bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 **Molecular changes during extended neoadjuvant letrozole treatment of**
2 **breast cancer: distinguishing acquired resistance from dormant tumours**

3 Cigdem Selli[1,2], Arran K. Turnbull[1,3], Dominic A. Pearce[1], Ang Li[1], Anu Fernando[1,3], Jimi
4 Wills[4], Lorna Renshaw[3], Jeremy S. Thomas[3], J. Michael Dixon[3], Andrew H. Sims[1]*
5
6 [1] Applied Bioinformatics of Cancer, University of Edinburgh Cancer Research UK Centre,
7 MRC Institute of Genetics and Molecular Medicine, Edinburgh, UK
8 [2] Department of Pharmacology, Faculty of Pharmacy, Ege University, 35040, Izmir, Turkey
9 [3]Edinburgh Breast Unit, Western General Hospital, Edinburgh, UK
10 [4]Mass Spectrometry Unit, MRC Institute of Genetics and Molecular Medicine, Edinburgh,
11 UK
12
13 *Corresponding author: Andrew H Sims, PhD, Tel: + 44 (0) 131 651 8595;
14 andrew.sims@ed.ac.uk
15 Cigdem Selli, cigdem.sellli@ed.ac.uk
16 Arran K. Turnbull, arran.turnbull@ed.ac.uk
17 Dominic A. Pearce, dominic.pearce@ed.ac.uk
18 Ang Li, s1667626@sms.ed.ac.uk
19 Anu Fernando, anu.fernando@ed.ac.uk
20 Jimi Wills, jimi.wills@ed.ac.uk
21 Lorna Renshaw, lorna.renshaw@nhslothian.scot.nhs.uk
22 Jeremy S. Thomas, jeremy.thomas@nhslothian.scot.nhs.uk
23 J. Michael Dixon, mike.dixon@ed.ac.uk
24

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 # Abstract

2 **Background:** The risk of recurrence for endocrine-treated breast cancer patients persists for
3 many years or even decades following surgery and apparently successful adjuvant therapy.
4 This period of dormancy and acquired resistance is inherently difficult to investigate,
5 previous efforts have been limited to *in vitro* or *in vivo* approaches. In this study, sequential
6 tumour samples from patients receiving extended neoadjuvant endocrine treatment were
7 characterised as a novel clinical model.

8 **Methods**: Consecutive tumour samples from 62 patients undergoing extended (4-45 months)
9 neoadjuvant aromatase inhibitor, letrozole, therapy were subjected to transcriptomic and
10 proteomic analysis, representing before (≤0), early-on (13-120 days) and long-term (>120
11 days) neoadjuvant letrozole treatment. Patients with at least a 40% initial reduction in tumour
12 size by 4 months of treatment were included. Of these, 42 patients with no subsequent
13 progression were classified as "dormant", and the remaining 20 patients as "acquired
14 resistant".

15 **Results**: Changes in gene expression in dormant tumours begin early and become more
16 pronounced at later timepoints. Therapy-induced changes in resistant tumours were common
17 features of treatment, rather than being specific to resistant phenotype. Comparative analysis
18 of long-term treated dormant and resistant tumours highlighted changes in epigenetics
19 pathways including DNA methylation and histone acetylation. DNA methylation marks 5-
20 methylcytosine and 5-hydroxymethylcytosine were significantly reduced in resistant tumours
21 compared to dormant tissues after extended letrozole treatment.

22 **Conclusions**: This is the first patient-matched gene expression study investigating long-term
23 aromatase inhibitor-induced dormancy and acquired resistance in breast cancer. Dormant
24 tumours continue to change during treatment whereas acquired resistant tumours more

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

closely resemble their diagnostic samples. Global loss of DNA methylation was observed in resistant tumours under extended treatment. Epigenetic alterations may lead to escape from dormancy and drive acquired resistance in a subset of patients supporting a potential role for therapy targeted at these epigenetic alterations in the management of endocrine resistant breast cancer.

Keywords: dormancy, endocrine treatment, epigenetics, letrozole, sequential samples, resistance, microarray, proteomics

Background

Approximately 70% of breast cancer patients who have oestrogen receptor alpha (ER) positive tumours receive adjuvant endocrine treatment. Five years of aromatase inhibitor therapy produces a 40% reduction in 10-year mortality [1]. However, while the annual risk of mortality for ER-negative breast cancer decreases following the first five years after diagnosis, the annual rate remains constant for ER+ patients [2]. In fact, women with ER+ early-stage disease treated with 5 years of adjuvant endocrine therapy have a persistent risk of recurrence and death from breast cancer for at least 20 years after diagnosis [3]. Molecular studies have demonstrated that nodal and distant metastasis are highly similar to their matched primary tumours, implicating a continuation of original cancer [4-6]. However, the time between treatment and recurrence is often greater than that which can be explained by normal cell-doubling rates [7], implying cancer cells remain dormant in the body before re-awakening.

Residual dormant cancer cells are hypothesised to persist either by withdrawing from the cell cycle and transitioning to a quiescence state or by continuing to proliferate at a reduced rate, counter-balanced by cell death [8]. Reawakened dormant cells may become detectable after reaching a detection threshold or reactivated via increased angiogenesis, and/or escape from

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 inhibitory microenvironment or immune effects [9, 10]. Dormancy is therefore considered a
2 major mechanism underlying resistance to therapy, where dormant cells survive despite anti-
3 proliferative endocrine treatment.

4 Resistance to endocrine therapy may occur at disease inception (*de novo* or innate resistance),
5 but a larger proportion of patients acquire resistance during treatment (acquired/secondary
6 resistance) [11]. Several mechanisms of endocrine resistance have been described previously
7 [12, 13]. However, the majority of these findings are based on preclinical data obtained from
8 cell lines and animal models. It is therefore difficult to know if these accurately reflect
9 molecular changes in patient tumours.

10 Expression profiling of clinical samples, measuring the effect of, or predicting response to
11 treatment has recently become feasible. However, experimental design issues, such as the
12 difficulty in obtaining paired samples for comparison particularly for longer time intervals,
13 makes it difficult to study changes within tumours [14]. For example, a previous study
14 investigating tamoxifen failure compared samples from patients requiring salvage surgery
15 with pre-treatment samples from an unrelated group of disease-free patients [15]. More
16 recently, sequential patient-matched samples have been successfully utilised to determine
17 treatment-induced dynamic changes in tumours at 2 weeks to 3 months, demonstrating the
18 effectiveness of this approach [16-18].

19 For a variety of reasons, including being unfit for surgery, a proportion of patients receiving
20 pre-surgical endocrine treatment do not have their tumours excised following 3-4 months of
21 treatment. These long-term endocrine-treated tumours represent a unique group that can
22 inform how tumours respond to extended oestrogen deprivation *in situ*. Having initially
23 shrunk in size, some tumours remain at a steady volume and appear dormant, whilst others
24 subsequently begin to regrow. We have utilised this unique cohort of sequential samples from

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 patients receiving extended-neoadjuvant endocrine treatment to characterise luminal breast
2 cancer dormancy and acquired resistance using as a novel clinical model.

3 # Methods

4 ## Patients and samples

5 Breast cancer patients were treated with neoadjuvant letrozole (Femara, 2.5 mg; Novartis
6 Pharma AG, Basel, Switzerland) for a minimum of four months, tumours were not removed
7 either because patients declined or were unfit for surgery. The study was approved by the
8 local regional ethics committee (07/S1103/26, August 2007) and all patients gave informed
9 consent. Clinical characteristics of the tumours are given in Table 1. Cohort size with
10 inclusion and exclusion criteria are given in Additional file 1: Figure S1. Patients with >40%
11 initial decrease in tumour size by 4 months of treatment were included in the study. Those
12 with no subsequent progression on imaging by the latest biopsy were classified as “dormant”,
13 otherwise, they were classified as “acquired resistant” (Fig. 1a-b). For patients whose latest
14 USS measurement was taken more than a month before surgery, changes in three widely used
15 proliferation markers (MKI67, PCNA and, MCM2) were used to assist classification.
16 Sequential tumour biopsies were taken with a 14-gauge needle before and after letrozole
17 treatment and at the time of surgery. Fresh samples were snap-frozen in liquid nitrogen and
18 each tumour sample confirmed to contain ≥ 50% cellularity and at least 60% tumour tissue
19 using H&E sections. Following pulverisation of tissue with a membrane disruptor (Micro-
20 Dismembrator U, Braun Biotech), phase separation was performed by guanidinium
21 thiocyanate-phenol-chloroform extraction (Qiazol Lysis Reagent).

22 ## Gene expression profiling and analysis

23 RNA was extracted from the aqueous phase by column-based purification (RNeasy mini kit,
24 Qiagen) and then labelled and hybridized (HumanHT-12 v4 Illumina BeadChip) according to

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 manufacturer's protocol (NuGEN) as previously described [19, 20]. Raw data was detection
2 ($p < 0.05$, ≥ 3 samples) and quality filtered, log2 transformed, and quantile normalized using
3 the Bioconductor lumi package [21]. Data is available from NCBI GEO under accession
4 GSE111563. The analysis also includes data from 14 patients (42 samples, GSE59515) and 9
5 patients (24 samples, GSE55374) from previous studies [16, 19]. Hierarchical clustering
6 analysis was performed using a complete linkage method. Pathway enrichment analysis and
7 visualisation were performed using ReactomePA [22]. Differential gene expression analysis
8 was performed with Rank Products [23]. The significance of differences was evaluated by
9 using unpaired Wilcoxon test for two groups and ANOVA with post-hoc Tukey HSD for
10 multiple comparisons.

11 **Proteomics analysis**

12 Proteins were isolated from the organic phase of Qiazol [24]. Pellets were sonicated and
13 dissolved in 1% SDS. Proteomics was performed using Thermo Q Exactive plus and Label-
14 free Quantitation (LFQ). Peptides obtained from samples were analysed in mass spectrometry
15 runs, serial samples from the patients were run on the same day. A modified version of Filter
16 Aided Sample Preparation (FASP) was performed using serial digests with lysC and trypsin
17 to generate two orthogonal fractions per sample [25, 26]. The mass spectrometry spectra
18 generated in each run was used for relative quantitation of individual peptides. Normalization
19 and quantifications of peptides were performed using MaxLFQ and MaxQuant [27]. A total
20 of 6251 protein groups were identified. Data was log2 transformed and missing values were
21 imputed as the minimum observed value in each sample. The data have been deposited to the
22 ProteomeXchange Consortium via the PRIDE [28] partner repository with the dataset
23 identifier PXD009328.

24 **Immunohistochemistry and scoring**

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 Formalin-fixed paraffin-embedded (FFPE) sections were processed using an automated
2 stainer (Leica Biosystems, Bond III). Heat-induced epitope retrieval for both antibodies was
3 done by 30-minute incubation in citrate based pH 6.0 epitope retrieval (ER1) solution
4 followed by incubation in 3.5 N HCl for 15 min at room temperature as suggested by Haffner
5 *et.al.* [29]. For 5-methylcytosine (5-mC) and 5-hydroxymethylcytosine (5-hmC) detection,
6 mouse monoclonal 5 methylcytosine specific [33D3] (Abcam, ab10805) and rabbit
7 polyclonal 5-hydroxylmethylcytosine (Active Motif, 39769) antibodies were used
8 respectively. Both antibodies were used at 1/1000 dilution and were incubated for 15 min.
9 Detection was performed by using secondary antibody-horseradish peroxidase (HRP)
10 conjugates and substrate-chromogen (DAB). After staining, slides were counterstained with
11 haematoxylin. Nuclear staining in epithelial cells was evaluated using an H-score obtained by
12 multiplying the intensity of the stain (0: no staining; 1: weak staining; 2: moderate staining;
13 3: intense staining) by the percentage of cells (H-score range, 0 to 300).

14 ## Results

15 **Long-term endocrine therapy as a model of dormancy and acquired resistance**

16 A cohort of 62 primary breast cancer patients receiving at least 4 months of endocrine
17 treatment (Fig. 1a) were stratified into two groups, 'dormant' and 'acquired resistant' based
18 on dynamic changes in tumour size and proliferation (Fig. 1b). Patient-matched sequential
19 samples were available at three time-points, before (pre, ≤0 days), early-on (early, 13-120
20 days) and long-term (long, >120 days) treatment. Dormant and acquired resistant samples
21 were distributed uniformly with respect to time on treatment, and duration at each time point
22 was not significantly different between response groups (Table 1). For long-term treatment,
23 the mean and range were 186 (121-884) days and 226 (121-1366) days, for dormant and
24 acquired resistant patients respectively (Fig. 1c).

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 There were no significant differences in patient clinico-pathological features between
2 response classes before treatment (Table 1). However, PAM50 intrinsic molecular subtypes
3 were found to change during endocrine treatment (Fig. 1d). These changes were consistent
4 with known associations with outcome, with all dormant tumours either remaining the same,
5 or switching to better prognosis luminal A or normal-like tumours. For resistant tumours,
6 however, 25% (5 out of 20) switched to a subtype of worse prognosis (Fig. 1d).

7 As expected, Kaplan-Meier survival analysis demonstrated significantly worse outcomes for
8 resistant compared to dormant patients (log rank, p = 0.026, Fig. 1e). Recurrence rates for
9 dormant and resistant patients were 21% (9/42) and 45% (9/20), respectively. Moreover,
10 resistant patients suffered significantly earlier recurrences compared to dormant patients (p =
11 0.05; range = 26-947 vs 136-2042 days; Fig. 1f).

12 **Distinct transcriptomic changes under long-term letrozole treatment**

13 Unsupervised analysis was performed to consider whether sequential samples displayed
14 greater similarity between response classes or treatment duration. Hierarchical clustering
15 using the 500 genes with the highest variance across all samples revealed two main
16 subclasses, seemingly driven by time on treatment, with resistant and dormant tumours
17 indistinguishable. When long-term treated samples were considered alone, two clusters did
18 emerge, the larger of which contained a majority of dormant samples (79%), whilst the
19 second had a roughly even proportion of dormant (48%) and resistant (52%) samples (Fig.
20 2a). Similarly, a multidimensional scaling (MDS) plot revealed consistent changes over time
21 in response to treatment for both dormant and acquired resistant samples (Fig. 2b), although
22 long-term dormant samples were much more distinguishable from pre-treatment samples than
23 the long-term acquired resistant samples (Fig. 2b).

24 Correlations between tumours from different individuals (inter-patient) remained similar at
25 each time point and were not different between response classes (not shown). However,

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

correlations between matched sequential samples (intra-patient) revealed that pre-treatment samples were significantly ($p = 0.01$) less similar to their long-term treated pairs (median = 0.89, range = 0.74-0.95) than their early-on treatment pairs (median = 0.91, range = 0.84-0.95) (Fig. 2c). However, when divided by dormancy status this finding was only significant ($p = 0.01$) for dormant patients (Fig. 2c), suggesting that dormant tumours continue to diverge transcriptionally, whereas acquired resistant tumours do not consistently differ after initial or extended treatment, as mirrored in the MDS representation (Fig. 2b).

Changes in genes/pathways following long-term letrozole treatment

Pairwise Rank Product analysis (pre- vs. long-term treatment, FDR < 0.01) of dormant patients identified 2319 genes significantly differentially expressed (1063 down- and 1256 up-regulated) (Additional file 2: Table S1). These genes were significantly enriched ($p < 0.01$) for a total of 62 and 26 pathways, respectively (Additional file 2: Table S2), including reductions in cell cycle, senescence, DNA methylation and an increase in extracellular matrix (ECM) organization. These findings are consistent with previous studies of patient-matched sequential endocrine treated samples [16-18]. Acquired resistant tumours displayed much fewer consistently differentially expressed genes (238, 63 down- and 175 up-regulated) between long-term treated and pre-treatment samples (Additional file 2: Table S3). Genes that were up-regulated in resistant patients were enriched for several of the same pathways as dormant tumours (ECM organization, Elastic fibre formation and Platelet degranulation), but down-regulated genes were much more variable (Additional file 2: Table S4; Fig. 3a-b).

Having determined that dormant and acquired resistant tumours have somewhat distinct changes during treatment at the molecular level, the question remained as to whether these changes tend to occur at earlier time points or were specific to long-term treatment. For dormant tumours, differential expression begins early-on, but becomes more pronounced at later timepoints (Fig. 3a). Down-regulated genes were most evident early-on treatment for

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 resistant patients, consistent with their initial response to treatment; whilst up-regulated genes
2 were most changed after long-term treatment, potentially suggesting that these genes may
3 mediate acquired resistance (Fig. 3b). We further examined whether differentially expressed
4 genes identified in each response class were shared (Fig. 3c-d). Both down- and up-regulated
5 genes identified in resistant tumours were significantly changed ($p < 0.01$) in dormant
6 patients (Fig. 3d). However, only up-regulated genes identified in dormant patients were
7 significantly up-regulated in resistant patients without any change in down-regulated genes
8 (Fig. 3c), implicating a partial lack of response to treatment at the molecular level in acquired
9 resistance patients.

10 **A potential role of epigenetic regulation in acquired resistance**

11 The above findings suggested that therapy-induced dynamic changes in genes and pathways
12 are common features of long-term treatment, rather than being specific to dormant or resistant
13 phenotypes. This led us to perform comparative analysis of dormant and acquired resistant
14 tumours at long-term time-point to identify any specific differences. Unpaired Rank Product
15 analysis (FDR < 0.01) revealed a total of 419 genes (170 down- and 249 up-regulated) to be
16 differentially expressed between long-term treated dormant and resistant tumours (Additional
17 file 2: Table S5; Fig. 4a). These genes were significantly enriched in 27 pathways ($p < 0.05$),
18 including several epigenetics-related pathways, including "DNA methylation", "PRC2
19 methylates histones and DNA", "histone acetyl transferases (HATs) acetylate histones",
20 "epigenetic regulation of gene expression" as well as senescence and cell cycle (Additional
21 file 2: Table S6; Fig. 4b). Examination of the expression of these genes alone, demonstrated
22 that they could partially separate dormant from majority of resistant tumours (Fig. 4c).
23 Single-sample Gene Set Enrichment Analysis (ssGSEA) [30] was performed to quantitatively
24 score the activity of differentially expressed genes in every sample. The differentially up-

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 regulated genes' score was significantly higher in acquired resistant compared to dormant
2 tumours under early-on ($p < 0.05$) as well as long-term ($p < 0.001$) treatment (Fig. 4d).

3 Our results prompted us to examine whether the changes we observed in clinical samples
4 were similarly changed in experimental models of resistant breast cancer cells. Oestrogen
5 receptor-positive MCF7 cells stably transfected with the aromatase gene (MCF7aro cells) and
6 long-term oestrogen-deprived (LTED) breast cancer cells have been widely used to
7 understand mechanisms of aromatase inhibitor resistance *in vitro*. Examining two publicly
8 available gene expression datasets (GSE10879 and GSE10911) demonstrated that genes
9 differentially expressed between acquired resistant and dormant tumours were significantly
10 enriched in aromatase inhibitor-resistant cells compared to sensitive/control cells (Fig. 5a). In
11 two out of three *in vitro* studies with dynamic gene expression data from LTED MCF7 cells
12 with an initial decrease in ssGSEA scores mimicking the dormancy/responsive state was
13 followed by a later-on increase representing acquired resistance (Fig. 5b), further validating
14 our results and emphasizing the utility of these *in vitro* models. Interestingly, no significant
15 difference was observed in tamoxifen- and fulvestrant-resistant MCF7 cells compared to
16 drug-sensitive control cells (Fig. 5c) suggesting the specificity of the results to aromatase
17 inhibitor therapy resistance.

18 In addition, proteomic analysis of a subset of samples was performed which revealed
19 differential expression in 656 proteins (279 down-, 377 up-regulated) between long-term
20 treated dormant and resistant tumours (Rank Product; $p < 0.05$; n = 10; Additional file 2:
21 Table S7; Fig. 5d). A total of 36 features including S100P and HIST2H3A (H3.2) overlapped
22 between proteomics and transcriptomics, validating the results with a different approach.

23 Next, we considered which drugs might be able to reverse the differences in gene expression
24 between dormant and resistant tumours using Connectivity Map (CMap), the pattern
25 matching software, and a collection of gene-expression profiles from cultured human cells

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 treated with various treatments [31]. Drugs that are negatively associated with the resistance
2 signature could potentially reverse the differences observed. A histone deacetylase (HDAC)
3 inhibitor Trichostatin A had the second lowest score (-0.90) and has previously been shown
4 to re-sensitise tamoxifen-unresponsive ER-negative breast cancer cells *in vitro* [32].
5 Clioquinol an anti-parasitic and apoptotic drug with HDAC inhibitory effects [33] was also
6 highlighted. Letrozole had a positive score of 0.89 further confirming the reliably of the
7 predictions and hypothetical scores calculated by CMap. Furthermore, differentially
8 expressed genes were uploaded to Enricher (ENCODE Histone modification 2015 dataset)
9 [34] to determine histone modification enrichment. Two H3 lysine methylation modifications
10 (H3K27me3 and H3K4me1) were enriched significantly (Adjusted $p = 0.0003$ and $p = 0.004$,
11 respectively) whereas no enrichment for histone acetylation was determined.

12 Finally, immunohistochemical evaluation of FFPE sections revealed significantly lower
13 global 5-mC and 5-hmC levels in resistant tumours compared to dormant tumours under
14 extended treatment (Fig. 6a-b). Significantly lower 5-hmC levels in acquired resistant
15 compared to dormant tumours were also observed at early on-treatment (Fig. 6b), suggesting
16 hypomethylation may be predictive of emergence from dormancy.

17 ## Discussion

18 Understanding the mechanisms underlying the maintenance of and escape from dormancy
19 have great importance considering most cancer-related deaths are caused by metastasis rather
20 than the primary tumour. In this study, we describe the first sequential patient-matched
21 clinical dataset of extended endocrine treatment in breast cancer. The results highlight the
22 difficulty of distinguishing dormant and resistant tumours, with dynamic molecular changes
23 of treatment being highly similar between the groups. However, comparative analysis

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 revealed a set of genes significantly up-regulated in resistant patients within the first months
2 of letrozole treatment suggesting a predictive role for changes in DNA methylation.

3 Failure to reduce proliferation after 2 weeks of endocrine treatment [16, 35] may well
4 identify patients that are innately resistant; however, acquired resistance remains a greater
5 challenge in terms of identifying biomarkers and appropriate alternative or combination
6 therapies [36]. Many of the transcriptomic changes identified in long-term treated dormant
7 tumours are shared by some, but not all resistant tumours, providing further evidence of
8 resistance heterogeneity [37] where dormant tumours share similar molecular changes, but
9 there are a variety of escape mechanisms that lead to acquired resistance.

10 In the present study, paired differential expression analysis demonstrated that dormant
11 tumours continue to change under long-term treatment. Some of the identified dormancy-
12 related pathways such as cell cycle arrest and senescence have established roles in metastasis
13 dormancy [38] further supporting the relevance of our clinical model, with senescence-
14 associated secretory phenotype (SASP) recently suggested to regulate breast cancer
15 dormancy and relapse [39]. As in short-term responsive tumours [16], ECM organization and
16 degradation were significantly up-regulated in dormant tumours. ECM remodelling and its
17 degradation by matrix metalloproteases (MMP) have previously been suggested to regulate
18 the switch between dormancy and metastatic growth [40].

19 The most up-regulated gene in resistant tumours *S100P*, previously shown to be an inducer of
20 breast cancer metastasis correlated with decreased survival [41]. Recently, *S100P*
21 hypomethylation in blood was demonstrated to be inversely correlated with tissue *S100P*
22 expression and significantly associated with breast cancer, implicating *S100P* as a potential
23 diagnostic marker [42]. High plasma *S100P* levels have also been correlated with poor
24 prognosis in metastatic breast cancer patients, with levels decreasing following treatment,
25 suggesting a role of S100P in dynamic monitoring of response [43]. In the present study,

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 S100P gene expression and protein levels were significantly higher in resistant tumours after
2 long-term treatment, as well as being differentially expressed before treatment supporting its
3 potential role as a therapeutic target [44] and a predictive marker.

4 Comparative analysis of dormant and resistant samples after extended treatment revealed a
5 set of genes enriched for DNA methylation and histone acetylation/deacetylation, indicating
6 the involvement of epigenetic regulation in escape from dormancy. Epigenetic alterations are
7 recognized to occur in breast cancer. DNA methyltransferase (DNMT) and HDAC inhibitors
8 have been shown to exert encouraging effects in the disease [45]. Recently, the potential role
9 of epigenetic changes in regulating dormancy and reactivation state has been suggested to
10 explain the reversible (on/off) nature of dormancy [46].

11 Breast cancer "CpG island methylator phenotype" (CIMP), revealed by genome-wide
12 methylation analysis of metastatic breast cancers where a large number of genes are
13 hypermethylated, and has been suggested to be informative for metastatic potential [47]. A
14 significant correlation between pre-treatment global DNA methylation with neoadjuvant
15 chemotherapy response in rectal cancer has been reported [48]. Though DNA
16 hypomethylation was the first epigenetic alteration identified in cancer, its molecular process
17 and effects are not well understood yet [49]. In addition, 5-hmC levels were shown to
18 correlate with differentiation status, with higher levels in more differentiated [29]. In
19 addition, alterations in DNA methylation in LTED MCF7 cells have been previously reported
20 [50]. Our results provide evidence for loss of global DNA methylation process in resistant
21 tumours and strengthen the case to use these models for further study. The global decrease in
22 5-mC may account for the observed reduction in 5-hmC levels, since 5-mC is converted to 5-
23 hmC. On the other hand, at early-on time point, 5-hmC levels were significantly reduced with
24 no significant change in 5-mC levels suggesting an independent role of 5-hmC mark.
25 Hypomethylated cancer cells have been suggested to be selected to form tumours with

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 increased malignancy [49]. We suggest that hypomethylation in resistant tumours may reflect
2 dedifferentiation process inducing stem-cell-like cell formation. Determining the time point
3 at which that hypomethylation starts, which would allow intervening before it starts to
4 prevent therapy resistance, needs further investigation.

5 The main genes significantly enriched for epigenetics-associated pathways in the present
6 study are core histone (H3, H4, H2B) genes. Well-known epigenetics-associated genes such
7 as DNA methyltransferase (DNMT) were not differentially expressed in the present study.
8 Therefore, it might be suggested that observed changes in histone gene levels may simply
9 reflect the high proliferation rate in resistant tumours. Although these histone genes are
10 replication-dependent and their levels increase during DNA replication [51], we suggest that
11 elevated levels may also result from disrupted DNA replication and packaging. Deregulation
12 of histone H2A and H2B were associated with anthracycline resistance in breast cancer cells
13 and reversed by HDAC small molecule inhibitors [52]. Furthermore, up-regulation of
14 replication-dependent core histone proteins has been suggested to be selective indicator of
15 ER-mediated MCF7 cell proliferation regardless of proliferation-rate [53]. Also, observed
16 global loss of DNA methylation in resistant tumours suggests dynamic regulation of gene
17 transcription under letrozole therapy. Therefore, histone up-regulation and alterations in
18 epigenetic pathways observed in our study may mediate reactivation of ER signalling in
19 resistant tumours, rather than simply mirroring the degree of proliferation.

20 Our results indicate alterations both in DNA methylation, and histone modifications
21 suggesting a cooperative interplay between them to mediate acquired resistance. HDAC
22 inhibitors, which have been shown to regulate DNA methylation [54], may be successful
23 clinically as second-line drugs alone or in combination following endocrine therapy failure as
24 there is growing evidence for their tumour selective action [55, 56]. A time-dependent role
25 for HDACs in leukaemia has been shown [57] and may also be critical in determining when

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 to start HDAC inhibition therapy to successfully treat endocrine resistant patients. Whether or
2 not the epigenetic alterations are triggers of reawakening and if the timely use of epigenetic
3 drugs can prevent acquired resistance warrants further investigation.

4 **Conclusions**

5 We have performed the first study of sequential tumour samples from breast cancer patients
6 receiving extended neoadjuvant endocrine treatment as a clinical model of dormancy and
7 acquired resistance. Our analysis suggests that molecular differences between dormant and
8 resistant tumours are initially subtle, becoming more obvious only after extended treatment.
9 This study emphasizes alterations in DNA methylation in the first months of treatment may
10 predict which patients will eventually develop acquired resistance. We provide valuable
11 evidence that epigenetic drugs such as HDAC inhibitors may have a role in treating endocrine
12 resistance.
13

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 **Abbreviations**

2 5-hmC: 5-hydroxymethylcytosine; 5-mC: 5-methylcytosine; CIMP: CpG island methylator
3 phenotype; cMap: Connectivity map; DAB: 3,3'-diaminobenzidine; DNMT: DNA
4 methyltransferase; ECM: extracellular matrix; ER: oestrogen receptor; FASP: Filter Aided
5 Sample Preparation; FDR: false discovery rate; FFPE: formalin-fixed paraffin embedded;
6 HATs: histone acetyl transferases; HCl: hydrochloric acid; HDAC: histone deacetylase; HRP:
7 horseradish peroxidase; LFQ: Label-free Quantitation; LTED: long-term oestrogen
8 deprivated; MDS: multidimensional scaling; MMP: matrix metalloproteases; SASP:
9 senescence-associated secretory phenotype; ssGSEA: Single-sample Gene Set Enrichment
10 Analysis.

11 **Declarations**

12 *Ethics approval and consent to participate*

13 All patients provided informed consent and sample collection was approved by the local
14 research ethics committee (Lothian Local Research Ethics Committee 03, REC Reference
15 number 07/S1103/26, approval date 13/08/2007).

16 *Consent for publication*

17 Not applicable

18 *Availability of data and material*

19 The microarray dataset generated during the current study is available in NCBI GEO under
20 accession GSE111563 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE111563].
21 The analysis also includes previously published microarray data under accession numbers
22 GSE59515 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59515], and
23 GSE55374 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE55374]. The
24 proteomics dataset generated during the current study is available in PRIDE with the

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

identifier PXD009328 [https://www.ebi.ac.uk/pride/archive/]. Publicly available resistant cell line gene expression datasets GSE10879 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE10879], GSE10911 [58], GSE20361 [59], GSE50820 [60], GSE75971 [61], GSE14986 [62], GSE74391 [63] were also analysed.

Competing interests

The authors declare that they have no potential conflicts of interest.

Funding

The work was supported Marie Skłodowska-Curie Individual Fellowship (H2020-MSCA-IF, 658170) to CS. AHS and JMD are very grateful for funding provided by Breast Cancer Now. The work was partly supported by Welcome Trust Institutional Fund (ISSF3) to CS and AHS.

Authors' contributions

AT, JMD and AHS conceived the study. CS and AT generated the transcriptome dataset. JW conducted proteomics and supported proteomics data analysis. DP, AL, and AHS provided help with the data analysis. AF and LR provided technical support with tissue collection and processing. LR, JST, and JMD co-ordinated the collection and assessment of clinical samples. CS analysed and interpreted the data and drafted the manuscript. AHS supervised the project and helped to write and edit the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

Authors' information

Not applicable

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. Lancet. 2015;386(10001):1341-1352.
2. Demicheli R, Ardoino I, Boracchi P, Coradini D, Agresti R, Ferraris C, et al. Recurrence and mortality according to estrogen receptor status for breast cancer patients undergoing conservative surgery. Ipsilateral breast tumour recurrence dynamics provides clues for tumour biology within the residual breast. BMC Cancer. 2010;10:656.
3. Pan HC, Gray R, Braybrooke J, Davies C, Taylor C, McGale P, et al. 20-Year Risks of Breast-Cancer Recurrence after Stopping Endocrine Therapy at 5 Years. N Engl J Med. 2017;377(19):1836-1846.
4. Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ. Gene expression profiles of primary breast tumors maintained in distant metastases. Proc Natl Acad Sci U S A. 2003;100(26):15901-15905.
5. Tang MH, Dahlgren M, Brueffer C, Tjitrowirjo T, Winter C, Chen Y, et al. Remarkable similarities of chromosomal rearrangements between primary human breast cancers and matched distant metastases as revealed by whole-genome sequencing. Oncotarget. 2015;6(35):37169-37184.
6. Kroigard AB, Larsen MJ, Thomassen M, Kruse TA. Molecular Concordance Between Primary Breast Cancer and Matched Metastases. Breast J. 2016;22(4):420-430.
7. Demicheli R, Terenziani M, Bonadonna G. Estimate of tumor growth time for breast cancer local recurrences: rapid growth after wake-up? Breast Cancer Res Treat. 1998;51(2):133-137.
8. Uhr JW, Pantel K. Controversies in clinical cancer dormancy. Proc Natl Acad Sci U S A. 2011;108(30):12396-12400.
9. Sosa MS, Bragado P, Aguirre-Ghiso JA. Mechanisms of disseminated cancer cell dormancy: an awakening field. Nat Rev Cancer. 2014;14(9):611-622.
10. Dittmer J. Mechanisms governing metastatic dormancy in breast cancer. Semin Cancer Biol. 2017;44:72-82.
11. Selli C, Dixon JM, Sims AH. Accurate prediction of response to endocrine therapy in breast cancer patients: current and future biomarkers. Breast Cancer Res. 2016;18(1):118.
12. Clarke R, Tyson JJ, Dixon JM. Endocrine resistance in breast cancer--An overview and update. Mol Cell Endocrinol. 2015;418 Pt 3:220-234.
13. Ma CX, Reinert T, Chmielewska I, Ellis MJ. Mechanisms of aromatase inhibitor resistance. Nature Reviews Cancer. 2015;15(5):261-275.
14. Sims AH, Bartlett JMS. Approaches towards expression profiling the response to treatment. Breast Cancer Res. 2008;10(6).
15. Vendrell JA, Robertson KE, Ravel P, Bray SE, Bajard A, Purdie CA, et al. A candidate molecular signature associated with tamoxifen failure in primary breast cancer. Breast Cancer Res. 2008;10(5):R88.
16. Turnbull AK, Arthur LM, Renshaw L, Larionov AA, Kay C, Dunbier AK, et al. Accurate Prediction and Validation of Response to Endocrine Therapy in Breast Cancer. J Clin Oncol. 2015;33(20):2270-2278.
17. Dunbier AK, Ghazoui Z, Anderson H, Salter J, Nerurkar A, Osin P, et al. Molecular Profiling of Aromatase Inhibitor-Treated Postmenopausal Breast Tumors Identifies Immune-Related Correlates of Resistance. Clin Cancer Res. 2013;19(10):2775-2786.
18. Patani N, Dunbier AK, Anderson H, Ghazoui Z, Ribas R, Anderson E, et al. Differences in the transcriptional response to fulvestrant and estrogen deprivation in ER-positive breast cancer. Clin Cancer Res. 2014;20(15):3962-3973.

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 19. Arthur LM, Turnbull AK, Webber VL, Larionov AA, Renshaw L, Kay C, et al.
2 Molecular changes in lobular breast cancers in response to endocrine therapy. Cancer Res.
3 2014;74(19):5371-5376.
4 20. Turnbull AK, Kitchen RR, Larionov AA, Renshaw L, Dixon JM, Sims AH. Direct
5 integration of intensity-level data from Affymetrix and Illumina microarrays improves
6 statistical power for robust reanalysis. BMC Med Genomics. 2012;5:35.
7 21. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray.
8 Bioinformatics. 2008;24(13):1547-1548.
9 22. Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway
10 analysis and visualization. Mol Biosyst. 2016;12(2):477-479.
11 23. Hong FX, Breitling R, McEntee CW, Wittner BS, Nemhauser JL, Chory J. RankProd:
12 a bioconductor package for detecting differentially expressed genes in meta-analysis.
13 Bioinformatics. 2006;22(22):2825-2827.
14 24. Likhite N, Warawdekar UM. A unique method for isolation and solubilization of
15 proteins after extraction of RNA from tumor tissue using trizol. J Biomol Tech.
16 2011;22(1):37-44.
17 25. Coleman O, Henry M, Clynes M, Meleady P. Filter-Aided Sample Preparation
18 (FASP) for Improved Proteome Analysis of Recombinant Chinese Hamster Ovary Cells.
19 Methods Mol Biol. 2017;1603:187-194.
20 26. Rappsilber J, Ishihama Y, Mann M. Stop and go extraction tips for matrix-assisted
21 laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in
22 proteomics. Anal Chem. 2003;75(3):663-670.
23 27. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized
24 p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol.
25 2008;26(12):1367-1372.
26 28. Vizcaino JA, Csordas A, Del-Toro N, Dianes JA, Griss J, Lavidas I, et al. 2016 update
27 of the PRIDE database and its related tools. Nucleic Acids Res. 2016;44(22):11033.
28 29. Haffner MC, Chaux A, Meeker AK, Esopi DM, Gerber J, Pellakuru LG, et al. Global
29 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell
30 compartments and in human cancers. Oncotarget. 2011;2(8):627-637.
31 30. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic
32 RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature.
33 2009;462(7269):108-U122.
34 31. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, et al. The
35 connectivity map: Using gene-expression signatures to connect small molecules, genes,
36 and disease. Science. 2006;313(5795):1929-1935.
37 32. Jang ER, Lim SJ, Lee ES, Jeong G, Kim TY, Bang YJ, et al. The histone deacetylase
38 inhibitor trichostatin A sensitizes estrogen receptor alpha-negative breast cancer cells to
39 tamoxifen. Oncogene. 2004;23(9):1724-1736.
40 33. Cao B, Li J, Zhu J, Shen M, Han K, Zhang Z, et al. The antiparasitic clioquinol
41 induces apoptosis in leukemia and myeloma cells by inhibiting histone deacetylase
42 activity. J Biol Chem. 2013;288(47):34181-34189.
43 34. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive
44 and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics.
45 2013;14:128.
46 35. Ellis MJ, Suman VJ, Hoog J, Goncalves R, Sanati S, Creighton CJ, et al. Ki67
47 Proliferation Index as a Tool for Chemotherapy Decisions During and After Neoadjuvant
48 Aromatase Inhibitor Treatment of Breast Cancer: Results From the American College of
49 Surgeons Oncology Group Z1031 Trial (Alliance). J Clin Oncol. 2017;35(10):1061-1069.

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

36. Jankowitz RC, Oesterreich S, Lee AV, Davidson NE. New Strategies in Metastatic Hormone Receptor-Positive Breast Cancer: Searching for Biomarkers to Tailor Endocrine and Other Targeted Therapies. Clin Cancer Res. 2017;23(5):1126-1131.
37. Miller WR, Larionov A. Changes in expression of oestrogen regulated and proliferation genes with neoadjuvant treatment highlight heterogeneity of clinical resistance to the aromatase inhibitor, letrozole. Breast Cancer Res. 2010;12(4):R52.
38. Zhang XHF, Giuliano M, Trivedi MV, Schiff R, Osborne CK. Metastasis Dormancy in Estrogen Receptor-Positive Breast Cancer. Clin Cancer Res. 2013;19(23):6389-6397.
39. Bartosh TJ. Cancer cell cannibalism and the SASP: Ripples in the murky waters of tumor dormancy. Mol Cell Oncol. 2017;4(1):e1263715.
40. Barkan D, Green JE, Chambers AF. Extracellular matrix: a gatekeeper in the transition from dormancy to metastatic growth. Eur J Cancer. 2010;46(7):1181-1188.
41. Wang G, Platt-Higgins A, Carroll J, de Silva Rudland S, Winstanley J, Barraclough R, et al. Induction of metastasis by S100P in a rat mammary model and its association with poor survival of breast cancer patients. Cancer Res. 2006;66(2):1199-1207.
42. Yang RX, Stocker S, Schott S, Heil J, Marme F, Cuk K, et al. The association between breast cancer and S100P methylation in peripheral blood by multicenter case-control studies. Carcinogenesis. 2017;38(3):312-320.
43. Peng C, Chen H, Wallwiener M, Modugno C, Cuk K, Madhavan D, et al. Plasma S100P level as a novel prognostic marker of metastatic breast cancer. Breast Cancer Res Treat. 2016;157(2):329-338.
44. Dakhel S, Padilla L, Adan J, Masa M, Martinez JM, Roque L, et al. S100P antibody-mediated therapy as a new promising strategy for the treatment of pancreatic cancer. Oncogenesis. 2014;3:e92.
45. Basse C, Arock M. The increasing roles of epigenetics in breast cancer: Implications for pathogenicity, biomarkers, prevention and treatment. Int J Cancer. 2015;137(12):2785-2794.
46. Crea F, Nur Saidy NR, Collins CC, Wang Y. The epigenetic/noncoding origin of tumor dormancy. Trends Mol Med. 2015;21(4):206-211.
47. Fang F, Turcan S, Rimner A, Kaufman A, Giri D, Morris LG, et al. Breast cancer methylomes establish an epigenomic foundation for metastasis. Sci Transl Med. 2011;3(75):75ra25.
48. Tsang JS, Vencken S, Sharaf O, Leen E, Kay EW, McNamara DA, et al. Global DNA methylation is altered by neoadjuvant chemoradiotherapy in rectal cancer and may predict response to treatment - A pilot study. Ejso. 2014;40(11):1459-1466.
49. De Smet C, Loriot A. DNA hypomethylation in cancer Epigenetic scars of a neoplastic journey. Epigenetics. 2010;5(3).
50. Pathiraja TN, Xi Y, Lee AV, Santen R, Gannon F, Kaipparettu B, et al. Estrogen Deprivation Results in Altered DNA Methylation Profile in Breast Cancer Cells - Role in Endocrine Resistance? Cancer Res. 2009;69(24):808s-808s.
51. Harris ME, Bohni R, Schneiderman MH, Ramamurthy L, Schumperli D, Marzluff WF. Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. Mol Cell Biol. 1991;11(5):2416-2424.
52. Braunstein M, Liao L, Lyttle N, Lobo N, Taylor KJ, Krzyzanowski PM, et al. Downregulation of histone H2A and H2B pathways is associated with anthracycline sensitivity in breast cancer. Breast Cancer Res. 2016;18(1):16.
53. Zhu Z, Edwards RJ, Boobis AR. Increased expression of histone proteins during estrogen-mediated cell proliferation. Environ Health Perspect. 2009;117(6):928-934.

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

54. Sarkar S, Abujamra AL, Loew JE, Forman LW, Perrine SP, Faller DV. Histone Deacetylase Inhibitors Reverse CpG Methylation by Regulating DNMT1 through ERK Signaling. Anticancer Res. 2011;31(9):2723-2732.
55. Lee JH, Choy ML, Ngo L, Foster SS, Marks PA. Histone deacetylase inhibitor induces DNA damage, which normal but not transformed cells can repair. Proc Natl Acad Sci U S A. 2010;107(33):14639-14644.
56. Bolden JE, Shi W, Jankowski K, Kan CY, Cluse L, Martin BP, et al. HDAC inhibitors induce tumor-cell-selective pro-apoptotic transcriptional responses. Cell Death Dis. 2013;4:e519.
57. Ceccacci E, Minucci S. Inhibition of histone deacetylases in cancer therapy: lessons from leukaemia. Br J Cancer. 2016;114(6):605-611.
58. Masri S, Lui K, Phung S, Ye J, Zhou D, Wang X, et al. Characterization of the weak estrogen receptor alpha agonistic activity of exemestane. Breast Cancer Res Treat. 2009;116(3):461-470.
59. Aguilar H, Sole X, Bonifaci N, Serra-Musach J, Islam A, Lopez-Bigas N, et al. Biological reprogramming in acquired resistance to endocrine therapy of breast cancer. Oncogene. 2010;29(45):6071-6083.
60. Milosevic J, Klinge J, Borg AL, Foukakis T, Bergh J, Tobin NP. Clinical instability of breast cancer markers is reflected in long-term in vitro estrogen deprivation studies. BMC Cancer. 2013;13.
61. Simigdala N, Gao Q, Pancholi S, Roberg-Larsen H, Zvelebil M, Ribas R, et al. Cholesterol biosynthesis pathway as a novel mechanism of resistance to estrogen deprivation in estrogen receptor-positive breast cancer. Breast Cancer Res. 2016;18.
62. Coser KR, Wittner BS, Rosenthal NF, Collins SC, Melas A, Smith SL, et al. Antiestrogen-resistant subclones of MCF-7 human breast cancer cells are derived from a common monoclonal drug-resistant progenitor. Proc Natl Acad Sci U S A. 2009;106(34):14536-14541.
63. Alves CL, Elias D, Lyng M, Bak M, Kirkegaard T, Lykkesfeldt AE, et al. High CDK6 Protects Cells from Fulvestrant-Mediated Apoptosis and is a Predictor of Resistance to Fulvestrant in Estrogen Receptor-Positive Metastatic Breast Cancer. Clin Cancer Res. 2016;22(22):5514-5526.

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 **Fig. 1** Long-term endocrine therapy as a clinical model to investigate breast cancer dormancy
2 and acquired resistance. **a** Extended (4-45 months) letrozole treatment was exploited as a
3 clinical model of breast cancer dormancy and acquired resistance. Sequential clinical samples
4 from the same patient with no surgery and extended treatment were used to model clinical
5 breast cancer dormancy and resistance. Before (pre, ≤0 days), early-on (early, 13-120 days)
6 and long-term (long, >120 days) neoadjuvant letrozole treatment. **b** Dynamic change in
7 tumour size by USS and mean expression of proliferation markers MKI67, PCNA and
8 MCM2 were used to classify patients into two categories, dormant (blue) and resistant (red).
9 Overall comparisons of classifications per patient based on USS and mean change in
10 proliferation markers with final classification are shown. **c** The duration of letrozole
11 treatment (days) for samples, each bar represents a sample. **d** Intrinsic subtype classification
12 by PAM50 at each timepoint. **e** Kaplan-Meier plot showing overall survival probability in
13 dormant vs resistant patients (log-rank test). **f** Density plot showing the distribution of time to
14 recurrence in dormant and resistant patients.

15 **Fig. 2** Distinct transcriptomic changes during long-term aromatase inhibitor treatment. **a**
16 Unsupervised hierarchical clustering with most variant 500 features. **b** Multidimensional
17 scaling (MDS) plot using most variant 500 genes across all timepoints. Each dot corresponds
18 to a sample and sizes represent the duration of treatment. **c** Intra-patient (comparison of
19 samples from the same patient) correlations are shown. Dormant (blue); resistant (red);
20 Before (pre, ≤0 days), early-on (early, 13-120 days) and long-term (long, >120 days)
21 neoadjuvant letrozole treatment; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

22 **Fig. 3** Long-term endocrine treatment is associated with cell cycle, senescence, epigenetic
23 regulation and ECM-associated pathways. Differentially expressed genes between pre-
24 treatment and long-term treated samples of dormant (**a**) and resistant (**b**) patients were
25 determined. Heat-maps showing change in down- and up-regulated genes' expression in

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

dormant (**a**) and resistant (**b**) samples. Each column represents a sample and each row a gene. Colours are log2 mean-centred values with red indicating high values and blue indicating low expression. Bar plots on top of heat-maps represent the time on treatment (days) for each sample. **c-d** Graphs show dynamic changes in mean expression of differentially expressed genes in response classes. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Fig. 4 Comparative analysis of dormant and resistant tumours. **a** Volcano plot showing differentially expressed genes between long-term treated dormant and resistant tumours (dormancy vs resistance genes). Some up- and down-regulated genes in resistant tumours are highlighted in red and blue, respectively. **b** Significantly enriched pathways for dormancy vs resistance genes ($p < 0.01$; ReactomePA). **c** Heatmap showing partial separation of long-term treated dormant and resistant samples using dormancy vs resistance genes. Colours are log2 mean-centred values with red indicating high and blue indicating low expression. Genes are sorted by FC values from most to least up/down-regulated. Samples are sorted by sum expression of up-regulated genes. **d** Comparison of ssGSEA scores of dormancy vs resistance up-regulated genes between dormant and acquired resistant tumours. Dormant (blue); resistant (red); Before (pre, ≤0 days), early-on (early, 13-120 days) and long-term (long, >120 days) neoadjuvant letrozole treatment; *** $p < 0.001$; * $p < 0.05$.

Fig. 5 Validation of results using *in vitro* gene expression data from resistant cell lines and proteomics analysis. **a** Normalised enrichment scores of differently up-regulated genes calculated using single sample gene set enrichment analysis (ssGSEA) in aromatase inhibitor -resistant cells. Scores were significantly higher (** $p < 0.01$, *** $p < 0.001$) in two aromatase inhibitor-resistant cell lines, MCF7:2A and MCF7:5C, which were clonally derived from MCF7 breast cancer cells following long-term estrogen deprivation (LTED) compared to control/sensitive MCF7 cells (n = 4). Anastrozole-resistant (Ana_R) and exemestane-resistant (Exe_R) MCF7aro cells had significantly higher scores compared to

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 control (n = 3). **b** Dynamic changes in enrichment scores of LTED MCF7 cells in three
2 different datasets. **c** Scores in tamoxifen-resistant (Tam_R) and fulvestrant-resistant (Fulv_R)
3 and drug-sensitive (control) MCF7 cells (n = 4, n = 10; sc: subclone). **d** Volcano plot
4 showing differentially expressed proteins between long-term treated dormant and resistant
5 tumours ($p < 0.05$). Some overlapping features between transcriptomics and proteomics
6 analysis and the most up- and down- proteins are highlighted in red and blue, respectively

7 **Fig. 6** Immunohistochemical evaluation of **a** 5-methylcytosine (5-mC) and **b** 5-
8 hydroxymethylcytosine (5-hmC) in FFPE sections from letrozole-treated samples.
9 Representative images in dormant and resistant tumours are shown. Boxplots show
10 distributions of semi-quantitative intensity scores of 5-mC (* $p < 0.05$, *** $p < 0.001$, n = 5-
11 12) and 5-hmC (*** $p < 0.001$, n = 5-13) levels in dormant and resistant patients. Early-on
12 (early, 13-120 days) and long-term (long, >120 days) neoadjuvant letrozole treatment.

13

bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 **Additional files**

2 **Additional file 1: FigureS1.** Consort diagram showing the cohort and sample sizes. Patient
3 and sample sizes in each group are shown with inclusion and exclusion criteria. (JPG 81 kb)

4 **Additional file 2: TableS1-S7.** Tables for lists of differentially expressed genes and proteins,
5 and enriched pathways. **Table S1**. Differentially expressed genes between long-term treated
6 and pre-treatment samples of dormant patients. **Table S2.** Enriched pathways for
7 differentially expressed genes between long-term treated and pre-treatment samples of
8 dormant patients. **Table S3.** Differentially expressed genes between long-term treated and
9 pre-treatment samples of acquired resistant patients. **Table S4.** Enriched pathways for
10 differentially expressed genes between long-term treated and pre-treatment samples of
11 acquired resistant patients. **Table S5.** Differentially expressed genes between long-term
12 treated dormant and resistant tumours. **Table S6.** Enriched pathways for differentially
13 expressed genes between long-term treated dormant and resistant tumours. **Table S7.**
14 Differentially expressed proteins between long-term treated dormant and resistant tumours.
15 (XLSX 283 kb)

16

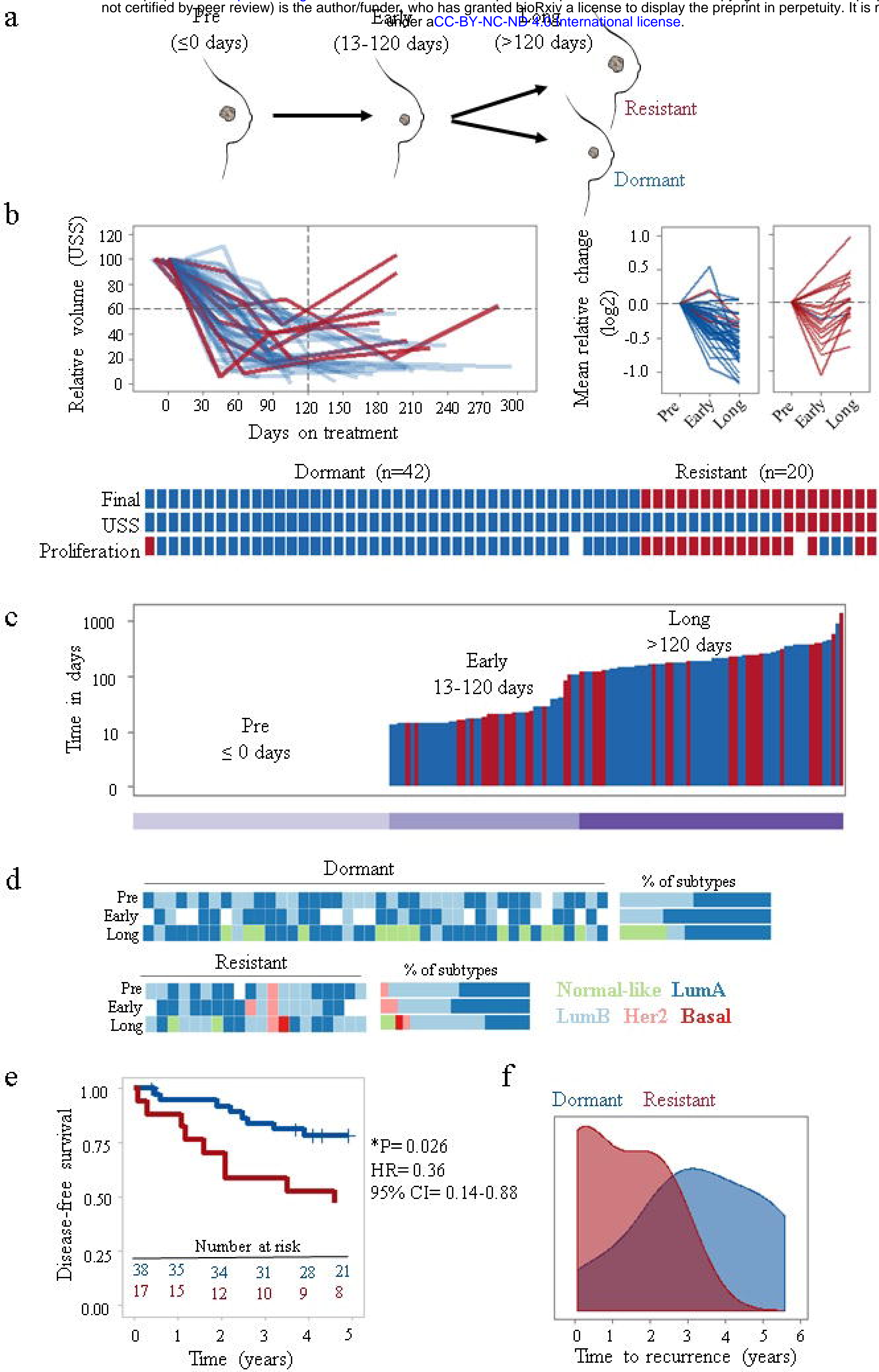
bioRxiv preprint doi: https://doi.org/10.1101/426155; this version posted September 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 **Table 1.** Patient characteristics.

	Dormant n (%)	**Resistant** n (%)	**Total** n	***p* value**[*]
Total no of patients	42	20	62	
Total no of samples	111	56	167	
Age at diagnosis				
Mean	75	72		
Range	53-87	56-89		
Tumour grade				0.39
1	6 (14.3)	1 (5.0)	7	
2	27 (64.3)	10 (50.0)	37	
3	8 (19.0)	6 (30.0)	14	
NA	1 (2.4)	3 (15.0)	4	
Tumour size				0.71
T1	5 (11.9)	4 (20.0)	9	
T2	19 (45.2)	9 (45.0)	28	
T3	2 (4.8)	2 (10.0)	4	
T4	11 (26.2)	4 (20.0)	15	
NA	5 (11.9)	1 (5.0)	6	
Nodal status				0.36
N0	27 (64.3)	11 (55.0)	38	
N1	8 (19.0)	7 (35.0)	15	
N2	1 (2.4)	1 (5.0)	2	
N3	1 (2.4)	0	1	
NX	1 (2.4)	0	1	
NA	4 (9.5)	1 (5.0)	5	
Metastasis status				1.00
M0	34 (80.9)	18 (90.0)	56	
M1	2 (4.8)	0	2	
MX	1 (2.4)	1 (5.0)	2	
NA	5 (11.9)	1 (5.0)	6	
ER score (Allred)				0.18
6	1 (2.4)	0	1	
7	6 (14.3)	6 (30.0)	12	
8	35 (83.3)	14 (70.0)	49	
HER status				0.69
Negative	35 (83.3)	12 (60.0)	47	
Positive	6 (14.3)	3 (15.0)	9	
NA	1 (2.4)	5 (25.0)	6	
Molecular subtype[#]				1.00
Luminal A	21 (50.0)	9 (45.0)	30	
Luminal B	20 (47.6)	9 (45.0)	29	
HER2 enriched	0	1 (5.0)	1	
Basal-like	0	0	0	
Normal-like	0	0	0	
NA	1 (2.4)	1 (5.0)	2	

2 [*]Fisher exact test ($p < 0.05$ two-tailed) [#]At diagnosis by PAM50 (genefu) NA: not available

not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



a

All samples

Long-term treated samples

12/11 8/31

Dormant Resistant

Pre Early Long

ER6 ER7 ER8

HER2+ HER2- NA

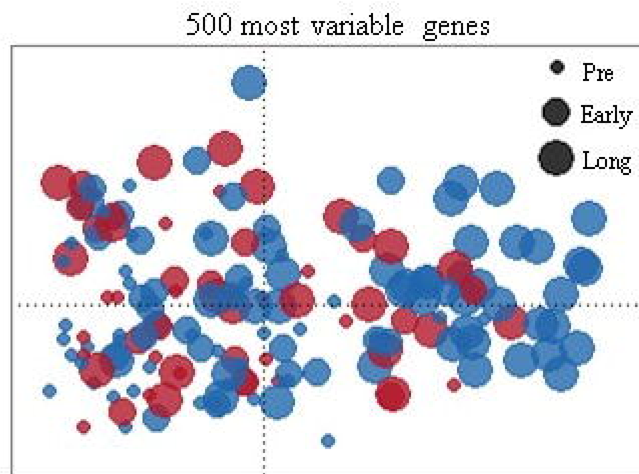
Rec+ Rec- NA

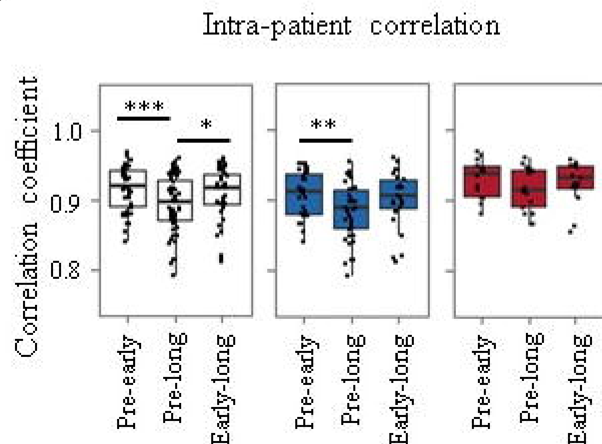
500 most variable genes

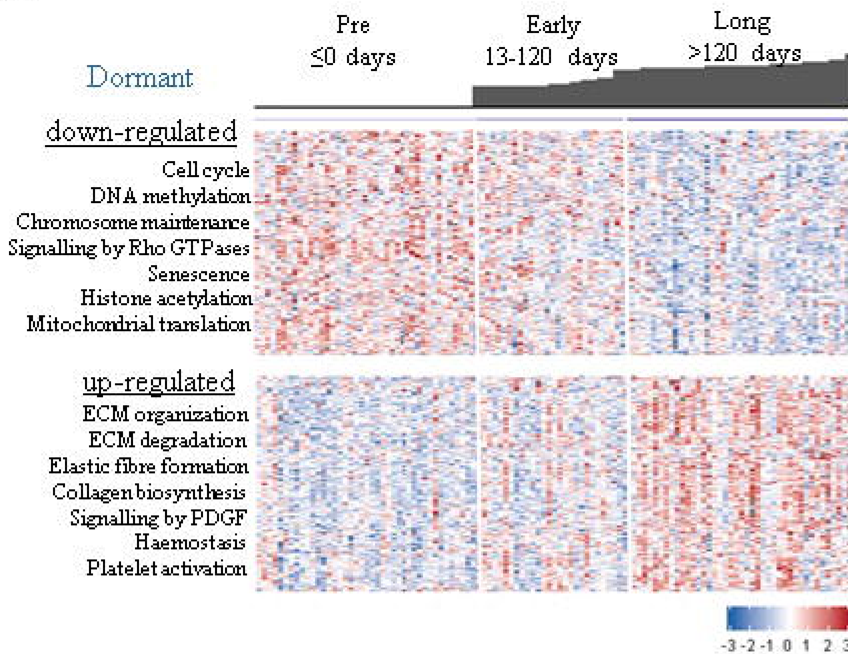
AURKA
TOP2A
HIST1H3C
HIST2H3C
MMP9
CDH11
COL3A1
COL11A1
TGFBR2
MMP2
IGF1
CD36
ADH1A
ADH1C

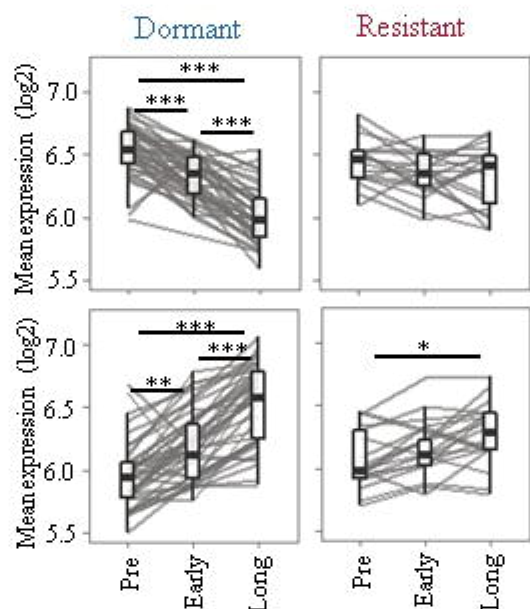
500 most variable genes

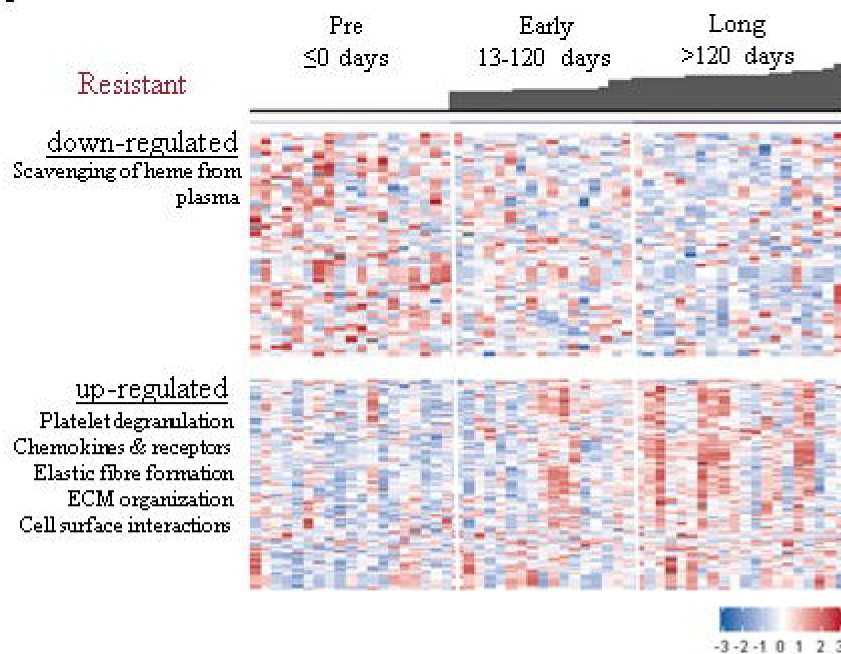
EPCAM
CD24
AGR2
MYB
BMP7
ESR1
CYP2A6
MMP2
IGF1
DUSP1
IL6
MMP2
IGF1
CDH2

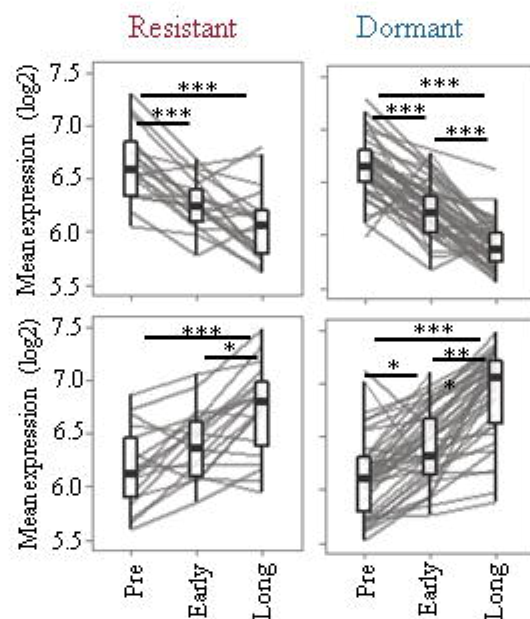


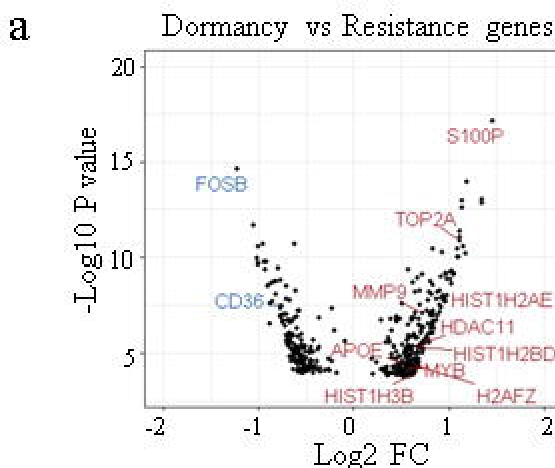


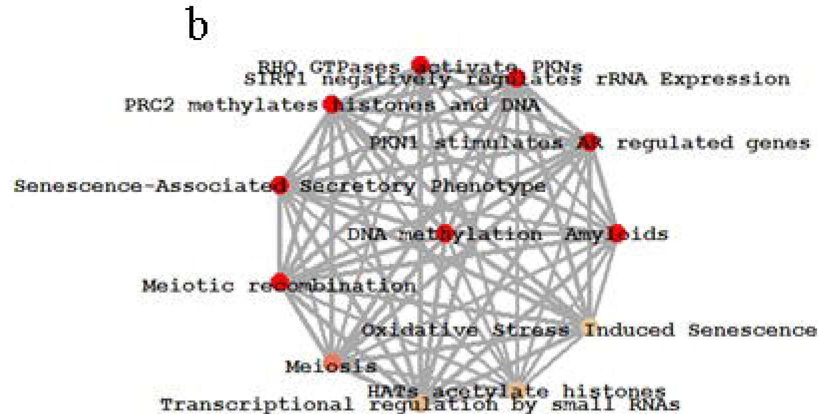

a
Pre
≤0 days
Early
13-120 days
Long
>120 days
Dormant
down-regulated
Cell cycle
DNA methylation
Chromosome maintenance
Signalling by Rho GTPases
Senescence
Histone acetylation
Mitochondrial translation
up-regulated
ECM organization
ECM degradation
Elastic fibre formation
Collagen biosynthesis
Signalling by PDGF
Haemostasis
Platelet activation
-3 -2 -1 0 1 2 3

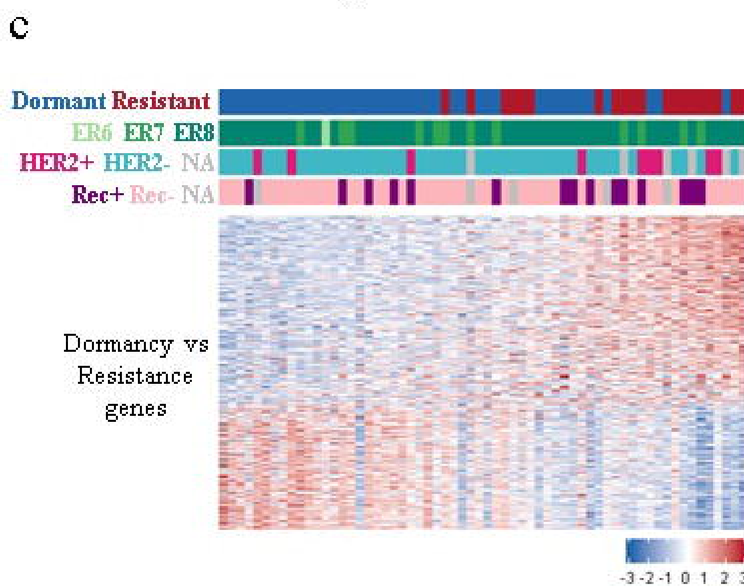

c
Dormant
Resistant
Mean expression (log2)
Pre
Early
Long

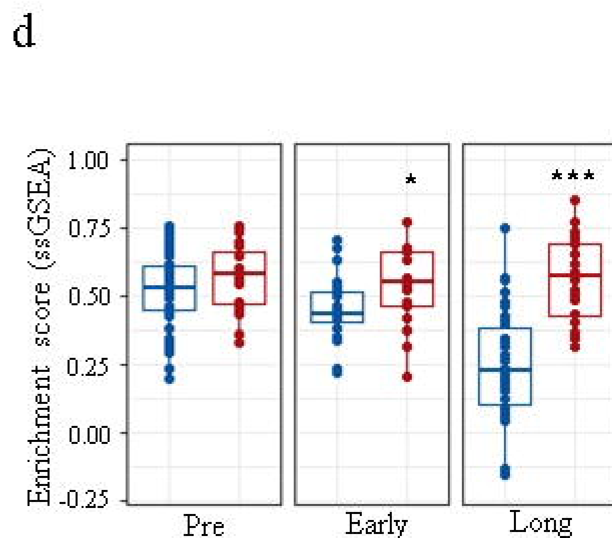

b
Pre
≤0 days
Early
13-120 days
Long
>120 days
Resistant
down-regulated
Scavenging of heme from plasma
up-regulated
Platelet degranulation
Chemokines & receptors
Elastic fibre formation
ECM organization
Cell surface interactions
-3 -2 -1 0 1 2 3

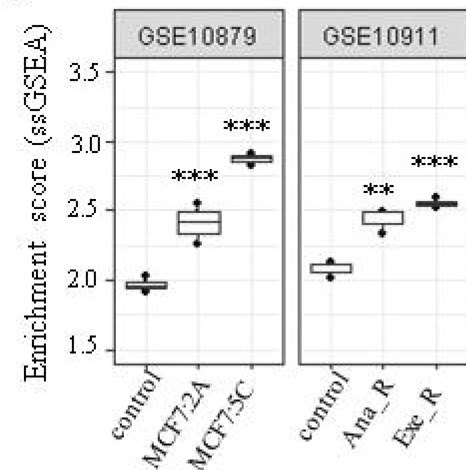

d
Resistant
Dormant
Mean expression (log2)
Pre
Early
Long

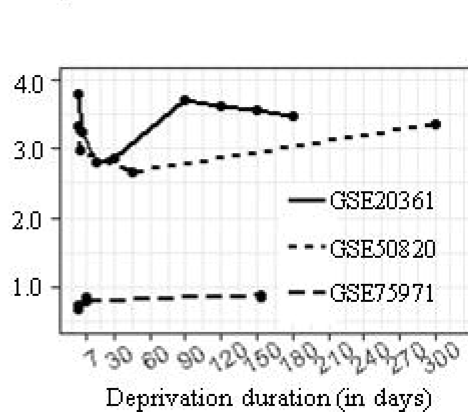

a
Dormancy vs Resistance genes
-Log10 P value
Log2 FC
S100P
FOSB
TOP2A
CD36
MMP9
HIST1H2AE
HDAC11
APOE
HIST1H2BD
MYB
HIST1H3B
H2AFZ

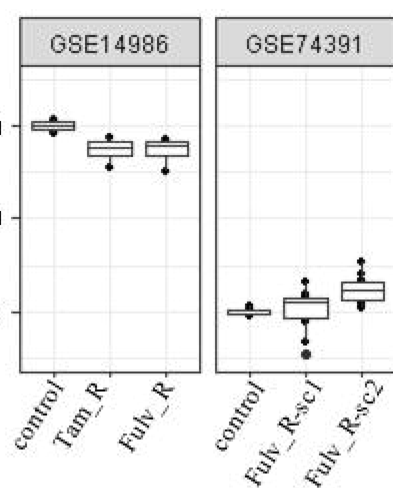

b
RHO GTPases activate PKNs
SIRT1 negatively regulates rRNA Expression
PRC2 methylates histones and DNA
PKN1 stimulates AR regulated genes
Senescence-Associated Secretory Phenotype
DNA methylation
Amyloids
Meiotic recombination
Oxidative Stress Induced Senescence
Meiosis
HATs acetylate histones
Transcriptional regulation by small RNAs

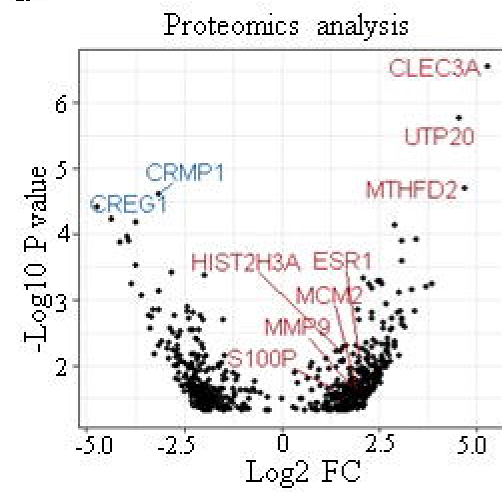

c
Dormant Resistant
ER6 ER7 ER8
HER2+ HER2- NA
Rec+ Rec- NA
Dormancy vs Resistance genes
-3 -2 -1 0 1 2 3


d
Enrichment score (ssGSEA)
*
***
Pre
Early
Long

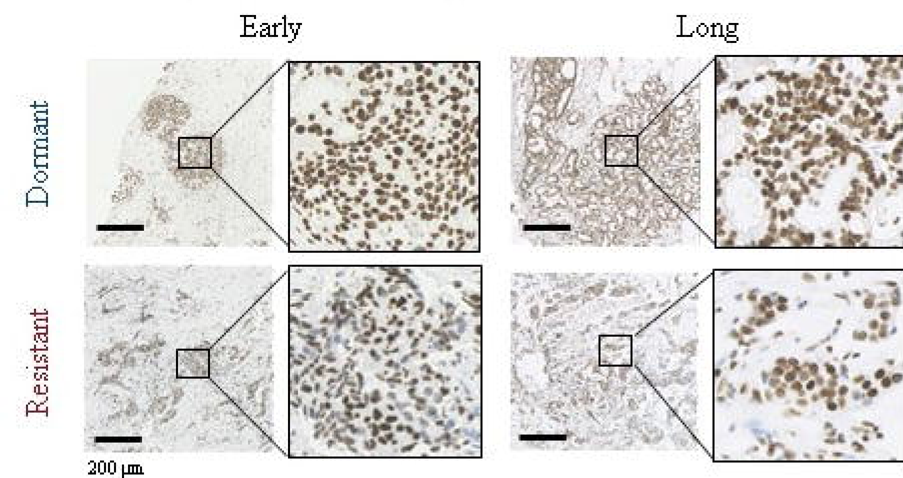

a
GSE10879
GSE10911
Enrichment score (ssGSEA)
3.5
3.0
2.5
2.0
1.5
***
***
**
***
control
MCF7:2A
MCF7:5C
control
Ana_R
Exe_R

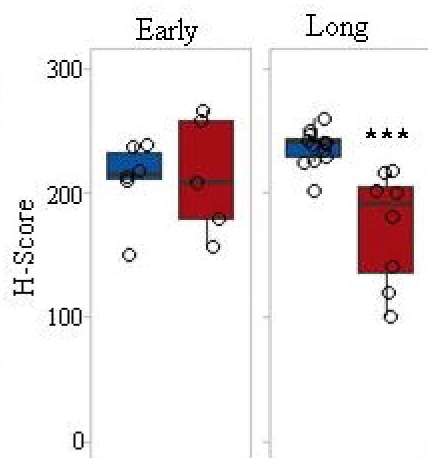

b
4.0
3.0
2.0
1.0
GSE20361
GSE50820
GSE75971
7 30 60 90 120 150 180 210 240 270 300
Deprivation duration (in days)


c
GSE14986
GSE74391
5.0
4.0
3.0
control
Tam_R
Fulv_R
control
Fulv_R-sc1
Fulv_R-sc2


d
Proteomics analysis
-Log10 P value
6
5
4
3
2
CLEC3A
UTP20
MTHFD2
CRMP1
CREG1
HIST2H3A
ESR1
MCM2
MMP9
S100P
-5.0
-2.5
0
2.5
5.0
Log2 FC

a 5-methylcytosine (5-mC)





b 5-hydroxymethylcytosine (5-hmC)