Early enrichment of *ESR1* mutations and the impact on gene expression in primary breast cancer treated with aromatase inhibitors in the pre-surgical setting

Mariana Ferreira Leal[1,2], Ben P Haynes[1], Eugene Schuster[2], Belinda Yeo[1], Maria Afentakis[1], Lila Zabaglo[1,2], Vera Martins[2], Richard Buus[2], Andrew Dodson[1], Maggie CU Cheang[3], Ian E Smith[1], Lesley-Ann Martin[2], Mitch Dowsett[1,2].

[1]Ralph Lauren Centre for Breast Cancer Research, Royal Marsden Hospital NHS Trust, London, United Kingdom, SW3 6JJ.

[2]The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London, United Kingdom, SW3 6JB.

[3]Clinical Trials and Statistic Unit, The Institute of Cancer Research, Sutton, United Kingdom, SM2 5NG

[4]Breast Unit, The Royal Marsden NHS Foundation Trust, London, United Kingdom, SW3 6JJ.

Running title: Early aromatase inhibitor therapy enriches *ESR1* mutations

Key-words: aromatase inhibitor**/** breast cancer / *ESR1* mutation / gene expression / neoadjuvant therapy.

Corresponding author: Mariana Ferreira Leal. Ralph Lauren Centre for Breast Cancer Research, 4th Floor Wallace Wing, Royal Marsden Hospital NHS Trust, 203 Fulham Road, London SW3 6JJ, United Kingdom. Phone: 0207 808 2619; FAX: 0207 808 2808; E-mail: mleal@icr.ac.uk

Financial support

This study was supported by fellowship from Le Cure to MFL and by a grant from The Breast Cancer Research Foundation. We thank Breast Cancer Now for funding this work as part of Programme Funding to the Breast Cancer Now Toby Robins Research Centre. We acknowledge support from the National Institute for Health Research through the National Cancer Research Network and the Royal Marsden / The Institute of Cancer Research Biomedical Research Centre.

Conflict of interest

MD receives honorary from speaker's bureau of Roche and Myriad Genetics, is a consultant and advisory board member of Radius, receives academic funding from Pfizer and has received remuneration from the ICR rewards to Inventors Scheme.

LAM receives academic funding from Radius, PUMA, Pfizer and AstraZeneca and receives honorary from the speakers Bureau of Pfizer

MCUC has a patent for Breast Cancer Classifier (US Patent No. 9,631,239) with royalties paid.

Word count: 4666

Figures: 5

Supplementary tables: 12

Supplementary figures: 7

Statement of translational relevance

Despite the efficacy of aromatase inhibitors (AI) for the treatment of post-menopausal woman with oestrogen-receptor positive (ER+) breast cancer (BC), over 20% of patients with early-stage disease will relapse. Few predictive biomarkers have been identified in treatment-naïve tumours most likely due to the requirement for early exposure to treatment in order to reveal rewiring events that in the long-term will drive resistance to therapy. Herein, we are the first to show the enrichment of *ESR1* mutation, a known mechanism of endocrine therapy resistance in metastatic ER+BC, in primary tumours after AI treatment. We also identified *ESR1* wild-type tumours with high residual proliferation and ligand-independent ER activity. Our data demonstrates that pre-surgical AI exposure enhances the ability to identify tumours dependent on classical ER signalling and reveals mechanisms of resistance, that can be targeted therapeutically in the adjuvant (post-surgical) setting with pertinent combination therapies.

ABSTRACT

Purpose: To investigate the presence of *ESR1* mutation in primary oestrogen-receptor positive breast cancer (ER+BC) treated with extended (>4 weeks) neoadjuvant (pre-surgical) aromatase inhibitor (NAI) therapy and to identify patients who may gain less benefit from aromatase inhibition (AI) alone based upon on-treatment changes in gene expression. **Experimental Design:** We evaluated ER, progesterone receptor and Ki67 by immunostaining, *ESR1* mutations by droplet-digital-PCR and expression of over 800 key BC genes in paired pre- and post-NAI tumour samples from 87 ER+BC patients. **Results:** Cell proliferation and oestrogen-regulated genes (*ERGs*) remained suppressed in most tumours indicative of persistent response to NAI. Enrichment of *ESR1* mutations was found in five tumours and predominantly in patients receiving therapy for >6 months. *ESR1* mutant tumours showed increased expression of *ESR1*-transcript and limited suppression of *ERGs* and proliferation associated genes in response to NAI. *ESR1* wild-type tumours with high residual proliferation (Ki67r≥10%; 15/87 tumours) showed lower *ESR1*/ER expression pre- and post-therapy and lower *ERGs*. Tumours with *ESR1* mutations or Ki67r≥10% showed less inhibition of oestrogen-response, cell-cycle and E2F-target genes. **Conclusion:** Ligand-independent ER-signalling, as a result of *ESR1* mutation or reduced ER-dependence, identified after extended NAI therapy, can guide early selection of patients who would benefit from combination therapy.

INTRODUCTION

Over 80% of patients diagnosed with breast cancer (BC) present with tumours which are oestrogen-receptor positive (ER+) and proliferate in response to the female hormone oestrogen (E) (1). Aromatase inhibitors (AIs) block the conversion of androgens to oestrogens and are first-line treatment for postmenopausal women with ER+BC. Despite their efficacy, over 20% of patients with early-stage disease will eventually relapse and those with metastatic disease will inevitably recur despite initial response to AI-therapy (2).

Currently, few mechanisms of resistance to AI therapy have been identified with most being attributed to crosstalk between ER and growth factor signalling pathways allowing tumour cells to circumvent the need for steroid hormone (3). Furthermore, although studies have shown that AIs lead to a marked remodelling of the clonal mutational landscape (4-6), few mutations have been shown to be enriched in the metastatic setting with the exception of *ESR1*. Mutations in *ESR1* have been observed in 30-40% of patients who progress on AI therapy, but only sporadically in patients who have not received AI for metastatic disease (6-11). The identification of new biomarkers and therapeutic strategies that can target early resistance, is therefore, of paramount importance.

Neoadjuvant (pre-surgical) AI (NAI) therapy, which is used to downstage primary tumours to enhance the likelihood of breast conserving surgery being a treatment option, provides an ideal opportunity to observe biological changes as a result of AI treatment. This can yield both prognostic and predictive information and facilitate the design of novel clinical trials targeting endocrine resistant disease (12). Most of the clinical trials exploiting this concept have been restricted to short-term pre-surgical exposure to AI, such as the PeriOperative Endocrine Therapy for Individualising Care (POETIC, CRUK/07/015) and Alliance for Clinical trial in Oncology (ACOSOG) Z1031B studies, where patients were treated for 2-4 weeks before surgery (13-15). Whilst informative, these studies do not address the long-term effect of NAI therapy that may be necessary to evaluate the full

impact of AI-induced phenotypic/genotypic alterations (16) or the effects that might limit response and lead to clinical resistance.

Here, we report a detailed study of the molecular alterations associated with extended (>4 weeks) NAI treatment in the index primary ER+ BC and show that *ESR1* mutations are enriched with longer duration of therapy and become a key mitogenic driver. Using the validated proliferation marker Ki67 to identify endocrine resistant tumours (17), we show that *ESR1* wild-type (*ESR1*^{Wt}) tumours with high residual proliferation after NAI therapy appear partially independent of "classical ER signalling" highlighting the high degree of heterogeneity in adaptive mechanisms circumventing E-deprivation.

MATERIALS AND METHODS

Patients

We retrospectively identified 109 post-menopausal women diagnosed with primary ER+ BC and treated with NAI for at least one month at the Royal Marsden Hospital (RMH) between 2003 and 2016. Inclusion criteria included the requirement for generic consent to conduct tissue-based research and the availability of both the diagnostic core-biopsy and paired surgical excision post-NAI. Exclusion criteria were: 1) multifocal disease; 2) previous BC in a period of 10 years; 3) involvement in a neoadjuvant clinical trial; 4) concomitant anticancer treatments including chemotherapy, biologic response modifiers, endocrine therapy (including steroids) and radiotherapy. Paired biopsies with >40% invasive cell areas were available from 87 patients (Supplementary Fig. S1).

Clinical and histological details are shown in Supplementary Table S1. Given the focus on the molecular characteristics in the excision, clinical response by RECIST (17) was characterized based on ultrasound changes between start of AI and surgery.

All the research was carried out in accordance with the provisions of the declaration of Helsinki of 1975. Ethical approval for the study was received from an NHS research ethics

committee (reference 17/EM/0145) and patients had to have given consent for their tissues to be used for ethically approved research.

Protein expression analysis by Immunohistochemistry

Immunohistochemistry (IHC) for ER, PgR and Ki67 was performed as previously described (18,19). H-Score was used for the assessment of ER and percentage-positivity for PgR. Ki67 percentage-positivity was scored as a continuous variable, according to the method described by the International Ki67 Working Group (20). Ki67 proliferation was used as an End-of-Neoadjuvant-Treatment Endpoint to identify index tumours that are endocrine resistant (17).

RNA and DNA extraction

Formalin-fixed paraffin-embedded (FFPE) tissue sections were microdissected before co-extraction of RNA and DNA using the AllPrep DNA/RNA FFPE Kit (Qiagen, Hilden, Germany), according to the manufacturer instructions with the exception of an extended overnight digestion for the DNA extraction. Nucleic acid quantification was done using high sensitivity RNA and DNA Qubit assays (Thermo Fisher Scientific, Carlsbad, CA, USA).

Gene expression analysis

Gene expression was evaluated using nCounter® FLEX Analysis System (NanoString™ Technologies, USA) with two panels (744 and 106 genes, including 30 in common; Supplementary Table S2). The panels included reference genes, PAM50 gene set (panel #1) and genes involved in the most important aspects of BC or with evidence of an association with AI resistance, including *ERGs*, proliferation, invasion, growth factor

receptors, PI3K-AKT-mTOR pathway, MAPK signalling, cholesterol metabolism, inflammation, and epithelial mesenchymal transition (EMT) genes. For three patients, gene expression analysis was conducted using only the smallest panel due to the low availability of RNA.

Raw counts were normalized by NanoStringNorm package in R (21). Briefly, the geometric mean of the counts from the six External RNA Controls Consortium (ERCC) positive controls to take into account the efficiency of the hybridization. Background correction was done by subtracting the geometric mean of the nine ERCC negative control probes. Data was scaled and normalized by nine reference genes (Supplementary Table S2), that were confirmed as representative of the lowest combined variation across the studied samples. Expression values were log2 transformed for statistical analysis.

Tumors were classified into one of the intrinsic subtypes (Luminal A, Luminal B, Basal-like and HER2-Enriched) based on the PAM50 classifier algorithm (22,23). *ERGs* expression was defined as the mean of *TFF1*, *GREB1*, *PDZK1* and *PGR* (24) and *PAGs* as the mean of 11 proliferation genes in the PAM50 gene set (*BIRC5*, *CCNB1*, *CDC20*, *CDCA1*, *CEP55*, *KNTC2*, *MKI67*, *PTTG1*, *RRM2*, *TYMS* and *UBE2C*). An E2F activation metagene was developed based on the 24-gene E2F signature devoid of cell cycle-associated genes described by Miller et al (25) (Supplementary Methods).

ESR1 mutation analysis

Seven hot-spot *ESR1* mutations within the ligand-binding domain were evaluated by droplet digital polymerase chain reaction (ddPCR). Initially, we screened all post-AI samples using two multiplexed reactions for the following mutations: 1) E380Q, L536R, Y537C and D538G; 2) S463P, Y537N and Y537S. ddPCR was performed with 5ng of DNA on an Automated droplet generator and QX100^{TM} system (Bio-Rad, Hercules, CA, USA). The results were validated using singleplex ddPCR. *ESR1* mutations were also assessed in the

pre-AI samples from those patients with a mutation in their residual tumour. *ESR1* mutation was considered positive with at least two mutant droplets detected. Mutation allele fraction was calculated as previously described (26).

Tumours with VAF <1% were validated by ddPCR after FACS to enrich the number of cytokeratin-positive neoplastic cells evaluated (Supplementary Methods). This approach was also used to confirm the lack of detectable mutations in pre-NAI samples.

ESR1 copy number

ESR1 copy number was evaluated by Fluorescence in situ Hybridization (FISH) in residual tumours harbouring *ESR1* mutations and in one pre-treatment sample. Dual colour FISH probes hybridizing at 6q25 (*ESR1*) and chromosome-6 (CEN6) were applied (ZytoLight, Germany). Briefly, four-micron FFPE sections were deparaffinised and incubated for 20 minutes in Target Retrieval Solution Citrate pH 6.1 (Agilent, US) at 98°C, followed by pepsin digestion for 15 min at 37°C and RNase A treatment for 30 minutes at 37°C. Co-denaturation was performed for 10 minutes at 75°C followed by hybridisation for 24 hours at 37°C. Sections were mounted in DAPI-containing Vectashield (VectorLabs, UK) and analysed using fluorescence microscopy (Leica Biosystems, Germany).

FISH scoring was performed by counting 40 representative non-overlapping nuclei. Average copy number ratio *ESR1*/CEN6 was assessed. A ratio ≥2 was rated amplification and ≥1.3 as copy number gain (27).

Data analysis

All analyses were performed using R v3.4.4. P value <0.05 was considered statistically significant. For gene expression analysis involving multiple comparisons, false discovery rate (FDR) was applied as indicated. GSEA was run using the GSEA v.3.0 software (http://software.broadinstitute.org/gsea) with 1,000 permutations.

RESULTS

Clinical and pathological profile of patients treated with extended NAI therapy

Paired pre- and post-NAI therapy tumours were available from 87 ER+ BC patients (Supplementary Fig. S1, Supplementary Table S1), in which key biomarkers ER, PgR, HER2 and Ki67 were assessed by IHC, together with the expression of 820 genes (Supplementary Table S2) using NanoString^{TM} technology (Figure 1A).

Mean age was 72.1 years (ranging between 50–93); 58 (66.7%) of patients had grade 2 and 23% had grade 3 ER+ tumours; 63 (72.4%) were ductal subtype (Supplementary Table S1). PgR expression was detected in 69 patients (79.3%) and five (5.7%) were classified as HER2+ by IHC and FISH (Figure 1B, Supplementary Table S1). Based on intrinsic subtypes (n=84), four (4.6%) tumours were HER2-enriched and two (2.3%) were basal-like (Figure 1B, Supplementary Table S1).

The mean ± standard deviation duration of treatment was 26.3±16.2 weeks (Figure 1A–B). Three patients (3.4%) received second- or third-line endocrine therapy after AI. One (1.1%) patient showed a complete response (CR) as measured by ultrasound, 55 (63.2%) partial response (PR), 13 (14.9%) stable disease (SD) and five (5.7%) progressive disease (PD, Figure 1B). Among the PRs, six (12.7%) showed >20% increased tumour volume before surgery compared with previous ultrasound, with five being treated for less than 32 weeks and one for 70 weeks. Among those with SD, six (46.2%) showed an initial objective response to AI therapy, which was subsequently followed by an increase of tumour volume.

Clinical response was not associated with clinical, pathological or protein biomarkers tested at diagnosis (pre-NAI) or surgery (post-NAI) ($p>0.05$, Chi-Squared test or T-test) with the exception pre-NAI PgR levels ($p=0.007$; T test) and expression of E-regulated genes (*ERGs*; $p=0.019$) (24) that were lower in SD/PD in comparison with CR/PR (Supplementary Fig. S2). In post-NAI tumours, proliferation associated genes (*PAGs*) were higher in SD/PD

compared with CR/PR (p=0.013, T-test; Supplementary Fig. S2). As expected, Ki67 abundance correlated with its transcript level and also with *PAGs* (r=0.59–0.77; p<0.001; Pearson correlation; Supplementary Fig. S3A-B). Based on this observation and the wealth of the data supporting the use of residual Ki67 (Ki67r) as a biomarker of benefit from AI in the adjuvant setting (13,15), Ki67r by IHC was used as a measure of response in this study.

Overall changes with AI treatment

Assessment of the on-treatment change in Ki67 (mean reduction: 21.7%) and *PAGs* (3.8-fold) showed that overall the majority of tumours responded to NAI at surgery (p<0.001; paired T test; Figure 2A; Supplementary Fig. S3A-B), with 55 (63.2%) tumours showing complete cell cycle arrest (28) based on Ki67r (Low-Ki67r; ≤2.7%) and 15 (17.2%) high residual proliferation (High-Ki67r; Ki67r≥10%).

In further confirmation of the response to NAI, PgR/*PGR* and ER/*ESR1* were significantly supressed on therapy (protein mean reduction: 41% and 21%, gene expression reduction: 4.4 and 1.8-fold, respectively; p<0.001; paired T test; Figure 2A; Supplementary Fig. S3A), with two (2.3%) tumours classified as ER- / PgR- and 64 (73.6%) ER+ / PgR- post-NAI. In keeping with this observation, *ERGs* post-NAI were similarly supressed (6.8-fold; p<0.001; Figure 2A; Supplementary Fig. S3C). There was no clear difference between HER2- and the small number of HER2+ (5.7%) tumours with regard to these biomarkers (Figure 1B, 2A, 3A; Supplementary Fig. S5A; Supplementary Fig. S3).

In addition, analysis of intrinsic subtypes showed that most of tumours were phenotypically luminal-A-like post-NAI (Figure 1A, 2B) and gene set enrichment analysis (GSEA) confirmed the inhibition of genes involved in E response and proliferation, including E2F targets (Figure 2C). Most of tumours also showed a significant reduction of an E2F activation signature (25) (p<0.001, paired T-test; Supplementary Fig. S4), which was also associated with SD/PD in post-NAI tumours (p=0.022; T-test; Supplementary Fig. S4).

Comparison of gene expression between pre- and post-NAI revealed that 554 genes were differentially expressed (FDR 5%; Supplementary Table S3). Hierarchical clustering based on the changes in expression of these genes with >25% change (410 genes) separated tumours into four main branches labelled A-D (Figure 2B). Branch A and B showed less inhibition of cell cycle genes and of genes involved in E response and contained 5 of 9 patients with recurrence. Branch B also showed less inhibition of genes involved in immune response, focal adhesion, MAPK and cytokine signalling and in this aspect was distinct from the other branches. Notably, this branch was enriched of tumours with poor response based on clinical response and with Ki67r≥10% and was also enriched for tumours with post-NAI *ESR1* mutations described in more detail below. Branch C was mainly characterized by down-regulation of E-related and proliferation associated genes, but also upregulation of genes involved in immune response (expanded immune pathways are shown in Supplementary Table S4) and making this branch was distinct from D. Overall branch C tumours changed more with treatment based on the intra-patient correlation score. Branch C contained three of the four patients with late distant recurrence (≥5 years). Both branches C and D showed greater upregulation of genes associated with focal adhesion, MAPK and cytokine signalling compared to branches A and B.

Early acquisition of *ESR1* mutation is associated with proliferation in tumours treated with NAI

Twelve (13.8%) tumours showed increased expression of *ERGs* after treatment (Figure 3A), which led us to investigate the presence of *ESR1* mutation in post-NAI tumour samples. Seven *ESR1* hot-spots mutations were identified in six tumours (Figure 1B, Figure 3A, Supplementary Table S5): five with D538G mutation (variant allele frequency; VAF: 0.2–27.6%), one with Y537N/D538G (VAF_{Y537N}: 12.3%; VAF_{D538G}: 27.6%) and one with Y537S (VAF: 17.3%). Those tumours with VAF <1% were validated by ddPCR after enrichment of

neoplastic cells (cytokeratin positive) using fluorescent activated cell sorting (FACS) (Supplementary Methods; Figure 3B).

Noteworthy, one case harboured a D538G *ESR1* mutation pre-NAI (VAF: 2%), which was further enriched in the post-NAI sample (VAF: 19.3%; Supplementary Table S5). To further determine if *ESR1* mutations could be detected in the pre-NAI biopsies, we used FACS to enrich tumour cells in the other five pre-NAI tumours from patients harbouring post-NAI mutation and were able to extract enough DNA to test for the presence of *ESR1* mutations by ddPCR in four specimens. No *ESR1* mutation was detected by doing this.

None of the patients with *ESR1* mutant (*ESR1*^{Mut}) tumours were among the ones treated with a second- or third-line therapy pre-surgery. Only one (1.1%) patient showed SD and local recurrence, which had a VAF in the recurrence similar to that detected in the primary tumour ($VAF_{post\text{-}NAI}$: 24.3%; $VAF_{recurrence}$: 22.1%; Supplementary Table S5). The remaining five patients achieved PR prior to surgery.

In pre-NAI samples, *ERGs* (Supplementary Fig. S5A), Ki67, *PAGs* and *ESR1* (Supplementary Table S6) expression did not differ between *ESR1*^{Mut} and *ESR1*^{Wt} tumours ($p>0.05$). However, *ESR1*^{Mut} tumours showed less suppression of *ERGs* ($p=0.002$, Mann-Whitney test) and *PAGs* ($p=0.039$) and greater *ESR1* ($p=0.016$; Supplementary Table S7-8) expression post-NAI compared with *ESR1*^{Wt} tumours (Figure 3A, 3C).

We further accessed the number of *ESR1* copies by FISH in the residual *ESR1*^{Mut} tumours (Supplementary Table S5, Figure 3D) and found one case (VAF: 0.23%) presenting copy number gain (>1.3 *ESR1*/*CEP6* ratio). Despite the copy number gain, this patient showed a reduction in the expression of both *ERGs* and proliferation after NAI probably reflecting the large majority of *ESR1* being wild-type; however, increased on-treatment *ESR1* expression was detected (fold-change: 1.96). Additionally, copy number analysis of the pre-NAI sample from this patient confirmed that the gain of *ESR1* copies was acquired with treatment; however, three copies of chromosome 6 were evident in both pre- and post-NAI

samples. Based on the frequencies, our data suggests that copy number gain preceded the mutation.

GSEA showed a lack of inhibition of E-response and less inhibition of proliferation-related pathways (including *E2F* targets) in *ESR1*^{Mut} tumours in comparison to *ESR1*^{Wt} (Figure 3E). The *E2F* metagene was similarly less supressed in *ESR1*^{Mut} (p=0.016, Mann-Whitney test; Figure 3F). Furthermore, dependence on *ESR1* as a mitogenic driver was confirmed by the observation that several genes linked with ER signalling, including *CCND1*, *RET* and *FOXM1* (p=0.023–0.047, Mann-Whitney test; Supplementary Fig. S5B), showed smaller change in response to NAI (Supplementary Table S7-8).

Of particular note post-NAI *ESR1*^{Mut} tumours were treated for longer with NAI in comparison with *ESR1*^{Wt} (p=0.011; Mann-Whitney test; Figure 3G). Furthermore, all five acquired mutations occurred in the third of patients with the longest duration of NAI 3rd tertile: > 191 days; >6 months), with a prevalence of 5/29 (17%) in this tertile. Taken together this data supports that *ESR1* mutations are enriched with extended NAI treatment in primary ER+ BC.

ESR1^{Wt} tumours with reduced dependence on classical E-signalling, gain less benefit from AI therapy

In order to identify putative resistance mechanisms independent of *ESR1* mutation, we analysed the molecular changes associated with high Ki67r in tumours harbouring *ESR1*^{Wt} (Supplementary Table S9-11). Overall, the expression profile between pre-NAI and post-NAI samples from tumours with High-Ki67r changed less than those from tumours with Low-Ki67r (p=0.023, T-test; Figure 4A and Figure 2B). Moreover, tumours with High-Ki67r tended to maintain their baseline intrinsic subtype (Figure 2B). Both these results are consistent with the molecular phenotype of these responsive tumours being refractory to the NAI therapy.

As expected, higher baseline expression of *ERGs* was correlated with reduced proliferation after treatment ($p<0.001$, $r=-0.38$, Supplementary Fig. S6) highlighting their dependence on ER signalling as the main mitogenic driver. Conversely, high Ki67r was associated with less inhibition of *ERGs* ($p=0.012$; T-test; Figure 4B; Supplementary Fig. S6), which was paralleled by less reduction in PgR abundance ($p=0.023$; Figure 4B). Furthermore, genes involved either directly or indirectly in cell cycle control were less inhibited in tumours with High-Ki67r in comparison with Low-Ki67r (Supplementary Fig. S7), including genes regulated by E, such as *CCND1* and *RET* ($p=0.01$ and $p=0.011$, respectively; Figure 4C; Supplementary Table S11). In addition, an ER dependent *E2F* activation signature (25) was less inhibited in High-Ki67r tumours ($p=0.002$–0.031; T-test, Figure 4D). Moreover, tumours with High-Ki67r did not show significant inhibition of pathways involved in E-early and -late response (Figure 4E).

Further interrogation of the data showed that tumours with High-Ki67r had lower *ESR1*/ER expression/abundance at diagnosis (pre-NAI) compared with those with Low-Ki67r (*ESR1*: $p=0.044$; ER: $p=0.013$ T-test; Figure 4F; Supplementary Table S9). This observation was paralleled by the lower ER abundance in High-Ki67r tumours compared with Low-Ki67r and Medium-Ki67r (2.7%>Ki67r<10%) tumours post-NAI ($p=0.021$, $p=0.025$, respectively; Figure 4F). These findings endorse the hypothesis that a subset of ER+ tumours are less dependent on classical ER-signalling at diagnosis.

Although there was a high degree of similarity in the gene expression profiles at diagnosis irrespective of Ki67r (FDR >10% for all genes, Supplementary Table S9), two key genes involved in the regulation of cell proliferation and inflammation (Supplementary Table S12), *CDK2* ($p_{pre\text{-}NAI}=0.028$, 1.3-fold; $p_{post\text{-}NAI}=0.001$, 1.4-fold, T-test in relation with Low-Ki67r) and *FGFR4* ($p_{pre\text{-}NAI}=0.007$, 4.34-fold, $p_{post\text{-}NAI}=0.013$, 3.93-fold), showed higher expression in those tumours with High-Ki67r at both time-points investigated (Figure 5A). In addition, both *CDK2* and *FGFR4* showed higher expression in tumours of patients with

SD/PD in comparison with CR/PR in both pre-NAI (p=0.017; p=0.012, respectively; T-test) and post-NAI (p=0.017, p=0.007, respectively; Figure 5B).

DISCUSSION

In this study, we focused on understanding mechanisms of resistance that emerge in primary ER+ BC treated with extended NAI therapy and the importance to evaluate paired pre- and post-treatment biopsy. Our study is the first to show the early enrichment of *ESR1* mutation in neoadjuvant setting. Here, we provide further insights into early mechanisms of endocrine resistance, which may inform on combination treatment either before or after surgery or in recurrent disease.

Whilst overall, we observed that cell proliferation was suppressed in most tumours, an observation paralleled by the reduced expression of E-responsive genes. At surgery, tumours tended to be more phenotypically luminal-A-like, but the degree to which this occurred varied across the patient cohort emphasising the high degree of heterogeneity.

ESR1 mutations were identified in 7% of tumours within the study and were enriched among patients treated with NAI for more than six months. This is consistent with previous studies, which have shown that *ESR1* mutations appear enriched almost exclusively in response to AI therapy (8-11,29,30). However, in this study, the reported frequency was lower than that noted in metastatic ER+ BC patients who had relapsed on AI therapy (11-54%) (8-11,29,30). Thus, our results indicate that the selection of *ESR1* mutations can occur frequently in primary as well as metastatic disease.

Nonetheless, *ESR1* mutations were detected at a higher frequency in our cohort compared with a previous study of NAI therapy, which reported 1.5% (16). One explanation for this discrepancy is the difference in technologies used to call the mutation status. In the latter study, mutations were identified by exome sequencing, whilst we used targeted ddPCR and microdissected tumours, allowing identification of VAFs as low as 0.2%, which we also

supported using FACS. We are the first to describe a temporal association of duration of AI as first-line treatment in a neoadjuvant setting with enrichment of *ESR1* mutation. Our findings support our previous studies in ER+BC cell lines that demonstrated enrichment of *ESR1* mutations with time post E-deprivation (31). Further study of the emergence of *ESR1* mutations during NAI therapy may improve our understanding of the tissue dynamics that underpin clinical relevance of treatment-dependent clonal selection during extended E-deprivation.

Our findings provide evidence that *ESR1* mutations are the mitogenic driver of AI resistance. Thus, tumours harbouring a *ESR1* mutation in their residual disease showed activation of genes involved in E-response and of pathways associated with proliferation, highlighted by smaller change of the E-regulated *CCND1*, the *ESR1* co-activator *FOXM1* (32) together with downstream *E2F* targets. Moreover, *ESR1*^{Mut} tumours showed increased expression of the oncogene *RET,* which has previously been associated with ligand-independent ER activity (33). Our findings provide functional evidence for the gain of *ESR1* mutations being a *bona fide* resistance mechanism to AI. These data provide further support for recent *in vitro* characterisations of *ESR1* mutations, which show that these mutations govern an altered cistrome leading to the engagement of E2-independent—ER-driven transcriptional programmes (31,34). They also support for the concept that selective ER down-regulators (SERD) or combination of AI with CDK4/6 inhibitors may provide greater benefit than AI alone in the adjuvant setting for patients with primary *ESR1*^{Mut} ER+ BC.

The present study was focused on molecular changes that underpin response in the index tumour and to gain a better understanding of the de novo and acquired resistance mechanisms as opposed to deriving a direct tool that predicts clinical response. There is strong evidence to support Ki67 as the primary endpoint of neoadjuvant endocrine therapy study from multiple previous clinical trials [Preoperative Anastrozole, Tamoxifen, or Combined with Tamoxifen (IMPACT); P024 study; American College of Surgeons Oncology Group (ACOSOG) Z1031] and the level of Ki67 after treatment had been associated with

recurrence-free and overall survival (17,31,38). Noteworthy, clinical response per se is poorly related to recurrence risk on adjuvant endocrine therapy in contrast to pCR with chemotherapy in ER- and HER2+ disease. Moreover, Ki67 can be used as a marker for endocrine resistant tumour to discriminate patients requiring more aggressive treatment (15).

We identified a subgroup of *ESR1*^{Wt} tumours in which proliferation remained high after NAI therapy with less inhibition of classical and non-classical ERGs. Although ER+, tumours with higher Ki67r showed lower *ESR1*/ER expression at diagnosis, confirming that tumours with decreased dependence of ER gain less benefit from AI therapy. Interestingly, previous clinical studies have shown patients with higher levels of ER abundance measured by ligand binding assays gained greater benefit from tamoxifen in the adjuvant setting (2). Indeed, the measure of *ESR1*/ER expression may also help in the prediction of patients who would gain greater response with extended NAI (35). Moreover, the decreased dependence on ER-signalling associated with high expression of several cyclins and E2F targets support, as noted above, that patients with this phenotype may benefit from the combined use of an endocrine agent with a CDK4/6 inhibitor targeting the RB/E2F regulon.

Notably, *RET* expression appeared to increase in tumours with high residual proliferation and decreased in tumours gaining greater benefit from AI therapy. As noted above, RET has been linked with resistance to therapy and its potential as a therapeutic target has been suggested (36,37).

Tumours showing less dependence on *ESR1*/ER signalling at baseline and higher residual proliferation also showed increased expression of cell cycle control and immune response genes pre-treatment. In this context, high expression of *CDK2* was evident in tumours with high Ki67r in both pre- and post-NAI therapy. CDK2 may act 1) as a direct mitogenic driver or 2) to phosphorylate ER leading to ligand-independent ER signalling (38,39). Although *CDK2* may be a contributing factor for AI resistance in primary ER+ BC, it is important to note that only minimal differences were observed at the transcriptional level

between groups based on Ki67r. Whilst CDK2 is an obvious therapeutic target no specific inhibitors have thus far been developed for clinical use.

Similarly, the growth factor receptor *FGFR4* showed higher expression in pre-NAI tumours with high Ki67r in our cohort. FGFR4 can stimulate the proliferation of BC cells via an ER-independent mechanism (40,41). Furthermore, increased expression of *FGFR4* has previously been associated with poor response to tamoxifen (42,43). FGFR4 is a possible targetable alteration (DGIdb: http://www.dgidb.org/) and FGFR4 inhibitors (FGF401, H3B-6527 and BLU554) have been tested in phase I and phase II trials targeting other cancer types, such as in the trials NCT02508467, NCT02834780 and NCT02325739.

Tumours with high Ki67r are a very heterogenous group in our study; however, our findings show that tumours with acquired resistance exhibit smaller changes in gene expression compared to sensitive tumours and that these tumour therefore more closely resemble their diagnostic samples, an observation in keeping with the study from Selli et al. (44). Taken together, this supports the notion that the pre-surgical exposure of ER+ tumours to AI markedly enhances the ability to reveal their dependence on classical ER signalling and therefore identify mechanisms of resistance.

It is important to underscore that four cases with recurrences were observed in the branch C of our hierarchical cluster analysis in which tumours with the greatest inhibition of proliferation and oestrogen signalling were grouped. We and others previously reported that, while patients whose tumours were rated as more highly oestrogen responsive at diagnosis had a lower risk of recurrence up to 5 years, their risk was greater with further follow-up after such treatment discontinuation (40,41). This is consistent with such patients only showing lower recurrence rates when the disease is controlled by endocrine therapy.

While our study had biological strength in tumours phenotypic characterisation, some limitations should be noted. Firstly, only about half of the patients had a follow-up of more than five years impairing our ability to directly link phenotypic/genotypic alterations with risk of recurrence. Historically, NAI treatment has been selected for postmenopausal woman with

large ER+ tumours or for those who may be too frail to undergo surgery. This patient population is often older, with limited long-term follow-up (12). Secondly, although we have a representative cohort of ER+BC treated with NAI, the largest to date with extended NAI (4,16), subgroup analysis was restricted due to lack of statistical power. Nonetheless, a significant strength of the study was our access to both pre- and post-NAI samples, which enabled us to conduct comparative gene expression profiling and mutation analysis to define the acquisition of *ESR1* mutations.

In summary, overall most tumours showed little evidence for the emergence of resistant disease after NAI therapy, highlighted mainly by the continued reduced expression of proliferation genes/proteins and several genes involved in E-response. Two main groups of tumours showing possible resistance to long-term NAI therapy were observed: 1) tumours with *ESR1* mutations that were enriched with longer exposure to AI and 2) *ESR1*^{Wt} tumours with relatively low *ESR1* expression at diagnosis and high Ki67r. In both groups ligand-independent ER signalling was detected and it can be used to inform on subsequent adjuvant treatment in early ER+ BC.

Acknowledgements

The authors are thankful for Ricardo Ribas critical reading.

REFERENCES

1. Dodson A, Parry S, Ibrahim M, Bartlett JM, Pinder S, Dowsett M*, et al.* Breast cancer biomarkers in clinical testing: analysis of a UK national external quality assessment scheme for immunocytochemistry and in situ hybridisation database containing results from 199 300 patients. The journal of pathology Clinical research **2018**;4(4):262-73 doi 10.1002/cjp2.112.

2. Early Breast Cancer Trialists' Collaborative G, Davies C, Godwin J, Gray R, Clarke M, Cutter D*, et al.* Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet **2011**;378(9793):771-84 doi 10.1016/S0140-6736(11)60993-8.
3. Patani N, Martin LA. Understanding response and resistance to oestrogen deprivation in ER-positive breast cancer. Mol Cell Endocrinol **2014**;382(1):683-94 doi 10.1016/j.mce.2013.09.038.
4. Miller CA, Gindin Y, Lu C, Griffith OL, Griffith M, Shen D*, et al.* Aromatase inhibition remodels the clonal architecture of estrogen-receptor-positive breast cancers. Nat Commun **2016**;7:12498 doi 10.1038/ncomms12498.
5. Ellis MJ, Ding L, Shen D, Luo J, Suman VJ, Wallis JW*, et al.* Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature **2012**;486(7403):353-60 doi 10.1038/nature11143.
6. Razavi P, Chang MT, Xu G, Bandlamudi C, Ross DS, Vasan N*, et al.* The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. Cancer Cell **2018**;34(3):427-38 e6 doi 10.1016/j.ccell.2018.08.008.
7. Schiavon G, Hrebien S, Garcia-Murillas I, Cutts RJ, Pearson A, Tarazona N*, et al.* Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. Science translational medicine **2015**;7(313):313ra182 doi 10.1126/scitranslmed.aac7551.
8. Fribbens C, O'Leary B, Kilburn L, Hrebien S, Garcia-Murillas I, Beaney M*, et al.* Plasma ESR1 Mutations and the Treatment of Estrogen Receptor-Positive Advanced Breast Cancer. J Clin Oncol **2016**;34(25):2961-8 doi 10.1200/JCO.2016.67.3061.
9. Lopez-Knowles E, Pearson A, Schuster E, Gellert P, Ribas R, Yeo B*, et al.* Molecular characterisation of aromatase inhibitor-resistant advanced breast cancer: the phenotypic effect of ESR1 mutations. Brit J Cancer **2018**.

10. Toy W, Shen Y, Won H, Green B, Sakr RA, Will M*, et al.* ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. Nat Genet **2013**;45(12):1439-45 doi 10.1038/ng.2822.
11. 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 doi 10.1038/ng.2823.
12. Yeo B, Dowsett M. Neoadjuvant endocrine therapy: Patient selection, treatment duration and surrogate endpoints. Breast **2015**;24 Suppl 2:S78-83 doi 10.1016/j.breast.2015.07.019.
13. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, Griffith C*, et al.* Short-term changes in Ki-67 during neoadjuvant treatment of primary breast cancer with anastrozole or tamoxifen alone or combined correlate with recurrence-free survival. Clinical cancer research : an official journal of the American Association for Cancer Research **2005**;11(2 Pt 2):951s-8s.
14. Gellert P, Segal CV, Gao Q, Lopez-Knowles E, Martin LA, Dodson A*, et al.* Impact of mutational profiles on response of primary oestrogen receptor-positive breast cancers to oestrogen deprivation. Nat Commun **2016**;7:13294 doi 10.1038/ncomms13294.
15. Ellis MJ, Suman VJ, Hoog J, Goncalves R, Sanati S, Creighton CJ*, et al.* Ki67 Proliferation Index as a Tool for Chemotherapy Decisions During and After Neoadjuvant Aromatase Inhibitor Treatment of Breast Cancer: Results From the American College of Surgeons Oncology Group Z1031 Trial (Alliance). J Clin Oncol **2017**;35(10):1061-9 doi 10.1200/JCO.2016.69.4406.
16. Guerrero-Zotano AL, Stricker TP, Formisano L, Hutchinson KE, Stover DG, Lee KM*, et al.* ER(+) Breast Cancers Resistant to Prolonged Neoadjuvant Letrozole Exhibit an E2F4 Transcriptional Program Sensitive to CDK4/6 Inhibitors. Clin Cancer Res **2018**;24(11):2517-29 doi 10.1158/1078-0432.CCR-17-2904.

17. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R*, et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer **2009**;45(2):228-47 doi 10.1016/j.ejca.2008.10.026.
18. Dowsett M, Allred C, Knox J, Quinn E, Salter J, Wale C*, et al.* Relationship between quantitative estrogen and progesterone receptor expression and human epidermal growth factor receptor 2 (HER-2) status with recurrence in the Arimidex, Tamoxifen, Alone or in Combination trial. J Clin Oncol **2008**;26(7):1059-65 doi 10.1200/JCO.2007.12.9437.
19. Zabaglo L, Salter J, Anderson H, Quinn E, Hills M, Detre S*, et al.* Comparative validation of the SP6 antibody to Ki67 in breast cancer. J Clin Pathol **2010**;63(9):800-4 doi 10.1136/jcp.2010.077578.
20. Leung SCY, Nielsen TO, Zabaglo L, Arun I, Badve SS, Bane AL*, et al.* Analytical validation of a standardized scoring protocol for Ki67: phase 3 of an international multicenter collaboration. NPJ Breast Cancer **2016**;2:16014 doi 10.1038/npjbcancer.2016.14.
21. Waggott D, Chu K, Yin S, Wouters BG, Liu FF, Boutros PC. NanoStringNorm: an extensible R package for the pre-processing of NanoString mRNA and miRNA data. Bioinformatics **2012**;28(11):1546-8 doi 10.1093/bioinformatics/bts188.
22. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T*, et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol **2009**;27(8):1160-7 doi 10.1200/JCO.2008.18.1370.
23. Cheang MC, Voduc KD, Tu D, Jiang S, Leung S, Chia SK*, et al.* Responsiveness of intrinsic subtypes to adjuvant anthracycline substitution in the NCIC.CTG MA.5 randomized trial. Clin Cancer Res **2012**;18(8):2402-12 doi 10.1158/1078-0432.CCR-11-2956.
24. Dunbier AK, Anderson H, Ghazoui Z, Folkerd EJ, A'Hern R, Crowder RJ*, et al.* Relationship between plasma estradiol levels and estrogen-responsive gene

expression in estrogen receptor-positive breast cancer in postmenopausal women. J Clin Oncol **2010**;28(7):1161-7 doi 10.1200/JCO.2009.23.9616.

25. Miller TW, Balko JM, Fox EM, Ghazoui Z, Dunbier A, Anderson H*, et al.* ERalpha-dependent E2F transcription can mediate resistance to estrogen deprivation in human breast cancer. Cancer Discov **2011**;1(4):338-51 doi 10.1158/2159-8290.CD-11-0101.

26. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ*, et al.* Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. Sci Transl Med **2015**;7(302):302ra133 doi 10.1126/scitranslmed.aab0021.

27. Laenkholm AV, Knoop A, Ejlertsen B, Rudbeck T, Jensen MB, Muller S*, et al.* ESR1 gene status correlates with estrogen receptor protein levels measured by ligand binding assay and immunohistochemistry. Mol Oncol **2012**;6(4):428-36 doi 10.1016/j.molonc.2012.04.003.

28. Ellis MJ, Tao Y, Luo J, A'Hern R, Evans DB, Bhatnagar AS*, et al.* Outcome prediction for estrogen receptor-positive breast cancer based on postneoadjuvant endocrine therapy tumor characteristics. J Natl Cancer Inst **2008**;100(19):1380-8 doi 10.1093/jnci/djn309.

29. O'Leary B, Cutts RJ, Liu Y, Hrebien S, Huang X, Fenwick K*, et al.* The Genetic Landscape and Clonal Evolution of Breast Cancer Resistance to Palbociclib plus Fulvestrant in the PALOMA-3 Trial. Cancer Discov **2018**;8(11):1390-403 doi 10.1158/2159-8290.CD-18-0264.

30. Chandarlapaty S, Chen D, He W, Sung P, Samoila A, You D*, et al.* Prevalence of ESR1 Mutations in Cell-Free DNA and Outcomes in Metastatic Breast Cancer: A Secondary Analysis of the BOLERO-2 Clinical Trial. JAMA Oncol **2016**;2(10):1310-5 doi 10.1001/jamaoncol.2016.1279.

31. Martin LA, Ribas R, Simigdala N, Schuster E, Pancholi S, Tenev T*, et al.* Discovery of naturally occurring ESR1 mutations in breast cancer cell lines modelling endocrine resistance. Nat Commun **2017**;8(1):1865 doi 10.1038/s41467-017-01864-y.

32. Sanders DA, Ross-Innes CS, Beraldi D, Carroll JS, Balasubramanian S. Genome-wide mapping of FOXM1 binding reveals co-binding with estrogen receptor alpha in breast cancer cells. Genome Biol **2013**;14(1):R6 doi 10.1186/gb-2013-14-1-r6.
33. Morandi A, Plaza-Menacho I, Isacke CM. RET in breast cancer: functional and therapeutic implications. Trends Mol Med **2011**;17(3):149-57 doi 10.1016/j.molmed.2010.12.007.
34. Jeselsohn R, Bergholz JS, Pun M, Cornwell M, Liu W, Nardone A*, et al.* Allele-Specific Chromatin Recruitment and Therapeutic Vulnerabilities of ESR1 Activating Mutations. Cancer Cell **2018**;33(2):173-86 e5 doi 10.1016/j.ccell.2018.01.004.
35. Dixon JM. Prospects of neoadjuvant aromatase inhibitor therapy in breast cancer. Expert Rev Anticancer Ther **2008**;8(3):453-63 doi 10.1586/14737140.8.3.453.
36. Morandi A, Martin LA, Gao Q, Pancholi S, Mackay A, Robertson D*, et al.* GDNF-RET signaling in ER-positive breast cancers is a key determinant of response and resistance to aromatase inhibitors. Cancer Res **2013**;73(12):3783-95 doi 10.1158/0008-5472.CAN-12-4265.
37. Andreucci E, Francica P, Fearns A, Martin LA, Chiarugi P, Isacke CM*, et al.* Targeting the receptor tyrosine kinase RET in combination with aromatase inhibitors in ER positive breast cancer xenografts. Oncotarget **2016**;7(49):80543-53 doi 10.18632/oncotarget.11826.
38. Trowbridge JM, Rogatsky I, Garabedian MJ. Regulation of estrogen receptor transcriptional enhancement by the cyclin A/Cdk2 complex. Proc Natl Acad Sci U S A **1997**;94(19):10132-7.
39. Rogatsky I, Trowbridge JM, Garabedian MJ. Potentiation of human estrogen receptor alpha transcriptional activation through phosphorylation of serines 104 and 106 by the cyclin A-CDK2 complex. J Biol Chem **1999**;274(32):22296-302.
40. Tiong KH, Tan BS, Choo HL, Chung FF, Hii LW, Tan SH*, et al.* Fibroblast growth factor receptor 4 (FGFR4) and fibroblast growth factor 19 (FGF19) autocrine

enhance breast cancer cells survival. Oncotarget **2016**;7(36):57633-50 doi 10.18632/oncotarget.9328.

41. Zhao X, Xu F, Dominguez NP, Xiong Y, Xiong Z, Peng H*, et al.* FGFR4 provides the conduit to facilitate FGF19 signaling in breast cancer progression. Mol Carcinog **2018**;57(11):1616-25 doi 10.1002/mc.22884.

42. Meijer D, Sieuwerts AM, Look MP, van Agthoven T, Foekens JA, Dorssers LC. Fibroblast growth factor receptor 4 predicts failure on tamoxifen therapy in patients with recurrent breast cancer. Endocr Relat Cancer **2008**;15(1):101-11 doi 10.1677/ERC-07-0080.

43. Wei W, You Z, Sun S, Wang Y, Zhang X, Pang D*, et al.* Prognostic implications of fibroblast growth factor receptor 4 polymorphisms in primary breast cancer. Mol Carcinog **2018**;57(8):988-96 doi 10.1002/mc.22819.

44. Selli C, Turnbull AK, Pearce DA, Li A, Fernando A, Wills J*, et al.* Molecular changes during extended neoadjuvant letrozole treatment of breast cancer: distinguishing acquired resistance from dormant tumours. Breast Cancer Res **2019**;21(1):2 doi 10.1186/s13058-018-1089-5.

FIGURE LEGENDS

Figure 1. Neoadjuvant E-deprivation therapy. A) Pre- and post-NAI samples were obtained from the same patients for IHC and molecular analysis. *For *ESR1* mutation analysis, firstly the presence of the mutation was investigated in all post-NAI specimens and, once detected in a patient, it was also evaluated in pre-NAI samples. IHC: immunohistochemistry; ddPCR: digital droplet PCR. **B)** Individual patient response to NAI-therapy. Each bar represents a patient and the length of the bar shows duration of therapy. The colour of the bar shows clinical response based on ultrasound; triangles mark the timing to progression determined as a 20% increase of the tumour volume in relation to the previous ultrasound. Tumours with *ESR1* mutation are marked with •, * or #. Waterfall plot is shown together with clinicopathological parameters, ER, PgR and Ki67 immunostaining and PAM50 intrinsic subtypes. CR: complete response to therapy (green); PR: partial response (blue); SD: Stable disease (yellow); PD: progressive disease (red). NA: no data available (gray). NET: neoadjuvant endocrine therapy. RMH: Royal Marsden Hospital.

Figure 2. Overall gene expression changes. A) Changes in proliferation (Ki67 and proliferation metagene), ER/*ESR1* and *ERGs* between pre-AI and post-AI paired tumours. *PAGs*: mean of 11 proliferation genes in the PAM50 gene set (analysis performed in 84 paired tumours); *ERGs*: mean of *TFF1*, *GREB1*, *PDZK1* and *PGR.* Arrow graphs represent the individual expression (left) and the mean expression with the 95% confidence interval of the mean difference (right) in pre- and post-NAI samples. Individual blue arrows mark *ESR1* wild-type HER2- tumours, yellow arrows *ESR1* wild-type HER2+ tumours and red arrow *ESR1* mutant HER2- tumours. P-values based on paired T-test are shown. **B)** Hierarchical clustering of gene expression difference between of pre- and post-NAI tumours in 84 sample pairs (samples with all genes evaluated). Only genes showing more than 25% change are shown (n=410). Gene (row) clusters are annotated by most significant terms generated from

compute overlaps analysis in Broad Institute GSEA website (http://software.broadinstitute.org/gsea/msigdb/annotate.jsp). Hierarchical cluster is showed together with the mean difference (log2) of branches A-D and with intra-patient correlation calculated by Pearson correlation test (all genes analysed). CR: complete response to therapy (green); PR: partial response (blue); SD: Stable disease (yellow); PD: progressive disease (red). NA: not available (gray); NET: neoadjuvant endocrine therapy. **C)** Pathway analysis using GSEA. Data were derived from the mean difference Post-NAI – Pre-NAI in each presented group.

Figure 3. *ESR1* mutation in NAI-treated primary ER+BC. A) Difference between post-AI and pre-AI tumours based on mean expression of *ERGs*. Individual values are shown for *ESR1* wild-type HER2- tumours (blue bars), *ESR1* wild-type HER2+ tumours (yellow bars) and *ESR1* mutant HER2- tumours (red bars). **B)** Representative image of *ESR1* mutation validation in tumours with variant allele frequency (VAF) <1% by digital droplet PCR followed by fluorescence activated cell sorting (FACS). Images of digital droplet PCR using DNA extracted after microdissection of invasive tumours cells (top) and after FACS by vimentin (middle; stromal cells) and cytokeratin-positive cells (bottom; cancer cells) are shown together with type of mutation and VAF. Blue dots: *ESR1* mutant alleles; Green dots: *ESR1* wild-type alleles. **C)** *ERGs*, *ESR1* and *PAGs* expression in *ESR1* wild-type (blue dots and arrows) and mutant tumours (red dots and arrows). A significant reduction of these biomarkers was only detected in *ESR1* wild-type tumours. Box plot graphs represent the expression difference (Post-NAI – Pre-NAI) with individual values also shown. Arrow graphs (right) represent the mean expression of each group in pre- and post-NAI samples. **D)** Representative images of dual probe *ESR1* (green) /*CEP6* (red) Fluorescence in vitro Hybridization (FISH) in tumours harbouring *ESR1* mutation. **E)** Pathway analysis using GSEA. Data were derived from the mean difference Post-NAI – Pre-NAI in each presented group. **F)** *ESR1* mutant tumours showed less inhibition of *E2F* activation metagene. **G)**

Higher frequency of *ESR1* mutation (red dots) in patients treated for longer period of NAI. P-values based on Mann-Whitney test (box plots) or Wilcoxon (arrow plots) are shown. *ERGs*: oestrogen-regulated genes – mean of *TFF1*, *GREB1*, *PDZK1* and *PGR. PAGs*: proliferation associated genes – mean of 11 proliferation genes in the PAM50 gene set (analysis performed in 84 paired tumours). Wt: *ESR1* wild-type tumours; Mut: tumours harbouring *ESR1* mutation. *ESR1* mutation type are highlighted. #Two residual tumours with *ESR1* mutation in less than 1% of cells (Case #2 and Case #6). *Patient with *ESR1* mutation detected in both pre-NAI and post-NAI samples.

Figure 4. Gene expression in *ESR1*^{Wt} tumours based on residual Ki67. A) Intra-patient correlation (comparison of pre- and post-NAI samples from the same patient); p-value based on Spearman correlation test. **B)** Less inhibition of classical *ERGs* and PgR protein abundance in tumours with High-Ki67r. **C)** Less effect of NAI in *CCND1* and *RET* expression in tumours with High-Ki67r. **D)** *E2F* activation metagene is less inhibited with NAI in tumours with High-Ki67r. **E)** Pathway analysis using GSEA. Data were derived from the mean difference Post-NAI – Pre-NAI in each presented group. **F)** *ESR1*/ER expression in pre-NAI, post-NAI and the mean change in tumours classified by Ki67r. High-Ki67r tumours had a relatively lower pre-NAI *ESR1*/ER expression and lower ER expression post-NAI. **(B-D,F)** Box plots represent on-treatment change (left), pre-NAI or post-NAI expression, as indicated. Arrow graphs (right) represent the mean expression of each group in pre-NAI and post-NAI samples. P-values based on T-test (box plots) or paired T-test (arrow plots) are shown. Light blue: low residual Ki67 (% of +ve cells ≤2.7%, n=53). Bright blue: medium level of residual Ki67 (>2.7% & ≤10%, n=15). Dark blue: high residual Ki67 (≥10%, n=13). *ERGs*: oestrogen-regulated genes – mean of *TFF1*, *GREB1*, *PDZK1* and *PGR.* Ki67r: residual Ki67 (post-neoadjuvant AI therapy).

Figure 5. *CDK2* and *FGFR4* expression in *ESR1*^{Wt} tumours of NAI-treated patients. A) *CDK2* and *FGFR4* expression in pre- and post-NAI together with the mean change in tumours classified by Ki67r. High-Ki67r tumours had a relatively higher pre-NAI *CDK2* and *FGFR4* expression before and after NAI therapy. Box plots represent on-treatment change, pre-NAI or post-NAI expression as indicated. Arrow graphs (right) represent the mean expression of each group in pre-NAI and post-NAI samples. Light blue: low residual Ki67 (% of +ve cells ≤2.7%, n=53). Bright blue: medium level of residual Ki67 (>2.7% & ≤10%, n=15). Dark blue: high residual Ki67 (≥10%, n=13). **B)** *CDK2* and FGFR4 showed higher expression in tumours of patients with SD/PD in comparison with CR/PR in both pre-NAI and post-NAI samples. CR: complete response to therapy (green); PR: partial response (blue); SD: Stable disease (yellow); PD: progressive disease (red). Light blue dots mark cases with PR that showed clinical signs of progression disease (>20% increase of the tumour volume in relation to the previous ultrasound). P-values based on T-test (box plots) or paired T-test (arrow plots) are shown. Ki67r: residual Ki67 (post-neoadjuvant AI therapy).

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Fig. S1. Consort diagram. [a]Multifocal disease confirmed in histopathology analysis. [b]Concomitant anticancer treatments included chemotherapy, biologic response modifiers, endocrine therapy (including steroids) and radiotherapy.

Supplementary Fig. S2. Pre- and post-NAI expression of proliferation markers (Ki67 and proliferation metagene), *ESR1*/ER and *ERGs* based on clinical response stratification. *PAGs*: mean of 11 proliferation genes in the PAM50 gene set; *ERGs*: oestrogen-regulated genes – mean of *TFF1*, *GREB1*, *PDZK1* and *PGR.* CR: complete response to therapy (green); PR: partial response (blue); SD: stable disease (yellow); PD: progressive disease (red). Light blue dots mark cases with PR that showed clinical signs of progression disease (>20% increase of the tumour volume in relation to the previous ultrasound). P-values based on T-test are shown.

Supplementary Fig. S3. Correlation between protein and gene expression. A) Ki67, PgR and ER expression measured by NanoString™ technology and IHC. **B)** Correlation between Ki67 protein expression and *PAGs*. **C)** Correlation between PgR protein ad *ERGs*. Individual blue dots mark *ESR1* wild-type HER2- tumours, yellow dot *ESR1* wild-type HER2+ tumours and red dots *ESR1* mutant HER2- tumours. Light colours: pre-NAI values; Dark colours: post-NAI values. P-values and coefficient of correlation (r) based on Pearson correlation test are shown. *PAGs*: mean of 11 proliferation genes in the PAM50 gene set; *ERGs*: oestrogen-regulated genes – mean of *TFF1*, *GREB1*, *PDZK1* and *PGR.*

Supplementary Fig. S4. E2F activation metagene in NAI-therapy. Overall inhibition of *E2F* activation metagene with NAI treatment and higher post-NAI expression of this signature in patients with stable disease / progressive disease (SD/PD) in comparison with

complete or partial response (CR/PR) in both pre- and post-NAI samples based on clinical response stratification. Arrow graphs represent the individual expression (left) and the mean expression with the 95% confidence interval of the mean difference (right) in pre-NAI and post-NAI samples. P-values based on T-test (box plots) or paired T-test (arrow plots) are shown.

Supplementary Fig. S5. Gene expression based on *ESR1* mutational status. A) Pre and post-NAI mean expression of oestrogen-regulated genes (*ERGs*). Individual values are shown for *ESR1* wild-type HER2- tumours (blue bars), *ESR1* wild-type HER2+ tumours (yellow bars) and *ESR1* mutant HER2- tumours (red bars). Light colours: pre-NAI values; Dark colours: post-NAI values. **B)** *CCND1*, *RET* and *FOXM1* expression in *ESR1* wild-type (blue dots and arrows) and mutant tumours (red dots and arrows). Less inhibition of these biomarkers was detected in *ESR1* mutant tumours. Box plot graphs represent the expression difference (Post-NAI – Pre-NAI) with individual values also shown. Arrow graphs (right) represent the mean expression of each group in pre-NAI and post-NAI samples. P-values based on Mann-Whitney test (box plots) or Wilcoxon (arrow plots) are shown. *ERGs*: oestrogen-regulated genes – mean of *TFF1*, *GREB1*, *PDZK1* and *PGR*. Wt: *ESR1* wild-type tumours; Mut: tumours harbouring *ESR1* mutation. *ESR1* mutation type are highlighted.

Supplementary Fig. S6. Correlation between *ERGs* and Ki67 expression in *ESR1*^{Wt} tumours. Light blue: low residual Ki67 (% of +ve cells ≤2.7%, n=53). Bright blue: medium level of residual Ki67 (>2.7% & ≤10%, n=15). Dark blue: high residual Ki67 (≥10%, n=13). P-values and coefficient of correlation (r) based on Pearson correlation test are shown.

Supplementary Fig. S7. Change in cyclins expression in *ESR1*^{Wt} tumours classified based on Ki67r. Box plots represent on-treatment change. Arrow graphs (right) represent

the mean expression of each group in pre- and post-NAI samples. Light blue: low residual Ki67 (% of +ve cells ≤2.7%, n=53). Bright blue: medium level of residual Ki67 (>2.7% & ≤10%, n=15). Dark blue: high residual Ki67 (≥10%, n=13). P-values and coefficient of correlation (r) based on Pearson correlation test are shown. P-values based on T-test (box plots) or paired T-test (arrow plots) are shown. Ki67r: residual Ki67 (post-neoadjuvant AI therapy).

SUPPLEMENTARY TABLE LEGENDS

Supplementary Table S1. Clinical and pathological profile in the studied population. [a]Volume measured by ultrasound = (a x b x c x π)/6. [b]Previous breast cancer at least 10 years a part; Patient under any cancer treatment in the breast cancer diagnosis were excluded; [c]HER2 status by IHC and/or FISH. NA not available; CR: complete response – disappearance of all target lesions; PR: partial clinical response – regression of at least 30% of tumour volume but without complete response; SD: stable disease – neither sufficient shrinkage to qualify for regression nor sufficient increase to qualify for progression; PD: progressive disease – increase of at least 20% of tumour volume.

Supplementary Table S2. Studied genes. Genes were selected based on their key role in breast cancer development or with evidence of an association with aromatase inhibitor resistance, including oestrogen-regulated genes, proliferation, invasion, growth factor receptors, PI3K-AKT-mTOR pathway, MAPK signalling, cholesterol metabolism, inflammation and epithelial mesenchymal transition. *Reference genes for gene expression analysis. Bold letters: Genes in common in both panels.

Supplementary Table S3. Differentially expressed genes between pre- and post-NAI samples. Red: Increased expression in Post-NAI samples in relation to matched Pre-AI samples; Blue: Reduced expression in Post-NAI samples in relation to matched Pre-AI samples. *p-value by paired T-test. **All genes had adjusted p-value < 0.05.

Supplementary Table S4. Genes involved in immune response that are upregulation in branch C of hierarchical clustering analysis. Genes are annotated by most significant terms generated from compute overlaps analysis in Broad Institute GSEA website (http://software.broadinstitute.org/gsea/msigdb/annotate.jsp).

Supplementary Table S5. *ESR1* mutation, copy number and clinical and pathological features. [a]Presence and absence of *ESR1* mutation was also confirmed by ddPCR after fluorescence activated cell sorting (FACS) to enrich the number of cytokeratin-positive neoplastic cells evaluated; [b]Volume measured by ultrasound = (a x b x c x π)/6; [c]Previous breast cancer at least 10 years a part; Patient under any cancer treatment in the breast cancer diagnosis were excluded; [d]HER2 status by IHC and/or FISH. VAF: Variant allele frequency; NA: not available; PR: partial clinical response - regression of at least 30% of tumour volume but without complete response; SD: stable disease - neither sufficient shrinkage to qualify for regression nor sufficient increase to qualify for progression.

Supplementary Table S6. Differentially expressed genes between *ESR1*^{Mut} and *ESR1*^{Wt} tumours pre-NAI therapy. Red: Increased expression in *ESR1* mutant carriers in relation to *ESR1* wild-type tumours; Blue: Reduced expression in *ESR1* mutant carriers in relation to *ESR1* wild-type tumours. *p-value by Mann-Whitney test.

Supplementary Table S7. Differentially expressed genes between *ESR1*^{Mut} and *ESR1*^{Wt} tumours post-NAI therapy. Red: Increased expression in *ESR1* mutant carriers in relation to *ESR1* wild-type tumours; Blue: Reduced expression in *ESR1* mutant carriers in relation to *ESR1* wild-type tumours. *p-value by Mann-Whitney test.

Supplementary Table S8. On-treatment change (post-NAI – pre-NAI) comparison between *ESR1*^{Mut} and *ESR1*^{Wt} tumours. Red: Increased change in *ESR1* mutant carriers in relation to *ESR1* wild-type tumours; Blue: Reduced change in *ESR1* mutant carriers in relation to *ESR1* wild-type tumours. *p-value by Mann-Whitney test.

Supplementary Table S9. Differentially expressed genes between tumours with Low-, Medium- and High-Ki67r pre-NAI. Red: Increased expression in relation to the reference group (Low or Medium Ki67r); Blue: Reduced expression in relation to the reference group (Low or Medium Ki67r). Bold letters: p-value <0.05 or adjusted p-value < 0.1. Ki67r Low: ≤2.7% of positive invasive cells; Ki67r Medium: >2.7% & <10% of positive invasive cells; Ki67r High: ≥10% of positive invasive cells. *p-value by T-test.

Supplementary Table S10. Differentially expressed genes between tumours with Low-, Medium- and High-Ki67r post-NAI. Red: Increased expression in relation to the reference group (Low or Medium Ki67r); Blue: Reduced expression in relation to the reference group (Low or Medium Ki67r). Bold letters: p-value <0.05 or adjusted p-value < 0.1. Ki67r Low: ≤2.7% of positive invasive cells; Ki67r Medium: >2.7% & <10% of positive invasive cells; Ki67r High: ≥10% of positive invasive cells. *p-value by T-test.

Supplementary Table S11. On-treatment change (post-NAI – pre-NAI) comparison based on residual Ki67. Red: Increased change in relation to the reference group (Low or Medium Ki67r); Blue: Reduced change in relation to the reference group (Low or Medium Ki67r). Bold letters: p-value <0.05 or adjusted p-value < 0.1. Ki67r Low: ≤2.7% of positive invasive cells; Ki67r Medium: >2.7% & <10% of positive invasive cells; Ki67r High: ≥10% of positive invasive cells. *p-value by T-test.

Supplementary Table S12. Gene set enrichment analysis based on genes showing higher expression in pre-NAI tumours with High-Ki67r in relation with Low-Ki67r. *Annotated by most significant terms generated from compute overlaps analysis in Broad Institute GSEA website (http://software.broadinstitute.org/gsea/msigdb/annotate.jsp).