a; Extended Data Fig.6; Supplementary Table 10).

The JNK pathway was upregulated during combination therapy in resistant tumor cells at day 14, but not in sensitive cells (Figure 4a) (high dose: $t=3.10$, $p<0.01$, low dose: $t=2.40$, $p<0.05$, resistant versus sensitive $t=2.79$, $p<0.05$). Tumor cells from all patients remaining after six months of combination therapy showed upregulation of JNK expression. This result indicates the acquired resistance in persistent cells, even in tumors that are initially sensitive (as shown in Figure 2c). Significant activation of the JNK pathway score was not seen in resistant tumors treated with endocrine therapy alone (Figure 4a left panel) ($t=0.81$, $p=0.44$), indicating its specific role in CDK inhibitor resistance.

Growth factor signaling can mimic estrogen action and ERK can phosphorylate ESR1, leading to estrogen independent activation (16–18). During combination therapy, resistant tumors became less dependent on ERK signaling (Extended Data Fig.6) ($t=-2.83$, $p<0.05$). In contrast, sensitive tumors maintained higher ERK signaling (no significant ERK change: $t=0.89$, $p=0.41$) perhaps due to the significant crosstalk between ERK and ER, including ER activation of growth factor signaling. Taken together with the activation of JNK signaling under combination therapy, the lack of ERK signal utilization further reflects the transition to an endocrine independent resistance state, with reduced reliance on ESR1/ERK crosstalk (17, 36).

An inverse expression pattern occurred between JNK and ERK pathway genes (MAPK gene set in Supplementary Table 13). Pathway level analysis also shows an inverse relationship (Figure 4b). This dichotomy was consistent in cells across treatments and timepoints (Extended Data Fig.6). An overall JNK activation phenotype score was constructed which integrates expression across MAPK genes, providing a pseudotime metric and placing cells on a continuum from JNK to ERK dominated signal transduction. Utilizing the collinearity of MAPK gene expression, we performed Uniform Manifold Approximation and Projection

(UMAP) dimension reduction of MAPK genes and used the major axis of variation to measure JNK activation (Extended Data Fig.7). High JNK activation was verified to correlate strongly with single cell upregulation of ASK1 and JNK1/3 genes, while low scores reflected ERK activation, including MEK4, MEKK1–3, and ERBB4 upregulation.

To understand how JNK signaling relates to resistance, we examined the link between JNK activation and the expression of the ribociclib target gene, cyclin dependent kinase 6 (CDK6). Across each subclone in all tumors, the average CDK6 expression was calculated along with JNK activation levels (Figure 4c, red line). In addition to expression levels in copy number neutral cells, we also separately analyzed JNK levels in subclones with amplified CDK6 (Figure 4c, blue line; no patients in arm B had CDK6 CNA). Using generalized additive models, we analyzed the nonlinear CDK6:JNK relationship in each treatment and CDK6 amplification group (Figure 4c).

Following combination therapy, JNK activation was concurrent with the upregulation of CDK6 in patients lacking pre-existing CDK6 copy number alterations (Figure 4c) ($F=130.10$, $p<0.001$). Subclonal populations with low JNK activation had low CDK6 expression ($t=12.39$, $p<0.001$), indicating that these cells were not in a proliferative state. In contrast, tumors with CDK6 copy number alterations had high CDK6 expression independent of their JNK activation status ($t=16.03$; $p<0.001$), indicating that genetic alteration removed the requirement for altered signal transduction. Overall, tumors with higher JNK activation decreased in size less during combination therapy ($t=13.84$, $p<0.001$).

Estrogen signaling was explored between subclone populations differing in JNK activation and CDK6 amplification. Subclonal populations with low ESR1 expression exhibited CDK6 upregulation (Figure 4c) ($t=-3.89$, $p<0.001$). In patients lacking CDK6 amplification (Figure 4c, filled circles), estrogen loss was linked to JNK activation ($t=-6.13$, $p<0.001$). Meanwhile, patients with CDK6 amplified tumors lost ESR1 expression independent of JNK activity (Figure 4c, open triangles) ($t=-0.20$, $p=0.84$). In the absence of CDK6 amplification, JNK activation provides an alternative pathway to estrogen independent proliferation under combined therapy. Increased expression of the anti-apoptotic MCL1 gene was also observed in persistent tumor cells after treatment ($t=2.68$, $p<0.05$), in line with studies showing JNK stabilization of MCL1 through phosphorylation (37, 38).

Receptor tyrosine kinase upregulation is common during treatment

As both JNK activation and CDK6 amplification allowed estrogen independence and/or potentiation, we examined GFRs that showed compensatory increases in expression. Receptor genes comprising ssGSEA signatures detected in the pathway analysis were identified along with genes listed in the cell surface protein atlas ($n=1406$) (39). For each patient, we compared pre- and post-treatment expression using analysis of variance and identified receptors with a) increased expression during treatment or b) that had initially higher expression in resistant tumors. Of these genes, we identified receptors with expression inversely correlated to estrogen receptor activity (Extended Data Fig.8; Supplementary Table 14).

Following this method, we found that Erb-B4 Receptor Tyrosine Kinase 4 (ERBB4) gene was significantly upregulated relative to baseline in 50% of tumors (n=13), including all but one patient with CDK6 copy number amplification (n=4/5). ERBB4 is upstream of MAPK signaling and an ER coregulator that can drive proliferation via ERK signaling (40). Tumors lacking CDK6 amplification showed less ERBB4 upregulation than tumors with CDK6 amplification (Figure 5a, top row) ($t=-13.06$, $p<0.001$), indicating that it is a mechanism of resistance to endocrine but not CDK inhibition. Consistent with this hypothesis, higher expression of ERBB4 was also seen in tumors after endocrine monotherapy ($t=11.88$, $p<0.001$).

Fibroblast Growth Factor Receptor 2 (FGFR2) was also upregulated in JNK activated cells following high dose intermittent combination therapy when patients lacked CDK6 amplification (Figure 5a, bottom row). FGF receptors can activate JNK MAPK signaling (41, 42). This relationship was most evident in JNK activated cells following intermittent high dose combination therapy ($t=18.38$, $p<0.001$), with 64% (7/11) of patients showing high expression of FGFR2 (Supplementary Table 14). An additional receptor upregulated over time in resistant cells was RAR Related Orphan Receptor A (RORA), a nuclear receptor that potentially modulates both aromatase enzyme (43) and the ribociclib target, CDK6.

Single cell relationships were constructed through a cluster tree and show that as subclones evolve during treatment, the acquisition of alternative receptors is accompanied by a loss of ESR1 (Figure 5b; Extended Data Fig.9) (Supplementary Information Dataset 2 shows single-cell copy number alterations of each patients' tumor). As an example, subclones with reduced ESR1 and upregulated ERBB4 emerged after endocrine therapy (P20) or combination therapy (P34) (Figure 5b). In addition, subclones with dysregulated ESR1 and upregulated FGFR2 emerged after endocrine therapy (P15) or disappeared after combination therapy (P21) (Figure 5b). As summarized in Figure 5c, in tumor cells with high estrogen signaling, potentiation of CDK4/6 activation can occur through ERBB4 and ERK upregulation and activation. Alternatively, cancer cells with diminished endocrine signaling can bypass CDK4/6 inhibition through upregulation of the JNK pathway. CDK6 amplification itself can potentiate its activity and correlates with cell proliferation. In sum, resistant cancer cell state reflects a phenotypic shift from ERK to JNK MAPK signaling and diminished estrogen signaling in tumors without CDK6 amplification.

Consequences of therapy: cell cycle rewiring to bypass CDK inhibition

Cancer cell proliferation during treatment was examined by measuring changes in cell cycle pathway activity. Cell cycle activity was initially inhibited by both endocrine therapy alone and combination therapy (Figure 6a) (Biocarta cell cycle pathway decline: $t=-2.728$, $p<0.05$). However, by the end of combination treatment, cell cycle activity had rebounded ($t=2.678$, $p<0.05$). Tumor cells that persisted following intermittent high dose combination therapy showed the largest initial reduction in cell cycle activity ($t=-3.290$, $p<0.05$), followed by the greatest proliferative rebound, suggesting stronger selection and cell cycle rewiring to bypass the G1 checkpoint and proliferate independent of estrogen deprivation. In contrast, tumors persistent to letrozole treatment alone exhibited the weaker initial and

subsequent reductions in cell cycle activity compared to shrinking tumors, reflective of innate resistance.

Cell cycle dysregulation is reflected in the changing distribution of cells across cell cycle phases and altered expression of cell cycle regulating genes. To reveal these alterations from single cell data, we extended the Markov model-based reCAT algorithm (44) to reconstruct the cancer cell cycle transitions (Extended Data Fig.10). Cell cycle states were identified by applying UMAP dimension reduction followed by clustering with a Gaussian mixture model. States were connected by finding the shortest path that visited all states. Extending this algorithm, the dynamics of gene expression around distinct cell cycle phases was delineated using cyclical generalized additive models (Figure 6b). Through the reconstruction of the cell cycle from scRNAseq data (Figure 6b), we recovered expected cell cycle stages, including a G1 checkpoint transition where cyclin D initially rises and is followed by CDK6 expression. Further, we observed that Retinoblastoma (RB1), the key G1 checkpoint protein, was downregulated in concert with increased expression of the E2F3 proliferation gene.

The fraction of cancer cells in each patient biopsy that were in the division (S/G2) phase was calculated. During combination therapies, an increasing fraction of each patient's cancer cells were found to be in the division and growth (S/G2) phase ($t=2.94$, $p<0.001$). The frequency of proliferating cells increased in persistent tumors, especially those receiving high dose combination therapy ($t=2.1$, $p<0.05$). In contrast, those receiving endocrine therapy alone exhibited fewer proliferating cells ($t=-2.08$, $p<0.05$), with no detectable difference between tumors that shrank or persisted while on therapy. This result indicates effective bypass of the ribociclib enhanced G1/S checkpoint in surviving subclonal populations.

Next, fluctuations in the expression of ESR1 and JNK genes around the cell cycle were recovered by applying the cyclical generalized additive models. By applying this approach to cells sampled at different timepoints and from patients given different therapies, we distinguished whether treatment altered expression at specific cell cycle stages or if genes expression was dysregulated independent of the cell cycle. ESR1 was expressed at consistent levels throughout the cell cycle (Figure 6d, top row). However, when looking across combination therapy arms, decreasing ESR1 expression was observed over time, accompanied by increasing expression of JNK1 and its target JUNB (Figure 6d, second row) ($t=2.57$, $p<0.05$ and $t= 10.10$, $p<0.001$) (Supplementary Table 15). Further, during combination treatment, we observed a decrease in cyclin dependent kinase inhibitor 2A (CDKN2A (coding for p14 and p16)) ($t=-3.07$, $p<0.005$) and an increase in CDK6 expression in the G1 to S/G2 phases ($t=4.81$, $p<0.001$) (Figure 6d, bottom two rows). Taken together, these observations support the role of estrogen independent JNK signaling in decreasing cell cycle inhibition prior to the G1 checkpoint, thereby permitting cell cycle reactivation.

Discussion

With their proven success in treatment of metastatic ER positive breast cancer (45–47), CDK 4/6 inhibitors are currently being tested in the treatment of early-stage breast cancer (6–9). We studied the evolution to resistant states in patient tumors during treatment with endocrine therapy alone or in combination with the CDK4/6 inhibitor, ribociclib. In patients treated with endocrine therapy alone, we see compensatory signaling between ESR1 and ERBB4 receptor tyrosine kinase, with activation of RTK and downstream ERK MAPK upregulation offsetting decreased ESR1 levels in resistant tumors. We uncovered convergence towards distinct pathways of resistance in patients treated with combined letrozole and ribociclib, including the evolution of estrogen independent proliferation through the upregulation of alternative growth factor receptors (e.g. ERBB4, FGFR2 and RORA) and JNK MAPK signaling including ASK1/MAP3K5 and JNK1–3 (MAPK8–10) (48, 49). 21% of tumors had genetic amplification of the CDK6 gene through copy number gain, directly enhancing its transcription, while those lacking CDK6 amplification initially upregulated JNK signaling during treatment, providing an alternative route to upregulation of CDK6 levels. Due to gene dropouts in scRNAseq data, transcription of all genes was not measured, and additional pathways and genes may also be dysregulated in this setting. It is possible that alternative pathways drive proliferation and response to CDK4/6 inhibitors in early versus late-stage breast cancer (50–53).

JNK has been found to drive aberrant tumor growth in a drosophila model system by modulating cell survival (54, 55), highlighting a potential role for JNK in tumorigenesis and cell-cell interactions. Further, JNK has been shown to activate cell cycle regulated proteins such as CDK4 (56, 57). Regulation of these pathways can drive proliferation of ER+ tumors without a requirement for estrogen. The CDK inhibitor abemaciclib, which recently demonstrated clinical benefit in early-stage patients, is less specific for CDK4/6 than ribociclib, has a different CDK4/CDK6 inhibition ratio, and may include targets such as JNK (58).

Changes in markers of proliferation after 2 weeks of endocrine therapy have been used as early indicators of efficacy (53, 59). Our results demonstrate that information beyond proliferation can be learned from early biopsies in neoadjuvant therapy trials. As a prognostic marker, changes in JNK activation, endocrine signaling and luminal/basal-like features, or cell cycle state could be tested at 2 weeks as a biomarker. Tumors that show an insufficient decrease in estrogen signaling and proliferation after two weeks of endocrine therapy alone may benefit from switching to another endocrine therapy or receiving growth factor receptor signaling inhibitors. Given the potential for additional therapy to increase selective pressure, it will be imperative to rationally determine optimal dosing and timing of drug treatment regimens to reverse or prevent a resistant state while minimizing side effects.

In conclusion, our analysis identifies mechanisms of how tumors circumvent endocrine and CDK4/6 inhibition through phenotypic and subclonal evolution. These mechanisms include shifts to alternative proliferative signaling pathways, bypassing dependence on ER and CDK6 activation. This approach provides a method to detect resistance mechanisms early

in cancer treatment to identify phenotypic targets directed at surviving tumor subclones in tumors.

Methods

Patient cohort and sample collection

Patient tumor core biopsies were collected prospectively under Clinical Trial [#NCT02712723](#), as a multicenter study led by Dr. Qamar Khan at the University of Kansas Medical Center (IND #127673) entitled: Femara (letrozole) plus ribociclib (LEE011) or placebo as neo-adjuvant endocrine therapy for women with ER-positive, HER2-negative early breast cancer (FELINE Trial). FELINE is a randomized, placebo controlled, multicenter investigator-initiated trial. Patients in this trial were enrolled from 10 centers in the United States. Postmenopausal women with pathologically confirmed non-metastatic, operable, invasive breast cancer and clinical tumor size of at least 2 cm were included. Invasive breast cancer had to be ER positive (66% of the cells positive or ER Allred score 6–8) and HER2 negative by ASCO-CAP guidelines. Patients were randomized between three treatment arms (1:1:1). Arm A received letrozole plus placebo, Arm B letrozole plus ribociclib 600 mg daily for 21 out of 28 days of each cycle and Arm C received letrozole plus ribociclib 400 mg continuously. Protocol therapy was continued until the day before surgery. Mammogram, MRI and ultrasound of the affected breast were performed at baseline and a mammogram and ultrasound was performed at completion of neoadjuvant therapy. MRI of the breast was performed after completion of 2 cycles of treatment (Day 1 of cycle 3). Sample collection for tissue was mandatory, providing three tumor core biopsies over the course of treatment: screening (Day 0), Cycle 1 Day 14 (Day 14), and end of trial (Day 180) using a 14-gauge needle. Immediately after collection, biopsy samples were snap frozen embedded in optimal cutting temperature (OCT). Informed consent was obtained from all patients following protocols approved by the institutional IRBs and in accordance with the Declaration of Helsinki. This study was approved by University of Kansas Institutional Review Board (protocol #CLEE011XUS10T). Further information on research design is available in the Nature Research Reporting Summary linked to this article (Supplementary Information Dataset 3 shows study protocols).

Statistics & Reproducibility

Trial sample size was determined based on existing information of the rate of PEPI scores of zero and powered to have a Type I error rate of 10% and Type II error rate of 20%. No samples were excluded due to patient specific factors and both the trial and experiments were blinded and randomized. Statistical tests are two-sided and multiple comparison p-value corrections were applied using Holm's conservative approach.

Exome sequencing, variant calling, and copy number alteration

Whole-exome sequencing was performed for 24 patients with cancer cells that are present at both pre- (Day 0) and post-treatment (Day 14 or Day 180). Genomic DNA of tumor and matched germline samples were captured using the SureSelect Human All Exon v7 (Agilent) or xGen Exome Research Panel v2 (IDT). Enriched DNA samples were sequenced on an Illumina NovaSeq 6000 with 150 bp paired-end reads. Sequence analyses

were performed with a Bioinformatics Expert SYstem (BETSY) (60). Briefly, sequences were processed by trimmomatic v0.39-1 (61) (MINLEN:15 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15) to trim adaptors and low-quality sequences. Trimmed sequences were aligned to the hg19 human genome with BWA-MEM v0.7.17 (62–64) and bam files were sorted by sambamba v0.6.8 (<https://lomeriteir.github.io/sambamba/>). PCR duplicates were marked using Picard v2.18.4 (65). Local realignment around the 1000 Genomes Phase I indels were performed with the Genome Analysis Toolkit (GATK v3.8) (66).

Somatic single-nucleotide variations (SNVs) and small indels (< 50 bp) were identified with MuTect2 (in GATK v3.8) (66), Strelka v2.9.2 (67), and VarScan2 v2.4.3-2 (68) using tumor-normal pairs. Variants with read depth ≥ 25 , alternative allele read depth ≥ 5 , variant allele frequency (VAF) ≥ 0.05 in tumor samples, and VAF $\geq 1\%$ in normal samples were characterized as somatic variants for further analysis. Somatic variants were annotated with ANNOVAR v2018Apr16 (69). Somatic copy number, tumor purity, and ploidy were estimated from WES using FACETS v0.5.14 (70) and Sequenza v2.1.2 (71) (Supplementary Information Dataset 1). Multiple runs with different parameters were performed for each tool. A best model was chosen for each patient through inspection of log-ratio and copy number profile of each patient. Copy number alteration of cancer driver genes were determined based on the median log-ratio of the segment (cnlr.median) and copy number estimated by FACETS. Copy number gain or loss were defined as genes located on a segment with cnlr.median ≥ 0.2 and copy number $>$ ploidy or cnlr.median ≤ -0.2 and copy number $<$ ploidy respectively. A loss of heterozygosity (LOH) was defined as genes located on a segment with minor allele completely lost (minor_cn = 0).

Clonal structure and evolution

Somatic mutations obtained from the above analysis were clustered by PyClone v0.13.1 using the Beta Binomial model (10,000 iterations). Tumor purity and copy number were used by PyClone to estimate cellular prevalence of somatic mutations and mutation clusters. Clonal evolution models were inferred by ClonEvol v0.99.11 (72) based on mutation clusters and cellular prevalence of somatic mutations. The truncal cluster was assigned to the cluster with 80% cellular prevalence in at least one sample. Mutation clusters with five or fewer mutations were discarded unless the cluster represented a truncal cluster. Mutation clusters showing similar changes across samples were merged when ClonEvol failed to predict clonal evolution models.

Evolution of subclonal diversity

Evolution of tumor heterogeneity during treatment was assessed by changes in subclonal diversity in each arm. The relative frequency of cancer subclones ($p_s(i, T)$) was calculated for tumor samples from each patient (i) at the first and last treatment timepoints (T). The overall subclonal heterogeneity of each tumor was measured using the Shannon diversity index. To disentangle the two key components of diversity (richness and dominance), we calculated richness by the number of subclones identified per sample and measured dominance as $1 - \text{Simpson's index} (\sum (p_s(i, T)^2))$.

Changes in tumor diversity (D) over time (T) and between treatment arms (A), were assessed using a random effects model:

$$D_i \sim \text{Normal}((\beta_0 + u_{0i}) + \beta_A A + \beta_T T + \beta_{AT} AT, \sigma_D^2)$$

$$u_{0i} \sim \text{Normal}(0, \sigma_0^2)$$

Pre-treatment differences in tumor diversity between arms were accounted for by the treatment-specific estimates of initial diversity (β_0 vs β_A). Subsequent changes in diversity were described by treatment-specific temporal trend terms (β_T and β_{AT}). Individual variability in pre-treatment tumor heterogeneity was accounted for by allowing the model intercepts to vary among patients (u_{0i}). Likelihood ratio tests were used to assess whether significant changes in tumor heterogeneity occurred during treatment. The likelihood of the above model was compared with that of nested null hypothesis models in which no change in tumor heterogeneity occurred (fixing β_T & $\beta_{AT} = 0$) or equal changes in tumor heterogeneity occurred across treatment groups (fixing $\beta_{AT} = 0$ but estimating β_T). The likelihood function for these models was:

$$L(\hat{D}, \sigma_D^2) = \prod_{i=1}^N \frac{e^{-\frac{(D_i - \hat{D}_i)^2}{2\sigma_D^2}}}{\sqrt{2\pi\sigma_D^2}}$$

where \hat{D}_i is the expected diversity of the linear predictor and N is the sample size (patients number=34).

Single nuclei RNA sequencing and processing

Tumor cell nuclei were isolated from OCT embedded core tumor biopsies using a modified lysis buffer containing 0.2% Igepal CA-630 as previously described (73). Single cell RNA-Sequencing (scRNA-Seq) was performed on single nuclei suspensions using either the Takara Bio ICELL8 Single Cell System (10 patients) or the 10X Genomics Chromium (35 patients), to prepare cDNA sequencing libraries (Supplementary Table 3). Samples processed on the ICELL8 Single Cell System (Takara Bio) were prepared using the SMARTer ICELL8 3' DE Reagent Kit V2 (Takara Bio, Cat#640167) from isolated nuclei. DAPI stained nuclei were diluted to 60,000 cell/mL in 1x PBS + 1% BSA + 1x Second Diluent + 0.2U SUPERase-In RNase Inhibitor and dispensed onto the ICELL8 3' DE Chip (Takara Bio, Cat#640143) using the ICELL8 MultiSample NanoDispenser. Single nuclei candidates were selected using the ICELL8 imaging system with CellSelect Software (Takara Bio) selecting for DAPI positive nuclei and reverse transcription and sequencing library preparation was performed according to manufacturer instructions. ICELL8 cDNA sequencing libraries were sequenced at a depth of 200K reads/cell on Illumina HiSeq 2500 (read #1=26nt and read #2=100nt). Samples processed on the 10X Genomics Chromium were processed using the Chromium Single Cell 3' V3 Kit (10X Genomics, Cat #1000075) using isolated nuclei. Single nuclei were diluted to 1,000 cells/ μ L in 1x PBS + 1.0% BSA

+ 0.2U/ μ L SUPERase-In RNase Inhibitor to generate GEM's prepared at 5,000 cells per sample. Barcoding, reverse transcription, and library preparation were performed according to manufacturer instructions. 10X Genomics generated cDNA libraries were sequenced on Illumina HiSeq 2500 or NovaSeq 6000 instruments using 150 bp paired-end sequencing at a median depth of 34K reads/cell. scRNA-Seq was performed at the Integrative Genomics Core at City of Hope, Fulgent Genetics, and the High Throughput Genomics Core at Huntsman Cancer Institute (HCI) of University of Utah. Sequence reads were processed with BETSY and CellRanger v3.0.2, which aligned reads to reference genome (GRCh38) using STAR v2.6.0 (74). Counts on gene transcripts were calculated by featureCounts using subread v1.5.2. A gene-barcode matrix was generated for each sample containing counts of unique molecular identifiers (UMIs) for each gene in each cell.

Copy number alteration and subclone analysis from scRNA

Copy number alterations of each cell were estimated based on the count matrix using R package 'infercnv' v1.0.2 (cutoff=0, min_cells_per_gene=100 or 500, cluster_by_groups=T, HMM=T, analysis_mode="subclusters"), which uses gene expression intensity to predict copy number changes in tumor cells compared to normal reference cells. A subset of 500 immune or stromal cells from this study were used as reference for 'infercnv' analysis. Cancer cells from each patient were clustered based on the estimated gene copy numbers in each cell using 'hclust' from R package 'fastcluster' v1.1.25 (75) (method='ward.D2'). Clusters with distinct copy number profiles were defined as subclones for each patient. Single-cell grouping was performed based on hierarchical cluster analysis.

Cell type classification

An integrative approach was used to classify cell types in samples from 45 patients. First, cell type of each cell was predicted using the R package 'SingleR' v1.0.1 to generate preliminary cell type classifications for all patients. Second, we applied Seurat v3.1.1.9023 Reciprocal PCA integration workflow to 34 patients with 10x scRNA data to integrate cells from different samples (24). Patients with ICELL8 scRNA data were analyzed with standard Seurat workflow. Cell clusters were identified using 'FindClusters' method (resolution=0.8) in R package Seurat v3.1.1.9023. Third, each cell cluster was defined as epithelial cells, stromal cells (fibroblasts, endothelial cells) or immune cells (macrophages, T cells, B cells) based on the most frequent cell type annotation by 'SingleR'. Expression profiles of marker genes validated the cell type classifications: epithelial cells (KRT19, CDH1), immune cells (PTPRC), stromal cells (HTRA1, FAP). Finally, epithelial cells were classified into cancer cells and normal epithelial cells based on presence and absence of copy number alteration. Cell type annotation was summarized in Supplementary Table 3.

Breast cancer intrinsic subtype prediction

Primary molecular subtypes of breast cancer (basal-like, HER2-enriched, luminal A, luminal B) for each cell was predicted based on the log(CPM+1) count matrix by R package 'genefu' v2.18.0 (74) with default parameters (sbt.model = scmod1). The predicted subtype that has the highest prediction probability was assigned as breast cancer intrinsic subtype of each cell. This analysis was performed only on patients with 10x scRNA data. Patients with ICELL8 scRNA have very few cancer cells, thereby were not included in this analysis.

Gene Set enrichment analysis

The count matrix of each cell type was filtered to keep genes expressed in at least 10 cells. The filtered count matrix was normalized by R package ‘zinbwave’ v 1.8.0 (23) with total number of counts and gene length and GC-content as covariates ($K=2$, $X=\sim\log(\text{total number of counts})$, $V=\sim\text{GC-content} + \log(\text{gene length})$, $\text{epsilon}=1000$, $\text{normalizedValues}=\text{TRUE}$). Single sample Gene Set Enrichment Analysis (ssGSEA) scores of 50 hallmark signatures (MSigDB, hallmark) (32) and 4725 curated pathway signatures (MSigDB, c2) (32) were calculated for each cell based on the normalized count matrix using R package ‘GSVA’ v1.30.0 ($\text{kcdf}=\text{“Gaussian”}$, $\text{method}=\text{“ssgsea”}$) (76).

Pathway analysis: identifying response related phenotypes

For each ssGSEA pathway in the C2-level and Hallmark pathway signatures, changes in cancer cell pathway activity (x) over time (T) was examined during each treatment arm of the trial (A). Phenotypic changes linked to treatment or differing between sensitive and resistant patients (R) were identified. A random effects model (77) with the following linear predictor and error structure was constructed for each pathway:

$$E(ssGSEA[x]_i) \sim (\beta_0 + u_{0i}) + \beta_{RR} + (\beta_T + u_{Ti})T + \beta_{RT}RT$$

$$ssGSEA[x]_i \sim \text{Normal}(E(ssGSEA[x]_i), \sigma_x^2)$$

$$u_{0i} \sim \text{Normal}(0, \sigma_0^2)$$

$$u_{Ti} \sim \text{Normal}(0, \sigma_T^2)$$

Initial differences in pathway activity between cancer cells from sensitive and resistant tumors, at the pre-treatment time point were captured by the group-specific intercepts (β_0 vs β_R). Subsequent changes in pathway activity were described by temporal effect terms of sensitive and resistant tumors (β_T and β_{RT}). Preexisting individual variability in gene expression and patient specific susceptibility of pathway activation to therapy, were accounted for by allowing the model intercept and temporal effect to vary among patients (i) (u_{0i} and u_{Ti}). Significant differences in pathway activity before or during treatment were identified between treatment arms and patient response groups, using likelihood ratio tests with multiple comparison corrections following Holm’s p-value correction. Compared to false discovery rate (FDR) correction, Holm’s adjustment was more conservative, avoiding spurious detection of response related phenotypes. The likelihood of the full model was compared against that of nested null models in which: i) no change in pathway activity occurred in sensitive or resistant tumors (fixing β_T & $\beta_{RT} = 0$) or ii) equal changes in pathway activity occurred in sensitive and resistant tumors (fixing $\beta_{RT} = 0$ but estimating β_T). The likelihood function for each of these models was:

$$L(\widehat{ssGSEA[x]}_i, \sigma_D^2) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(\widehat{ssGSEA[x]}_i - \overline{ssGSEA[x]}_i)^2}{2\sigma_x^2}},$$

where $\overline{ssGSEA[x]}_i$ is the expected pathway activity ($E(ssGSEA[x]_i)$) and N is the number of single cells.

Significant pathway activation in sensitive and resistant tumors (non-zero β_T and β_{RT} parameters) was assessed using a two-tailed t-tests (sample size based on patient number; per arms (A=11, B=11, C=12)). The Satterthwaite method was applied to perform degree of freedom, t-statistic and p-value calculations for the hierarchical regression model coefficients, using the 'lmerTest' R package (78). The hierarchical regression model explicitly described the paired structure in our dataset, with earlier and later samples per patient, and this non-independence of cells within a sample determined the effective residual degrees of freedom. Statistical tests of pathway activation related tumor growth thus controlled for this pairing of samples per patient tumor. Detected ssGSEA signatures were classified into functional categories (Supplementary Table 10). Genes contributing to each pathway category were identified from the MSigDB and used for downstream analyzes of each process.

Assessing the loss of estrogen receptor expression

Changes in single cell ESR1 expression during treatment were assessed using a hierarchical generalized additive model (79). The nonlinear trajectory of ESR1 expression during the trial was described by a thin plate spline. Baseline differences in patients' ESR1 expression was accounted for by a patient specific random intercept term. Single cell ESR1 expression over time was assessed separately for patients in each treatment arm.

MAPK signaling gene network co-regulation

To investigate the co-regulation of signal transduction states of cancer cells, we analyzed the pairwise correlation of MAPK gene expression. Hierarchical clustering of the correlation matrices showed the dichotomy between the expressions of kinases acting in the JNK versus the ERK pathways of signal transduction.

The primary phenotypic variation in MAPK signaling across cancer cells was determined by performing dimension reduction of MAPK gene expression, using UMAP (confirmed by Principle component analysis). To assess phenotypic variation across patients, without large patient samples biasing results, the dataset was initially down sampled to 100 cells per biopsy. The UMAP model was then trained and used to calculate phenotypic scores of the full dataset. This analysis confirmed that JNK activation status was the primary component of heterogeneity in MAPK signaling state across cancer cells (Extended Data Fig.7). A single cell JNK activation score (high JNK, low ERK) was determined, using the major axis of phenotypic variation, due to the collinearity of gene expression between JNK and ERK genes. For each patient, average scRNAseq gene expression was calculated for subclonal cancer cell populations with different levels of JNK activation (n=40 levels).

Relating CDK6 expression to JNK activation

We characterized the subclonal relationship between JNK activation and the expression of the key ribociclib target gene, CDK6, accounting for genetic differences in CDK6 copy number amplification status. Resistance phenotypes were examined specifically, by focusing on the transcription profiles of cells at end of treatment. Using generalized additive models, a nonlinear smooth function was fitted describing how the average CDK6 expression changed as the JNK activation of subclonal populations increased. A separate relationship was identified for each arm of therapy and for patients with and without baseline CDK6 genetic copy number amplification (identified by CNA analysis). The significance of CDK6:JNK relationships in each group were assessed using likelihood ratio tests comparing the full model to a null model without a CDK6:JNK association.

Cell surface receptor compensation for ER loss

The Cell Surface Protein Atlas (39) provides a database of known cell receptor proteins. Genes encoding each protein were identified from the Ensembl database (80) (n=1406). For each patient, we identified genes with: i) significantly higher/lower than average initial expression, or b) altered expression during treatment, using an ANOVA test. Receptor genes identified consistently across patients were identified as those overlapping gene detected by our pathway analysis as well as being detectable in multiple patients using the ANOVA approach. Genes activated to compensate for the loss of estrogen signaling during treatment were determined by correlating estrogen pathway activity with alternative receptor gene expression for cells in each treatment arm and response group.

Cell cycle reconstruction

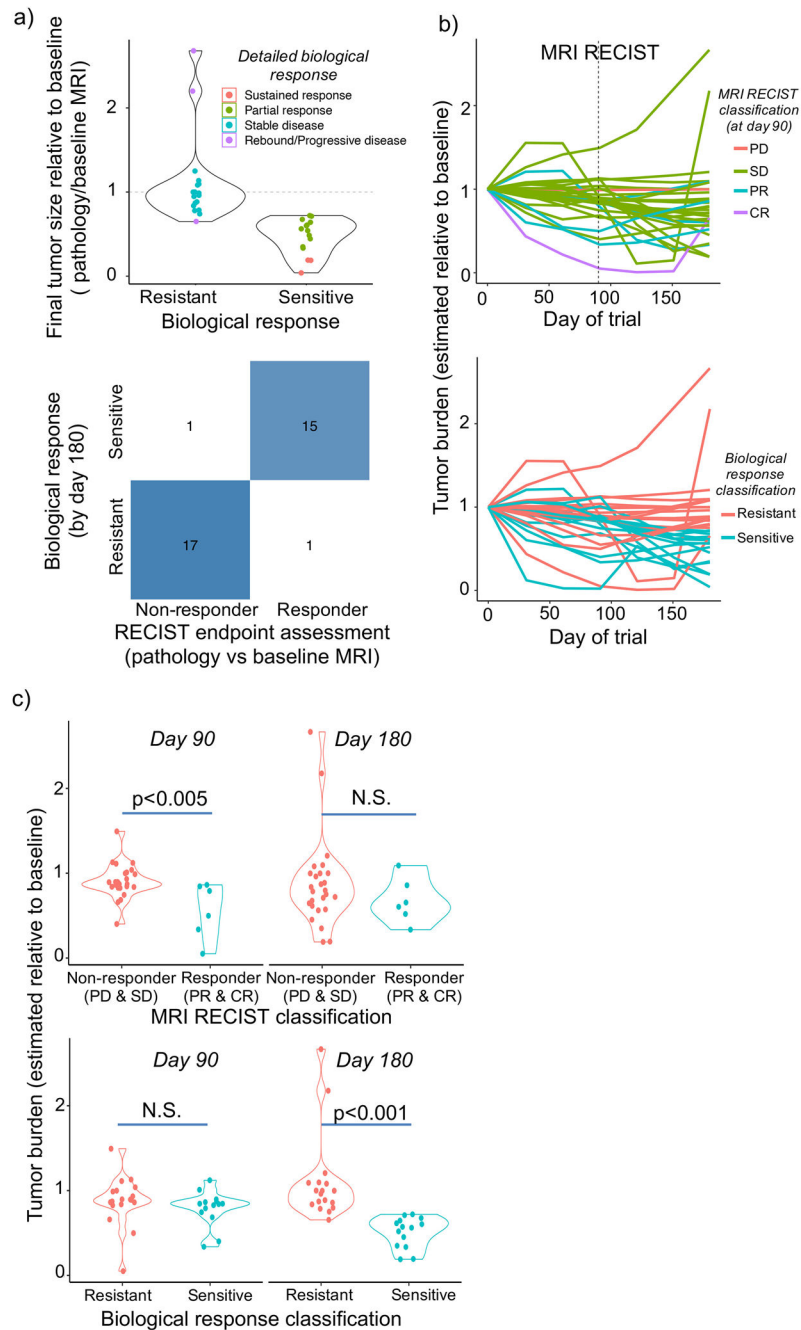
We next assessed the cell cycle consequences of endocrine and combination therapy. As individual marker genes are insufficient to resolve cell cycle state, we reconstructed the cell cycle, using the reCAT model of the cell cycle pseudo-time transitions (44) and a set of cell cycle genes and obtained from the Biocarta cell cycle pathway (48 genes). This signature was repeatedly detected in the pathway analysis and its predicted changes in cell cycle activity mirrored KI-67 antigen expression changes and tumor size dynamics.

Discrete cell cycle states were first identified by applying a Gaussian mixture model to the cell cycle gene set. The reconstruction of cell cycle transitions was then formulated as a traveling salesman problem and the shortest cycle that passes through each cell cycle stage was identified based on subclonal distances on a UMAP, using the arbitrary insertion traveling salesman algorithm (81) (Extended Data Fig.10).

Fluctuations in gene expression throughout the cell cycle (including non-cell cycle genes) were recovered using cyclical generalized additive models. The expected expression of each gene, as cells transition through the cell cycle, was described using a smooth and cyclical cubic spline function relating gene expression to cell stage. Three distinct phases of the cell cycle (G0, G1, S/G2) were identified by reclassifying cell states, by applying a Gaussian mixture model to the reconstructed gene expression of cell cycle genes in each stage. Cell frequencies in each cell cycle phase were calculated and compared across tumors between arms and over time. Treatment induced differences in the frequency of cells from a sample

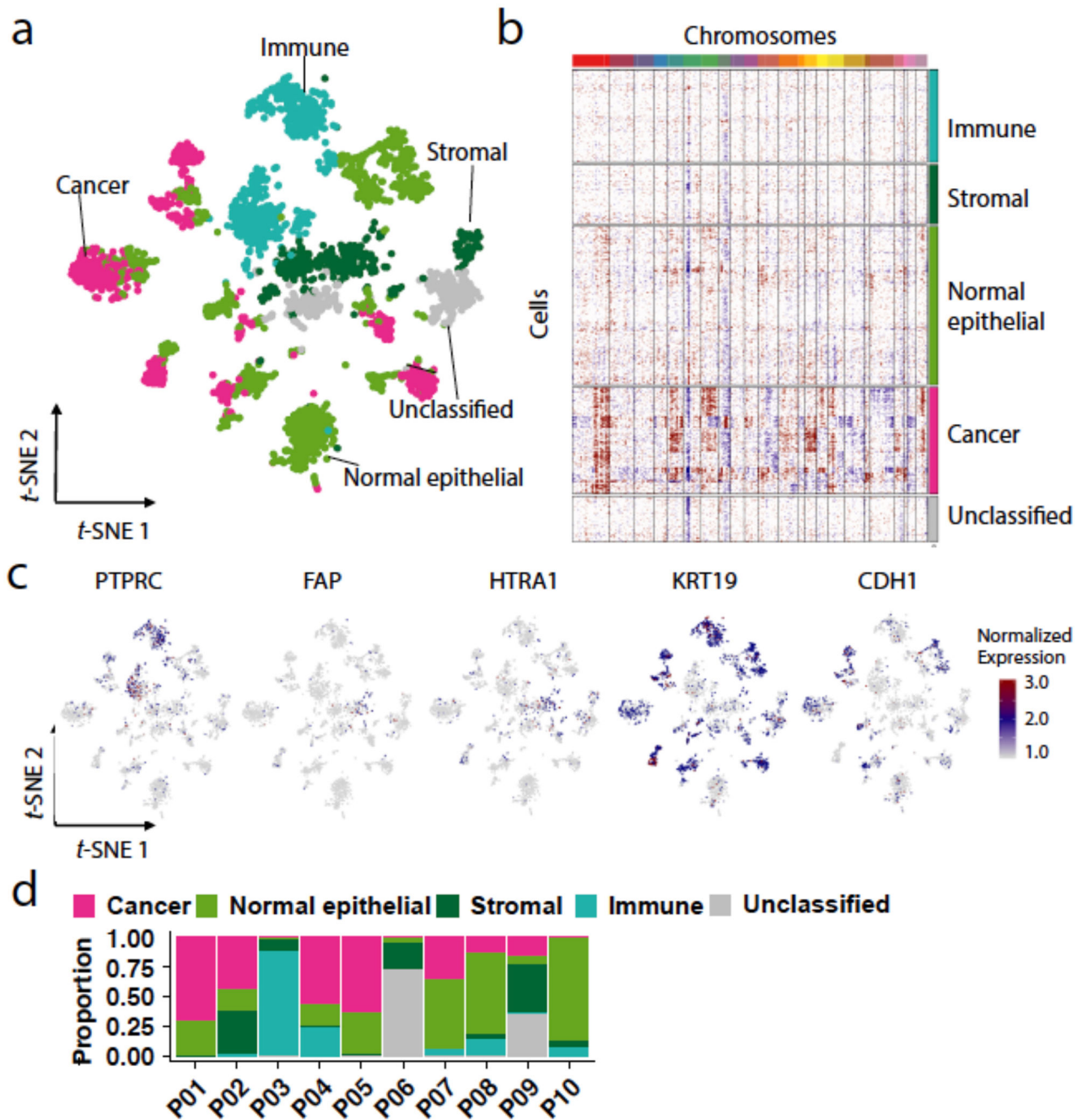
found to be in the proliferative S/G2 phase was assessed using a quasi-binomial generalized additive model (79).

Extended Data



Extended Data Fig. 1. Classification of patient tumors as sensitive or resistant to treatment, reflecting changes in tumor size observed at pathology relative to baseline. Reconstructed trajectories of tumor burden are consistent with results of RECIST 1.1 MRI assessment at day 90 and allow sensitive and resistant tumors to be

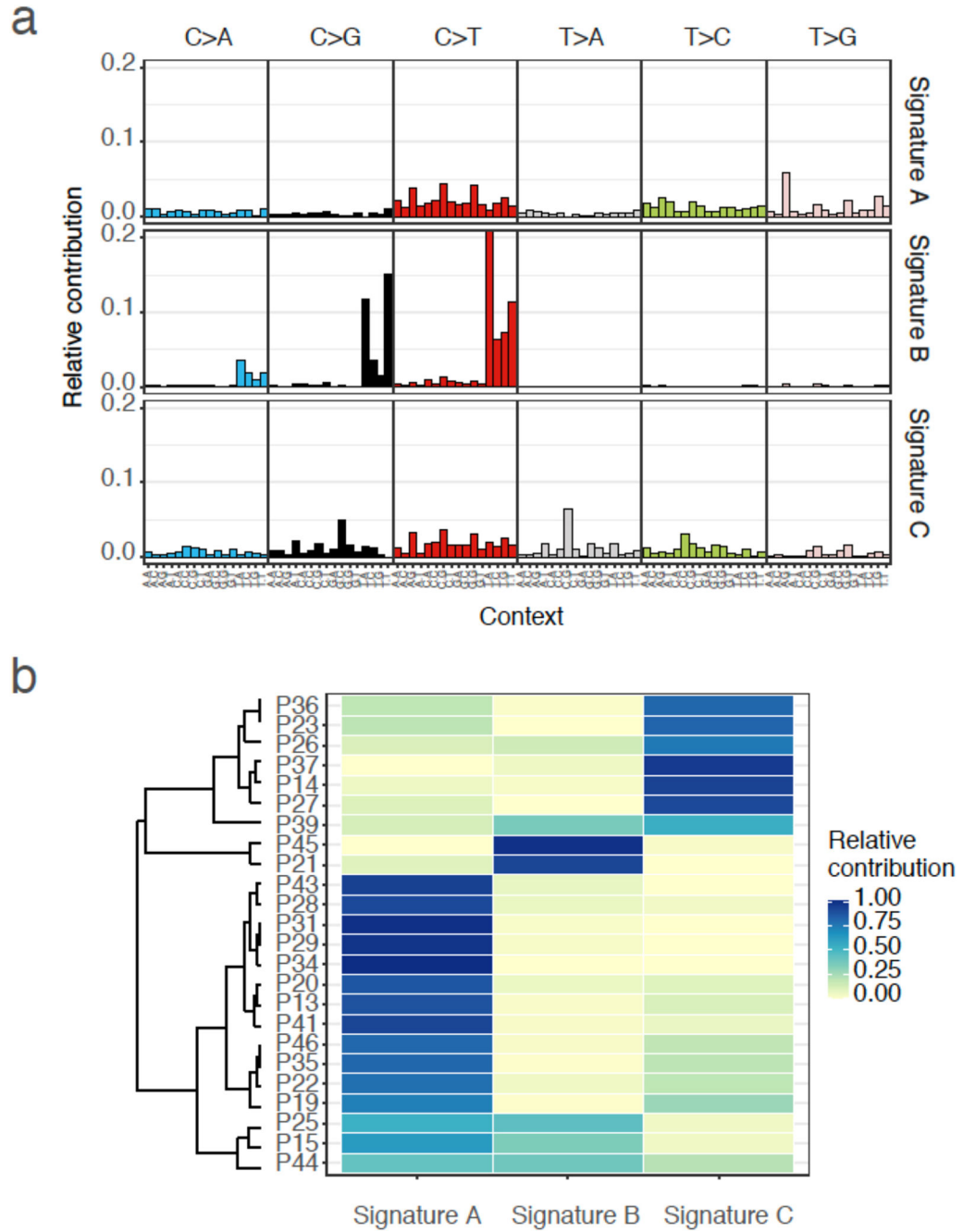
distinguished at end of treatment (day 180). a, Changes in tumor size during therapy for tumors classified as sensitive or resistant. Tumor growth (y-axis) calculated directly from data as the proportion tumor remaining at end of trial (final observed tumor size at pathology/baseline MRI tumor measurement). Values <1 indicate tumor shrinkage, whilst values >1 indicate an increase in size (Dashed horizontal line = no change in size during trial). A detailed biological response classification was determined by classifying tumors with similar trajectories using a Gaussian mixture model (colors). Sustained or partial responses were grouped and defined as sensitive tumors, whilst those with stable, progressive or rebound disease were classified as resistant tumors. The changes in tumor size are highly significantly different between resistance categories (two-sided ANOVA test: $t=4.45$, $p<0.001$). Violins show the distinct distribution of tumor growth observed across patients. Heatmap shows the strong agreement in the end of treatment classification obtained by classifying trajectories of tumor growth vs simple pathology/baseline MRI RECIST assessment of change in size during trial. Number patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: $P=11$, (S=6, R=5); Intermittent high dose ribociclib: $P=12$ (S=6, R=6) Continuous low dose ribociclib: $P=11$ (S=4, R=7). **b,** Spiderplots show the reconstructed trajectories of tumor size (relative to day 0) during the trial, as inferred using all available clinical measurements of patients' tumor size. Predicted tumor sizes at day 90 match the RECIST assessments of tumor response (top panels) whilst trajectories of tumor burden distinguish sensitive (shrinking) and resistant (persistent) tumor through to the end of the trial (bottom panels). Number patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: $P=11$, (S=6, R=5); Intermittent high dose ribociclib: $P=12$ (S=6, R=6) Continuous low dose ribociclib: $P=11$ (S=4, R=7). **c,** Inferred change in tumor size between the start- midpoint (left panel) or start-end (right panel) of the trial, in patient response groups classified by either RECIST assessment at trial midpoint (top row) or the biological response classification from tumor trajectories (bottom row). RECIST assessments distinguish response/non-response at day 90 but not day 180, whilst the biological response classification does distinguish resistance or sensitivity at day 180 (two-sided ANOVA test: MRI day 180 p -value= 0.38 and Biological response day 90 p -value= 0.34). Number patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: $P=11$, (S=6, R=5); Intermittent high dose ribociclib: $P=12$ (S=6, R=6) Continuous low dose ribociclib: $P=11$ (S=4, R=7).



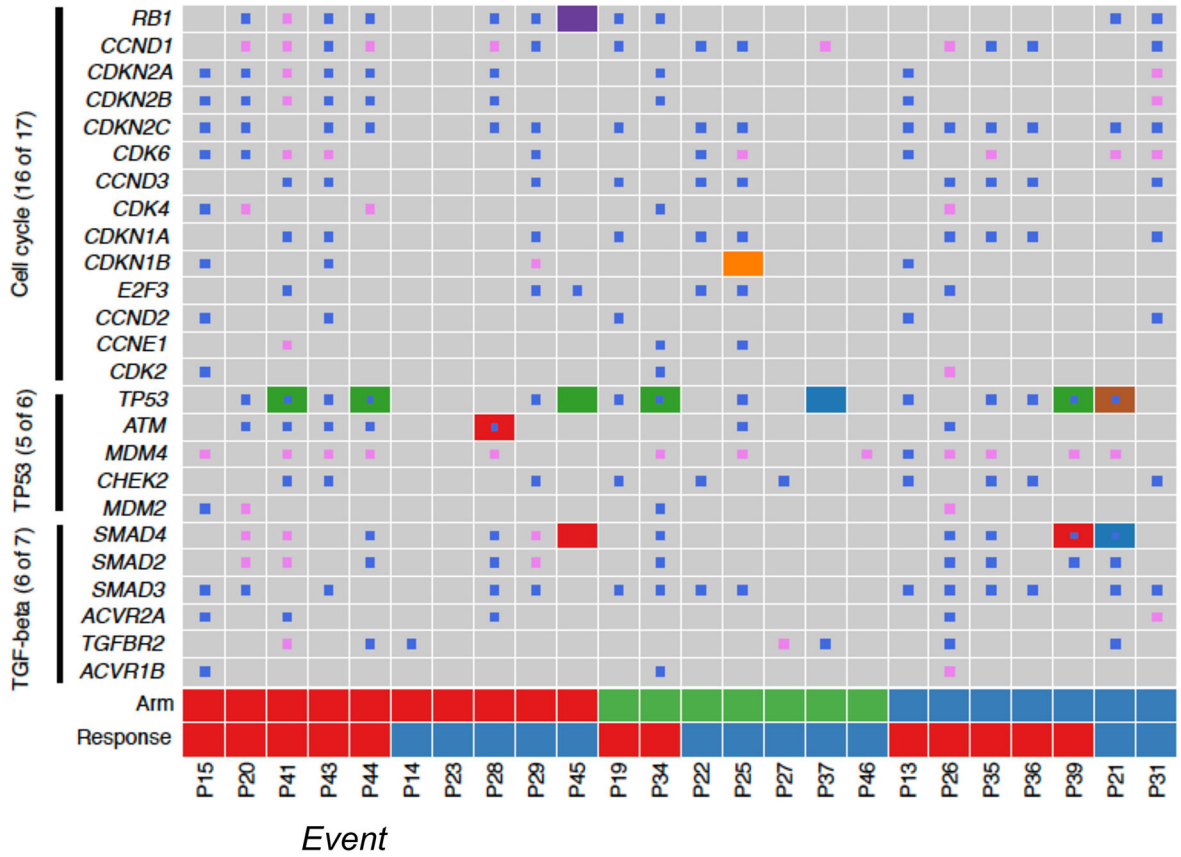
Extended Data Fig. 2. Landscape of tumor and microenvironment of 10 patients with single nucleus isolated by ICELL8 platform.

a, t -SNE plot of 3,484 cells. Cells were classified into cancer cells, normal epithelial cells, immune cells, stromal cells, and unclassified cells, which are indicated by colors and labels. The 3,484 cells are from 7 patients (3 from the Intermittent high dose arm and 4 from the Continuous low dose ribociclib arm). **b**, Gene copy number profile in cancer cells and neighboring normal cells. Blue color indicates copy number loss and red color indicates copy number gain. **c**, Expression of marker genes of cancer cells and normal epithelial cells

(KRT19, CDH1), stromal cells (FAP, HTRA1), and immune cells (PTPRC). **d**, Proportion of cancer cells and neighboring normal cells in each patient.

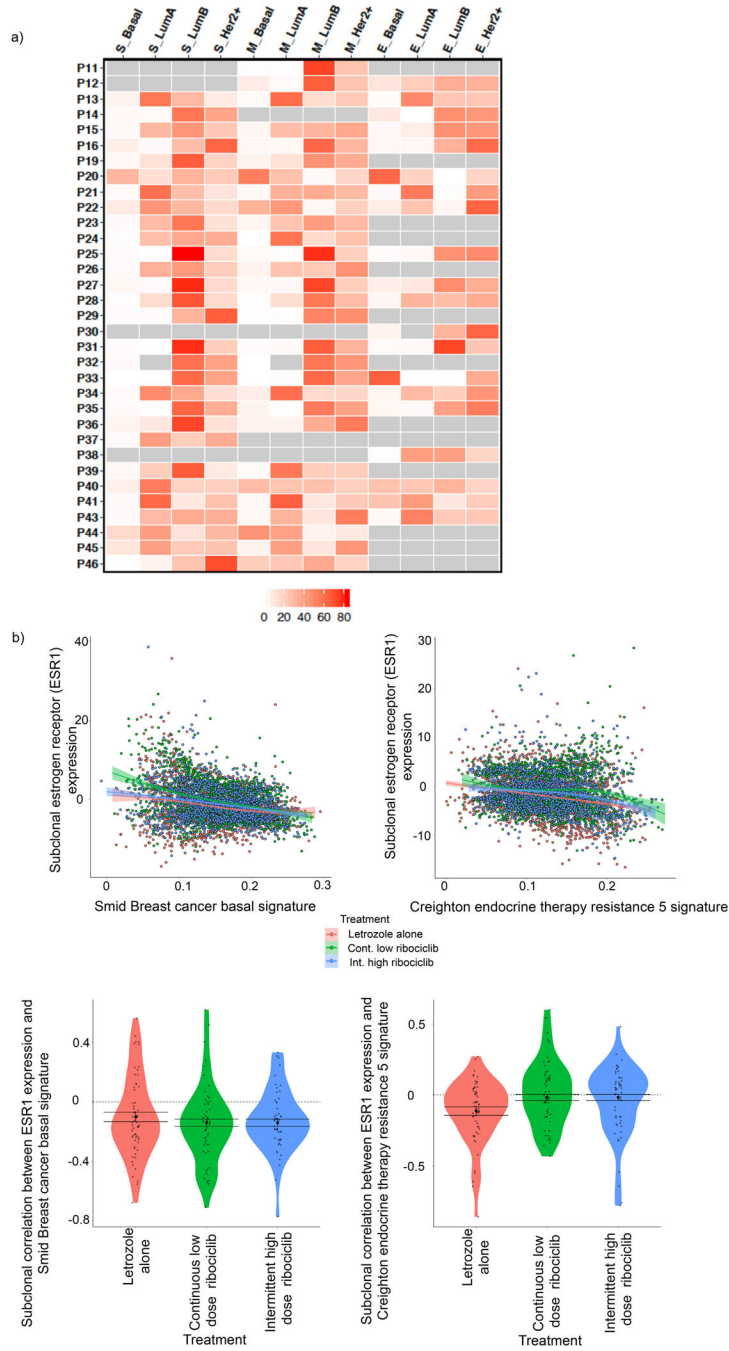


Extended Data Fig. 3. Mutational signature in 24 patients with whole-exome sequencing data. **a**, Relative contribution of trinucleotide changes to three *de novo* mutational signatures identified in 24 patients. **b**, Relative contribution of each mutational signature to mutations in each patient.



- In frame deletion
 - Frame shift deletion
 - Splice site
 - Nonsense mutation
 - Missense mutation
 - Frame shift insertion
 - Deletion
 - Amplification
 - Multi hit
 - Complex event
- Treatment*
- Letrozole alone
 - Continuous low dose ribociclib
 - Intermittent high dose ribociclib
- Response*
- Resistant
 - Sensitive

Extended Data Fig. 4. Mutated genes in three frequently altered oncogenic pathways. Genes are grouped by oncogenic pathway. Presence of gene mutations in each patient is colored as indicated in the legend. Treatment arm and clinical response (Response: sensitive, resistant) are indicated in final two rows of the plot (colors indicated in legend).



Extended Data Fig. 5. Intrinsic subtype of 35 patients with single nucleus isolated by 10x genomics platform and reduced subclonal estrogen receptor (ESR1) expression at end of therapy as correlated with increased basal-like pathway and Creighton endocrine therapy resistance signatures, independent of treatment.

a, Intrinsic subtyping. Each row represents a patient and each column represents an intrinsic subtype at three timepoints. The proportion of cancer cells in each intrinsic subtype was indicated by colors ranging from 0 to 85. Patient samples without cancer cells were indicated by gray. **b,** Reduced subclonal estrogen receptor (ESR1) expression. Top row shows the ESR1 expression and basal-like (left) and endocrine resistance (right) pathway

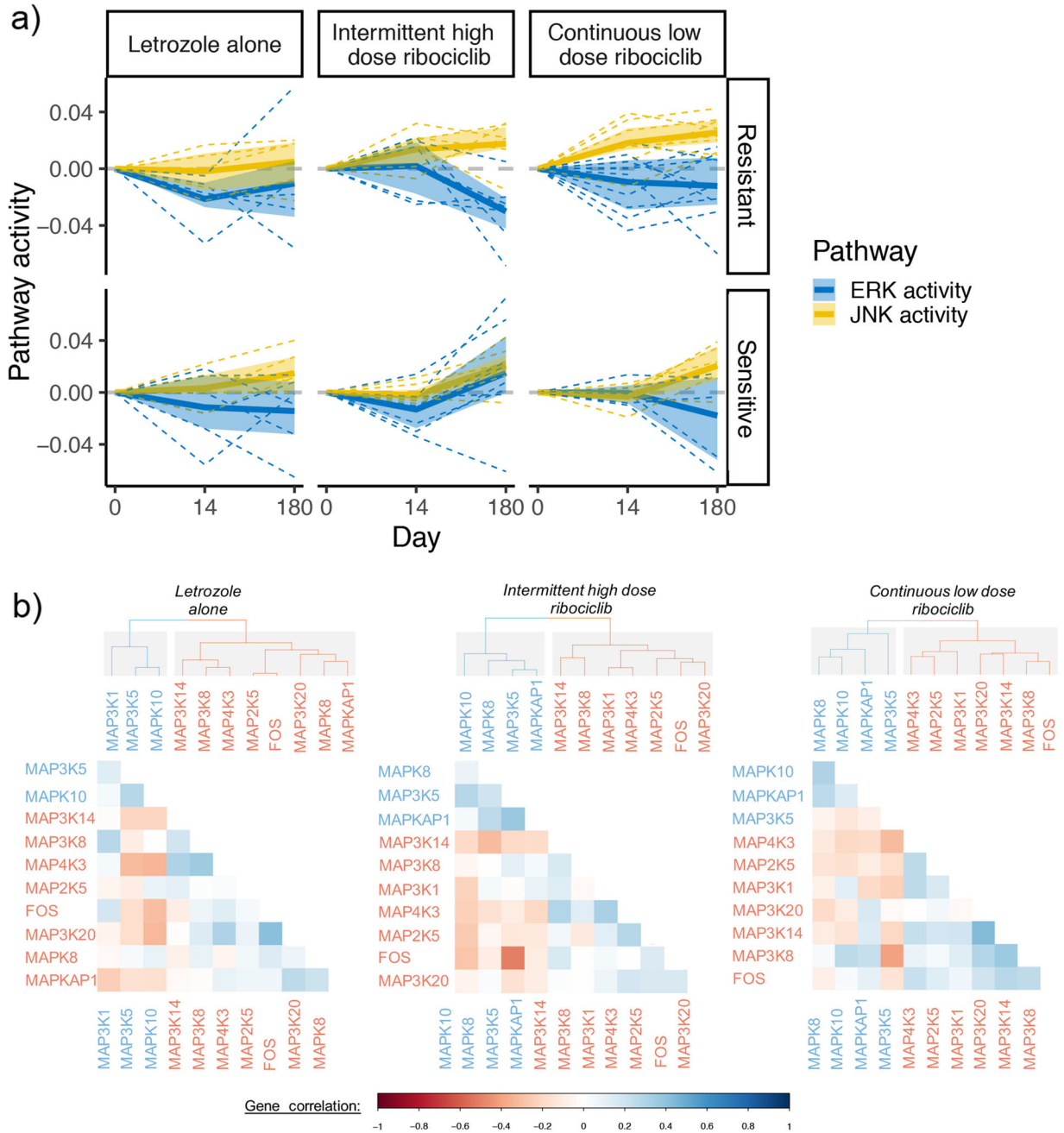
signatures across subclonal cancer populations with differing MAPK activation (points) and the coloration signifies the treatment received. Fitted lines show the overall trend between ESR1 expression and pathway activity (shaded regions show 95% confidence bands). Bottom row shows the correlation between ESR1 expression and basal-like (left) and endocrine resistance (right) pathway signatures for each cancer subclone present at end of trial, in patients treated with different therapies (colors). Black points and error bars signifies the mean and confidence interval for the correlation between ESR1 and pathway activity under each treatment. Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



Extended Data Fig. 6. Divergence of JNK and ERK signalling pathway activity during treatment with combination therapy, especially in resistant tumors and heatmaps of the correlation between MAPK gene expression in each treatment arm (columns), showing the dichotomy between JNK and ERK activating genes across treatments.

A, JNK and ERK expression (color=pathway) during treatment (columns) in sensitive and resistant tumors (rows). Pathway trends determined across patients using hierarchical regression (solid lines). Inter-patient variability in pathway activity shown by dashed lines indicating patient specific responses and shaded regions showing confidence intervals of model estimates (JNK ssGSEA pathway=St JNK MAPK and ERK pathway=Biocarta ERK). Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by

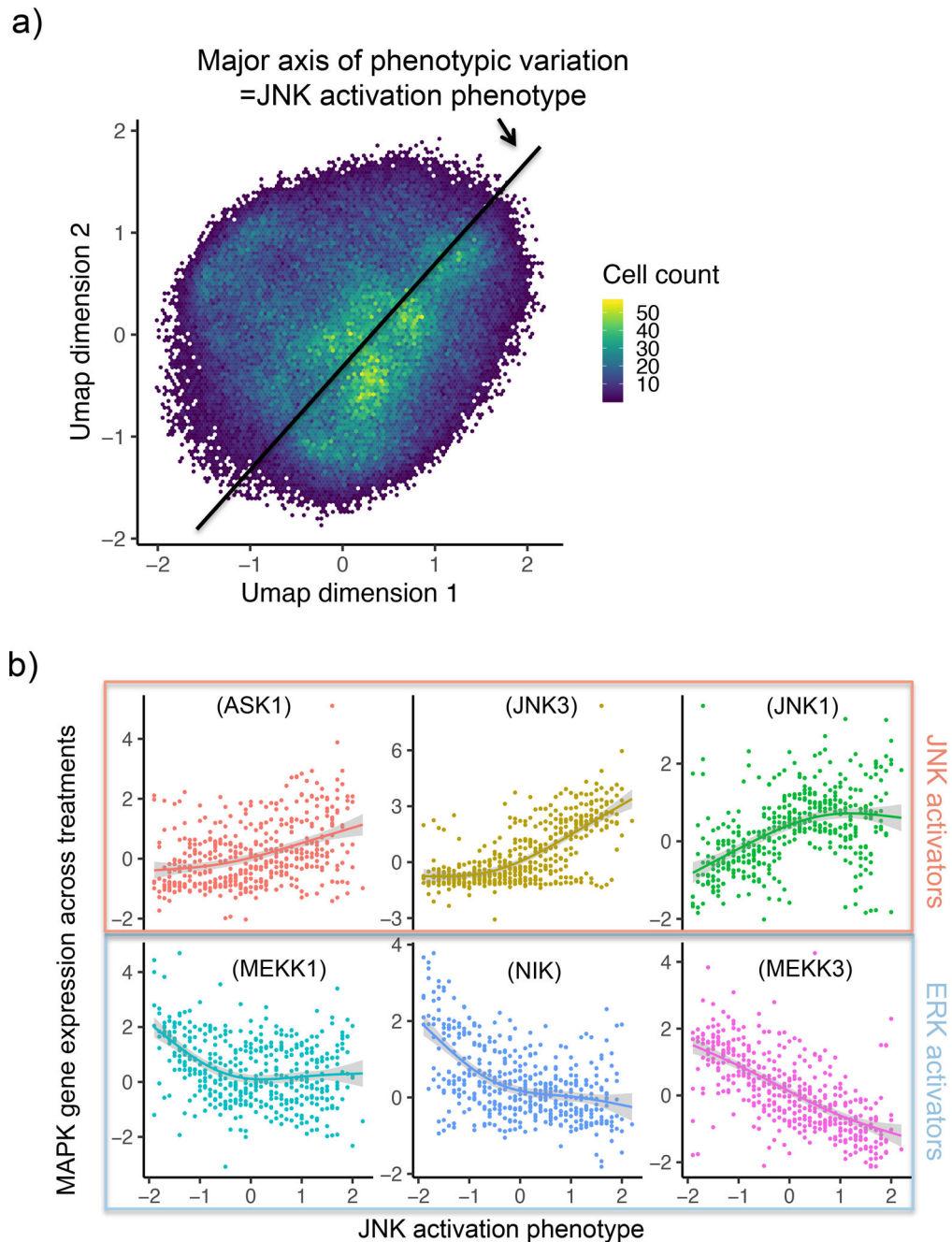
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7. **B**, Dendrograms show the collinearity of MAPK gene expression following each endocrine or combination therapies (columns).



Extended Data Fig. 7. Construction of the overall JNK activation phenotype score, utilizing this collinearity of gene expression between ERK and JNK genes

a, UMAP dimension reduction of MAPK genes, showing the bivariate Gaussian distribution of UMAP values, centered around the major axis of phenotypic variation (black line). The

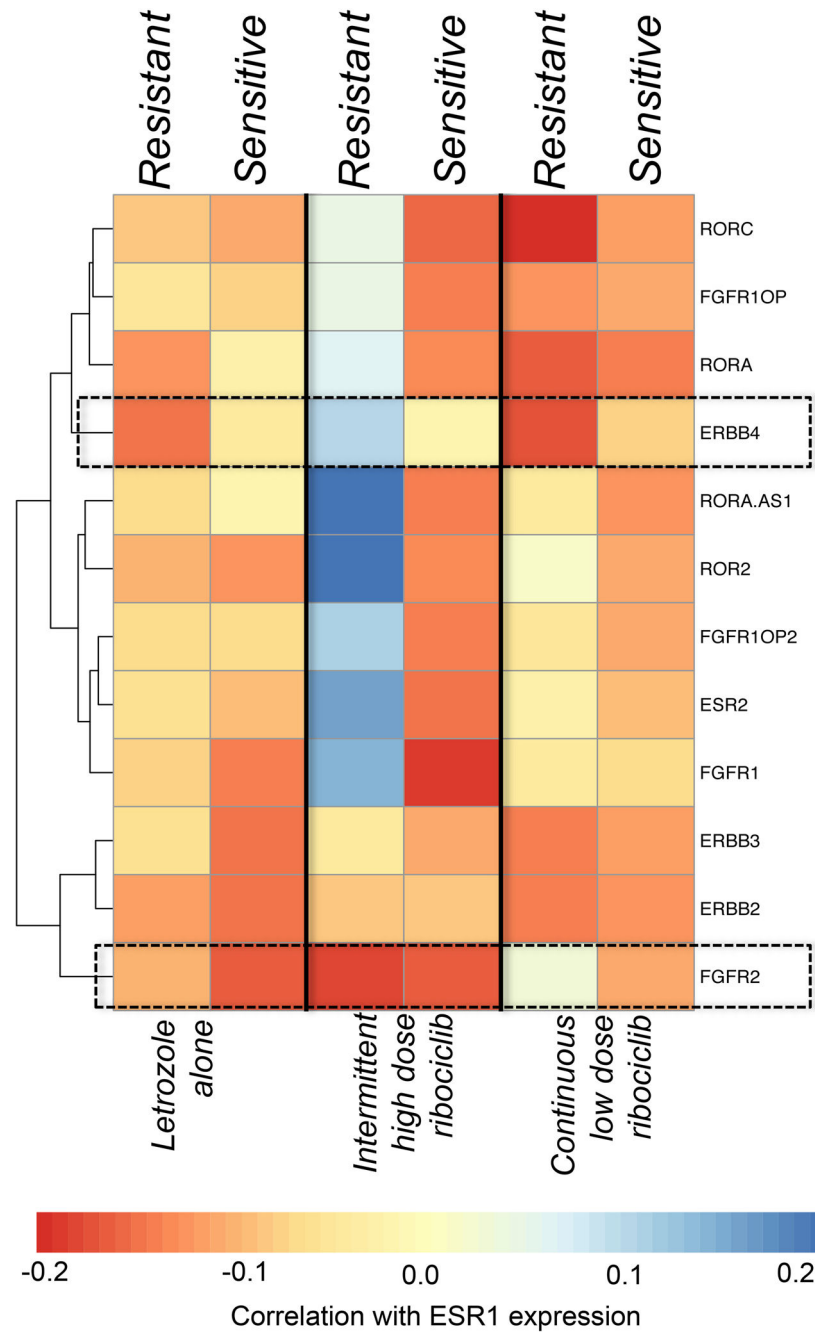
frequency of cells found in different parts of the UMAP phenotype space is shown by the color gradient. The major axis of phenotypic variation (the JNK activation phenotype) is identified as the first principle component in the UMAP phenotype space. **b**, Relationship between the JNK activation phenotype and expression of MAPK genes that are known a JNK activators (red) or ERK activators (blue) across subclonal cancer populations. Loess smooths are added showing the positive relationship between the JNK phenotype score and key JNK activators and the negative association between ERK activators and the JNK phenotype. Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7.

Author Manuscript

Author Manuscript

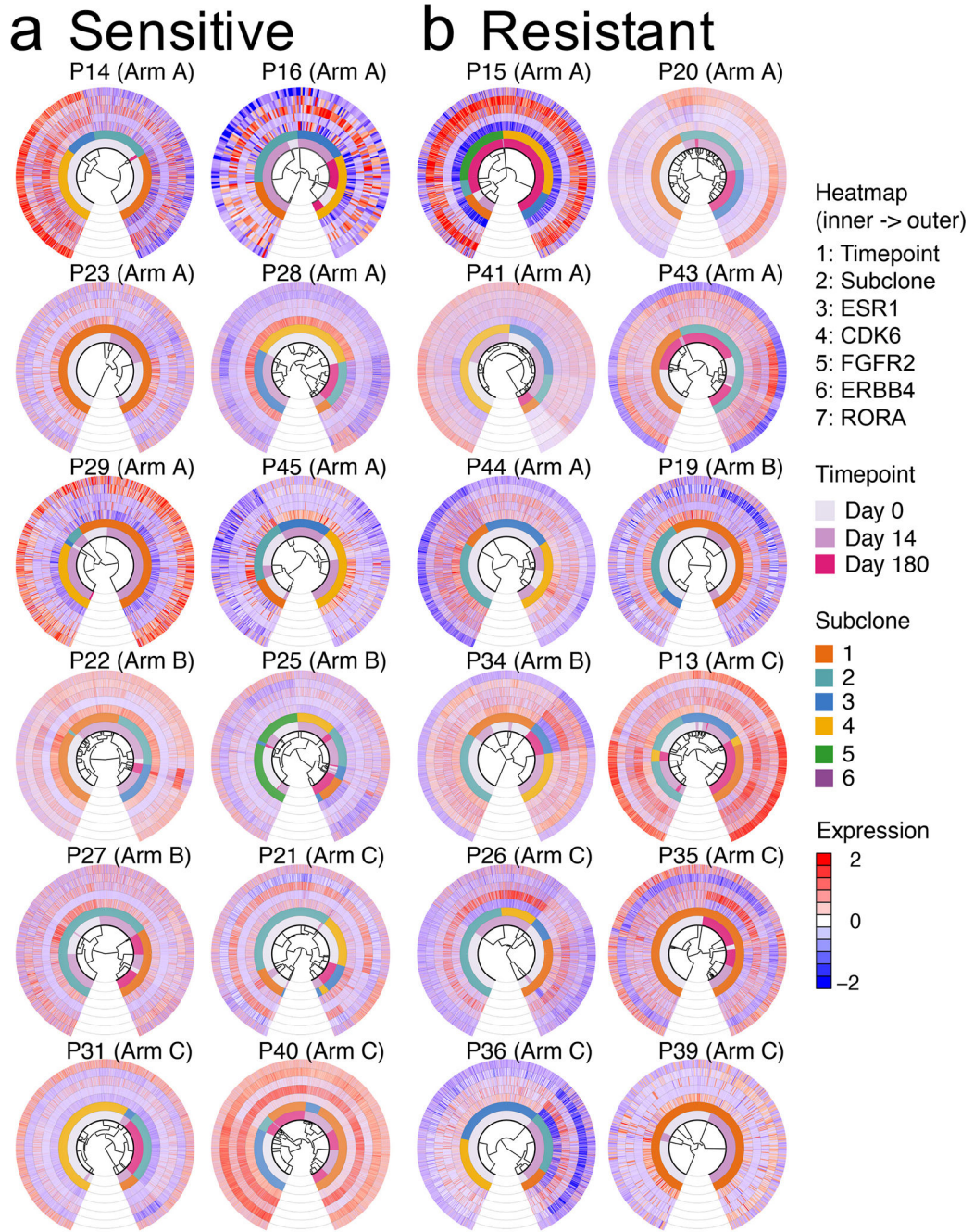
Author Manuscript

Author Manuscript



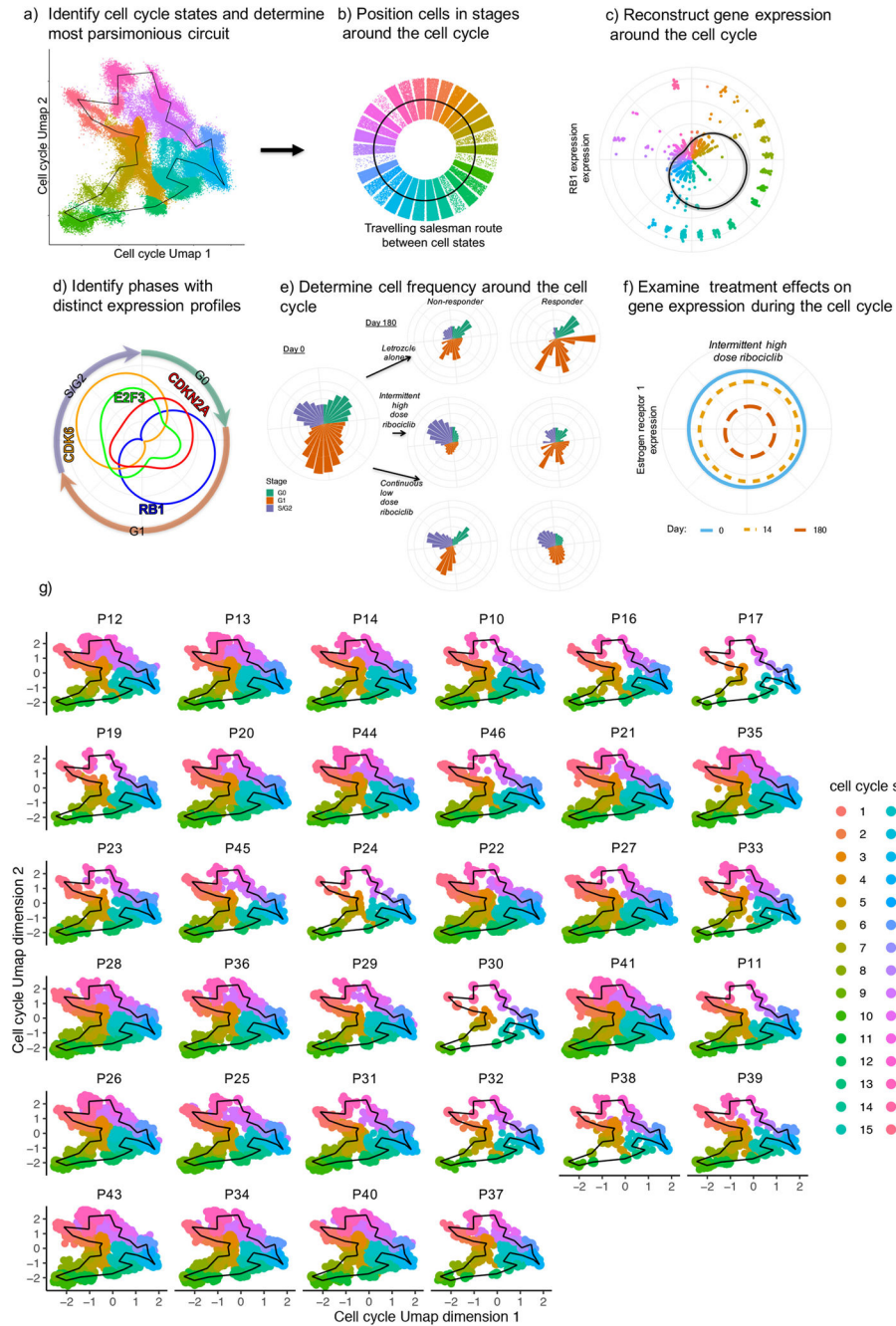
Extended Data Fig. 8. Correlation of growth factor receptors expression with estrogen pathway activity (Hallmark estrogen response early) in cancer cells from sensitive and resistant tumors under each therapy.

Strong negative correlations identify genes that are upregulated as estrogen signaling is lost. Specifically, tumors resistant to intermittent high dose and continuous low dose show compensatory activation of FGFR2 and ERBB4 respectively. Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7.



Extended Data Fig. 9. Transcriptional heterogeneity of key resistant genes.

a, sensitive and **b**, resistant tumors. For each patient's tumor cells, a single-cell phylogenetic tree is shown at the center of circos plot. Cell annotation (timepoint and subclone) as well as expression of key resistance genes (ESR1, CDK6, FGFR2, ERBB4, RORA) are shown as heatmap. Phylogenetic tree of cells were constructed based on the distance between cell gene copy number profile. Subclones were inferred based on gene copy number profile. Zinbwave normalized gene expression were centered and scaled.



Extended Data Fig. 10. Reconstruction of cell cycle, fluctuations in gene expression during the cell cycle, distinct cell cycle phases, frequencies of cells throughout the cell cycle and shifts in gene expression within the cell cycle during therapy.

a, Single cell RNA seq gene expression profiles of cell cycle genes are extracted and used to perform dimension reduction with the UMAP algorithm. Cell cycle states (colors) with differing expression were identified using a Gaussian mixture model and the transitions between these states determined by the shortest distance to travel through each state and return to the original (Traveling salesman route=black line). **b**, Cells states ordered along the traveling salesman route. **c**, Example of fluctuations in gene expression of cells around

the cell cycle (distance of points from origin=RB1 expression; colors=cell cycle state) Reconstruction of the fluctuation in average gene expression is predicted using a cyclical generalised additive model (black line with shaded confidence bands). **d**, Reconstructed fluctuations (coloured curves) in expression of genes around the cell cycle are used to classify distinct phases of the cell cycle (annotated by arrows around). Here we show four examples of key cell cycle genes, which influence the classification of cell cycle phases (G0, G1, S/G2). **e**, The frequency of cells in each stage of the cell cycle (height of bars) was counted and used to examine changes in the fraction of sampled cells in each phases cell cycle phase over time and between treatment and response groups. **f**, During treatment, the changes in gene expression fluctuations around the cell cycle were examined. Distance of the curve from the origin indicates gene expression and colored curves shows expression at different timepoints. **g**, Consistent cell cycle stages present across patients. For each patient (subpanel), single cell RNAseq gene expression profiles for cell cycle genes were extracted and the fitted UMAP model used to project cells onto the lower dimensional cell phenotype space (UMAP dimensions 1 and 2). Cell cycle stages (colors) with differing expression, identified using the Gaussian mixture model, were overlaid, showing that all patients have cells that are distributed across the cell cycle phenotype space. The traveling salesman route (black line) shows the transitions between these stages, as determined by the shortest distance to travel through each state and return to the original.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the anonymous patients from the trial that made this study possible. A.H.B, J.G., J.C., J.T.C, P.C., and F.A. were supported by the National Cancer Institute of the National Institutes of Health under award number U54CA209978. The content is solely the authors responsibility and does not necessarily represent the official views of the NIH. The High-Throughput Genomics Shared Resource was supported by the NIH Award Number P30CA042014. J.T.C. was supported by a Cancer Prevention Research Institute of Texas Core Facility Support Award (RP170668).

Competing Interests Statement

Ruth O'Regan participates on the advisory board for Cyclacel, PUMA, Biotheranostics, Lilly, Pfizer, Genentech, Novartis; declares research funding from Pfizer, Novartis, Seattle Genetics, PUMA. Priyanka Sharma declares research funding from Novartis, Merck, Bristol Myers Squibb. Consulting: Seattle Genetics, Merck, Novartis, AstraZeneca, Immunomedics, Exact Biosciences. Laura Spring participates on the advisory board for Novartis, Lumicell, Puma Biotechnology and AvroBio. Cynthia Ma declares research funding from Pfizer, Puma; Consulting: Eisai, Athenex, OncoSignal, Agendia, Biovica, AstraZeneca, Seattle Genetics. Kari Wisinski declares research funding and clinical trial involvement with Novartis, Eli Lilly, Astra Zeneca, Sanofi and Pfizer. He participated on an advisory board for Eisai, Pfizer and Astra Zeneca. Kevin Kalinsky is a medical advisor to Immunomedics, Pfizer, Novartis, Eisai, Eli-Lilly, Amgen, Merck, Seattle Genetics and Astra Zeneca; receives institutional support from Immunomedics, Novartis, Incyte, Genentech/Roche, Eli-Lilly, Pfizer, Calithera Biosciences, Acetylon, Seattle Genetics, Amgen, Zentalis Pharmaceuticals, and CytomX Therapeutics; and his spouse is employed by Grail and previously by Array Biopharma and Pfizer. Anne O'Dea Consults for the Pfizer, PUMA Biotechnology, Astra Zeneca, and Daiichi Sankyo. Qamar Khan declares research funding from Novartis. All other authors have no conflicts of interest to disclose.

Data availability

Raw single cell RNA-seq data are available through GEO under accession code GSE158724 at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE158724>. The DNA-Seq data is available from dbGaP at phs002287.v1.p1 at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs002287.v1.p1. Source data is provided with this manuscript as individual excel files (one per figure) and listed in the Inventory. Code for Figures 1–6 is available on our GitHub repository at https://github.com/U54Bioinformatics/FELINE_project. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

References

1. Harbeck N, Gnant M. Breast cancer. *Lancet*. 2017;389(10074):1134–50. Epub 2016/11/21. doi: 10.1016/s0140-6736(16)31891-8. [PubMed: 27865536]
2. Rani A, Stebbing J, Giamas G, Murphy J. Endocrine Resistance in Hormone Receptor Positive Breast Cancer-From Mechanism to Therapy. *Front Endocrinol (Lausanne)*. 2019;10:245. Epub 2019/06/11. doi: 10.3389/fendo.2019.00245. [PubMed: 31178825]
3. Zhao H, Zhou L, Shanguan AJ, Bulun SE. Aromatase expression and regulation in breast and endometrial cancer. *J Mol Endocrinol*. 2016;57(1):R19–33. Epub 2016/04/14. doi: 10.1530/jme-15-0310. [PubMed: 27067638]
4. Osborne CK. Tamoxifen in the treatment of breast cancer. *N Engl J Med*. 1998;339(22):1609–18. Epub 1998/11/26. doi: 10.1056/nejm199811263392207. [PubMed: 9828250]
5. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet*. 2005;365(9472):1687–717. Epub 2005/05/17. doi: 10.1016/s0140-6736(05)66544-0. [PubMed: 15894097]
6. Company ELA. Endocrine Therapy With or Without Abemaciclib (LY2835219) Following Surgery in Participants With Breast Cancer (monarchE) 2017 [updated 9 18, 2020; cited 2020 October 13]. Available from: <https://www.clinicaltrials.gov/ct2/show/NCT03155997>.
7. Pharmaceuticals N. A Trial to Evaluate Efficacy and Safety of Ribociclib With Endocrine Therapy as Adjuvant Treatment in Patients With HR+/HER2– Early Breast Cancer (NATALEE) 2018 [updated 10 2, 2020; cited 2020 October 13]. Available from: <https://clinicaltrials.gov/ct2/show/NCT03701334>.
8. Bertagnolli MME, DeMichele A, Gnant M. PALbociclib CoLaborative Adjuvant Study (PALLAS) 2015 [updated 3 9, 2020; cited 2020 October 13]. Available from: <https://clinicaltrials.gov/ct2/show/NCT02513394>.
9. PENELOPE-B TRIAL OF IBRANCE® (PALBOCICLIB) IN EARLY BREAST CANCER DID NOT MEET PRIMARY ENDPOINT: Pfizer; 2020 [updated 10 9, 2020; cited 2020 October 13]. Available from: <https://www.pfizer.com/news/press-release/press-release-detail/penelope-b-trial-ibrancer-palbociclib-early-breast-cancer>.
10. Weintraub SJ, Chow KN, Luo RX, Zhang SH, He S, Dean DC. Mechanism of active transcriptional repression by the retinoblastoma protein. *Nature*. 1995;375(6534):812–5. Epub 1995/06/29. doi: 10.1038/375812a0. [PubMed: 7596417]
11. Finn RS, Aleshin A, Slamon DJ. Targeting the cyclin-dependent kinases (CDK) 4/6 in estrogen receptor-positive breast cancers. *Breast Cancer Res*. 2016;18(1):17. Epub 2016/02/10. doi: 10.1186/s13058-015-0661-5. [PubMed: 26857361]
12. Portman N, Alexandrou S, Carson E, Wang S, Lim E, Caldon CE. Overcoming CDK4/6 inhibitor resistance in ER-positive breast cancer. *Endocr Relat Cancer*. 2019;26(1):R15–R30. Epub 2018/11/06. doi: 10.1530/ERC-18-0317. [PubMed: 30389903]
13. Sabbah M, Courilleau D, Mester J, Redeuilh G. Estrogen induction of the cyclin D1 promoter: Involvement of a cAMP response-like element. *Proceedings of the National Academy of Sciences*. 1999;96(20):11217–22. doi: 10.1073/pnas.96.20.11217.

14. Razavi P, Chang MT, Xu G, Bandlamudi C, Ross DS, Vasan N, Cai Y, Bielski CM, Donoghue MTA, Jonsson P, Penson A, Shen R, Pareja F, Kundra R, Middha S, Cheng ML, Zehir A, Kandath C, Patel R, Huberman K, Smyth LM, Jhaveri K, Modi S, Traina TA, Dang C, Zhang W, Weigelt B, Li BT, Ladanyi M, Hyman DM, Schultz N, Robson ME, Hudis C, Brogi E, Viale A, Norton L, Dickler MN, Berger MF, Iacobuzio-Donahue CA, Chandarlapaty S, Scaltriti M, Reis-Filho JS, Solit DB, Taylor BS, Baselga J. The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. *Cancer Cell*. 2018;34(3):427–38 e6. Epub 2018/09/12. doi: 10.1016/j.ccell.2018.08.008. [PubMed: 30205045]
15. Formisano L, Lu Y, Servetto A, Hanker AB, Jansen VM, Bauer JA, Sudhan DR, Guerrero-Zotano AL, Croessmann S, Guo Y, Ericsson PG, Lee K-m, Nixon MJ, Schwarz LJ, Sanders ME, Dugger TC, Cruz MR, Behdad A, Cristofanilli M, Bardia A, O’Shaughnessy J, Nagy RJ, Lanman RB, Solovieff N, He W, Miller M, Su F, Shyr Y, Mayer IA, Balko JM, Arteaga CL. Aberrant FGFR signaling mediates resistance to CDK4/6 inhibitors in ER+ breast cancer. *Nature Communications*. 2019;10(1):1373. doi: 10.1038/s41467-019-09068-2.
16. Kato S, Endoh H, Masuhiro Y, Kitamoto T, Uchiyama S, Sasaki H, Masushige S, Gotoh Y, Nishida E, Kawashima H, Metzger D, Chambon P. Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase. *Science*. 1995;270(5241):1491–4. Epub 1995/12/01. doi: 10.1126/science.270.5241.1491. [PubMed: 7491495]
17. Filardo EJ, Quinn JA, Bland KI, Frackelton AR Jr. Estrogen-induced activation of Erk-1 and Erk-2 requires the G protein-coupled receptor homolog, GPR30, and occurs via trans-activation of the epidermal growth factor receptor through release of HB-EGF. *Mol Endocrinol*. 2000;14(10):1649–60. Epub 2000/10/24. doi: 10.1210/mend.14.10.0532. [PubMed: 11043579]
18. Bi R, Foy MR, Vouimba R-M, Thompson RF, Baudry M. Cyclic changes in estradiol regulate synaptic plasticity through the MAP kinase pathway. *Proceedings of the National Academy of Sciences*. 2001;98(23):13391–5. doi: 10.1073/pnas.241507698.
19. Khan Q Letrozole Plus Ribociclib or Placebo as Neo-adjuvant Therapy in ER-positive, HER2-negative Early Breast Cancer (FELINE) 2016 [updated 8 29, 2019; cited 2020 September 11]. Available from: <https://clinicaltrials.gov/ct2/show/NCT02712723>.
20. Griffiths JI, Chen J, O’Dea A, Sharma P, Winblad O, Trivedi M, Kalinsky K, Wisinski K, O’Regan R, Makhoul I, Spring L, Bardia A, Yuan Y, Jahanzeb M, Adler F, Cohen A, Bild A, Khan Q. Reconstructing tumor trajectories during therapy through integration of multiple measurement modalities. *bioRxiv*. 2021:2021.01.14.426737. doi: 10.1101/2021.01.14.426737.
21. Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Böhm-Vélez M, Pisano ED, Jong RA, Evans WP, Morton MJ, Mahoney MC, Larsen LH, Barr RG, Farria DM, Marques HS, Boparai K. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *Jama*. 2008;299(18):2151–63. Epub 2008/05/15. doi: 10.1001/jama.299.18.2151. [PubMed: 18477782]
22. Marinovich ML, Houssami N, Macaskill P, Sardanelli F, Irwig L, Mamounas EP, von Minckwitz G, Brennan ME, Ciatto S. Meta-analysis of magnetic resonance imaging in detecting residual breast cancer after neoadjuvant therapy. *J Natl Cancer Inst*. 2013;105(5):321–33. Epub 2013/01/09. doi: 10.1093/jnci/djs528. [PubMed: 23297042]
23. Risso D, Perraudeau F, Gribkova S, Dudoit S, Vert JP. A general and flexible method for signal extraction from single-cell RNA-seq data. *Nat Commun*. 2018;9(1):284. Epub 2018/01/20. doi: 10.1038/s41467-017-02554-5. [PubMed: 29348443]
24. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive Integration of Single-Cell Data. *Cell*. 2019;177(7):1888–902 e21. Epub 2019/06/11. doi: 10.1016/j.cell.2019.05.031. [PubMed: 31178118]
25. Tickle TI GC, Brown M, Haas B. inferCNV of the Trinity CTAT Project 2019 [cited 2020 September 14]. Available from: <https://github.com/broadinstitute/inferCNV>.
26. van der Maaten L, Hinton G. Visualizing Data using t-SNE. *Journal of Machine Learning Research*. 2008;9:2579–605.
27. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR, Butte AJ, Bhattacharya M. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature Immunology*. 2019;20(2):163–72. doi: 10.1038/s41590-018-0276-y. [PubMed: 30643263]

28. Roth A, Khattra J, Yap D, Wan A, Laks E, Biele J, Ha G, Aparicio S, Bouchard-Côté A, Shah SP. PyClone: statistical inference of clonal population structure in cancer. *Nat Methods*. 2014;11(4):396–8. Epub 2014/03/19. doi: 10.1038/nmeth.2883. [PubMed: 24633410]
29. Brady SW, McQuerry JA, Qiao Y, Piccolo SR, Shrestha G, Jenkins DF, Layer RM, Pedersen BS, Miller RH, Esch A, Selitsky SR, Parker JS, Anderson LA, Dalley BK, Factor RE, Reddy CB, Boltax JP, Li DY, Moos PJ, Gray JW, Heiser LM, Buys SS, Cohen AL, Johnson WE, Quinlan AR, Marth G, Werner TL, Bild AH. Combating subclonal evolution of resistant cancer phenotypes. *Nat Commun*. 2017;8(1):1231. Epub 2017/11/03. doi: 10.1038/s41467-017-01174-3. [PubMed: 29093439]
30. Kim C, Gao R, Sei E, Brandt R, Hartman J, Hatschek T, Crosetto N, Foukakis T, Navin NE. Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. *Cell*. 2018;173(4):879–93 e13. Epub 2018/04/24. doi: 10.1016/j.cell.2018.03.041. [PubMed: 29681456]
31. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, Schinzel AC, Sandy P, Meylan E, Scholl C, Fröhling S, Chan EM, Sos ML, Michel K, Mermel C, Silver SJ, Weir BA, Reiling JH, Sheng Q, Gupta PB, Wadlow RC, Le H, Hoersch S, Wittner BS, Ramaswamy S, Livingston DM, Sabatini DM, Meyerson M, Thomas RK, Lander ES, Mesirov JP, Root DE, Gilliland DG, Jacks T, Hahn WC. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*. 2009;462(7269):108–12. Epub 2009/10/23. doi: 10.1038/nature08460. [PubMed: 19847166]
32. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739–40. Epub 2011/05/07. doi: 10.1093/bioinformatics/btr260. [PubMed: 21546393]
33. Chen SY, Feng Z, Yi X. A general introduction to adjustment for multiple comparisons. *J Thorac Dis*. 2017;9(6):1725–9. Epub 2017/07/26. doi: 10.21037/jtd.2017.05.34. [PubMed: 28740688]
34. Iwase H, Greenman JM, Barnes DM, Bobrow L, Hodgson S, Mathew CG. Loss of heterozygosity of the oestrogen receptor gene in breast cancer. *Br J Cancer*. 1995;71(3):448–50. Epub 1995/03/01. doi: 10.1038/bjc.1995.91. [PubMed: 7880722]
35. Herynk MH, Fuqua SA. Estrogen receptor mutations in human disease. *Endocr Rev*. 2004;25(6):869–98. Epub 2004/12/08. doi: 10.1210/er.2003-0010. [PubMed: 15583021]
36. Arpino G, Wiechmann L, Osborne CK, Schiff R. Crosstalk between the estrogen receptor and the HER tyrosine kinase receptor family: molecular mechanism and clinical implications for endocrine therapy resistance. *Endocr Rev*. 2008;29(2):217–33. Epub 2008/01/25. doi: 10.1210/er.2006-0045. [PubMed: 18216219]
37. Thomas LW, Lam C, Edwards SW. Mcl-1; the molecular regulation of protein function. *FEBS Letters*. 2010;584(14):2981–9. doi: 10.1016/j.febslet.2010.05.061. [PubMed: 20540941]
38. Hirata Y, Sugie A, Matsuda A, Matsuda S, Koyasu S. TAK1–JNK Axis Mediates Survival Signal through Mcl1 Stabilization in Activated T Cells. *The Journal of Immunology*. 2013;190(9):4621–6. doi: 10.4049/jimmunol.1202809. [PubMed: 23547112]
39. Bausch-Fluck D, Hofmann A, Bock T, Frei AP, Cerciello F, Jacobs A, Moest H, Omasits U, Gundry RL, Yoon C, Schiess R, Schmidt A, Mirkowska P, Hartlova A, Van Eyk JE, Bourquin JP, Aebersold R, Boheler KR, Zandstra P, Wollscheid B. A mass spectrometric-derived cell surface protein atlas. *PLoS One*. 2015;10(3):e0121314. Epub 2015/04/22. doi: 10.1371/journal.pone.0121314. [PubMed: 25894527]
40. Zhu Y, Sullivan LL, Nair SS, Williams CC, Pandey AK, Marrero L, Vadlamudi RK, Jones FE. Coregulation of estrogen receptor by ERBB4/HER4 establishes a growth-promoting autocrine signal in breast tumor cells. *Cancer Res*. 2006;66(16):7991–8. Epub 2006/08/17. doi: 10.1158/0008-5472.Can-05-4397. [PubMed: 16912174]
41. Tiong KH, Mah LY, Leong CO. Functional roles of fibroblast growth factor receptors (FGFRs) signaling in human cancers. *Apoptosis*. 2013;18(12):1447–68. Epub 2013/08/01. doi: 10.1007/s10495-013-0886-7. [PubMed: 23900974]
42. Brewer JR, Mazot P, Soriano P. Genetic insights into the mechanisms of Fgf signaling. *Genes Dev*. 2016;30(7):751–71. Epub 2016/04/03. doi: 10.1101/gad.277137.115. [PubMed: 27036966]
43. Odawara H, Iwasaki T, Horiguchi J, Rokutanda N, Hirooka K, Miyazaki W, Koibuchi Y, Shimokawa N, Iino Y, Takeyoshi I, Koibuchi N. Activation of aromatase expression by retinoic

acid receptor-related orphan receptor (ROR) alpha in breast cancer cells: identification of a novel ROR response element. *J Biol Chem*. 2009;284(26):17711–9. Epub 2009/05/15. doi: 10.1074/jbc.M109.009241. [PubMed: 19439415]

44. Liu Z, Lou H, Xie K, Wang H, Chen N, Aparicio OM, Zhang MQ, Jiang R, Chen T. Reconstructing cell cycle pseudo time-series via single-cell transcriptome data. *Nat Commun*. 2017;8(1):22. Epub 2017/06/21. doi: 10.1038/s41467-017-00039-z. [PubMed: 28630425]
45. Hortobagyi GN, Stemmer SM, Burris HA, Yap YS, Sonke GS, Paluch-Shimon S, Campone M, Petrakova K, Blackwell KL, Winer EP, Janni W, Verma S, Conte P, Arteaga CL, Cameron DA, Mondal S, Su F, Miller M, Elmeliegy M, Germa C, O'Shaughnessy J. Updated results from MONALEESA-2, a phase III trial of first-line ribociclib plus letrozole versus placebo plus letrozole in hormone receptor-positive, HER2-negative advanced breast cancer. *Ann Oncol*. 2018;29(7):1541–7. Epub 2018/05/03. doi: 10.1093/annonc/mdy155. [PubMed: 29718092]
46. Finn RS, Martin M, Rugo HS, Jones S, Im SA, Gelmon K, Harbeck N, Lipatov ON, Walshe JM, Moulder S, Gauthier E, Lu DR, Randolph S, Diéras V, Slamon DJ. Palbociclib and Letrozole in Advanced Breast Cancer. *N Engl J Med*. 2016;375(20):1925–36. Epub 2016/12/14. doi: 10.1056/NEJMoa1607303. [PubMed: 27959613]
47. Goetz MP, Toi M, Campone M, Sohn J, Paluch-Shimon S, Huober J, Park IH, Trédan O, Chen SC, Manso L, Freedman OC, Garnica Jaliffe G, Forrester T, Frenzel M, Barriga S, Smith IC, Bourayou N, Di Leo A. MONARCH 3: Abemaciclib As Initial Therapy for Advanced Breast Cancer. *J Clin Oncol*. 2017;35(32):3638–46. Epub 2017/10/03. doi: 10.1200/jco.2017.75.6155. [PubMed: 28968163]
48. Ichijo H, Nishida E, Irie K, ten Dijke P, Saitoh M, Moriguchi T, Takagi M, Matsumoto K, Miyazono K, Gotoh Y. Induction of apoptosis by ASK1, a mammalian MAPKKK that activates SAPK/JNK and p38 signaling pathways. *Science*. 1997;275(5296):90–4. Epub 1997/01/03. doi: 10.1126/science.275.5296.90. [PubMed: 8974401]
49. Tobiume K, Matsuzawa A, Takahashi T, Nishitoh H, Morita K, Takeda K, Minowa O, Miyazono K, Noda T, Ichijo H. ASK1 is required for sustained activations of JNK/p38 MAP kinases and apoptosis. *EMBO Rep*. 2001;2(3):222–8. Epub 2001/03/27. doi: 10.1093/embo-reports/kve046. [PubMed: 11266364]
50. Piccart-Gebhart M, Holmes E, Baselga J, de Azambuja E, Dueck AC, Viale G, Zujewski JA, Goldhirsch A, Armour A, Pritchard KI, McCullough AE, Dolci S, McFadden E, Holmes AP, Tonghua L, Eidtmann H, Dinh P, Di Cosimo S, Harbeck N, Tjulandin S, Im YH, Huang CS, Diéras V, Hillman DW, Wolff AC, Jackisch C, Lang I, Untch M, Smith I, Boyle F, Xu B, Gomez H, Suter T, Gelber RD, Perez EA. Adjuvant Lapatinib and Trastuzumab for Early Human Epidermal Growth Factor Receptor 2-Positive Breast Cancer: Results From the Randomized Phase III Adjuvant Lapatinib and/or Trastuzumab Treatment Optimization Trial. *J Clin Oncol*. 2016;34(10):1034–42. Epub 2015/11/26. doi: 10.1200/jco.2015.62.1797. online at www.jco.org. Author contributions are found at the end of this article. [PubMed: 26598744]
51. von Minckwitz G, Procter M, de Azambuja E, Zardavas D, Benyunes M, Viale G, Suter T, Arahmani A, Rouchet N, Clark E, Knott A, Lang I, Levy C, Yardley DA, Bines J, Gelber RD, Piccart M, Baselga J. Adjuvant Pertuzumab and Trastuzumab in Early HER2-Positive Breast Cancer. *N Engl J Med*. 2017;377(2):122–31. Epub 2017/06/06. doi: 10.1056/NEJMoa1703643. [PubMed: 28581356]
52. Ruíz-Borrego M, Guerrero-Zotano A, Bermejo B, Ramos M, Cruz J, Baena-Cañada JM, Cirauqui B, Rodríguez-Lescure Á, Alba E, Martínez-Jáñez N, Muñoz M, Antolín S, Álvarez I, Del Barco S, Sevillano E, Chacón JI, Antón A, Escudero MJ, Ruiz V, Carrasco E, Martín M. Phase III evaluating the addition of fulvestrant (F) to anastrozole (A) as adjuvant therapy in postmenopausal women with hormone receptor-positive HER2-negative (HR+/HER2-) early breast cancer (EBC): results from the GEICAM/2006–10 study. *Breast Cancer Res Treat*. 2019;177(1):115–25. Epub 2019/06/04. doi: 10.1007/s10549-019-05296-8. [PubMed: 31152327]
53. Johnston S, Puhalla S, Wheatley D, Ring A, Barry P, Holcombe C, Boileau JF, Provencher L, Robidoux A, Rimawi M, McIntosh SA, Shalaby I, Stein RC, Thirlwell M, Dolling D, Morden J, Snowdon C, Perry S, Cornman C, Batten LM, Jeffs LK, Dodson A, Martins V, Modi A, Osborne CK, Pogue-Geile KL, Cheang MCU, Wolmark N, Julian TB, Fisher K, MacKenzie M, Wilcox M, Huang Bartlett C, Koehler M, Dowsett M, Bliss JM, Jacobs SA. Randomized Phase II Study Evaluating Palbociclib in Addition to Letrozole as Neoadjuvant Therapy in Estrogen

- Receptor-Positive Early Breast Cancer: PALLET Trial. *J Clin Oncol*. 2019;37(3):178–89. Epub 2018/12/14. doi: 10.1200/jco.18.01624. [PubMed: 30523750]
54. La Marca JE, Richardson HE. Two-Faced: Roles of JNK Signalling During Tumourigenesis in the Drosophila Model. *Frontiers in Cell and Developmental Biology*. 2020;8(42). doi: 10.3389/fcell.2020.00042.
 55. Pinal N, Calleja M, Morata G. Pro-apoptotic and pro-proliferation functions of the JNK pathway of Drosophila: roles in cell competition, tumorigenesis and regeneration. *Open Biology*. 2019;9(3):180256. doi: doi:10.1098/rsob.180256. [PubMed: 30836847]
 56. Colleoni B, Paternot S, Pita JM, Bisteau X, Coulonval K, Davis RJ, Raspé E, Roger PP. JNKs function as CDK4-activating kinases by phosphorylating CDK4 and p21. *Oncogene*. 2017;36(30):4349–61. Epub 2017/04/04. doi: 10.1038/onc.2017.7. [PubMed: 28368408]
 57. Zhang JY, Tao S, Kimmel R, Khavari PA. CDK4 regulation by TNFR1 and JNK is required for NF- κ B-mediated epidermal growth control. *Journal of Cell Biology*. 2005;168(4):561–6. doi: 10.1083/jcb.200411060.
 58. Hafner M, Mills CE, Subramanian K, Chen C, Chung M, Boswell SA, Everley RA, Liu C, Walmsley CS, Juric D, Sorger PK. Multiomics Profiling Establishes the Polypharmacology of FDA-Approved CDK4/6 Inhibitors and the Potential for Differential Clinical Activity. *Cell Chem Biol*. 2019;26(8):1067–80.e8. Epub 2019/06/11. doi: 10.1016/j.chembiol.2019.05.005. [PubMed: 31178407]
 59. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, Griffith C, Boeddinghaus I, Salter J, Detre S, Hills M, Ashley S, Francis S, Walsh G. Short-term changes in Ki-67 during neoadjuvant treatment of primary breast cancer with anastrozole or tamoxifen alone or combined correlate with recurrence-free survival. *Clin Cancer Res*. 2005;11(2 Pt 2):951s–8s. Epub 2005/02/11. [PubMed: 15701892]
 60. Chen X, Chang JT. Planning bioinformatics workflows using an expert system. *Bioinformatics*. 2017;33(8):1210–5. Epub 2017/01/06. doi: 10.1093/bioinformatics/btw817. [PubMed: 28052928]
 61. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20. Epub 2014/04/04. doi: 10.1093/bioinformatics/btu170. [PubMed: 24695404]
 62. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754–60. Epub 2009/05/20. doi: 10.1093/bioinformatics/btp324. [PubMed: 19451168]
 63. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26(5):589–95. Epub 2010/01/19. doi: 10.1093/bioinformatics/btp698. [PubMed: 20080505]
 64. Li H Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:13033997. 2013.
 65. Picard Toolkit: Broad Institute, GitHub Repository; 2019 [cited 2020 September 18]. Available from: <http://broadinstitute.github.io/picard/>.
 66. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297–303. Epub 2010/07/21. doi: 10.1101/gr.107524.110. [PubMed: 20644199]
 67. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Källberg M, Chen X, Kim Y, Beyter D, Krusche P, Saunders CT. Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods*. 2018;15(8):591–4. Epub 2018/07/18. doi: 10.1038/s41592-018-0051-x. [PubMed: 30013048]
 68. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 2012;22(3):568–76. Epub 2012/02/04. doi: 10.1101/gr.129684.111. [PubMed: 22300766]
 69. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164. Epub 2010/07/06. doi: 10.1093/nar/gkq603. [PubMed: 20601685]

70. Shen R, Seshan VE. FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res.* 2016;44(16):e131. Epub 2016/06/09. doi: 10.1093/nar/gkw520. [PubMed: 27270079]
71. Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, Li Q, Szallasi Z, Eklund AC. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann Oncol.* 2015;26(1):64–70. Epub 2014/10/17. doi: 10.1093/annonc/mdu479. [PubMed: 25319062]
72. Dang HX, White BS, Foltz SM, Miller CA, Luo J, Fields RC, Maher CA. ClonEvol: clonal ordering and visualization in cancer sequencing. *Ann Oncol.* 2017;28(12):3076–82. Epub 2017/09/28. doi: 10.1093/annonc/mdx517. [PubMed: 28950321]
73. Sei E, Bai S, Navin N. Dissociation of Nuclear Suspensions from Human Breast Tissues 2018 [cited 2021 March 17]. Available from: 10.17504/protocols.io.t3aeqie.
74. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923–30. Epub 2013/11/15. doi: 10.1093/bioinformatics/btt656. [PubMed: 24227677]
75. Müllner D fastcluster: Fast Hierarchical, Agglomerative Clustering Routines for R and Python. *Journal of Statistical Software.* 2013;53(no 9):1–18.
76. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* 2013;14:7. Epub 2013/01/18. doi: 10.1186/1471-2105-14-7. [PubMed: 23323831]
77. Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MH, White JS. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol.* 2009;24(3):127–35. Epub 2009/02/03. doi: 10.1016/j.tree.2008.10.008. [PubMed: 19185386]
78. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. 2017. 2017;82(13):26. Epub 2017-11-29. doi: 10.18637/jss.v082.i13.
79. Pedersen EJ, Miller DL, Simpson GL, Ross N. Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ.* 2019;7:e6876. Epub 2019/06/11. doi: 10.7717/peerj.6876. [PubMed: 31179172]
80. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, Bhai J, Billis K, Boddu S, Marugán JC, Cummins C, Davidson C, Dodiya K, Fatima R, Gall A, Giron CG, Gil L, Grego T, Haggerty L, Haskell E, Hourlier T, Izuogu OG, Janacek SH, Juettemann T, Kay M, Lavidas I, Le T, Lemos D, Martinez JG, Maurel T, McDowall M, McMahon A, Mohanan S, Moore B, Nuhn M, Oheh DN, Parker A, Parton A, Patricio M, Sakthivel MP, Abdul Salam AI, Schmitt BM, Schuilenburg H, Sheppard D, Sycheva M, Szuba M, Taylor K, Thormann A, Threadgold G, Vullo A, Walts B, Winterbottom A, Zadissa A, Chakiachvili M, Flint B, Frankish A, Hunt SE, G II, Kostadima M, Langridge N, Loveland JE, Martin FJ, Morales J, Mudge JM, Muffato M, Perry E, Ruffier M, Trevanion SJ, Cunningham F, Howe KL, Zerbino DR, Flicek P. Ensembl 2020. *Nucleic Acids Res.* 2020;48(D1):D682–d8. Epub 2019/11/07. doi: 10.1093/nar/gkz966. [PubMed: 31691826]
81. Hahsler M, Hornik K. TSP—Infrastructure for the Traveling Salesperson Problem. 2007. 2007;23(2):21. Epub 2007-11-21. doi: 10.18637/jss.v023.i02.

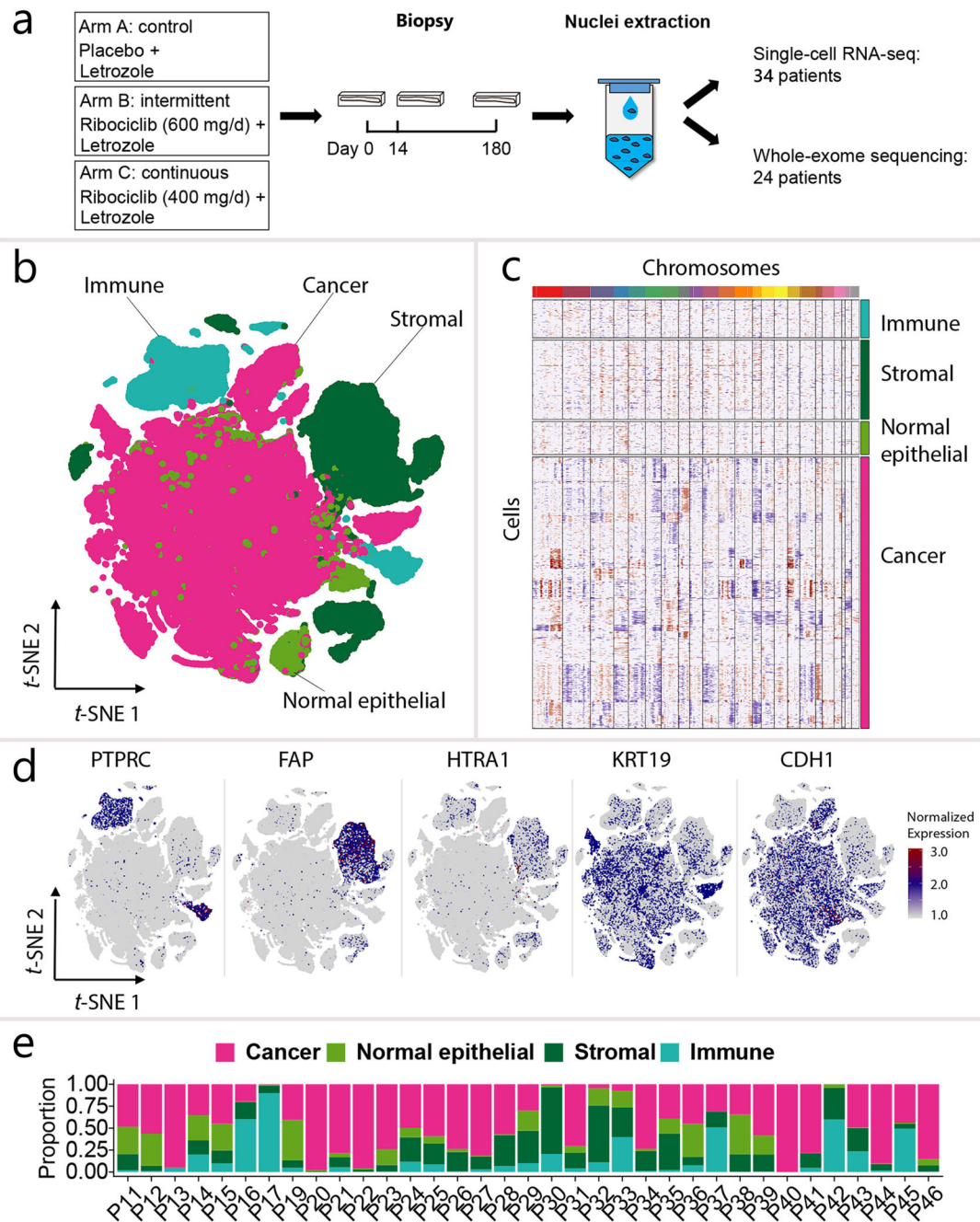


Figure 1. Landscape of tumor and macroenvironment of early-stage ER+ breast cancer patients in FELINE trial.

a, Schematic diagram of single-cell RNA-seq workflow. The data of 34 patients generated using the 10x genomics platform are shown in Fig.1 b–d. Number of cancer cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7. The data of 10 patients generated using ICELL8 platform are shown in Extended Data Fig.2. **b**, Distinction of different cell types is shown by a *t*-SNE plot of single cells of 34 ER+ breast cancer patients

(color=cell type). **c**, Gene copy number profile in cancer cells and neighboring normal cells. Blue color indicates copy number loss and red color indicates copy number gain. **d**, Expression of marker genes of cancer cells and normal epithelial cells (KRT19, CDH1), stromal cells (FAP, HTRA1), and immune cells (PTPRC). **e**, Proportion of cancer cells and neighboring normal cells in each patient.

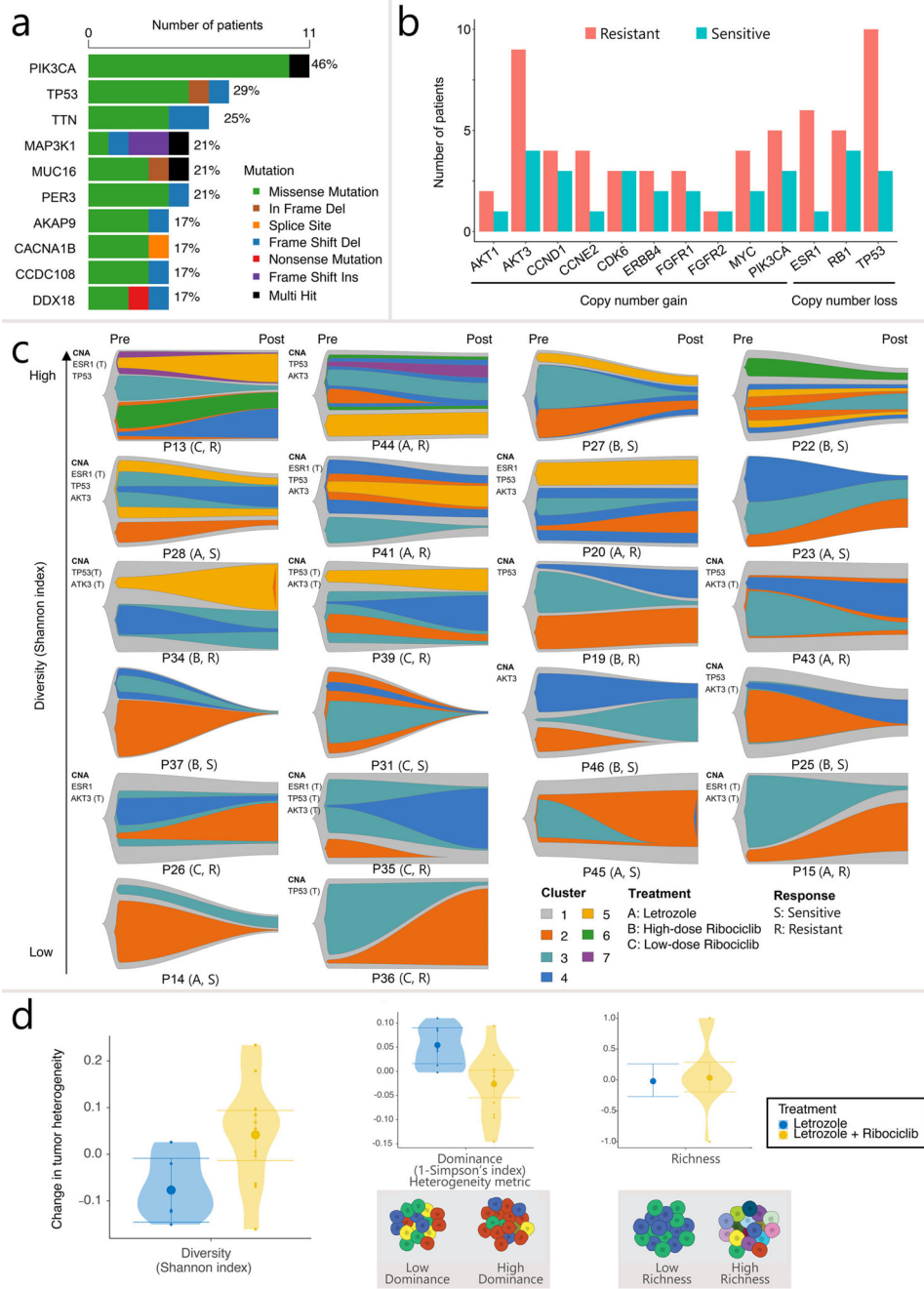


Figure 2. Evolution of genomic mutations in response to endocrine or combination therapy. **a**, Frequently mutated genes. Cancer driver genes were ranked based on number of patients carrying large impact somatic mutations. Mutations were counted if they are present in one or both biopsies. The top 10 frequently mutated genes are shown. **b**, Copy number alteration of key resistant genes in sensitive and resistant tumors. Gene copy number alterations were counted for a patient if one or both biopsies carry the alteration. **c**, Clonal evolution in response to endocrine therapy or combination therapy. Patient tumors were ranked by initial diversity (Shannon index at Day 0). The height of fishplot scaled to reflect changes of tumor size during treatment. Two biopsies (pre- and post-treatment) were sequenced and

used to perform analyses of clonal evolution for each tumor. Day 180 biopsies were used as post-treatment samples except P36, P44, P45, and P46. Only Day 14 biopsies were available for sequencing for these four patients. Copy number alteration (CNA) of ESR1 loss, TP53 loss, and AKT3 gain are labelled for each patient when present at one or both biopsies. A “T” in parentheses indicates CNA is truncal in this tumor defined based on FACETS copy number analysis using WES and inferCNV copy number profiles using scRNAseq. Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7. **d**, Change of tumor heterogeneity in response to endocrine therapy or combination therapy, as measured by Shannon index of overall diversity, and broken down in to the diversity components of subclonal dominance and richness. Violin curves show the between patient variability in the change of heterogeneity during the trial (small points show observed changes in diversity). Hierarchical regression models identified the average change in tumor heterogeneity (large points) during endocrine (blue) or combination therapy (yellow) (uncertainty quantified by 95% confidence interval error bars). Schematic diagrams show the distinction between differences in dominance and richness.

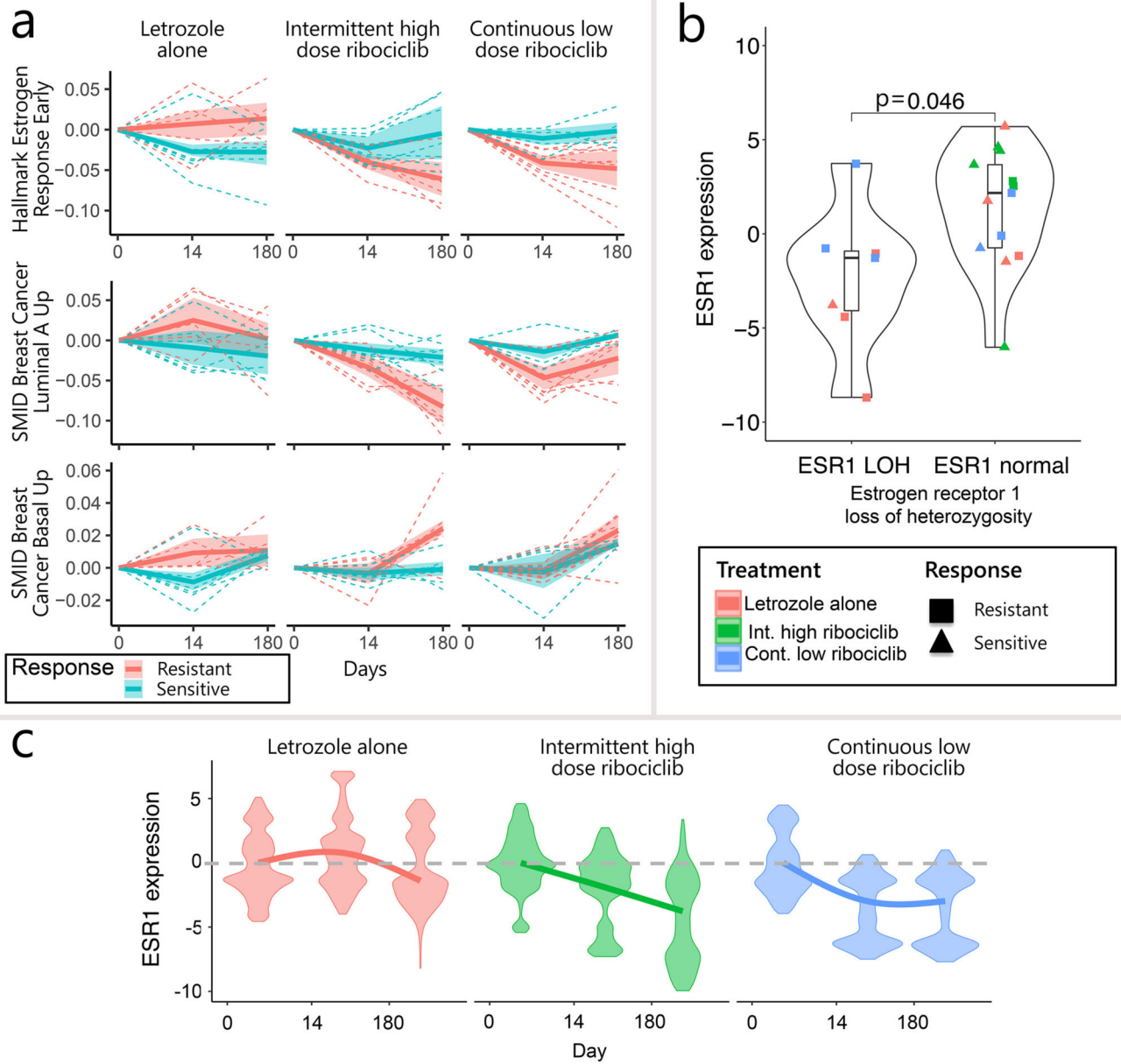


Figure 3. Accelerated evolution of estrogen independence during combination therapy.

a, Changes in single cell estrogen receptor activity and signatures of estrogen dependent (luminal) and estrogen independent (basal) phenotypes during treatment (columns), in tumors resistant (persistent=red) or sensitive (shrinking=blue) to therapy. Pathway trends across tumors were determined using hierarchical regression (solid lines). Tumor's pathway trajectories are shown by dashed lines and confidence intervals of model estimates shown by shaded regions. Pathway activity measured using the ssGSEA signatures (Hallmark estrogen response early, Smid breast cancer luminal A up and Smid breast cancer basal up respectively). Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11,

S=4, R=7. **b**, Loss of heterozygosity (LOH) of the estrogen receptor gene (ESR1) associated with reduced average expression in pre-treatment biopsies (two-sided Mann-Whitney test). Violin curves show the distribution of average ESR1 expression in tumors with or without LOH. Points indicate patient specific averages (color indicates treatment and shape signifies tumor response). Box and whisker plots indicate the median and upper/lower quantiles of patient tumor expression, with whiskers signifying the data range. **c**, Reduction in ESR1 expression accelerated under combination treatments, compared to endocrine therapy (columns: letrozole vs ribociclib treatments). Violin curves show the distribution of single cell expression during each treatment (color). Expression normalized relative to the baseline average (grey dashed line). Hierarchical generalized additive models predict the changes in expression during treatment, accounting for initial patient specific difference (colored curves). Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7.

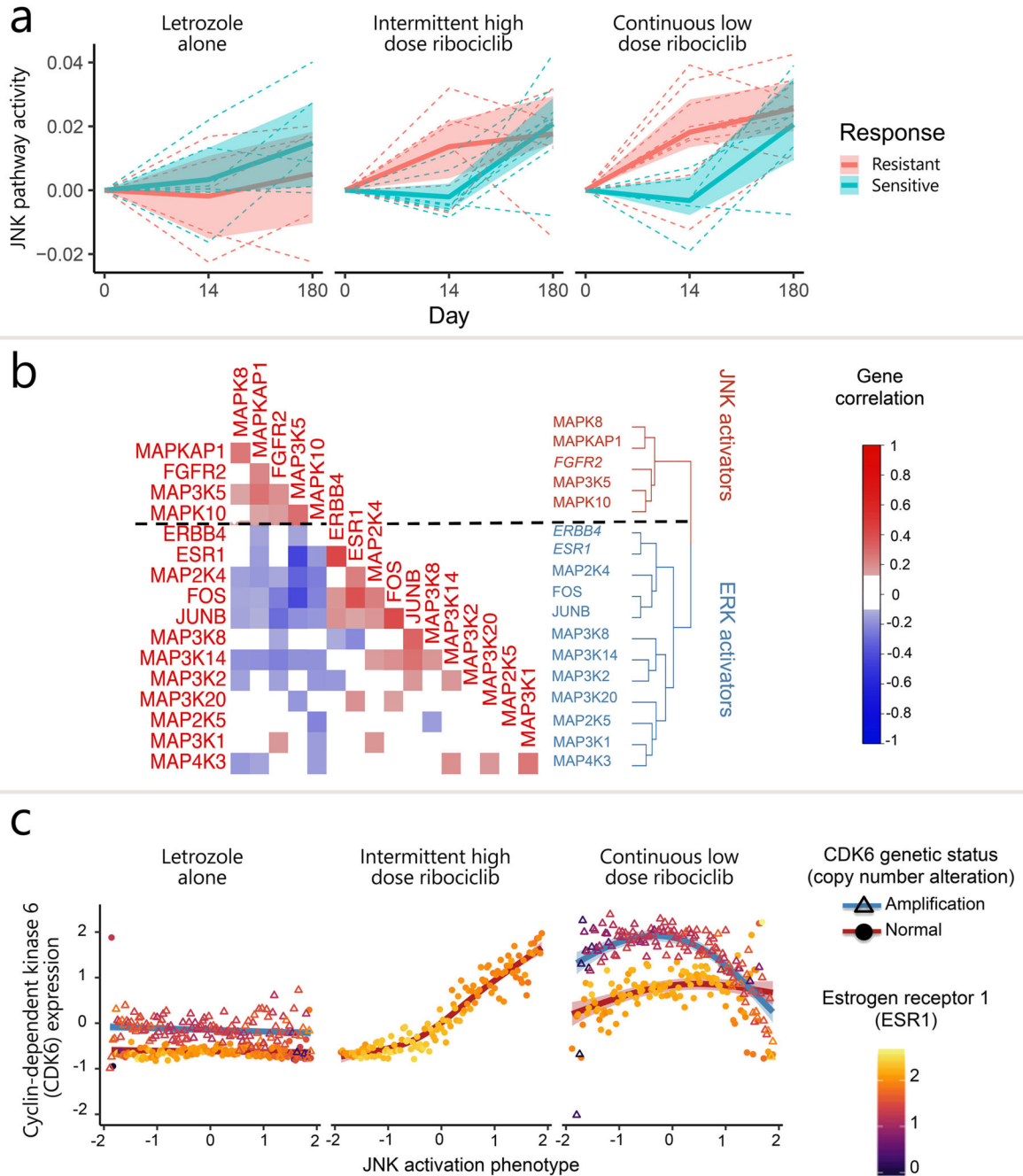


Figure 4. JNK pathway activation occurs during the emergence of combination therapy resistance and is associated with estrogen independence and increased CDK6 expression.
a, JNK pathway activity increases during treatments (columns) in resistant tumors (red) compared to sensitive tumors (blue) at day 14, and in cancer cells of most tumors by end of treatment (day 180). Pathway trends determined using hierarchical regression (solid lines). Tumor’s pathway trajectories are shown (dashed lines) along with confidence intervals of model estimates (shaded regions) (JNK ssGSEA pathway=St JNK MAPK). Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12,

S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7. **b**, Heatmap of correlation between MAPK gene expression, showing the dichotomy between JNK and ERK activating genes across treatments (ESR1, ERBB4 and FGFR2 receptors added to indicate their relationship to MAPK genes). Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7. **c**, End of trial expression of CDK6 (ribociclib target gene) in subclonal tumor populations with differing levels of JNK activation (phenotype integrates across MAPK gene expression) and for patients with CDK6 genetic amplification (triangles and blue curves) or normal copy number (circles and red curves). Generalized additive models describe the relationship between JNK signaling activity and CDK6 expression at end of therapy (curves), with shaded regions indicating model confidence bands. Average estrogen receptor (ESR1) expression is shown for each subclonal population with differing JNK activation (point color gradient). Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7.

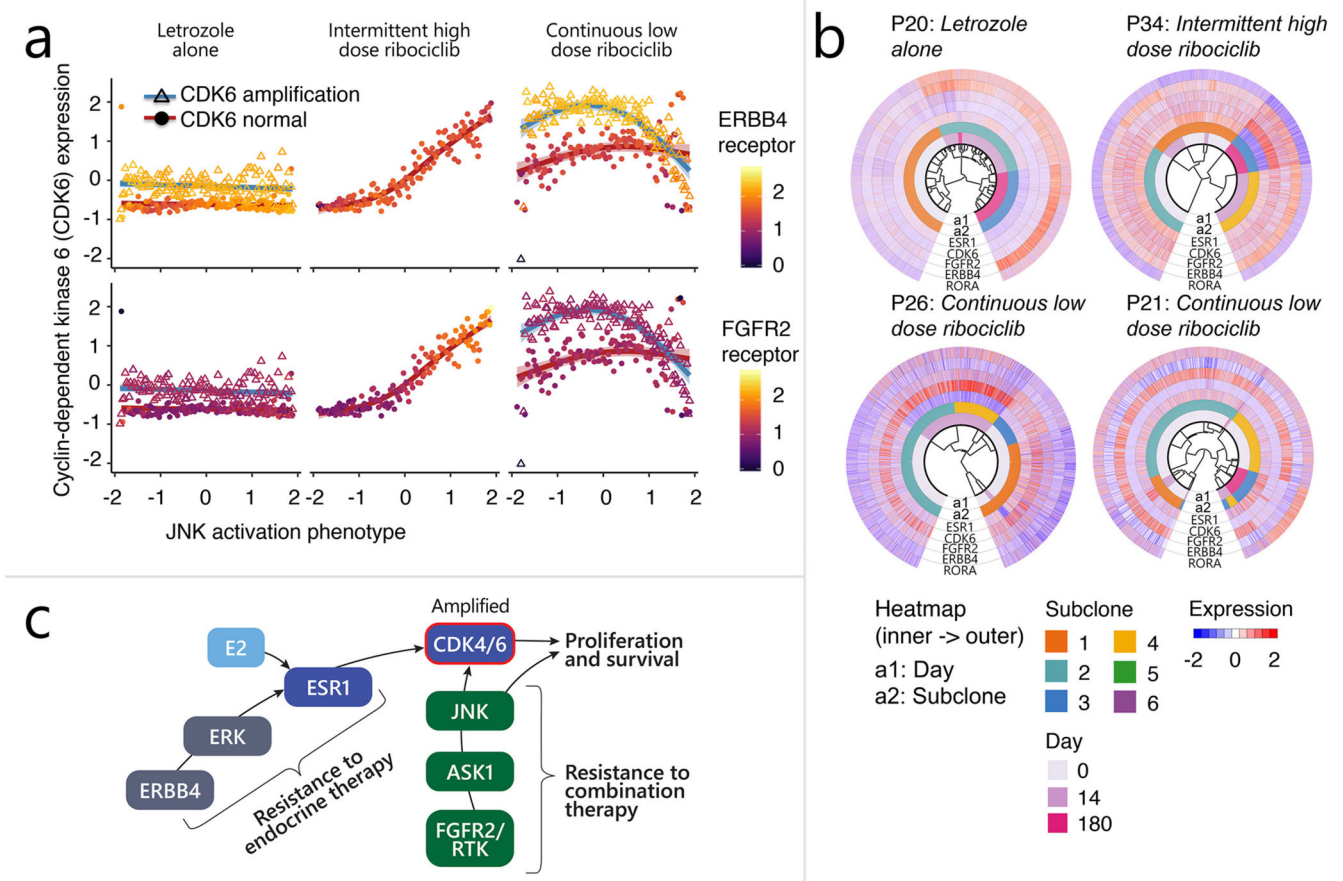


Figure 5. Activation of ERBB4 and FGFR2 as resistance mechanisms to endocrine and combination therapy.

a, Average end of trial growth factor receptor expression (point color gradient) of ERBB4 (top) and FGFR2 (bottom) is shown for subclonal populations with differing JNK activation (x-axis) and for tumors with genetic amplification (triangles and blue curves) or normal copy number (circles and red curves). Cyclin-dependent kinase 6 (CDK6; ribociclib target gene) expression is shown for each population and generalized additive models (curves) describe the relationship between JNK signaling activity and CDK6 expression at end of therapy, (shaded regions indicate model confidence bands). Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7. **b**, Hierarchical clustering of tumors showing dysregulated ESR1 as well as upregulated ERBB4 and/or FGFR2. For each patient’s tumor biopsies, a hierarchical clustering tree is shown at the center of circos plots. Cell annotation (timepoint and subclone) as well as expression of key resistance genes (ESR1, CDK6, FGFR2, ERBB4, RORA) were shown as heatmap. **c**, Schematic diagram showing resistance mechanisms driven by upregulation of ERBB4 and CDK6 amplification (red circle signifies amplification) or alternative signaling via FGFR2/RTK’s and JNK signal transduction.

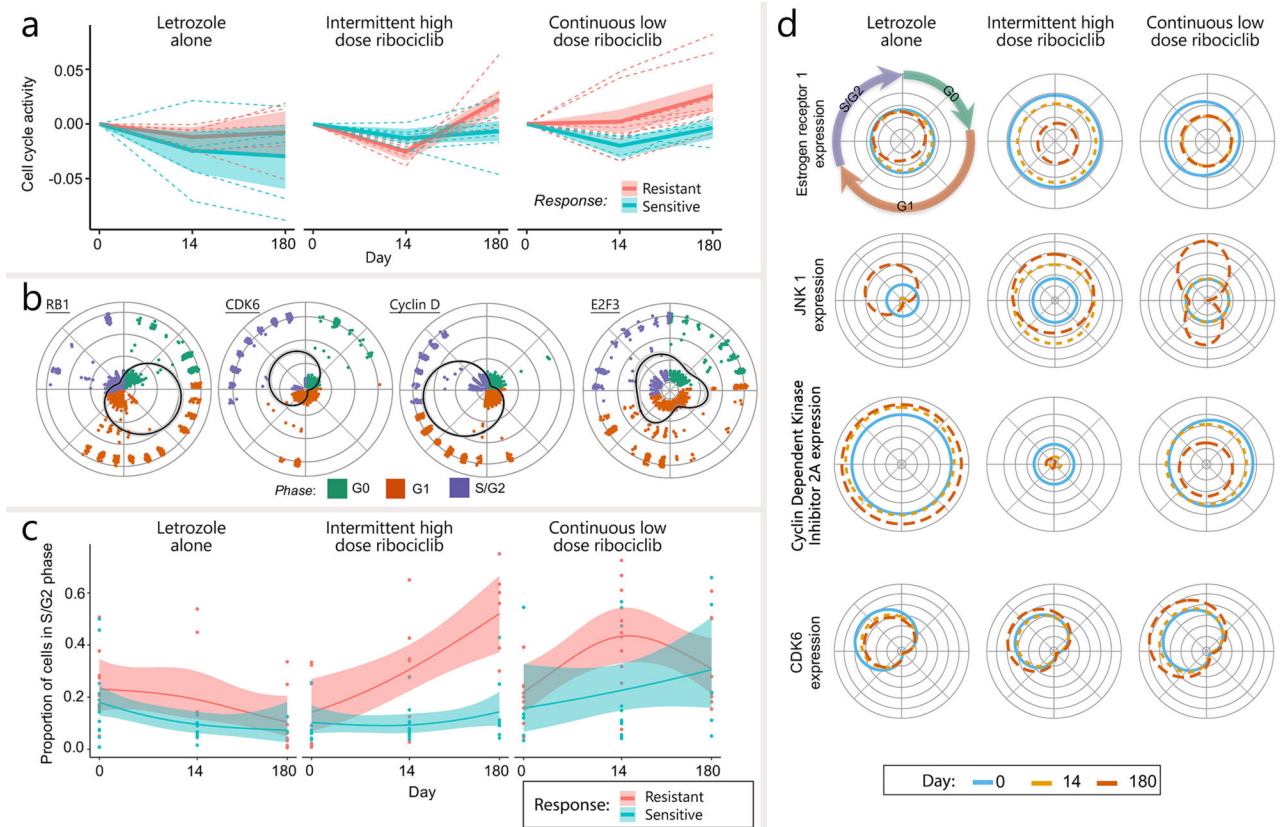


Figure 6. Cell cycle reactivates during combination therapy follows the loss of estrogen receptor expression, activation of JNK1, repression of the cell cycle inhibitor cyclin-dependent kinase 2A and upregulation of CDK6 during the G1 checkpoint phase.

a, Cell cycle activity of resistant (red) and sensitive (blue) tumors during treatment (columns= regimens). Trend in cell cycle activity, measured by the ssGSEA biocarta cell cycle pathway, are determined by hierarchical regression (solid lines). Tumor specific trajectories are shown (dashed lines) along with confidence intervals of model estimates (shaded regions). Number of cells (n) and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7. **b**, Visualization of the pseudotime cell cycle reconstruction obtained using the Markov model-based reCAT algorithm. The dynamics of key cell cycle gene expression across stages of the cell cycle (black lines) were recovered from cell specific gene expression (points), using cyclical generalized additive models. The recovered fluctuations of cell cycle gene expression were used to identify three distinct cell cycle phases (colors; G0, G1 and S/G2), using a Gaussian mixture model. Cell cycle stages (clusters of cells) are colored by cell cycle phase and the distance of the point from the origin signifies the cells expression in that stage. Cell cycle orientation is consistent and comparable across circular plots. **c**, The proportion of S/G2 phase cells (passed the G1 checkpoint) in samples from resistant and sensitive tumors (color) during each treatment (column). Logistic generalized additive models describe trends in S/G2 phase cell frequency over time (solid lines) and heterogeneity across tumors (shaded regions signify confidence bands). Number of cells (n)

and patients (P) with sensitive (S) versus resistant (R) tumors by arm = Letrozole alone: n= 46986, P=11, S=6, R=5; Intermittent high dose ribociclib: n= 27790, P=12, S=6, R=6; Continuous low dose ribociclib: n= 34543, P=11, S=4, R=7. **d**, Changes in ESR1, JNK1, CDKN2A expression around the cell cycle and during treatment (columns). Colored lines show the expected gene expression of cells throughout the cell cycle prior (blue), during (orange) and after treatment (red). The distance from the center of the circle shows gene expression at point in the cell cycle.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript