Comprehensive mutation and copy number profiling in archived circulating breast cancer tumor cells reveals heterogeneous resistance mechanisms

Costanza Paoletti, Andi K. Cani, Jose M. Larios, Daniel H. Hovelson, Kimberly Aung, Elizabeth P. Darga, Emily M. Cannell, Paul J. Baratta, Chia-Jen Liu, David Chu, Maryam Yazdani, Allen R. Blevins, Valeria Sero, Nahomi Tokudome, Dafydd G. Thomas, Christina Gersch, Anne F. Schott, Yi-Mi Wu, Robert Lonigro, Dan R. Robinson, Arul M. Chinnaiyan, Farideh Z. Bischoff, Michael D. Johnson, Ben H. Park, Daniel F. Hayes, James M. Rae, Scott A. Tomlins

SUPPLEMENTARY MATERIALS

Supplementary Materials and Methods

Preclinical evaluation of cultured breast cancer cells spiked into normal human blood

Cultured human breast cancer cell line BT-474 was obtained from the laboratory of Dr. Stephen Ethier and routinely maintained in DMEM +10% FBS at 37°C in 5% CO_2. The identity of the cells was confirmed by standard Short Tandem Repeat profiling (February 2011). BT-474

(*n*=300) were freshly harvested and spiked into pooled de-identified healthy human whole blood (WB) collected in a CellSave Preservative tube (Janssen Diagnostics, LLC). Tumor cell isolation and enumeration was performed using the CellSearch® system with the CellSearch® CXC kit, as previously described (1). The elapsed time between CTC enrichment and purification varied from 2 years (cartridges were stored at RT), 6 months (cartridge stored at 4°C), and 3 days (cartridge stored at 4°C). This evaluation was also performed to see the effects of long term storage on the quality of the cells. As described below for patient specimens, BT-474 cells were flushed from the CellSearch® cartridges and purified using DEPArray™ technology. Isolated single or groups of CTC were subjected to whole genome amplification (WGA) before sequencing analysis.

Patient population

Twenty-eight eligible patients with MBC, who were enrolled in the MI-ONCOSEQ protocol at the University of Michigan Comprehensive Cancer Center (2), signed a separate informed consent approved by the University of Michigan IRB to be enrolled in a companion protocol designated Mi-CTC-ONCOSEQ that allowed blood collection for cfDNA (data partially reported elsewhere) (3) and CTC analyses. All patients enrolled into this study had WES performed on their metastatic research tissue biopsy using the Illumina HiSeq 2500 platform. Thirteen of these 24 patients who had ≥5 CTC/7.5 ml WB by CellSearch® enumeration, and who had at least one DEPArray™-purified CTC with high quality DNA determined by the Ampli1™ quality control (QC) kit (as described in methods below), were enrolled in the present study.

Blood specimens were collected into fixative containing 10 ml vacutainer tubes (CellSave) and processed within 96 hours. These specimens were collected from a range of 7 days prior, to 344 days after the research biopsy. For two patients (#4, who had <5 CTC/7.5 ml at time of enrollment and #24, who had ≥5/7.5 ml CTC at time of enrollment), blood for CTC-DNA was also collected at the time of disease progression after 690 days and 465 days from the baseline samples, respectively.

CTC enrichment and enumeration

CTC were enriched from WB using ferrofluid particles coated with EpCAM antibody and enumerated using the CellSearch® System, according to manufacturer's instructions (Janssen Diagnostics, LLC Raritan, NJ, United States), and as previously described (1). Briefly, cells were stained with anti-cytokeratin (CK) 8, 18, and 19- Fluorescein isothiocyanate (FITC) labeled antibodies, the nuclear dye 4',6'-diamino-2-phenylindole (DAPI), and CD45 antibody labeled with allophycocyanin (APC). CTC were defined by visual inspection as CK positive/DAPI positive/CD45 negative. After processing, CellSearch® cartridges were stored protected from light at 4°C for an average of 322 days (range 37-827 days) before CTC purification (**Supplementary Table S3**).

Cell Purification

Because CellSearch® technology enriches, but does not purify CTC from WBC, each CellSearch® cartridge was subsequently processed to recover single and pooled populations of pure CTC using the DEPArray™ system (Di-Electro-Phoretic Array; Menarini Silicon Biosystems, S.p.A., Bologna, Italy) according to the manufacturer's instructions. After the cells

were flushed from the cartridges using the manipulation buffer wash provided by the manufacturer, the cell pellet was re-suspended and placed into the DEPArray™ 300k chip which contains electronically controlled electrodes used to generate the electrophoretic cages. The DEPArray™ single cell sorting system separates individual cells into separate virtual magnetic nests/cages, based on immunofluorescent staining criteria and cell morphology (CTC: CK-FITC positive, DAPI positive, and CD45-APC negative; WBC: CK-FITC negative, DAPI positive, and CD45-APC positive). After imaging of the cell within each virtual magnetic nest/cage generated by DEPArray™, individual cells of interest were routed for isolation and recovery into 0.2 ml tubes (4).

CTC whole genome amplification and WGA quality control

DNA from individual CTC or WBC from the same specimen, was isolated and subjected to whole genomic amplification (WGA) using *Ampli*1™ WGA kit (Menarini Silicon Biosystems, S.p.A.) which includes digestion with a restriction enzyme (MseI) according to the manufacturer's instructions (5). Subsequent DNA quality control was performed using *Ampli*1™ QC kit as described by Polzer *et al* (5). Only cells with high quality DNA were further investigated for down-stream sequence analysis. Normal control DNA derived from matched WBC was processed and tested for QC following the same process and criteria.

CTC Genomic profiling and data analysis

Targeted next-generation sequencing (NGS) was performed essentially as previously described (6-8). Briefly, 20 ng of WGA DNA (quantified with the Qubit Fluorometer 3.0, ThermoFisher,

Waltham, MA) for each single, or equally pooled CTC, or WBC DNA was used for targeted, multiplexed PCR-based NGS (Ampliseq, Ion Torrent). Libraries with barcode incorporation were constructed using the DNA component of the Oncomine Comprehensive Assay (OCP), a panel comprised of 2,531 amplicons targeting 130 genes covering 260,717 bases (ThermoFisher, Waltham, MA). Genes were selected based on pan-solid tumor NGS and copy number profiling data analysis that prioritized somatic, recurrently altered oncogenes, tumor suppressor genes, genes present in high level copy gains/losses and known/investigational therapeutic targets (7). Libraries for spiked BT-474 cells were constructed using the Cancer Hotspot v2 Assay (CHPv2), a panel comprised of 207 amplicons targeting regions containing recurrent alterations in 50 onco- and tumor suppressor genes (ThermoFisher, Waltham, MA). Because the WGA method we used digests the DNA with the MseI restriction enzyme, some amplicons in each panel were adversely affected based on the presence of the MseI restriction site (for OCP see **Supplementary Table S1;** for CHPv2, as previously described (9)). Library preparation, template, and sequencing using the Ion Torrent Proton were performed according to the manufacturer's instructions. Data analysis was performed using Torrent Suite 4.0.2 – 5.0.2, with alignment by TMAP using default parameters, and variant calling with the Ion Torrent Variant Caller plugin (version v4.0-r76860 – v5.0.2.1) using default low-stringency somatic variant settings. Variant annotation, filtering and prioritization was performed essentially as previously described using validated in-house pipelines (6-8). Briefly, called variants were filtered to remove synonymous or non-coding variants, those with flow corrected read depths (FDP) less than 10, flow corrected variant allele containing reads (FAO) less than 10, variant allele frequencies (FAO/FDP) less than 0.10, extreme skewing of forward/reverse flow corrected reads calling the variant (FSAF/FSAR <0.2 or >5), or indels within homopolymer runs >4. Called

variants were filtered using a panel-specific, in house blacklist. Variants reported at population allele frequencies >0.5% in ESP6500 or 1000 Genomes or those reported in ESP6500 or 1000 Genomes at population allele frequencies <0.5% but with observed variant read frequencies >0.40 and <0.60 or >0.9 were considered germ line variants unless occurring at a known cancer hot-spot. Variants located at the last mapped base (or outside) of amplicon target regions, those with the majority of supporting reads harboring additional mismatches or indels (likely sequencing error), those in repeat–rich regions (likely mapping artifacts), and those occurring exclusively in one amplicon if overlapping amplicons cover the variant were excluded. High confidence somatic variants passing the above criteria were then visually confirmed in Integrated Genome Viewer (IGV) as were the same regions in samples from the same patient where the variant was not called to confirm both coverage and absence of the variant. We have previously confirmed that these filtering criteria identify variants that pass Sanger Sequencing validation with >95% accuracy (6,8,10). We prioritized potential driving alterations using the Oncomine annotation in the Oncomine Knowledge Base, which uses pan-cancer NGS data to identify genes as oncogenes or tumor suppressors, based on overrepresentation of hot-spot or deleterious mutations, respectively. Variants in oncogenes are then considered gain-of-function (GoF) if at a hot-spot, and variants in tumor suppressors are considered loss-of-function (LoF) if deleterious or at a hot-spot (6,7,10).

Copy number analysis from total amplicon read counts provided by the coverageAnalysis plugin (version v4.4.2.2 - v5.0.2.0) was performed essentially as previously described using a validated approach (6-8,11) with adaptations for single cell sequencing. Specifically, for copy number analysis, we retained only well-performing amplicons (with >100 reads in all individual or pooled WBC samples,795 total amplicons) in order to exclude amplicons lost due to their

containing the MseI restriction site. To identify CNAs, normalized read counts per amplicon for each sample were divided by the normalized mean read count of the corresponding amplicons from the WBC samples, yielding a copy-number ratio for each amplicon. Gene-level copy-number estimates were determined by taking the coverage-weighted mean of the per amplicon ratios, with expected error determined by the amplicon-to-amplicon variance (11). Genes with a log2 copy-number estimate < -2 or > 2 were considered to have high-level loss or gain, respectively.

Confirmation of ESR1, CDH1 and BRCA2 mutations by Sanger Sequencing

Bidirectional Sanger Sequencing to validate patient #2 *CDH1* (p.Q641X) and *ESR1* (p.Y537S, p.D538G, p.A569S) and patient #17 *BRCA2* (p.Q1931X) somatic variants was performed on selected samples. WGA CTC/WBC Genomic DNA (10 ng) was used as template in PCR amplifications with Invitrogen Platinum PCR Supermix Hi-Fi (ThermoFisher, Waltham, MA) with the suggested denaturation and cycling conditions. Primer sequences were designed with the PrimerQuest tool from IDT (Coralville, IA) (*CDH1* Q641 forward: CAGACCTTCCTCCCAATACATC, reverse: AAGGGAAGCATGGCAGTT; *ESR1* Y537/D538 forward: CTTTCTGTGTCTTCCCACCTAC, reverse: TTTGGTCCGTCTCCTCCA; *ESR1* A569 forward: CTGCTGGAGATGCTGGAC, reverse: AGGGAAACCCTCTGCCT; *BRCA2* Q1931 forward: TTGTTACGAGGCATTGGATGA, reverse: ACTGACTTATGAAGCTTCCCTATAC) (**Supplementary Table S8**). PCR products were subjected to bidirectional Sanger Sequencing by the University of Michigan DNA Sequencing Core after purification with the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and

sequencing traces were analyzed using SeqMan Pro software (DNASTAR) and visually inspected.

Tissue Analysis

Specimen collection and processing

Sequencing of clinical samples was performed under IRB-approved studies at the University of Michigan. Patients were enrolled and consented for integrative tumor sequencing in MI-ONCOSEQ Protocol, HUM00046018, as previously described (12). Needle biopsies were snap frozen in Optimal Cutting Temperature compound, and a longitudinal section was cut. Frozen sections stained with hematoxylin and eosin were reviewed by pathologists to identify cores with the highest tumor content. Remaining portions of each needle biopsy core were retained for nucleic acid extraction.

DNA extraction and library preparation for exome sequencing.

Genomic DNA from frozen needle biopsies was isolated using the Qiagen DNeasy Blood and Tissue kit, according to the manufacturer's instructions. Exome libraries of matched tumor/normal paired DNA were generated using the Illumina TruSeq DNA Sample Prep kit, per manufacturer's instructions as previously described (12). Briefly, 1–3 μg of DNA was sheared to a 250 bp peak target size, and end repaired followed by A-base addition and ligation of Illumina indexed adaptors. Agarose gel purified 300–350 bp fragments were amplified using Illumina index primers for eight cycles and purified. One μg of library was hybridized to the Agilent SureSelect Human All Exon v4 chip (Agilent Technologies) and hybridized exon fragments were captured and amplified with the Illumina index primers for nine additional PCR cycles. Purified

PCR products were analyzed for quality and quantity with the Agilent 2100 Bioanalyzer (DNA 1000 kit).

WES was performed on Illumina HiSeq 2000 or HiSeq 2500 (paired-end). Sequencing quality was assessed with FastQC. For each sequencing lane, we examined per-base quality scores across the length of the reads with lanes deemed passing if the per-base quality is >Q20 in >85% of the reads for bases 1–100. We also assessed alignment quality (Picard package) by monitoring duplication rates and potential chimeric reads from ligation artifacts.

Tissue mutation analysis.

Analysis was performed as previously described (12). Briefly, FASTQ converted sequence files were processed through an in-house pipeline for WES analyses of paired cancer and normal genomes. Sequencing reads were aligned to reference genome hg19 with Novoalign multithreaded (v2.08.02, Novocraft) and converted into BAM files using SAMtools (v0.1.18) (13). Mutation analysis was performed using VarScan2 (v2.3.2) (14) with the pileup files created by SAMtools mpileup for tumor and matched normal samples, simultaneously performing pairwise comparisons of base call and normalized sequence depth at each position. For SNV detection, filtering parameters including coverage, variant read support, variant frequency, P value, base quality, the presence of homopolymers and strandedness were applied. For indel analysis, Pindel (v0.2.4) was used and indels common to both tumor and matched normal samples were classified as germline, whereas indels present in tumor but not in normal samples were classified as somatic. Finally, a list of candidate indels and somatic mutations was generated by excluding synonymous SNVs. Detected variants were functionally annotated with ANNOVAR (15).

Tumor content for each tumor exome library was estimated from the sequence data by fitting a binomial mixture model with two components to the set of most likely SNV candidates from two-copy genomic regions as determined by copy number analysis, as previously described (12). Copy number aberrations were quantified for each gene as the segmented, normalized, GC content-corrected, log2-transformed exon coverage ratio between each tumor sample and its matched normal sample (16).

ddPCR analysis of tissue and cfDNA

DNA extraction from FFPE blocks

Formalin fixed paraffin embedded (FFPE) blocks and slides were obtained from University of Michigan Department of Pathology archives. A representative block from each specimen was chosen and DNA extracted using a commercial kit (Qiagen QiAMP FFPE tissue kit) according to the manufacturer's instructions.

Isolation of cfDNA for ddPCR

Blood sample and plasma preparation were performed as previously described (17). Blood was collected in DNA BC tube (Streck) and centrifuged within 7 days. Plasma was obtained by a double-spin centrifuge protocol of whole blood to remove cellular concomitants and DNA was extracted using the Qiagen Circulating Nucleic Acid Kit (Qiagen) per the manufacture's protocol as previously described (3).

ddPCR for FFPE DNA and cfDNA

Genomic and plasma DNA samples were subjected to short cycle PCR amplification for the *ESR1* p.Y537S/p.D538G and p.A569S locus with the primers shown in **Supplementary**

Table S8. Amplified DNA was purified using the QIAquick PCR Purification kit (Qiagen). The Bio-Rad QX200 platform was then used for ddPCR per the manufacturer's protocols with the ddPCR primers and probes listed in **Supplementary Table S8**. The results are reported as fractional abundance of mutant DNA alleles to total (mutant plus wild-type) DNA alleles for each sample assayed.

Functional analysis of novel *ESR1* mutation

Steroids and drugs

17β-estradiol, tamoxifen, fulvestrant, polybrene and caffeine were obtained from Sigma Aldrich. 4-hydroxytamoxifen and endoxifen were obtained from Toronto Research Chemical.

Plasmids and mutagenesis

The lentiviral vector for wild-type *ESR1* coding sequence, pCDH-ESR1, was obtained from stock used in (12). The viral packaging plasmids psPAX2 and pMD2.G were gifts from Didier Trono (Addgene plasmids # 12260 and 12259, respectively). The previously constructed plasmid pCDH-ESR1 was mutated with Quick Change Lightning Kit (Agilent Technologies). The alanine residue at position 569 of ER was mutated to a serine residue according the manufacturer's protocol. The sequences for the sense and anti-sense primers for site-directed mutagenesis were as follows, sense: 5'-GAGCCCGCA GTGGACAAGTGGCTTTGG-3'; anti-sense: 5'-CCAAAGCCACTTGTCCACTGCGGGCTC-3' (the underlined nucleotides are mutated). Mutagenesis was confirmed by Sanger Sequencing.

Cell Culture

MCF-7 and 293T cells were obtained from the Lombardi Comprehensive Cancer Center and routinely maintained in DMEM (ThermoFisher) supplemented with 5% fetal bovine serum (Valley Biomedical). Lentivirus containing the p.A569S *ESR1* transgene was packaged in 293T cells. 293T cells were plated at a density of 2 x 10^6 cells/10 cm dish. 24 hours after plating, cells were transfected with 8ug of pCDH-*ESR1*-A569S plasmid, 5ug psPAX2 and 2ug pMD2.G plasmids via calcium chloride. Transfection medium was allowed to remain on the cells for 8 hours and then replaced with fresh DMEM/5% FBS medium and cells were incubated overnight. On the following day, 12 ml of DMEM/5% FBS supplemented with 0.5mM caffeine was added to the dish for virus production. 48 hours later, the supernatant was carefully removed from the dish, filtered through a 0.45 um filter and frozen down in 1 ml aliquots.

MCF-7 cells were seeded into a 6-well plate at a density of 0.25 x 10^6 cells per well and allowed to attach overnight. 24 hours after plating, cells were virally transduced with 1ml of viral supernatant supplemented with 4ug/ml polybrene for 8 hours and then restored in fresh DMEM/5% FBS. 48 hours after viral transduction, cells were stably selected in DMEM/5% FBS medium supplemented with 1ug/ml puromycin for one week. p.A569S *ESR1* transgene expression was confirmed by Sanger Sequencing of genomic DNA and cDNA reverse transcribed from 1ug of DNase-treated total RNA. For growth assays in defined hormone conditions, cells were depleted of exogenous steroids as previously described (18). Steroid-depleted parental and *ESR1* p.A569S MCF-7 cells were seeded into 96-well plates at a density of 1,000 cells/well in 0.1 ml and allowed to attach overnight. Cells were treated with 17β-estradiol (Sigma Aldrich) or ethanol control alone or in combination with tamoxifen, 4-hydroxytamoxifen,

endoxifen or fulvestrant. Cell number was assessed by crystal violet stain five days after hormone treatment as previously published (18).

Supplementary Figure Legends

Supplementary Figure S1. Clinical timelines. Clinical timelines and treatment from first diagnosis until enrollment into MI-CTC-ONCOSEQ for the 11 metastatic breast cancer patients (No available timeline for patient #14). Each bar represents the timeframe of treatment. Colors represent different therapies: Surgery=■; RT=■; NED= ; Tamoxifen=■; Letrozole =■; Anastrazole=■; Exemestane=■; Fulvestrant =■; Therapy other than chemo-or endocrine-therapy =■; chemotherapy =■; oophorectomy = ■; Estrace = ■; Everolimus = *; Palbociclib = #.

Supplementary Figure S2. Sanger sequencing of WGA CTC DNA was fully concordant with sequencing data. A. Sanger sequencing of for CDH1 (Q641X), ESR1 (Y537S, D538G, A569S) and BRCA2 (Q1931X). NGS variant fraction (V.F.) shown when performed. Codon of affected aminoacid is underlined. Red rectangles denote Pt #2 single CTC with novel ESR1 mutation and Pt #17 single CTC with BRCA2 mutation. Only forward Sanger traces are shown, reverse traces were all in concordance (not shown); **B.** Sanger sequencing results of cDNA from parental and A569S over-expressing MCF-7 cells; **C.** Western blot confirming the presence of A569S mutation in overexpressed cells compared to Wild Type.

Supplementary Figure S3. MCF-7 (ESR1 A569S) are not estrogen independent over 5 days in hormone-free conditions when assessed by crystal violet assay. Cells were depleted of exogenous steroids and treated with 17β-estradiol or ethanol control in triplicates. Cell number was assessed by crystal violet stain five days after treatment.

Supplementary Figure S4. The tamoxifen metabolites 4-hydroxytamoxifen and endoxifen have no increased agonistic effect in MCF-7 cells over-expressing ER-A569S.

Supplementary Figure S5. Inhibition of estradiol-stimulated MCF-7 parental or A569S-overexpressing cells. Cells were withdrawn from estrogen and treated with 50pM estradiol in combination with increasing concentrations of **A.** tamoxifen, **B.** 4-hydroxytamoxifen, **C.** endoxifen or **D.** fulvestrant, in triplicates. Cell growth was assessed six days after treatment by crystal violet staining and plotted as %-growth vs. vehicle control.

Supplementary Table Legends

Supplementary Table S1. Oncomine Comprehensive Panel (OCP) target genes by presence of the MseI restriction site

Supplementary Table S2. NGS of BT-474 cells spiked into blood and purified identifies expected *TP53* mutation

Supplementary Table S3. Results of CTC purification by DEPArray

Supplementary Table S4: Mean CTC Next Generation Sequencing parameters

Supplementary Table S5. NGS-identified prioritized mutations in individual and pooled CTC samples

Supplementary Table S6. 2 x 2 contingency tables of concordance for alterations in CTC vs tissue

Supplementary Table S7. Patient #2 analysis of primary, clinical metastatic tissue, research biopsy, pt-DNA by ddPCR

Supplementary Table S8. Primer/probe sequences used in Sanger sequencing and ddPCR

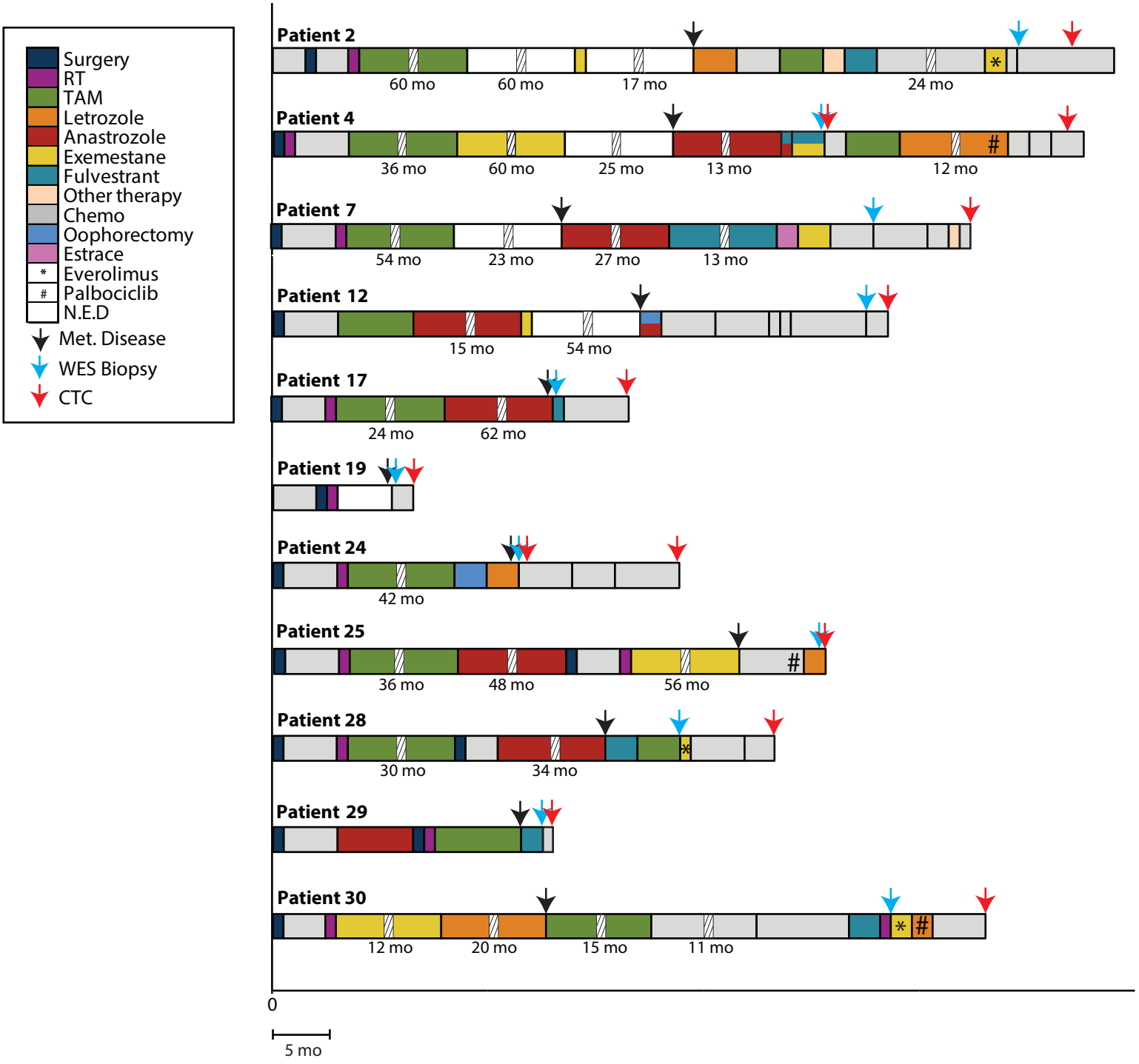
Supplementary References

1. Paoletti C, Muniz MC, Thomas DG, Griffith KA, Kidwell KM, Tokudome N, et al. Development of circulating tumor cell-endocrine therapy index in patients with hormone receptor-positive breast cancer. Clin Cancer Res 2015;21(11):2487-98.
2. Roychowdhury S, Iyer MK, Robinson DR, Lonigro RJ, Wu YM, Cao X, et al. Personalized oncology through integrative high-throughput sequencing: a pilot study. Sci Transl Med 2011;3(111):111ra21.
3. Chu D, Paoletti C, Gersch C, VanDenBerg DA, Zabransky DJ, Cochran RL, et al. ESR1 Mutations in Circulating Plasma Tumor DNA from Metastatic Breast Cancer Patients. Clin Cancer Res 2016;22(4):993-9.
4. Peeters DJ, De Laere B, Van den Eynden GG, Van Laere SJ, Rothe F, Ignatiadis M, et al. Semiautomated isolation and molecular characterisation of single or highly purified tumour cells from CellSearch enriched blood samples using dielectrophoretic cell sorting. Br J Cancer 2013;108(6):1358-67.
5. Polzer B, Medoro G, Pasch S, Fontana F, Zorzino L, Pestka A, et al. Molecular profiling of single circulating tumor cells with diagnostic intention. EMBO Mol Med 2014;6(11):1371-86.
6. Cani AK, Hovelson DH, McDaniel AS, Sadis S, Haller MJ, Yadati V, et al. Next-Gen Sequencing Exposes Frequent MED12 Mutations and Actionable Therapeutic Targets in Phyllodes Tumors. Mol Cancer Res 2015;13(4):613-9.
7. Hovelson DH, McDaniel AS, Cani AK, Johnson B, Rhodes K, Williams PD, et al. Development and validation of a scalable next-generation sequencing system for assessing relevant somatic variants in solid tumors. Neoplasia 2015;17(4):385-99.
8. Warrick JI, Hovelson DH, Amin A, Liu CJ, Cani AK, McDaniel AS, et al. Tumor evolution and progression in multifocal and paired non-invasive/invasive urothelial carcinoma. Virchows Arch 2015;466(3):297-311.
9. Salvianti F, Rotunno G, Galardi F, De Luca F, Pestrin M, Vannucchi AM, et al. Feasibility of a workflow for the molecular characterization of single cells by next generation sequencing. Biomol Detect Quantif 2015;5:23-9.
10. McDaniel AS, Stall JN, Hovelson DH, Cani AK, Liu CJ, Tomlins SA, et al. Next-Generation Sequencing of Tubal Intraepithelial Carcinomas. JAMA Oncol 2015.
11. Grasso C, Butler T, Rhodes K, Quist M, Neff TL, Moore S, et al. Assessing copy number alterations in targeted, amplicon-based next-generation sequencing data. J Mol Diagn 2015;17(1):53-63.
12. Robinson DR, Wu YM, Vats P, Su F, Lonigro RJ, Cao X, et al. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. Nat Genet 2013;45(12):1446-51.
13. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25(16):2078-9.
14. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012;22(3):568-76.
15. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38(16):e164.
16. Lonigro RJ, Grasso CS, Robinson DR, Jing X, Wu YM, Cao X, et al. Detection of somatic copy number alterations in cancer using targeted exome capture sequencing. Neoplasia 2011;13(11):1019-25.

17. Beaver JA, Jelovac D, Balukrishna S, Cochran RL, Croessmann S, Zabransky DJ, et al. Detection of cancer DNA in plasma of patients with early-stage breast cancer. Clin Cancer Res 2014;20(10):2643-50.
18. Johnson MD, Zuo H, Lee KH, Trebley JP, Rae JM, Weatherman RV, et al. Pharmacological characterization of 4-hydroxy-N-desmethyl tamoxifen, a novel active metabolite of tamoxifen. Breast Cancer Res Treat 2004;85(2):151-9.

Supplementary Figures and Tables

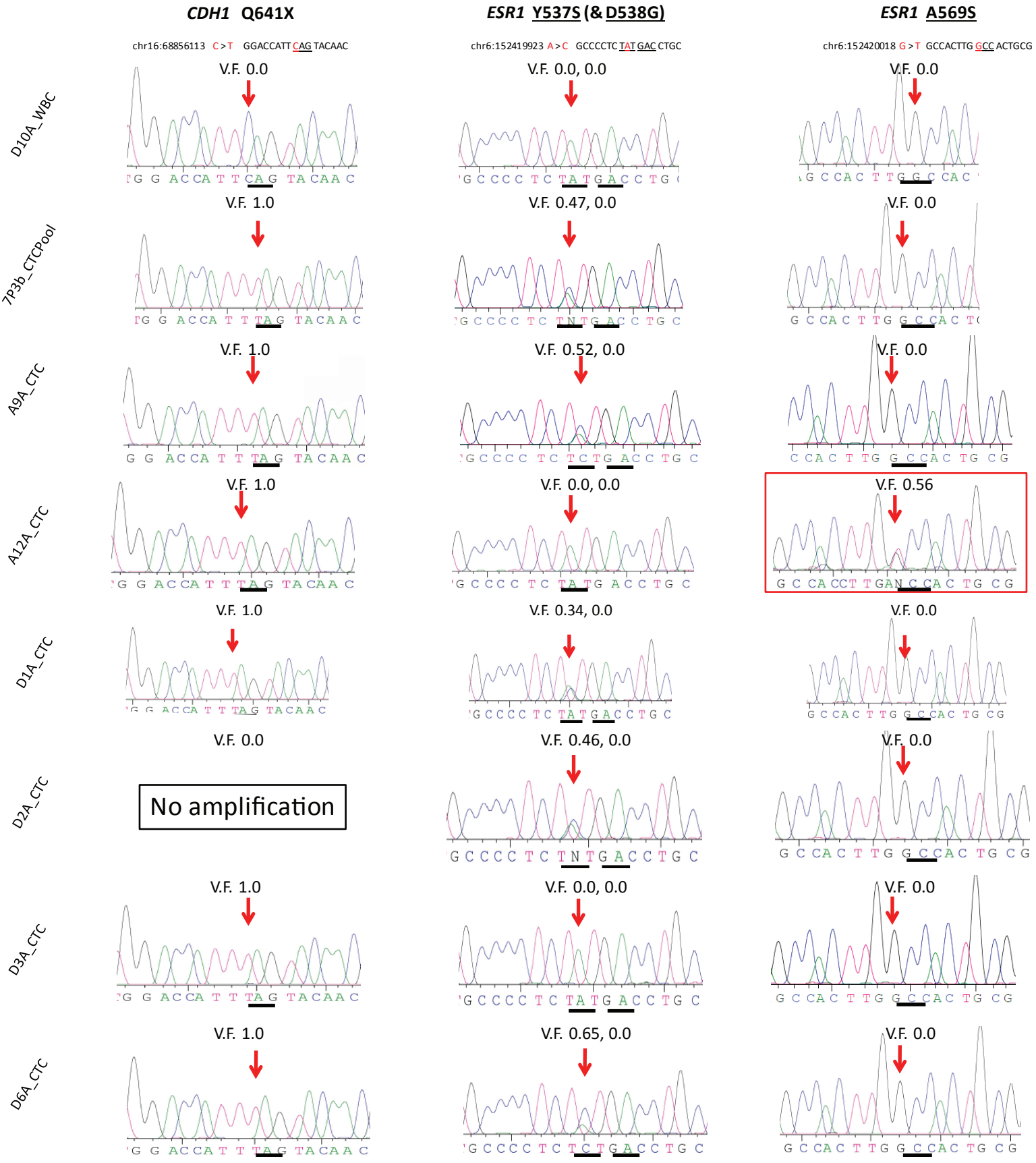
Supplementary Figure S1



Supplementary Figure S2

A.

Patient 02



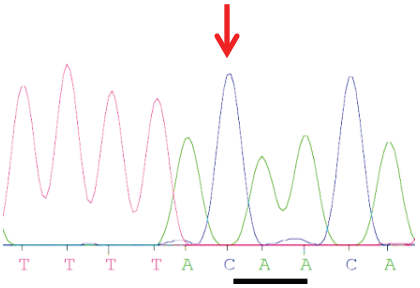
Patient 17

BRCA2 Q1931X

chr13:32914283 C>T TTTTA CAA CA

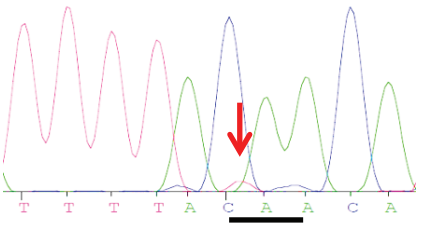
No amplification

Pt17 A5 CTC

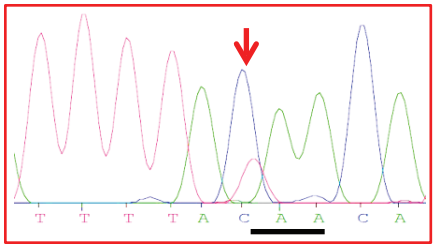


Pt17_A6_CTC

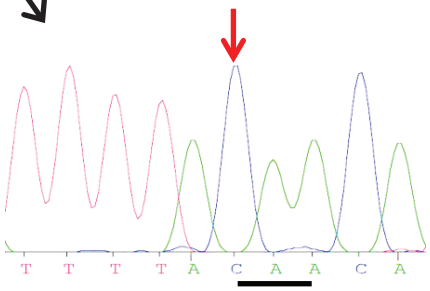
V.F. 0.10



Pt17(4P)_CTC_Pool
A5_A6_A7_A8

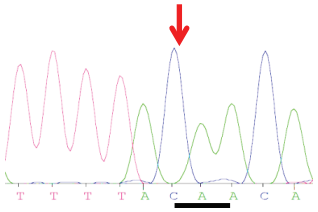


Pt17_A7_CTC



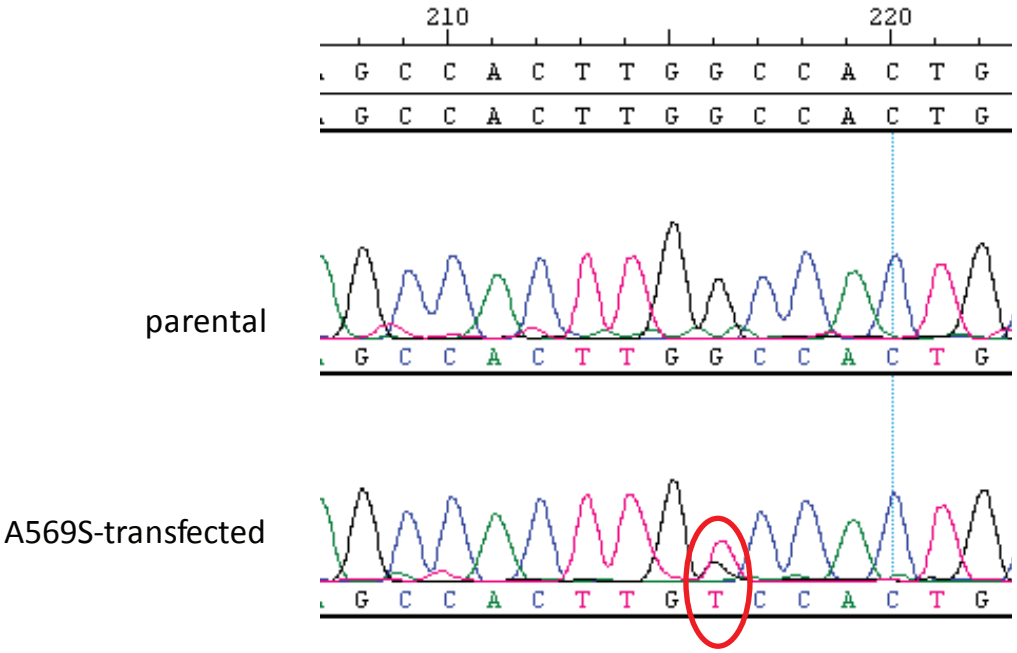
Pt17_A8_CTC

V.F. 0.0

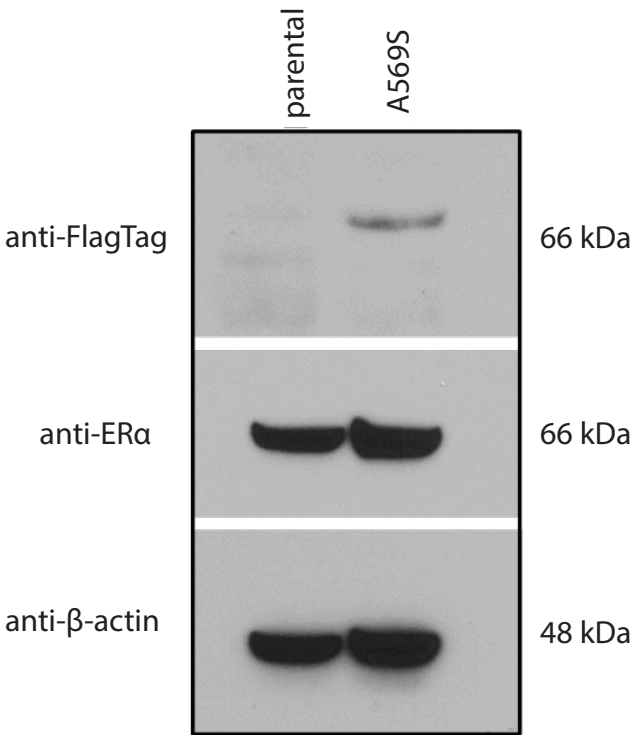


Pt17_WBC_Pool

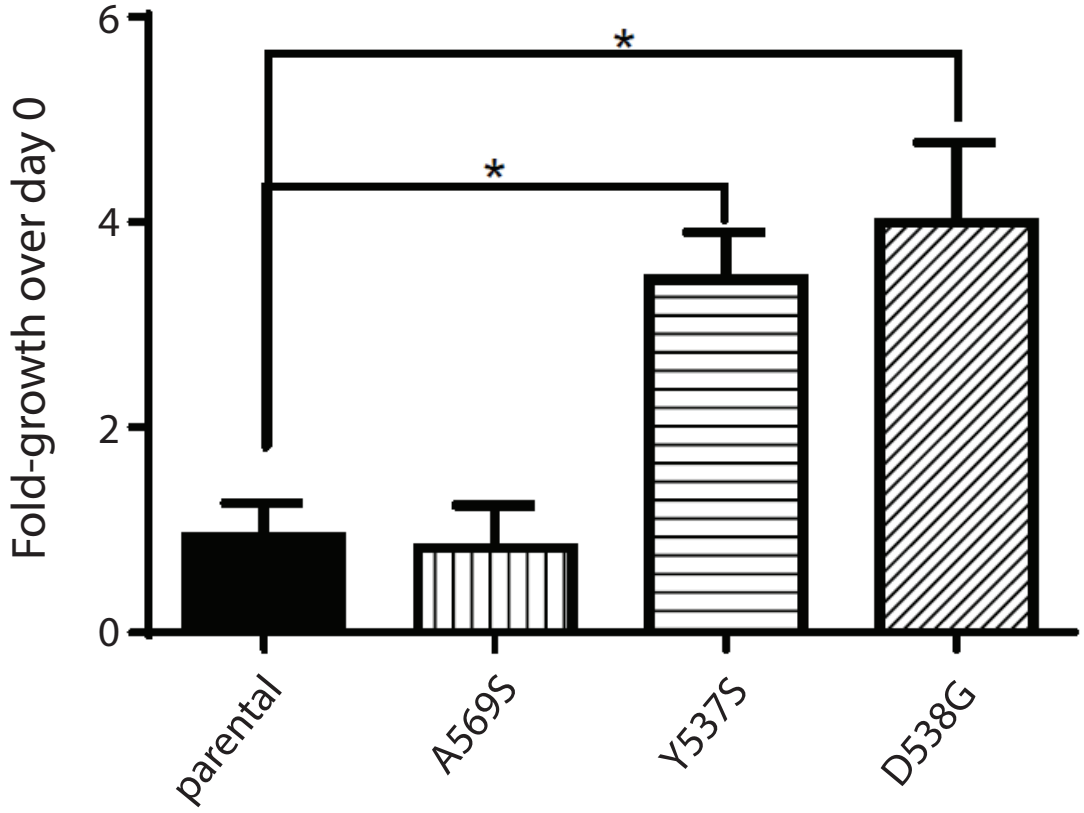
B.


210
220
G C C A C T T G G C C A C T G
parental
A569S-transfected

C.


parental
A569S
anti-FlagTag
66 kDa
anti-ERα
66 kDa
anti-β-actin
48 kDa

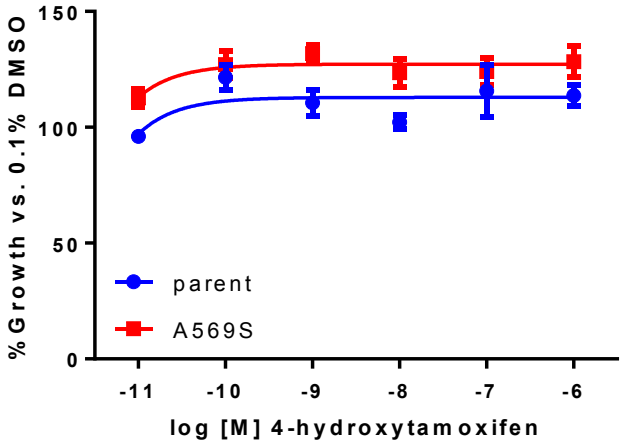
Supplementary Figure S3



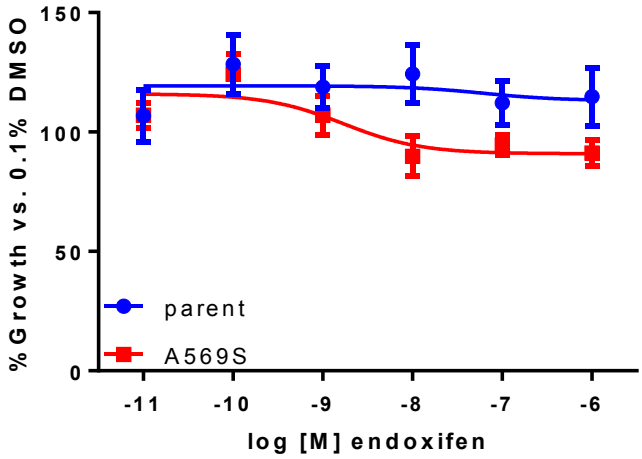
* $p < 0.0001$

Supplementary Figure S4

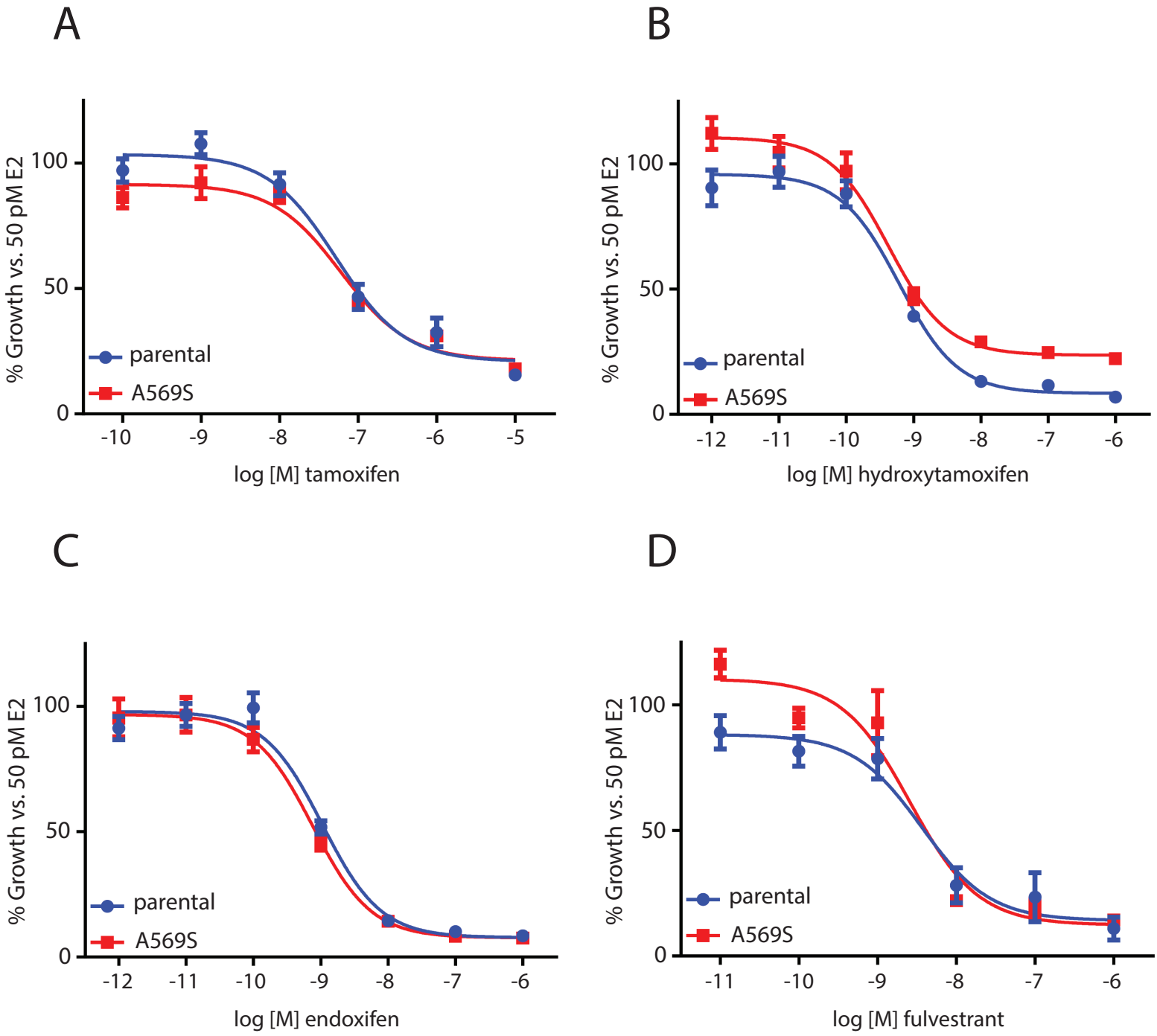
A



B



Supplementary Figure S5



Supplementary Table S1. Oncomine Comprehensive Panel (OCP) target genes by presence of the MseI restriction site

Unaffected Genes	*ARAF, BTK, FOXL2, GNA11, IDH2, IFITM1, IFITM3, KNSTRN, MAP2K2, MAPK1, MLH1, MPL, RHOA, STAT3, HNF1A, HRAS, JAK1, MAX, PPP2R1A, SRC, CSF1R, ESR1, MED12, NKX2-8, NRAS, SMO*
Partially Affected Genes	*ABL1, ACVRL1, AKT1, ALK, APC, APEX1, AR, ATM, ATP11B, BAP1, BCL2L1, BCL9, BIRC2, BIRC3, BRAF, BRCA1, BRCA2, CBL, CCND1, CCNE1, CD274, CD44, CDH1, CDK4, CDK6, CDKN2A, CSNK2A1, CTNNB1, DCUN1D1, DDR2, DNMT3A, EGFR, ERBB2, ERBB3, ERBB4, FBXW7, FGFR1, FGFR2, FGFR3, FGFR4, FLT3, GAS6, GATA2, GATA3, GNAS, IDH1, IGF1R, IL6, JAK3, KDR, KIT, KRAS, MAP2K1, MCL1, MDM2, MDM4, MET, MSH2, MTOR, MYC, MYCL, MYCN, MYD88, MYO18A, NF1, NF2, NKX2-1, NOTCH1, PDCD1LG2, PDGFRA, PIK3CA, PIK3R1, PNP, PPARG, PTCH1, PTEN, PTPN11, RB1, RET, RPS6KB1, SMAD4, SMARCB1, SOX2, STK11, TERT, TET2, TIAF1, TP53, TSC1, TSC2, VHL, WT1, ZNF217*
Completely Affected Genes	*CHEK2, EZH2, GNAQ, JAK2, MAGOH, NFE2L2, NPM1, PAX5, RAC1, RAF1, RHEB, SF3B1, SPOP, U2AF1, XPO1*

Genes included in the Oncomine Comprehensive Panel (OCP) divided in three categories based on the presence of the MseI restriction site in at least one amplicon.

Supplementary Table S2. NGS of BT-474 cells spiked into blood and purified identifies known *TP53* mutation

Cartridge Number	Number of CTC Single (S) Pooled (P)	Elapsed time between CellSearch® and DEPArray	Storage condition	Genome integrity index (GII)	*TP53* (p.E285K) Variant frequency
1	1 (S)	3 days	4°C	4	1
	1 (S)	3 days		4	0.8
	1 (S)	3 days		4	1
2	1 (S)	6 months	4°C	3	1
	1 (S)	6 months		2	1
	1 (S)	6 months		3	1
	9 (P)	6 months		3	1
3	1 (S)	2 years	RT	1	*
	1 (S)	2 years		1	*

NGS of CellSearch®-enriched and DEPArray™-purified spiked in cells from the BT-474 breast cancer cell line with the CHPv2 panel. Number of cells per sample, storage time and conditions, DNA quality (4=high) and the detected variant frequency of the expected mutation are shown. RT= room temperature; * Low library quality and uniformity of coverage over the target regions

Supplementary Table S3. Results of CTC purification by DEPArray

Patient #	Timepoint	CTC Enumeration	Unassigned events	Processed by CellSearch	Processed by DEPArray	Time Between Assays (Days)	DEPArray High Quality CTC	DEPArray Low Quality CTC	DEPArray Total CTC	# of single cells sequenced	#of cells sequenced as pool(s)	Sequenced	% of purified CTC	% of high quality CTC/purified CTC	% of high quality CTC/total CTC
CTC-002 A	BL	78	63	1/22/2014	12/18/2014	330	14	4	18	7	6	Yes	23%	78%	18%
CTC-002 B	BL	57	40	1/22/2014	4/28/2016	827	10	12	22	10	12	Yes	39%	45%	18%
CTC-002 C	BL	57	61	1/22/2014	4/28/2016	827	9	13	22	10	10	Yes	39%	41%	16%
CTC-003	BL	6	55	1/23/2014	6/12/2015	505	0	2	2	0	0	No	33%	0%	0%
CTC-007 A	BL	15	1311	2/21/2014	4/15/2015	418	1	2	3	0	3	Yes	20%	33%	7%
CTC-007 B	BL	18	1513	2/20/2014	12/22/2014	305	2	5	7	1		Yes	39%	29%	11%
CTC-008 A	BL	40	379	2/21/2014	3/31/2015	403	2	16	18	0	3	Yes	45%	11%	5%
CTC-008 B	BL	41	361	2/21/2014	12/22/2014	304	0	17	17		0	No	41%	0%	0%
CTC-009	BL	12	68	2/24/2014	4/14/2015	414	0	10	10	1	0	Yes	83%	0%	0%
CTC-010	BL	14	147	2/26/2014	4/14/2015	412	0	16	16	0	0	No	114%	0%	0%
CTC-012	BL	13	31	3/12/2014	4/2/2015	386	1	5	6	1	0	Yes	46%	17%	8%
CTC-014	BL	16	45	4/17/2014	1/5/2015	263	8	1	9	0	3	Yes	56%	89%	50%
CTC-015	BL	5	39	5/2/2014	6/12/2015	406	0	0	0	0	0	No	0%	0%	0%
CTC-017	BL	21	242	7/16/2014	4/1/2015	259	4	16	20	0	4	Yes	95%	20%	19%
CTC-019 A	BL	317	6250	7/16/2014	12/29/2014	166	1	6	7	0	3	Yes	2%	14%	0%
CTC-019 B	BL	344	5404	7/16/2014	3/30/2015	257	2	14	16			Yes	5%	13%	1%
CTC-024	BL	275	71	11/11/2014	12/18/2014	37	14	8	22	0	13	Yes	8%	64%	5%
CTC-025	BL	8	75	7/14/2015	4/5/2016	266	5	8	13	5	5	Yes	163%	38%	63%
CTC-028*	BL	142	329	9/25/2015	4/5/2016	193	3	13	16	2	2	Yes	23%	19%	2%
CTC-029*	BL	81	58	10/9/2015	4/5/2016	179	15	1	16	4	8	Yes	40%	94%	19%
CTC-030*	BL	962	441	10/7/2015	5/23/2016	229	34	0	34	5	4	Yes	7%	100%	4%
CTC-004	P	9	38	12/16/2015	4/5/2016	111	6	7	13	3	3	Yes	144%	46%	67%
CTC-024*	P	126	291	2/19/2016	5/23/2016	94	5	5	10	2	3	Yes	16%	50%	4%
Mean													**47%**	**35%**	**14%**

CellSearch®-enriched cells were purified into individual cells by DEPArray™. Patient # and CellSearch® cartridge (by letter) are shown in left-most column. CTC enumeration, processing dates, DNA quality and sequencing status are shown. BL = baseline; P = progression; *For these samples, only 1/2 cartridge was purified by DEPArray, % of purified CTC was adjusted accordingly

Supplementary Table S4: Mean CTC Next Generation Sequencing parameters

Mapped Reads	Reads On Target	Mean Sequencing Depth	Sequencing Uniformity	Variants
3,683,759	96.93%	1,508	41.05%	151

Mean sequencing parameters for 91 samples including mapped reads (total number of targeted reads per sample), reads on target (percentage of total reads that mapp to regions targeted by the sequencing panel), mean sequencing depth (mean number of reads covering each targeted base), sequencing uniformity (percentage of targeted bases that have at least 0.2X of the mean sequencing depth), variants (total number of sequencing variants detected per sample by the somatic low stringency setting).

Supplementary Table S5. NGS-identified prioritized mutations in individual and pooled CTC samples

SAMPLE	Gene	Location	REFERENCE	ALTERNATE	Aminoacid	RefSeq	Exonic Function	Vartaiont Fraction	Varaint Reads	Total Reads
Pt02 5P R1-R7 C	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	246	246
Pt02 5P R9-R18 C	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	0.71	550	771
Pt02 6P A3-D5 A	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	1995	1995
Pt02 6P R1-R6 B	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	0.93	569	611
Pt02 6P R7-R20 B	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	21	21
Pt02 7P 3b A	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	1652	1652
Pt02 A12 A	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	1997	1997
Pt02 A9 A	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	1993	1993
Pt02 D1 A	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	1993	1993
Pt02 D3 A	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	1994	1994
Pt02 D6 A	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	1996	1996
Pt02 R1 C	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	441	441
Pt02 R11 B	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	12	12
Pt02 R11 C	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	0.52	438	846
Pt02 R12 C	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	1676	1676
Pt02 R13 B	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	120	120
Pt02 R14 B	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	37	37
Pt02 R2 B	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	262	262
Pt02 R3 B	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	1584	1584
Pt02 R7 B	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	0.91	21	23
Pt02 R3 C	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	489	489
Pt02 R9 C	CDH1	chr16:68856113	C	T	p.Q641X	NM_004360	stopgainSNV	1.00	116	116
Pt02 R17 B	CDH1	chr16:68863698	G	T	p.E813X	NM_004360	stopgainSNV	0.83	149	180
Pt02 5P R1-R7 C	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1980	1980
Pt02 5P R9-R18 C	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	0.44	867	1976
Pt02 6P A3-D5 A	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	0.98	1964	1997
Pt02 6P R1-R6 B	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	0.33	665	1997
Pt02 6P R7-R20 B	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1990	1990
Pt02 7P 3b A	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	0.99	1981	1998
Pt02 A12 A	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1990	1990
Pt02 D1 A	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1991	1991
Pt02 D2 A	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1989	1989
Pt02 D3 A	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1990	1990
Pt02 D6 A	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1989	1989
Pt02 R1 B	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1984	1984
Pt02 R11 B	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1980	1980
Pt02 R11 C	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	0.52	1029	1986
Pt02 R12 C	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1983	1983
Pt02 R13 B	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1981	1981
Pt02 R14 B	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1991	1991
Pt02 R2 B	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1974	1974
Pt02 R20 B	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1989	1989
Pt02 R3 C	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1980	1980
Pt02 R4 C	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1989	1989
Pt02 R5 B	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1977	1977
Pt02 R6 C	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1981	1981
Pt02 R7 B	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1985	1985
Pt02 R7 C	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1976	1976
Pt02 R9 C	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	1.00	1991	1991
Pt02 R18 C	CDH1	chr16:68835618	C	T	p.S70F	NM_004360	nonsynonymousSNV	0.79	11	14
Pt02 A12 A	ESR1	chr6:152420018	G	T	p.A569S	NM_000125	nonsynonymousSNV	0.56	1117	1994
Pt02 5P R1-R7 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.41	819	1998
Pt02 5P R9-R18 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.40	797	2000
Pt02 6P A3-D5 A	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.35	693	1997
Pt02 6P R1-R6 B	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.15	299	2000
Pt02 6P R7-R20 B	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.46	911	1998
Pt02 7P 3b A	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.47	947	1997
Pt02 A9 A	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.52	1046	1995
Pt02 D1 A	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.34	676	1998
Pt02 D2 A	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.46	920	1995

SAMPLE	Gene	Location	REFERENCE	ALTERNATE	Aminoacid	RefSeq	Exonic Function	Vartaiont Fraction	Varaint Reads	Total Reads
Pt02 D6 A	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.65	1288	1994
Pt02 R1 B	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	1.00	715	715
Pt02 R1 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.37	730	1999
Pt02 R11 B	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.10	631	5645
Pt02 R11 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.49	988	1999
Pt02 R12 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.48	956	2000
Pt02 R13 B	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.68	1357	1999
Pt02 R14 B	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.52	1030	2000
Pt02 R17 B	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	1.00	1961	1961
Pt02 R2 B	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	1.00	2000	2000
Pt02 R20 B	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.51	1021	2000
Pt02 R3 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.87	965	1110
Pt02 R4 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.45	909	1999
Pt02 R6 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.31	623	1996
Pt02 R7 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.39	782	2000
Pt02 R9 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	1.00	1998	1998
Pt02 R18 C	ESR1	chr6:152419923	A	C	p.Y537S	NM_000125	nonsynonymousSNV	0.83	10	12
Pt04 (3P)	PIK3CA	chr3:178936091	G	A	p.E545K	NM_006218	nonsynonymousSNV	0.53	190	361
Pt04 R3	PIK3CA	chr3:178936091	G	A	p.E545K	NM_006218	nonsynonymousSNV	1.00	1138	1996
Pt04 (3P)	TP53	chr17:7577506	C	G	p.D259H	NM_000546	nonsynonymousSNV	0.43	857	1821
Pt04 R3	TP53	chr17:7577506	C	G	p.D259H	NM_000546	nonsynonymousSNV	1.00	1986	1138
Pt04 (3P)	TP53	chr17:7577099	C	G	p.R280T	NM_000546	nonsynonymousSNV	0.16	292	1986
Pt04 R3	TP53	chr17:7577099	C	G	p.R280T	NM_000546	nonsynonymousSNV	0.65	1294	1998
Pt12 A9	PIK3CA	chr3:178952085	A	G	p.H1047R	NM_006218	nonsynonymousSNV	0.85	1703	1997
Pt12 A9	TP53	chr17:7577538	C	T	p.R248Q	NM_000546	nonsynonymousSNV	0.73	1457	1999
Pt14 (3P)	HNF1A	chr12:121431414	G	T	p.W206C	NM_000545	nonsynonymousSNV	0.17	270	1569
Pt17 (4P)	BRCA2	chr13:32914283	C	T	p.Q1931X	NM_000059	stopgainSNV	0.10	88	878
Pt19 (3P)	PTCH1	chr9:98211431	C	A	p.E1242X	NM_000264	stopgainSNV	0.28	390	1378
Pt24B (4P2)	CDH1	chr16:68867274	G	T	p.E841X	NM_004360	stopgainSNV	0.14	135	952
Pt24 R2	CDH1	chr16:68855944	TC	T	p.I584fs	NM_004360	frameshiftdeletion	1.00	1007	1007
Pt24 R3	CDH1	chr16:68855944	TC	T	p.I584fs	NM_004360	frameshiftdeletion	1.00	1978	1978
Pt24B (4P1)	CDH1	chr16:68855944	C	-	p.I584fs	NM_004360	frameshiftdeletion	0.69	1361	1986
Pt24B (4P2)	CDH1	chr16:68855944	C	-	p.I584fs	NM_004360	frameshiftdeletion	0.77	1536	1992
Pt24B (5P)	CDH1	chr16:68855944	C	-	p.I584fs	NM_004360	frameshiftdeletion	0.79	1568	1977
Pt24P (3P)	CDH1	chr16:68855944	TC	T	p.I584fs	NM_004360	frameshiftdeletion	0.45	749	1648
Pt24B (4P2)	NOTCH1	chr9:139390716	G	T	p.S2492X	NM_017617	stopgainSNV	0.18	350	1998
Pt24 R2	TP53	chr17:7578464	CGGGTGCCGGGCGG	C	p.152_156del	NM_000546	frameshiftdeletion	1.00	1971	1971
Pt24 R3	TP53	chr17:7578464	CGGGTGCCGGGCGG	C	p.152_156del	NM_000546	frameshiftdeletion	0.98	1932	1971
Pt24B (4P1)	TP53	chr17:7578464	GGGTGCCGGGCGG	-	p.152_156del	NM_000546	frameshiftdeletion	0.30	578	1938
Pt24B (4P2)	TP53	chr17:7578464	GGGTGCCGGGCGG	-	p.152_156del	NM_000546	frameshiftdeletion	0.88	1702	1941
Pt24B (5P)	TP53	chr17:7578464	GGGTGCCGGGCGG	-	p.152_156del	NM_000546	frameshiftdeletion	0.94	1831	1952
Pt24P (3P)	TP53	chr17:7578464	CGGGTGCCGGGCGG	C	p.152_156del	NM_000546	frameshiftdeletion	0.51	1003	1958
Pt24B (4P2)	TP53	chr17:7578203	C	A	p.V216L	NM_000546	nonsynonymousSNV	0.36	286	784
Pt24B (5P)	TET2	chr4:106196820	G	T	p.V1718L	NM_001127208	nonsynonymousSNV	0.90	1207	1334
Pt24B (4P1)	TET2	chr4:106196820	G	T	p.V1718L	NM_001127208	nonsynonymousSNV	0.53	1303	2449
Pt24B (4P2)	TET2	chr4:106196820	G	T	p.V1718L	NM_001127208	nonsynonymousSNV	0.70	346	497
Pt25 R4	TP53	chr17:7578395	T	C	p.H179R	NM_000546	nonsynonymousSNV	0.99	3084	3113
Pt25 (3P)	TP53	chr17:7578395	T	C	p.H179R	NM_000546	nonsynonymousSNV	0.61	1057	1723
Pt28 (2P)	TP53	chr17:7577547	C	A	p.G245V	NM_000546	nonsynonymousSNV	1.00	1987	1987
Pt28 R3	PIK3CA	chr3:178936083	G	A	p.E542K	NM_006218	nonsynonymousSNV	0.20	86	437
Pt28 (2P)	PIK3CA	chr3:178936083	G	A	p.E542K	NM_006218	nonsynonymousSNV	0.55	296	539
Pt28 R3	TP53	chr17:7577547	C	A	p.G245V	NM_000546	nonsynonymousSNV	1.00	1987	1987
Pt30 R1	PIK3CA	chr3:178952085	A	G	p.H1047R	NM_006218	nonsynonymousSNV	0.52	330	636
Pt30 R2	PIK3CA	chr3:178952085	A	G	p.H1047R	NM_006218	nonsynonymousSNV	0.37	212	567
Pt30 R3	PIK3CA	chr3:178952085	A	G	p.H1047R	NM_006218	nonsynonymousSNV	0.54	360	661
Pt30 R5	PIK3CA	chr3:178952085	A	G	p.H1047R	NM_006218	nonsynonymousSNV	0.65	589	900
Pt30 R10	PIK3CA	chr3:178952085	A	G	p.H1047R	NM_006218	nonsynonymousSNV	0.49	311	635
Pt30 (4P)	PIK3CA	chr3:178952085	A	G	p.H1047R	NM_006218	nonsynonymousSNV	0.48	260	536
Pt30 R1	ESR1	chr6:152419926	A	G	p.D538G	NM_000125	nonsynonymousSNV	1.00	2829	2843
Pt30 R2	ESR1	chr6:152419926	A	G	p.D538G	NM_000125	nonsynonymousSNV	0.29	1050	3634
Pt30 R3	ESR1	chr6:152419926	A	G	p.D538G	NM_000125	nonsynonymousSNV	1.00	2589	2599
Pt30 R5	ESR1	chr6:152419926	A	G	p.D538G	NM_000125	nonsynonymousSNV	1.00	3137	3151
Pt30 (4P)	ESR1	chr6:152419926	A	G	p.D538G	NM_000125	nonsynonymousSNV	1.00	2550	2562

All high confidence, non-synonymous variants identified in CTC samples are shown. For each variant, the gene, location (hg19), reference and variant (Alternate) alleles, amino acid (AA) change, RefSeq ID, variant type, variantiant read frequency (flow corrected variant over total reads) and flow-corrected variant and total reads separately are shown. See supplementary methods for variant filtering and prioritization. Each variant location was visually checked in IGV (Integrated Genome Viewer) in all samples from each patient who harbored at least one sample with that variant in order to confirm adequate coverage and

Supplementary Table S6. 2 x 2 contingency concordance tables for alterations in CTC vs tissue

A. CTC vs. tissue mutation concordance

		Tissue			
		Y	N	Total	
CTC	Y	14	9	23	61%
	N	9	0	9	
	Total	23	9	32	
		61%			

B. CTC vs. tissue CNA concordance

		Tissue			
		Y	N	Total	
CTC	Y	19	12	31	61%
	N	7	0	7	
	Total	26	12	38	
		73%			

Legend: Cohort-wide numbers of alterations for **A**. Mutations and **B**. Copy number alterations (CNA) that were detected by either method, showing numbers and percentages of alterations concordant in either direction. Y= Yes (detected), N= No (not detected)

Supplementary Table S7. Patient #2 analysis of primary, clinical metastatic tissue, research biopsy, pt-DNA by ddPCR

Time of sampling	Initial Diagnosis		1st clin. met. 2009		Later metastasis		Later metastasis			
Date	7/1/97		7/1/09		7/19/13		1/24/14			
Site	Breast Primary		Omental		Pleural effusion		Blood			
Type of specimen	FFPE		FFPE		cytospin		CTC		cfDNA	
Assay	ddPCR	WES	ddPCR	WES	ddPCR	WES	Targeted NGS	Sanger	ddPCR	WES
ESR1										
Y537S	WT		mut (45%)		mut (45%)	mut (44.5%)	mut 26 /32 (81%)	mut 4/6 (67%)	mut (0.87%)	
D538G	mut (0.01%)		mut (0.01%)		mut (0.02%)	WT	mut 0/32 (0%)	mut 0/6 (0%)	mut (0.01%)	
A569S	WT		WT (few droplets +**)		WT (few droplets +**)	WT	mut 1/32 (3%)	mut 1/6 (17%)	WT	

For each timepoint, specimen type and detection technique, presence/absence (and variant frequency) are shown for all three ESR1 mutations. FFPE = Formalin-Fixed Paraffin-Embedded; ddPCR = droplet digital PCR; WES = whole exome sequencing; mut = mutation present; WT = Wild Type; **a few droplets mutated for specific mutation, but below the predetermined threshold. Shaded area = assay not performed.

With this approach, we detected, albeit at extremely low levels, an ESR1 p.D538G hotspot mutation that was absent from both CTC and tissue by the respective NGS techniques applied. ddPCR droplets carrying this mutation were present at VF 0.0001 – 0.0002 (0.01 – 0.02%) but consistently appeared in the primary tumor, two consecutive metastases, and later in plasma derived cfDNA collected concurrently with CTC. In concordance with CTC and tissue profiling, ddPCR of tissues and cfDNA detected the ESR1 p.Y537S mutation only in post-ET specimens. Interestingly, the novel ESR1 p.A569S mutation detected in one CTC was also only present in a few droplets, but below the predetermined detection threshold [Chu, Clin Cancer Res 2016; Beaver, Clin Cancer Res 2014] in post-ET specimens.

Supplementary Table S8. Primer/probe sequences used in Sanger sequencing and ddPCR

Pre-amplification Primers Sequences	**Forward**	**Reverse**
ESR1 Y537S/D538G	TCCCAGCTCCCATCCTAAAGTG	CACTGCGGGCTCTACTTCAT
ESR1 A569S	TGCTGCTGGAGATGCTGGAC	CCTCCCCCGTGATGTAATACT
ddPCR Primers Sequences	**Forward**	**Reverse**
ESR1 Y537S/D538G	GCATGAAGTGCAAGAACGTG	AGACGGACCAAAGCCACTT
ESR1 A569S	TACATGCGCCCACTAGCCGT	GCAAGGAATGCGATGAAGTA
ddPCR Probes Sequences	**Mutant**	**Wild type**
ESR1 Y537S	TCTcTGACCTGCTGCTGGAGATGCT	TCTATGACCTGCTGCTGGAGATGCT
ESR1 D538G	TCTATGgCCTGCTGCTGGAGATGCT	TCTATGACCTGCTGCTGGAGATGCT
ESR1 A569S	ACCAAAGCCACTTGtCCACTGCGG	ACCAAAGCCACTTGGCCACTGCGG