Article

Plasma Proteome Signature to Predict the Outcome of Breast Cancer Patients Receiving Neoadjuvant Chemotherapy

Sungchan Gwark [1,†], Hee-Sung Ahn [2,3,†], Jeonghun Yeom [3], Jiyoung Yu [2], Yumi Oh [2,4], Jae Ho Jeong [5], Jin-Hee Ahn [5], Kyung Hae Jung [5], Sung-Bae Kim [5], Hee Jin Lee [6], Gyungyub Gong [6], Sae Byul Lee [7], Il Yong Chung [7], Hee Jeong Kim [7], Beom Seok Ko [7], Jong Won Lee [7], Byung Ho Son [7], Sei Hyun Ahn [7], Kyunggon Kim [2,3,4,8,9,*] and Jisun Kim [7,*]

1 Department of Surgery, Ewha Womans University Mokdong Hospital, Ewha Womans University College of Medicine, Seoul 07985, Korea; drgwark@ewha.ac.kr
2 Asan Institute for Life Sciences, Asan Medical Center, Seoul 05505, Korea; hsahn_2021@amc.seoul.kr (H.-S.A.); scarlet202@amc.seoul.kr (J.Y.); yum7102@amc.seoul.kr (Y.O.)
3 Convergence Medicine Research Center, Asan Institute for Life Sciences, Asan Medical Center, Seoul 05505, Korea; nature8309@amc.seoul.kr
4 Department of Biomedical Sciences, University of Ulsan College of Medicine, Seoul 05505, Korea
5 Department of Oncology, Asan Medical Center, University of Ulsan College of Medicine, Seoul 05505, Korea; jaeho.jeong@amc.seoul.kr (J.H.J.); drjiny@amc.seoul.kr (J.-H.A.); khjung@amc.seoul.kr (K.H.J.); sbkim3@amc.seoul.kr (S.-B.K.)
6 Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul 05505, Korea; backlila1@amc.seoul.kr (H.J.L.); gygong@amc.seoul.kr (G.G.)
7 Department of Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul 05505, Korea; newstar153@amc.seoul.kr (S.B.L.); doorkeeper1@amc.seoul.kr (I.Y.C.); heejeong_kim@amc.seoul.kr (H.J.K.); spdoctor@amc.seoul.kr (B.S.K.); leejw@amc.seoul.kr (J.W.L.); brdrson@amc.seoul.kr (B.H.S.); ahnsh@amc.seoul.kr (S.H.A.)
8 Clinical Proteomics Core Laboratory, Convergence Medicine Research Center, Asan Medical Center, Seoul 05505, Korea
9 Bio-Medical Institute of Technology, Asan Medical Center, Seoul 05505, Korea
* Correspondence: kkkon1@amc.seoul.kr (K.K.); jisunkim@amc.seoul.kr (J.K.); Tel.: +82-2-1688-7575 (K.K.); +82-2-3010-3935 (J.K.)
† These authors contributed equally to this work.

Simple Summary: The prognostic impact of plasma protein biomarkers in breast cancer patients treated with neoadjuvant chemotherapy (NCT) was evaluated using a proteomics approach. Three biomarkers were identified among differentially expressed proteins. The plasma concentration of APOC3 was higher in the pathological complete response (pCR) group, whereas MBL2, ENG, and P4HB were upregulated in the non-pCR group. Univariate survival analysis was performed to identify protein biomarkers that could classify patients into low- and high-risk groups. The results showed that MBL2 and P4HB were statistically significantly associated with disease-free survival (log-rank test $p < 0.05$); P4HB was statistically significantly associated with overall survival (log-rank test $p < 0.05$), whereas MBL2 was statistically significantly associated with distant metastasis-free survival (log-rank test $p < 0.05$). The results demonstrated that protein markers from non-invasive liquid biopsy sampling correlate with pCR and survival in breast cancer patients receiving NCT. Further investigation of these protein markers may help clarify their role in predicting prognosis and thus their therapeutic potential for preventing metastasis.

Abstract: The plasma proteome of 51 non-metastatic breast cancer patients receiving neoadjuvant chemotherapy (NCT) was prospectively analyzed by high-resolution mass spectrometry coupled with nano-flow liquid chromatography using blood drawn at the time of diagnosis. Plasma proteins were identified as potential biomarkers, and their correlation with clinicopathological variables and survival outcomes was analyzed. Of 51 patients, 20 (39.2%) were HR+/HER2-, five (9.8%) were HR+/HER2+, five (9.8%) were HER2+, and 21 (41.2%) were triple-negative subtype. During a median follow-up of 52.0 months, there were 15 relapses (29.4%) and eight deaths (15.7%). Four potential biomarkers were identified among differentially expressed proteins: APOC3 had higher plasma

concentrations in the pathological complete response (pCR) group, whereas MBL2, ENG, and P4HB were higher in the non-pCR group. Proteins statistically significantly associated with survival and capable of differentiating low- and high-risk groups were MBL2 and P4HB for disease-free survival, P4HB for overall survival, and MBL2 for distant metastasis-free survival (DMFS). In the multivariate analysis, only MBL2 was a consistent risk factor for DMFS (HR: 9.65, 95% CI 2.10–44.31). The results demonstrate that the proteomes from non-invasive sampling correlate with pCR and survival in breast cancer patients receiving NCT. Further investigation may clarify the role of these proteins in predicting prognosis and thus their therapeutic potential for the prevention of recurrence.

1. Introduction

Neoadjuvant chemotherapy (NCT) provides several benefits for locally advanced breast cancer (BC) patients. Down-staging of tumors may increase the probability of breast conservation with better cosmesis [1–3]. Additionally, it allows in vivo monitoring of the response of tumors to therapy, which could be helpful for predicting pathological responses [4–9]. Tumors that respond well to a given therapy show better outcomes, and pathological complete response (pCR) is a surrogate factor for survival in the neoadjuvant setting [10–12]. However, the value of pCR for predicting prognosis in different subtypes of BC, especially in estrogen receptor (ER)-positive and HER2-negative tumors, is under debate. Because of substantial inconsistencies between clinical, radiological, and pathologic responses [13], extensive clinical/laboratory research has focused on achieving a more accurate prediction of the treatment response.

The Oncotype DX® test provides genomic-based personalized information that enables the design of individualized treatments for patients with ER-positive primary BC [14]. The assay provides information on clinicopathological variables and has been validated for use in predicting the response to chemotherapy in ER-positive and HER2-negative BC. However, the assay requires a surgical specimen (formalin-fixed paraffin-embedded (FFPE) >10 × 10 μm) and two to three weeks of processing time, with costs higher than USD 4000. In addition, its indication is limited to the ER-positive and HER2-negative subtype. Plasma cancer biomarkers can provide useful information relevant to the diagnosis of several malignancies. In BC, several serum tumor biomarkers have been proposed, including the MUC-1 antigen (CA 15-3), the onco-fetal protein carcinoembryonic antigen (CEA), the oncoprotein HER-2/neu, and the cytokeratin tissue polypeptide specific antigen (TPS). However, because of their low sensitivity and specificity, the clinical utility of these serum markers is limited [15–18]. Compared with tissue biopsy, a quick and easy 'liquid biopsy' provides a less invasive and simpler method to assess tumor response through simple blood collection [19,20]. Circulating tumor cells (CTCs), which are detected by liquid biopsy, serve as a prognostic factor in metastatic BC, where a high CTC level is associated with poor prognosis [21,22]. However, analysis of CTCs in non-metastatic tumors has shown conflicting results [23–27], including in BC patients treated with NCT [28–31].

In translational research, proteomic-based BC research uses clinical samples such as plasma and tissues, to improve BC care through screening, early or companion diagnostics studies, predicting prognosis, subtyping, predicting metastasis and therapeutic responses, and discovering new drug targets [32–37]. Some protein signatures associated with the response to NCT were identified using FFPE tissues. PYCR1 and ALDH18A1 were identified by comparing the proteome of tissue before and after NCT with that of normal tissue [38]. The combination of four proteins, RAC2, RAB6A, BIEA, and IPYR, showed the best performance for predicting recurrence after NCT in triple-negative BC (TNBC) patients [39]. CD45 was identified as a predictor of pCR to neoadjuvant HER2-targeted therapy using the spatial proteomics approach [40]. The VEGF inhibition response predictor score, which is derived using nine-protein signatures, predicts the response to bevacizumab NCT in

HER2-negative BC [41]. Analysis of blood samples identified differentially expressed proteins related to the pathological criteria for predicting patient responses to NCT [42–46], and blood exosomes [47] have been analyzed using proteomic approaches to predict the response to NCT.

Recent proteomic-centric multiomics studies analyzed BC clinical samples to classify BC into subtypes using protein and phosphorylated protein information, which led to the identification of novel subtypes such as basal-like and luminal B tumors by infiltration of immunological components [48]. Mass spectrometry-based proteomics that include analysis of post-translational modifications such as phosphorylation or acetylation, combined with next-generation DNA and RNA sequencing profiles, may provide a more comprehensive description of breast tumors [49–51]. Proteogenomic approaches highlight the potential of proteomics for clinical research in cancer through the identification of targetable signaling pathways and more precise curation of the biological signatures of tumor heterogeneity. Specifically, different genetic backgrounds may affect the inhibitory relationship between target kinases and tumor suppressors by post-translational modification.

Reliable biomarkers related to treatment response and survival in BC patients receiving NCT are currently lacking, and comorbidities are important for chemotherapy indication and regimen selection [52]. Proteome analysis may provide more reliable and direct information that can be used for monitoring the treatment response and for selecting specific treatment agents. Additionally, plasma proteome analysis can overcome the limitations of current biomarkers and tools by providing faster processing times (<2 days) and by using small samples (40 μL plasma). The aim of the present study was to identify potential biomarkers for predicting the response to therapy and for predicting recurrence by performing whole proteome analyses of the plasma of BC patients undergoing neoadjuvant systemic therapy.

2. Materials and Methods

2.1. Study Participants and Surveillance

Inclusion criteria for the baseline database of the prospective cohort were, (1) women aged >20 years, (2) pathologically diagnosed primary invasive breast cancer, (3) no distant metastasis at time of diagnosis (no de novo stage IV disease), (4) had undergone preoperative systemic treatment with curative intent, and (5) had agreed and signed consent. All patients had 4 mL blood drawn at time of diagnosis before chemotherapy. Among the 60 initial consecutively enrolled patients treated with NCT at Asan Medical Center in Seoul, Korea, between February 2014 and April 2017, 51 were eligible for final analysis with sufficient protein extracted for analysis. All patients' survival outcomes were updated, including loco-regional recurrences, distant metastasis and death information. The study was approved by the Institutional Review Board (IRB) of Asan Medical Center (Seoul, Korea; IRB-e no. 2013-1048), and was performed in compliance with the REMARK criteria [53]. Written informed consent was obtained from all participants. All experiments were performed in accordance with the relevant guidelines and regulations.

All patients of the study received standard treatment, and regular surveillance was performed. The initial diagnostic and follow-up work-up included mammography, breast ultrasound imaging, magnetic resonance imaging, chest X-rays, blood sampling, and clinical examination. ER and progesterone receptor expression were evaluated based on the Allred score [54]. HER2 status was considered negative if the immunohistochemistry score was 1+, or if the score was 2+ and the result of fluorescence or silver in situ hybridization for HER2 amplification was negative [55]. Clinical and histopathological staging was based on the 7th edition of the Cancer Staging Manual of the American Joint Committee on Cancer [56]. The clinical treatment response was evaluated by both physical examination and imaging assessments at each treatment timeline (baseline, after the first treatment, and after completing the course of NCT). Tumor response was assessed by the Response Evaluation Criteria In Solid Tumors (RECIST 1.1) [57,58].

2.2. Sample Preparation

For the proteomic analyses, 51 clinical plasma samples were prepared for LC-MS analysis. Plasma samples were loaded onto a MARS14 column (100 × 4.6 mm; Agilent Technology, Palo Alto, CA, USA) on a Shimadzu binary HPLC system (20A Prominence; Shimadzu, Tokyo, Japan) in order to deplete 14 highly abundant proteins; the unbound fraction was lyophilized with a cold trap (CentriVap Cold Traps; Labconco, Kansas City, MO, USA). Dried samples were resuspended in 400 µL of 5% SDS in 50 mM TEAB (pH 7.55), and dithiothreitol was added to a final concentration of 20 mM for 10 min at 95 °C to reduce disulfide bonds. Reduced samples were then incubated with 40 mM iodoacetamide for 30 min at room temperature in the dark. By a 10-fold dilution of 12% phosphoric acid, acidified samples were loaded onto S-Trap mini columns (ProtiFi, Farmingdale, NY, USA; Cat. No: CO2-mini-80). We treated suspension-trapping (S-trap) proteolysis according to the manufacturer's protocol, followed by the addition of 10 µg Lys-C/trypsin mixture and incubation for 16 h at 37 °C [59]. The eluted peptide mixture was lyophilized using a cold trap and stored at −80 °C until use.

2.3. Nano-LC-ESI-MS/MS Analysis

The LC system was an Dionex UltiMate 3000 RSLCnano system (Thermo Fisher Scientific, Waltham, MA, USA). Mobile phase A was 0.1% formic acid and 5% DMSO in water and mobile phase B was 0.1% formic acid, 5% DMSO and 80% acetonitrile in water. Samples were reconstituted with 25 µL of mobile phase A, injected with a full sample loop injection of 5 µL into a C18 Pepmap trap column (20 × 100 µm i.d., 5 µm, 100 Å; Thermo Fisher Scientific), and separated in Acclaim™ Pepmap 100 C18 column (500 × 75 µm i.d., 3 µm, 100 Å; Thermo Fisher Scientific) over 200 min (250 nL/min) at 50 °C. The column was priory equilibrated with 95% mobile phase A and 5% mobile phase B. A gradient of 5–40% B for 150 min, 40–95% for 2 min, 95% for 23 min, 95–5% B for 10 min, and 5% B for 15 min were applied. The LC system was coupled to a Q Exactive plus mass spectrometer (Thermo Fisher Scientific) with a nano-ESI source. The instrument was operated in the data-dependent mode. One scan cycle included one MS1 scan at a resolution of 70,000 at m/z 400 followed by 20 MS2 scans in higher energy collisional dissociation mode to fragment the 20 most abundant precursor ions identified in the MS1 spectrum. The target value for MS1 by Orbitrap was 3×10^6 with a maximum injection time of 100 ms. The ion target value for MS2 was set to 1×10^6 with a maximum injection time of 50 ms and a resolution of 17,500 at m/z 400. The dynamic exclusion was enabled with the following settings: repeat count = 1 and exclusion duration = 20 s. All MS data were deposited in the Proteomics Identification Database (PRIDE) archive under PXD028251 [60].

2.4. Protein Identification by Database Search

Individual raw files acquired MS analysis and were retrieved against the reviewed Human Uniprot-SwissProt protein database (released on May 2017) [61] using the SEQUEST-HT on Proteome Discoverer (Version 2.2, Thermo Fisher Scientific). Search parameters used were as follows: 10-ppm tolerance for precursor ion mass and 0.02 Da for fragmentation mass. Trypsin peptides tolerate up to two false cleavages. Carbamidomethylation of cysteines was set as fixed modification and N-terminal acetylation and methionine oxidation were set as variable modifications. The false discovery rate (FDR) was calculated using the target-decoy search strategy, and the peptides within 1% of the FDR were selected using the post-processing semi-supervised learning tool Percolator [62] based on the SEQUEST result. Label-free quantitation (LFQ) of proteins was calculated using the precursor ion peak intensity for unique and razor peptides of each protein and excluded peptides with methionine oxidation.

2.5. Differential Data Analysis by Normalization and Filling Missing Data

The normalization method by endogenous normalization proteins, which is mainly used in LFQ, was performed [63,64]. In this study, four proteins (C6, HPX, KNG1, and

SERPINC1) were selected by NormFinder software [65] because the difference between the pCR and non-pCR group was the smallest. Since the quantitative values of the four proteins depend on the characteristics of the plasma sample, they were scaled by dividing the median values of the corresponding proteins in all samples. After that, the geometric mean of the adjusted ratio values of the four proteins for each sample is calculated, and this is defined as the normalization scaling factor (NSF) for that sample. The normalized quantitative values of the remaining proteins except for the four proteins were derived by dividing the raw protein quantitative values in each sample by the NSF. The details of this method are described in previous studies [66–68].

Proteins were selected based on >80% of quantified proteins in all samples, and the missing data were filled by the local least squared imputation method, calculating through correlation with 100% quantified proteins at the raw abundance [69]. After that, normalization was performed.

2.6. Statistical Clinical Model Generation Based on Feature Selection

Feature selection was performed to identify the optimal subset from four proteins [apolipoprotein C3 (APOC3), endoglin (ENG), mannose-binding lectin 2 (MBL2), and prolyl 4-hydrolase beta (P4HB)] to classify patients into two NCT response groups using a random forest (RF)-based backward elimination process [70]. This process consisted of the following two steps: first, 10,000 decision trees containing four variables were randomly generated, and area under the curve (AUC) values were calculated. The AUC values were used to determine the optimal number of proteins using out-of-bag error estimation, yielding a value of three. Second, 50 iterations and three-fold cross-validation were performed using the three selected variables to calculate the predictive importance of each variable included in the model. Three proteins (>0.5 probability of selection) were selected. Data preprocessing, including centering and scaling, were performed before model building. The training and validation sets were divided into thirds using whole data. In the training set, a linear kernel support vector machine (SVM) model [71] with optimized cost parameters was generated by three-fold cross-validation with three repeats. The RF model [72] was optimized with a mtry parameter by three repeats of three-fold cross-validation with 10,000 trees and nodeSize = 5. Machine learning (ML) model prediction values were obtained in the validation set (without the training set samples), and ROC analysis was performed. This process was performed by randomly changing sets 100 times.

2.7. Mining Public Microarray Data

Microarray gene expression data (series accession number: GSE22513 [73–75] and GSE22093 [76,77]) were downloaded from the Gene Expression Omnibus database [78]. The GEO2R interactive web tool was used to extract three identifiers that matched the three selected genes according to the platform record and their expression values. When there were two or more probes for one gene, the median value was estimated.

2.8. Statistical Methods

Survival analyses were performed including DFS, DMFS, and OS. DFS was defined as the time from the date of enrollment into the study to the first date of any type of recurrence. DMFS and overall survival (OS) were defined as the time elapsed between the date of enrollment into the study and the date of distant metastasis or the date of death from any cause, respectively. Statistical analysis was performed using IBM SPSS Statistics Version 26. The univariate Kaplan–Meier method was used to estimate survival probabilities. Multivariate Cox proportional hazards regression analyses were performed for each proteome using the following clinical parameters: patient age at diagnosis; clinical tumor stage; nodal status; hormone receptor (HR) status; and HER2 status. A *p*-value of <0.05 was considered to be statistically significant.

Proteome data evaluation was performed by statistical language R 3.6.0 and RStudio 1.1.456 with the several packages that contained ggplot2 for displaying violin and volcano

plots, mixOmics for PLA-DA [79], RVAideMemoire for calculating VIP scores, stats for applying the t-test, pcaMethods for missing data imputation, survival for survival analysis, survminer for determining cutoff values by maximally selected rank statistics (minimal proportion of observations per group: 20%), and GEOquery for downloading GEO sets.

3. Results

3.1. Baseline Characteristics

The patient characteristics of each subgroup and the NCT regimen of the patients are summarized in Table 1 (detailed in Supplementary Table S1). Twenty-eight patients (47.5%) were HR+/HER2-, five patients (8.5%) were HR+/HER2+, six patients (10.2%) were HER2+, and 20 patients (33.9%) had TNBC. The mean age was 59 years (range, 32–66 years; median age, 46 years). Thirty-eight patients (74.5%) were lymph node-positive. TNBC patients had a significantly higher tumor nuclear grade (Grade 3, 57.1%). Forty-seven patients (92.2%) underwent chemotherapy with an anthracycline-based regimen. According to RECIST criteria [57], 49 patients (96.1%) demonstrated a partial or complete response (PR or CR), two patients (3.9%) showed progressive disease, and none showed stable disease. Twenty-six of the 51 patients underwent breast-conserving surgery followed by radiation therapy (100.0%, 26/26), and 25 had a mastectomy. Of those 25 patients, 20 (80.0%) patients with tumor stage ≥3 or nodal stage ≥2 selectively received radiation therapy. All HR-positive patients received hormonal therapy after surgery. Fifteen patients (29.4%, 15/51) achieved pCR. The pCR rate was significantly higher in HER2+ or triple-negative tumors (HER2+, 40%, 2/5; HR+/HER2+, 60%, 3/5; and triple-negative, 38.1%, 8/21) than in HR+/HER2- patients (10.0%, 2/20), consistent with previous studies [12,80].

Table 1. Patient characteristics and the NCT regimen.

Variables	HR+/HER2- (n = 20)	HR+/HER2+ (n = 5)	HER2+ (n = 5)	Triple-Negative (n = 21)	p
Age at diagnosis (range)	32–58	41–66	45–59	35–53	0.463
≤40	18 (90.0%)	3 (60.0%)	4 (80.0%)	17 (81.0%)	
>40	2 (10.0%)	2 (40.0%)	1 (20.0%)	4 (19.0%)	
Clinical T stage					0.206
T1	0 (0%)	0 (0%)	0 (0%)	0 (0%)	
T2	11 (55.0%)	2 (40.0%)	5 (100%)	15 (71.4%)	
T3	8 (40.0%)	2 (40.0%)	0 (0%)	6 (28.6%)	
T4	1 (5.0%)	1 (20.0%)	0 (0%)	0 (0%)	
Lymph node status					0.473
Negative	6 (30.0%)	0 (0%)	2 (40.0%)	5 (23.8%)	
Positive	14 (70.0%)	5 (100%)	3 (60.0%)	16 (76.2%)	
Nuclear grade					0.001
G1 and G2	19 (95.0%)	5 (100%)	4 (80.0%)	9 (42.9%)	
G3	1 (5.0%)	0 (0%)	1 (20.0%)	12 (57.1%)	
Tumor response (RECIST)					0.295
CR	3 (15.0%)	1 (20.0%)	0 (0%)	7 (33.3%)	
PR	17 (85.0%)	4 (80.0%)	5 (100%)	12 (57.2%)	
SD	0 (0%)	0 (0%)	0 (0%)	0 (0%)	
PD	0 (0%)	0 (0%)	0 (0%)	2 (9.5%)	
Type of Surgery (adjuvant RT)					0.419
BCS (26/26)	8 (40.0%)	4 (80.0%)	3 (60.0%))	11 (52.4%)	
Mastectomy (20/25)	12 (60.0%)	1 (20.0%)	2 (40.0%)	10 (47.6%)	
Pathological response					0.047
pCR	2 (10.0%)	3 (60.0%)	2 (40.0%)	8 (38.1%)	
non-pCR	18 (90.0%)	2 (40.0%)	3 (60.0%)	13 (61.9%)	
NCT regimen					
Anthracycline based (AC#4, AC#4 > D#4, FEC#4 > D#4)				47 (92.2%)	
NCT02032277 * (Veliparib/Placebo + Carboplatin/Placebo + Paclitaxel)				4 (7.8%)	

AC: adriamycin and cyclophosphamide, BCS: breast conserving surgery, CR: complete response, D: docetaxel, FEC: fluorouracil, epirubicin, and cyclophosphamide, N/A: not applicable, NCT: neoadjuvant chemotherapy, pCR: pathological complete response, PD: progressive disease, PR: partial response, RECIST: Response Evaluation Criteria in Solid Tumors, SD: stable disease. * The results of the trial have not yet been reported.

3.2. Proteome Results from Clinical Plasma Samples by LC-MS/MS

A workflow was established for biomarker identification in BC patients with or without pCR after NCT (Figure 1A). To identify prognostic marker candidates for pCR, clinical plasma samples were collected from 51 BC patients, including 15 with and 36 without pCR after NCT. Depleted plasma samples from the 51 study participants were used to analyze constitutive proteins via single LC-MS/MS runs, which led to the identification of 594 proteins. Among them, 548 proteins were quantified in one or more samples using a label-free quantification method. Among these, four relatively stable abundant proteins (C6, HPX, KNG1, and SERPINC1) were used to normalize the raw abundance of the other candidates, which were quantified in at least 80% of the samples [81] (Supplementary Figure S1A and Supplementary Table S2). Before normalization, missing values were imputed [82]. After normalization, 254 common proteins out of 305 proteins showed a significant positive correlation with the plasma concentrations of the published Plasma Proteome Database [83] ($\rho = 0.657$; Pearson's correlation coefficient, permutation $p < 0.001$; Figure S1B). Partial least squares-discriminant analysis (PLS-DA) indicated that the pCR and non-pCR groups were separated into two components, component 1 (6%) and component 2 (17%) (Figure 1B). VIP score-ordered contributions are shown in Figure 1C. The top 26 proteins had VIP scores >1.5. Statistical analysis was performed to identify pCR prediction marker candidates. Four signature proteins annotated by molecular functional terms and processes were selected for building the clinical model. Finally, three biomarkers were selected and used in the survival analysis, including recurrence, death, and metastasis events.
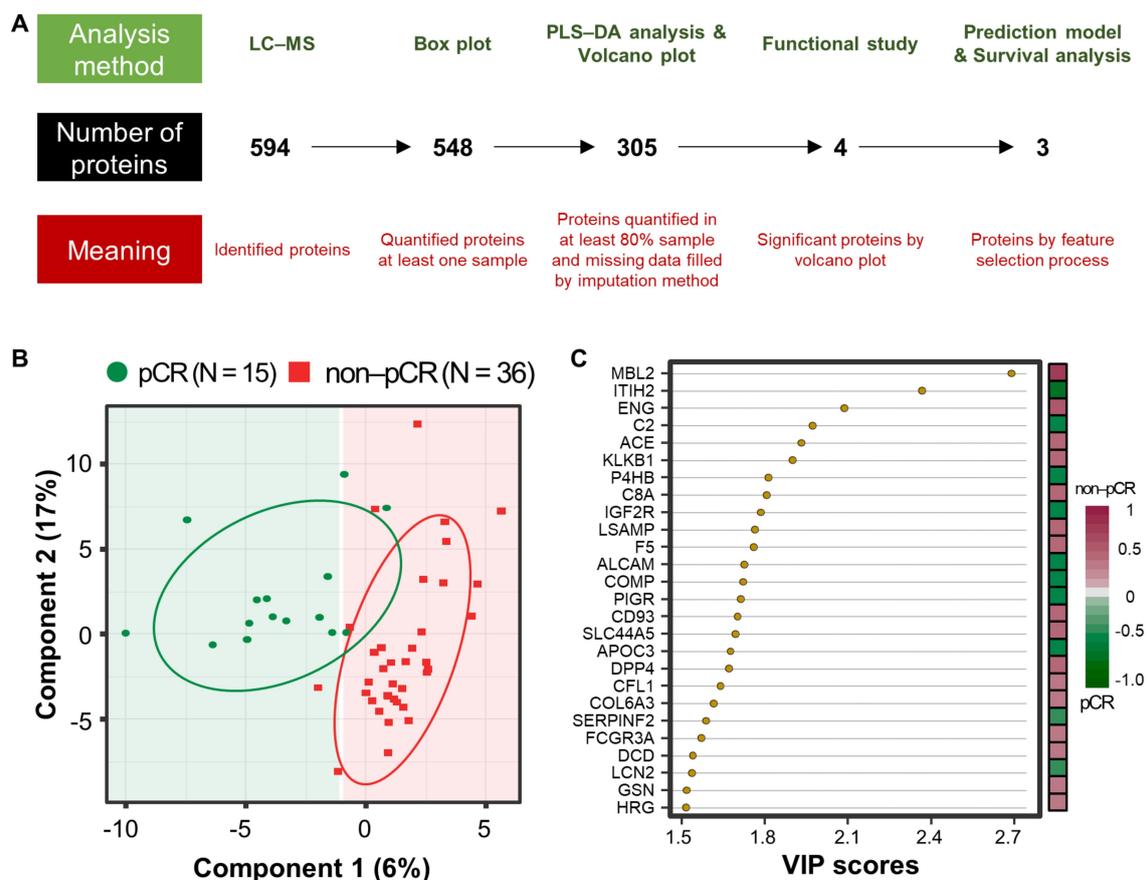

Figure 1. Analysis workflow and partial least-squares discriminant analysis (PLS-DA) of plasma proteomes in 51 breast cancer (BC) patients. (**A**) The analysis method is shown at the top, the number of proteins is shown in the middle, and the meaning of the step is shown at the bottom. PLS-DA score plot (**B**) and top26 variable importance in projection (VIP) score (>1.5) plot derived from PLS-DA analysis (**C**) in 15 patients with pathological complete response (pCR; green) after neoadjuvant therapy and 36 patients with non-pCR (red).

3.3. Differentially Abundant Plasma Proteins between pCR and Non-pCR BC Patients

Statistical analysis was performed using the Student's t-test to identify differentially abundant plasma proteins (DAPs) between the two groups. A volcano plot was drawn to represent log2 fold-changes against negative log10 p-values. We identified a single upregulated protein in the pCR group and three proteins in the non-pCR group ($p < 0.05$ and |fold-change| > 2; Figure 2A and Supplementary Table S3). We examined whether the abundance of the four proteins was related to the subtypes as confounding factors. Each protein was stratified by pCR status, and the quantitative differences according to subtype (HER2 and HR positive or negative) were statistically analyzed ($p > 0.05$; Supplementary Figure S2). The results confirmed that the subtype did not affect the four proteins as a confounding factor.
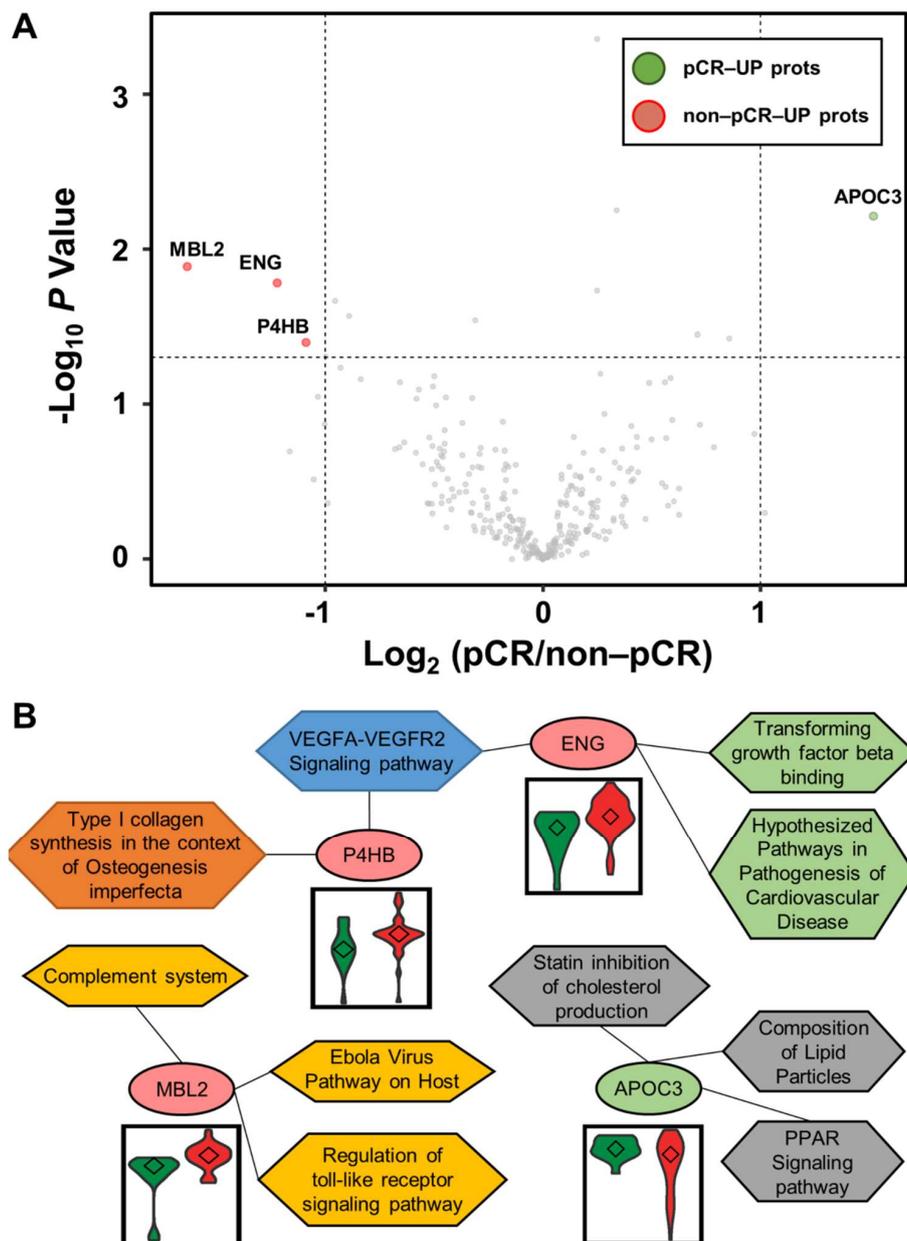

Figure 2. Volcano plot for DAPs altered by pathological complete response (pCR) and functional interpretation. (**A**) Log_2 fold changes and the corresponding *p*-values of all proteins between pCR group ($n = 15$) and non-pCR group ($n = 36$) are presented as volcano plot. Proteins upregulated with more than a twofold change with a *p*-value < 0.05 are depicted in red circles and those downregulated with identical fold change and *p*-value are in green circles. Gray circles show plasma proteins that did not show statistically significant differences. (**B**) Association between WikiPathways and proteins, and violin plots of the corresponding proteins between the two groups (pCR: green; non-pCR: red).

Functional annotation of the proteins was performed using Enrichr [84]. Significant differences between the two groups were identified using WikiPathways (Figure 2B). P4HB and ENG were involved in the "VEGFA-VEGFR2 signaling pathway". ENG was also involved in "transforming growth factor beta binding" and "hypothesized pathways in pathogenesis of cardiovascular disease". P4HB was also associated with "type I collagen synthesis in the context of osteogenesis imperfecta". MBL2 was related to "complement system", "Ebola virus pathway on host", and "regulation of toll-like receptor signaling pathway". APOC3 was linked to "PPAR signaling pathway", "composition of lipid particles", and "statin inhibition of cholesterol production". In addition, we focused on the TNBC subtype, which shows the greatest long-term clinical benefit from pCR in BC [85]. Statistical analysis was performed as described above by dividing patients into pCR and non-pCR groups only in the TNBC subtype (Supplementary Figure S3). In all BC patients, one highly abundant protein, MBL2, was identified in the non-pCR group, and three highly abundant proteins, DCD, KNG1-2, and TLN1, were identified in the non-pCR group. Two highly abundant proteins, ALCAM and MAN1A1, were identified in the pCR group.

3.4. Multivariate Analysis for Predicting pCR Outcome

Multivariate analysis was performed using ML classifiers based on random forest (RF) [72] and SVM [71] to improve the predictive performance for distinguishing pCR from non-pCR patients. First, feature selection was performed with the four significant proteins by AUC-based RF backward elimination [70] according to a probability of selection >0.5 (Table 2) and independently performed 305 proteins as input (Supplementary Table S4). The RF and SVM models were built with three proteins (MBL2, ENG, and P4HB). To avoid overfitting, threefold cross-validation was performed three times to generate 10,000 decision trees from the RF model, and a linear SVM model was applied. To confirm the robustness of the ML models, the sample was randomly trisected 100 times, and the model was then built with 2/3 of the sample and validated with 1/3 of the sample. Evaluation of the performance of the classifiers showed that the median AUC values for SVM and RF were 0.861 (95% CI: 0.845–0.873) and 0.861 (95% CI: 0.830–0.867), respectively (Figure 3).

Table 2. Selected feature proteins by AUC-based RF backward elimination.

Uniprot Accession No.	Gene Name	Importance	Prob. Select *	Selection	Univariate AUC
P11226	MBL2	6.105	0.96	Y	0.807
P17813	ENG	5.556	0.85	Y	0.739
P07237	P4HB	3.522	0.58	Y	0.722
P02656	APOC3	NA	NA	N	0.654

* Probability of selection for each variable.

The three plasma biomarkers were also expressed in tissues of BC patients. The mRNA expression levels of five proteins obtained from BC tissues by fine needle aspiration prior to NCT were analyzed. The data were obtained from two publicly available GEO datasets (GSE22513 [73–75] and GSE22093 [76,77]). ML models were built as described above. In GSE22513, the median AUC values for SVM and RF were 0.631 (95% CI: 0.613–0.643) and 0.646 (95% CI: 0.633–0.669), respectively. In GSE22093, the median AUC values for SVM and RF were 0.709 (95% CI: 0.684–0.713) and 0.658 (95% CI: 0.645–0.666), respectively.
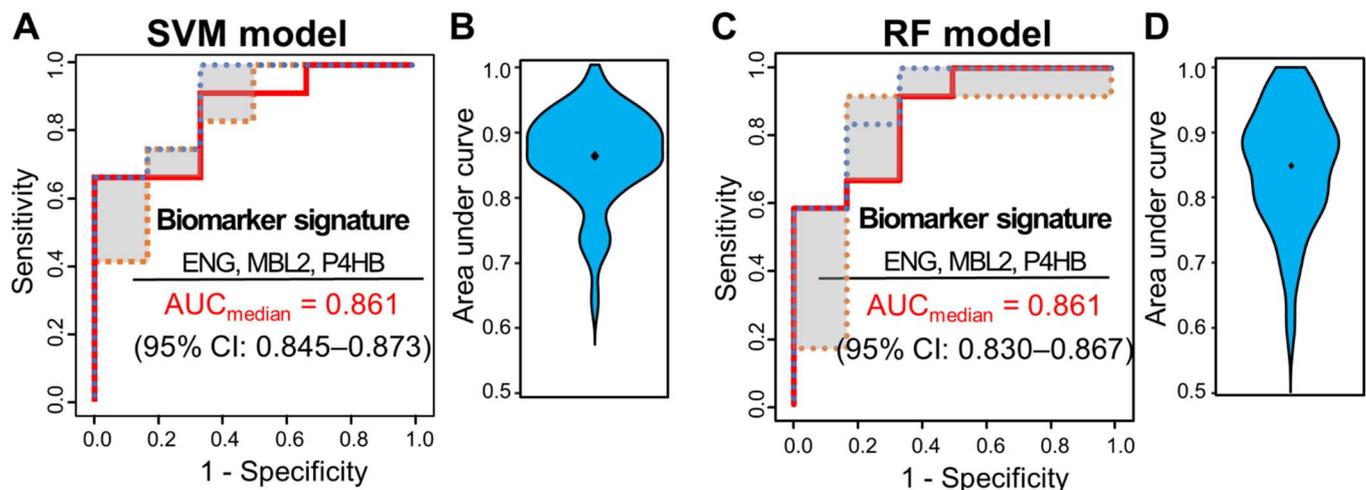
Figure 3. ROC curves of SVM and RF classifiers for three selected proteins (ENG, MBL2, and P4HB). Capability of the two classifiers in a set of 51 samples, 15 from patients with pCR and 36 from patients with non-pCR. (**A**) ROC curves of SVM classifiers generated through 100 repeats of threefold cross-validation steps. (**B**) ROC curves of SVM classifiers generated through 100 repeats of threefold cross-validation steps. ROC curves were obtained by plotting the 25th, 50th, and 75th quantiles of the sensitivities for each value of 1-specificity. (**B**) Violin plots of 100 area under the curve (AUC) values in the SVM model. (**C**) ROC curves of RF classifiers generated through 100 repeats of threefold cross-validation steps. ROC curves were obtained by plotting the 25th, 50th, and 75th quantiles of the sensitivities for each value of 1-specificity. (**D**) Violin plots of 100 AUC values in the RF model.

3.5. Survival Analysis

During a median follow-up of 52.0 months, 15 relapses (29.4%) and eight deaths (15.7%) were observed. To determine the correlation of single plasma proteins with long-term clinical indicators such as DFS, OS, and DMFS, we performed univariate survival analysis for the three proteins in the model (MBL2, ENG, and P4HB), and the remaining 302 proteins were quantified. The Kaplan–Meier method was used to select the cutoff values based on the maximally selected rank statistics. At first, pCR was statistically a better prognostic factor than non-PCR for DFS, OS, and DMFS (log-rank test $p < 0.05$). MBL2 and P4HB for DFS, P4HB for OS, and MBL2 for DMFS were statistically significant in dividing patients into low-risk and high-risk groups (log-rank test $p < 0.05$; Figure 4, Supplementary Figure S4A,B). Among the remaining 302 proteins quantified, the prognosis with respect to the three survival results in the two patient groups was separated by a threshold of protein quantification values: 84 proteins for DFS, 46 proteins for OS, and 96 proteins for DMFS were statistically significant for classifying patients into high-risk and low-risk groups (log-rank test $p < 0.05$; Supplementary Table S5). We also analyzed the prognosis of patients according to Miller-Payne grades in the patients with partial response, but couldn't find a significant correlation (Figure S5). In the multivariate Cox analysis of survival with following factors: patient age at diagnosis; clinical tumor stage; nodal status; hormone receptor (HR) status; and HER2 status, no factor showed significant correlation. However, in DMFS analysis with proteins, MBL2 was identified as the only consistent risk factor (HR: 9.65, 95% CI: 2.10–44.31, $p = 0.004$; Table 3). In other survival analyses including DFS and OS, none of the proteins demonstrated a significant correlation. All three protein (MBL2, ENG, P4HB) levels were significantly increased as the pathological stages elevated (Supplementary Table S6).
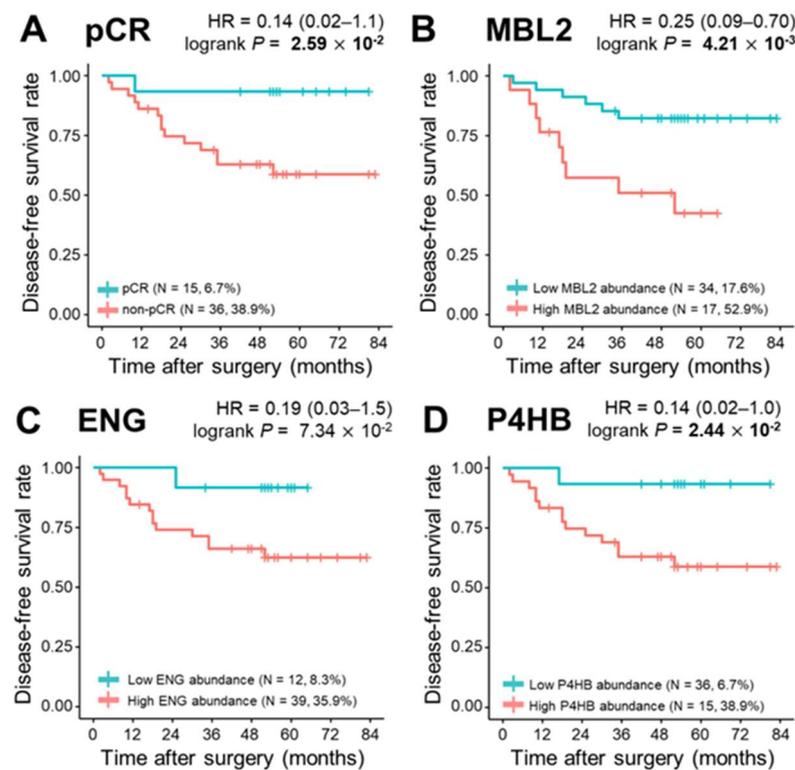
Figure 4. Kaplan–Meier plots of pathological complete response (pCR) and three proteins, MBL2, ENG, and P4HB. (**A**) Categorization of patients into pCR and non—pCR risk groups (pCR, $n = 15$, 6.7%; non-pCR, $n = 36$, 38.9%; $p = 2.59 \times 10^{-2}$). Classification of patients into risk groups according to (**B**) MBL2 abundance: low abundance group ($n = 34$, 17.6%) and high abundance group ($n = 17$, 52.9%), $p = 4.21 \times 10^{-3}$; (**C**) ENG abundance: low abundance group ($n = 12$, 8.3%) and high abundance group ($n = 39$, 35.9%), $p = 7.34 \times 10^{-2}$; and (**D**) P4HB abundance: low abundance group ($n = 36$, 6.7%) and high abundance group ($n = 15$, 38.9%), $p = 2.44 \times 10^{-2}$. Statistical significance was determined using the log-rank test. p-values < 0.05 are displayed in bold.

Table 3. Multivariate Cox analysis of DMFS including the MBL2 proteome.

	DMFS	
	Multivariate HR (95% CI)	*p*
Patient age (>40 vs. ≤40)	1.30 (0.33–5.06)	0.709
Tumor size (≤5 cm vs. >5 cm)	2.55 (0.56–11.65)	0.226
Node negative vs. positive	* 2.1×10^5	0.963
HR positive vs. negative	2.95 (0.63–13.88)	0.172
HER2 negative vs. positive	0.61 (0.07–5.44)	0.660
MBL2 abundance (low vs. high)	9.65 (2.10–44.31)	0.004

Patients were divided into two risk groups according to MBL2 abundance: low abundance group ($n = 34$, 17.6%) and high abundance group ($n = 17$, 52.9%); HR, hormone receptor. * No events in the node negative group.

4. Discussion

Despite considerable advances in our understanding of BC biology, the design of therapeutic approaches is dependent on and guided by molecular profiling that categorizes tumors according to HR and HER2 status [86]. Despite the improvement of treatment strategies related to HR and HER2 status, recent emerging global trends show increased BC mortality rates [87], which are attributed to treatment resistance and highly proliferative BC variants within these subtypes [88]. Thus, the identification of novel markers that can detect resistance and prognosis is an important issue.

The pCR can be a potential surrogate marker with a prognostic value for predicting survival in the HER2-positive and triple-negative subtypes [11,89–91]. In these sub-

types, patients who achieve pCR have a better prognosis than those who fail to achieve pCR [10–12]. Thus, predicting tumor response before or during NCT is important for evaluating patients' prognosis. However, in luminal subtype BC, reliable factors associated with tumor response and prognosis are relatively rare.

In luminal subtype BC, a genomic assay based on tissue biopsy samples such as the Oncotype DX® provides information on clinicopathological factors and is a powerful precision medicine tool for cancer patients. However, its application is limited to the HR-positive and HER2-negative BC subtypes [14]. Moreover, the assay requires a relatively large specimen and is limited by high costs and long processing times.

Another strategy to obtain information for tumor assessment is a 'liquid biopsy'. Compared with tissue biopsy, a quick and easy 'liquid biopsy' allows for a less invasive and simpler assessment of tumor response through simple blood collection [19]. Plasma cancer biomarkers and the number of circulating cancer cells provide useful information that is relevant to cancer diagnosis. CTCs are potential biomarkers of prognosis after treatment [24,26,92,93]. Although studies show that CTC detection is a potential prognostic factor in metastatic BC [21,22,94,95], its clinical value and prognostic impact in non-metastatic BC patients, especially those treated with NCT, are under debate with conflicting results [23,92,96,97].

Proteomics is the analysis method that is similar to CTC analysis and also another form of "non-invasive analysis" as liquid biopsy. Proteomic profiles derived from liquid biopsy samples can provide more direct profiling of diseases. Thus, proteomics is expected to yield promising results for the early diagnosis of cancer and for evaluating the efficacy of antitumor therapy. Current clinical proteomics methods for cancer management are focused on biomarker discovery and validation [98]. Because proteins function through specific pathways rather than individually, anti-cancer strategies can be designed by targeting the biomarker that affects specific pathways related to cancer development [99,100].

In the present study, we analyzed the plasma proteomes of locally advanced BC patients receiving NCT. We used a biomarker discovery workflow system to identify candidate protein biomarkers of tumor response and prognosis in BC patients using LC-MS/MS and three proteins were selected for validation. These proteins were used to generate models that independently predicted the risk of recurrence regardless of tumor subtype. The three protein markers were P4HB, ENG, and MBL2. P4HB is a protein disulfide-isomerase that is one of the core genes in the beta subunit of prolyl 4-hydroxylase [101]. P4HB is related to the carcinogenesis and development of multiple tumors. P4HB is highly expressed in colon cancer, and knockdown of P4HB promotes cancer cell apoptosis [102]. In liver cancer, knockdown of P4HB inhibits the migration and invasion of HepG2/ADR cells, as demonstrated by culturing HepG2 cells (a human hepatocellular carcinoma cell line) in the presence of increasing concentrations of adriamycin [103]. P4HB is also highly expressed in renal clear cell carcinoma and significantly related to poor OS [104]. These findings indicate that P4HB may serve as a potential molecular marker for the diagnosis and treatment of cancer. We demonstrated that low serum P4HB level is significantly associated with better DFS and OS, which is consistent with the findings reported by Yang et al. [105]. The study demonstrated that downregulation of P4HB represses the promoting effects of overexpressed COL10A1 on the proliferation, migration, and invasion of BC cells and, conversely, upregulation of P4HB promotes BC cell proliferation and clone-forming ability, as well as increasing BC cell migration and invasion [105].

ENG (also known as CD105) is a receptor for transforming growth factor β that is expressed at high levels on the cell surface of tumor blood vessels and tumor stromal components [106]. ENG shows affinity for "newly forming" angiogenic endothelium, whereas CD34 and CD31 react not only with angiogenic vessels, but also with the endothelium of normal vessels; ENG is thus superior to CD34 and CD31 for the evaluation of tumor angiogenesis [107]. Elevated expression of ENG is often observed in the actively proliferating endothelium [108,109]. There is a significant correlation between markers of cell proliferation such as Ki-67 and cyclin-A [108]. Thus, ENG is a potential marker of tumor-associated

angiogenesis and prognosis [109]. Li et al. showed that plasma ENG levels are elevated in BC patients at risk of metastasis, and ENG overexpression is significantly correlated with metastatic disease, suggesting the value of ENG for predicting metastasis [110]. Kumar et al. demonstrated that the reactivity of ENG in blood vessels of BC tissues correlates with a poor prognosis [111]. Although the results did not reach statistical significance, we also observed that patients with high ENG levels had poor survival rates.

MBL2 is an activator of the lectin pathway and a crucial component of the innate immune system, and inflammatory reactions are critical for tumor progression and can promote human carcinogenesis [112,113]. MBL can inhibit tumor progression via the complement system and through MBL-dependent cell-mediated cytotoxicity [114–116]. However, recent studies show controversial results. Holm et al. showed that high plasma levels of MBL2 are a marker of poor survival in colorectal cancer patients [117]. Yitting et al. reported that the MBL complement activation pathway is activated in patients with colorectal cancer compared with healthy controls; however, MBL pathway deficiency rates are similar between patients and healthy controls [118]. Additionally, local expression of MBL2 genes is higher in women with ovarian cancer than in controls [119]. In this study, plasma MBL2 levels were higher in the non-pCR group, and high plasma MBL2 levels were associated with poor survival (DFS and DMFS). Taken together, these results suggest that MBL2 can have a protective effect against tumors, as well as a tumorigenic effect.

The present study had several limitations. First, the sample size was small, which limits the statistical power. To overcome any possible overfitting issue associated with the small sample size, cross-validation was used for model development. In addition, the validation study included additional plasma samples collected under IRB. Second, because the study sample was heterogeneous regarding the distribution of tumor subtypes, representing the general BC population is difficult. However, considering the prevalence of each subtype of BC, the present results are valuable for discovering prognosis-related protein signatures. Third, although tumor-derived proteomes are present at high concentrations in the blood of cancer patients, abundant proteins can be derived from cellular sources other than the tumor, which is the major limitation of "liquid biopsy". In terms of homeostasis, the plasma proteome can reflect differences in the immune or inflammatory status of patients, which may affect the response to chemotherapy. The plasma proteome is thus a critical indicator of the chemotherapy response.

Despite the limitations listed above, the findings of this study show that certain proteomes are associated with chemotherapy response and prognosis in patients receiving NCT.

5. Conclusions

This study demonstrated that proteins from non-invasive liquid biopsy sampling correlate with pCR and survival in BC patients receiving NCT. Among them, potential druggable targets were identified. Plasma protein analyses identified differentially expressed proteins between groups with distant metastasis, independently from the achievement of pCR. Quantitative protein analyses by liquid biopsy may provide a means to predict response and recurrence with minimal amounts of sample, at a lower cost, and with faster times. Further investigation of these proteomes may reveal their role in predicting prognosis, which could serve as a novel therapeutic strategy.

Supplementary Materials: The following materials are available online at https://www.mdpi.com/article/10.3390/cancers13246267/s1. Figure S1: (A) Boxplots of normalized plasma protein abundances in the 51 samples (15 BC patients with pathological complete response (pCR) after neoadjuvant therapy and 36 patients with non-pCR) measured by LC-MS analysis. (B) Scatterplot of 254 plasma proteins between log2 plasma concentration in the Plasma Proteome Database (bottom) and normalized log2 abundance (Pearson correlation coefficient (ρ): 0.657 and *p*-value: 9.9×10^{-5}). Figure S2: (A) Boxplot of four proteins (MBL2, ENG, P4HB, and APOC3) in groups according to pCR and HER2 presence (B). Boxplot of four proteins (MBL2, ENG, P4HB, and APOC3) in groups divided by pCR and hormone receptor presence. Figure S3: Volcano plots are depicted with the fold-change of each protein abundance; the *p*-value was calculated by performing a t-test. In the triple-negative

BC (TNBC) subtype (*n* = 21), the averages of the plasma proteomic abundance data in the pCR group (*n* = 8) were compared with the averages of the data for the non-pCR group (*n* = 13) with TNBC. The red circle shows four plasma proteins showing significant increases in the non-pCR group. The green circle shows two plasma proteins with significant decreases in the non-pCR group. Gray circles are plasma proteins with no statistical significance. Figure S4: Kaplan–Meier plots of pCR in relation to three proteins (MBL2, ENG, and P4HB) against overall survival (A) and distant metastasis-free survival (B). Figure S5: Kaplan-Meier plots of Miller–Payne grades in disease-free survival (A), distant meta-free survival (B), and overall survival (C). Statistical significance was determined using the log-rank test. *p*-values < 0.05 are displayed in bold. Table S1. Demographic and clinical characteristics of the 51 breast cancer patients with or without pCR. Table S2. Normalized abundance of 305 plasma proteins in 51 breast patients with or without pCR. Table S3. Results of volcano plot analysis between the pCR group and the non-pCR group. Table S4. Selected feature proteins by AUC-based RF backward elimination with 305 proteins as input. Table S5. Univariate survival analysis of recurrence, metastasis, and death. Table S6. Association between three biomarkers with clinical and pathological stage.

Author Contributions: Conceptualization, J.K. and K.K.; methodology, J.Y. (Jiyoung Yu), J.Y. (Jeonghun Yeom), Y.O., S.G., J.K., H.-S.A. and K.K.; validation, J.K. and K.K.; formal analysis, S.G., H.-S.A.; investigation, S.G. and H.-S.A.; resources, S.B.L., J.K., I.Y.C., H.J.K., B.S.K., J.W.L., B.H.S., S.H.A., J.H.J., J.-H.A., K.H.J. and S.-B.K.; pathologic reports, H.J.L. and G.G.; data curation, S.G., H.-S.A., K.K. and J.K.; writing—original draft preparation, S.G., H.-S.A. and K.K.; writing—review and editing, K.K. and J.K.; visualization, H.-S.A.; supervision, S.B.L., I.Y.C., H.J.K., B.S.K., J.W.L., B.H.S. and S.H.A.; project administration, J.K.; funding acquisition, J.K. and K.K. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Asan Medical Center, Seoul, Korea (IRB number: 2013-1048 approved on 7 November 2013).

Informed Consent Statement: The patients/participants provided written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fisher, B.; Brown, A.; Mamounas, E.; Wieand, S.; Robidoux, A.; Margolese, R.G.; Cruz, A.B., Jr.; Fisher, E.R.; Wickerham, D.L.; Wolmark, N.; et al. Effect of preoperative chemotherapy on local-regional disease in women with operable breast cancer: Findings from National Surgical Adjuvant Breast and Bowel Project B-18. *J. Clin. Oncol.* **1997**, *15*, 2483–2493. [CrossRef] [PubMed]
2. van der Hage, J.A.; van de Velde, C.J.; Julien, J.P.; Tubiana-Hulin, M.; Vandervelden, C.; Duchateau, L. Preoperative chemotherapy in primary operable breast cancer: Results from the European Organization for Research and Treatment of Cancer trial 10902. *J. Clin. Oncol.* **2001**, *19*, 4224–4237. [CrossRef]
3. Caudle, A.S.; Bedrosian, I.; Milton, D.R.; DeSnyder, S.M.; Kuerer, H.M.; Hunt, K.K.; Mittendorf, E.A. Use of Sentinel Lymph Node Dissection After Neoadjuvant Chemotherapy in Patients with Node-Positive Breast Cancer at Diagnosis: Practice Patterns of American Society of Breast Surgeons Members. *Ann. Surg. Oncol.* **2017**, *10*, 2925–2934. [CrossRef]
4. Meisamy, S.; Bolan, P.J.; Baker, E.H.; Bliss, R.L.; Gulbahce, E.; Everson, L.I.; Nelson, M.T.; Emory, T.H.; Tuttle, T.M.; Yee, D. Neoadjuvant chemotherapy of locally advanced breast cancer: Predicting response with in vivo ^{1}H MR spectroscopy—A pilot study at 4 T. *Radiology* **2004**, *233*, 424–431. [CrossRef]
5. Cerussi, A.; Hsiang, D.; Shah, N.; Mehta, R.; Durkin, A.; Butler, J.; Tromberg, B.J. Predicting response to breast cancer neoadjuvant chemotherapy using diffuse optical spectroscopy. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4014–4019. [CrossRef]
6. Sharma, U.; Danishad, K.K.A.; Seenu, V.; Jagannathan, N.R. Longitudinal study of the assessment by MRI and diffusion-weighted imaging of tumor response in patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy. *NMR Biomed. Int. J. Devoted Dev. Appl. Magn. Reson. In Vivo* **2009**, *22*, 104–113. [CrossRef] [PubMed]

7. Manton, D.; Chaturvedi, A.; Hubbard, A.; Lind, M.; Lowry, M.; Maraveyas, A.; Pickles, M.; Tozer, D.; Turnbull, L. Neoadjuvant chemotherapy in breast cancer: Early response prediction with quantitative MR imaging and spectroscopy. *Br. J. Cancer* **2006**, *94*, 427. [CrossRef] [PubMed]

8. Rajan, R.; Poniecka, A.; Smith, T.L.; Yang, Y.; Frye, D.; Pusztai, L.; Fiterman, D.J.; Gal-Gombos, E.; Whitman, G.; Rouzier, R. Change in tumor cellularity of breast carcinoma after neoadjuvant chemotherapy as a variable in the pathologic assessment of response. *Cancer Interdisc. Int. J. Am. Cancer Soc.* **2004**, *100*, 1365–1373. [CrossRef]

9. Baek, H.-M.; Chen, J.-H.; Nie, K.; Yu, H.J.; Bahri, S.; Mehta, R.S.; Nalcioglu, O.; Su, M.-Y. Predicting pathologic response to neoadjuvant chemotherapy in breast cancer by using MR imaging and quantitative ^{1}H MR spectroscopy. *Radiology* **2009**, *251*, 653–662. [CrossRef]

10. Liedtke, C.; Mazouni, C.; Hess, K.R.; André, F.; Tordai, A.; Mejia, J.A.; Symmans, W.F.; Gonzalez-Angulo, A.M.; Hennessy, B.; Green, M. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J. Clin. Oncol.* **2008**, *26*, 1275–1281. [CrossRef] [PubMed]

11. Houssami, N.; Macaskill, P.; von Minckwitz, G.; Marinovich, M.L.; Mamounas, E. Meta-analysis of the association of breast cancer subtype and pathologic complete response to neoadjuvant chemotherapy. *Eur. J. Cancer* **2012**, *48*, 3342–3354. [CrossRef]

12. Carey, L.A.; Dees, E.C.; Sawyer, L.; Gatti, L.; Moore, D.T.; Collichio, F.; Ollila, D.W.; Sartor, C.I.; Graham, M.L.; Perou, C.M. The triple negative paradox: Primary tumor chemosensitivity of breast cancer subtypes. *Clin. Cancer Res.* **2007**, *13*, 2329–2334. [CrossRef] [PubMed]

13. Schott, A.F.; Roubidoux, M.A.; Helvie, M.A.; Hayes, D.F.; Kleer, C.G.; Newman, L.A.; Pierce, L.J.; Griffith, K.A.; Murray, S.; Hunt, K.A.; et al. Clinical and radiologic assessments to predict breast cancer pathologic complete response to neoadjuvant chemotherapy. *Breast Cancer Res. Treat.* **2005**, *92*, 231–238. [CrossRef]

14. Baehner, F.L. The analytical validation of the Oncotype DX Recurrence Score assay. *Ecancermedicalscience* **2016**, *10*, 675. [CrossRef]

15. Safi, F.; Kohler, I.; Beger, H.G.; Röttinger, E. The value of the tumor marker CA 15-3 in diagnosing and monitoring breast cancer. A comparative study with carcinoembryonic antigen. *Cancer* **1991**, *68*, 574–582. [CrossRef]

16. Lumachi, F.; Basso, S.M.M. Serum tumor markers in patients with breast cancer. *Expert Rev. Anticancer Ther.* **2004**, *4*, 921–931. [CrossRef]

17. Bayo, J.; Castaño, M.; Rivera, F.; Navarro, F. Analysis of blood markers for early breast cancer diagnosis. *Clin. Transl. Oncol.* **2017**, *20*, 467–475. [CrossRef] [PubMed]

18. Martínez-Trufero, J.; de Lobera, A.R.; Lao, J.; Puértolas, T.; Artal-Cortés, A.; Zorrilla, M.; Alonso, V.; Pazo, R.; Valero, M.I.; Ríos-Mitchell, M.J. Serum markers and prognosis in locally advanced breast cancer. *Tumori J.* **2005**, *91*, 522–530. [CrossRef]

19. Alix-Panabieres, C.; Pantel, K. Liquid Biopsy: From Discovery to Clinical Application. *Cancer Discov.* **2021**, *11*, 858–873. [CrossRef]

20. Alix-Panabieres, C.; Pantel, K. Circulating tumor cells: Liquid biopsy of cancer. *Clin. Chem.* **2013**, *59*, 110–118. [CrossRef] [PubMed]

21. Cristofanilli, M.; Budd, G.T.; Ellis, M.J.; Stopeck, A.; Matera, J.; Miller, M.C.; Reuben, J.M.; Doyle, G.V.; Allard, W.J.; Terstappen, L.W.; et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N. Engl. J. Med.* **2004**, *351*, 781–791. [CrossRef]

22. Hayes, D.F.; Cristofanilli, M.; Budd, G.T.; Ellis, M.J.; Stopeck, A.; Miller, M.C.; Matera, J.; Allard, W.J.; Doyle, G.V.; Terstappen, L.W. Circulating tumor cells at each follow-up time point during therapy of metastatic breast cancer patients predict progression-free and overall survival. *Clin. Cancer Res.* **2006**, *12*, 4218–4224. [CrossRef] [PubMed]

23. Van Dalum, G.; Van der Stam, G.J.; Tibbe, A.G.; Franken, B.; Mastboom, W.J.; Vermes, I.; De Groot, M.R.; Terstappen, L.W. Circulating tumor cells before and during follow-up after breast cancer surgery. *Int. J. Oncol.* **2015**, *46*, 407–413. [CrossRef] [PubMed]

24. Rack, B.; Schindlbeck, C.; Jückstock, J.; Andergassen, U.; Hepp, P.; Zwingers, T.; Friedl, T.W.P.; Lorenz, R.; Tesch, H.; Fasching, P.A.; et al. Circulating Tumor Cells Predict Survival in Early Average-to-High Risk Breast Cancer Patients. *JNCI J. Natl. Cancer Inst.* **2014**, *106*, dju066. [CrossRef] [PubMed]

25. Franken, B.; de Groot, M.R.; Mastboom, W.J.; Vermes, I.; van der Palen, J.; Tibbe, A.G.; Terstappen, L.W. Circulating tumor cells, disease recurrence and survival in newly diagnosed breast cancer. *Breast Cancer Res.* **2012**, *14*, R133. [CrossRef]

26. Bidard, F.C.; Mathiot, C.; Delaloge, S.; Brain, E.; Giachetti, S.; de Cremoux, P.; Marty, M.; Pierga, J.Y. Single circulating tumor cell detection and overall survival in nonmetastatic breast cancer. *Ann. Oncol.* **2010**, *21*, 729–733. [CrossRef]

27. Lucci, A.; Hall, C.S.; Lodhi, A.K.; Bhattacharyya, A.; Anderson, A.E.; Xiao, L.; Bedrosian, I.; Kuerer, H.M.; Krishnamurthy, S. Circulating tumour cells in non-metastatic breast cancer: A prospective study. *Lancet Oncol.* **2012**, *13*, 688–695. [CrossRef]

28. Hall, C.; Karhade, M.; Laubacher, B.; Anderson, A.; Kuerer, H.; DeSynder, S.; Lucci, A. Circulating tumor cells after neoadjuvant chemotherapy in stage I–III triple-negative breast cancer. *Ann. Surg. Oncol.* **2015**, *22*, 552–558. [CrossRef]

29. Fei, F.; Du, Y.; Di, G.; Wu, J.; Shao, Z. Are changes in circulating tumor cell (CTC) count associated with the response to neoadjuvant chemotherapy in local advanced breast cancer? A meta-analysis. *Oncol. Res. Treat.* **2014**, *37*, 250–254. [CrossRef]

30. Kasimir-Bauer, S.; Bittner, A.-K.; König, L.; Reiter, K.; Keller, T.; Kimmig, R.; Hoffmann, O. Does primary neoadjuvant systemic therapy eradicate minimal residual disease? Analysis of disseminated and circulating tumor cells before and after therapy. *Breast Cancer Res.* **2016**, *18*, 20. [CrossRef]

31. Gwark, S.; Kim, J.; Kwon, N.-J.; Kim, K.-Y.; Kim, Y.; Lee, C.H.; Kim, Y.H.; Kim, M.S.; Hong, S.W.; Choi, M.Y. Analysis of the serial circulating tumor cell count during neoadjuvant chemotherapy in breast cancer patients. *Sci. Rep.* **2020**, *10*, 17466. [CrossRef] [PubMed]

32. Bertucci, F.; Birnbaum, D.; Goncalves, A. Proteomics of breast cancer—Principles and potential clinical applications. *Mol. Cell. Proteom.* **2006**, *5*, 1772–1786. [CrossRef]

33. Galvao, E.R.; Martins, L.M.; Ibiapina, J.O.; Andrade, H.M.; Monte, S.J. Breast cancer proteomics: A review for clinicians. *J. Cancer Res. Clin. Oncol.* **2011**, *137*, 915–925. [CrossRef]

34. Gast, M.C.; Schellens, J.H.; Beijnen, J.H. Clinical proteomics in breast cancer: A review. *Breast Cancer Res. Treat.* **2009**, *116*, 17–29. [CrossRef]

35. Goncalves, A.; Bertucci, F. Clinical application of proteomics in breast cancer: State of the art and perspectives. *Med. Princ. Pract.* **2011**, *20*, 4–18. [CrossRef]

36. Lumachi, F.; Chiara, G.B.; Foltran, L.; Basso, S.M. Proteomics as a Guide for Personalized Adjuvant Chemotherapy in Patients with Early Breast Cancer. *Cancer Genom. Proteom.* **2015**, *12*, 385–390.

37. Tyanova, S.; Albrechtsen, R.; Kronqvist, P.; Cox, J.; Mann, M.; Geiger, T. Proteomic maps of breast cancer subtypes. *Nat. Commun.* **2016**, *7*, 10259. [CrossRef]

38. Shenoy, A.; Belugali Nataraj, N.; Perry, G.; Loayza Puch, F.; Nagel, R.; Marin, I.; Balint, N.; Bossel, N.; Pavlovsky, A.; Barshack, I.; et al. Proteomic patterns associated with response to breast cancer neoadjuvant treatment. *Mol. Syst. Biol.* **2020**, *16*, e9443. [CrossRef]

39. Gonzalez-Angulo, A.M.; Hennessy, B.T.; Meric-Bernstam, F.; Sahin, A.; Liu, W.; Ju, Z.; Carey, M.S.; Myhre, S.; Speers, C.; Deng, L.; et al. Functional proteomics can define prognosis and predict pathologic complete response in patients with breast cancer. *Clin. Proteom.* **2011**, *8*, 11. [CrossRef] [PubMed]

40. McNamara, K.L.; Caswell-Jin, J.L.; Joshi, R.; Ma, Z.; Kotler, E.; Bean, G.R.; Kriner, M.; Zhou, Z.; Hoang, M.; Beechem, J.; et al. Spatial proteomic characterization of HER2-positive breast tumors through neoadjuvant therapy predicts response. *Nat. Cancer* **2021**, *2*, 400–413. [CrossRef]

41. Haugen, M.H.; Lingjaerde, O.C.; Hedenfalk, I.; Garred, O.; Borgen, E.; Loman, N.; Hatschek, T.; Borresen-Dale, A.L.; Naume, B.; Mills, G.B.; et al. Protein Signature Predicts Response to Neoadjuvant Treatment With Chemotherapy and Bevacizumab in HER2-Negative Breast Cancers. *JCO Precis. Oncol.* **2021**, *5*, 286–306. [CrossRef]

42. Bonneterre, J.; Revillion, F.; Desauw, C.; Blot, E.; Kramar, A.; Fournier, C.; Hornez, L.; Peyrat, J.P. Plasma and tissue proteomic prognostic factors of response in primary breast cancer patients receiving neoadjuvant chemotherapy. *Oncol. Rep.* **2013**, *29*, 355–361. [CrossRef]

43. He, J.; Whelan, S.A.; Lu, M.; Shen, D.; Chung, D.U.; Saxton, R.E.; Faull, K.F.; Whitelegge, J.P.; Chang, H.R. Proteomic-based biosignatures in breast cancer classification and prediction of therapeutic response. *Int. J. Proteom.* **2011**, *2011*, 896476. [CrossRef] [PubMed]

44. Mazouni, C.; Baggerly, K.; Hawke, D.; Tsavachidis, S.; Andre, F.; Buzdar, A.U.; Martin, P.M.; Kobayashi, R.; Pusztai, L. Evaluation of changes in serum protein profiles during neoadjuvant chemotherapy in HER2-positive breast cancer using an LC-MALDI-TOF/MS procedure. *Proteomics* **2010**, *10*, 3525–3532. [CrossRef]

45. Michlmayr, A.; Bachleitner-Hofmann, T.; Baumann, S.; Marchetti-Deschmann, M.; Rech-Weichselbraun, I.; Burghuber, C.; Pluschnig, U.; Bartsch, R.; Graf, A.; Greil, R.; et al. Modulation of plasma complement by the initial dose of epirubicin/docetaxel therapy in breast cancer and its predictive value. *Br. J. Cancer* **2010**, *103*, 1201–1208. [CrossRef]

46. Yang, T.; Fu, Z.; Zhang, Y.; Wang, M.; Mao, C.; Ge, W. Serum proteomics analysis of candidate predictive biomarker panel for the diagnosis of trastuzumab-based therapy resistant breast cancer. *Biomed. Pharm.* **2020**, *129*, 110465. [CrossRef] [PubMed]

47. Koh, Y.Q.; Ng, D.Q.; Ng, C.C.; Boey, A.; Wei, M.; Sze, S.K.; Ho, H.K.; Acharya, M.; Limoli, C.L.; Chan, A. Extracellular Vesicle Proteome of Breast Cancer Patients with and Without Cognitive Impairment Following Anthracycline-based Chemotherapy: An Exploratory Study. *Biomark Insights* **2021**, *16*, 11772719211018204. [CrossRef]

48. Johansson, H.J.; Socciarelli, F.; Vacanti, N.M.; Haugen, M.H.; Zhu, Y.; Siavelis, I.; Fernandez-Woodbridge, A.; Aure, M.R.; Sennblad, B.; Vesterlund, M.; et al. Breast cancer quantitative proteome and proteogenomic landscape. *Nat. Commun.* **2019**, *10*, 1600. [CrossRef] [PubMed]

49. De Marchi, T.; Pyl, P.T.; Sjostrom, M.; Klasson, S.; Sartor, H.; Tran, L.; Pekar, G.; Malmstrom, J.; Malmstrom, L.; Nimeus, E. Proteogenomic Workflow Reveals Molecular Phenotypes Related to Breast Cancer Mammographic Appearance. *J. Proteome Res.* **2021**, *20*, 2983–3001. [CrossRef] [PubMed]

50. Krug, K.; Jaehnig, E.J.; Satpathy, S.; Blumenberg, L.; Karpova, A.; Anurag, M.; Miles, G.; Mertins, P.; Geffen, Y.; Tang, L.C.; et al. Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. *Cell* **2020**, *183*, 1436–1456.e31. [CrossRef]

51. Mertins, P.; Mani, D.R.; Ruggles, K.V.; Gillette, M.A.; Clauser, K.R.; Wang, P.; Wang, X.; Qiao, J.W.; Cao, S.; Petralia, F.; et al. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **2016**, *534*, 55–62. [CrossRef]

52. Moo, T.A.; Sanford, R.; Dang, C.; Morrow, M. Overview of Breast Cancer Therapy. *PET Clin.* **2018**, *13*, 339–354. [CrossRef]

53. McShane, L.M.; Altman, D.G.; Sauerbrei, W.; Taube, S.E.; Gion, M.; Clark, G.M. Reporting recommendations for tumor marker prognostic studies (REMARK). *J. Natl. Cancer Inst.* **2005**, *97*, 1180–1184. [CrossRef]

54. Allred, D.C.; Harvey, J.M.; Berardo, M.; Clark, G.M. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod. Pathol.* **1998**, *11*, 155–168.

55. Wolff, A.C.; Hammond, M.E.H.; Hicks, D.G.; Dowsett, M.; McShane, L.M.; Allison, K.H.; Allred, D.C.; Bartlett, J.M.S.; Bilous, M.; Fitzgibbons, P.; et al. Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Update. *Arch. Pathol. Lab. Med.* **2014**, *138*, 241–256. [CrossRef]

56. Edge, S.B.; Compton, C.C. The American Joint Committee on Cancer: The 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann. Surg. Oncol.* **2010**, *17*, 1471–1474. [CrossRef]

57. Eisenhauer, E.A.; Therasse, P.; Bogaerts, J.; Schwartz, L.H.; Sargent, D.; Ford, R.; Dancey, J.; Arbuck, S.; Gwyther, S.; Mooney, M. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur. J. Cancer* **2009**, *45*, 228–247. [CrossRef]

58. Therasse, P.; Arbuck, S.G.; Eisenhauer, E.A.; Wanders, J.; Kaplan, R.S.; Rubinstein, L.; Verweij, J.; Van Glabbeke, M.; van Oosterom, A.T.; Christian, M.C. New guidelines to evaluate the response to treatment in solid tumors. *J. Natl. Cancer Inst.* **2000**, *92*, 205–216. [CrossRef]

59. HaileMariam, M.; Eguez, R.V.; Singh, H.; Bekele, S.; Ameni, G.; Pieper, R.; Yu, Y. S-Trap, an Ultrafast Sample-Preparation Approach for Shotgun Proteomics. *J. Proteome Res.* **2018**, *17*, 2917–2924. [CrossRef]

60. Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D.J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; et al. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **2019**, *47*, D442–D450. [CrossRef]

61. Bateman, A. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489. [CrossRef]

62. Kall, L.; Canterbury, J.D.; Weston, J.; Noble, W.S.; MacCoss, M.J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **2007**, *4*, 923–925. [CrossRef] [PubMed]

63. Martinez-Aguilar, J.; Molloy, M.P. Label-free selected reaction monitoring enables multiplexed quantitation of S100 protein isoforms in cancer cells. *J. Proteome Res.* **2013**, *12*, 3679–3688. [CrossRef] [PubMed]

64. Wisniewski, J.R.; Hein, M.Y.; Cox, J.; Mann, M. A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. *Mol. Cell. Proteom.* **2014**, *13*, 3497–3506. [CrossRef]

65. Andersen, C.L.; Jensen, J.L.; Orntoft, T.F. Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* **2004**, *64*, 5245–5250. [CrossRef]

66. Ahn, H.S.; Ho, J.Y.; Yu, J.; Yeom, J.; Lee, S.; Hur, S.Y.; Jung, Y.; Kim, K.; Choi, Y.J. Plasma Protein Biomarkers Associated with Higher Ovarian Cancer Risk in BRCA1/2 Carriers. *Cancers* **2021**, *13*, 2300. [CrossRef]

67. Ahn, H.S.; Kim, J.H.; Jeong, H.; Yu, J.; Yeom, J.; Song, S.H.; Kim, S.S.; Kim, I.J.; Kim, K. Differential Urinary Proteome Analysis for Predicting Prognosis in Type 2 Diabetes Patients with and without Renal Dysfunction. *Int. J. Mol. Sci.* **2020**, *21*, 4236. [CrossRef]

68. Ahn, H.S.; Sohn, T.S.; Kim, M.J.; Cho, B.K.; Kim, S.M.; Kim, S.T.; Yi, E.C.; Lee, C. SEPROGADIC—Serum protein-based gastric cancer prediction model for prognosis and selection of proper adjuvant therapy. *Sci. Rep.* **2018**, *8*, 16892. [CrossRef]

69. Kim, H.; Golub, G.H.; Park, H. Missing value estimation for DNA microarray gene expression data: Local least squares imputation. *Bioinformatics* **2005**, *21*, 187–198. [CrossRef]

70. Calle, M.L.; Urrea, V.; Boulesteix, A.L.; Malats, N. AUC-RF: A new strategy for genomic profiling with random forest. *Hum. Hered.* **2011**, *72*, 121–132. [CrossRef]

71. Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

72. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

73. Adams, S.; Chakravarthy, A.B.; Donach, M.; Spicer, D.; Lymberis, S.; Singh, B.; Bauer, J.A.; Hochman, T.; Goldberg, J.D.; Muggia, F.; et al. Preoperative concurrent paclitaxel-radiation in locally advanced breast cancer: Pathologic response correlates with five-year overall survival. *Breast Cancer Res. Treat.* **2010**, *124*, 723–732. [CrossRef] [PubMed]

74. Bauer, J.A.; Chakravarthy, A.B.; Rosenbluth, J.M.; Mi, D.; Seeley, E.H.; De Matos Granja-Ingram, N.; Olivares, M.G.; Kelley, M.C.; Mayer, I.A.; Meszoely, I.M.; et al. Identification of markers of taxane sensitivity using proteomic and genomic analyses of breast tumors from patients receiving neoadjuvant paclitaxel and radiation. *Clin. Cancer Res.* **2010**, *16*, 681–690. [CrossRef] [PubMed]

75. Lehmann, B.D.; Bauer, J.A.; Chen, X.; Sanders, M.E.; Chakravarthy, A.B.; Shyr, Y.; Pietenpol, J.A. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Investig.* **2011**, *121*, 2750–2767. [CrossRef]

76. Iwamoto, T.; Bianchini, G.; Booser, D.; Qi, Y.; Coutant, C.; Shiang, C.Y.; Santarpia, L.; Matsuoka, J.; Hortobagyi, G.N.; Symmans, W.F.; et al. Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer. *J. Natl. Cancer Inst.* **2011**, *103*, 264–272. [CrossRef]

77. Shen, K.; Song, N.; Kim, Y.; Tian, C.; Rice, S.D.; Gabrin, M.J.; Symmans, W.F.; Pusztai, L.; Lee, J.K. A systematic evaluation of multi-gene predictors for the pathological response of breast cancer patients to chemotherapy. *PLoS ONE* **2012**, *7*, e49529. [CrossRef]

78. Barrett, T.; Edgar, R. Gene expression omnibus: Microarray data storage, submission, retrieval, and analysis. *Methods Enzym.* **2006**, *411*, 352–369. [CrossRef]

79. Rohart, F.; Gautier, B.; Singh, A.; Le Cao, K.A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* **2017**, *13*, e1005752. [CrossRef]

80. Rouzier, R.; Perou, C.M.; Symmans, W.F.; Ibrahim, N.; Cristofanilli, M.; Anderson, K.; Hess, K.R.; Stec, J.; Ayers, M.; Wagner, P. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin. Cancer Res.* **2005**, *11*, 5678–5685. [CrossRef] [PubMed]

81. Dziura, J.D.; Post, L.A.; Zhao, Q.; Fu, Z.; Peduzzi, P. Strategies for dealing with missing data in clinical trials: From design to analysis. *Yale J. Biol. Med.* **2013**, *86*, 343–358.

82. Karpievitch, Y.V.; Dabney, A.R.; Smith, R.D. Normalization and missing value imputation for label-free LC-MS analysis. *BMC Bioinform.* **2012**, *13* (Suppl. 16), S5. [CrossRef] [PubMed]

83. Nanjappa, V.; Thomas, J.K.; Marimuthu, A.; Muthusamy, B.; Radhakrishnan, A.; Sharma, R.; Ahmad Khan, A.; Balakrishnan, L.; Sahasrabuddhe, N.A.; Kumar, S.; et al. Plasma Proteome Database as a resource for proteomics research: 2014 update. *Nucleic Acids Res.* **2014**, *42*, D959–D965. [CrossRef]

84. Kuleshov, M.V.; Jones, M.R.; Rouillard, A.D.; Fernandez, N.F.; Duan, Q.; Wang, Z.; Koplev, S.; Jenkins, S.L.; Jagodnik, K.M.; Lachmann, A.; et al. Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **2016**, *44*, W90–W97. [CrossRef] [PubMed]

85. Cortazar, P.; Zhang, L.; Untch, M.; Mehta, K.; Costantino, J.P.; Wolmark, N.; Bonnefoi, H.; Cameron, D.; Gianni, L.; Valagussa, P.; et al. Pathological complete response and long-term clinical benefit in breast cancer: The CTNeoBC pooled analysis. *Lancet* **2014**, *384*, 164–172. [CrossRef]

86. Koboldt, D.; Fulton, R.; McLellan, M.; Schmidt, H.; Kalicki-Veizer, J.; McMichael, J.; Fulton, L.; Dooling, D.; Ding, L.; Mardis, E. Comprehensive molecular portraits of human breast tumours. *Nature* **2012**, *490*, 61–70.

87. Heer, E.; Harper, A.; Escandor, N.; Sung, H.; McCormack, V.; Fidler-Benaoudia, M.M. Global burden and trends in premenopausal and postmenopausal breast cancer: A population-based study. *Lancet Glob. Health* **2020**, *8*, e1027–e1037. [CrossRef]

88. Redig, A.J.; McAllister, S.S. Breast cancer as a systemic disease: A view of metastasis. *J. Intern. Med.* **2013**, *274*, 113–126. [CrossRef]

89. Fehm, T.; Hoffmann, O.; Aktas, B.; Becker, S.; Solomayer, E.F.; Wallwiener, D.; Kimmig, R.; Kasimir-Bauer, S. Detection and characterization of circulating tumor cells in blood of primary breast cancer patients by RT-PCR and comparison to status of bone marrow disseminated cells. *Breast Cancer Res.* **2009**, *11*, R59. [CrossRef]

90. Von Minckwitz, G.; Untch, M.; Blohmer, J.-U.; Costa, S.D.; Eidtmann, H.; Fasching, P.A.; Gerber, B.; Eiermann, W.; Hilfrich, J.; Huober, J. Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J. Clin. Oncol.* **2012**, *30*, 1796–1804. [CrossRef]

91. Broglio, K.R.; Quintana, M.; Foster, M.; Olinger, M.; McGlothlin, A.; Berry, S.M.; Boileau, J.-F.; Brezden-Masley, C.; Chia, S.; Dent, S. Association of pathologic complete response to neoadjuvant therapy in HER2-positive breast cancer with long-term outcomes: A meta-analysis. *JAMA Oncol.* **2016**, *2*, 751–760. [CrossRef]

92. Pierga, J.-Y.; Bidard, F.-C.; Mathiot, C.; Brain, E.; Delaloge, S.; Giachetti, S.; de Cremoux, P.; Salmon, R.; Vincent-Salomon, A.; Marty, M. Circulating tumor cell detection predicts early metastatic relapse after neoadjuvant chemotherapy in large operable and locally advanced breast cancer in a phase II randomized trial. *Clin. Cancer Res.* **2008**, *14*, 7004–7010. [CrossRef] [PubMed]

93. Wülfing, P.; Borchard, J.; Buerger, H.; Heidl, S.; Zänker, K.S.; Kiesel, L.; Brandt, B. HER2-positive circulating tumor cells indicate poor clinical outcome in stage I to III breast cancer patients. *Clin. Cancer Res.* **2006**, *12*, 1715–1720. [CrossRef] [PubMed]

94. Budd, G.T.; Cristofanilli, M.; Ellis, M.J.; Stopeck, A.; Borden, E.; Miller, M.C.; Matera, J.; Repollet, M.; Doyle, G.V.; Terstappen, L.W. Circulating tumor cells versus imaging—predicting overall survival in metastatic breast cancer. *Clin. Cancer Res.* **2006**, *12*, 6403–6409. [CrossRef]

95. Bidard, F.-C.; Vincent-Salomon, A.; Sigal-Zafrani, B.; Dieras, V.; Mathiot, C.; Mignot, L.; Thiery, J.-P.; Sastre-Garau, X.; Pierga, J.-Y. Prognosis of women with stage IV breast cancer depends on detection of circulating tumor cells rather than disseminated tumor cells. *Ann. Oncol.* **2008**, *19*, 496–500. [CrossRef] [PubMed]

96. Camara, O.; Rengsberger, M.; Egbe, A.; Koch, A.; Gajda, M.; Hammer, U.; Jörke, C.; Rabenstein, C.; Untch, M.; Pachmann, K. The relevance of circulating epithelial tumor cells (CETC) for therapy monitoring during neoadjuvant (primary systemic) chemotherapy in breast cancer. *Ann. Oncol.* **2007**, *18*, 1484–1492. [CrossRef] [PubMed]

97. Riethdorf, S.; Müller, V.; Loibl, S.; Nekljudova, V.; Weber, K.; Huober, J.; Fehm, T.; Schrader, I.; Hilfrich, J.; Holms, F. Prognostic impact of circulating tumor cells for breast cancer patients treated in the neoadjuvant "Geparquattro" trial. *Clin. Cancer Res.* **2017**, *23*, 5384–5393. [CrossRef]

98. Wulfkuhle, J.D.; McLean, K.C.; Paweletz, C.P.; Sgroi, D.C.; Trock, B.J.; Steeg, P.S.; Petricoin, E.F., III. New approaches to proteomic analysis of breast cancer. *Proteom. Int. Ed.* **2001**, *1*, 1205–1215. [CrossRef]

99. Sawyers, C.L. The cancer biomarker problem. *Nature* **2008**, *452*, 548–552. [CrossRef]

100. Vogelstein, B.; Kinzler, K.W. Cancer genes and the pathways they control. *Nat. Med.* **2004**, *10*, 789–799. [CrossRef]

101. Zhang, J.; Guo, S.; Wu, Y.; Zheng, Z.-c.; Wang, Y.; Zhao, Y. P4HB, a novel hypoxia target gene related to gastric cancer invasion and metastasis. *BioMed Res. Int.* **2019**, *2019*, 9749751. [CrossRef]

102. Zhou, Y.; Yang, J.; Zhang, Q.; Xu, Q.; Lu, L.; Wang, J.; Xia, W. P4HB knockdown induces human HT29 colon cancer cell apoptosis through the generation of reactive oxygen species and inactivation of STAT3 signaling. *Mol. Med. Rep.* **2019**, *19*, 231–237. [CrossRef]

103. Ma, X.; Wang, J.; Zhuang, J.; Ma, X.; Zheng, N.; Song, Y.; Xia, W. P4HB modulates epithelial-mesenchymal transition and the β-catenin/Snail pathway influencing chemoresistance in liver cancer cells. *Oncol. Lett.* **2020**, *20*, 257–265. [CrossRef] [PubMed]

104. Xie, L.; Li, H.; Zhang, L.; Ma, X.; Dang, Y.; Guo, J.; Liu, J.; Ge, L.; Nan, F.; Dong, H. Autophagy-related gene P4HB: A novel diagnosis and prognosis marker for kidney renal clear cell carcinoma. *Aging* **2020**, *12*, 1828. [CrossRef] [PubMed]

105. Yang, W.; Wu, X.; Zhou, F. Collagen Type X Alpha 1 (COL10A1) Contributes to Cell Proliferation, Migration, and Invasion by Targeting Prolyl 4-Hydroxylase Beta Polypeptide (P4HB) in Breast Cancer. *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* **2021**, *27*, e928919-1–e928919-17. [CrossRef] [PubMed]

106. Li, C.; Hampson, I.N.; HAMPSON, L.; KUMAR, P.; Bernabeu, C.; Kumar, S. CD105 antagonizes the inhibitory signaling of transforming growth factor βl on human vascular endothelial cells. *FASEB J.* **2000**, *14*, 55–64. [CrossRef]

107. Tanaka, F.; Otake, Y.; Yanagihara, K.; Kawano, Y.; Miyahara, R.; Li, M.; Yamada, T.; Hanaoka, N.; Inui, K.; Wada, H. Evaluation of angiogenesis in non-small cell lung cancer: Comparison between anti-CD34 antibody and anti-CD105 antibody. *Clin. Cancer Res.* **2001**, *7*, 3410–3415.

108. Miller, D.W.; Graulich, W.; Karges, B.; Stahl, S.; Ernst, M.; Ramaswamy, A.; Sedlacek, H.H.; Müller, R.; Adamkiewicz, J. Elevated expression of endoglin, a component of the TGF-β-receptor complex, correlates with proliferation of tumor endothelial cells. *Int. J. Cancer* **1999**, *81*, 568–572. [CrossRef]

109. Nassiri, F.; Cusimano, M.D.; Scheithauer, B.W.; Rotondo, F.; Fazio, A.; Yousef, G.M.; Syro, L.V.; Kovacs, K.; Lloyd, R.V. Endoglin (CD105): A review of its role in angiogenesis and tumor diagnosis, progression and therapy. *Anticancer Res.* **2011**, *31*, 2283–2290.

110. Li, C.; Guo, B.; Wilson, P.B.; Stewart, A.; Byrne, G.; Bundred, N.; Kumar, S. Plasma levels of soluble CD105 correlate with metastasis in patients with breast cancer. *Int. J. Cancer* **2000**, *89*, 122–126. [CrossRef]

111. Kumar, S.; Ghellal, A.; Li, C.; Byrne, G.; Haboubi, N.; Wang, J.M.; Bundred, N. Breast carcinoma: Vascular density determined using CD105 antibody correlates with tumor prognosis. *Cancer Res.* **1999**, *59*, 856–861.

112. Coussens, L.M.; Werb, Z. Inflammation and cancer. *Nature* **2002**, *420*, 860–867. [CrossRef] [PubMed]

113. Hussain, S.P.; Hofseth, L.J.; Harris, C.C. Radical causes of cancer. *Nat. Rev. Cancer* **2003**, *3*, 276–285. [CrossRef] [PubMed]

114. Ohta, M.; Kawasaki, T. Complement-dependent cytotoxic activity of serum mannan-binding protein towards mammalian cells with surface-exposed high-mannose type glycans. *Glycoconj. J.* **1994**, *11*, 304–308. [CrossRef]

115. Nakagawa, T.; Ma, B.Y.; Uemura, K.; Oka, S.; Kawasaki, N.; Kawasaki, T. Role of mannan-binding protein, MBP, in innate immunity. *Anticancer Res.* **2003**, *23*, 4467–4471. [PubMed]

116. Nakagawa, T.; Kawasaki, N.; Ma, Y.; Uemura, K.; Kawasaki, T. Antitumor activity of mannan-binding protein. *Methods Enzymol.* **2003**, *363*, 26–33.

117. Holm, M.; Joenväärä, S.; Saraswat, M.; Mustonen, H.; Tohmola, T.; Ristimäki, A.; Renkonen, R.; Haglund, C. Identification of several plasma proteins whose levels in colorectal cancer patients differ depending on outcome. *FASEB BioAdvances* **2019**, *1*, 723–730. [CrossRef]

118. Ytting, H.; Jensenius, J.; Christensen, I.; Thiel, S.; Nielsen, H. Increased activity of the mannan-binding lectin complement activation pathway in patients with colorectal cancer. *Scand. J. Gastroenterol.* **2004**, *39*, 674–679. [CrossRef]

119. Swierzko, A.S.; Szala, A.; Sawicki, S.; Szemraj, J.; Sniadecki, M.; Sokolowska, A.; Kaluzynski, A.; Wydra, D.; Cedzynski, M. Mannose-Binding Lectin (MBL) and MBL-associated serine protease-2 (MASP-2) in women with malignant and benign ovarian tumours. *Cancer Immunol. Immunother.* **2014**, *63*, 1129–1140. [CrossRef]