

RESEARCH

Open Access



ESR1 mutations and CDK4/6 inhibitor choice shape clonal selection and adaptive cell states during acquired resistance

Cristina Guarducci^{1,2,3†}, Daniel Abravanel^{1†}, Douglas Russo^{1,4}, Kavya Prasad⁵, Jingxin Fu¹, Zsuzsanna Nagy^{1,2}, Vishwajit Rao^{1,2}, Agostina Nardone^{1,2}, Yanan Kuang¹, Avery Feit^{1,2}, Wen Ma^{1,2}, Gabriella Cohen Feit^{1,2}, Francisco Hermida-Prado^{1,2}, Capucine Heraud^{1,2}, Miguel Munoz Gomez^{1,2}, Diana Elisa Garcia Cortes¹, Patrick Kurnia¹, Jorge Gomez Tejeda Zanutto^{1,5}, Rong Li², Xintao Qiu², Simona Cristea^{4,5}, Ariel Feiglin⁶, Yu-Chen Cheng^{7,8}, Nancy U. Lin¹, Sara M. Tolaney¹, Luca Malorni^{3,9}, Piotr Sicinski^{10,11,12}, Kornelia Polyak^{1,13}, Cloud Paweletz¹, Henry Long^{1,2}, Chip Stewart⁵, Franziska Michor^{4,5,7,8,14,15}, Gad Getz^{5,16}, Myles Brown^{1,2} and Rinath Jeselsohn^{1,2,5*}

Abstract

Background In estrogen receptor-positive (ER+) breast cancer, CDK4/6 inhibitors (CDK4/6is) combined with endocrine therapy (ET) are standard first-line treatment for metastatic disease. However, most patients eventually develop resistance. Activating *ESR1* mutations are a prevalent mechanism of acquired resistance to ET and are enriched after ET plus a CDK4/6 inhibitor (CDK4/6i), but their role in the clonal evolution and adaptive mechanisms of acquired resistance to CDK4/6 inhibition, independent of ET, is unknown. In addition, whether different CDK4/6is impose distinct selective pressures and divergent resistance states remains elusive.

Methods To investigate the clonal dynamics, cell states and cellular plasticity during acquired CDK4/6i resistance in mutant versus wild-type (WT) *ESR1*, we performed high-complexity DNA barcoding (ClonTracer library) with longitudinal sampling and multi-omic profiling in an isogenic MCF7 model expressing WT ER or Y537S mutant ER. We also evaluated the clonality of the *ESR1* mutations in clinical samples with CDK4/6i resistance.

Results We showed that *ESR1* mutations are enriched in clinical tumors with acquired resistance to CDK4/6is, and in paired biopsies expanded to near clonality after treatment. We demonstrated progressive clonal selection with both divergent and partially convergent evolutionary trajectories. The *ESR1* mutation substantially reshapes clonal and epigenetic evolution during palbociclib resistance but had a weaker impact under abemaciclib selection. Overall, clonal evolution and cell states in palbociclib and abemaciclib resistance were distinct. Single-cell RNA-seq revealed transcriptional heterogeneity highlighting cellular plasticity during passaging of cells and selection. Finally, in vivo barcoding of mammary xenograft, local recurrences, and distant metastases demonstrated site-specific

[†]Cristina Guarducci and Daniel Abravanel contributed equally to this work.

*Correspondence:
Rinath Jeselsohn
Rinath_jeselsohn@dfci.harvard.edu

Full list of author information is available at the end of the article



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

clonal outgrowth in mutant ER metastases, and partial overlap between metastatic and CDK4/6i-resistant subclones, supporting the dual role of specific populations in therapeutic resistance and metastatic colonization.

Conclusions High-resolution lineage tracing and multi-omic studies demonstrate that CDK4/6i resistance is shaped by clonal selection and adaptive remodeling of cell states, with the *ESR1* mutation status and the specific inhibitor acting as key determinants of evolutionary trajectories. These findings suggest that both variables should be considered when designing sequential and combination treatment strategies to overcome CDK4/6i resistance.

Keywords Estrogen receptor-positive breast cancer, Mutant estrogen receptor, CDK4/6 inhibitor resistance, Clonal dynamics, Cell plasticity

Background

Endocrine therapy (ET) in combination with a CDK4/6 inhibitor (CDK4/6i) is the cornerstone treatment in estrogen receptor positive (ER+) breast cancer [1]. These therapies converge on the cyclin D1/CDK4/6/Rb axis, blocking the G1-S cell cycle transition and inducing cell cycle arrest, which accounts, at least in part, for their synergistic effects [2–4]. Despite the substantial benefits from this treatment combination, nearly all patients with metastatic disease will ultimately develop resistance. Genetic alterations that are direct drivers of acquired resistance to CDK4/6is, such as *RBI* mutations, occur in only a small proportion of patients [5, 6]. Other recurrent alterations that indirectly drive resistance were also reported, including mutations in *FAT1* [7], *PIK3CA* [8] and *FGFR1* amplifications [9]. In addition, epigenetic remodeling with consequent upregulation of resistance-associated signaling pathways [10] is likely to represent another mechanism of CDK4/6i resistance.

The *ESR1* ligand binding domain (LBD) mutations, which confer constitutive activity and resistance to ET [11, 12], were not shown to be associated with primary (intrinsic) resistance to CDK4/6i-ET combinations [13–15]. However, these mutations, particularly the Y537S *ESR1* mutation, that exhibits the most robust ligand independent and antagonistic resistance phenotypes among the *ESR1* LBD mutations [16], were enriched following the acquisition of resistance to the combination of the CDK4/6i palbociclib and ET [6]. Moreover, clinical trials showed that in patients treated with ET plus a CDK4/6i, who had evidence of an emerging *ESR1* mutation, better targeting of the *ESR1* mutations by switching from an aromatase inhibitor to a selective estrogen receptor degrader (SERD) (fulvestrant or camizestrant) prolonged treatment efficacy [17, 18]. These clinical observations suggest that the *ESR1* mutations have a role in acquired resistance to CDK4/6i-ET combinations. However, their mechanistic role in resistance to the CDK4/6i component isolated from ET remains unresolved. It is unknown whether these mutations directly alter clonal selection, reshape chromatin landscapes, or influence the adaptive cellular states that permit survival under CDK4/6 inhibition. Furthermore, whether different CDK4/6is

impose distinct selective pressures giving rise to divergent resistant states has not been investigated. This question is particularly relevant since multiple CDK4/6is are approved for clinical use, and prior studies have shown that abemaciclib has a broader target profile, with activity against CDKs beyond CDK4 and CDK6 [19].

In this study we investigated the impact of the Y537S *ESR1* mutation and specific CDK4/6is on clonal evolution, genetic diversity and cell state plasticity, and how they shape the fitness landscape of acquired resistance.

Methods

Clinical sample tissue collection, whole exome sequencing and RNA sequencing

Tissues were collected from patients treated at the Dana-Farber Cancer Institute as described in detail previously and the current cohort is an expansion of the previously published cohort [20]. Metastatic core biopsies were obtained from patients, and samples were immediately snapfrozen in optimal cutting temperature (OCT) and stored in -80°C . Archival formalin-fixed, paraffin embedded (FFPE) blocks of primary tumor samples were also obtained. A blood sample was obtained during treatment, and whole blood was stored at -80°C until DNA extraction from peripheral blood mononuclear cells (for germline DNA) was performed.

Tissue samples were collected from patients with ER+ metastatic breast cancer who were treated with a CDK4/6i, with or without ET, and provided metastatic tumor biopsies ($N=112$ biopsies, $N=89$ patients). Tumor biopsies were classified as CDK4/6i-sensitive if obtained prior to CDK4/6i treatment from patients who subsequently experienced disease progression after >6 months of therapy ($N=28$ pre-treatment biopsies), acquired resistant if obtained from patients who progressed following >6 months treatment with a CDK4/6i ($N=53$ post-treatment biopsies), and intrinsically resistant if obtained from patients who progressed within six months of treatment initiation ($N=31$ pre-treatment biopsies) (details in Additional file 1: Table S1) [21]. This cohort included 23 pre-treatment sensitive tumors matched with post-treatment acquired resistant tumors from the same patient.

Tumor samples from this cohort were subjected to whole exome sequencing (WES) and RNA sequencing (RNA-seq). Clinical annotations were generated from the electronic medical record under the supervision of a board-certified medical oncologist and a cancer registrar (Additional file 1: Table S1). All samples included in this study were voluntarily donated by patients who provided informed consent under an institutional review board (IRB)-approved protocol (Dana-Farber/Harvard Cancer Center no. 05–246), which includes permission for sample acquisition, clinical data abstraction, sample analysis and data sharing. Analysis of biospecimens at the Broad Institute was performed under Broad Institute protocol number 15–370B.

The intrinsic resistant and sensitive cohorts were limited to biopsies that were obtained within 1 year prior to starting treatment with a CDK4/6i. Intrinsic resistance was defined as tumors that developed resistance within 6 months of treatment with a CDK4/6i in the metastatic setting. Sensitive tumors were from patients that had disease stability or responsive disease for at least 6 months in the metastatic setting. Samples with acquired resistance included biopsies that were acquired within 30 days prior to progression and up to 1 year after disease progression. These were samples that were obtained from patients that had disease stability or responsive disease for at least 6 months in the metastatic setting and later developed progressive disease.

These categories were defined based on the 6th and 7th International consensus guidelines for the management of advanced breast cancer (ABC7 guidelines) [21]. The cohort included 23 matched pairs consisting of a pre-treatment, CDK4/6i-sensitive tumor biopsy and a post-treatment, acquired resistant biopsy from the same patient.

Transcriptional analyses were performed only on samples from patients with acquired resistance who had a biopsy within 30 days of disease progression while on treatment (excluding patients who underwent a biopsy after the CDK4/6i was discontinued). Samples with non-pathogenic *ESR1* mutation were excluded from the transcriptional analysis.

RNA-seq reads were mapped to the human genome (hg19) with STAR aligner [22] with default parameters. Transcriptome quality was assessed using RNA-SeQC 2 [23] and gene expression quantification was conducted using RSEM [24]. Samples with < 8,000 unique genes were removed from subsequent analyses. Gene expression was measured using upper quartile normalized log₂(TPM + 1) values of expressed genes and corrected for tissue type (Frozen/OCT vs. FFPE) with receptor status as a covariate using ComBat [25] in the *sva* package. Single sample gene set enrichment analysis was performed using the Hallmark gene sets [26, 27].

WES were captured using Illumina technology and the sequencing data processing and analysis was performed using the Picard and Terra pipelines at the Broad Institute. The Picard pipeline (<http://picard.sourceforge.net>) was used to produce a BAM file with aligned reads. This includes alignment to the GRCh37 human reference sequence using the BWA aligner [28] and estimation and recalibration of base quality score with the Genome Analysis Toolkit (GATK) [29]. A custom-made cancer genomics analysis pipeline was used to identify somatic alterations using the Terra platform (<https://app.terra.bio/>). The CGA WES Characterization pipeline developed at the Broad Institute was used to call, filter and annotate somatic mutations and copy number variation (CNV). GATK CNV was used for the generation of accurate relative copy number profiles from the WES data and reference/alternate read counts at heterozygous single nucleotide polymorphism (SNP) sites present in both the normal and tumor samples. After accurate proportional coverage profiles were generated for a sample, Allelic CapSeg tool [30] was used to generate a segmented allelic copy ratio profile. Allelic copy number profile and mutational call data were modeled jointly by ABSOLUTE [31] to produce purity for the samples, a discrete copy number profile, and compute cancer cell fractions (CCF). Tumor samples which had a purity of 8% or more were used for downstream analysis. The clonal structure and clonal evolution of *ESR1* mutations was inferred with PhylogicNDT [32].

Additional validation RNA-seq analyses were performed in a smaller sample set of ER+ and HER2-negative metastatic breast cancer biopsies with acquired resistance to palbociclib [33]. RNA-seq data for these patient samples are available in dbGaP (study accession phs001285.v2.p1).

Cell lines

MCF7 cells engineered to express a doxycycline (DOX)-inducible-hemagglutinin (HA)-tagged Y537S mutant ER [13] were grown in DMEM supplemented with 10% fetal bovine serum (FBS), 10 µg/ml penicillin–streptomycin–glutamine (PSG) and 500 µg/ml of geneticin (full medium), either without DOX (ER-WT MCF7 cells) or with 1 µg/ml DOX to induce the expression of the Y537S ER mutation (ER-mutant MCF7 cells). This model was extensively characterized in reference 13.

DOX-untreated engineered MCF7 cells were transduced with the ClonTracer library [34], a high-complexity 30 bp-long semi-random DNA barcode library which contains ~73 million unique barcodes (Addgene #69830). The packaging of the library into lentiviruses and the infection of the cells were performed following the manufacturer's instructions. Specifically, 20 million DOX-untreated MCF7 cells were infected with the

ClonTracer library. To enhance virus infection, cells were maintained in full medium supplemented with 8 µg/ml of polybrene for 48 h. After transduction, the selection of the barcoded cells was performed by maintaining them in full medium supplemented with 2 µg/ml of puromycin for 96 h. In order to ensure that each cell was infected with one barcode, the multiplicity of infection (MOI) was set at 0.1. The selected barcoded cells (parental cells) were expanded for 3 passages in full medium supplemented with 1 µg/ml of puromycin and the complexity of the library was assessed by barcode sequencing. Next, the parental cells were further expanded and divided to be maintained in full medium supplemented with 1 µg/ml of puromycin, with or without DOX.

To generate three evolutionary replicates of ER-WT and ER-mutant palbociclib-resistant and abemaciclib-resistant cells, DOX-treated and untreated MCF7 cells were cultured in triplicates over time and exposed to escalating doses of the drugs (50 nM, 100 nM, 200 nM, 350 nM, 500 nM, 750 nM and 1 µM) every three passages. Cells were defined as resistant when their ability to proliferate in the presence of palbociclib or abemaciclib 1 µM was restored. Palbociclib-resistant and abemaciclib-resistant cells were regularly maintained in full medium supplemented with 1 µg/ml of puromycin, with or without DOX, and with palbociclib or abemaciclib 1 µM, respectively. In parallel, as controls, ER-WT and ER-mutant MCF7 cells were cultured in vehicle (0.01% dimethyl sulfoxide, DMSO). To ensure equal barcode representation across replicates and minimize stochastic barcode losses over time, 30 million cells were plated at every passage. To assess the barcoding library distribution over time, resistant models were sequenced at three time points: an early time point (after 3 passages in full medium with palbociclib or abemaciclib at 50 nM), an intermediate time point (after 3 passages in full medium with palbociclib or abemaciclib at 200 nM), and at the time of resistance to palbociclib or abemaciclib. Control cells were sequenced at the same timepoints.

WT-ER was overexpressed in WT-ER PalboR MCF7 cells by lentiviral transduction using a pLVX-ESR1_HA_blast plasmid (an *ESR1* overexpression construct carrying a C-terminal 3×HA tag). This plasmid was custom made by GenScript using the pLVX-IRES-puro vector. The pLVX-ESR1_HA_blast plasmid was packaged into viral particles in HEK293T cells using the pMD2.G (Addgene #12259) and psPAX (Addgene #12260) lentiviral packaging vectors. Transduced cells were selected and maintained in full medium supplemented with 5 µg/ml blasticidin for 10 days. After selection, the WT-ER palbociclib-resistant MCF7 cells overexpressing WT-ER (ER-OE PalboR MCF7 cells) were maintained in full medium supplemented with 1 µg/ml puromycin, 5 µg/ml blasticidin, and 1 µM palbociclib.

Cells were authenticated by short tandem repeat profiling (Bio-Synthesis, USA) and regularly tested for mycoplasma contamination using the MycoAlert® Mycoplasma Detection Kit (Lonza) according to the manufacturer's instructions.

Single cell clone isolation

Single cell clones were generated by serial limiting dilutions of either ER-WT or ER-mutant vehicle-treated and palbociclib-resistant MCF7 cells, seeding 1 cell per well in 96-well plates.

To validate that the single cell clones originated from a single cell, single cell subclones were generated by serial limiting dilutions of either ER-WT or ER-mutant palbociclib-resistant clonal cells, seeding 1 cell per well in 96-well plates.

ER-WT and ER-mutant vehicle-control clones were regularly cultured in full medium supplemented with 0.01% DMSO, 1 µg/ml of puromycin, with or without DOX. ER-WT and ER-mutant palbociclib-resistant clones and subclones were regularly cultured in full medium supplemented with palbociclib 1 µM, 1 µg/ml of puromycin, with or without DOX.

Droplet digital PCR (ddPCR) for validation of single cell clones

DNA was extracted from vehicle-control and palbociclib-resistant clones using the Blood & Cell Culture DNA Midi Kit (Qiagen) and quantified by Nanodrop (Thermo Fisher Scientific). The allelic fraction of the barcodes enriched in a clone was assessed using the ddPCR™ Copy Number Variation Assay (Biorad) and the QX200 ddPCR system (Biorad). Data were analyzed with the QuantaSoft software (Biorad). The number of copies of the *ASL* gene/clone was assessed with a pre-designed assay comprising both a specific *ASL* HEX probe and *ASL*-specific primers (dHsaCP2506678, Biorad) and used as a reference. The number of copies of the top enriched barcodes in each clone was checked using the primers and the FAM probes reported in Additional file 2: Table S2 and determined based on the barcode/*ASL* ratio.

Xenograft studies

Ovariectomized NSG (Jackson, RRID:IMSR JAX:005557) mice that were 5 weeks old were placed on 2000 ppm doxycycline diet and then implanted with 5 million MCF7 cells stably expressing the DOX inducible Y537S *ESR1* mutation and the ClonTracer library. Specifically, cells were resuspended in 50% matrigel and injected in the right mammary fat pad. Weekly body weight and tumor caliper measurements were taken for the remainder of the study.

Eight days after cell injection, surgery was performed in 3 mice (#816, #818, #823) to remove the mammary

fat pad tumor (average tumor size: 92 mm³) and assess the complexity of the library of the injected cell population. Survival surgery was performed in the remaining 10 mice 28 days after cell injection (average tumor size: 350 mm³). Primary tumors were collected from all the mice and snap frozen in liquid nitrogen. After the removal of the early mammary tumor, local tumor re-growth was monitored by caliper assessments of mammary fat pads. Metastatic tumor growth was monitored with MRI imaging every 3 weeks. Mice were euthanized if they showed significant weight loss due to tumor burden, if tumors exceeded 2 cm in size, or if MRI scans revealed substantial tumor burden. Recurrent tumors grown in the mammary fat pad and distant metastases were collected and snap frozen in liquid nitrogen.

All mice were maintained and all procedures were conducted in accordance with local guidelines and therapeutic interventions approved by the Animal Care and Use Committees of Dana-Farber Cancer Institute, protocol 04–111. Mice were housed up to 5 mice/cage. The room temperature was kept at 20–26 °C and humidity 30–70%. The animals were under a 12 h:12 h light–dark cycle and had ad libitum access to water and irradiated Laboratory Rodent Diet. Animals had at least 3 days of acclimation prior to study initiation. At the experimental endpoint, mice were euthanized by carbon dioxide inhalation using a gradual fill method, followed by a secondary physical method (cervical dislocation) to ensure death, in accordance with institutional animal care guidelines and the American Veterinary Medical Association (AVMA) Guidelines for the Euthanasia of Animals.

Drugs

Palbociclib (cat. #S1579) and abemaciclib (cat. #S7158) for in vitro studies were purchased from Selleckchem.

Growth assays

Cells were seeded in triplicate in 96-well plates (5,000 cells/well), and 24 h later (day 0) treated with palbociclib, abemaciclib, or vehicle (0.01% DMSO). At specific timepoints cells were stained with 0.4 µg/ml hoechst for nuclear staining and 4 µg/ml propidium iodide to label dead cells. The number of live and dead cells following drug treatments was counted using the Celigo image Cytometer (Nexcelom). For the IC₅₀ studies, cells were counted at day 5 and the number of live cells was normalized against the number of cells treated with the respective drug's vehicle. IC₅₀ values were next calculated in Graphpad Prism using the log(inhibitor) vs. response – Variable slope (four parameters) function. For the GR metrics, cells were counted at day 0 and at day 6. GR₅₀ and GR_{max} values were next calculated using the R metrics package (v1.36) (<https://bioconductor.org/packages/GRmetrics/>). For the growth studies, the number of live

cells at different timepoints was normalized to the number of live cells counted at day 0.

Functional analysis of ER-WT and ER-mutant impact on acquired CDK4/6i resistance

To assess the impact of the Y537S *ESR1* mutation in the acquired resistance to CDK4/6is, two million ER-mutant and ER-WT palbociclib-resistant MCF7 cells that were maintained in full medium with and without DOX, respectively, and with palbociclib 1 µM, were seeded in 10 cm dishes and cultured both with and without DOX for 4 weeks. At each passage, cells were counted and two million cells were replated. Cell numbers were normalized to the re-plated cell number and plotted.

To determine whether restoring the expression of WT-ER in ER-WT palbociclib-resistant cells is sufficient to resensitize the cells to CDK4/6is, ER-WT palbociclib-resistant MCF7 cells (no-DOX) were transduced to overexpress WT-ER (as above in the cell line section). Two million ER-WT palbociclib-resistant MCF7 cells, with and without ER-WT overexpression, were kept in their respective medium supplemented with palbociclib 1 µM. At each passage, cells were counted and two million cells were plated back. Cell numbers were normalized to the re-plated cell number and plotted.

Western blots

Whole cell protein extraction was performed scraping the cells with RIPA buffer (Boston BioProducts) supplemented with protease and phosphatase inhibitors (Sigma Aldrich). Cell pellets were incubated 30 min at 4 °C while rotating, sonicated for 1 min and centrifuged for 15 min at 4 °C at 14,000 g. Supernatants were assessed by BCA assay (Thermo Fisher Scientific) to measure the protein concentration of the extracts, following the manufacturer's instructions. Samples were subjected to SDS-PAGE by using NuPAGE 4 to 12% Bis–Tris gels (Life Technologies) and blotted on Trans-Blot Turbo Midi Nitrocellulose membranes (Bio-Rad Laboratories). Primary antibodies: Rb (4H1) (CST #9309, 1:1,000), ER (D6R2W) (CST #13258, 1:1,000) and GAPDH (D16H11) (Brand CST #5174, 1:1,000). Secondary antibodies: anti-rabbit (Invitrogen #31458, 1:2,500) and anti-mouse (Invitrogen #SA1-100, 1:2,500).

Barcodes sequencing

Genomic DNA was extracted from 10 million cells or frozen pulverized tumors using Blood & Cell Culture DNA Kit (Qiagen). Library preparation was performed using 2 µg of genomic DNA, Titanium Taq DNA Polymerase (Takara Bio) and primers following the PCR conditions reported in Additional file 2: Table S2.

The PCR product (~183 bp) was purified using the QIAquick PCR purification kit (Qiagen) following the

manufacturer's protocol, quantified by Qubit dsDNA high-sensitivity assay (ThermoFisher Scientific) and sequenced at the Molecular Biology Core Facility of Dana-Farber Cancer Institute by Illumina NS500 Single-End 75 bp sequencing, using the Illumina custom primer GCGACCACCGAGATCTACACTGACTGCAGTCTGAGTCTGACAG.

Barcode sequencing analyses

The barcode composition analysis was performed using the ClonTracer algorithm as previously described [34] and as reported in Addgene (Addgene #69830, <https://www.addgene.org/pooled-library/clontracer/>).

Barcode-level analyses

Low-frequency barcodes, defined as barcodes with fractions less than a chosen threshold fraction across all samples for models included in an analysis, were removed prior to the generation of the barcode fraction heatmaps (threshold=0.01), the hierarchical clustering of all the bulk samples (threshold=0.001), and the study of the dynamics of cell clusters within each condition (threshold=0.001). These low-frequency barcodes were retained for barcode-level entropy calculations.

For barcode fraction heatmaps, Euclidean distances between barcodes (samples) were calculated on rows (columns) of the (barcodes x samples) natural-log transformed barcode fraction matrix and hierarchically clustered with complete linkage. A pseudocount of $1e-12$ was added to all barcode fractions prior to log transforming for distance calculations.

Barcode distributions were compared across pairs of samples using the log2 unit Jensen-Shannon (JS) divergence using the philentropy (v0.9.0) R package [35]. Matrices of JS divergences among samples were hierarchically clustered using complete linkage. The JS divergences between samples and their respective parental control were compared across three timepoints to assess the progression of clonal selection using two-sided paired t-tests adjusted for multiple comparisons using the Holm procedure with pairing by evolutionary replicate.

Clonal selection among the barcodes at the endpoint for each sample was compared by calculating the normalized, natural-log unit Kullback–Leibler (KL) divergence of the sample barcode fraction distribution against a reference discrete uniform distribution over the same number of barcodes and normalizing by the total number of barcodes using the philentropy (v0.9.0) R package [35].

Normalized KL divergences at the barcode level were compared between models using two-sided, unpaired, unequal variances t-tests using rstatix (v0.7.2) (<https://doi.org/10.32614/CRAN.package.rstatix>) and adjusting for multiple testing with the Holm procedure [36].

Cellular population-level analyses

Cellular populations per model were chosen by taking the (barcode x sample) barcode fraction matrix across one parental model and all 3 timepoints (Early, Middle, End) for all 3 evolutionary replicates of a model (10 samples total) and fitting a Poisson GLM per row and extracting the Poisson Pearson residual (PPR) matrix. Euclidean distances between barcodes were calculated on rows of the PPR matrix and hierarchically clustered with complete linkage. The resulting dendrogram was cut to get 8 cellular populations per model.

Clonal selection among the 8 cellular populations at the End timepoint for each sample per condition was evaluated by calculating the normalized, natural-log KL divergence of the cellular population fraction distribution (rescaled to 1) against a reference discrete uniform distribution over 8 elements using the philentropy (v0.9.0) R package [35].

Normalized KL divergences at the cellular population level were compared between models using unpaired, t-tests with unequal variances, using rstatix (v0.7.2) (<https://doi.org/10.32614/CRAN.package.rstatix>), and were adjusted for multiple testing with the Holm procedure where appropriate [36].

All barcoding analyses were done in R (v4.3.2, v4.4.1) using tidyverse (v2.0.0) (<https://doi.org/10.32614/CRAN.package.tidyverse>) packages and readxl (v1.4.5) (<https://doi.org/10.32614/CRAN.package.readxl>). Plots were created in R, using ggplot2 (v3.5.2) (<https://doi.org/10.32614/CRAN.package.ggplot2>), RColorBrewer (v1.1–3) (<https://doi.org/10.32614/CRAN.package.RColorBrewer>), pheatmap (v1.0.12) (<https://doi.org/10.32614/CRAN.package.pheatmap>), and ggpubr (v0.6.0) (<https://doi.org/10.32614/CRAN.package.ggpubr>), or in GraphPad Prism (version 10.0.0).

WES of cell models

Genomic DNA was extracted from 10 million cells or frozen pulverized tumors using Blood & Cell Culture DNA Kit (Qiagen). Quality controls, library preparation and WES were performed at Novogene Corporation Inc. Sequencing libraries were generated using Agilent SureSelectXT2 Homo Sapiens All Exon V6 Kit (Agilent Technologies, CA, USA) based on the manufacturer's recommendations. WES was conducted by Illumina Paired-End 150 bp sequencing.

Cell line WES analysis

Mutation calling

Mutation calling was performed as described here: <https://assets-eu.researchsquare.com/files/rs-4270758/v1/784d15e4-765a-4b1e-a02b-7c93680b707c.pdf?c=1716564499>. Briefly, somatic single nucleotide variants were called using MuTect [37], followed by filters to remove

any mutations due to OxoG artifacts. A BLAT realignment filter was used downstream to remove any mutations having potentially mismatched reads. In our analysis we used the parental cell line as a control sample. For details, refer to code in GitHub (link provided in the *Data availability* section).

CNV

We used the GATK-CNV (<https://github.com/gatk-workflows/gatk4-somatic-cnvs>) segmentation algorithm to assign copy-number segments along the exomes. To obtain gene-level copy ratio calls, a read-count normalization was performed by calculating separately for vehicle-control cells, drug-resistant cells and parental cells, the proportion of all reads that mapped to a gene of interest. The samples had systematic differences in coverage; for instance, the samples treated with abemaciclib had greater average depth than most samples treated with palbociclib. To mitigate sample-specific variance differences while preserving within-sample copy-number patterns, we applied per-sample z-score normalization prior to visualization. The normalized proportion of reads that mapped to a gene computed for the vehicle-control and drug-resistant cells was then divided by that one computed for the parental cells, to further normalize the gene-level read count. Code implementation can be found in GitHub (link provided in the *Data availability* section).

Bulk RNA-seq of cell models

Total RNA was extracted from cells seeded in 6-well plates using TRIzol (Life Technologies) and RNeasy Mini Kit (Qiagen). Quality controls, library preparation and RNA-seq were performed at the Molecular Biology Core Facility of Dana-Farber Cancer Institute. RNA-seq libraries were prepared using KAPA mRNA HyperPrep Kit (Roche). The RNA was sequenced by Illumina Nova-seq6000 Paired-End 150 bp sequencing. RNA extraction and sequencing were performed at least in technical duplicates.

RNA-seq analysis

Sequenced reads were aligned to the hg19 reference genome assembly and samples were analyzed using the pipeline VIPER [38]. Alignment to the hg19 human genome was done using STAR v2.7.0f followed by transcript assembly using cufflinks (v2.2.1) [39] and RseQC (v2.6.2) [40]. Batch effect adjustments were made using ComBat [25]. Differential expression analysis was done using DEseq2 (v1.18.1) [41].

Single cell-RNA sequencing (scRNA-seq)

To perform scRNA-seq analyses we used single-cell suspensions from parental cells and ER-WT and ER-mutant

vehicle-control, palbociclib-resistant and abemaciclib-resistant cells. Specifically, for vehicle-control, palbociclib-resistant and abemaciclib-resistant conditions, the scRNA-seq was performed on the evolutionary replicates #3. ScRNA-seq analyses were also performed on the ER-WT palbociclib-resistant clone #3_15.

ScRNA-seq libraries were prepared at the Translational Immunogenic Lab at Dana-Farber Cancer Institute, according to 10× Genomics Chromium Next GEM Single Cell 3' v3.1 protocol and using cell stock concentrations with a likelihood of around 8,000 transcriptomes recovering. The quality of the libraries was checked by Bioanalyzer using the High Sensitivity DNA kit (Agilent). The RNA was sequenced by Illumina HiSeq or Novaseq6000 Paired-End 150 bp sequencing at Novogene Corporation Inc.

ScRNA-seq analysis

Samples were aligned with the nf-core/scrnaseq (v2.7.1) [42] pipeline from the collection of nf-core workflows [43] using the Bioconda [44] and Biocontainers [45] environments and Nextflow (v23.04.1). Alignment was done using the default Salmon Alevin [46, 47] pseudoaligner to the hg38 reference genome and GENCODE GTF, using pre-built reference assemblies and annotations from refgenie (v0.12.1) [48, 49] with custom options ‘-profile singularity’, ‘-protocol 10X3’, and ‘-skip-emptydrops’.

Salmon Alevin outputs were imported into R using fishpond (v2.10.0) [50] in mode ‘loadFry(outputFormat = "snRNA")’ and empty droplets and dead cells were filtered out by exploratory data analysis of QC metrics and manually choosing thresholds to keep cells with percent mitochondrial reads $\leq 25\%$, minimum number of genes with nonzero reads per cell ≥ 100 , and minimum total number reads per cell $\geq 10,000$. DropletUtils (v1.24.0) emptyDrops function was used to predict empty droplets, keeping cells with an FDR ≤ 0.01 [51, 52]. Count matrices were converted using SeuratObject (v5.0.2) (<https://doi.org/10.32614/CRAN.package.SeuratObject>) into Seurat (v5.2.1) [53] objects.

Dimensionality reduction was done by Poisson generalized linear model principal components analysis (GLM-PCA) [54, 55] on the gene count matrix as input and considering the top 2,500 genes with the highest deviance across all cells using scry R package (v1.12.0) (<https://doi.org/10.18129/B9.bioc.scry>) as input to fastglmPCA [56] (v0.1–103) (<https://doi.org/10.32614/CRAN.package.fastglmPCA>) with 50 GLM-PCs chosen. A UMAP dimensionality reduction was calculated on the first 50 GLM-PCA principal components (GLM-PCs) as input using the uwot R package (v0.2.3) (<https://doi.org/10.32614/CRAN.package.uwot>).

Cell–cell distances within each sample were calculated as Euclidean distances on the first 50 GLM-PCs on all

pairwise comparisons of 1,000 sampled random cells per sample.

Cell cycle phases were predicted using Seurat's Cell Cycle Scoring function on expression values from Seurat's default `NormalizeData` (normalization.method = 'Log-Normalize') normalization.

Clustering was done using kmeans clustering on the first 50 GLM-PCs. A combination of exploratory plots of cell embeddings, silhouette scores, expression of marker gene sets, and clustering tree diagrams examining size, stability, and separation of clusters were used to choose the number of biologically meaningful clusters per sample using the `cluster` (v2.1.8) (<https://doi.org/10.32614/CRAN.package.cluster>) and `clustree` (v0.5.1) R packages [57]. In downstream analyses split by cell cycle phase, clusters were subsequently subdivided into cells of a particular predicted cell cycle phase.

Gene set testing was performed by calculating the quasi-Poisson GLM Wald statistic for each gene for comparing expression rates either between pairs of samples or between clusters within a sample, where each gene for each sample has its own estimated overdispersion term using `Matrix` (v1.7–3) (<https://doi.org/10.32614/CRAN.package.Matrix>) and `MatrixGenerics` (v1.16.0) (<https://doi.org/10.18129/B9.bioc.MatrixGenerics>). Genes were ranked by these Wald statistics prior to pre-ranked correlation-adjusted mean rank gene set testing using `cameraPR` [58] from `limma` (v3.60.6) [59] on the Hallmark gene collection with `msigdb` [60] (v7.5.1) (<https://doi.org/10.32614/CRAN.package.msigdb>). For gene set testing between two models, genes were ranked by the Wald statistic comparing the difference of expression rates between those two models. For gene set testing of one cluster compared to multiple others (marker gene set setting), we rank genes based on all of the Wald statistics comparing that fixed cluster to the others. For a gene, if all such Wald statistics are positive, we take their minimum as the ranking-statistic for that gene, if all negative, we take their maximum as the ranking-statistic for that gene, and if they are mixed, we take zero as the ranking-statistic for that gene. This prioritizes marker genes that have expression rates that are larger or smaller than all other clusters as distinctive markers and places genes that are mixed among the other clusters in the middle of the ranked gene list. For within-sample across-cluster testing, we omitted any cluster where there were fewer than 50 genes, and for across-cluster comparisons within a cell cycle phase, we compared clusters of cells for a single predicted cycle phase (S, G1, G2M) at a time. Multiple testing of gene sets was accounted for by controlling the false discovery rate (FDR) within families of tests using the Benjamini–Hochberg procedure [61].

Copy number inference was run on the gene count matrix with the R package `inferCNV` (v1.20.0) (`inferCNV`

of the Trinity CTAT Project. <https://github.com/broadinstitute/inferCNV>), considering the parental cells as the reference population and using the following options: `infercnv::run(cutoff=0.1, cluster_by_groups=TRUE, denoise=TRUE, HMM=FALSE)` and `infercnv::plot_cnv(cluster_by_groups=TRUE, x.center=1, x.range="auto", hclust_method='ward.D', color_safe_pal=FALSE, png_res=300, dynamic_resize=0)`.

All scRNA-seq analyses were performed in R (v4.3.2, v4.4.1) using `tidyverse` (v2.0.0) (<https://doi.org/10.32614/CRAN.package.tidyverse>) packages. Plots were created using `ggplot2` (v3.5.2) (<https://doi.org/10.32614/CRAN.package.ggplot2>) and `RColorBrewer` (v1.1–3) (<https://doi.org/10.32614/CRAN.package.RColorBrewer>).

ATAC-sequencing (ATAC-seq)

Nuclei extraction from vehicle-control, palbociclib-resistant and abemaciclib-resistant MCF7 cells, either ER-WT or ER-mutant, and transposition reactions were performed as previously described [62].

Briefly, cells were collected and washed. Cells pellets were resuspended and incubated for 3 min in 50 μ L of cold ATAC-seq resuspension buffer (RSB; 10 mmol/L Tris–HCl pH 7.4, 10 mmol/L NaCl, and 3 mmol/L MgCl₂ in water) containing 0.1% NP40, 0.1% Tween-20, and 0.01% digitonin. After the incubation, cells were resuspended in 1 mL of cold RSB containing 0.1% Tween-20 but no NP40 or digitonin and centrifuged. Nuclei were then resuspended in 50 μ L of transposition mix (25 μ L 2 $\times$ TD buffer, 2.5 μ L transposase (100 nmol/L final), 16.5 μ L PBS, 0.5 μ L 1% digitonin, 0.5 μ L 10% Tween-20, and 5 μ L water). Transposition reactions were incubated at 37 °C for 30 min with 1,000 rpm mixing. Reactions were cleaned up with Qiagen MinElute Kit following the manufacturer's instruction.

Library preparation was performed as previously described [63] using 5 uM Customized Nextera PCR primers, Invitrogen 10X SYBR Green I, NEBNext High-Fidelity 2X PCR Master Mix (New England Labs).

The amplified libraries were purified using Qiagen PCR Cleanup Kit according to the manufacturer's protocol and sequenced at Novogene.

ATAC-seq analysis

Analysis was done using COBRA [64] and CHIPS [65]. Sequence tags were aligned to the hg19 reference genome using the Burrows-Wheeler Aligner (BWA) [28]; only uniquely mapped, non-redundant reads were retained. These reads were used to generate binding sites with Model-Based Analysis of ChIP-Seq 2 (MACS v2.1.1.20160309) [66], with an FDR threshold of 0.01. DESeq2 [41] was used to identify differentially accessible regions between drug-resistant vs vehicle-treated cells, ER-mutant vs ER-WT cells, and abemaciclib-resistant

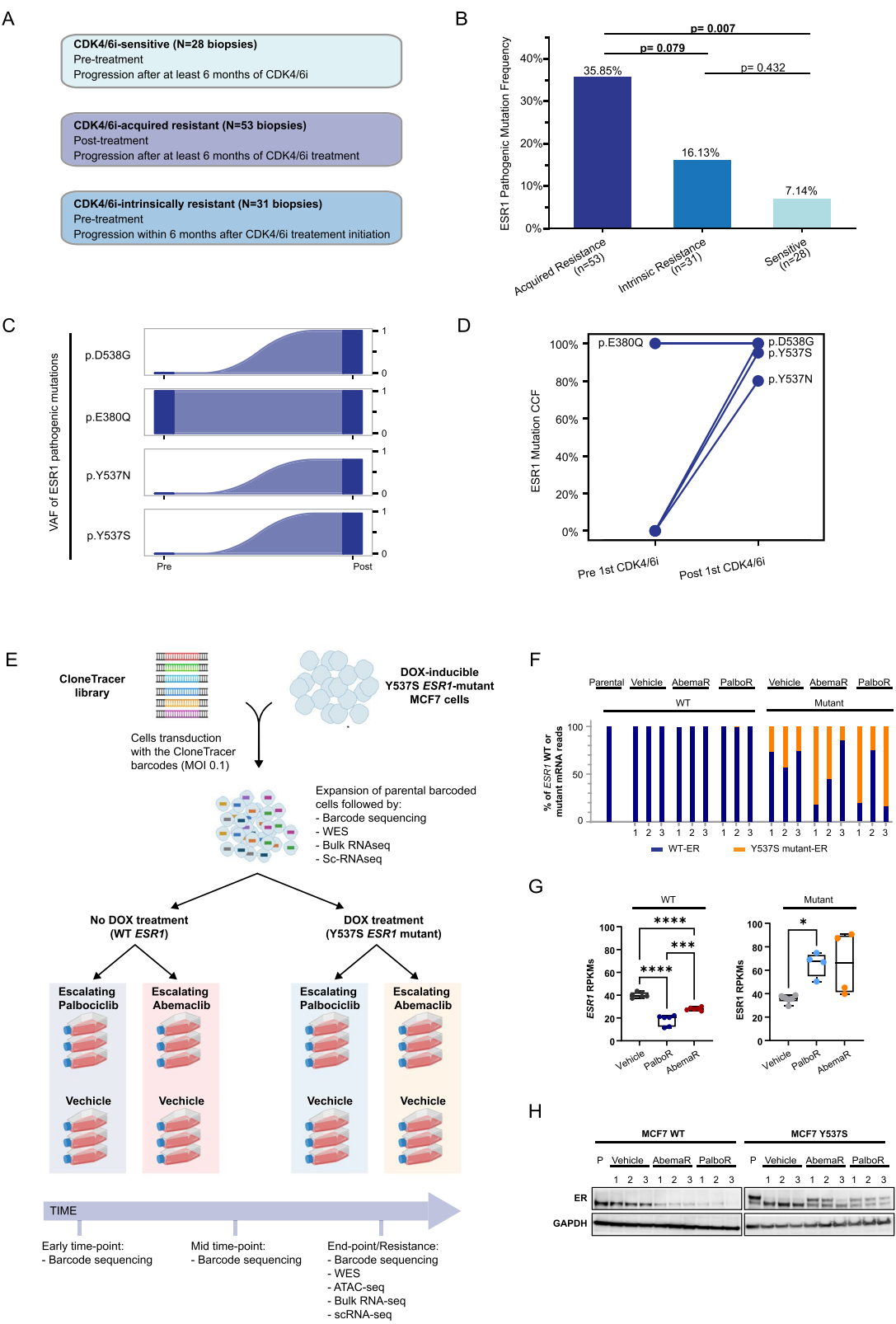


Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 *ESR1* mutation enrichment in clinical samples and cell models after acquired CDK4/6i resistance in ER+ breast cancer. **A** Patient cohort and tumor biopsy classification based on patient sensitivity to CDK4/6i-treatment (palbociclib or abemaciclib). Sensitive biopsies were collected within 1 year prior to starting treatment with a CDK4/6i from patients that had disease stability or responsive disease for at least 6 months in the metastatic setting. Intrinsic resistant biopsies were obtained from patients that developed resistance within 6 months of treatment with a CDK4/6i in the metastatic setting and were collected within 1 year prior to starting the treatment. Acquired resistance biopsies were obtained within 30 days prior to progression and up to 1 year after disease progression from patients that had disease stability or responsive disease for at least 6 months in the metastatic setting and later developed progressive disease. **B** Percentage of *ESR1*-mutant tumor biopsies in each subgroup: CDK4/6i-sensitive, CDK4/6i-intrinsic resistance, and CDK4/6i-acquired resistance. Fisher's exact test was applied for comparisons between subgroups. **C** Variant allele fraction (VAF) of pathogenic *ESR1* mutations in four matched tumor biopsies, collected at baseline (pre) and after acquired resistance to 1st line palbociclib treatment (post). **D** Cancer cell fraction (CCF) of pathogenic *ESR1* mutations in the four matched tumor biopsies (as in panel C) at baseline (pre 1st CDK4/6i) and after the development of acquired resistance (post 1st CDK4/6i). **E** Schema of the in vitro lineage tracing of the clonal evolution in CDK4/6i resistance using DNA-barcoded ER-WT and ER-mutant MCF7 cells. MCF7 cells with stable expression of a DOX inducible HA-tagged Y537S *ESR1* construct were transduced with the ClonTracer library at a multiplicity of infection (MOI) of 0.1 and divided into DOX-treated and untreated flasks. The DNA barcoded ER-WT and ER-mutant cells were further subdivided into treatment groups cultured with escalating doses of palbociclib or abemaciclib and the vehicle of the drugs (vehicle-controls). Each condition included three evolutionary replicates. The distribution of the barcoding library was assessed in the parental cells and in all the vehicle-control and drug-treated models, at different time points: an early time point (after 3 passages in full medium with palbociclib or abemaciclib at 50 nM), an intermediate time point (after 3 passages in full medium with palbociclib or abemaciclib at 200 nM), and at the time of resistance to palbociclib or abemaciclib. Vehicle-control cells were analyzed at the same number of passages as their respective drug-treated counterparts. Parental cells and all the resistant cell models were characterized by multi-omic analyses. **F** Percentage of *ESR1*-WT and Y537S *ESR1*-mutant mRNA reads estimated in bulk RNA-seq data from ER-WT parental MCF7 cells and from evolutionary replicates (#1, #2, #3) of ER-WT and ER-mutant palbociclib-resistant (PalboR), abemaciclib-resistant (AbemaR), and vehicle-maintained MCF7 cells. Percentages were computed within each sample. **G** Quantification of total *ESR1* mRNA in ER-WT and ER-mutant PalboR, AbemaR, and vehicle-maintained MCF7 cells, reported as Reads Per Kilobase per Million mapped reads (RPKM). The *ESR1* mRNA level was estimated from bulk RNA-seq data. Each boxplot shows the quartile distribution of the total *ESR1* mRNA in each experimental condition, with dots representing individual technical replicates. Statistical comparisons among conditions were performed using a one-way ANOVA followed by Tukey's multiple comparisons test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$). **H** Protein expression of ER in whole cell lysates from parental cells and from evolutionary replicates (#1, #2, #3) of ER-WT and Y537S ER-mutant vehicle-control, PalboR and AbemaR MCF7 cells. ER expression in parental cells was assessed after the cells were cultured in full medium with or without DOX for 72 h. In the ER-mutant cells, the expression level of the endogenous ER is represented by the lower ER band, while the expression level of the exogenous DOX-induced Y537S-mutant ER is represented by the upper HA-tagged ER band. GAPDH expression was used as loading control

vs palbociclib-resistant cells. Gained or lost peaks were defined with the threshold of log2 fold change of >1 or <-1 , respectively, and adjusted p -value <0.05 . Binding and Expression Target Analysis (BETA basic) (v1.0.7) [67] was used to integrate differentially accessible chromatin sites with differential gene expression data to infer direct target genes.

Results

ESR1 mutations are enriched in acquired CDK4/6i and ET resistance

To determine the clinical relevance of *ESR1* mutations in CDK4/6i resistance, we assembled a cohort of patients with ER+ metastatic breast cancer who were treated with a CDK4/6i, with or without ET, and provided metastatic tumor biopsies ($N=112$ biopsies, $N=89$ patients) as part of an IRB-approved tissue collection protocol [20]. Tumor biopsies were classified as CDK4/6i-sensitive if obtained prior to CDK4/6i treatment from patients who subsequently experienced disease progression after >6 months of therapy ($N=28$ pre-treatment biopsies), acquired resistant if obtained from patients who progressed following >6 months treatment with a CDK4/6i ($N=53$ post-treatment biopsies), and intrinsically resistant if obtained from patients who progressed within six months of treatment initiation ($N=31$ pre-treatment biopsies) (Fig. 1A) (Additional file 1: Table S1). This cohort included 23 pre-treatment sensitive tumors

matched with post-treatment acquired resistant tumors from the same patient.

Samples were subjected to whole exome sequencing (WES), and comparison of the frequency of *ESR1* pathogenic mutations within each subgroup showed that the *ESR1* mutations were enriched in biopsies of acquired resistance compared with sensitive tumors (p -value = 0.007, Fisher's test). Although *ESR1* mutations were numerically more frequent in intrinsically resistant tumors than in sensitive tumors, this difference was not statistically significant (p -value = 0.432) (Fig. 1B). The cohort included four matched pre-and post-CDK4/6i treatment biopsy pairs in which a pathogenic *ESR1* mutation was detected in at least one sample. We analyzed the variant allele fraction (VAF) of the *ESR1* mutations and performed phylogenetic analyses on these paired samples (Fig. 1C-D). In three pairs, the *ESR1* mutation was not detected in the pre-treatment sample but emerged after treatment with a VAF of near 1 and a cancer cell fraction of above 80%, consistent with near clonal representation after treatment. In one patient, the *ESR1* E380Q mutation was present as a truncal event. Taken together, these findings support enrichment of *ESR1* mutations in tumors with CDK4/6i-acquired resistance and treatment-associated expansion of *ESR1*-mutant subclones. Based on these observations, we next investigated the functional role of *ESR1* mutations in models of acquired CDK4/6i resistance.

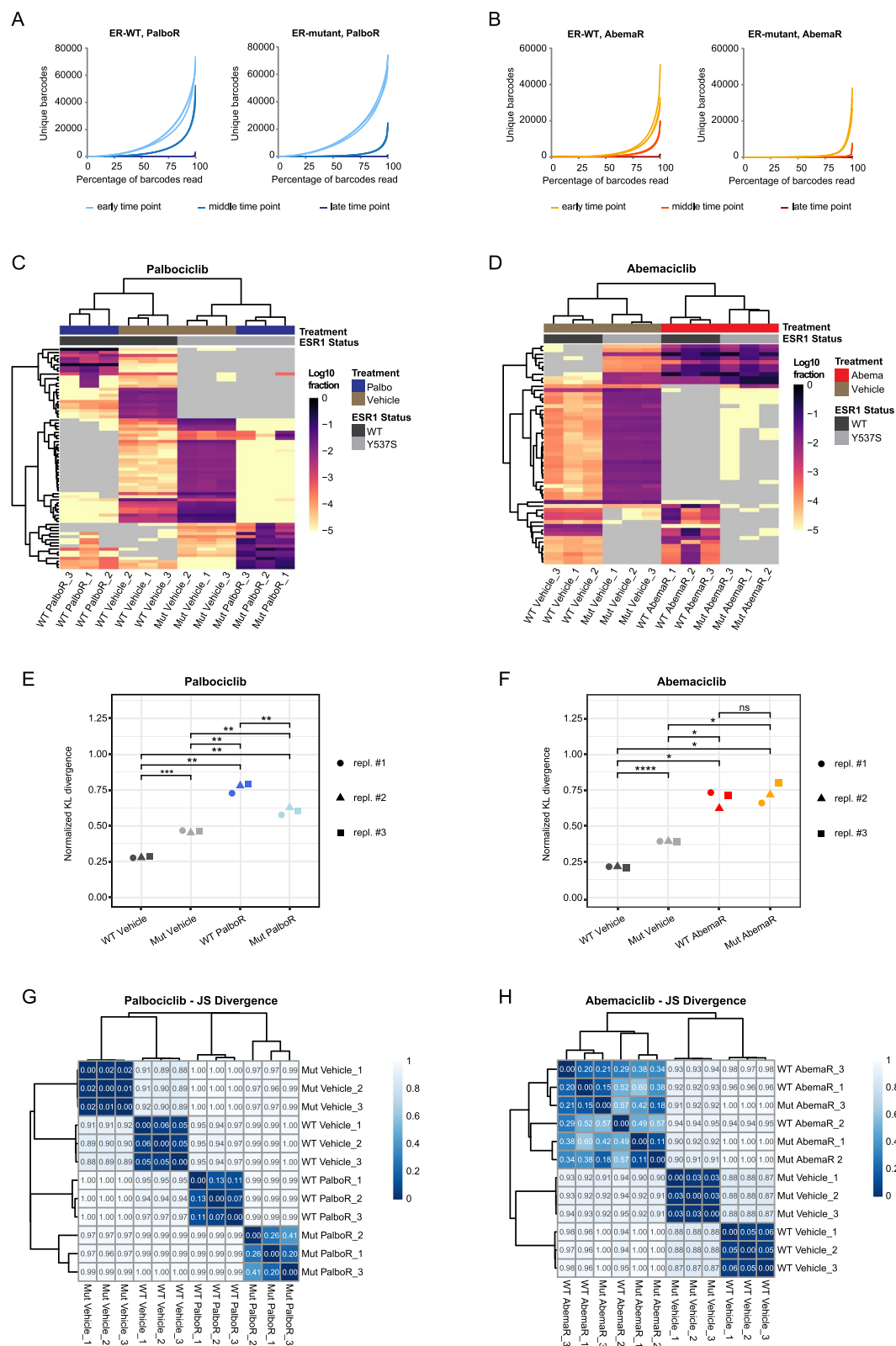


Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Clonal evolution of ER-WT and ER-mutant palbociclib-resistant and abemaciclib-resistant cells. **A–B** Cumulative curves showing the progressive selection of barcodes in ER-WT and Y537S ER-mutant evolutionary replicates of palbociclib-resistant (PalboR) (**A**) and abemaciclib-resistant (AbemaR) (**B**) MCF7 cells. The barcodes were sequenced at three timepoints (early time-point: palbociclib and abemaciclib 50 nM; middle time-point: palbociclib and abemaciclib 200 nM; late time-point: defined as resistance with cells growing in 1 μ M palbociclib or abemaciclib). **C–D** Unsupervised hierarchical clustering of vehicle-control and PalboR (**C**) and vehicle-control and AbemaR (**D**) evolutionary replicates ($N=3$) of ER-WT and Y537S ER-mutant MCF7 cells, based on barcode fractions. The clustering includes only barcodes with a fraction ≥ 0.01 in at least one sample. Barcode fraction is shown as log10 barcode fraction in a yellow–purple color gradient (yellow: low enrichment; purple: high enrichment). Barcodes with a fraction < 0.01 or not detected are shown in gray. **E–F** Normalized Kullback–Leibler (KL) divergence of barcode distribution in vehicle-control and PalboR (**E**) and vehicle-control and AbemaR (**F**) evolutionary replicates ($N=3$) of ER-WT and Y537S ER-mutant MCF7 cells, relative to a discrete uniform distribution. Lower values indicate more similar distributions, and higher values indicate more divergent distributions. Statistical comparisons among conditions were performed using two-sided unequal-variances t-tests with each set of comparisons adjusted for multiple testing using the Holm procedure (n.s., not significant; *, p-adjusted ≤ 0.05 ; **, p-adjusted ≤ 0.01 ; ***, p-adjusted ≤ 0.001 ; ****, p-adjusted ≤ 0.0001). **G–H** Jensen-Shannon (JS) divergences in vehicle-control and PalboR (**G**) and vehicle-control and AbemaR (**H**) evolutionary replicates ($N=3$) of ER-WT and Y537S ER-mutant MCF7 cells, based on barcode fraction distributions. The JS divergence between two barcode distributions is indicated by the blue–white color gradient (blue: similar distributions; white: divergent distributions) and is reported in each square. Rows and columns are hierarchically clustered by their divergences. Numbers 1, 2 and 3 indicate distinct evolutionary replicates

The Y537S *ESR1* mutation shapes the clonal evolution of CDK4/6i-resistant cells

To study the clonal evolution during acquired resistance to CDK4/6is, we employed the ClonTracer DNA barcoding system [34] with three primary objectives: (i) track clonal dynamics during the acquisition of resistance to CDK4/6i, (ii) assess the impact of the *ESR1* Y537S mutation on these dynamics, and (iii) compare the clonal dynamics in acquired resistance to palbociclib versus abemaciclib, two distinct CDK4/6is. MCF7 cells with stable expression of a doxycycline (DOX) inducible hemagglutinin (HA)-tagged Y537S *ESR1* construct [13] were transduced with the ClonTracer barcoding library at a MOI of 0.1 to minimize multiple barcode integrations per cell (Fig. 1E). Following confirmation of the even distribution and high diversity of the barcoded population (average of unique barcodes = 303,756, 95% CI = 300,796 to 306,715) (Additional file 3: Fig. S1A) (Additional file 4: Table S3), cells were expanded and divided into DOX-treated and untreated flasks, and further subdivided into treatment groups with vehicle-control, escalating doses of palbociclib or escalating doses of abemaciclib. Each condition included three evolutionary replicates plated at a 100-fold barcode representation to limit stochastic losses of barcodes during passaging and to ensure comparable starting barcode representation across replicates. Each replicate constitutes an independent realization of an evolutionary trajectory initiated from the same ancestral, barcoded population and subjected to identical selective pressures.

At baseline, expression of the Y537S *ESR1* mutation had a limited impact on sensitivity to either palbociclib or abemaciclib. As expected, exposure to escalating doses of CDK4/6is led to the acquisition of resistance (Additional file 5: Table S4) (Additional file 3: Fig. S1B–H). *ESR1* Y537S mRNA measurements confirmed expression of ER mutant transcript in all DOX-treated cells (Fig. 1F). Total *ESR1* mRNA levels were increased in palbociclib and abemaciclib-resistant cells expressing the *ESR1* Y537S mutation compared to the vehicle-treated

counterparts. In contrast, ER-WT resistant cells had lower total *ESR1* mRNA levels compared to their vehicle-control (Fig. 1G) (Additional file 6: Table S5). At the protein level, cells without DOX treatment did not express mutant ER, as expected. In these cells, WT-ER expression decreased after acquisition of CDK4/6i resistance. In parental cells, mutant ER expression was confirmed after 72 h of DOX exposure. However, prolonged culture in vehicle led to loss of mutant ER expression, consistent with prior observations that the Y537S *ESR1* mutation does not provide a fitness advantage in full medium conditions. In contrast, Y537S *ESR1* expressing cells treated with palbociclib or abemaciclib retained mutant ER expression (Fig. 1H). These differences in ER expression in distinct conditions provide initial evidence supporting a role of the Y537S *ESR1* mutation in the development of CDK4/6i resistance.

Long-term treatment of cells with vehicle or increasing doses of palbociclib and abemaciclib led to gradual selection of specific cellular subpopulations, as shown by barcode sequencing at the three time points (Fig. 2A–B and Additional file 3: Fig. S2A–B), consistent with recent studies showing that adaptive resistance to target therapies occurs along a continuum [68, 69]. Relative to vehicle-control, CDK4/6i treatment led to stronger enrichment of specific barcodes. Using a cut-off defined by the maximum single-barcode fraction in the parental population (fraction = 0.000172), we found that in ER-WT cells an average of 738 barcodes (ranging from 728 to 746 across the three replicates) were enriched in the vehicle condition, and an average of 270 barcodes (ranging from 234 to 290 across the three replicates) were enriched in the vehicle-treated ER mutant cells. In palbociclib-resistant cells, 88 barcodes (ranging from 59 to 107 across the three replicates) and 115 barcodes (ranging from 78 to 142 across the three replicates) were enriched in the ER-WT and ER-mutant conditions, respectively. Given the barcode complexity at baseline prior to treatment (303,756 barcodes), this corresponds to 0.029% and 0.038% of parental cells contributing to the ER-WT and

ER-mutant palbociclib-resistant cells, respectively. Using the same threshold, abemaciclib-resistant cells showed enrichment of 46 barcodes (ranging from 40 to 50 across the three replicates) in the ER-WT cells and 29 barcodes (ranging from 18 to 38 across the three replicates) in the ER-mutant cells. These values indicate that 0.015% and 0.0097% of parental cells contributed to ER-WT and ER-mutant abemaciclib-resistant populations, respectively (Additional file 4: Table S3).

To begin to decipher differences in the barcode selection in WT and mutant ER cells with palbociclib and abemaciclib resistance, we performed unsupervised hierarchical clustering of barcode fractions, restricting the analysis to barcodes with a fraction ≥ 0.01 in at least one sample. In all conditions, the three evolutionary replicates had the strongest level of similarity (Fig. 2C–D). The impact of the expression of the Y537S mutation on the barcode selection differed between palbociclib and abemaciclib resistance. In the palbociclib- and vehicle-treated cells, the primary clustering split grouped the cells by *ESR1* mutational status, followed by treatment type (vehicle-control versus palbociclib) (Fig. 2C). Conversely, in the abemaciclib- and vehicle-treated cells, treatment was the dominant driver of clustering, followed by *ESR1* mutational status (Fig. 2D). Consistent with these patterns, the normalized Kullback–Leibler (KL) divergence, which quantifies the extent to which the barcode distribution in each condition differs from that of the parental population and thus reflects changes in clonal population structure, was higher in palbociclib-resistant WT-ER than in palbociclib-resistant mutant-ER cells (adjusted p -value < 0.01 , t-test with each comparison adjusted for multiple testing using the Holm procedure) (Fig. 2E). This finding indicates that the *ESR1* mutation is associated with altered clonal dynamics during the evolution of palbociclib resistance, with reduced dependence on marked enrichment of a limited number of subclones. In contrast, the normalized KL divergence was similar between WT-ER and mutant-ER abemaciclib-resistant cells (adjusted p -value > 0.05 , t-test with each comparison adjusted for multiple testing using the Holm procedure), suggesting a more limited effect of the mutation on changes in clonal population structure during the acquisition of abemaciclib resistance (Fig. 2F), consistent with the broader target profile of abemaciclib [19]. In both drug-resistant conditions, the normalized KL divergence was higher compared to the vehicle-control treated cells, consistent with greater shifts in barcode composition and reduced clonal diversity after the acquisition of CDK4/6i resistance. Collectively, these results provide evidence of clonal selection during the acquisition of CDK4/6i resistance, however, they do not rule out the contributions of cell plasticity and adaptive transitions during resistance.

These results also support a more prominent role for *ESR1* Y537S in the remodeling of the clonal population structure during palbociclib resistance than abemaciclib resistance.

Although enriched barcodes aggregated the evolutionary replicates, the replicates did not fully overlap. In this experimental design, divergence in adaptive trajectories across evolutionary replicates is the expected null outcome. In the palbociclib-resistant models, only 15.4% of the enriched barcodes in ER-WT cells and 55.6% of enriched barcodes in ER-mutant cells were shared across all the three replicates (ER-WT: average sum of the fractions of the shared barcodes was 0.79; ER-mutant: average sum of the fractions of the shared barcodes was 0.81) (Additional file 3: Fig. S2C–D). Similarly, in the abemaciclib-resistant models, 22.7% of the barcodes enriched in ER-WT cells and 44.4% of the barcodes enriched in ER-mutant cells were shared across all the three evolutionary replicates (ER-WT: average sum of the fractions of the shared barcodes was 0.83; ER-mutant: average sum of the fractions of the shared barcodes was 0.43) (Additional file 3: Fig. S2E–F).

To quantify the differences in clonal population structure across all cell models, we calculated pairwise Jensen–Shannon (JS) divergence between sample barcode distributions, with lower values indicating compositional similarity and higher values indicating greater divergence. JS divergence was generally low within the evolutionary replicates for ER-WT palbociclib-resistant cells (range: 0.07–0.13) and mutant palbociclib-resistant cells (range: 0.2–0.41), whereas it was higher for the ER-WT abemaciclib-resistant (range, 0.2–0.52) and mutant abemaciclib-resistant (range: 0.11–0.42) samples (Fig. 2G–H). These findings indicate that the degree of clonal population restructuring differed between the palbociclib- and abemaciclib-resistant models, consistent with distinct evolutionary responses to these two CDK4/6is.

We also calculated the JS divergence between the samples over three time points (early, middle and end), and their respective parental control, which revealed a gradual increase in JS divergence from baseline over the course of resistance acquisition (two-sided paired t-tests, with each set of timepoint comparisons adjusted for multiple testing within a mutational status and treatment using the Holm procedure) (Additional file 3: Fig. S2G–H). Together, these results support the co-existence of divergent evolutionary trajectories and collateral profiles alongside convergence, the latter consistent with selection within a specific constrained fitness landscape that is shaped by the underlying mutational background and occurs in a gradual manner [70].

We next directly compared palbociclib and abemaciclib-resistant evolutionary replicates (Fig. 3A). Pairwise

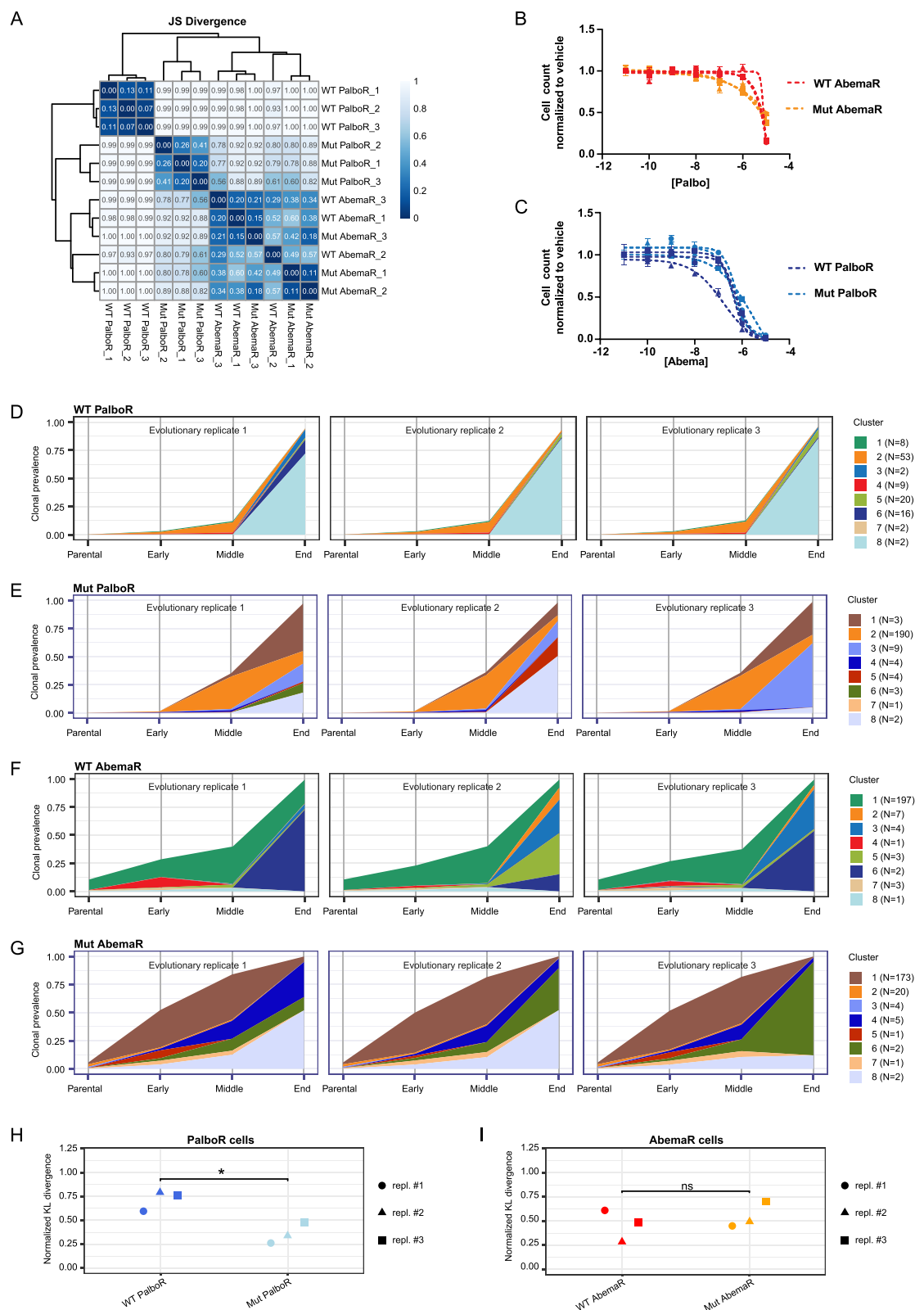


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Clonal evolution of abemaciclib-resistant versus palbociclib-resistant cells with ER-WT or ER-mutant expression. **A** Jensen-Shannon (JS) divergences between palbociclib-resistant (PalboR) and abemaciclib-resistant (AbemaR) evolutionary replicates (#1, #2, #3) of ER-WT and Y537S ER-mutant MCF7 cells, based on barcode fraction distributions. The JS divergence between two barcode distributions is indicated by the blue-white color gradient (blue: similar distributions; white: divergent distributions) and is reported in each square. Rows and columns are hierarchically clustered by their divergences. **B-C** Dose-response curves in ER-WT and Y537S ER-mutant AbemaR (**B**) and PalboR (**C**) MCF7 cells after 5 days of palbociclib and abemaciclib treatment, respectively. Experiments were performed in triplicates and data were reported as average $\pm$ SEM. **D-G** Dynamics of cell populations identified by clusters of enriched barcodes detected in ER-WT and Y537S ER-mutant PalboR and AbemaR MCF7 evolutionary replicates. For each condition, we identified eight clusters of barcodes with similar fractions over time and consistent across the three evolutionary replicates. Each cluster is shown in a different color. Color codes are independent across conditions. The number of barcodes included in each cluster is reported in brackets. The plots show the changes in clonal prevalence over time, calculated as the cumulative barcode fraction of the barcodes assigned to each cluster. Low-frequency barcodes that were not enriched above a threshold (0.001) at any timepoint are not plotted. **H-I** Normalized Kullback-Leibler (KL) divergence of the clusters distribution in PalboR (**H**) and in AbemaR (**I**) evolutionary replicates, relative to a discrete uniform distribution. Lower values indicate more similar distributions and higher values indicate more divergent distributions. Statistical comparisons among conditions were performed using a t-test for each treatment (n.s., not significant; *, $p \leq 0.05$)

JS divergences between barcode distributions in the palbociclib versus abemaciclib-resistant models was high overall in ER-WT (range: 0.93–0.99) and mutant (range: 0.6–0.92) samples. Moreover, the overlap between the enriched barcodes in palbociclib and abemaciclib resistance was limited: 2.9% and 22.7% in ER-WT and mutant cell lines, respectively, using the union of the barcodes in the three replicates and a barcode fraction of >0.01 (Additional file 3: Fig. S2I–J). Functionally, abemaciclib-resistant cells were cross-resistant to palbociclib, whereas palbociclib-resistant cells remained sensitive to abemaciclib, although with elevated IC_{50} values (1.6×10^{-7} M– 3×10^{-6} M) (Fig. 3B–C) (Additional file 5: Table S4). These results further support differences in clonal selection and divergent resistance mechanisms to these two CDK4/6is.

Inferred cellular populations reveal *ESR1* mutant and CDK4/6i specific dependent dynamics

In a barcoded cell population, it is unlikely that every unique barcode represents a truly distinct cellular state. Moreover, despite using a low MOI, a fraction of cells is likely to be infected with more than one barcode. To address these limitations, we inferred cellular populations by grouping barcodes with similar temporal trajectories, tracking changes in barcode fractions over time. Using this approach, we identified 8 cellular populations in each model. These populations were reproducible across the three evolutionary replicates within each model but showed substantial variability and limited overlap across models (Additional file 3: Fig. S3A–F). Comparing population dynamics across models, *ESR1* Y537S expression was associated with earlier enrichment in vehicle-treated and palbociclib or abemaciclib-resistant cell populations (Fig. 3D–G and Additional file 3: Fig. S3G–J). Consistent with the single-barcode analyses, we observed a stronger selection of specific populations in the ER-WT vs ER-mutant palbociclib-resistant models and a higher normalized KL divergence relative to a discrete uniform in the ER-WT cellular populations compared

to *ESR1* Y537S cellular populations (DM = 0.35, 95% CI = [0.11, 0.60], $t(4.00) = 4.03$, p -value = $1.6e-02$), whereas in the abemaciclib-resistant models, the distribution of the cellular populations after the acquisition of resistance between ER-WT and *ESR1* Y537S populations were comparable (DM = -0.09 , 95% UMCi = [$-0.44, 0.26$], $t(3.84) = -0.72$, p -value = $5.1e-01$) (DM, difference in means; 95% CI, 95% confidence interval; two-sided unequal-variances t-test) (Fig. 3H–I). Collectively, these results underscore the impact of the *ESR1* Y537S on clonal dynamics and the differences between acquired palbociclib versus abemaciclib resistance.

Whole exome sequencing reveals selection of genetically distinct cellular populations during the acquisition of CDK4/6i resistance

Since resistance to targeted therapies is often attributed to the selection of a driver mutation, such as the enrichment of *RB1* mutations observed in clinical samples of palbociclib resistance [5, 6, 71], we performed WES of the vehicle-control, and palbociclib or abemaciclib-resistant models. To enhance the sensitivity of this analysis, single cell clones were derived from palbociclib-resistant and vehicle-control populations. The single cell origin of the resistant clones was confirmed by barcode sequencing and ddPCR (Additional file 3: Fig. S4A–D). To search for single nucleotide variations (SNVs) and INDELs unique to the resistant cells, we compared the vehicle-control and palbociclib-resistant cells to the parental MCF7 cells. Although we detected several mutations unique to the vehicle-control and resistant cells, none of these are known to be pathogenic and are likely passenger mutations (Additional file 3: Fig. S5A and Additional file 7: Table S6). However, unsupervised hierarchical clustering of the cells based on the unique mutations resembled the barcode clustering: (i) the evolutionary replicates clustered together, (ii) ER-mutant and ER-WT abemaciclib-resistant cells clustered together, (iii) ER-WT palbociclib-resistant cells segregated from the ER mutant cells (Fig. 4A–B). These findings indicate that

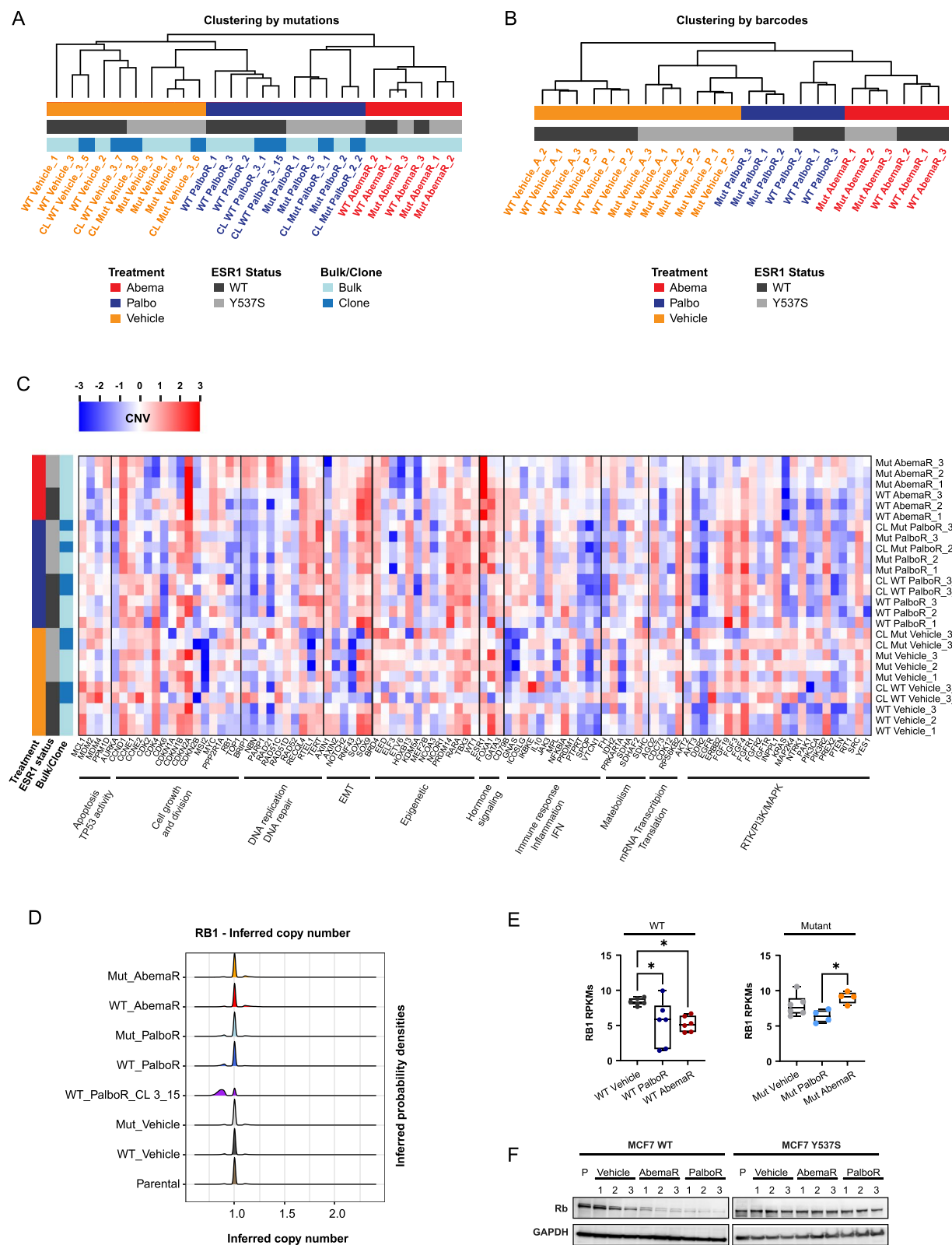


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Genetic landscape of ER-WT and ER-mutant cells resistant to palbociclib or abemaciclib. **A** Hierarchical clustering of ER-WT and Y537S ER-mutant palbociclib-resistant (PalboR), abemaciclib-resistant (AbemaR), and vehicle-control MCF7 cells, based on single-nucleotide variants (SNVs) and INDELs that are not detected in parental cells. The heatmap includes data from both the evolutionary replicates (bulk cells) and single-cell-derived clones ($N=8$). **B** Hierarchical clustering of ER-WT and Y537S ER-mutant PalboR, AbemaR, and vehicle-control MCF7 cells, based on the fraction of barcodes with a fraction ≥ 0.001 . The heatmap includes only data from bulk cells. Both vehicle-controls of PalboR (Vehicle_P) and AbemaR (Vehicle_A) are included. **C** Copy number variations (CNVs) of breast cancer related genes in ER-WT and Y537S ER-mutant PalboR, AbemaR and vehicle-control MCF7 cells. The analysis includes data from both bulk cell populations and single-cell-derived clones. The genes are listed based on their biological functions. **D** *RB1* CNV inferred from single-cell RNA-seq data of ER-WT and Y537S ER-mutant PalboR, AbemaR and vehicle-control MCF7 cells (evolutionary replicate #3 for each condition) and single-cell RNA-seq data of an ER-WT PalboR clone derived from the ER-WT PalboR evolutionary replicate #3 (clone #3_15). The plot displays the estimated distributions of *RB1* CNVs across all the cells in each sample. The x axis represents the relative inferred copy number per cell. CNV values less than one represent a loss and CNV values greater than one represent an amplification compared to the averaged parental reference sample. The y axis represents the inferred probability densities per sample. **E** *RB1* mRNA level in MCF7 ER-WT (left) and Y537S ER-mutant (right) PalboR, AbemaR, and vehicle-control cells, reported as Reads Per Kilobase per Million mapped reads (RPKM). The *RB1* mRNA level was estimated from bulk RNA-seq data. Each boxplot shows the quartile distribution of the *RB1* mRNA level in each experimental condition, with dots representing individual technical replicates. Statistical comparisons among conditions were performed using a one-way ANOVA followed by Tukey's multiple comparisons test (*, $p < 0.05$). **F** Protein expression of Rb in whole cell lysates from ER-WT and Y537S ER-mutant parental, vehicle-controls, PalboR and AbemaR MCF7 cells. Rb expression in parental cells was assessed after the cells were cultured in full medium with or without DOX for 72 h. GAPDH expression level was used as loading control. (WT PalboR: ER-WT PalboR; Mut PalboR: ER-mutant PalboR; WT AbemaR: ER-WT AbemaR; Mut AbemaR: ER-mutant AbemaR; WT Vehicle: ER-WT vehicle; Mut Vehicle: ER-mutant vehicle; numbers 1, 2 and 3 indicate distinct bulk evolutionary replicates; Clone (CL) IDs: the first number indicates the bulk cell population from which the clone was derived, while the second number identifies the specific clone from that population)

CDK4/6i-resistant cells acquire a reproducible genetic footprint across evolutionary replicates, consistent with convergent genetic selection, while also suggesting that resistance may arise through divergent evolutionary trajectories shaped in part by cellular plasticity.

To further search for genetic mechanisms of acquired resistance in addition to the *ESR1* mutation, we analyzed copy number variations (CNV) profiles in vehicle-control and resistant cells, normalized to the parental cell population (Fig. 4C). Because matched germline sequencing is not available, the analysis was restricted to cancer genes frequently altered in metastatic breast cancer [72] (Additional file 7: Table S6). We did not identify CNVs that consistently explained CDK4/6i resistance across all models. However, *RB1* loss was detected in an ER-WT palbociclib-resistant clone (WT PalboR_clone 3_15). The barcode corresponding to this clone was rare in the parental cell population (fraction = 4.74×10^{-6}) but was enriched across all ER-WT palbociclib-resistant replicates: its barcode fraction was 0.40 in the parental ER-WT palbociclib-resistant population (evolutionary replicate #3) and 0.33 and 0.39 in evolutionary replicates #1 and #2, respectively. Importantly, this barcode was not enriched in the other resistant cells (Additional file 4: Table S3).

Copy number inference from single cell RNA-seq (scRNA-seq) independently confirmed *RB1* loss in the palbociclib-resistant clone 3_15 (Fig. 4D and Additional file 3: Fig. S5B). Bulk RNA-seq showed transcriptional downregulation of *RB1* in the ER-WT palbociclib and abemaciclib-resistant models (Fig. 4E). Consistently, Rb protein level was reduced in these models (Fig. 4F). These findings underscore the value of sequencing individual resistant clones or applying deep sequencing approaches to detect genetic mechanisms of resistance that may be missed by bulk analyses and suggest that *RB1* alterations

may be more prevalent but under-detected due to technical limitations. Furthermore, *RB1* genomic loss and reduced Rb expression were observed only in ER-WT cells, consistent with a report suggesting a trend towards mutual exclusivity between *ESR1* and *RB1* mutations [5].

Epigenetic and transcriptional plasticity in the acquisition of CDK4/6i resistance

Our data suggests that resistance to CDK4/6is is a multifactorial process involving cooperating convergent and divergent trajectories, this is consistent with recent studies showing that drug resistance involves genetic and epigenetic adaptive changes [68, 73]. To characterize resistant cell states and adaptive transcriptional programs while assessing the impact of the Y537S *ESR1* mutation and the specific CDK4/6i, we performed ATAC-seq and RNA-seq in all cell models. Principal components analysis (PCA) of the open chromatin sites clustered the evolutionary replicates and separated the CDK4/6i-resistant cells from the vehicle-control cells. Among palbociclib-resistant cells, ER-WT and ER-mutant further segregated, whereas the abemaciclib-resistant ER-WT and ER-mutant cells clustered together (Fig. 5A). These patterns were consistent with the barcode-based clustering (Fig. 4B), supporting marked differences in epigenetic states between ER-mutant and ER-WT palbociclib-resistant cells, but less distinction in the abemaciclib-resistant ER-WT and ER-mutant models.

PCA of bulk RNA-seq data separated cells by treatment condition (palbociclib resistance, abemaciclib resistance, and vehicle treatment) and within each treatment group (ER-WT cells separated from ER-mutant cells) (Fig. 5B). Unlike the ATAC-seq data, where abemaciclib-resistant cells with WT-ER and mutant-ER showed no significant differences in chromatin accessibility, RNA-seq revealed distinct transcriptomic profiles in these models.

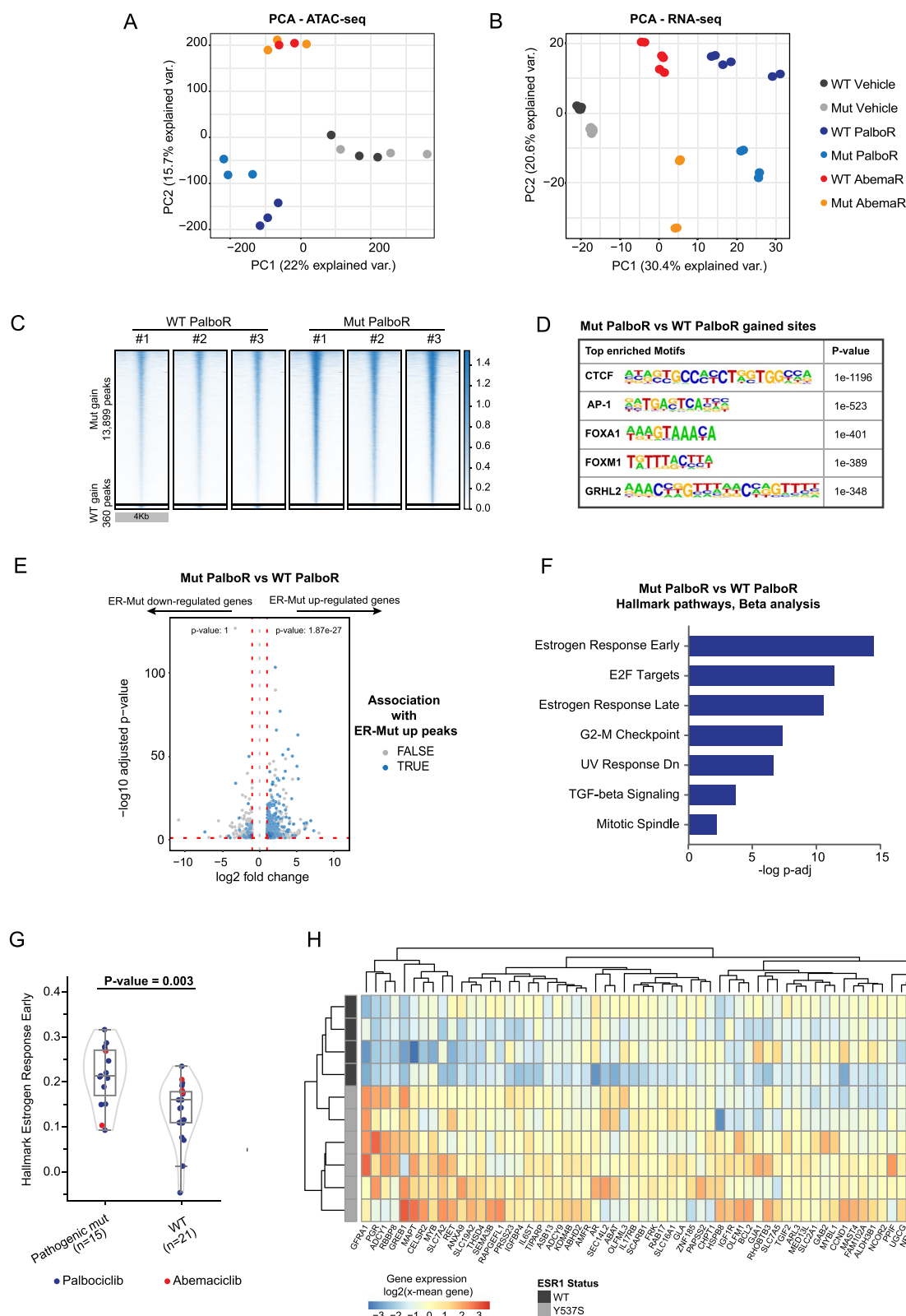


Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 Chromatin accessibility and transcriptional profiles of ER-WT and ER-mutant cells resistant to palbociclib or abemaciclib. **A** Principal component analysis (PCA) plot of bulk ATAC-seq data from ER-WT and Y537S ER-mutant palbociclib-resistant (PalboR), abemaciclib-resistant (AbemaR), or vehicle-control MCF7 cells. **B** PCA plot of bulk RNA-seq data from ER-WT and Y537S ER-mutant PalboR, AbemaR, or vehicle-control MCF7 cells. **C** Tornado plots of differential accessible chromatin sites ($\log_2FC < -1$, $\log_2FC > 1$; $\text{padj} < 0.05$, DESeq2) identified by comparing Y537S ER-mutant vs ER-WT PalboR MCF7 cells. The x-axis shows the width of differentially accessible genomic regions centered on a peak (± 2 kb), the y-axis shows the detected differentially accessible genomic regions, ranked by signal intensity. Numbers 1, 2 and 3 on top of the plot indicate the bulk evolutionary replicates. **D** Top significant motifs enriched in chromatin regions with increased accessibility in Y537S ER-mutant vs ER-WT PalboR MCF7 cells identified by HOMER. **E** Association between differentially expressed genes ($\log_2FC < -1$, $\log_2FC > 1$; $\text{padj} < 0.05$, DESeq2) and chromatin sites with increased accessibility ($\log_2FC > 1$; $\text{padj} < 0.05$, DESeq2) in Y537S ER-mutant vs ER-WT PalboR MCF7 cells. The volcano plot shows the differential gene expression between Y537S ER-mutant vs ER-WT PalboR MCF7 cells, with each dot representing a gene. Blue dots (true) indicate genes that are differentially expressed and are associated with chromatin sites with increased accessibility. Grey dots (false) indicate genes that are differentially expressed but are not associated with chromatin sites with increased accessibility. The p -values indicate the significance of the association between increased chromatin accessibility by ATAC-seq and up or down gene expression by RNA-seq, calculated by Binding and Expression Target (BETA) basic analysis. **F** Analysis of the Hallmark gene sets associated with chromatin regions with increased accessibility and upregulated in Y537S ER-mutant versus ER-WT PalboR MCF7 cells based on BETA analysis. **G** Expression of Hallmark ER-response early pathway in biopsies from patients included in our cohort described in Fig. 1A. The analysis was performed on biopsies of ER-WT and ER-mutant metastatic tumors that acquired resistance to palbociclib (blue dots) or abemaciclib (red dots). Only ER-mutant biopsies obtained within 30 days of treatment discontinuation due to progressive disease were included. Samples with non-pathogenic *ESR1* mutations were excluded. ER-WT biopsies = 15. Y537S ER mutant biopsies = 21. Comparison between ER-mutant and ER-WT biopsies was performed by Wilcoxon test. **H** Unsupervised heatmap of metastatic breast cancer samples from patients with acquired resistance to palbociclib [33], based on the expression of the leading-edge ER-response early genes upregulated in the Y537S ER-mutant vs ER-WT MCF7 PalboR cell models. The expression level of the genes is shown in the heatmap

This discrepancy highlights the distinction between the chromatin state, a measure of transcriptional potential, and the plasticity of the transcriptional output, which is shaped by additional layers of gene regulation.

To further delineate the differences between CDK4/6i resistance in ER-mutant and ER-WT cells, we compared chromatin accessibility between ER-mutant and ER-WT palbociclib-resistant models [41]. Relative to ER-WT resistant cells, ER-mutant lost 360 accessible sites and gained 13,899 sites ($\log_2FC < -1$, $\log_2FC > 1$; $\text{padj} < 0.05$) (Fig. 5C) (Additional file 8: Table S7). Gained sites were enriched for CTCF, AP-1, FOXA1, FOXM1, and GRHL2 motifs, consistent with activation of transcription factor programs linked to the ER transcriptional axis [74–78] (Fig. 5D). Differential gene expression analysis comparing the ER-mutant vs ER-WT palbociclib-resistant cells identified 206 upregulated genes and 125 downregulated genes ($\log_2FC < -1$, $\log_2FC > 1$; $\text{padj} < 0.05$) (Additional file 6: Table S5) [41]. Applying the Binding and Expression Target Analysis (BETA algorithm) [67], we observed a significant association between the gained accessible sites and the upregulated genes (Fig. 5E) (Additional file 8: Table S7). Pathway analysis of genes upregulated in the ER-mutant cells compared to the ER-WT counterparts and associated with the gained accessibility, indicated predominant enrichment of estrogen response and cell proliferation related genes (Fig. 5F).

To assess the clinical relevance, we compared transcriptomes of the metastatic breast cancer biopsies harboring pathogenic *ESR1* mutations ($N=15$) to those with *ESR1* WT ($N=21$) from the cohort described above, including the subgroup of acquired resistance to CDK4/6is (Fig. 5G). This analysis showed the upregulation of the Hallmark Estrogen Response Early pathway in tumor biopsies with pathogenic *ESR1* mutations

(p -value = 0.003, Wilcoxon test), consistent with the cell line data. As an additional validation in a smaller sample set of ER+ and HER2-negative metastatic breast cancer biopsies with acquired resistance to palbociclib from a clinical trial [33], leading-edge Hallmark Estrogen Response early genes upregulated in ER-mutant versus ER-WT palbociclib-resistant cells were similarly upregulated in the tumors harboring a pathogenic *ESR1* mutation compared to the ER-WT tumors (Fig. 5H). Collectively, these results indicate that in palbociclib resistance *ESR1* mutations are associated with a unique chromatin state linked to increased ER signaling.

To investigate chromatin-state changes during the acquisition of resistance to CDK4/6is, we compared the chromatin accessibility in resistant cells versus their vehicle-control counterparts (Additional file 8: Table S7) [41]. Compared to the vehicle-treated cells, both ER-WT and mutant palbociclib-resistant cells predominantly gained accessible sites, with relatively few sites lost (Additional file 3: Fig. S6A–B). The gained sites were enriched for motifs of transcription factors that are known to interact with the ER transcriptional axis, including AP-2 gamma, FOXA1 and GRHL2 [77–79] (Additional file 3: Fig. S6C–D), further supporting a central role for the ER axis in palbociclib treatment. Chromatin reprogramming was also observed in the abemaciclib-resistant models (Additional file 3: Fig. S6E–F). Motif analysis of gained accessibility in abemaciclib-resistant versus vehicle-control revealed the enrichment of motifs related to interferon and immune signaling, such as NF- κ B, IRF and AP-1 [80] (Additional file 3: Fig. S6G–H). Across all four resistant versus vehicle-control counterpart comparisons, BETA analysis identified a significant association between gained accessible sites and transcriptionally upregulated genes, supporting functional consequences of chromatin

remodeling (Additional file 3: Fig. S6I-L). Overall, these findings underscore the epigenetic and transcriptional plasticity as an adaptive mechanism during the acquisition of CDK4/6i resistance, while also revealing distinct adaptive states in palbociclib versus abemaciclib resistance.

Single cell analysis reveals transcriptionally heterogeneous cellular populations

Because the lineage tracing revealed enrichment of distinct cellular subpopulations in vehicle and CDK4/6i treated cells, we posited that both parental and resistant cells would display single cell transcriptional heterogeneity. To characterize this aspect of cell adaptation during acquired resistance, we performed single cell RNA-seq (scRNA-seq) analyses. Low dimensional embeddings separated cell models apart from the parental and ER-WT vehicle-treated cells, which overlapped (Fig. 6A). Interestingly, the intra-sample cell-cell Euclidean distance was broadly similar across models, including the cell line derived from a single cell (WT_PalboR_clone 3_15) (Additional file 3: Fig. S7A), consistent with transcriptional plasticity during passaging. Within each model, cell cycle associated genes partitioned cells into distinct subpopulations with variable distributions of the cells across cell cycle phases (Fig. 6B-C). Therefore, to identify transcriptional programs underlying adaptive states that support CDK4/6i resistance, we performed pathway analyses comparing the CDK4/6i-resistant models to their corresponding vehicle-treated counterpart, accounting for the cell cycle phase (Fig. 6D). We identified several consistently upregulated pathways: EMT was elevated in palbociclib-resistant models, whereas interferon-gamma and interferon-alpha response pathways were upregulated in abemaciclib-resistant models. Additionally, early and late estrogen response pathways were consistently upregulated only in the ER-mutant palbociclib-resistant model. Marker gene analysis highlighted mediators of treatment resistance, such as high expression of *CCNE1* in ER-mutant abemaciclib-resistant cells and increased expression of *MDM2*, *MDM4* and *IGF1R* in WT-ER abemaciclib-resistant cells (Additional file 3: Fig. S7B). Direct comparisons between ER-mutant and ER-WT resistant models further supported the transcriptional impact of mutant ER on CDK4/6i resistance, with distinct effects in palbociclib versus abemaciclib resistance (Additional file 3: Fig. S7C).

We next examined the scRNA-seq data of each cell model and observed structured transcriptional heterogeneity. These patterns were independent of the cell cycle states, evidence for biological meaningful transcriptional heterogeneity (Fig. 6E-J and Additional file 3: Fig. S7D-E). Pathway analysis comparing each cluster to the remaining cells within each model indicated that these transcriptional states were biologically distinct. As an example, in

the ER-mutant palbociclib-resistant cells one cluster was enriched for TNF-alpha signaling, a second was enriched in MYC targets and cell proliferation pathways, and a third was enriched in ER response pathways. Similarly, the cell line derived from a single palbociclib-resistant cell harboring an *RBI* deletion (Fig. 6H) consisted of three transcriptionally defined cell states; cluster 1 had higher expression of genes related to ER, TNF-alpha and TGF-beta, cluster 2 exhibited higher expression of genes involved in EMT, and cluster 3 was enriched in genes of the mTOR signaling and cell proliferation pathways. Importantly, these differences remained evident even after accounting for the cell cycle phase (Additional file 3: Fig. S8). In aggregate, our data suggests that transcriptional heterogeneity and plasticity may enable cells to adapt to treatment or other stressors. We showed that even in a population derived from a single cell harboring a defined genetic mechanism of resistance there is transcriptional heterogeneity, consistent with the coexistence of genetic and epigenetic mechanisms of resistance.

Mutant ER shapes acquisition but is dispensable for maintenance of palbociclib resistance, while ER loss drives resistance in WT-ER

Our results show that the Y537S *ESR1* mutation is retained after acquisition of CDK4/6i resistance and impacts clonal evolution and cells state programs, with more pronounced effects in palbociclib compared to abemaciclib resistance. We next asked whether *ESR1* Y537S is required to maintain palbociclib resistance. Silencing of the mutant transgene by DOX withdrawal did not diminish growth in the presence of palbociclib, indicating that mutant *ESR1* is not essential to maintain palbociclib resistance, and that additional pathways sustain the resistance state (Fig. 7A). This is consistent with clinical observations that many patients do not benefit from single-agent oral SERD after progression on a CDK4/6i, and that targeting other pathways is necessary [81, 82].

Because WT-ER resistant cells showed reduced ER expression, we next tested whether loss of ER is required for resistance. First, we induced the expression of the *ESR1* mutant trans-gene in the palbociclib-resistant ER-WT cells treating them with DOX. This induction restored sensitivity to palbociclib, suggesting that ER loss is important for the palbociclib resistance (Fig. 7B). To distinguish whether restored sensitivity depends specifically on mutant ER, which regulates canonical and non-canonical ER transcriptional targets [13], or restoration of ER signaling more broadly, we rescued ER-WT expression in the ER-WT palbociclib-resistant cells, which showed loss of ER (Fig. 7C). This was sufficient to restore palbociclib sensitivity, supporting ER loss as a key driver of palbociclib resistance in the setting of WT-ER (Fig. 7D-E).

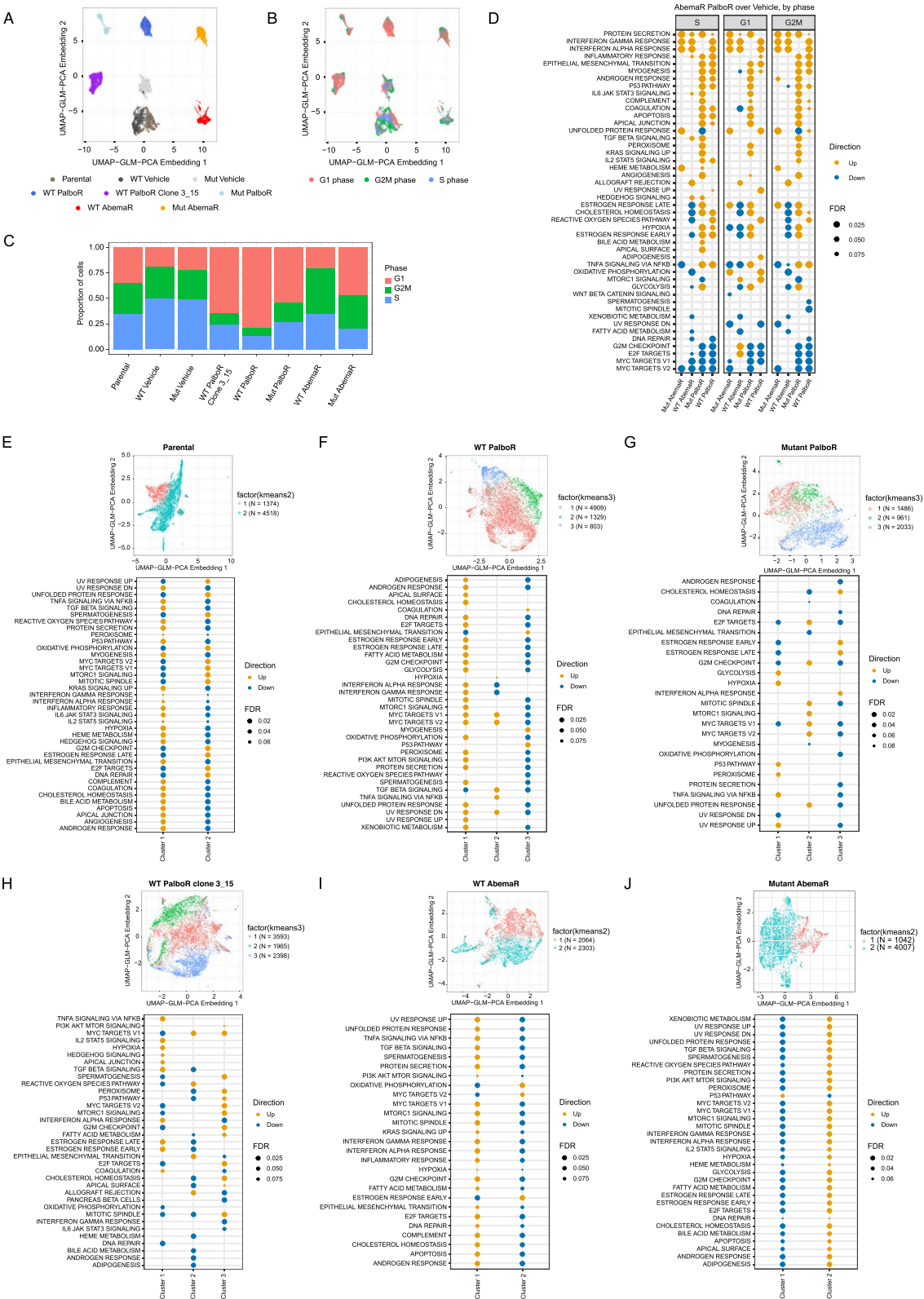


Fig. 6 (See legend on next page.)

(See figure on previous page.)

Fig. 6 Transcriptional heterogeneity of ER-WT and ER-mutant parental, palbociclib-resistant, abemaciclib-resistant and vehicle-control MCF7 cells. **A–B** UMAP showing the sample clustering of ET-WT parental, ER-WT and Y537S ER-mutant palbociclib-resistant (WT PalboR and Mut PalboR), ER-WT and Y537S ER-mutant abemaciclib-resistant (WT AbemaR and Mut AbemaR), ER-WT and Y537S ER-mutant vehicle-control (WT Vehicle and Mut Vehicle) MCF7 cells (evolutionary replicates #3) and ER-WT palbociclib-resistant clone 3.15 (WT_PalboR Clone 3_15). In the UMAP shown in **A** colors represent different samples, while in the UMAP shown in **B** colors represent the cell cycle phases. **C** Proportion of cells in each of the cell cycle phases within each cell model. **D** Gene set testing of Hallmark pathways comparing ER-WT and Y537S ER-mutant palbociclib-resistant (PalboR) and abemaciclib-resistant (AbemaR) models to the vehicle-treated counterparts, within each cell cycle phase separately. Multiple testing was adjusted for each tested condition comparison (column) using the Benjamini Hochberg procedure. Gene sets with an $FDR \leq 10\%$ are plotted. **E–J** Cellular sub-populations identified within ER-WT parental cells (**E**), ER-WT PalboR cells (**F**), Y537S ER-mutant PalboR cells (**G**), ER-WT PalboR 3.15 clone (**H**), ER-WT AbemaR cells (**I**) and ER-mutant AbemaR cells (**J**). The top panels show the UMAP visualization of the sub-populations detected in each cell model. The bottom panels show the results of gene set testing of Hallmark pathways comparing each sub-population to the others. Multiple testing was adjusted for each tested cluster comparison (column) using the Benjamini Hochberg procedure to control the false discovery rate. Gene sets with an $FDR \leq 10\%$ are plotted

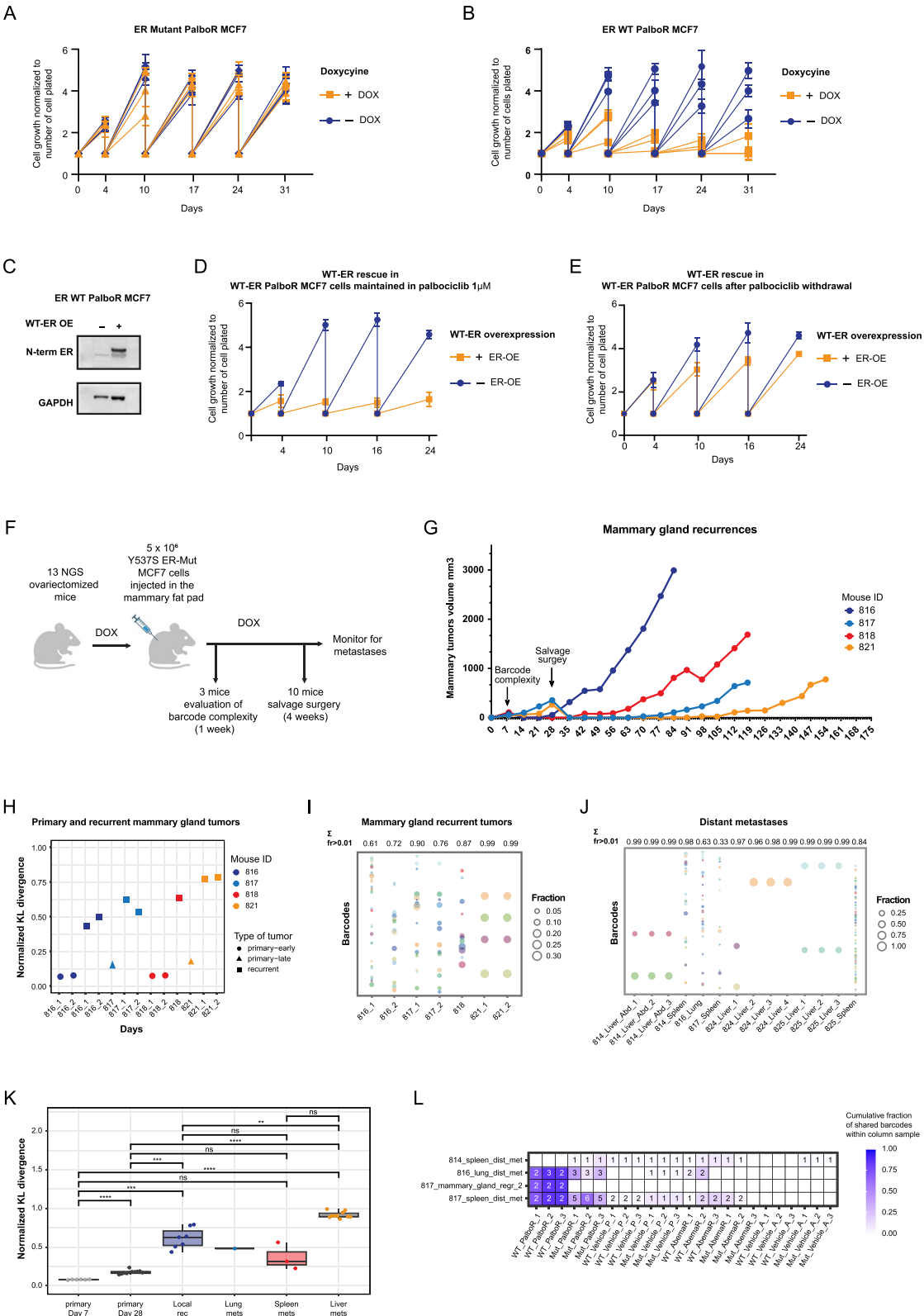
Subpopulations with a shared fitness for resistance and metastases

We previously demonstrated that activating *ESR1* mutations drive metastases in vivo [13]. To place these findings within the same clonal-dynamics framework used for our CDK4/6 inhibitor resistance studies, we profiled barcode selection in mammary gland xenografts, local recurrences, and distant metastases. We also asked whether the cellular populations that expand during distant metastases in vivo overlap with those selected during the development of CDK4/6i resistance in vitro, which would suggest shared fitness advantages across these evolutionary contexts. MCF7 cells stably expressing the DOX inducible Y537S *ESR1* mutation and the ClonTracer library were injected into the mammary fat pads of ovariectomized mice and mice were treated with DOX. Tumors from three mice were harvested at 1 week to assess the barcode complexity, while tumors from the remaining ten mice were resected after 4 weeks (salvage surgery) (Fig. 7F). Mice were then monitored for local and distant recurrences. Barcode sequencing of the 1-week tumors confirmed an even distribution of the barcodes with a barcode complexity of more than 80,000 unique barcodes (average of unique barcodes = 81,493, 95% CI = 75,949 to 87,937) (Additional file 4: Table S3). We confirmed that the tumors contained barcodes that were found to be enriched in the CDK4/6i-resistant cells (Additional file 3: Fig. S9A–D), indicating that the clones capable of expanding under therapeutic pressure were present within the initial population. By 4 weeks, the mammary xenografts showed substantial clonal selection, with ~30,000 unique barcodes remaining (average of unique barcodes = 30,413, 95% CI = 26,651 to 34,176) (Additional file 4: Table S3).

We next analyzed local recurrences and distant metastases. The growth rates and clonal selection patterns of the mammary gland recurrent tumors were highly variable (Fig. 7G–I). For example, a slow growing tumor that emerged after day 105 (#821) exhibited strong clonal selection with high normalized KL divergence (#821_1: normalized KL = 0.78, #821_2: normalized KL = 0.79) and was dominated by four barcodes contributing to the tumor, while a fast-growing tumor (#816) had a

lower normalized KL divergence (#816_1: normalized KL = 0.44, #816_2: normalized KL = 0.51). The clonal diversity was variable across distant metastatic lesions and were dependent on the site of the metastasis (Fig. 7J–K). Liver metastases exhibited strong enrichment of 1–2 dominant barcodes, consistent with high normalized KL divergence compared to primary-early (DM = -0.84, 95% UMCI = [-0.87, -0.81], $t(9.22) = -62.54$, $\text{padj} < 1e-06$) and compared to primary-late (DM = -0.74, 95% UMCI = [-0.77, -0.71], $t(14.67) = -47.78$, $\text{padj} < 1e-06$) samples (DM: difference in means, 95% UMCI: 95% unadjusted marginal confidence interval, padj : adjusted p -value, two-sided unequal-variances t -test with each comparison adjusted for multiple testing with the Holm procedure). In contrast, spleen and lung metastases retained more diverse barcode representation and lower normalized KL divergence values. Spleen metastases had KL divergences closer to primary-early (DM = -0.29, 95% UMCI = [-0.72, 0.14], $t(2.00) = -2.87$, $\text{padj} = 3.1e-01$) and primary-late samples (DM = -0.19, 95% UMCI = [-0.62, 0.24], $t(2.03) = -1.90$, $\text{padj} = 3.1e-01$).

To determine whether the clonal populations selected in vivo were related to those enriched during drug resistance, we compared barcodes enriched in the CDK4/6i-resistant cell lines with those detected in recurrent and metastatic tumors. We observed significant overlap between barcodes enriched in palbociclib-resistant ER-WT cells and those found in the local recurrent tumor and spleen metastases from mouse #817, as well as lung metastases from mouse #816. We also observed overlap between barcodes enriched in ER-mutant palbociclib-resistant cells and those present in spleen metastases from mouse #817 and lung metastases from mouse #816 (Fig. 7L and Additional file 3: Fig. S9E). Together, these data extend our clonal-dynamics framework from in vitro resistance to in vivo progression, showing that tumor evolution is shaped by both context-specific selection and shared clonal fitness programs. In this model, genomic resistance-associated changes may contribute to the fitness of specific clones, while the metastatic niche further constrains which barcoded populations ultimately expand, thereby linking genetic evolution, clonal selection, and organ-specific outgrowth.



(See figure on previous page.)

Fig. 7 In vivo clonal dynamics in Y537S *ESR1* mutant metastases. **A** Cell growth of palbociclib-resistant (PalboR) doxycycline (DOX)-inducible Y537S ER-mutant MCF7 cells maintained with DOX (yellow) or after DOX withdrawal (blue). Growth is normalized to the number of cells plated back at each passage (2 million cells). **B** Cell growth of PalboR ER-WT MCF7 cells maintained without DOX (blue) or after DOX treatment (yellow) to induce the expression of the Y537S mutant ER. Growth is normalized to the number of cells plated back at each passage (2 million cells). **C** Whole cell lysate western blotting for ER in WT-ER PalboR cells characterized by loss of endogenous ER without exogenous overexpression (OE) of WT-ER and with exogenous OE of WT-ER. **D** Cell growth of PalboR ER-WT MCF7 cells with (yellow) or without (blue) WT-ER OE. Growth is normalized to the number of cells plated back at each passage (2 million cells). **E** Cell growth of PalboR ER-WT MCF7 cells with (yellow) or without (blue) WT-ER OE, after removal of palbociclib. Growth is normalized to the number of cells plated back at each passage (2 million cells). **F** Schema of the in vivo experiment. Ovariectomized NSG mice were implanted with DNA barcoded (ClonTracer) MCF7 cells stably expressing a DOX-inducible Y537S *ESR1* mutation. Eight days after cell injection, three mice underwent surgery to remove the mammary fat pad tumor and assess the complexity of the barcoded cell population. The remaining ten mice underwent survival surgery 28 days after cell injection and were monitored for local tumor regrowth and metastases development. **G** Tumor growth of local recurrences that developed in 4 mice after resection of primary tumors at week 1 (resection to test the barcode complexity) or at week 4 (salvage resection). **H** Normalized Kullback–Leibler (KL) divergence of barcode distribution against a discrete uniform distribution, computed for each available sample of local recurrences (squares) and the respective primary-early tumor (primary tumor resected at week 1; dots) or primary-late tumor (primary tumor resected at week 4; triangles). Lower values indicate more similar distributions and higher values indicate more divergent distributions. Colors represent the mouse number. Sample IDs: the first number indicates the mouse, while the second number identifies different samples from the same local recurrence or primary tumor. If no second number is present, only one sample was available. **I–J** Bubble plots showing the enrichment of barcodes with a fraction > 0.01 in each sample of local recurrences (**I**) and in metastases (**J**). Each color represents a unique enriched barcode. The size of the circle is proportional to the fraction of the barcode within a sample. The cumulative fraction of barcodes with a fraction > 0.01 is indicated above each sample. Sample IDs: the first number indicates the mouse, while the second number refers to different samples of the same local recurrence (**I**) or distant metastasis (**J**). If no second number is present, only one sample was available. **K** Normalized KL divergence of barcode distribution against a discrete uniform distribution in primary tumors (after surgery at week 1 or at week 4), local recurrences and in lung, spleen and liver distant metastases. Lower values indicate more similar distributions, and higher values indicate more divergent distributions. Statistical comparisons among conditions were performed using a t-test adjusting for multiple testing using the Holm procedure (n.s., not significant; *, p-adjusted ≤ 0.05; **, p-adjusted ≤ 0.01; ***, p-adjusted ≤ 0.001; ****, p-adjusted ≤ 0.0001). **L** Contingency plot showing the overlap of barcodes with a fraction > 0.01 between evolutionary replicates of ER-WT or Y537S ER-mutant palbociclib-resistant (PalboR), abemaciclib-resistant (AbemaR) or vehicle-treated cell models and distant metastases. The number reported in each square represents the number of barcodes shared between two samples. The color intensity reflects the cumulative fraction of the shared barcodes within each cell model. (WT PalboR: ER-WT PalboR; Mut PalboR: ER-mutant PalboR; WT AbemaR: ER-WT AbemaR; Mut AbemaR: ER-mutant AbemaR; WT Vehicle_P: ER-WT vehicle control of ER-WT PalboR cells; WT Vehicle_A: ER-WT vehicle control of ER-WT AbemaR cells; Mut Vehicle_P: ER-mutant vehicle control of ER-mutant PalboR cells; Mut Vehicle_A: ER-mutant vehicle control of ER-mutant AbemaR cells)

Discussion

Multiple mechanisms of resistance to CDK4/6is in ER+ breast cancer have been identified [83, 84]. These include activation of upstream effectors, downstream bypass mechanisms and alterations in the CDK4/6-cyclin D1-Rb axis. Here, we combined high-complexity barcoding with longitudinal sampling and multi-omic profiling to resolve how clonal selection and cell plasticity unfolds during CDK4/6i resistance, how this process is shaped by an *ESR1* mutation, and whether different CDK4/6is drive distinct evolutionary outcomes. Three key insights emerge from our study: (i) *ESR1* mutations are enriched in clinical samples of acquired CDK4/6i resistance and can expand to near clonality, (ii) *ESR1* Y537S measurably reshapes clonal and epigenetic evolution under palbociclib but has a weaker impact under abemaciclib, and (iii) palbociclib and abemaciclib impose distinct selective pressures leading to divergent resistant states with limited clonal overlap and asymmetric cross-resistance.

The ClonTracer profiling demonstrated marked selection of barcodes and a decrease in clonal diversity under drug pressure. However, evolutionary replicate overlap of enriched barcodes was incomplete, indicating that multiple evolutionary routes to resistance can co-exist with potentially multiple collateral sensitivities. At the same time, replicate clustering and within-condition similarity suggested convergence that can arise from

biological constraints influenced by the drug and genetic background. Importantly, even within dominant resistant clones, single-cell transcriptomic analyses reveal substantial transcriptional heterogeneity, including coexisting cell-cycle, estrogen-responsive, EMT-like, and stress-associated programs. This heterogeneity persists even in populations derived from a single resistant cell, highlighting cell plasticity and the divergence in CDK4/6i resistance. Additionally, increasing evidence suggests that therapeutic resistance does not arise through a binary transition from drug sensitivity to resistance but rather emerges along a continuum of progressively adaptive cellular states shaped by selective pressure [68]. Our findings strongly support this model in the context of CDK4/6i resistance. Longitudinal sequencing of the high-complexity DNA barcoding demonstrated that resistance is driven by progressive enrichment barcodes. Together, these observations align with a model in which genetic predisposition and transcriptional plasticity cooperate to enable incremental gains in fitness along a resistance landscape, challenging binary sensitive/resistant classifications and underscoring the dynamic, multifactorial nature of acquired CDK4/6i resistance.

The role of the Y537S *ESR1* mutation in ET resistance is well established [11]. There is also evidence that *ESR1* mutations have a role in resistance to the combination of ET and palbociclib [6, 17, 18]. However, it is unknown

if this effect is due to ET resistance only or whether the *ESR1* mutations also contribute to resistance to CDK4/6is. Our study indicates that the expression of the Y537S *ESR1* mutation, which is the activating *ESR1* mutation with the most robust phenotypes [13, 16], influences cell adaptation, selection pressures and clonal evolution under treatment, but it is not required to maintain resistance once resistance is established. Specifically, Y537S *ESR1* mutation is associated with sustained mutant ER expression and broader clonal diversity in palbociclib-resistant models, suggesting reduced selection pressure for subclones harboring other drivers of resistance. In contrast, its impact on abemaciclib resistance was less pronounced, with similar clonal diversity between ER-WT and ER-mutant models. These two drugs also have distinct clinical properties, including their side effect profiles [85], the demonstration of single agent activity with abemaciclib [86], and the ability of abemaciclib to reduce disease recurrence in the adjuvant setting [87]. Both our findings and these clinical observations likely reflect differences in the inhibitory profiles of the two drugs. Palbociclib is a relatively selective CDK4/6 inhibitor, whereas abemaciclib exhibits greater affinity for CDK4, and displays broader activity against additional CDKs, including CDK1, CDK2 and CDK9 inhibition [19]. Notably, because CDK9 signaling does not directly converge with the ER axis, its inhibition by abemaciclib may explain the diminished influence of mutant ER on the acquisition of resistance to abemaciclib compared with palbociclib resistance. Furthermore, the barcode selection differed between palbociclib and abemaciclib irrespective of the ER mutant status, highlighting distinct resistance mechanisms. This effect is further supported by our observation that abemaciclib-resistant cells exhibit cross-resistance to palbociclib, whereas palbociclib-resistant cells remain sensitive to abemaciclib, albeit with higher IC_{50} values. These findings align with a previous pre-clinical study [88] and the post-MONARCH clinical trial, which demonstrated clinical benefit from abemaciclib treatment after progression on palbociclib treatment suggesting non-overlapping resistance mechanisms between these inhibitors [89].

Our integrated ATAC-seq and RNA-seq profiling revealed therapy-induced shifts in chromatin accessibility and gene expression programs, underscoring an interplay between genetic and epigenetic mechanisms of resistance. Principal component analysis of ATAC-seq data separated drug-resistant cells from vehicle-treated controls, indicating widespread chromatin remodeling during resistance. In the palbociclib-resistant condition, Y537S ER-mutant cells formed a distinct epigenetic cluster apart from ER-WT clones, whereas abemaciclib-resistant cells showed a more consolidated chromatin state regardless of *ESR1* mutational status. This finding mirrors the clonal barcode patterns and suggests that the *ESR1* mutation drives unique

chromatin reprogramming under palbociclib pressure that is not observed with abemaciclib. In addition, the changes in chromatin state further underscore the cell plasticity during resistance and the importance of targeting multiple pathways after CDK4/6i resistance.

The complexity of clonal evolution was further evident when we extended our barcode analysis to an in vivo setting. Using barcoded MCF7 cells implanted as mammary fat pad xenografts, we examined how clones evolved during tumor growth, regression, and metastasis in the presence of an active Y537S mutant ER. Even in the absence of drug treatment, we found substantial clonal selection during tumor progression and spread. The dependence of clonal dominance on metastatic site, as demonstrated by the strong clonal selection in the liver versus greater diversity in lung and spleen, suggests organ-specific constraints on colonization. Overlap between barcodes enriched in palbociclib-resistant states in vitro and those enriched in select recurrent and metastatic lesions suggest that some subpopulations have shared fitness advantages relevant to both therapeutic resistance and metastatic competence.

Our study has a few limitations. We used an engineered inducible *ESR1* Y537S system enabling controlled lineage tracing, but this system cannot capture all biologically relevant selective pressures. Specifically, interactions between cancer cells and the tumor micro-environment cells are not accounted for. Future studies should extend lineage tracing to in vivo drug treatment in immune competent models. Since the barcodes in the ClonTracer library are not expressed we could not link all the enriched subpopulations with their specific transcriptional profiles. In addition, given the relatively small number of evolutionary replicates, we cannot definitively rule out the possibility that the observed convergence is influenced, at least in part, by sampling bias.

Conclusions

Our study shows that in ER+ breast cancer the evolution of resistance to CDK4/6is is jointly shaped by clonal selection and adaptive state remodeling, supporting the need for combination treatments after the acquisition of resistance. The Y537S *ESR1* mutation and the specific CDK4/6is are both major determinants of the evolutionary resistance trajectories. The *ESR1* Y537S mutation promotes distinct palbociclib-associated chromatin and transcriptional programs during acquisition of resistance but becomes dispensable for resistance maintenance once alternative sustaining pathways are established. Palbociclib and abemaciclib drive distinct evolutionary outcomes with limited clonal overlap and impose different selective landscapes. These findings underscore the need to consider ER mutational status and inhibitor-specific evolutionary pressures when designing sequential approaches and combination strategies.

Abbreviations

AbemaR	Abemaciclib-resistant
ATAC-seq	ATAC sequencing
AVMA	American veterinary medical association
CCF	Cancer cell fraction
CDK4/6i	CDK4/6 inhibitors
CNV	Copy number variation
ddPCR	Droplet digital PCR
DOX	Doxycycline
EMT	Epithelial to mesenchymal transition
ER	Estrogen receptor
ER_OE	ER overexpression
ET	Endocrine therapy
FFPE	Formalin-fixed, paraffin embedded
HA	Hemagglutinin
INDEL	Insertion-deletion
IRB	Institutional review board
JS	Jensen-Shannon
KL	Kullback-Leibler
LBD	Ligand binding domain
MOI	Multiplicity of infection
MRI	Magnetic resonance imaging
NSG	NOD SCID Gamma
OCT	Optimal cutting temperature
PalboR	Palbociclib-resistant
PCA	Principal component analysis
RNA-seq	RNA sequencing
RPKM	Reads per kilobase million
scRNA-seq	Single cell-RNA sequencing
SERD	Selective estrogen receptor degrader
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variation
TPM	Transcripts per million
UMAP	Uniform manifold approximation and projection
VAF	Variant allele fraction
WES	Whole exome sequencing
WT	Wild-type

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-026-01690-2>.

- Additional file 1: Table S1. Patient samples characteristics.
- Additional file 2: Table S2. Methods barcoding library and ddPCR.
- Additional file 3: Supplementary figures (Figure S1–S7).
- Additional file 4: Table S3. Barcoding data.
- Additional file 5: Table S4. IC50 data.
- Additional file 6: Table S5. RNA-seq data.
- Additional file 7: Table S6. WES.
- Additional file 8: Table S7. Annotated ATAC-seq data.

Acknowledgements

We thank the members of our laboratories for their critical inputs, in particular Shengqing Stan Gu and Xiaoqing Wang for discussions on the DNA barcoding system, Rafael A. Irizarry, Phillip B. Nicol, and Jeremy M. Simon for discussions on scRNA-seq methodology, and Mendy Miller, Kiara Pontious and Brian Danysh for discussions on the whole exome sequencing analyses.

Authors' contributions

R. Jeselsohn, M. Brown, F. Michor and C. Guarducci designed the study. R. Jeselsohn, M. Brown, F. Michor, K. Polyak, P. Sicinski, G. Getz, C. Stewart, C. Paweletz, H. Long, N.U. Lin, S.M. Tolane and L. Malorni provided senior supervision for the study. C. Guarducci, Z. Nagy, A. Nardone, Y. Kuang, W. Ma, G.C. Feit, F. Hermida-Prado, C. Heraud, M. Munoz-Gomez and V. Rao performed all the in vitro experiments under the supervision of R. Jeselsohn, C. Paweletz

and H. Long. C. Guarducci and A. Nardone processed all the samples collected from mice. D.L. Abravanel, J. Gomez Tejeda Zandudo, J. Fu, D.E. Garcia Cortes and P. Kurnia conducted patient's studies and provided clinical information. D. Russo, A. Feit, S. Cristea, Y.C. Chen, P. Kurnia, R. Li, X. Qiu and K. Prasad performed all the sequencing analyses under the supervision of A. Feiglin, F. Michor, H. Long, C. Stewart, G. Getz and R. Jeselsohn. C. Guarducci was responsible for final data curation, validation and investigation and prepared all the figures under the supervision of R. Jeselsohn. R. Jeselsohn and C. Guarducci wrote the manuscript with input from all co-authors. All authors read and approved the final manuscript.

Funding

CG was supported by the American-Italian Cancer Foundation Post-Doctoral Research Fellowship 2018–2020 and the Associazione Italiana Ricerca sul Cancro and Fondazione CR Firenze (AIRC fellowship #19288, year 2017) during her postdoctoral training in R.J.'s lab. DLA received support from Cancer Couch Foundation, Hope Scarves, Jamieson Family Fund for Early Career Breast Cancer Researchers, and the Elaine and Eduardo Saverin Foundation; KP was supported by NIH-NCI P01 CA250959. PS was supported by NIH-NCI P01 CA250959, P01 CA250959-04 and R35 CA305086. CPP received funds from Expect Miracles Foundation and the Robert A. and Renée E. Belfer Family Foundation. RJ was supported by NIH-NCI 5R01CA237414, DFCI Susan F. Smith Center for Women's Cancers, Breast & Gynecologic Cancer Innovation Awards, Melanie Rosay fund for Breast Cancer research, Maor Foundation, and the Elaine and Eduardo Saverin Foundation.

Data availability

Genomic data for cell lines in this study have been deposited in NCBI's Gene Expression Omnibus (GEO) under the following accession numbers: overall study (GSE325486; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE325486>), bulk RNA-seq (GSE324093; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE324093>), scRNA-seq (GSE326012; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE326012>), and ATAC-seq (GSE324094; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE324094>) [90]. Code and rendered plot outputs for barcoding and scRNA-seq analyses are located on GitHub: https://github.com/douglas-russo/GuarducciAbravanel2026_barcoding-scRNAseq. The WES analysis code is available on GitHub: https://github.com/getzlab/CDK46_barcoding_paper. Data and code are also available on Zenodo: <https://zenodo.org/records/20414461>. RNA-seq data for samples from the clinical trial in reference 33 in Fig. 5H are available in dbGaP (study accession phs001285.v2.p1). For tumor genomic data and clinical metadata in Fig. 1B–D and Fig. 5G, qualified researchers can apply for access to these summarized data by contacting Daniel Abravanel (Daniel_Abravanel@dfci.harvard.edu) and signing a data usage agreement upon reasonable request. Data requests will be reviewed to determine whether the request is subject to any intellectual property or confidentiality obligations. Data sharing will require a data/material transfer agreement, which may take several months to complete.

Declarations

Ethics approval and consent to participate

All samples included in this study were voluntarily donated by patients who provided informed consent under an IRB-approved protocol (Dana-Farber/Harvard Cancer Center no. 05–246) and patient consent in accordance with the Helsinki Declaration. The protocol includes permission for sample acquisition, clinical data abstraction, sample analysis and data sharing. Analysis of biospecimens at the Broad Institute was performed under Broad Institute protocol number 15-370B.

All mice were maintained in accordance with local guidelines and therapeutic interventions approved by the Animal Care and Use Committees of Dana-Farber Cancer Institute Protocol # 04-111.

Consent for publication

Not applicable.

Competing interests

AN is an AstraZeneca employee; JGTZ reported ownership of stocks in the biotechnology exchange-traded funds CNCR, IDNA, IBB, and XBI; owns stocks in Novo Nordisk and GRAIL; and previously owned stocks in Adaptive Biotechnologies, 2seventy bio, and bluebird bio; NUL has received research

support paid to her institution from Genentech, Pfizer, Merck, Seattle Genetics, Zion Pharmaceuticals, Olema Pharmaceuticals, AstraZeneca. She has received honoraria for consulting work from Seattle Genetics, Daiichi-Sankyo, AstraZeneca, Olema Pharmaceuticals, Janssen, Blueprint Medicines, Stemline/Menarini, Eisai, Shorla Oncology. She holds stock options in Artera Inc (unrelated to breast cancer or brain metastases research) and receives royalties from UpToDate (book chapter). She receives travel support from Olema, AstraZeneca and Daiichi Sankyo; SMT reports consulting or advisory roles for Novartis, Pfizer/SeaGen, Merck, Eli Lilly, AstraZeneca, Genentech/Roche, Eisai, Bristol Myers Squibb/Systimmune, Daiichi Sankyo, Gilead, Blueprint Medicines, Reveal Genomics, Sumitovant Biopharma, Artios Pharma, Menarini/Stemline, Aadi Bio, Bayer, Jazz Pharmaceuticals, Natera, Tango Therapeutics, eFFECTOR, Hengrui USA, Cullinan Oncology, Circle Pharma, Arvinas, BioNTech, Launch Therapeutics, Zuellig Pharma, Johnson&Johnson/Ambrx, Bicycle Therapeutics, BeiGene Therapeutics, Mersana, Summit Therapeutics, Avenzo Therapeutics, Aktis Oncology, Celcuity, Boehringer Ingelheim, Samsung Bioepis, Olema; research funding Genentech/Roche, Merck, Exelixis, Pfizer, Lilly, Novartis, Bristol Myers Squibb, AstraZeneca, NanoString Technologies, Gilead, SeaGen, OncoPep, Daiichi Sankyo, Menarini/Stemline, Jazz Pharmaceuticals; and travel support from Lilly, Gilead, Jazz Pharmaceuticals, Pfizer, Arvinas, Roche; LM reports advisory role for AstraZeneca, Pfizer, Lilly, Novartis, Seagen, Roche and Menarini, fundings from Pfizer, Novartis and travel grants from Pfizer and Gilead; PS has been a consultant at Novartis, Genovis, Guidepoint, The Planning Shop, ORIC Pharmaceuticals, Cedilla Therapeutics, Syros Pharmaceuticals, Blueprint Medicines, Curie Bio, Differentiated Therapeutics, Excientia, Ligature Therapeutics, Merck, Redesign Science, Scorpion Therapeutics, Incyte, Serinus Biosciences, Meliora Therapeutics, Sibylla Biotech, Exo Therapeutics and Astex Pharmaceuticals; KP serves on the Scientific Advisory Board of Ideaya Biosciences, is an adviser to Curie.Bio and Redona Therapeutics, holds equity in Antares Therapeutics and stock options in Ideaya Biosciences, receives sponsored research funding from Novartis, and received honoraria from Astra-Zeneca in the past 12 months; CP is a consultant for XSphera Biosciences and has stock and other ownership interests in XSphera Biosciences. Received honoraria from Thermo Fisher and Agilent and has sponsored research agreements with Daiichi Sankyo, Bicycle Therapeutics, Transcenta, Bicara Therapeutics, AstraZeneca, Janssen Pharmaceuticals, Array Biopharma, Takeda Oncology, Bristol Meyers Squibb, TargImmune, and Mirati Therapeutics; FM is a co-founder of and has equity in Harbinger Health, has equity in Zephyr AI, and serves as a consultant for both companies. She is also on the board of directors of Recursion Pharmaceuticals. She declares that none of these relationships are directly or indirectly related to the content of this manuscript. GG receives research funds from IBM, Pharmacyclics/Abbvie, Bayer, Genentech, Calico, Ultima Genomics, Inocras, Google, Kite, and Novartis and is also an inventor on patent applications filed by the Broad Institute related to MSMuTect, MSMutSig, POLYSOLVER, SignatureAnalyzer-GPU, MSEye, MinimuMM-seq, and DLBclass. He is a founder, consultant, and holds privately held equity in Scorpion Therapeutics; he is also a founder of, and holds privately held equity in, Predicta Biosciences; and holds privately held equity in Antares Therapeutics. RJ reports consulting or advisory role for Eli Lilly, Astra Zeneca, Pfizer, Novartis, Carrick Therapeutics, GE Health, and Luminex and research funding from Pfizer, Eli Lilly, and Novartis. All other authors declare no competing interests.

Author details

¹Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, 450 Brookline Avenue, Boston, MA, USA

²Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA, USA

³Translational Research Unit, Department of Oncology, Hospital of Prato, Azienda USL Toscana Centro, Prato, Italy

⁴Department of Data Science, Dana-Farber Cancer Institute, Boston, MA, USA

⁵Cancer Program, Eli and Edythe L. Broad Institute of MIT and Harvard, Cambridge, MA, USA

⁶Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA

⁷Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

⁸Center for Cancer Evolution, Dana-Farber Cancer Institute, Boston, MA, USA

⁹Department of Oncology, Hospital of Prato, Azienda USL Toscana Centro, Prato, Italy

¹⁰Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA

¹¹Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, MA, USA

¹²Department of Histology and Embryology, Center for Biostructure Research, Medical University of Warsaw, Warsaw, Poland

¹³Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

¹⁴The Ludwig Center, Harvard University, Cambridge, MA, USA

¹⁵Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA, USA

¹⁶Cancer Center and Department of Pathology, Massachusetts General Hospital, Boston, MA, USA

Received: 18 August 2025 / Accepted: 2 June 2026

Published online: 27 June 2026

References

1. Morrison L, Loibl S, Turner NC. The CDK4/6 inhibitor revolution - a game-changing era for breast cancer treatment. *Nat Rev Clin Oncol*. 2024;21(2):89–105.
2. Hui R, Cornish AL, McClelland RA, Robertson JF, Blamey RW, Musgrove EA, et al. Cyclin D1 and estrogen receptor messenger RNA levels are positively correlated in primary breast cancer. *Clin Cancer Res*. 1996;2(6):923–8.
3. Finn RS, Dering J, Conklin D, Kalous O, Cohen DJ, Desai AJ, et al. PD 0332991, a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines in vitro. *Breast Cancer Res*. 2009;11(5):R77.
4. Fanning SW, Hodges-Gallagher L, Myles DC, Sun R, Fowler CE, Plant IN, et al. Specific stereochemistry of OP-1074 disrupts estrogen receptor alpha helix 12 and confers pure antiestrogenic activity. *Nat Commun*. 2018;9(1):2368.
5. Kan Z, Wen J, Bonato V, Webster J, Yang W, Ivanov V, et al. Real-world clinical multi-omics analyses reveal bifurcation of ER-independent and ER-dependent drug resistance to CDK4/6 inhibitors. *Nat Commun*. 2025;16(1):932.
6. O'Leary B, Cutts RJ, Liu Y, Hrebien S, Huang X, Fenwick K, et al. The genetic landscape and clonal evolution of breast cancer resistance to palbociclib plus fulvestrant in the PALOMA-3 trial. *Cancer Discov*. 2018;8(11):1390–403.
7. Li Z, Razavi P, Li Q, Toy W, Liu B, Ping C, et al. Loss of the FAT1 tumor suppressor promotes resistance to CDK4/6 inhibitors via the Hippo pathway. *Cancer Cell*. 2018;34(6):893–905.
8. Tsuji J, Li T, Grinshpun A, Coorens T, Russo D, Anderson L, et al. Clinical efficacy and whole-exome sequencing of liquid biopsies in a phase Ib/II study of bazedoxifene and palbociclib in advanced hormone receptor-positive breast cancer. *Clin Cancer Res*. 2022;28(23):5066–78.
9. Formisano L, Lu Y, Servetto A, Hanker AB, Jansen VM, Bauer JA, et al. Aberrant FGFR signaling mediates resistance to CDK4/6 inhibitors in ER+ breast cancer. *Nat Commun*. 2019;10(1):1373.
10. Watt AC, Cejas P, DeCristo MJ, Metzger-Filho O, Lam EYN, Qiu X, et al. CDK4/6 inhibition reprograms the breast cancer enhancer landscape by stimulating AP-1 transcriptional activity. *Nat Cancer*. 2021;2(1):34–48.
11. Jeselsohn R, Yelensky R, Buchwalter G, Frampton G, Meric-Bernstam F, Gonzalez-Angulo AM, et al. Emergence of constitutively active estrogen receptor- α mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clin Cancer Res*. 2014;20(7):1757–67.
12. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M, et al. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat Genet*. 2013;45(12):1439–45.
13. Jeselsohn R, Bergholz JS, Pun M, Cornwell M, Liu W, Nardone A, et al. Allele-specific chromatin recruitment and therapeutic vulnerabilities of ESR1 activating mutations. *Cancer Cell*. 2018;33(2):173–86 e5.
14. Fribbens C, O'Leary B, Kilburn L, Hrebien S, Garcia-Murillas I, Beaney M, et al. Plasma ESR1 mutations and the treatment of estrogen receptor-positive advanced breast cancer. *J Clin Oncol*. 2016;34(25):2961–8.
15. Lloyd MR, Brett JO, Carmeli A, Weipert CM, Zhang N, Yu J, et al. CDK4/6 inhibitor efficacy in ESR1-mutant metastatic breast cancer. *NEJM Evid*. 2024;3(5):EVID02300231.
16. Katzenellenbogen JA, Mayne CG, Katzenellenbogen BS, Greene GL, Chandrapaty S. Structural underpinnings of oestrogen receptor mutations in endocrine therapy resistance. *Nat Rev Cancer*. 2018;18(6):377–88.

17. Bidard FC, Hardy-Bessard AC, Dalenc F, Bachelot T, Pierga JY, de la Motte Rouge T, et al. Switch to fulvestrant and palbociclib versus no switch in advanced breast cancer with rising ESR1 mutation during aromatase inhibitor and palbociclib therapy (PADA-1): a randomised, open-label, multicentre, phase 3 trial. *Lancet Oncol*. 2022;23(11):1367–77.
18. Bidard FC, Mayer EL, Park YH, Janni W, Ma C, Cristofanilli M, et al. First-line camizestrant for emerging ESR1-mutated advanced breast cancer. *N Engl J Med*. 2025. <https://doi.org/10.1056/nejmoa2502929>.
19. Hafner M, Mills CE, Subramanian K, Chen C, Chung M, Boswell SA, et al. 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.
20. Wander SA, Cohen O, Gong X, Johnson GN, Buendia-Buendia JE, Lloyd MR, et al. The genomic landscape of intrinsic and acquired resistance to cyclin-dependent kinase 4/6 inhibitors in patients with hormone receptor-positive metastatic breast cancer. *Cancer Discov*. 2020;10(8):1174–93.
21. Cardoso F, Paluch-Shimon S, Schumacher-Wulf E, Matos L, Gelmon K, Aapro MS, et al. 6th and 7th International consensus guidelines for the management of advanced breast cancer (ABC guidelines 6 and 7). *Breast*. 2024;76:103756.
22. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
23. Graubert A, Aguet F, Ravi A, Ardlie KG, Getz G. RNA-SeQC 2: efficient RNA-seq quality control and quantification for large cohorts. *Bioinformatics*. 2021;37(18):3048–50.
24. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
25. Zhang Y, Parmigiani G, Johnson WE. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genom Bioinform*. 2020;2(3):lqaa078.
26. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–50.
27. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*. 2009;462(7269):108–12.
28. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754–60.
29. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297–303.
30. Landau DA, Carter SL, Stojanov P, McKenna A, Stevenson K, Lawrence MS, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013;152(4):714–26.
31. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol*. 2012;30(5):413–21.
32. Leshchiner I, Mroz EA, Cha J, Rosebrock D, Spiro O, Bonilla-Velez J, et al. Inferring early genetic progression in cancers with unobtainable premalignant disease. *Nat Cancer*. 2023;4(4):550–63.
33. Gomez Tejeda Zanutto J, Barroso-Sousa R, Jain E, Jin Q, Li T, Buendia-Buendia JE, et al. Exemestane plus everolimus and palbociclib in metastatic breast cancer: clinical response and genomic/transcriptomic determinants of resistance in a phase I/II trial. *Nat Commun*. 2024;15(1):2446.
34. Bhang HE, Ruddy DA, Krishnamurthy Radhakrishna V, Caushi JX, Zhao R, Hims MM, et al. Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. *Nat Med*. 2015;21(5):440–8.
35. Drost H. Philentropy: information theory and distance quantification with R. *J Open Source Softw*. 2018;3(26):765–765.
36. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. 1979;6:65–70.
37. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013;31(3):213–9.
38. Cornwell M, Vangala M, Taing L, Herbert Z, Koster J, Li B, et al. VIPER: visualization pipeline for RNA-seq, a Snakemake workflow for efficient and complete RNA-seq analysis. *BMC Bioinformatics*. 2018;19(1):135.
39. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28(5):511–5.
40. Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics*. 2012;28(16):2184–5.
41. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
42. Marques de Almeida F, Peltzer A, Sturm G, heyfl, Botvinnik O, nf-core bot, et al. nf-core/scrnaseq: 2.7.1. Zenodo; 2024. <https://doi.org/10.5281/zenodo.3568187>.
43. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, et al. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol*. 2020;38(3):276–8.
44. Gruning B, Dale R, Sjoedin A, Chapman BA, Rowe J, Tomkins-Tinch CH, et al. Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat Methods*. 2018;15(7):475–6.
45. da Veiga Leprevost F, Gruning BA, Alves Aflitos S, Rost HL, Uszkoreit J, Barsnes H, et al. BioContainers: an open-source and community-driven framework for software standardization. *Bioinformatics*. 2017;33(16):2580–2.
46. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017;14(4):417–9.
47. Srivastava A, Malik L, Smith T, Sudbery I, Patro R. Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. *Genome Biol*. 2019;20(1):65.
48. Stolarczyk M, Reuter VP, Smith JP, Magee NE, Sheffield NC. Refgenie: a reference genome resource manager. *Gigascience*. 2020;9(2):giz149.
49. Stolarczyk M, Xue B, Sheffield NC. Identity and compatibility of reference genome resources. *NAR Genom Bioinform*. 2021;3(2):lqab036.
50. Zhu A, Srivastava A, Ibrahim JG, Patro R, Love MI. Nonparametric expression analysis using inferential replicate counts. *Nucleic Acids Res*. 2019;47(18):e105.
51. Griffiths JA, Richard AC, Bach K, Lun ATL, Marioni JC. Detection and removal of barcode swapping in single-cell RNA-seq data. *Nat Commun*. 2018;9(1):2667.
52. Lun ATL, Riesenfeld S, Andrews T, Dao TP, Gomes T, participants in the 1st Human Cell Atlas J, et al. EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol*. 2019;20(1):63.
53. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol*. 2024;42(2):293–304.
54. Townes FW, Hicks SC, Aryee MJ, Irizarry RA. Feature selection and dimension reduction for single-cell RNA-seq based on a multinomial model. *Genome Biol*. 2019;20(1):295.
55. Townes FW, Hicks SC, Aryee MJ, Irizarry RA. Author correction: feature selection and dimension reduction for single-cell RNA-seq based on a multinomial model. *Genome Biol*. 2020;21(1):179.
56. Weine E, Carbonetto P, Stephens M. Accelerated dimensionality reduction of single-cell RNA sequencing data with fastglmpca. *Bioinformatics*. 2024. <https://doi.org/10.1093/bioinformatics/btae494>.
57. Zappia L, Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *Gigascience*. 2018;7(7):gly083.
58. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res*. 2012;40(17):e133.
59. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
60. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417–25.
61. Benjamini YH. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B (Methodol)*. 1995;57(1):289–300.
62. Nardone A, Qiu X, Spisak S, Nagy Z, Feiglin A, Feit A, et al. A distinct chromatin state drives therapeutic resistance in invasive lobular breast cancer. *Cancer Res*. 2022;82(20):3673–86.
63. Buenostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013;10(12):1213–8.
64. Qiu X, Feit AS, Feiglin A, Xie Y, Kesten N, Taing L, et al. CoBRA: containerized bioinformatics workflow for reproducible ChIP/ATAC-seq analysis. *Genomics Proteomics Bioinformatics*. 2021;19(4):652–61.

65. Taing L BG, Cousins C, et al. CHIPS: A Snakemake pipeline for quality control and reproducible processing of chromatin profiling data. *F1000Research*. 2021;10:517.
66. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.
67. Wang S, Sun H, Ma J, Zang C, Wang C, Wang J, et al. Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nat Protoc*. 2013;8(12):2502–15.
68. Franca GS, Baron M, King BR, Bossowski JP, Bjornberg A, Pour M, et al. Cellular adaptation to cancer therapy along a resistance continuum. *Nature*. 2022;631(8022):876–83.
69. Vander Velde R, Yoon N, Marusyk V, Durmaz A, Dhawan A, Miroshnychenko D, et al. Resistance to targeted therapies as a multifactorial, gradual adaptation to inhibitor specific selective pressures. *Nat Commun*. 2020;11(1):2393.
70. Card KJ, Crozier D, Durmaz A, Gray J, Creary J, Stocks A, et al. Evolution under vancomycin selection drives divergent collateral sensitivity patterns in *Staphylococcus aureus*. *Proc Natl Acad Sci U S A*. 2025;122(39):e2507962122.
71. Jeselsohn R, Fu J, Ren Y, Mahtani R, Ma C, DeMichele A, et al. Abstract PS12-02: exploration of ctDNA dynamics in the PACE trial: a randomized phase II study of fulvestrant, palbociclib, and avelumab for HR+/HER2- metastatic breast cancer. *Cancer Res*. 2024;84(9_Supplement):PS12-02-PS12-02.
72. Razavi P, Chang MT, Xu G, Bandlamudi C, Ross DS, Vasan N, et al. The genomic landscape of endocrine-resistant advanced breast cancers. *Cancer Cell*. 2018;34(3):427–38 e6.
73. Dhimolea E, de Matos Simoes R, Kansara D, Al'Khafaji A, Bouyssou J, Weng X, et al. An embryonic diapause-like adaptation with suppressed Myc activity enables tumor treatment persistence. *Cancer Cell*. 2021;39(2):240–56 e11.
74. Ross-Innes CS, Brown GD, Carroll JS. A co-ordinated interaction between CTCF and ER in breast cancer cells. *BMC Genomics*. 2011;12:593.
75. Malorni L, Giuliano M, Migliaccio I, Wang T, Creighton CJ, Lupien M, et al. Blockade of AP-1 potentiates endocrine therapy and overcomes resistance. *Mol Cancer Res*. 2016;14(5):470–81.
76. Lupien M, Eeckhoutte J, Meyer CA, Wang Q, Zhang Y, Li W, et al. FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell*. 2008;132(6):958–70.
77. Chi D, Singhal H, Li L, Xiao T, Liu W, Pun M, et al. Estrogen receptor signaling is reprogrammed during breast tumorigenesis. *Proc Natl Acad Sci U S A*. 2019;116(23):11437–43.
78. Eeckhoutte J, Carroll JS, Geistlinger TR, Torres-Arzayus MI, Brown MA. A cell-type-specific transcriptional network required for estrogen regulation of cyclin D1 and cell cycle progression in breast cancer. *Genes Dev*. 2006;20(18):2513–26.
79. Cyr AR, Kulak MV, Park JM, Bogachek MV, Spanheimer PM, Woodfield GW, et al. TFAP2C governs the luminal epithelial phenotype in mammary development and carcinogenesis. *Oncogene*. 2015;34(4):436–44.
80. Ji Z, He L, Regev A, Struhl K. Inflammatory regulatory network mediated by the joint action of NF- κ B, STAT3, and AP-1 factors is involved in many human cancers. *Proc Natl Acad Sci U S A*. 2019;116(19):9453–62.
81. Bidard FC, Kaklamani VG, Neven P, Streich G, Montero AJ, Forget F, et al. Elacestrant (oral selective estrogen receptor degrader) versus standard endocrine therapy for estrogen receptor-positive, human epidermal growth factor receptor 2-negative advanced breast cancer: results from the randomized phase III EMERALD trial. *J Clin Oncol*. 2022;40(28):3246–56.
82. Jhaveri KL, Neven P, Casalnuovo ML, Kim SB, Tokunaga E, Aftimos P, et al. Imlunestrant with or without abemaciclib in advanced breast cancer. *N Engl J Med*. 2025;392(12):1189–202.
83. Migliaccio I, Guarducci C, Malorni L. CDK4/6 inhibitor resistance in ER+ breast cancer. *Adv Exp Med Biol*. 2025;1464:475–93.
84. Jeselsohn R, Schiff R, Grinshpun A. Restoring order at the cell cycle border: co-targeting CDK4/6 and CDK2. *Cancer Cell*. 2021;39(10):1302–5.
85. Spring LM, Wander SA, Andre F, Moy B, Turner NC, Bardia A. Cyclin-dependent kinase 4 and 6 inhibitors for hormone receptor-positive breast cancer: past, present, and future. *Lancet*. 2020;395(10226):817–27.
86. Dickler MN, Tolane SM, Rugo HS, Cortes J, Dieras V, Patt D, et al. MONARCH 1, a phase II study of abemaciclib, a CDK4 and CDK6 inhibitor, as a single agent, in patients with refractory HR(+)/HER2(-) metastatic breast cancer. *Clin Cancer Res*. 2017;23(17):5218–24.
87. Johnston SRD, Harbeck N, Hegg R, Toi M, Martin M, Shao ZM, et al. Abemaciclib combined with endocrine therapy for the adjuvant treatment of HR+, HER2-, node-positive, high-risk, early breast cancer (monarchE). *J Clin Oncol*. 2020;38(34):3987–98.
88. Navarro-Yepes J, Kettner NM, Rao X, Bishop CS, Bui TN, Wingate HF, et al. Abemaciclib is effective in palbociclib-resistant hormone receptor-positive metastatic breast cancers. *Cancer Res*. 2023;83(19):3264–83.
89. Kalinsky K, Bianchini G, Hamilton E, Graff SL, Park KH, Jeselsohn R, et al. Abemaciclib plus fulvestrant in advanced breast cancer after progression on CDK4/6 inhibition: results from the phase III postMONARCH trial. *J Clin Oncol*. 2025;43(9):1101–12.
90. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002;30(1):207–10.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.