Published in final edited form as:
Nat Genet. 2013 December ; 45(12): 1446–1451. doi:10.1038/ng.2823.

Activating ESR1 mutations in hormone-resistant metastatic breast cancer

Dan R Robinson[1,2], **Yi-Mi Wu**[1,2], **Pankaj Vats**[1,2], **Fengyun Su**[1,2], **Robert J Lonigro**[1,3], **Xuhong Cao**[1,4], **Shanker Kalyana-Sundaram**[1,2], **Rui Wang**[1,2], **Yu Ning**[1,2], **Lynda Hodges**[1], **Amy Gursky**[1,2], **Javed Siddiqui**[1,2], **Scott A Tomlins**[1,2], **Sameek Roychowdhury**[5], **Kenneth J Pienta**[6], **Scott Y Kim**[7], **J Scott Roberts**[8], **James M Rae**[3,9], **Catherine H Van Poznak**[9], **Daniel F Hayes**[9], **Rashmi Chugh**[9], **Lakshmi P Kunju**[1,2], **Moshe Talpaz**[9], **Anne F Schott**[9], and **Arul M Chinnaiyan**[1,2,3,4,10,11]

[1]Michigan Center for Translational Pathology, University of Michigan Medical School, Ann Arbor, MI 48109, USA

[2]Department of Pathology, University of Michigan Medical School, Ann Arbor, MI 48109, USA

[3]Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, MI 48109, USA

[4]Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, MI 48109, USA

[5]Department of Internal Medicine, Ohio State University, Columbus, Ohio 43210, USA

[6]Brady Urological Institute, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

[7]Center for Bioethics and Social Science in Medicine, University of Michigan, Ann Arbor, MI 48109, USA

[8]Department of Health Behavior & Health Education, University of Michigan, Ann Arbor, MI 48109, USA

[9]Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI 48109, USA

ACCESSION CODES

Sequence data have been deposited at the dbGAP, which is hosted by the National Center for Biotechnology Information (NCBI), under accession dbGAP phs000602.v1.p1, and CSER Clinical Sequencing Exploratory Research Program for the NIH-NHGRI grant (1UM1HG006508).

AUTHOR CONTRIBUTIONS

D.R.R., Y.M.W., and A.M.C. conceived the experiments. D.R.R., Y.M.W., X.C., R.W., F.S., and Y.N. performed exome and transcriptome sequencing. P.V., R.J.L., S.K.S., and D.R.R. carried out bioinformatics analysis of high throughput sequencing data for somatic mutation, copy number and tumor content determination, gene expression and gene fusion analysis. D.R.R., Y.M.W. and F.S. generated ESR1 constructs and carried out in vitro experiments. L.H. coordinated patients for clinical research. J.S. and A.G. collected and processed clinical tissue samples for next generation sequencing. L.P.K. and S.A.T. provided pathology review. J.M.R. provided experimental analysis. C.H.V.P., D.F.H., R.C., and A.F.S. enrolled patients and provided clinical data and consultation at tumor boards. D.R.R., X.C., Y.M.W., P.V., R.J.L., S.K.S., J.S.R., S.R., M.T., K.J.P., and A.M.C. developed the integrated clinical sequencing protocol. D.R.R., Y.M.W., and A.M.C. prepared the manuscript, which was reviewed by all authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

[10]Department of Urology, University of Michigan Medical School, Ann Arbor, MI 48109 USA

[11]Center for Computational Medicine and Biology, University of Michigan, Ann Arbor, MI 48109, USA

Breast cancer is the most prevalent cancer in women, and over two-thirds of cases express estrogen receptor-a (ER-a, encoded by ESR1). Through a prospective clinical sequencing program for advanced cancers, we enrolled 11 patients with ER-positive metastatic breast cancer. Whole-exome and transcriptome analysis showed that six cases harbored mutations of ESR1 affecting its ligand-binding domain (LBD), all of whom had been treated with anti-estrogens and estrogen deprivation therapies. A survey of The Cancer Genome Atlas (TCGA) identified four endometrial cancers with similar mutations of ESR1. The five new LBD-localized ESR1 mutations identified here (encoding p.Leu536Gln, p.Tyr537Ser, p.Tyr537Cys, p.Tyr537Asn and p.Asp538Gly) were shown to result in constitutive activity and continued responsiveness to anti-estrogen therapies in vitro. Taken together, these studies suggest that activating mutations in ESR1 are a key mechanism in acquired endocrine resistance in breast cancer therapy.

Advances in high-throughput sequencing technologies are beginning to establish a molecular taxonomy for a spectrum of human diseases and has facilitated a move toward ‘precision medicine’ (refs. 1,2). With regard to oncology, defining the mutational landscape of a patient’s tumor will lead to more precise treatment and management of individuals with cancer. Comprehensive clinical sequencing programs for cancer patients have been initiated at a variety of medical centers, including our own [3,4]. In addition to the potential for identifying ‘actionable’ therapeutic targets in cancer patients, these clinical sequencing efforts may also shed light on acquired resistance mechanisms developed against targeted therapies [5–7].

ER is the primary therapeutic target in breast cancer and is expressed in 70% of cases [8]. Drugs directly antagonizing ER, such as tamoxifen and fulvestrant, are a mainstay of breast cancer treatment; however, approximately 30% of ER-positive breast cancers exhibit de novo resistance, whereas 40% acquire resistance to these therapies [9]. In addition to anti-estrogen therapies, patients with ER-positive breast cancer are also treated with aromatase inhibitors such as letrozole and exemestane [10]. Aromatase inhibitors block the peripheral conversion of androgens into estrogen and, in post-menopausal women, lead to over a 98% decrease in circulating levels of estrogen. As with anti-estrogens, treatment with aromatase inhibitors results in the development of resistance, but this is presumably due to different mechanisms, as patients with breast cancer who develop resistance to aromatase inhibitors often still respond to anti-estrogen therapies [11]. The molecular mechanisms of endocrine resistance in ER-positive breast cancer continues to be an active area of research [12].

Our institutional review board (IRB)-approved clinical sequencing program, called MI-ONCOSEQ (the Michigan Oncology Sequencing Program), enrolls patients with advanced cancer across all histologies[3]. Since April 2011, we have enrolled over 200 patients in this program, which involves obtaining a current tumor biopsy with matched normal samples (blood and/or buccal swab). Samples are then subjected to integrative sequencing, which

includes whole-exome sequencing of the tumor and matched normal sample, transcriptome sequencing and, as needed, low-pass whole-genome sequencing [3]. This combination of DNA and RNA sequencing technologies allows one to be relatively comprehensive with regard to the mutational landscape of coding genes, including analysis of point mutations, indels, amplifications, deletions, gene fusions or translocations, and outlier gene expression profiles. These results are generated within a 5- to 7-week time frame and are presented at an institutional 'precision medicine tumor board' to deliberate upon potentially actionable findings.

As part of the MI-ONCOSEQ program, we enrolled and sequenced 11 patients with metastatic ER-positive breast cancer (Table 1 and Supplementary Table 1). A diverse array of aberrations were identified in individual patients, some of which are potentially actionable, including mutations in PIK3CA (n = 4), BRCA1 aberrations (n = 2), FGFR2 aberrations (n = 2)[13], NOTCH2 frameshift deletion (n = 1), cyclin and associated cyclin-dependent kinase aberrations (n = 3) and MDM2 amplification and overexpression (n = 1). Aberrations were also frequently found in the tumor suppressor TP53 (n = 6), the DNA mismatch repair gene MSH2 (n = 1) and in epigenetic regulators (n = 2), including ARID2, ARID1A and SMARCA4, among others. The complete spectra of somatic mutations with associated alterations in expression levels and copy number in the index patients are given in Supplementary Figure 1 and Supplementary Tables 2 and 3. Two of the index patients, MO_1031 and MO_1051, exhibited a high level of mutations consistent with 'signature B' identified in a whole-genome study of mutational processes in breast cancer[14]. There were 39 gene fusions identified in the 6 index patients, with 11 encoding in-frame fusion proteins (Supplementary Fig. 2 and Supplementary Tables 4 and 5), including an activating FGFR2-AFF3 fusion[13].

The most notable observation in the mutational landscapes of these treated patients with ER-positive breast cancer was the finding of nonsynonymous mutations in ESR1 affecting the LBD (n = 6). The six index patients MO_1031, MO_1051, MO_1069, MO_1129, MO_1167 and MO_1185 had mutations encoding p.Leu536Gln, p.Tyr537Ser, p.Asp538Gly, p.Tyr537Ser, p.Asp538Gly and p.Tyr537Ser alterations in the LBD, respectively. The respective mutation in each case was detected by whole-exome sequencing of the tumor relative to the matched normal sample and was corroborated by whole-transcriptome sequencing, as ESR1 was expressed at moderate to high levels (Supplementary Table 2). The clinical histories of the index patients are depicted in timelines in Figure 1. For three of the patients (MO_1051, MO_1069 and MO_1129), we had access to primary diagnostic material and showed that the ESR1 mutations were not present at an earlier stage, indicating that they were acquired after endocrine therapy (Fig. 1 and Supplementary Table 2). Interestingly, all of the index patients were treated with anti-estrogens (tamoxifen and/or fulvestrant) and aromatase inhibitors (letrozole, anastrozole and/or exemestane). Two of the patients also had an oophorectomy. Comparison of the mutations present in each primary versus post-treatment pair showed a substantial number of shared mutations in both samples of the pair, including activating mutations in PIK3CA in two of the cases. Thus, it is clear that the index cases presented with recurrent disease of the original primary tumor surviving in an estrogen-deprived state and having acquired ESR1 mutations. Of note, neither ESR1 amplifications nor gene fusions were observed in these cases.

The five new LBD alterations of ESR1 identified in this study are depicted in Figure 2. Each occurred in the vicinity of the synthetic alterations of ESR1 that are inverted in response to tamoxifen and involve p.Met543Ala and p.Leu544Ala alterations (Inv-mut-AA2)[15] and served as a positive control for our subsequent in vitro studies. We next investigated the occurrence of ESR1 mutations in a range of breast cancer types. Here we took advantage of data from the TCGA Project, which has generated whole-exome sequences for 27 tumor types across at least 4,000 individual samples. As expected, LBD-disrupting mutations of ESR1 were not detected in the 390 ER-positive breast cancers sequenced by TCGA, as these were primary resection samples before hormonal treatment[16], nor did we detect ESR1 mutations in a cohort of 80 triple-negative breast carcinoma transcriptomes (D.R.R., Y.-M.W., X.C., S.K.-S., A.M.C. et al., unpublished data).

As the LBD-disrupting mutations of ESR1 we identified were somatic and were acquired after treatment, we next assessed whether the encoded proteins were dependent on estrogen for activation. We cloned into expression vectors each of the five ESR1 mutants identified in this study (encoding p.Leu536Gln, p.Tyr537Ser, p.Asp538Gly, p.Tyr537Cys and p.Tyr573Asn alterations) and subsequently cotransfected these constructs into HEK293T cells with an estrogen response element (ERE)-luciferase reporter system. We then exposed steroid hormone–deprived cells to β-estradiol for 24 h and assessed ERE reporter levels. Surprisingly, unlike wild-type ESR1, which had little ERE reporter activity in the absence of ligand, all five of the ESR1 mutants had strong constitutive activation of the ERE reporter that was not markedly enhanced with β-estradiol (Fig. 3). This finding suggested that each of the mutations developed in the context of evolution during an estrogen-deprived state. Consistent with this idea, a whole-genome sequencing study of 46 patients with ER-positive breast cancer enrolled in 2 aromatase inhibitor trials did not identify any of these ESR1 mutations in the pretreatment samples analyzed[17].

Next, we assessed whether anti-estrogen therapies affected the functional activity of the LBD mutants. As effects on inhibition can be influenced by the levels of ectopic ER expression, we performed a dose response study with expression plasmid and selected a dose of 50 ng for the following experiments[18] (Supplementary Fig. 3). As expected, wild-type ESR1 was inhibited in a dose-dependent fashion by the anti-estrogens 4-hydroxytamoxifen, fulvestrant and endoxifen (Fig. 4 and Supplementary Figs. 4–6). In addition, the mutant corresponding to the synthetic ESR1 mutation (Inv-mut-AA2) was activated in a dose-dependent fashion by these anti-estrogens (Fig. 4), which has been reported previously[15]. Interestingly, ESR1 with each of the five LBD alterations identified in this study was inhibited by tamoxifen and fulvestrant in a dose-dependent fashion and did not exhibit the inverted response to anti-estrogens that the synthetic Inv-mut-AA2 mutant did. One could speculate that the corresponding mutations did not arise under selective pressure of anti-estrogen treatment but rather in the context of an estrogen deprivation setting, such as treatment with aromatase inhibitors and/or oophorectomy. The IC_{50} (half-maximal inhibitory concentration) values for both 4-hydroxytamoxifen and fulvestrant were two- to fourfold higher for all mutants compared to wild-type ESR1. Fulvestrant exhibited greater maximal inhibition than 4-hydroxytamoxifen for all the mutants tested (Supplementary Figs. 4 and 5).

The ESR1 alterations identified in this study cluster near the beginning of helix 12 (Fig. 2). Structural studies have demonstrated a key role for the position of helix 12 in the response of the ER to agonists and antagonists[19], and Tyr537 has been postulated to form a capping motif contributing to the activity of the receptor[20]. Specifically, the p.Tyr537Ser mutant has been reported to have higher affinity for estrogen than wild-type ESR1 and interacts with the SRC1 coactivator in the absence of ligand[21,22]. Several studies using experimental mutagenesis have implicated the same three residues identified here as critical determinants of the transcriptional activity of the receptor [21,23,24].

As estrogen therapy has been shown to have a positive effect in treating aromatase inhibitor–resistant advanced breast cancers, we tested the effect of low- to high-dose estrogen on the activity of the mutants in the transient luciferase reporter assays (Supplementary Fig. 7)[25,26]. The results did not suggest that the effectiveness of this therapy is mediated through direct control of the transcriptional activity of these mutants, if encoded by the responding patients.

Although the primary intent of our broad-based clinical sequencing program is to identify actionable and/or driver mutations in advanced cancers, this study demonstrates how such prospective, real-time sequencing efforts can also shed light on resistance mechanisms that develop against targeted therapies. A number of resistance mechanisms have been suggested to function in the evasion of endocrine treatment, including activation of the mTOR and phosphoinositide 3-kinase (PI3K) pathways, among others [9,27]. Although the total number of ER-positive breast cancers we have sequenced is modest, we have done so in a comprehensive fashion in terms of delineating mutational landscapes and incorporating both DNA and RNA sequencing. This analysis identified de novo driver mutations and/or potentially acquired mutations in breast cancer such as mutations resulting in PI3K activation, PAK1 amplification and FGFR fusion and amplification, which have been described previously [13,28,29]. Among potential new mechanisms described, we identified profound focal amplification of MDM2 (which encodes a negative regulator of p53 that is targetable) and copy gains of GNRHR (encoding gonadotropin-releasing hormone receptor), which may be related to past endocrine therapy.

As the LBD-disrupting mutations of ESR1 identified in this study result in constitutive activity, the encoded mutant proteins can function in the absence of ligand and maintain ER signaling. In 1997, an ESR1 mutation affecting the LBD, encoding a p.Tyr537Asn alteration, was detected in a single individual with stage IV metastatic breast cancer who had been treated with diethylstibesterol, but, since then, this mutation has been considered to be very rare[30]. With the advent of widespread aromatase inhibitor therapy, we suggest that alteration of the ESR1 LBD is likely a common mechanism of resistance that develops in low-estrogen states. Interestingly, LBD-disrupting mutations of ESR1 were detected somatically in 4 of 373 cases of endometrial cancer[31]. We speculate that the four TCGA endometrial tumors that harbor LBD-affecting mutations likely came from patients with concurrent breast cancer, as tamoxifen treatment is known to be associated with higher incidence of this tumor type and such patients also often receive estrogen deprivation treatment[32].

Our study suggests that it is unlikely that these LBD alterations develop in the context of anti-estrogen treatment, as the mutated ESR1 variants continue to be responsive to direct ER antagonists such as tamoxifen and fulvestrant. This finding is consistent with clinical reports showing that patients that develop resistance to aromatase inhibitors still respond to anti-estrogen treatment[11]. Although this prospective clinical sequencing study was not designed to characterize a specific cancer type or treatment resistance mechanism, future studies comprising larger cohorts of breast cancer patients with disease that recurs after varied endocrine treatments will more precisely delineate the incidence of this acquired resistance mechanism. The focused nature of these mutations and their role in aromatase inhibitor resistance suggest the possibility of monitoring patients undergoing treatment using circulating tumor DNA methods [33,34]. In this manner, treatment could be shifted to head off evolving tumor resistance.

ONLINE METHODS

Clinical study and specimen collection

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 (Michigan Oncology Sequencing Protocol, HUM00046018). Medically qualified patients 18 years or older with advanced or refractory cancer were eligible for the study. Informed consent detailed the risks of integrative sequencing and includes up-front genetic counseling. Informed consent was obtained from all subjects included in this study. Biopsies were arranged for safely accessible tumor sites. Needle biopsies were snap frozen in OCT (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.

Extraction of DNA and RNA

Genomic DNA from frozen needle biopsies and blood was isolated using the Qiagen DNeasy Blood and Tissue kit, according to the manufacturer's instructions. Total RNA was extracted from frozen needle biopsies using the Qiazol reagent with disruption using a 5-mm bead on a Tissuelyser II (Qiagen) and was purified using a miRNeasy kit (Qiagen) with DNase I digestion, according to the manufacturer's instructions. RNA integrity was verified on an Agilent 2100 Bioanalyzer using RNA Nano reagents (Agilent Technologies).

Preparation of next-generation sequencing libraries

Transcriptome libraries were prepared using 1–2 μg of total RNA. Polyadenylated RNA was isolated using Sera-Mag oligo(dT) beads (ThermoScientific) and fragmented with the Ambion Fragmentation Reagents kit. cDNA synthesis, end repair, A-base addition and ligation of the Illumina indexed adaptors were performed according to Illumina's TruSeq RNA protocol. Libraries were selected for DNA fragments of 250–300 bp in size on a 3% Nusieve 3:1 agarose gel (Lonza), recovered using QIAEX II gel-extraction reagents (Qiagen) and PCR amplified using Phusion DNA polymerase (New England BioLabs). Amplified libraries were purified using AMPure XP beads (Beckman Coulter). Library

quality was measured on an Agilent 2100 Bioanalyzer by product size and concentration. Paired-end libraries were sequenced with the Illumina HiSeq 2000 platform (2 × 100-nucleotide read length). Reads that passed the chastity filter of Illumina BaseCall software were used for subsequent analysis.

Exome libraries of matched pairs of tumor and normal genomic DNA were generated using the Illumina TruSeq DNA Sample Prep kit, following the manufacturer's instructions. In brief, 1–3 μg of each genomic DNA sample was sheared using a Covaris S2 to a peak target size of 250 bp. Fragmented DNA was concentrated using AMPure XP beads, and end repair, A-base addition and ligation of Illumina indexed adaptors were performed. Adaptor-ligated libraries were electrophoresed on 3% Nusieve agarose gels, and fragments of 300–350 bp were recovered using QIAEX II gel-extraction reagents. Recovered DNA was amplified using Illumina index primers for eight cycles and purified using AMPure XP beads, and DNA concentration was determined using a Nanodrop spectrophotometer. Libraries (1 μg) were hybridized to the Agilent SureSelect Human All Exon v4 chip at 65°C for 60 h, following the manufacturer's protocol (Agilent Technologies). Targeted exon fragments were captured on Dynal M-280 streptavidin beads (Invitrogen) and enriched by amplification with the Illumina index primers for nine additional PCR cycles. PCR products were purified with AMPure XP beads and analyzed for quality and quantity using an Agilent 2100 Bioanalyzer and DNA 1000 reagents.

We used the publicly available software FastQC to assess sequencing quality. For each lane, we examined per-base quality scores across the length of the reads. Lanes were deemed passing if the per-base quality score box plot indicated that >85% of the reads had >Q20 for bases 1–100. In addition to raw sequence quality, we also assessed alignment quality using the Picard package. This allows monitoring of duplication rates and chimeric reads that may result from ligation artifacts, crucial statistics for interpreting the results of copy number and structural variant analysis.

Gene fusion detection

Paired-end transcriptome sequencing reads were aligned to the human reference genome (GRCh37/hg19) using an RNA sequencing (RNA-seq) spliced read mapper Tophat2 (ref. 35) (Tophat 2.0.4) with the '–fusion-search' option turned on to detect potential gene fusion transcripts. In the initial process, Tophat2 internally deploys an ultrafast short-read alignment tool, Bowtie (Version 0.12.8), to map the transcriptome data. Potential false-positive fusion candidates were filtered out using the 'Tophat-Post-Fusion' module. Further, fusion candidates were manually examined for annotation and ligation artifacts. Junction reads supporting the fusion candidates were realigned using the BLAT alignment tool to confirm fusion breakpoints. Full-length sequence of each fusion gene was constructed on the basis of supporting junction reads and evaluated for potential ORFs using an ORF Finder. For gene fusions with robust ORFs, the amino acid sequences of the fused proteins were explored using the Simple Modular Architecture Research Tool (SMART) to examine the gain or loss of known functional domains in the fusion proteins.

Gene expression

BAM 'accepted_hits.bam' files, which were generated by the Tophat mapping module, were used to quantify the expression data through Cufflinks[36] (Version 2.0.2), an isoform assembly and RNA-seq quantification package. The structural features of 56,369 transcripts from the Ensembl resource (Ensembl 66) were used as an annotation reference to quantify the expression of individual transcripts and isoforms. The 'Max Bundle Length' parameter was set to '10000000', and 'multi-read-correct' was flagged on to perform an initial estimation procedure to more accurately weight reads mapping to multiple locations in the genome.

Mutation analysis

Whole-exome sequencing was performed on an Illumina HiSeq 2000 or HiSeq 2500 instrument in paired-end mode, and primary base call files were converted into FASTQ sequence files using the bcl2fastq converter tool bcl2fastq-1.8.4 in the CASAVA 1.8 pipeline. FASTQ sequence files were then processed through an in-house pipeline constructed for whole-exome sequence analyses of paired cancer and normal genomes. Sequencing reads were aligned to reference genome build hg19 (GRCh37) using Novoalign multithreaded (Version 2.08.02, Novocraft) and converted into BAM files using SAMtools (Version 0.1.18)[37]. Sorting and indexing of BAM files used Novosort threaded (Version 1.00.01), and duplicate reads were removed using Picard (Version 1.74). Mutation analysis was performed using VarScan2 algorithms (Version 2.3.2)[38] 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 (Version 0.2.4) was used on tumor and matched normal samples, and indels common to both 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 as well as of somatic and/or germline mutations was generated by excluding synonymous SNVs. ANNOVAR[39] was used to functionally annotate the detected genetic variants, and positions are based on Ensembl 66 transcript sequences.

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. The set of candidates used for estimation consisted of coding variants that (i) were supported by at least 3 variant fragments in the tumor sample, (ii) were not supported by variant fragments in the matched benign sample, with at least 16 fragments of coverage, (iii) were not present in dbSNP, (iv) were within a targeted exon or within 100 bp of a targeted exon, (v) were not in homopolymer runs of 4 or more bases and (vi) exhibited no evidence of amplification or deletion. To filter out regions of possible amplification or deletion, we used exon coverage ratios to infer copy number changes, as described below. Resulting SNV candidates were not used for the estimation of tumor content if the segmented log ratio exceeded 0.2 in absolute value. Candidates on the Y chromosome were also eliminated because they were unlikely to exist in two-copy genomic regions. Using this set of candidates, we fit a binomial mixture model with two components

using the R package flexmix, version 2.3–8. One component consisted of SNV candidates with very low variant fractions, presumably resulting from recurrent sequencing errors and other artifacts. The other component, consisting of the set of likely true SNVs, was informative of tumor content in the tumor sample. Specifically, under the assumption that most or all of the observed SNV candidates in this component are heterozygous SNVs, we expect the estimated binomial proportion of this component to represent one-half of the proportion of tumor cells in the sample. Thus, the estimated binomial proportion obtained from the mixture model was doubled to obtain an estimate of tumor content.

Copy number aberrations were quantified and reported for each gene as the segmented, normalized, $\log_2$-transformed exon coverage ratio between each tumor sample and its matched normal sample[40]. To account for observed associations between coverage ratios and variation in GC content across the genome, lowess normalization was used to correct per-exon coverage ratios before segmentation analysis. Specifically, mean GC percentage was computed for each targeted region, and a lowess curve was fit to the scatterplot of $\log_2$ coverage ratios versus mean GC content across the targeted exome using the lowess function in R (version 2.13.1) with smoothing parameter f = 0.05.

Partially redundant sequencing of areas of the genome affords the ability for cross-validation of findings. We cross-validated exome-based point mutation calls by manually examining the genomic and transcriptomic reads covering the mutation using the UCSC Genome Browser. Likewise, gene fusion calls from the transcriptome data can be further supported by structural variant detection in the genomic sequence data, as well as by copy number information derived from genome and exome sequencing.

Chemicals and reagents

β-estradiol, (Z)-4-hydroxytamoxifen, (E/Z)-endoxifen hydrochloride hydrate and fulvestrant were purchased from Sigma-Aldrich.

Plasmids and cloning

cDNA for wild-type ESR1 was PCR amplified from a breast cell line MCF7 (ATCC) with the introduction of a sequence encoding an N-terminal Flag tag. cDNAs encoding the relevant mutations of ESR1 were generated by site-directed mutagenesis (QuikChange, Agilent Technologies), and full-length constructs were fully sequenced. All ESR1 variants were placed in the lentiviral vector pCDH (System Biosciences) for eukaryotic expression.

ERE-luciferase reporter assays

For cell transfection experiments, HEK293T cells (ATCC) were plated at a density of 1–2 × 10^5 cells per well (24-well plates) in phenol red–free DMEM containing 10% FBS and antibiotics. Once cells attached, the medium was replaced with DMEM containing 10% charcoal/dextran-treated FBS (HyClone), and cells were cultured overnight. The next day, cells were transiently cotransfected with ESR1 expression plasmid (50 ng/well) and luciferase reporter constructs (25 ng/well; SABiosciences) using FuGene 6 reagent (Promega). The ER-responsive luciferase plasmid encoding the firefly luciferase reporter gene is driven by a minimal CMV promoter and tandem repeats of the estrogen

transcriptional response element (ERE). A second plasmid constitutively expressing Renilla luciferase served as an internal control for normalizing transfection efficiencies (Cignal ERE Reporter, SABiosciences). After transfection for 18 h, cells were serum starved for a few hours before treatment with β-estradiol or anti-estrogen drugs. Cells were harvested 18 h after treatment, and luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega). IC_{50} values were computed using the GraphPad Prism application to fit a four-parameter dose response curve.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Dan Miller, Terrance Barrette, and Doug Gibbs for hardware and database management, Karen Giles for assistance with manuscript preparation, physicians Max Wicha, Lori Pierce, David Smith, Kenneth Levin, Felix Feng for referring patients, and Christine Betts and Jyoti Athanikar for assistance with Tumor Boards. We also thank the larger MI-ONCOSEQ team including clinical research coordinator Erica Williams, pathologist Rohit Mehra, genetic counselors Jessica Everett, Shanna Gustafson, and Victoria Raymond, and radiologists E. Higgins, E. Caoili, and R. Dunnick. This project is supported in part by the Prostate Cancer Foundation for funding of our sequencing infrastructure, the NCI Early Detection Research Network (U01 CA111275), the NIH-NHGRI (1UM1HG006508), Department of Defense W81XWH-12-1-0080 and a Department of Defense Era of Hope Scholar Award. A.M.C. is also supported by the Alfred A. Taubman Institute, the American Cancer Society, the Howard Hughes Medical Institute, and a Doris Duke Charitable Foundation Clinical Scientist Award.

References

1. Chin L, Andersen JN, Futreal PA. Cancer genomics: from discovery science to personalized medicine. Nat Med. 2011; 17:297–303. [PubMed: 21383744]
2. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet. 2010; 11:685–96. [PubMed: 20847746]
3. Roychowdhury S, et al. Personalized oncology through integrative high-throughput sequencing: a pilot study. Sci Transl Med. 2011; 3:111ra121.
4. Welch JS, et al. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. JAMA. 2011; 305:1577–84. [PubMed: 21505136]
5. Gorre ME, et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. Science. 2001; 293:876–80. [PubMed: 11423618]
6. Korpal M, et al. An F876L Mutation in Androgen Receptor Confers Genetic and Phenotypic Resistance to MDV3100 (Enzalutamide). Cancer Discov. 2013
7. Joseph JD, et al. A clinically relevant androgen receptor mutation confers resistance to 2nd generation anti-androgens enzalutamide and ARN-509. Cancer Discov. 2013
8. Ariazi EA, Ariazi JL, Cordera F, Jordan VC. Estrogen receptors as therapeutic targets in breast cancer. Curr Top Med Chem. 2006; 6:181–202. [PubMed: 16515478]
9. Riggins RB, Schrecengost RS, Guerrero MS, Bouton AH. Pathways to tamoxifen resistance. Cancer Lett. 2007; 256:1–24. [PubMed: 17475399]
10. Lonning PE, Eikesdal HP. Aromatase inhibition 2013: clinical state of the art and questions that remain to be solved. Endocr Relat Cancer. 2013; 20:R183–201. [PubMed: 23625614]
11. Ingle JN, et al. Fulvestrant in women with advanced breast cancer after progression on prior aromatase inhibitor therapy: North Central Cancer Treatment Group Trial N0032. J Clin Oncol. 2006; 24:1052–6. [PubMed: 16505423]
12. Osborne CK, Schiff R. Mechanisms of endocrine resistance in breast cancer. Annu Rev Med. 2011; 62:233–47. [PubMed: 20887199]

13. Wu YM, et al. Identification of Targetable FGFR Gene Fusions in Diverse Cancers. Cancer Discov. 2013; 3:636–647. [PubMed: 23558953]
14. Nik-Zainal S, et al. Mutational processes molding the genomes of 21 breast cancers. Cell. 2012; 149:979–93. [PubMed: 22608084]
15. Feil R, Wagner J, Metzger D, Chambon P. Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. Biochem Biophys Res Commun. 1997; 237:752–7. [PubMed: 9299439]
16. TCGA. Comprehensive molecular portraits of human breast tumours. Nature. 2012; 490:61–70. [PubMed: 23000897]
17. Ellis MJ, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature. 2012; 486:353–60. [PubMed: 22722193]
18. Huang HJ, Norris JD, McDonnell DP. Identification of a negative regulatory surface within estrogen receptor alpha provides evidence in support of a role for corepressors in regulating cellular responses to agonists and antagonists. Mol Endocrinol. 2002; 16:1778–92. [PubMed: 12145334]
19. Shiau AK, et al. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. Cell. 1998; 95:927–37. [PubMed: 9875847]
20. Skafar DF. Formation of a powerful capping motif corresponding to start of "helix 12" in agonist-bound estrogen receptor-alpha contributes to increased constitutive activity of the protein. Cell Biochem Biophys. 2000; 33:53–62. [PubMed: 11322512]
21. Carlson KE, Choi I, Gee A, Katzenellenbogen BS, Katzenellenbogen JA. Altered ligand binding properties and enhanced stability of a constitutively active estrogen receptor: evidence that an open pocket conformation is required for ligand interaction. Biochemistry. 1997; 36:14897–905. [PubMed: 9398213]
22. Weis KE, Ekena K, Thomas JA, Lazennec G, Katzenellenbogen BS. Constitutively active human estrogen receptors containing amino acid substitutions for tyrosine 537 in the receptor protein. Mol Endocrinol. 1996; 10:1388–98. [PubMed: 8923465]
23. Pearce ST, Liu H, Jordan VC. Modulation of estrogen receptor alpha function and stability by tamoxifen and a critical amino acid (Asp-538) in helix 12. J Biol Chem. 2003; 278:7630–8. [PubMed: 12496244]
24. Zhao C, et al. Mutation of Leu-536 in human estrogen receptor-alpha alters the coupling between ligand binding, transcription activation, and receptor conformation. J Biol Chem. 2003; 278:27278–86. [PubMed: 12736255]
25. Ellis MJ, et al. Lower-dose vs high-dose oral estradiol therapy of hormone receptor-positive, aromatase inhibitor-resistant advanced breast cancer: a phase 2 randomized study. JAMA. 2009; 302:774–80. [PubMed: 19690310]
26. Swaby RF, Jordan VC. Low-dose estrogen therapy to reverse acquired antihormonal resistance in the treatment of breast cancer. Clin Breast Cancer. 2008; 8:124–33. [PubMed: 18621608]
27. Sokolosky ML, et al. Involvement of Akt-1 and mTOR in sensitivity of breast cancer to targeted therapy. Oncotarget. 2011; 2:538–50. [PubMed: 21730367]
28. Kan Z, et al. Diverse somatic mutation patterns and pathway alterations in human cancers. Nature. 2010; 466:869–73. [PubMed: 20668451]
29. Shrestha Y, et al. PAK1 is a breast cancer oncogene that coordinately activates MAPK and MET signaling. Oncogene. 2012; 31:3397–408. [PubMed: 22105362]
30. Barone I, Brusco L, Fuqua SA. Estrogen receptor mutations and changes in downstream gene expression and signaling. Clin Cancer Res. 2010; 16:2702–8. [PubMed: 20427689]
31. Kandoth C, et al. Integrated genomic characterization of endometrial carcinoma. Nature. 2013; 497:67–73. [PubMed: 23636398]
32. Fisher B, et al. Endometrial cancer in tamoxifen-treated breast cancer patients: findings from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14. J Natl Cancer Inst. 1994; 86:527–37. [PubMed: 8133536]
33. Dawson SJ, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med. 2013; 368:1199–209. [PubMed: 23484797]

34. Diehl F, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med. 2008; 14:985–90. [PubMed: 18670422]

35. Kim D, Salzberg SL. TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. Genome Biol. 2011; 12:R72. [PubMed: 21835007]

36. Trapnell C, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012; 7:562–78. [PubMed: 22383036]

37. Li H, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–9. [PubMed: 19505943]

38. Koboldt DC, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 2012; 22:568–76. [PubMed: 22300766]

39. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38:e164. [PubMed: 20601685]

40. Lonigro RJ, et al. Detection of somatic copy number alterations in cancer using targeted exome capture sequencing. Neoplasia. 2011; 13:1019–25. [PubMed: 22131877]

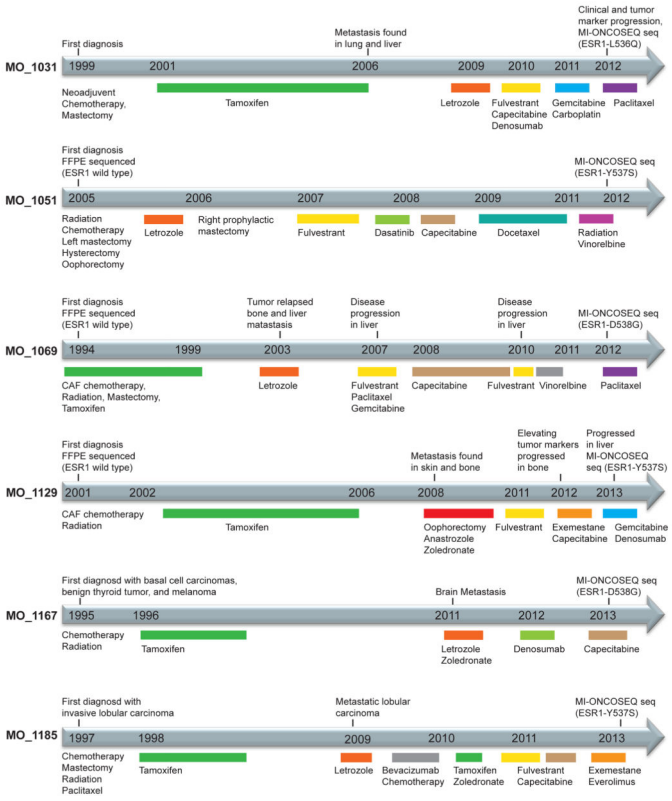

Figure 1.

Clinical timelines of the six index ER-positive metastatic breast cancer patients harboring ESR1 mutations. Shown are patients' histories of clinical treatments from first diagnosis until the enrollment on the MI-ONCOSEQ study. Each bar represents the timeframe of a treatment.

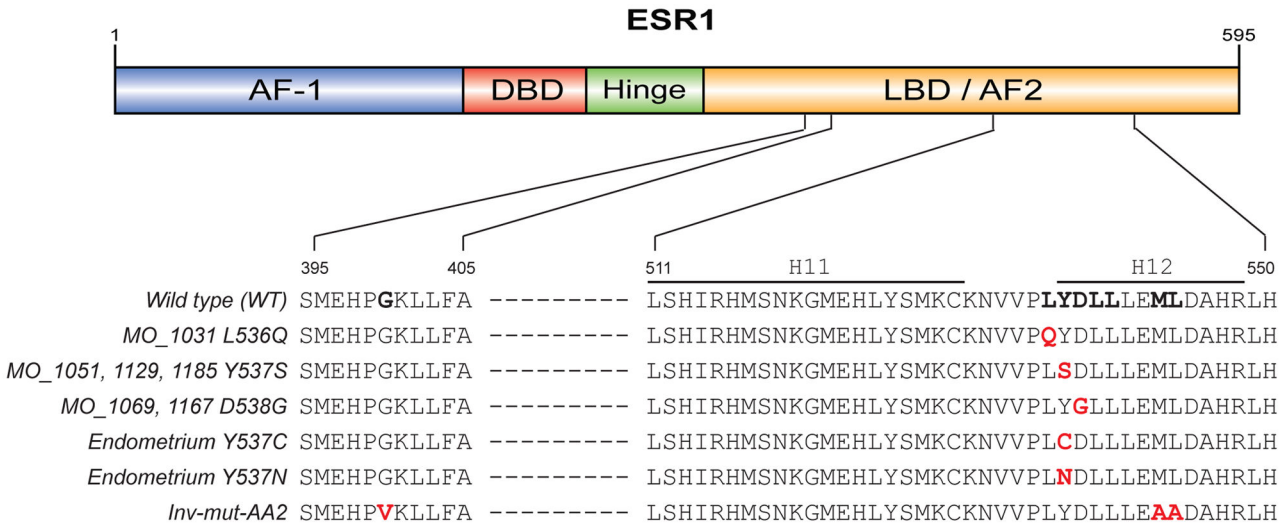

Figure 2.

Schematic representation of ESR1 mutations identified in this study. The structural domains of ESR1 are illustrated on top, including the transcription activation function-1 domain (AF-1), the DNA-binding domain (DBD), the hinge domain, and the ligand-binding domain (LBD/AF-2). Changed residues identified in mutants are marked in red, and the reference residues are bolded in the wild type sequence. Endometrium p.Tyr537Cys (Y537C) and p.Tyr537Asn (Y537N) are two mutations discovered in endometrial cancer from the TCGA study. Inv-mut-AA2 represents a ligand activity inversion mutant of ESR1 which renders the receptor with inverted response to anti-estrogen and estrogen. H11, helix 11; H12, helix 12.

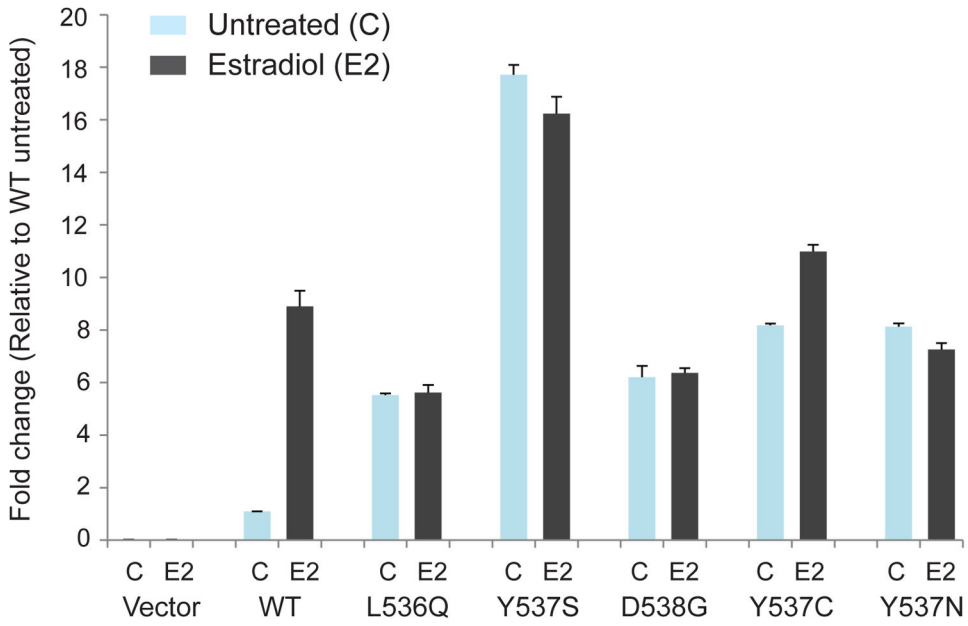

Figure 3.

Acquired ESR1 mutations are constitutively active. HEK-293T cells were co-transfected with an ERE-firefly luciferase reporter plasmid, a plasmid constitutively expressing Renilla luciferase as an internal control, and various ESR1 constructs as illustrated in Fig 2. Steroid hormone-deprived cells were either untreated (C) or stimulated with 5 nM of β-estradiol (E2) for 24 hrs. Firefly luciferase levels were normalized using corresponding Renilla luciferase levels for each condition. Fold change of ESR1 transcription activity was calculated using untreated wild type as control for each condition. Data shown are mean of triplicate. Amino acid mutations in respective ESR mutants are indicated. WT, wild-type ESR1.

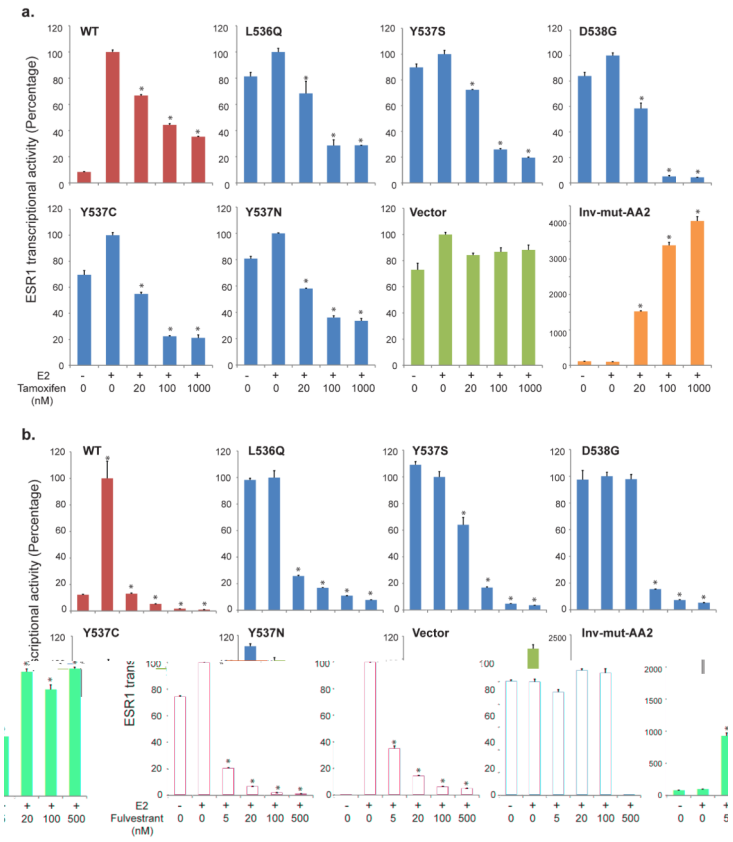

Figure 4.

Acquired ESR1 mutations maintain sensitivity to antiestrogen therapies. As described in Fig 3, HEK-293T cells were co-transfected with an ERE-firefly luciferase reporter plasmid, a plasmid constitutively expressing Renilla luciferase, and various ESR1 constructs as indicated. Steroid hormone-deprived cells were either untreated or treated with increasing doses of antiestrogen drugs tamoxifen (A) or fulvestrant (B) in the presence of 5 nM of β-estradiol (E2) for 24 hrs. Percentage change of ESR1 transcription activity was calculated using E2-treated cells as control for each tested construct. Data shown are mean of triplicate. Error bars indicate s.d. *, P values <0.001.

Table 1

Clinical sequencing of eleven metastatic ER-positive breast cancer cases.

Case	Age	ER/PR/ERBB2	Treatments[a]	#SNV/#Fusion	Genetic aberrations[b]
MO_1031	41	+ / + / −	Tamoxifen, Letrozole, Fulvestrant	266 / 18	*ESR1* (p.Leu536Gln), gene copy gains of *FGFR1*, *FGFR2*, *CCND1*, and *GNRHR*
MO_1051	31	+ / − / −	Oophorectomy, Letrozole, Fulvestrant	248 / 5	*ESR1* (p.Tyr537Ser), *PIK3CA* (p.His1047Arg), *TP53* (p.Gly199Glu), *FGFR2-AFF3* fusion
MO_1069	62	+ / + / −	Tamoxifen, Letrozole, Fulvestrant	74 / 9	*ESR1* (D538G), *ARID2* (p.Glu245*), gene copy losses of *TP53*, *BRCA1*, *RB1*, *ARID1A*, and *SMARCA4*
MO_1129	44	+ / + / −	Tamoxifen, oophorectomy, Anastrozole, Fulvestrant, Exemestane	32 / 3	*ESR1* (p.Tyr537Ser), *PIK3CA* (p.Glu542Lys), gene copy gains of *CCND1* and *PAK1*
MO_1030	78	+ / + / −	Tamoxifen (short), Anastrozole, Fulvestrant	26 / 2	*PIK3CA* (p.Glu545Ala), *TP53* copy loss
MO_1068	65	+ / − / −	Tamoxifen, Anastrozole	83 / 10	*PIK3CA* (p.His1047Arg), *TP53* (p.Glu51*), *MSH2* copy loss
MO_1090	52	+ / + / −	Tamoxifen, Anastrozole	28 / 11	No significant drivers identified
MO_1107	46	+ / + / −	Tamoxifen, oophorectomy, Anastrozole, Fulvestrant, Exemestane	63 / 12	*BRCA1* (c.5385_5386insC), frameshift deletions in *TP53*, *SMARCA4*, and *NF1*
MO_1167	60	+ / − / −	Tamoxifen, Letrozole	47 / 3	*ESR1* (p.Asp538Gly)
MO_1185	58	+ / + / −	Tamoxifen, Letrozole, Fulvestrant, Exemestane	88 / 1	*ESR1* (p.Tyr537Ser), *CDH1* (p.Gln641*), *NOTCH2* (frameshift deletion), *TP53* copy loss
TP_2004[c]	52	+ / − / −	Tamoxifen (short)	29 / 22	*MDM2* gene amplification, gene copy losses of *CDKN2A* and *CDKN2B*

Notes:

[a] Only anti-estrogen related treatments are listed in table. Patients also received chemotherapies, radiation, or mastectomy in the interim between diagnosis and MI-ONCOSEQ sequencing.

[b] Amino acid substitutions caused by nonsynonymous somatic mutations are marked in parentheses.

[c] TP_2004 is a male patient.