


# Proteomics on malignant pleural effusions reveals ER $\alpha$ loss in metastatic breast cancer associates with SGK1–NDRG1 deregulation

Isabel Mayayo-Peralta<sup>1</sup>, Donna O. Debets<sup>2</sup>, Stefan Prekovic<sup>1</sup>, Karianne Schuurman<sup>1</sup>, Suzanne Beerthuijzen<sup>1</sup>, Mathilde Almekinders<sup>3</sup>, Joyce Sanders<sup>3</sup>, Cathy B. Moelans<sup>4</sup>, Sandra Saleiro<sup>5</sup>, Jelle Wesseling<sup>3,6,7</sup>, Paul J. van Diest<sup>4</sup>, Rui Henrique<sup>8,9,10</sup> , Carmen Jerónimo<sup>8,9,10</sup> , Maarten Altelaar<sup>2</sup>  and Wilbert Zwart<sup>1,11</sup> 

1 Division of Oncogenomics, Oncode Institute, The Netherlands Cancer Institute, Amsterdam, The Netherlands

2 Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research, Utrecht Institute for Pharmaceutical Sciences, Utrecht University and Netherlands Proteomics Centre, The Netherlands

3 Department of Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

4 Department of Pathology, University Medical Center Utrecht, The Netherlands

5 Lung Cancer Clinics, Portuguese Oncology Institute of Porto, Portugal

6 Division of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

7 Department of Pathology, Leiden University Medical Center, The Netherlands

8 Cancer Biology and Epigenetics Group, Research Center of the Portuguese Oncology Institute-Porto, Portugal

9 Department of Pathology, Portuguese Oncology Institute of Porto, Portugal

10 Department of Pathology and Molecular Immunology, Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto, Portugal

11 Laboratory of Chemical Biology and Institute for Complex Molecular Systems, Department of Biomedical Engineering, Eindhoven University of Technology, The Netherlands

## Keywords

breast cancer metastasis; NDRG1; oestrogen receptor; proteomics; receptor conversion; SGK1

## Correspondence

W. Zwart, Division of Oncogenomics, Oncode Institute, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

Fax: +31 (0) 205122029

Tel: +31 (0) 205122101

E-mail: [w.zwart@nki.nl](mailto:w.zwart@nki.nl)

and

M. Altelaar, Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research, Utrecht Institute for Pharmaceutical Sciences, Utrecht University and Netherlands Proteomics Centre, Padualaan 8, 3584 CH

Breast cancer (BCa) is a highly heterogeneous disease, with hormone receptor status being a key factor in patient prognostication and treatment decision-making. The majority of primary tumours are positive for oestrogen receptor alpha (ER $\alpha$ ), which plays a key role in tumorigenesis and disease progression, and represents the major target for treatment of BCa. However, around one-third of patients with ER $\alpha$ -positive BCa relapse and progress into the metastatic stage, with 20% of metastatic cases characterised by loss of ER $\alpha$  expression after endocrine treatment, known as ER $\alpha$ -conversion. It remains unclear whether ER $\alpha$ -converted cancers are biologically similar to bona fide ER $\alpha$ -negative disease and which signalling cascades compensate for ER $\alpha$  loss and drive tumour progression. To better understand the biological changes that occur in metastatic BCa upon ER $\alpha$  loss, we performed (phospho)proteomics analysis of 47 malignant pleural effusions derived from 37 BCa patients, comparing ER $\alpha$ -positive, ER $\alpha$ -converted and ER $\alpha$ -negative cases. Our data revealed that the loss of ER $\alpha$ -dependency in this metastatic site leads to only a partial switch to an ER $\alpha$ -negative molecular phenotype, with preservation of a luminal-like

## Abbreviations

AR, androgen receptor; BCa, breast cancer; CAA, chloroacetamide; DDA, data-dependent acquisition; DIA, data-independent acquisition; EDTA, ethylenediamine tetracetic acid; ER $\alpha$ , oestrogen receptor alpha; ER $\alpha$ -C, oestrogen receptor alpha conversion; ER $\alpha$ -N, oestrogen receptor alpha negative; ER $\alpha$ -P, oestrogen receptor alpha positive; FFPE, formalin-fixed paraffin embedded; GR, glucocorticoid receptor; H&E, haematoxylin-and-eosin; IHC, immunohistochemistry; KSEA, kinase-substrate enrichment analyses; MPE, malignant pleural effusion; NDRG1, n-myc downregulated gene 1; PCA, principle component analysis; PR, progesterone receptor; RTK, receptor tyrosine kinase; SDC, sodium deoxycholate; SGK1, serum/glucocorticoid regulated kinase 1; SHRs, steroid hormone receptors; TCEP, phosphinehydrochloride; TNBC, triple negative breast cancer.

Utrecht, The Netherlands  
E-mail: [m.altelaar@uu.nl](mailto:m.altelaar@uu.nl)

Isabel Mayayo-Peralta and Donna O. Debets  
contributed equally to this work

(Received 2 May 2023, revised 6  
September 2023, accepted 17 October  
2023, available online 2 November 2023)

doi:10.1002/1878-0261.13540

proteomic landscape. Furthermore, we found evidence for decreased activity of several key kinases, including serum/glucocorticoid regulated kinase 1 (SGK1), in ER $\alpha$ -converted metastases. Loss of SGK1 substrate phosphorylation may compensate for the loss of ER $\alpha$ -dependency in advanced disease and exposes a potential therapeutic vulnerability that may be exploited in treating these patients.

## 1. Introduction

Breast cancer (BCa) is a highly heterogeneous disease comprising distinct molecular subtypes and is the most commonly diagnosed cancer type in women worldwide [1,2]. Approximately 70% of BCas are positive for the expression of Oestrogen Receptor alpha (ER $\alpha$ ), which drives tumour growth and progression. Although endocrine therapies targeting the ER $\alpha$  signalling axis are highly effective, approximately 30% of BCa patients do not respond to treatment and develop metastatic disease [3]. Extensive efforts have been put into genomic profiling of advanced BCa, identifying different genetic alterations that drive endocrine-therapy resistance, including mutations in the *ESR1* gene (encoding for ER $\alpha$ ), Receptor Tyrosine Kinase (RTK) pathways, SWI/SNF complex members and other transcriptional modulators such as FOXA1 and CTCF [4–10]. Importantly, around 10–20% of advanced BCa cases are characterised by the loss of ER $\alpha$  expression, hereafter referred as ER $\alpha$ -conversion (ER $\alpha$ -C), which occurs as a consequence of prolonged exposure to endocrine therapies [4,11–13]. ER $\alpha$ -C is suggested to occur through epigenetic suppression mechanisms [14].

In contrast to extensive research on the biological features of ER $\alpha$ -positive (ER $\alpha$ -P) metastatic disease [4,11], little is currently known about the implications of ER $\alpha$ -C on disease progression and therapy response. Thus, an in-depth characterisation of the molecular landscape of ER $\alpha$ -C is needed. Due to loss of ER $\alpha$  expression, ER $\alpha$ -C tumours are no longer sensitive to ER $\alpha$ -targeting endocrine therapies and may depend on other signalling cascades to regulate cell proliferation. Therefore, these pathways may be exploited as potential new therapies for ER $\alpha$ -C metastatic disease.

To gain better insights into the proteomic alterations in cellular signalling that occur upon ER $\alpha$  loss in metastatic BCa, we performed whole proteome and phospho-proteome analyses of 47 metastatic BCa tumours obtained from 37 metastatic BCa patients, including 15 of ER $\alpha$ -P,

14 ER $\alpha$ -C and 18 ER $\alpha$ -N (triple-negative BCa cases). Strikingly, our data revealed that loss of ER $\alpha$  expression in advanced disease does not result in a complete switch to an ER $\alpha$ -N molecular phenotype, as ER $\alpha$ -C tumours retained luminal-like features. These findings highlight the importance of in-depth profiling of metastatic BCa tumours beyond subtyping by use of classical BCa biomarkers. Importantly, we found evidence for inactivation of several key kinases upon transition to ER $\alpha$ -C. Particularly, loss of SGK1 activity and phosphorylation of its substrate NDRG1 may play a role in loss of ER $\alpha$ -dependency during disease progression. These findings suggest that modulation of this pathway could be a promising strategy for future clinical interventions.

## 2. Materials and methods

### 2.1. Sample collection

Build-up fluid in the pleural cavity was collected between November 2013 and July 2017, at the Netherlands Cancer Institute and at the Portuguese Oncology Institute, including a total of 47 MPEs (15 ER $\alpha$ -P, 14 ER $\alpha$ -C, 18 ER $\alpha$ -N) from 37 patients. Samples were collected and processed immediately after drainage as described before [12]. In short, pleural fluid was centrifuged at 1600 g for 8 min, and erythrocytes were lysed, using the corresponding buffer (5 mM KHCO<sub>3</sub>, 75 mM NH<sub>4</sub>Cl, 400  $\mu$ L 500 mM EDTA, up to 500 mL MilliQ, pH: 7.4) for 10 min at room temperature. Cells were then either resuspended using 10% DMSO solution and stored at  $-80^{\circ}\text{C}$ , or fixed with formalin and paraffin-embedded for future immunohistochemistry stainings. This study was approved by the institutional review boards of the Netherlands Cancer Institute (CFMPB411) and of the Portuguese Oncology Institute of Porto, Portugal. Informed written consent was provided by all participants in the study. The conducted research adhered to all relevant guidelines and regulations. The study methodologies conformed to the standards set by the Declaration of Helsinki.

## 2.2. Tissue processing for immunohistochemistry

Immunostainings were processed on the Ventana Bechmark ULTRA autostainer (Roche) according to the manufacturer's recommendations with antibody detection using the OptiView DAB IHC Detection Kit (Roche Tissues Diagnostics). Briefly, formalin fixed paraffin embedded (FFPE) tissue sections of 3  $\mu$ m thick were cut, deparaffinised, blocked, pretreated with the TRIS-ethylenediamine tetracetic acid (EDTA)-boric acid pH 8.4 buffer Cell Conditioning 1 (Roche; CD4: 64 min, CD8: 24 min, CD20: 40 min and CD68: 24 min) and stained with monoclonal antibodies directed against CD4 (Cellmarque; #104R-16, clone SP35, 1 : 20, incubation time 32 min), CD8 (Dako; #M7103, clone C8/144B, 1 : 100, incubation time 32 min), CD20 (Dako; #M755, clone L26, 1 : 400, incubation time 16 min) or CD68 (Roche; #790-2931, clone KP1, ready to use, incubation time 32 min). Stainings for ER $\alpha$  were performed as previously described [13] An experienced pathologist assessed the percentage of tumour cells and the intensity of IHC staining in tumour cells. The scoring of the samples was done using the online platform Slidescore ([www.slidescore.com](http://www.slidescore.com)).

## 2.3. Sample preparation for (phospho) proteomics

Samples were frozen and kept at  $-80^{\circ}\text{C}$  until cell lysis. Cell lysis was performed using SDC lysis buffer (1% sodium deoxycholate (SDC), 10 mM tris(2-carboxyethyl) phosphinehydrochloride (TCEP), 40 mM chloroacetamide (CAA) and 100 mM TRIS, pH 8.0 supplemented with phosphatase inhibitors (PhosSTOP; Roche) and protease inhibitors (complete mini EDTA-free; Roche) as described previously [14]. Proteins were digested using a two-step digestion. First, proteins were digested for 2 h using LysC (Wako Chemicals Europe GmbH, Neuss, Germany) at  $37^{\circ}\text{C}$  (enzyme to protein ratio 1 : 75). Next, trypsin (Sigma-Aldrich) was added in an enzyme to protein ratio 1 : 50 and digestion was performed overnight at  $37^{\circ}\text{C}$ . Next, samples were acidified, and desalted using Oasis HLB  $\mu$ Elution (Waters, Etten-Leur, The Netherlands). Samples were aliquoted for full proteome analysis (40  $\mu$ g), in-house build library (30  $\mu$ g), phosphopeptide enrichment (200  $\mu$ g). Phosphopeptides were enriched using the AssayMAP Bravo Platform (Agilent Technologies) using Fe(III)-NTA cartridges (Agilent Technologies) as described previously [14]. Samples were dried and stored at  $-20^{\circ}\text{C}$  until LC-MS analysis.

## 2.4. Library build: high pH fractionation

A peptide-fragment library was built in-house to match against the acquired DIA spectra. Therefore, 10  $\mu$ g of each sample was pooled for the full peptide library and 20  $\mu$ g of each sample was pooled for the phosphopeptide library. Pooled samples were fractionated on a high-pH reversed-phase C18 column (Kinetex 5u Evo C18 100A,  $150 \times 2.1$  mm; Phenomenex) coupled to an Agilent 1100 series HPLC over a 50 min gradient. Fractions were concatenated to 20 fractions for the proteome library and seven fractions for the phosphoproteome library. Samples were dried down and stored at  $-20^{\circ}\text{C}$ . Full peptide library samples were read for MS analysis. The phospho-peptide library samples were enriched for phosphorylated peptides as described previously [14].

## 2.5. LC-MS/MS analysis

Two different LC-MS/MS analysis were performed. First, a data-dependent analysis (DDA) was performed on the fractionated samples to build a library. Second, sample spectra were acquired using data-independent analysis (DIA). Samples were reconstituted in 10% FA, including  $1 \times$  iRT (Biognosys) for retention time alignment. Peptides were separated prior to MS/MS analysis on an Agilent 1290 Infinity System (Agilent Technologies), directly coupled to a Q Exactive HF (ThermoFischer Scientific). Peptides were trapped on a trap column (100  $\mu$ m i.d. by 2 cm, packed with C18 resin) at a flowrate of  $5 \mu\text{L}\cdot\text{min}^{-1}$  using buffer A (0.1% FA). Peptides were subsequently pushed onto the analytical column (75  $\mu$ m i.d. by 50 cm, packed with 2.7  $\mu$ m Poroshell C18 material) and eluted by increasing buffer B (80% ACN 0.1% FA) from 13% to 44% during 120 min for the full peptide measurements. Gradient was set from 8% to 32% during 120 min for the phosphopeptides. The subsequent MS analysis differed for the library build and the sample measurement. MS settings for the DDA library were as follows: MS1 resolution was set to 60 000 with AGC target of  $3e6$  and scan range from 395 to 1005  $m/z$ . The top 15 precursors were chosen for fragmentation and MS2 measurement. MS2 resolution was set to 30 000, AGC target to  $1e5$  and max injection time to 65 ms. Normalised collision energy was set to 27. MS settings for DIA were as follows: MS1 resolution was set to 60 000, AGC target was  $3e6$  and maximum injection time was 65 milliseconds and scan range from 395 to 1005  $m/z$ . Fragment spectra were measured at a 30 000 resolution, AGC target of  $1e6$  and automatic maximum injection time. Loop count was 30 and isolation window 20  $m/z$ , ranging from 410 to 990  $m/z$ . Normalised collision energy was set to 27.

## 2.6. Data processing

DDA MS files were searched using Mascot search engine and the human Swissprot database in Proteome Discoverer (version 2.3.0.522). Trypsin was set as cleavage enzyme and up to two missed cleavages were allowed. Oxidation (M) and Acetylation (N-terminus) were set as dynamic modifications. Carbamidomethyl (C) was set as a fixed modification. FDR for peptides and PSMs was set to 1%. For the phosphopeptide search, phosphorylation (STY) was added as a dynamic modification and up to three missed cleavages were allowed. Minimum site probability for the phosphosite localisation was set to 75%.

The acquired MSF-files from the P.D. search were used to build a library in Skyline. The cut-off score was set to 0.99 and ion match tolerance to 0.05  $m/z$ . The six most intense product ions were picked (with a minimum of four transitions per precursor) and only scans within 5 min of the predicted RT were included. Peptide length was set from 7 to 36 amino acids. A minimum of two peptides per protein was set. Included structural modifications were: carbamidomethyl (C), Oxidation (M), Acetyl (N-term) and for the phospho analysis phospho (STY). Decoy peptides were generated by sequence shuffle.

Raw DIA MS files were searched against the acquired library using Skyline and subsequently mPROPHET scoring was used to discern true hits from decoys [15]. A  $q$ -value cut-off of 0.01 was used. Quantification was performed by integration of the chromatographic peak, using only high-quality features. Data were Log<sub>2</sub>-transformed and normalisation was performed by median subtraction. Missing values were imputed using missforest imputation method [16]. Heatmaps were generated in R using Euclidean distance on  $z$ -scored data. Statistical testing was performed in R. All data were plotted using GGPLOT2 (v.3.3.5) or PHEATMAP (v.1.0.12).

## 2.7. Kinase substrate enrichment analyses

Kinase substrate enrichment analyses (KSEA) were performed to learn about changes in activity of a specific kinase by interrogating differential phosphorylation of all kinases' substrates [17]. KSEAAPP (v.0.99.0) was used to perform the analyses.

## 2.8. Genetic dependency analyses

The dependency data used in this study was obtained from DepMap 21Q4 public (<https://depmap.org/portal/>). Dependency data for *NDRG1*, *RB* and *OCA-2*

genes across BCa cell lines were downloaded. All data were plotted using GGPLOT2 (v.3.3.5), GGPUBR (v.0.4.0) and PUPILLOMETRYR (v.0.0.4).

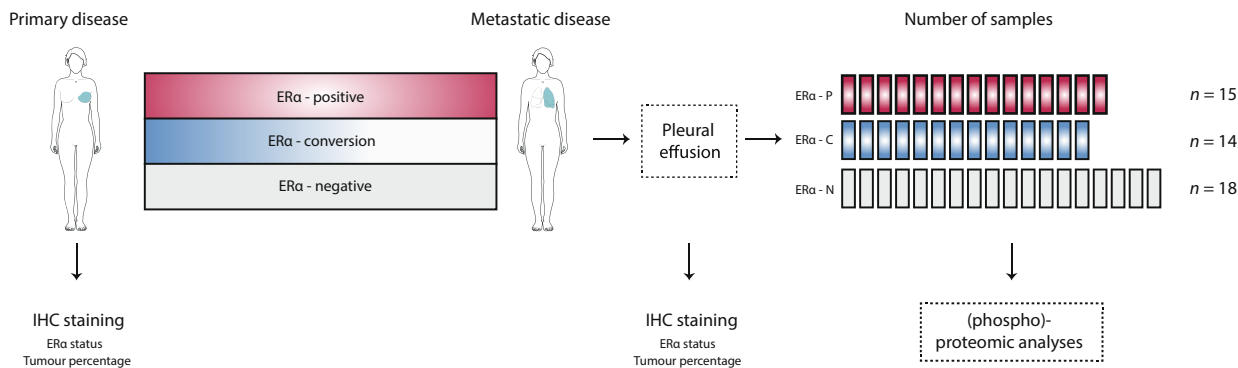
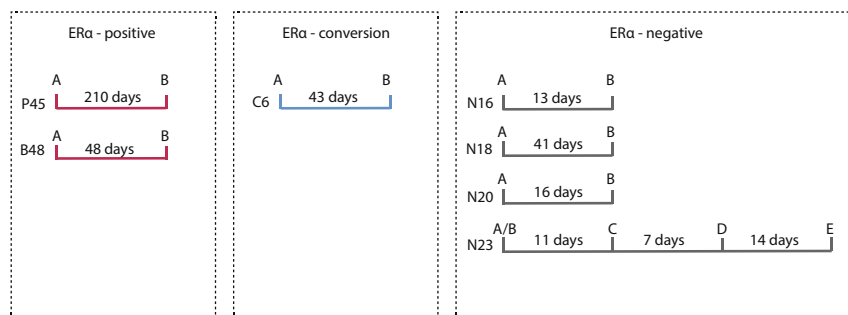
## 2.9. Gene set over-representation analyses

In order to carry out gene set over-representation analyses, we made use of the GSEA browser and used C2 curated pathways. The analysis was carried on differentially expressed proteins ( $P$ -value < 0.05, log<sub>2</sub>difference > 1.5 or < -1.5) between ER $\alpha$ -C and ER $\alpha$ -P samples.

## 3. Results

### 3.1. Design of the study: malignant pleural effusions from advanced metastatic BCa

To gain insights into BCa disease progression and endocrine treatment resistance due to loss of ER $\alpha$  expression, we performed proteomics and phosphoproteomics in a cohort of patients with metastatic BCa disease. All advanced BCa tumours included in this study metastasized to the pleural cavity, resulting in malignant pleural effusions (MPEs). As palliative care, MPEs were drained out to facilitate the patient's breathing and improve quality-of-life in this highly progressive phase of the disease. MPEs were collected by centrifugation (see Section 2) and used for this study (Fig. 1A). Patients who were included in this study are detailed in Table S1. The study comprised three distinct group of patients: ER $\alpha$ -positive (ER $\alpha$ -P) ( $n = 15$ ), ER $\alpha$ -conversion (ER $\alpha$ -C) ( $n = 14$ ) and ER $\alpha$ -negative (ER $\alpha$ -N) ( $n = 18$ ) cases. ER $\alpha$ -P and ER $\alpha$ -C were characterised by expression of ER $\alpha$  (50–100% ER $\alpha$ -positive tumour-cells) in primary disease, and by maintenance (5–100% ER $\alpha$ -positive tumour cells, average = 58%) or loss (< 1% ER $\alpha$ -positive cells) of ER $\alpha$ -expression, respectively, in metastatic disease (Table S1). In addition, as a control we used samples from ER $\alpha$ -N, which had no ER $\alpha$  expression throughout the course of the disease. In total, we collected 47 MPE samples, derived from 37 BCa patients. For seven patients, multiple MPEs were collected over a course of time (Fig. 1B). For six of these patients (two ER $\alpha$ -P, one ER $\alpha$ -C, three ER $\alpha$ -N), two tumour samples were collected, over a period ranging from 2 weeks to 10 months; these are referred to as repeat collections. For one patient (ER $\alpha$ -N), the data set contained five repeat collections, obtained over a period of 4 weeks (Fig. 1B). For all samples, Haematoxylin-and-Eosin (H&E) staining as well as Immunohistochemistry (IHC) for ER $\alpha$  were performed to enable

**(A)** Study design**(B)** Patients with multiple pleural effusions- repeat collections

**Fig. 1.** Study design and phospho-proteomics data collection. (A) Study design. Primary and metastatic breast cancer were collected and stained for tumour cell percentage and Oestrogen Receptor alpha (ER $\alpha$ ) status. Malignant pleural effusion samples were processed for total and phospho-proteomics. (B) Patients with multiple pleural effusions, constituting repeat sample collections. Seven patients had samples collected at different time points. Time elapsed between collections is depicted in days.

assessment of tumour percentage (ranging between 35% and 90% in all samples) and status of ER $\alpha$ , progesterone receptor (PR) and HER2 (Table S1). In addition, information of all treatments prescribed in the adjuvant setting, treatment at time of pleural effusion, ILC/IDC status and interval time between primary tumour and disease progression and as well from progression to date of pleural effusion, is provided (Table S1). On these well-annotated patient samples, we performed whole proteome and phospho-proteome analyses.

### 3.2. Characterisation of proteomic data and exclusion of samples with immune cell infiltration

For (phospho)proteomic analyses, MPE cells were lysed, proteins were extracted and digested. The

generated peptides were analysed by LC-MS/MS using Data-Independent Acquisition (DIA). The acquired DIA files were matched against an in-house built library. This library was generated using Data-Dependent Acquisition (DDA) of pooled and fractionated samples. For the phosphorylation analysis, phospho-enrichment was performed prior to LC-MS/MS analysis. An overview of the workflow is depicted in Fig. S1A. We obtained whole proteome and phospho-proteome coverage of all patient samples, with a total of 6440 proteins and 10 772 phosphopeptides quantified across all tumours (Fig. S1B,C).

Malignant pleural effusions are cell suspensions that contain multiple cell types, including not only metastatic BCa cells, but also reactive mesothelial cells and different types of immune cells [18]. Since the goal of this study was to gain insights into the molecular signature of metastatic tumour cells, we wanted to first



ensure that the non-tumour component of MPEs would not influence our downstream (phospho)proteomic analyses. In order to achieve this, we first performed unsupervised hierarchical clustering of all proteomic samples and found that most samples with a low tumour cell percentage cluster together, therefore suggesting similarity in their proteomic landscapes (Fig. S2A). Furthermore, principal component analysis (PCA) showed clustering of most of these same samples in the second component (Fig. S2B), indicating that the cell-type composition affected the general proteomic profile and consequently drove sample clustering. In addition, the samples with low tumour cell content displayed strong enrichment of immune cell signature proteins (ImSig database) [19], implying high immune cell infiltration (Fig. S2C,D). Based on these observations, eight samples—including three ER $\alpha$ -P (P43/P35/P50), four ER $\alpha$ -C (C5/C6B/C11/C12) and one ER $\alpha$ -N (N30)—with an immune cell dominated proteomic signal and with low tumour cell content were excluded for further analysis (Fig. S2C,D). Moreover, we performed IHC analyses for immune markers, including CD4, CD8, CD20, CD68, and observed that samples excluded from the analyses were significantly enriched for CD4-infiltrating cells (Fig. S2E). With this, we confirmed that there was an enrichment of infiltrating immune cells in the excluded samples which could have had an impact on downstream analyses of the (phospho)proteomic data sets.

### 3.3. Repeat collections from the same patient reveal patient-specific proteomic signatures that remain stable over time

We next assessed whether the proteomic profile of advanced BCa tumours changed over time during disease progression. To do this, we included repeat MPEs from the same patient that were collected over time (Fig. 1B). Interestingly, the proteomic profile of these metastatic tumours was highly reproducible; each patient displayed a patient-specific proteomic signature that remained stable over time between sample collections and differed from other individuals (Fig. 2A, Fig. S3A). Importantly, we observed that repeat collections of the same patient were more similar as compared to samples between different patients (Fig. 2A, B). Moreover, the high correlation of repeat collections of the same patient was also illustrated by tight grouping of patients following unsupervised clustering analysis of the total proteomics data set (Fig. 2C), as well as grouping in PCA space (Fig. 2D, Fig. S3B).

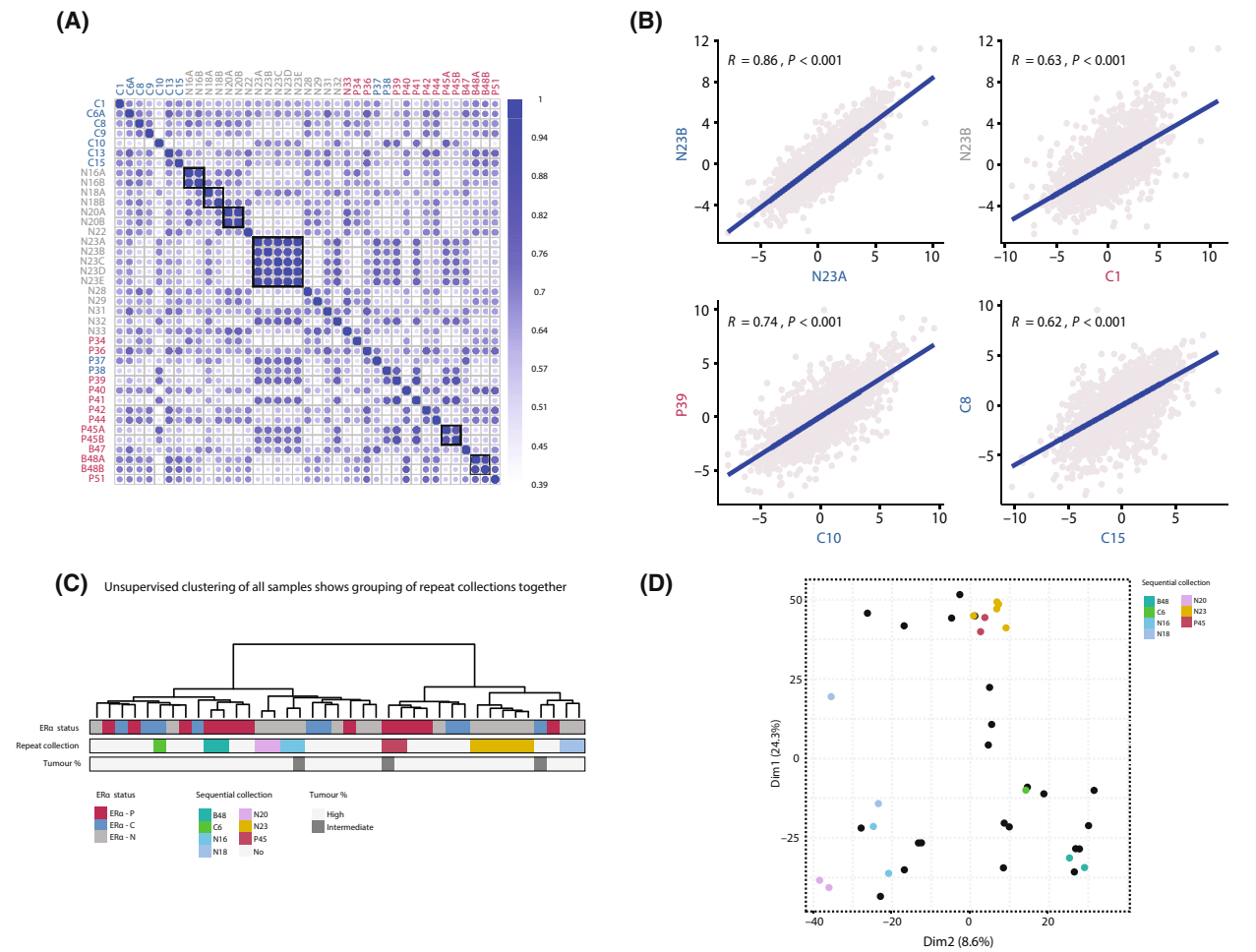
The same observations were made in the phospho-proteomic landscape (Fig. S3C). These findings suggest

that MPEs from BCa remains stable over time with a limited subclonal selection. However, repeat collections were isolated over short periods of time, suggesting that time may have not been sufficient for enrichment of different of cell populations by clonal selection under treatment pressure.

### 3.4. A subset of ER $\alpha$ -C tumours retains a proteomic signature similar to ER $\alpha$ -P tumours

We next aimed to identify proteomic differences between metastatic tumours, based on ER $\alpha$  status. First, we performed a quantitative comparison of protein expression of ER $\alpha$ -P and ER $\alpha$ -N metastatic tumours. Our analyses revealed 288 differentially expressed proteins ( $P$ -value < 0.05,  $\log_2$  difference > 1 or < -1) between the two groups (Fig. 3A). Notably, we found that proteins involved in triple-negative breast cancer (TNBC), such as GSTP1 and EPCAM [20,21], were enriched in ER $\alpha$ -N metastatic tumours, while proteins typically upregulated in ER $\alpha$ -P disease, such as ALCAM and KRT18 [22], were enriched in ER $\alpha$ -P as well as the ER $\alpha$ -C metastatic tumours (Fig. 3A,B). Interestingly, we found that the expression of those proteins in the ER $\alpha$ -C group was similar to the ER $\alpha$ -P samples, deviating from the ER $\alpha$ -N samples (Fig. 3B). These findings suggest that despite the loss of ER $\alpha$  expression, ER $\alpha$ -C metastatic tumours may have a proteomic landscape like the one of ER $\alpha$ -P metastatic tumours, and not that of ER $\alpha$ -N tumours.

The above-mentioned data indicate that on proteomic level ER $\alpha$ -P tumours resemble ER $\alpha$ -C tumours. To determine whether these observations are also consistent on a global level, we performed unsupervised clustering of differentially expressed proteins ( $P$ -value < 0.05,  $\log_2$  difference > 1.5 or < -1.5) and we observed that ER $\alpha$ -N patients clustered away from ER $\alpha$ -P/C patients (Fig. 3C). Importantly, we observed that there were two major groups of differentially expressed proteins, comprising a group of proteins mainly expressed in ER $\alpha$ -N/C patients (group 1) and a group of proteins that are more abundantly expressed in ER $\alpha$ -P/C patients (group 2) (Fig. 3C). Interestingly, by performing over-representation analyses we observed that proteins from group 1 belonged to pathways expressed in basal BCa (Fig. 3D). On the other hand, the same analyses on proteins from group 2 demonstrated enrichment of proteins contributing to luminal BCa pathways (Fig. 3D). Overall, our data suggest that a subset of ER $\alpha$ -C tumours has a similar proteomic landscape to ER $\alpha$ -P rather than ER $\alpha$ -N, even though ER $\alpha$ -C tumours have lost the expression



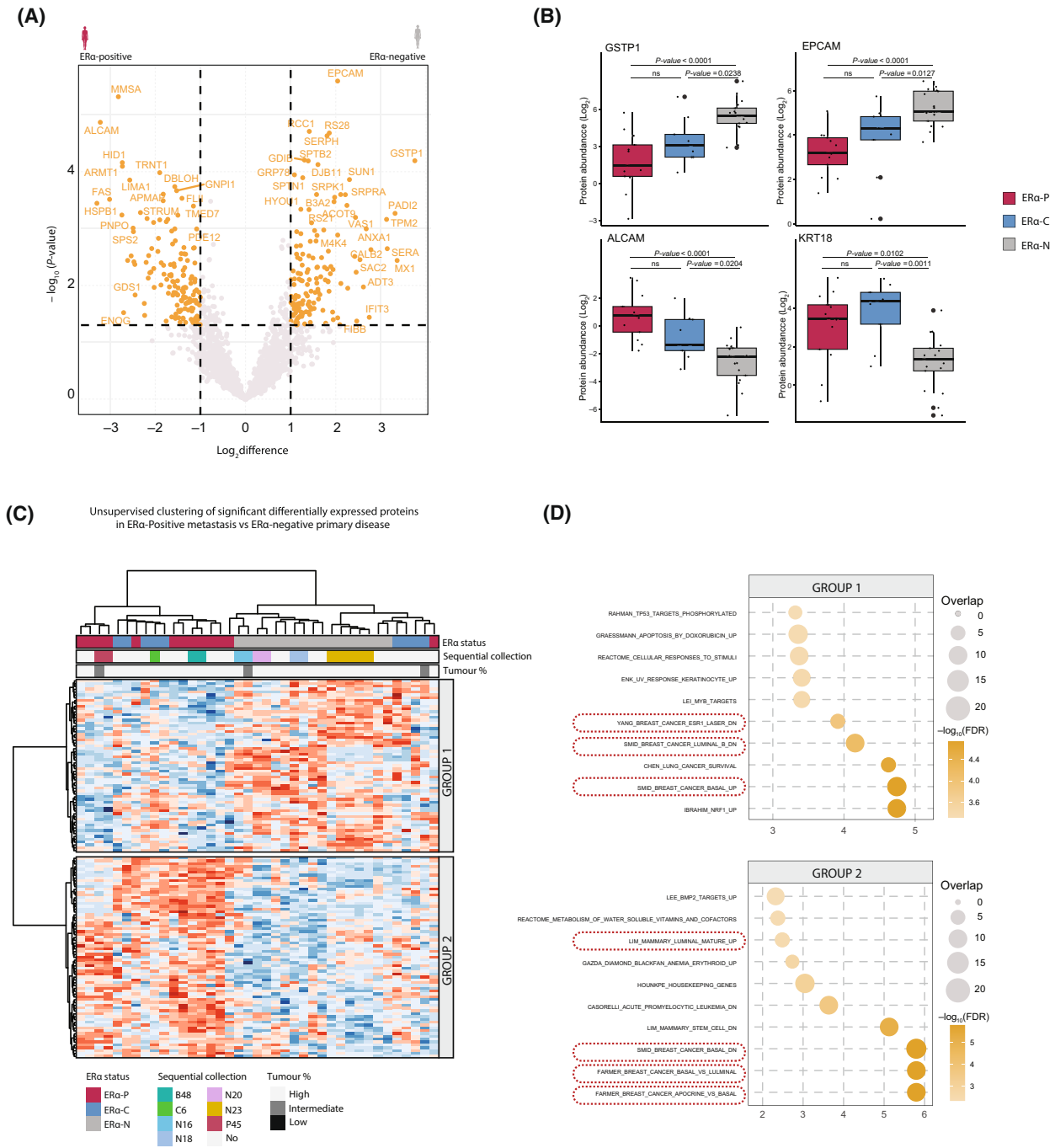
**Fig. 2.** Repeat MPE collections from the same patient revealed patient-specific proteome landscapes. (A) Correlation plot of all samples on proteomics data. Repeat collections are highlighted in the black boxes. Patient identifiers are coloured by Oestrogen Receptor alpha (ER $\alpha$ )-status. (B) Correlation plots of individual patients demonstrating a higher level of correlation for samples from the same patient, relative to inter-patient analyses. Pearson correlation is provided. (C) Unsupervised hierarchical clustering of all proteins in the data set grouped samples derived from the same patients together. The column colour bars indicate the patient group (ER $\alpha$ -C (converted), ER $\alpha$ -P (positive), ER $\alpha$ -N (negative)), whether samples are repeat collections and tumour percentage. (D) Principle Component Analysis plot, coloured by repeat collection. Samples from the same patient cluster together.

of ER $\alpha$  during the progression of the disease. These findings expand our understanding of the molecular mechanisms of BCa metastasis.

### 3.5. Decrease in kinase activity in ER $\alpha$ -converted cancers contributes to tumour progression

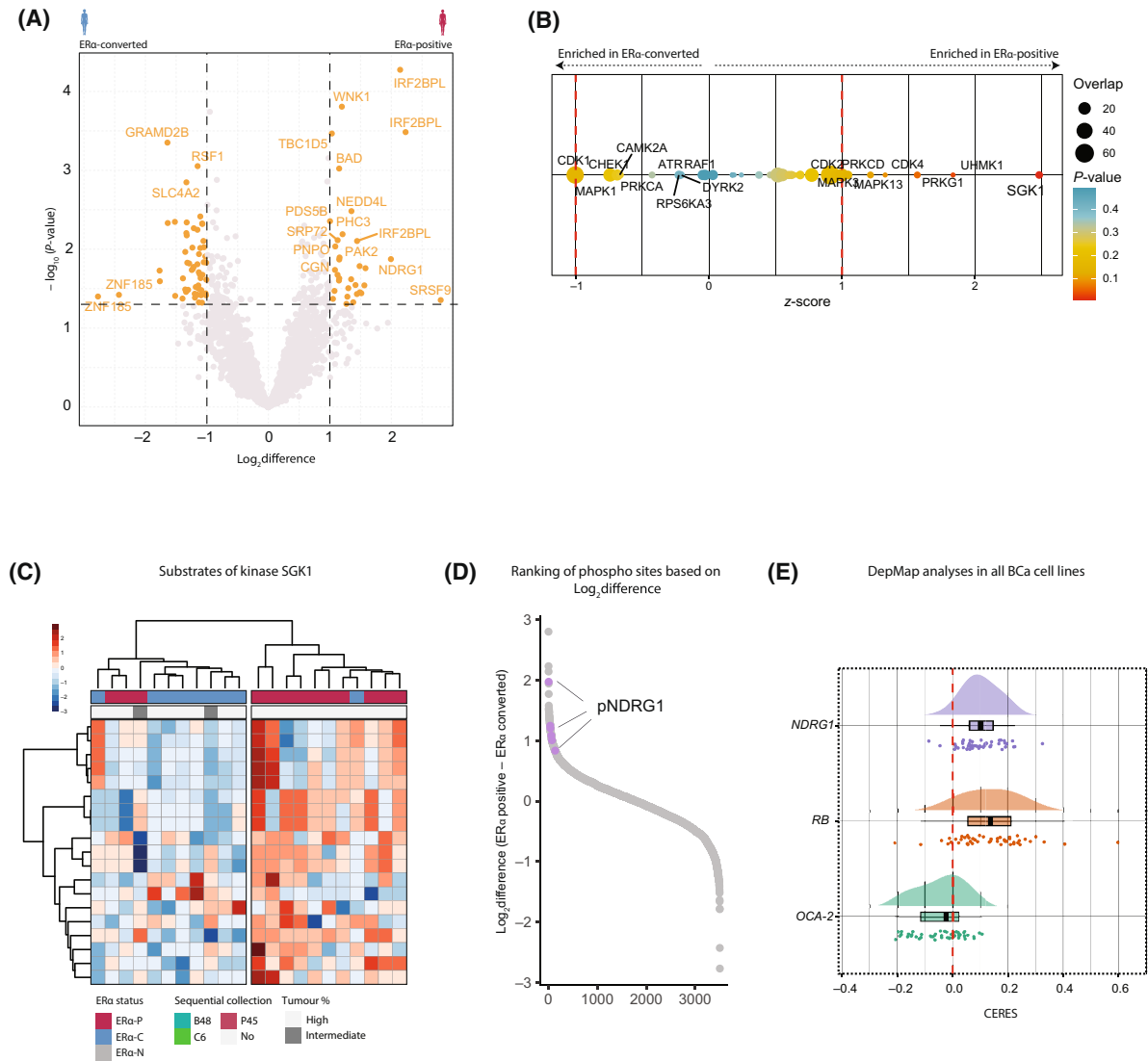
In order to understand whether alternative signalling cascades compensate for the loss of ER $\alpha$  in ER $\alpha$ -C tumours, we first compared the phospho-proteomic profiles of ER $\alpha$ -P and ER $\alpha$ -C samples and identified 86 phospho-sites differentially expressed ( $P$ -value  $< 0.05$ ,  $\log_2$ -difference  $> 1$  or  $< -1$ ) (Fig. 4A, Fig. S4A). We next wanted to further investigate the biological

consequences of changes in the phospho-proteomic landscape. To address this, we performed kinase-substrate enrichment analyses (KSEA) [17], which enabled identification of differentially activated kinases between patient groups. We observed that ER $\alpha$ -P were enriched for activity of different kinases including SGK1, UHKM1 and PRKG1 kinases (Fig. 4B). The latter suggests that pathways regulated by the aforementioned kinases were downregulated in ER $\alpha$ -C samples. As the SGK1 was the top enriched kinase, we specifically focused on interrogating differential phosphorylation on SGK1-downstream targets. For this, we performed unsupervised hierarchical clustering of ER $\alpha$ -P and ER $\alpha$ -C of all identified substrates



**Fig. 3.** ER $\alpha$ -C metastatic tumours retained ER $\alpha$ -P proteomic signature. (A) Volcano-plot of  $-\log_{10}(P\text{-value})$  and  $\log_2$  difference between ER $\alpha$ -N (negative) compared to ER $\alpha$ -P (positive) metastatic breast cancer samples. Proteins coloured in orange were considered significantly changing ( $P\text{-value} < 0.05$  and  $\log_2$  difference  $> 1$ ). (B) Boxplots of significantly changing proteins per ER $\alpha$ -status. One-way ANOVA test was used.  $P$ -values are depicted. (C) Unsupervised hierarchical clustering of significantly differentially expressed proteins ( $P\text{-value} < 0.05$  and  $\log_2$  difference  $> 1.5$ ). Unsupervised hierarchical clustering of all proteins in the data set grouped samples derived from the same patients together. The column colour bars indicate the patient group (ER $\alpha$ -C, ER $\alpha$ -P, ER $\alpha$ -N), whether samples are repeat collections and tumour percentage. (D) Gene-set overrepresentation analyses, proteins from C (group 1 and group 2) were used.





**Fig. 4.** Decrease in kinase activity in ER $\alpha$ -converted cancers contributes to tumour progression. (A) Phospho-proteomics data. Volcano-plot of  $-\log_{10}(P\text{-value})$  and  $\text{Log}_2$  difference between ER $\alpha$ -C (converted) compared to ER $\alpha$ -P (positive) metastatic breast cancer samples. Phospho-proteins coloured in orange were considered significantly changing ( $P\text{-value} < 0.05$  and  $\text{Log}_2$  difference  $> 1$ ). (B) Kinase substrate enrichment analyses (KSEA) on differentially phospho-peptides between ER $\alpha$ -C and ER $\alpha$ -P metastatic breast cancer samples. (C) Unsupervised hierarchical clustering of phosphorylation substrates of SGK1 kinase. The column colour bars indicate the patient group (ER $\alpha$ -C, ER $\alpha$ -P), and tumour percentage. (D) Phospho-peptides ranked based on  $\text{Log}_2$  difference between ER $\alpha$ -P and ER $\alpha$ -C samples. All the phospho-NDRG1 peptides are highlighted in purple. (E) DepMap analyses of NDRG1 essentiality in breast cancer cells. Expression of RB (tumour suppressor) or OCA-2 (no proliferation effects) are used as control. Essentiality scores are plotted (CERES).

of the kinase SGK1 (Fig. 4C). Interestingly, we found that ER $\alpha$ -C samples cluster away from ER $\alpha$ -P samples, exhibiting a general loss of phosphorylation for all identified SGK1 targets (Fig. 4C). Next, we ranked all differentially phosphorylated sites between ER $\alpha$ -P and ER $\alpha$ -C and observed that the phosphorylation sites of a well-described SGK1 target, NDRG1, were among the most highly differentially phosphorylated sites (Fig. 4D). Indeed, when correlating individual

patients from either ER $\alpha$ -C or ER $\alpha$ -P groups, we observed an enrichment for phosphorylation of NDRG1 in ER $\alpha$ -P samples (Fig. S4B). Downregulation of activity of the SGK1-NDRG1 signalling has been reported to be involved in tumour metastasis and invasion [23]. These data suggest that ER $\alpha$ -C samples lose activity of NDRG1, subsequently leading to downstream activation of tumour cell survival cascades. In line with this, we made use of the DepMap

database to study the essentiality of NDRG1 in proliferation of BCa cells (Fig. 4E). Importantly, we observed that depletion of NDRG1 resulted in a proliferation advantage of BCa cells, comparable to the effect observed when perturbing expression of the well-known tumour suppressor Rb1, indicating a role of NDRG1 as a tumour suppressor in BCa (Fig. 4E). Further follow-up studies should confirm the role of NDRG1 in ER $\alpha$ -C disease. Based on our findings, we propose that re-activation of the SGK1-NDRG1 axis in ER $\alpha$ -C samples may be exploited as a new therapeutic target in ER $\alpha$ -C patients.

#### 4. Discussion

Endocrine resistance poses a major clinical problem in the treatment of BCa, and substantial efforts have been made to identify players responsible for disease progression and metastasis [11]. Although acquisition of specific genomic features in endocrine resistant metastatic BCa have been extensively reported [4,9,10,24–26], including mutations of the *ESR1* gene, few studies have interrogated the implications of ER $\alpha$  loss in the metastatic setting, which accounts for around 20% of all endocrine-resistant BCa cases [4,27,28]. In this study, we aimed to gain insights into the proteomic features of advanced BCa upon loss of ER $\alpha$ -expression, to leverage discovery of potential therapeutic vulnerabilities for these tumours. To achieve this, we conducted proteomic and phospho-proteomic analyses of metastatic BCa tumours, in relation to loss of ER $\alpha$ .

First, our study demonstrates that multiple tumour collections from the same individual display highly similar proteomic profiles, indicating that these metastatic lesions likely consist of a relatively stable tumour cell population that remains largely unaltered throughout the course of the disease. It is important to note that the limitations of the technology used and the short time intervals (from 2 weeks to 7 months, see Table S1) between the collections MPEs may account for the absence of tumour subpopulations being enriched during treatment due to clonal selection under therapy pressure. In line with this, a previous study [29] used orthogonal approaches such as single-cell and bulk RNA-seq, as well as DNA-seq to track subclonal evolution in metastatic BCa patient samples. In this study, due to longer period times (2–15 years; Table S1) between MPE collection, they were able to identify subclones that became more prominent over time and investigated the efficacy of drugs targeting the enriched subclones.

We demonstrated that tumours that lose ER $\alpha$ -expression during disease progression, retained a proteomic signature more similar to ER $\alpha$ -P metastases (derived from ER $\alpha$ -P primary disease), as opposed to ER $\alpha$ -N metastases (derived from ER $\alpha$ -N primary disease). Our proteomics data further suggest that part of the ER $\alpha$ -related translational programme is still active in some of ER $\alpha$ -C tumours since ER $\alpha$ -target proteins did not diminish to levels of ER $\alpha$ -N tumours. However, ER $\alpha$ -targeting therapies would not be efficient in these patients due to the lack of ER $\alpha$  expression, which suggests that targeting downstream targets may be a potential therapeutic route. We hypothesise that the similar proteomic landscape between ER $\alpha$ -P and ER $\alpha$ -C tumours could be due to activity of other steroid hormone receptors (SHRs), such as Androgen Receptor (AR), PR, or Glucocorticoid Receptor (GR), now compensating for loss of ER $\alpha$  action. Previous research has shown that these SHRs are able to bind to similar areas in the genome as ER $\alpha$  and consequently induce similar transcriptional programmes, both *in vitro* and in patient samples [30–36]. In the absence of ER $\alpha$ -expression, these ER $\alpha$ -binding sites could be hijacked by AR, PR, or GR to maintain expression of genes contributing to tumour cell survival [30–36]. Due to the technical limitations of mass spectrometry, no peptides were detectable in relation to the mentioned SHRs. Consequently, further mechanistic investigations are imperative to formally test this hypothesis.

We identified the activity of kinase SGK1 to be enriched in ER $\alpha$ -P samples when compared to ER $\alpha$ -C samples. In line with this, we observed loss of phosphorylation of NDRG1, a known target of SGK1, in ER $\alpha$ -C samples [37]. NDRG1 has been described as a metastasis suppressor across different tumour types [23]. The phosphorylation of NDRG1 by SGK1 has been shown to be crucial to halt downstream oncogenic pathways [38]. When active, NDRG1 results in inhibition of cell proliferation, migration and metastasis [23]. In the scope of our investigation, centred around the analysis of notably aggressive samples, it is plausible to propose the hypothesis that the diminished functionality of SGK1 in ER $\alpha$ -C leads to the reduction in NDRG1 phosphorylation in a context-dependent manner. This absence, in turn, could facilitate the maintenance of cellular proliferation pathways in ER $\alpha$ -C samples, while alternative pathways might govern tumour growth in the other two groups. Through *in silico* analyses, we observed that loss of NDRG1 expression resulted in a growth advantage across a panel of BCa cell lines. Future mechanistic studies are

warranted to explore the context-specific characteristics underlying the disruption of the SGK1-NDRG1 axis. Our findings could suggest that modulation of NDRG1 expression might be a novel therapeutic avenue for ER $\alpha$ -C samples. Cellular stress as a result of iron depletion, among others, leads to the increase of NDRG1 expression [39]. Different studies used iron-binding agents to upregulate the expression of NDRG1, resulting in strong antitumour efficacy, which could be exploited for the treatment of ER $\alpha$ -C patients [23,40].

Endocrine therapy resistant metastases can grow out in various organs beyond the pleura, including solid metastasis such as the bone, brain and liver. It is important to emphasise that our observations concerning the diminished reliance on the SGK1-NDRG1 axis in ER $\alpha$ -C samples may hold relevance primarily within the context of the specific metastatic site under investigation. Consequently, further investigations encompassing samples from diverse metastatic sites are imperative to determine the universality of these resistance mechanisms in ER $\alpha$ -C samples.

The development of endocrine resistance can arise either from primary resistance or be acquired over the course of treatment. The loss ER $\alpha$  expression in relapsed breast cancers typically arises after prolonged endocrine therapy, contributing to acquired resistance. As a result, our study's findings are most likely relevant primarily to cases involving acquired resistance.

## 5. Conclusions

Altogether, our data show that metastatic tumours that lost ER $\alpha$  expression do not simply adopt an ER $\alpha$ -N molecular phenotype. This may have far-reaching implications for future clinical decision-making and emphasises the importance of deep profiling of the metastatic tumour. Furthermore, we identified downregulation of the SGK1 kinase activity in ER $\alpha$ -C advanced BCa, which could be further explored for new therapeutic approaches. Further research is warranted to understand the SGK1-NDRG1 role in BCa disease.

## Acknowledgments

This work was supported the Netherlands Organization for Scientific Research (NWO-VIDI-016.156.401 to WZ). We thank members of the Zwart and Bergman labs for valuable feedback, suggestions and input. We thank the Core Facility Molecular Pathology and Biobanking (CFMPB) from NKI for technical support and sample processing. Finally, we would like to thank

all the patients and clinical staff who were involved in this study.

## Conflict of interest

The authors declare no conflict of interest.

## Author contributions

IM-P and WZ were involved in conceptualisation; WZ was responsible for project funding. IM-P, KS, SB, CJ, SS, RH and CBM collected and processed all the MPE samples for (phospho)proteomics and IHC analyses. MA, PJD, JS and JW scored the IHC samples. DOD performed the (phospho)proteomic analyses and initial data processing. IM-P and SP processed and analysed the phospho(proteomics) data. IM-P, DOD, MA and WZ wrote the manuscript with input from all co-authors.

## Peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/1878-0261.13540>.

## Data accessibility

Proteomics (Table S2) and phosphoproteomics (Table S3) for all tumour samples is provided in the Supporting Information section.

## References

- Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, et al. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics*. 2006;7:96.
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394–424.
- Pan H, 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;377:1836–46.
- 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;34:427–438.e6.
- 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. *JAMA Oncol.* 2016;**2**:1310–5.
- 6 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**:2961–8.
  - 7 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**:1446–51.
  - 8 Yates LR, Knappskog S, Wedge D, Farmery JHR, Gonzalez S, Martincorena I, et al. Genomic evolution of breast cancer metastasis and relapse. *Cancer Cell.* 2017;**32**:169–184.e7.
  - 9 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**:1439–45.
  - 10 Toy W, Weir H, Razavi P, Lawson M, Goepfert AU, Mazzola AM, et al. Activating *ESR1* mutations differentially affect the efficacy of ER antagonists. *Cancer Discov.* 2017;**7**:277–87.
  - 11 Hanker AB, Sudhan DR, Arteaga CL. Overcoming endocrine resistance in breast cancer. *Cancer Cell.* 2020;**37**:496–513.
  - 12 Schrijver WAME, Schuurman K, van Rossum A, Dutch Distant Breast Cancer Metastases Consortium, Peeters T, Ter Hoeve N, et al. Loss of steroid hormone receptors is common in malignant pleural and peritoneal effusions of breast cancer patients treated with endocrine therapy. *Oncotarget.* 2017;**8**:55550–61.
  - 13 Schrijver W, Schuurman K, van Rossum A, Droog M, Jeronimo C, Salta S, et al. *FOXA1* levels are decreased in pleural breast cancer metastases after adjuvant endocrine therapy, and this is associated with poor outcome. *Mol Oncol.* 2018;**12**:1884–94.
  - 14 Post H, Penning R, Fitzpatrick MA, Garrigues LB, Wu W, MacGillavry HD, et al. Robust, sensitive, and automated phosphopeptide enrichment optimized for low sample amounts applied to primary hippocampal neurons. *J Proteome Res.* 2017;**16**:728–37.
  - 15 Reiter L, Rinner O, Picotti P, Hüttenhain R, Beck M, Brusniak MY, et al. MProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat Methods.* 2011;**8**:430–5.
  - 16 Stekhoven DJ, Buhlmann P. MissForest – non-parametric missing value imputation for mixed-type data. *Bioinformatics.* 2012;**28**:112–8.
  - 17 Wiredja DD, Koyutürk M, Chance MR. The KSEA app: a web-based tool for kinase activity inference from quantitative phosphoproteomics. *Bioinformatics.* 2017;**33**:3489–91.
  - 18 DeLong P, Carroll RG, Henry AC, Tanaka T, Ahmad S, Leibowitz MS, et al. Regulatory T cells and cytokines in malignant pleural effusions secondary to mesothelioma and carcinoma. *Cancer Biol Ther.* 2005;**4**:342–6.
  - 19 Nirmal AJ, Regan T, Shih BB-J, Hume DA, Sims AH, Freeman TC. ImSig: a resource for the identification and quantification of immune signatures in blood and tissue transcriptomics data. *bioRxiv.* 2016. <https://doi.org/10.1101/077487>
  - 20 Soysal SD, Muenst S, Barbie T, Fleming T, Gao F, Spizzo G, et al. EpCAM expression varies significantly and is differentially associated with prognosis in the luminal B HER2+, basal-like, and HER2 intrinsic subtypes of breast cancer. *Br J Cancer.* 2013;**108**:1480–7.
  - 21 Louie SM, Grossman EA, Crawford LA, Ding L, Camarda R, Huffman TR, et al. *GSTP1* is a driver of triple-negative breast cancer cell metabolism and pathogenicity. *Cell Chem Biol.* 2016;**23**:567–78.
  - 22 Piao D, Jiang T, Liu G, Wang B, Xu J, Zhu A. Clinical implications of activated leukocyte cell adhesion molecule expression in breast cancer. *Mol Biol Rep.* 2012;**39**:661–8.
  - 23 Bae D-H, Jansson PJ, Huang ML, Kovacevic Z, Kalinowski D, Lee CS, et al. The role of *NDRG1* in the pathology and potential treatment of human cancers. *J Clin Pathol.* 2013;**66**:911–7.
  - 24 Harrod A, Fulton J, Nguyen VTM, Periyasamy M, Ramos-Garcia L, Lai CF, et al. Genomic modelling of the *ESR1* Y537S mutation for evaluating function and new therapeutic approaches for metastatic breast cancer. *Oncogene.* 2017;**36**:2286–96.
  - 25 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**:1865. <https://doi.org/10.1038/s41467-017-01864-y>
  - 26 Jeselsohn R, Bergholz JS, Pun M, Cornwell MI, Liu W, Nardone A, et al. Allele-specific chromatin recruitment and therapeutic vulnerabilities of *ESR1* activating mutations. *Cancer Cell.* 2018;**33**:173–186.e5.
  - 27 Hoefnagel LD, van de Vijver MJ, van Slooten HJ, Wesseling P, Wesseling J, Westenend PJ, et al. Receptor conversion in distant breast cancer metastases. *Breast Cancer Res.* 2010;**12**:R75.
  - 28 Shiino S, Kinoshita T, Yoshida M, Jimbo K, Asaga S, Takayama S, et al. Prognostic impact of discordance in hormone receptor status between primary and recurrent sites in patients with recurrent breast cancer. *Clin Breast Cancer.* 2016;**16**:e133–40.
  - 29 Brady SW, McQuerry JA, Qiao Y, Piccolo SR, Shrestha G, Jenkins DF, et al. Combating subclonal evolution of resistant cancer phenotypes. *Nat Commun.* 2017;**8**:1231.
  - 30 Severson TM, Kim Y, Joosten SEP, Schuurman K, van der Groep P, Moelans CB, et al. Characterizing steroid hormone receptor chromatin binding landscapes in



- male and female breast cancer. *Nat Commun.* 2018;**9**:482. <https://doi.org/10.1038/s41467-018-02856-2>
- 31 Robinson JLL, MacArthur S, Ross-Innes CS, Tilley WD, Neal DE, Mills IG, et al. Androgen receptor driven transcription in molecular apocrine breast cancer is mediated by FoxA1. *EMBO J.* 2011;**30**:3019–27.
- 32 Mohammed H, Russell IA, Stark R, Rueda OM, Hickey TE, Tarulli GA, et al. Progesterone receptor modulates ER $\alpha$  action in breast cancer. *Nature.* 2015;**523**:313–7. <https://doi.org/10.1038/nature14583>
- 33 Yang F, Ma Q, Liu Z, Li W, Tan Y, Jin C, et al. Glucocorticoid receptor:MegaTrans switching mediates the repression of an ER $\alpha$ -regulated transcriptional program. *Mol Cell.* 2017;**66**:321–331.e6.
- 34 D'Amato NC, Gordon MA, Babbs B, Spoelstra NS, Carson Butterfield KT, Torkko KC, et al. Cooperative dynamics of AR and ER activity in breast cancer. *Mol Cancer Res.* 2016;**14**:1054–67. <https://doi.org/10.1158/1541-7786.MCR-16-0167>
- 35 Need EF, Selth LA, Harris TJ, Birrell SN, Tilley WD, Buchanan G. Research resource: interplay between the genomic and transcriptional networks of androgen receptor and estrogen receptor  $\alpha$  in luminal breast cancer cells. *Mol Endocrinol.* 2012;**26**:1941–52. <https://doi.org/10.1210/me.2011-1314>
- 36 Mayayo-Peralta I, Prekovic S, Zwart W. Estrogen receptor on the move: cistromic plasticity and its implications in breast cancer. *Mol Aspects Med.* 2021;**78**:100939.
- 37 McCaig C, Potter L, Abramczyk O, Murray JT. Phosphorylation of NDRG1 is temporally and spatially controlled during the cell cycle. *Biochem Biophys Res Commun.* 2011;**411**:227–34.
- 38 Murakami Y, Hosoi F, Izumi H, Maruyama Y, Ureshino H, Watari K, et al. Identification of sites subjected to serine/threonine phosphorylation by SGK1 affecting N-myc downstream-regulated gene 1 (NDRG1)/Cap43-dependent suppression of angiogenic CXC chemokine expression in human pancreatic cancer cells. *Biochem Biophys Res Commun.* 2010;**396**:376–81.
- 39 Lane DJR, Saletta F, Rahmanto YS, Kovacevic Z, Richardson DR. N-myc downstream regulated 1 (NDRG1) is regulated by eukaryotic initiation factor 3a (eIF3a) during cellular stress caused by iron depletion. *PLoS One.* 2013;**8**:e57273.
- 40 Yu Y, Rahmanto YS, Richardson D. Bp44mT: an orally active iron chelator of the thiosemicarbazone class with potent anti-tumour efficacy. *Br J Pharmacol.* 2012;**165**:148–66.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Fig. S1.** (Phospho)-proteomics workflow and characteristics of proteomics data set. A. Schematic workflow of phospho-proteomic analyses. Protein extraction and digestion was performed, followed by mass spectrometry analysis using data-independent acquisition (DIA) mode. Spectra were matched against an in-house library created from sample pooling, high pH fractionation and data-dependent acquisition (DDA). For the phosphoproteomics data set, a phosphopeptide enrichment step was added. B. Number of quantified proteins per sample in the total proteomics data set. C. Number of quantified proteins per sample in the phospho-proteomics data set.

**Fig. S2.** Distinct proteomic signature indicative of immune cell contamination. A. Unsupervised hierarchical clustering of all proteins in data set groups samples based on tumour cell percentage. The column colour bars indicate the patient group (ER $\alpha$ -C (converted), ER $\alpha$ -P (positive), ER $\alpha$ -N (negative)), whether samples are repeat collections and tumour percentage. B. Principle Component Analyses groups samples with the lowest tumour cell content in the second component (coloured in yellow). C. Unsupervised clustering of immune cell signature proteins (derived from the ImSig database) grouped metastatic breast cancer samples with low tumour cell content together. These samples show high abundance of immune cell proteins. The column colour bars indicate the patient group (ER $\alpha$ -C, ER $\alpha$ -P and ER $\alpha$ -N), whether samples are repeat collections and tumour percentage. D. PCA plot showing samples that were excluded for downstream analyses (coloured in red). E. Immunohistochemistry staining for CD4, CD8, CD20 and CD68. Samples excluded or not from downstream analyses are indicated with a filled and empty dot, respectively. t-Test was performed.

**Fig. S3.** Repeat collections from the same patient reveal a similar (phospho) proteomic landscape. A. Correlation plot of samples with repeat collections. Repeat collections are highlighted in the black boxes. Patient names are coloured by ER $\alpha$ -status. Samples 6A and 6B are excluded from this as 6B had immune infiltration. B. Principle Component Analysis plot showing grouping of patients after excluding samples with low immune infiltration. Samples are coloured based on group (ER $\alpha$ -C (converted) = blue, ER $\alpha$ -P (positive) = red, ER $\alpha$ -N (negative) = grey) C. Phospho-proteomics data. Correlation plots of individual patients demonstrating a higher level of correlation for samples from the same patient, relative to interpatient analyses. Pearson's correlation is provided.

**Fig. S4.** Characterisation of (phospho)proteomic data. A. Principle Component Analysis plot of phospho-



proteomic samples after excluding samples with low immune infiltration. Samples are coloured by group (ER $\alpha$ -C (converted) = blue, ER $\alpha$ -P (positive) = red). B. Phospho-proteomics data. Correlation plots of individual patients showing a higher enrichment of pNRDG1 peptides in ER $\alpha$ -P samples than ER $\alpha$ -C. Pearson correlation is provided.

**Table S1.** Clinopathological data for all pleural effusion samples. For all three patient groups (ER $\alpha$ -C (converted), ER $\alpha$ -P (positive), ER $\alpha$ -N (negative)),

relevant pathological information, patient follow-up and treatment history is provided.

**Table S2.** Proteomics data. Missing-forest imputed data on all proteomics samples (noncentred) used in this study, for ER $\alpha$ -C (converted), ER $\alpha$ -P (positive) and ER $\alpha$ -N (negative) pleural effusions.

**Table S3.** Phospho-proteomics data. Missing-forest imputed data on all phospho-proteomics samples (noncentred) used in this study, for ER $\alpha$ -C (converted) and ER $\alpha$ -P (positive) pleural effusions.