Proteomic Analysis of Breast Cancer Subtypes Identifies Stromal Contributions that Dictate Aggressive Malignant Behavior

Jordan B. Burton[1], Philippe Gascard[2], Deng Pan[2], Joanna Bons[1], Rosemary Bai[2], Chira Chen-Tanyolac[2], Joseph A. Caruso[2], Christie L. Hunter[3], Birgit Schilling[1*], Thea D. Tlsty[2*]

1 Buck Institute for Research on Aging, Novato, CA, USA

2 University of California San Francisco, San Francisco, CA, USA

3 SCIEX, Redwood City, CA, USA

*For correspondence:

thea.tlsty@ucsf.edu, Thea D. Tlsty, PhD, 513 Parnassus Ave., HSW 513, University of California, San Francisco, San Francisco, CA, 94113, USA and bschilling@buckinstitute.org, Birgit Schilling, PhD, Buck Institute for Research on Aging, 8001 Redwood Boulevard, Novato, CA 94945, USA.

Running Title: Proteomics Reveals Stromal Role in Aggressive Breast Cancer

ABSTRACT

Breast cancer manifests as multiple subtypes with distinct patient outcomes and treatment strategies. Here, we optimized proteomic analysis of Formalin-Fixed Paraffin-Embedded (FFPE) specimens from patients diagnosed with five breast cancer subtypes, luminal A, luminal B, Her2, triple negative (TNBC) and metaplastic breast cancers (MBC), and from disease-free individuals undergoing reduction mammoplasty (RM). We identified and quantified ~6,000 protein groups (with >2 peptides per protein) with significant changes in over 26% of proteins comparing each cancer subtype with control RM. Stringent statistical filters allowed us to deeply mine 576 significant conserved protein changes shared by all subtypes and protein changes unique to each subtype. The most aggressive subtype, MBC, revealed exacerbated stromal stress responses, as illustrated by a collagenolytic extracellular matrix (ECM) and immune participation biased towards neutrophils and eosinophils. Immunostaining of breast tissue sections confirmed differences across subtypes, in particular, a dramatic upregulation of SERPINH1, neutrophil-specific myeloperoxidase and eosinophil cationic protein in MBC. In summary, deep proteomic, digitalized protein abundance profiles, generated from FFPE breast cancer tissues, revealed significant changes in ECM and cellular proteins providing insight into clinically relevant states.

INTRODUCTION

Breast cancer afflicts approximately 1 in 8 women over their lifetime in the United States (1). It is becoming increasingly apparent that improvement of patient outcomes

and customization of effective treatment plans for precision medicine will rely on a better understanding of the cellular and intricate molecular changes in tissue ecology that manifest in each breast cancer subtype. Historically, classification of breast cancer subtypes was based upon epithelial traits, i.e. expression (or not) of Estrogen Receptor (ER), Progesterone Receptor (PR) and human epidermal growth factor receptor 2 (Her2). This led to the identification of four subtypes of breast cancers: ER+/PR+/Her2- (Luminal A; LumA), ER+/PR+/Her2+ (Luminal B; LumB), ER-/PR-/Her2+ (Her2) and ER-/PR-/Her2- (Triple Negative Breast Cancer; TNBC) (2).

Although originally considered as part of the TNBC subtype because of its marker profile, Metaplastic Breast Cancer (MBC) is now considered a fifth subtype based upon its remarkable intra-tumoral heterogeneity that illustrates its unique cell plasticity (lineage promiscuity). Indeed, MBCs present with unique regions of metaplasia (not seen in other breast cancer subtypes) that can adopt either spindle, squamous, chondroid or even osseous differentiation (3). For example, MBCs can contain morphologically and immunohistochemically validated sectors of multiple tissue types, such as pancreatic, bone, and brain (Supplemental Fig. S1*A*). Mirroring its histological heterogeneity, MBC is also characterized by genetic heterogeneity that further distinguishes it from TNBC (4). Thus, MBC is highly enriched for *PIK3CA*/*PIK3R1* and *Ras-Map kinase* pathway aberrations compared to TNBC (61% vs.14% and 25% vs 7%, respectively). In addition, MBC typically harbors a high frequency of *TP53* (64%) and *TERT* promoter (25%) mutations. Complementing these findings, a recent study analyzing snap frozen tumor samples documented heterogeneous proteomic profiles for MBC depending on histological type (squamous, spindle or sarcomatoid), when

compared to disease-free breast and TNBC (5). Although paraffinized FFPE samples often pose additional challenges and difficulties for proteomic workflows compared to analyses of fresh samples, in our study we were able to analyze the five breast cancer subtypes in great analytical depth.

The epithelial distinctions that typify MBC also extend to stromal distinctions although these are currently understudied in each of the breast cancer subtypes. Notably, MBC frequently expresses immune checkpoint markers indicative of an immunosuppressive microenvironment (6). Moreover, a recent study has established that stromal heterogeneity might explain the increased incidence of MBC in the African American population (7). These recent provocative reports from Kalaw et al., 2020 (6) and Kumar et al., 2023 (7) extend the pioneering work of Finak et al., 2008 (8) who first showed that stromal gene expression could predict clinical outcome in breast cancer. Indeed, molecular profiling of microdissected stroma from primary breast tumors enabled the Park team (8) to identify a stroma-derived prognostic predictor (SDPP) that stratified disease outcome independently of standard clinical prognostic factors and published predictors. The SDPP relies on a panel of genes involved in immune, angiogenic and hypoxic responses, supporting an instructive role of the stroma in tumorigenesis. This instructive function of the stroma on epithelial cell plasticity, i.e., on the epithelial cell propensity to generate metaplastic lesions, has been documented in other metaplastic cancers, such as lung squamous cell carcinoma (9), esophageal adenocarcinoma (10) and gastric adenocarcinoma (11).

In this study, we contrasted proteomic signatures across human breast cancer subtypes, including LumA, LumB, Her2, TNBC and MBC, with disease-free breast

tissues, to identify protein signatures common to all cancer subtypes, but also signatures specific to each subtype, with a particular focus on MBC. Technologically, we demonstrate a highly efficient workflow that combined FFPE protein extraction protocols, specifically tailored and adjusted for breast tissues that exhibit high fat content, with quantitative and comprehensive proteomic data-independent acquisitions (12-15). Our goal was to gain biological insights into the contribution of the stroma (including both cellular and acellular components, such as the ECM) to the progression of these breast cancer subtypes. We identified unique (or enriched) signatures associated with the most extreme disruption in tissue homeostasis observed in MBC and validated these signatures by immunohistochemistry. Such candidate markers and signatures may, in the future, lead to potential therapeutic targets.

EXPERIMENTAL PROCEDURES

Breast Cancer Human Cohorts

Tissues were obtained from the Cooperative Human Tissue Network Western Division (Nashville, TN) and mid-Atlantic Division (Charlottesville, VA) for FFPE tissue blocks. A panel of formalin fixed paraffin-embedded human breast tissue specimen consisting of 7x ER+/PR+/Her2- (luminal A), 7x ER+/PR+/Her2+ (luminal B), 7x ER-/PR-/Her2+ (Her2+), 7x ER-/PR-/Her2- (triple negative; TN) and 7x metaplastic breast cancers (MBCs) as well as 7x disease-free reduction mammoplasty (RM) specimen used as reference, were processed for proteomics analysis. The age range of patients

for these samples was from 22 – 92 years with an average age of 57 years. All obtained human tissue samples were FFPE conserved. Further human cohort details, including sex, race, specimen weight, metastatic outcome, necrosis (%), lesion (%), stroma (%), cellularity (%), acellular (%), stage, grade, and additional comments can be found in Supplemental Table S1.

FFPE Breast Tissue Multiplex Immunohistochemistry (mIHC)

Slides with five-micron thick breast tissue sections were baked at 60°C overnight, deparaffinized in two baths of xylene(s) (10 min each) and rehydrated in graded ethanol: 100%, 100%, 95%, 85%, 70% (5 min each), and finally in distilled water. Endogenous peroxidase and autofluorescence quenching was achieved by incubating slides in PBS containing 4.5% H2O2 and 10mM NaOH under a bright light for 45 min. Heat-induced antigen retrieval was carried out in citrate buffer, pH 6.0 (Sigma-Aldrich) at 95°C for 10 min. Non-specific antibody binding was blocked using Background Sniper buffer (Biocare Medical). Tissue sections were incubated for 1 hour at room temperature (RT) with a first primary antibody (anti-SERPINH1/HSP47, Abcam, rabbit monoclonal antibody clone EPR4217, cat#ab109117, dilution 1/2000; or anti-MPO, Abcam, rabbit monoclonal antibody clone SP72, cat#ab93665, dilution 1/4000; or anti-ECP (RNase3), Cell Signaling Technology, rabbit monoclonal antibody clone E6U5M, cat#24357S, dilution 1/5000) in antibody diluent reagent (ThermoFisher, cat#50-750-38) then for 30 min in pre-diluted MACH 2 Horseradish peroxidase (HRP) labeled goat anti-rabbit micro-polymer secondary antibody (Biocare Medical, cat# RHRP520). The slides

were washed in TNT buffer (0.1 M Tris-HCl pH 7.5, 0.15M NaCl, and 0.05% Tween-20), and the signal was developed using Tyramide Signal Amplification (TSA) Plus Fluorescein solution (Akoya Biosciences, cat# NEL741001KT) for 2 min. Removal of the primary antibody complex was achieved by heating in 95°C citrate buffer pH 6.0 for 5 min. Tissue sections were blocked again with Background Sniper, incubated overnight at 4°C with a second set of primary antibodies (anti-vimentin, Cell Signaling Technology, rabbit monoclonal antibody clone D21H3, cat#5741, dilution 1/2500; or anti-ITGAM/CD11b, Cell Signaling Technology, rabbit monoclonal antibody clone D6X1N, cat#49420, dilution 1/2000) in antibody diluent reagent then for 30 min in pre-diluted MACH 2 conjugated anti-rabbit secondary antibody. After washes in TNT buffer, the signal was developed using TSA Plus Cyanine 3 (Cy3) solution (Akoya Biosciences, cat# NEL744001KT) for 3 min. Removal of the second primary antibody complex, blocking and incubation for 1h at RT with CD45 antibody (Cell Signaling Technology, rabbit monoclonal antibody clone D9M8I, cat#13917, dilution 1/2000) for the MPO module were performed as described above. The signal was developed using TSA Plus Cyanine 5 (Cy5) solution (Akoya, cat#NEL745001KT) for 12 min. After washes in TNT buffer, nuclei were counterstained with 3 µM DAPI in PBS for 5 min, washed in distilled water, and mounted with Vectashield HardSet Mounting Medium (Vector Laboratories, cat#H-1400). Slides stained with the 3 mIHC marker modules described above were imaged using a BZ-X800 fluorescence microscope (Keyence). Each multiplex-stained specimen was first scanned at low resolution using 4x magnification, and then imaged at high resolution at a 20x magnification. Unless specimens were too small, at least ten independent 20x fields, randomly distributed throughout the tumor region to account for

intratumor heterogeneity, were imaged. For sections of disease-free breast tissue, ductal and lobule containing fields were randomly selected in the tissue section. ImageJ (version 2.14.0/1.54f) was used for image preparation, and CellSeg (16), a Convolutional Neural Network trained model for nuclear segmentation, was used to facilitate nuclear mask generation and staining intensity quantification in nuclear and perinuclear areas. Parameters for CellSeg were set as follows: overlap = 80, threshold = 15, boost = auto, increase factor = 2, growth pixels = 2, growth method = 'Sequential', and should compensate = True. Mean pixel value for each protein of interest was calculated for downstream analysis. Pixel intensity thresholds to identify “positive” cells were as follows: 'SERPINH1': 36, 'Vimentin': 20, 'MPO': 32, 'RNAse3': 32, and 'ITGAM': 25. Pairwise comparisons of number of double-positive cells (SERPINH1/VIM+, MPO+/ITGAM+ or RNAse3+/ITGAM+) per mm2 of tissue area between each breast cancer subtype and disease-free breast were carried out using Prism (version 10.2.0). Two-tailed p-values were calculated using Mann-Whitney tests.

Reagents for Proteomic Analysis

Acetonitrile (AH015) and water (AH365) for HPLC were obtained from Burdick & Jackson (Muskegon, MI). Qiagen supplied the Qproteome FFPE Tissue Kit (37623). Reagents for protein chemistry purchased from Sigma Aldrich (St. Louis, MO) include heptane (34873), methanol (34860), ethanol (459844), beta-mercaptoethanol (444203), chloroform (366927), sodium dodecyl sulfate (436143), triethylammonium bicarbonate buffer 1.0M, pH 8.5 (T7408), iodoacetamide (I1149), dithiothreitol (D9779), phosphoric

acid (49685), and formic acid (94318-50mL-F). BCA protein assays (23227) and PNGase F (A39245) were obtained from Thermo Fisher Scientific (Waltham, MA). Promega (Madison WI) provided sequencing grade trypsin (V5117). S-trap mini columns were obtained from Protifi (Fairport, NY). Empore C18 resin as 47 mm disks (051115-08807) from 3M (St. Paul, MN) were used for Stage Tips generated in house (3 disk per Stage Tip). Biognosys (Schlieren, Switzerland) supplied the iRT peptides (iRT-Kit).

FFPE Breast Tissue Deparaffinization and Protein Extraction

Three 20 µm-thick serial sections were cut from a 1 cm^2 block of FFPE tissue, excess paraffin was cut from the periphery of the tissue area, and each tissue was placed into a 1.5 mL tube. The sections were deparaffinized by adding 0.5 mL of heptane and vortexing vigorously for 10 seconds, followed by incubation at room temperature for 1 hour. Next, 25 µL of methanol were added and the samples were vortexed and centrifuged at 9,000 x g for 2 minutes to pellet the tissue. The supernatant was removed and discarded using a pipet, and the pellet was air-dried for 5 minutes. Then, 100 µL of Extraction Buffer EXB Plus (Qiagen) supplemented with beta-mercaptoethanol were added to the tube containing the pellet and the mixture was vortexed before and while incubation on ice for 5 minutes (17). The sample was subsequently incubated in a Thermomixer at 100 ºC for 20 minutes, followed by incubation at 80 ºC for 2 hrs with agitation at 750 rpm. Each sample was cooled at 4 ºC for 1 minute and centrifuged at 14,000 xg for 15 minutes at 4 ºC. The supernatant,

which contained extracted proteins, was transferred to a new 1.5 mL tube. The average concentration of proteins in the supernatant was determined to be approximately 2 mg/mL using a BCA protein assay.

Aliquots of 40 µL of the supernatant were taken for additional sample cleanup and digestion (17). The samples were mixed (vortexed) with 200 µL of methanol, and centrifuged at 9000 x g for 10 seconds. Next, 50 µL of chloroform were added to the samples, which were mixed and centrifuged at 9000 x g for 10 seconds to pellet the precipitated proteins. The top layer was removed, and 150 µL of water were added to the samples, which were mixed vigorously and centrifuged at 9000 x g for 1 minute to form an interphase containing protein. This process was repeated with the addition of 150 µL of water, vortexing, and centrifuging at 9000 x g for 1 minute. The top layer was removed and discarded, and 300 µL of methanol were added to the samples, followed by vigorous mixing and centrifuging at 9000 x g for 2 minutes. The supernatant was removed from the protein pellet and discarded, and an additional 300 µL of methanol were added to the samples, which were mixed and centrifuged at 9000 x g for 2 minutes. The supernatant was again removed and discarded, and 1 mL of ethanol was added to each sample, which were mixed and centrifuged at 9000 x g for 2 minutes to pellet the protein. The supernatant was removed and discarded, and the process was repeated with the addition of ethanol. The final pellet was resuspended in 50 µL of 5% sodium dodecyl sulfate in 50 mM triethylammonium bicarbonate (TEAB) through vigorous mixing and incubated at room temperature for 1 hour.

Proteolytic Digestion

Samples were treated with 20 mM dithiothreitol in 50 mM TEAB, pH 7, at 56 °C for 10 minutes. The samples were then cooled to room temperature and allowed to sit for an additional 10 minutes. Next, the samples were alkylated with 40 mM iodacetamide in 50 mM TEAB (pH 7) at room temperature for 30 minutes in the dark. The samples were then acidified by adding 6.5 µL of 12% phosphoric acid, and a fine precipitate was formed after the addition of 500 µL S-trap buffer and mixing by inversion. The entire mixture was added to S-trap mini spin columns (Protifi) and centrifuged at 4,000 xg for 20 seconds or until fully eluted. The samples were then washed with an additional 400 µL of S-trap buffer. Next, the samples were incubated with sequencing grade trypsin (Promega) dissolved in 50 mM TEAB (pH 7) at a 1:25 enzyme:protein ratio for 1 hour at 47 °C. Additional trypsin was added at the same ratio and the proteins were allowed to digest at 37 °C overnight. The peptides were eluted from the S-trap columns and dried by centrifugal evaporation, and then resuspended in 100 µL of 50 mM TEAB (pH 7). The peptides were then deglycosylated by the addition of 3 µL 1,500 U PNGase F and incubated at 37 °C for 3 hours. The samples were brought to a final concentration of 1% formic acid in water using 10% formic acid and desalted using in-house Stage Tips. The desalted elutions were dried by centrifugal evaporation and resuspended in 20 µL of 0.2% formic acid in water. Finally, indexed Retention Time Standards (iRT, Biognosys) were added to each sample according to the manufacturer's instructions (18).

Mass Spectrometric Acquisitions, Data Processing, Quantification, and Statistics

Reverse-phase HPLC-MS/MS data was acquired using a Waters M-Class HPLC (Waters, Massachusetts, MA) connected to a ZenoTOF 7600 system (SCIEX, Redwood City, CA) with an OptiFlow Turbo V Ion Source (SCIEX) equipped with a microelectrode (19). The chromatographic solvent system consisted of 0.1% formic acid in water (solvent A) and 99.9% acetonitrile, 0.1% formic acid in water (solvent B). Digested peptides (4 µL) were loaded onto a Luna Micro C18 trap column (20x0.30 mm, 5 µm particle size; Phenomenex, Torrance, CA) over a period of 2 minutes at a flow rate of 10 µL/min using 100% solvent A. Peptides were eluted onto a Kinetex XB-C18 analytical column (150x0.30 mm, 2.6 µm particle size; Phenomenex) at a flow rate of 5 µL/min using a 120 minute microflow gradient, with each gradient ranging from 5 to 32% solvent B using the following paradigm: digested peptides were loaded at 5% B and separated using a 120 min linear gradient from 5 to 32% B, followed by an increase to 80% B for 1 min, a hold at 80% B for 2 min, a decrease to 5% B for 1 min, and a hold at 5% B for 6 min. The total HPLC acquisition time was 130 minutes. The following MS parameters were used for all acquisitions: ion source gas 1 at 10 psi, ion source gas 2 at 25 psi, curtain gas at 30 psi, CAD gas at 7 psi, source temperature at 200°C, column temperature at 30°C, polarity set to positive, and spray voltage at 5000 V.

Acquisition and Quantification by Data-Independent Acquisitions (DIA-MS)

All human samples were acquired in data-independent acquisition mode (DIA) analysis with two technical replicates for each biological sample replicate as described in detail in the online methods. Briefly, the DIA-MS method on the ZenoTOF 7600

system is comprised of A survey MS1 scan (mass range: 395-1005 m/z), with an accumulation time of 100 ms, a declustering potential of 80 V, and a collision energy of 10 V. MS2 scans were acquired using 80 variable width windows across the precursor ion mass range (399.5-1000.5 m/z), with an MS2 accumulation time of 25 ms, dynamic collision energy enabled, charge state 2 selected, and Zeno pulsing enabled (total cycle time 2.5 seconds, Supplemental Table S6).

All data files were processed with Spectronaut v16 (version 16.0.220524.5300; Biognosys) performing a spectral library using the *Homo sapiens* pan-human spectral library with 20,526 entries, accessed on 08/28/2015 (20). Dynamic data extraction parameters and precision iRT calibration with local non-linear regression were used. Trypsin/P was specified as the digestion enzyme, allowing for specific cleavages and up to two missed cleavages. Methionine oxidation (+15.995 Da), protein N-terminus acetylation (+42.011 Da), and deamidation of asparagine (+0.984 Da) were set as dynamic modifications, while carbamidomethylation of cysteine (+57.021 Da) was set as a static modification. Protein group identification (grouping for protein isomers) required at least 2 unique peptides and was performed using a 1% false discovery rate (FDR) estimated with PyProphet for peptide spectrum match, peptide and protein level. Protein quantification was based on the peak areas of extracted ion chromatograms (XICs) of 3 – 6 MS2 fragment ions, specifically b- and y-ions, with automatic normalization and 1% q-value data filtering applied. Relative protein abundance changes were compared using the Storey method with paired t-tests and p-values corrected for multiple testing using group wise testing corrections (21). The quantification significance level was as

follows: q-value less than 0.001 and absolute Log2(fold-change) greater than 0.58 when comparing breast cancer types to disease-free RM specimen.

Further Data Processing and Visualizations

For biological pathway analysis the Database for Annotation, Visualization, and Integrated Discovery (DAVID, accessed 12/02/2022) was utilized for over-representation analysis (ORA) of significantly altered quantifiable proteins ($|\log2(FC)| \geq 0.58$ & q-value < 0.001) to determine which KEGG pathway gene ontology terms were enriched in these samples (22). The ggplot2 package in R (version 4.0.5; RStudio, version 1.4.1106) was used to visualize significantly enriched biological processes from each comparison (23). Localization of Organellar Proteins by Isotope Tagging (LOPIT) maps were generated with the pRolocData and pRoloc Bioconductor packages in R (24, 25). The LOPIT map was created using the t-SNE machine learning algorithm to reduce the multi-dimensional human U2OS cell dataset and cluster proteins by similarities, overlaying the points with fold change data comparing breast cancer subtypes to disease-free tissue (24-26). Weighted gene correlation network analysis was performed using the WGCNA package in R (27). The power was set to 5 with a minimum module size of 7 proteins per module. Modules of interest were selected based on the correlation to the breast cancer subtype.

Experimental Design and Statistical Rationale

In this study, we used patient-derived FFPE tissues (n = 42) consisting of 7x luminal A, 7x luminal B, 7x Her2+, 7x TN and 7x MBC cancer tissues, as well as 7 disease-free RM specimen used as reference. Multiplex immunohistochemistry experiments were conducted with one technical replicate per biological replicate. Stained slides were imaged using a BZ-X800 fluorescence microscope (Keyence). As described above, unless specimens were too small, at least ten independent, randomly distributed 20x fields were imaged. ImageJ (version 2.14.0/1.54f) was used for image preparation, and CellSeg (16), a Convolutional Neural Network trained model for nuclear segmentation, was used to facilitate nuclear mask generation and staining intensity quantification in nuclear and perinuclear areas. Mean pixel value for each protein of interest was calculated for downstream analysis. Pixel intensity thresholds to identify “positive” cells were as follows: 'SERPINH1': 36, 'Vimentin': 20, 'MPO': 32, 'RNAse3': 32, and 'ITGAM': 25. Pairwise comparisons of number of double-positive cells (SERPINH1/VIM+, MPO+/ITGAM+ or RNAse3+/ITGAM+) per mm^2 of tissue area between each breast cancer subtype and disease-free breast were carried out using Prism (version 10.2.0). Two-tailed p-values were calculated using Mann-Whitney tests.

Proteomic experiments were conducted with two technical replicate acquisitions per biological replicate. Indexed retention time peptide standards (Biognosys) were spiked into the samples before LC-MS/MS analysis in DIA mode on a ZenoTOF 7600 mass spectrometer. One DIA cycle (2.5 s) was composed of the acquisition of one MS1 scan, followed by the acquisition of 80 variable windows (5-25 m/z) covering the full mass range (m/z 399.5-1000.5) with an overlap of 1 m/z. DIA data were processed in Spectronaut v16 using a peptide-centric approach and a panhuman library containing

149,066 precursors and 10,316 human proteins to retrieve MS2 XIC-based quantification information, (20) as described above, and significantly altered protein groups in each comparison versus disease-free RMs were obtained using a paired t test followed by p-value correction for multiple testing using the Storey method (21).

RESULTS

Tissue Cohort, Sample Preparation and Analysis

Our study utilized a comprehensive proteomic quantification, profiling and digitalization approach: data-independent acquisition (DIA) to analyze formalin-fixed paraffin-embedded (FFPE) tissue specimens representative of five distinct breast cancer subtypes (Fig. 1, *A* and *B*) which include Luminal A, Luminal B, HER-2-expressing, Triple Negative and metaplastic breast cancer (LumA, LumB, Her2+, TNBC and MBC, respectively) (12, 13, 28). Archived FFPE tissue samples are an invaluable source of clinical research material that often cannot otherwise be obtained. However, the extraction of soluble proteins from FFPE tissues remains challenging because the fixation process can result in the cross-linking of proteins, leading to low yields, protein degradation, and reduced accessibility of membrane-bound and low-abundant proteins (29). This is especially true of ECM proteins which are often extensively crosslinked. These factors can make it difficult to achieve deep proteome profiling of FFPE tissues. We have optimized extraction procedures to improve mass spectrometric acquisition as described in Materials and Methods. Briefly, FFPE specimens were deparaffinized, followed by Folch-extraction (chloroform/methanol/water extraction) prior to subsequent

proteolytic digestion using trypsin and de-glycosylation with PNGase F (Fig. 1*C*) (17). It should be noted that previously reported methods for deparaffinization were not efficient enough for our specific breast tissue samples to provide fully delipidated samples for optimal MS analysis. However, the addition of multiple Folch extraction steps dramatically improved sample quality by successfully removing any residual lipid content associated with the fatty and adipocyte rich nature of the breast tissue. As a result of the above specifically tailored protocol to extract proteins from FFPE embedded human breast tissue we were now able to identify and quantify 5,858 protein groups (with each protein identified with at least 2 unique peptides), which provided us very deep proteome coverage (for detailed results see below). The 5,858 identified protein groups were localized to multiple organelles, as illustrated by the top 15 cellular compartments in Supplemental Fig. S2**.** Indeed, the analysis of whole tissue from FFPE specimens allowed isolation of proteins that were not simply restricted to the ECM and cytoskeleton (such as when enriching for ECM as we previously have reported for lung cancer (14)), but that also extended to structural proteins (cytoplasmic, ER, Golgi apparatus and nuclear-associated proteins) as well as functional proteins (such as enzymes) typically found intracellularly.

A panel of FFPE human breast tissue specimens (Supplemental Table S1) consisted of 7 LumA, 7 LumB, 7 Her2+, 7 TNBC and 7 MBC diagnosed patients, as well as 7 disease-free reduction mammoplasty control specimens, and was processed for proteomic analysis (Supplemental Fig. S1, *B*-*G*). All samples were acquired on a high-resolution ZenoTOF 7600 mass spectrometer (SCIEX) using a 2-hour microflow HPLC gradient. Comprehensive, label-free data-independent acquisition (DIA) resulted in the

generation of a very high-quality dataset with efficient statistics when comparing the different conditions (Fig. 1*C*) (12, 13, 28). The DIA-MS data were processed in Spectronaut v16 (Biognosys) using a comprehensive Pan-Human spectral library for proteins (20), resulting in the identification and quantification of a total of 5,858 protein groups. The report of protein groups is more stringent as it combines and 'groups' protein isoforms (proteoforms), thus reducing redundancies. In addition, each protein identification required 2 or more unique quantifiable peptides abiding to stringent proteomic guidelines (Supplemental Table S2-*A*) (30). Significantly altered protein groups were determined and filtered with an absolute log2-fold change ≥0.58 and a very stringent q-value ≤0.001 (quantification confidence) when comparing each breast cancer subtype specimen with disease-free breast specimens, respectively. The dataset consisted of 7 biological replicates for each of the 6 conditions (5 breast cancer subtypes and disease-free breast tissues) accounting for a total of 42 tissue samples. Further processing and visualization of the data with volcano plots revealed that 26% of proteins significantly changed when comparing each breast cancer subtype with the control disease-free breast tissues, respectively (Fig. 2, *A*-*E* and Supplemental Table S2, *B*-*F*, cancer subtype vs. control, note: all 7 bioreplicates per each subtype were compared against all 7 bioreplicates from the disease-free controls with paired t-tests and p-values corrected for multiple testing as described in the methods). We further deeply mined the data sets and investigated similarities (Supplemental Fig. S3*A*) and unique differences (Supplemental Fig. S3*B*) observed in each of the five breast cancer subtypes (Fig. 2*F*).

Conserved Protein Profile Changes Across All Breast Cancer Subtypes

The label-free DIA-MS approach provided exhaustive proteome quantification and profiling, that revealed 576 significantly altered protein groups with conserved signatures across all breast cancer subtypes (Fig. 2*F* and Supplemental Fig. S3*A*). These shared significant protein changes between all subtypes were related to tumor growth, such as increased transcription and protein translation, changes in metabolic pathways reflecting shifts in lipid and sugar metabolism and mitochondrial function, stress responses, DNA repair and lastly and significantly, large-scale ECM remodeling (Supplemental Table S3-A, ECM signatures are described in more detail below).

Cell Intrinsic Processes: Metabolism, Morphology, Proliferation, DNA Repair

In general, biological processes that were upregulated in all breast cancer subtypes (Fig. 2*G*) reflected increased energetic and biosynthetic demands for rapid cancer cell proliferation. For example, the upregulation of lactate dehydrogenase A (LDHA), a key checkpoint of anaerobic glycolysis, in the context of a universal stress response, such as the upregulation of heat shock proteins HSP90 and HSP47/SERPINH1, typified such biological processes known to propagate pro-tumorigenic signals in all breast cancers (31, 32). These phenotypes are examples of proteomic results that reflect the biology intrinsic to epithelial cells.

Widespread Changes in Cell State Observed Across All Breast Cancer Subtypes

Many of the protein changes observed in our analysis were not confined to epithelial cells but rather extended to stromal components. For example, downregulation of alcohol dehydrogenase 1B (ADH1B) and decorin (DCN) across most breast cancer subtypes (Supplemental Table S3-A) also documented additional dramatic changes in stromal cell landscape, in this instance fibroblasts. Indeed, the loss of ADH1B has been recently associated with acquisition of a carcinoma associated fibroblast (CAF) state contributing to an interleukin 6 tumor promoting phenotype in colon cancer (33). Loss of stromal CD36, a cell surface scavenger receptor, was observed in all cancer subtypes (Supplemental Table S3-A) and typified the acquisition of a pro-tumorigenic tissue microenvironment that extends to all stromal components through a multicellular coordinated program (34, 35). In addition to altering fibroblast production of collagens and macrophage changes in efferocytosis, the loss of CD36 was recently demonstrated to disrupt vascularization (36), as well as alter lipid metabolism and associated changes in adipose tissue inflammation (37). This measurement validated our previous findings reporting a downregulation of CD36 across human breast cancer subtypes (34) and more recently in lung squamous cell carcinoma (LSCC)(14) and esophageal adenocarcinoma (EAC) (10).

Extracellular Matrix Remodeling and Signaling

Importantly, we observed shared changes in proteins that suggest alterations in ECM composition and remodeling, which are extrinsic to epithelial cells. ECM proteins

are part of a highly dynamic tissue environment. Through control of tissue stiffness, viscoelasticity and degradability, ECM proteins play a pivotal role in regulation of cell behaviors (38). Alterations in ECM have been implicated as key drivers of tumorigenesis and metastasis in all cancers. Examples include breast cancer and chronic inflammation-associated cancers of the lung, i.e. LSCC (14, 39, 40). Consistent with these observations, our proteomic analysis revealed extensive remodeling of the ECM content in the breast tumor microenvironment of all subtypes. Supplemental Fig. S4*A* highlights the 2,023 protein groups quantified with the gene ontology term "Extracellular", of which 263 proteins are also annotated in the Human Matrisome Database (MatrisomeDB (41), Supplemental Fig. S4*B* and Supplemental Table S4). Of these 263 proteins, 42 proteins were shared in their altered profile amongst the 5 breast cancer subtypes (Fig. S4*C*). The widespread significant downregulation of small leucine-rich, desmosome, and basement membrane proteins (Fig. 3, *A*-*C*) (42), like perlecan (HSPG2), desmoplakin (DSP), lumican (LUM), and a multitude of collagens (COLs), implied detrimental degradation of normal ECM composition in all breast cancers. Concurrently, upregulation of factors known to increase epithelial motility and invasion, such as vitronectin (VTN), versican (VCAN), transforming growth factor beta induced (TGFBI), and cathepsin B (CTSB), are postulated to further contribute to loss of tissue integrity and increased tumor cell dissemination. Notably, of all collagens (COL), COL12A1, a structural protein previously implicated in breast cancer metastasis through spatial mass spectrometry imaging (43), was the only collagen to be upregulated in all breast cancer subtypes in our dataset. The upregulation of cross-linking enzymes of the lysyl oxidase (LOX) family, as well as the increase in selected matrix metalloproteases

(MMPs), MMP8, MMP9, MMP14, also reflected an extensive tissue microenvironment remodeling favoring tumor cell invasion. Interestingly, it has been recently reported that cleaved collagen I activates a mitochondrial biogenesis pathway, i.e. the discoidin domain receptor 1 (DDR1)-NF-κB-p62-NRF2 signaling axis, that promotes the growth of pancreatic ductal adenocarcinoma (PDAC), while in contrast, intact collagen I triggers the degradation of DDR1 and limits the growth of PDAC, a highly desmoplastic, aggressive and metastatic cancer (44). Thus, increased degradation of collagens and MMP activation, defined as a collagenolytic stroma, may therefore contribute to greater aggressivity of some breast cancer subtypes. While MMP level changes were shared amongst the breast cancer subtypes, the observed changes were not equal and the MMP profile observed in MBC was the one most consistent with a highly aggressive clinical phenotype often observed in chronic inflammation-associated cancers (CIACs) (Supplemental Fig. S5*A*).

Stress Response and Senescence-Associated Secretory Phenotype (SASP)

Injured and diseased tissues often display stress responses characterized by a disruption of homeostasis and a senescence-associated secretory phenotype (SASP). In this context, we detected a dramatic upregulation of proteins involved in stress response and influencing ECM processing (Fig. 4). Thus, SERPINH1 (HSP47), a collagen chaperone previously identified and validated by our team as upregulated in LSCC and EAC (14), displayed a dramatic increase in levels across all breast cancer subtypes when compared to control tissues. The dramatic upregulation of SERPINH1 in

all breast cancer subtypes, observed to the greatest extent in MBC, was confirmed by immunohistochemical analysis of an extended cohort of specimens including those subjected to proteomic analysis (Fig. 5*A*), while the disease-free tissue showed almost no IHF signal at all. Furthermore, co-staining with the mesenchymal marker vimentin (VIM) confirmed that cells expressing high levels of SERPINH1 were stromal components. Quantitative analysis of VIM+SERPINH1+ cells confirmed a higher density of such cells in MBC and to a lesser extent in TNBC compared to other breast cancer subtypes and disease-free breast (Fig. 5*B* and Supplemental Table S5). Endogenous 'danger' signal molecules released by damaged and dying cells, described as damage-associated molecular patterns (DAMPs), activate the innate immune system. Our dataset identified a large set of DAMPs upregulated across breast cancer subtypes including VCAN, TNC, HSPA1A, HSPA5, HSPA8, HSP90AA1, HSP90AB1, HSP90B1 (Fig. 4, *B* and *C*). In contrast with the above-described proteins, which exhibit an upregulation when the cell experiences stress, we also observed a consistent downregulation of some proteins under the same conditions. For example, expression of CD36 (discussed above), was repressed across all breast cancer subtypes, with the most significant decrease noted in MBC tissues when compared to disease-free control (Supplemental Table S3).

Both epithelial and stromal cells can senesce prematurely in a pathological context and secrete high levels of proteases, inflammatory cytokines, immune modulators and growth factors (45). Many proteins that are upregulated and identified as stress response proteins, for example SERPINF1, overlap with those described to be secreted by senescent cells (see SASP Atlas (45)). A distinct set of SASP proteins that was

significantly altered in the breast cancer subtypes is provided in Fig. 4*C*. Together, these shared alterations fundamentally change the biochemical and biomechanical properties of the matrix and tissue, generating a tumor-supportive landscape tailored for malignant progression. Our findings underscore the vital role of abnormal signaling and ECM remodeling in driving cancer pathogenesis and highlight common tumor adaptations that support malignancy regardless of tissue origin or cancer subtype.

Identification of Protein Profile Changes Unique to Breast Cancer Subtypes

We observed that the human breast cancer subtypes each showed specific and unique protein marker signatures assessing the significantly altered proteins unique to each comparison of breast cancer subtype to disease-free tissue (Fig. 2*F*, Supplemental Fig. S3*B*. and Supplemental Table S3, *B-F*). As a first step to identify breast cancer subtype-specific signatures, we conducted a Weighted Gene Co-expression Network Analysis (WGCNA) to study biological networks based on pairwise correlations between variables. The WGCNA analysis was implemented to identify modules of highly correlated proteins, potentially sharing common biological functions or regulatory mechanisms (27). These modules were then associated with different breast cancer subtypes, elucidating biological processes specific to each subtype (Supplemental Fig. S6). The modules that showed strong correlation to one or more breast cancer subtypes were selected for further investigation (Fig. 6*A*) as described below. Differentially abundant proteins in each WGCNA module are depicted in Fig. 6, *B-D*.

WGCNA Analysis of Luminal A and Luminal B Breast Cancers

The 'pale turquoise' WGCNA module, specifically associated with LumA and LumB breast cancer subtypes (Fig. 6*B*), featured upregulated proteins implicated in diverse cellular processes. YES proto-oncogene 1, src family tyrosine kinase (YES1), participates in the regulation of critical cellular processes like proliferation, migration, and survival (46). Proteins such as slingshot protein phosphatase 3 (SSH3) and dCTP pyrophosphatase 1 (DCTPP1) are integral to the regulation of the cell cycle and promote cancer cell proliferation (47, 48). Solute carrier family 25 member 24 (SLC25A24) plays a role in cellular metabolism, i.e. mitochondrial function (49), and molecular transport. Heat shock protein family B (Small) member 1 (HSPB1) is correlated with the presence of estrogen and progesterone receptors, and plays a role in Epithelial-Mesenchymal Transition (EMT), cellular stress responses and may be involved in breast cancer metastasis (50). Dynactin subunit 3 (DCTN3) directs cellular trafficking and when upregulated plays a role in breast cancer progression (51). Cysteine rich protein 1 (CRIP1) is involved in cellular growth processes and plays a role in homologous recombination repair, increasing resistance to chemotherapy (52). Apolipoprotein O (APOO) is critical for maintaining lipid transport, and could be involved in breast cancer development (53, 54).

Three proteins in the LumA/LumB module exhibit the most dramatic changes in protein levels between Luminal A and B and the other subtypes. Microtubule associated protein tau (MAPT), a protein that is critical for maintaining cellular structure, has been

associated with better patient survival in breast cancer, and was upregulated in LumA and LumB breast cancer subtypes which typically feature a better prognosis for patients and actively downregulated in those with poor prognosis (55). A dramatic difference was observed in MAPT abundance when compared to more aggressive cancers. Particularly interesting, annexin A9 (ANXA9), was also upregulated only in LumB and LumA subtypes. ANXA9 has been associated with cell signaling, homeostasis, and protects against the immunosuppressive tumor microenvironment, improving patient prognosis (56). The protein coding for 4-aminobutyrate aminotransferase (ABAT), a positive marker for estrogen-receptor α-positive inflammatory breast cancer, is crucial for its role in cellular metabolism and molecular transport (57, 58).

WGCNA Analysis of HER2+ Breast Cancer

Two WGCNA modules, 'light cyan' and 'dark orange', (Fig. 6*C*) were specific to HER2-positive breast cancer. The 'light cyan' module featured ribosomal proteins (RPLs), essential components in protein synthesis, a process often amplified in rapidly dividing cancer cells (59). Two proteins, which were part of the ‘dark orange’ module, were the most upregulated proteins solely and specifically in the HER2 breast cancer subtype. Notably, the genes encoding these two proteins, migration and invasion enhancer 1 (MIEN1) and growth factor receptor-bound protein 7 (GRB7), are immediately adjacent to the ERBB2 gene that encodes HER2. Thus, frequent amplification of HER2 in this subtype provides a probable explanation for their increase. MIEN1 underscores the aggressive nature of HER2 positive cancers due to its

involvement in cell migration and invasion (60). The upregulation of GRB7 highlights the importance of dysregulated signaling in cancer progression due to its role in cell signaling pathways controlling cell growth and survival (61). The 'dark orange' module also included proteins such as perilipin-2 (PLIN2), associated with lipid metabolism indicative of the altered metabolic state known as the Warburg effect often seen in cancer cells (62).

WGCNA Analysis of Triple Negative and Metaplastic Breast Cancer

The 'white' WGCNA module (Fig. 6*D*), primarily consisted of proteins related to neutrophil activity, and drew particular attention due to several significant features. Neutrophil activity was reported in a prior proteomic study in fresh frozen tissues comparing MBC sub-classes to TNBC (5), where it was indicated as a characteristic MBC signature linked with immune-mediated ECM remodeling. Our FFPE tissue dataset expanded on the previous work by demonstrating that such an extensive neutrophil recruitment to the tumor site was a prominent characteristic of MBC, as it was not only less prominent in TNBC, but also not detected in the other subtypes of breast cancer (Fig. 6*D*). Such dramatic differences in neutrophil recruitment across breast cancer subtypes was validated by immunohistochemistry as discussed below. Additional proteins identified in this ‘white module’, such as proteinase 3 (PRTN3), defensin alpha 1 (DEFA1), and MMP9, have been shown to play a role in neutrophil plasticity (63). These findings suggest a 'fluid cell state' where not only the epithelial cells, but also stromal cells, such as neutrophils, exhibit plasticity (64). Notably, this

remarkable upregulation of neutrophil-specific myeloperoxidase (MPO) and eosinophil-specific eosinophil cation protein (RNAse3) in MBCs, observed to a much lower extent in other breast cancer subtypes, was confirmed by immunohistochemical analysis of the specimens subjected to proteomic analysis (Fig. 7, *A* and *B*). Furthermore, co-staining with the myeloid marker integrin alpha M (ITGAM), a protein also identified in our proteomic analysis as specifically upregulated in MBC, confirmed that most of the cells expressing high levels of MPO or RNAse3 were indeed neutrophils and eosinophils. Quantitative analysis of MPO+ITGAM+ cells and RNAse3+ITGAM+ cells confirmed a higher density of such cells in MBC and to a lesser extent in TNBC compared to other breast cancer subtypes and disease-free breast (Fig. 7*C* and Supplemental Table S5).

The 'royal blue' WGCNA module (Fig. 6*D*) encompassed proteins associated with unique aspects of MBC such as epigenetic plasticity, ectopic tissue formation, metastasis, tumor stiffness, and cellular survival under stress. The cartilage-associated protein (CRTAP) was upregulated predominantly in MBC (Supplemental Table S2), a subtype known for frequent ectopic bone tissue formation in the breast tumor and its propensity to metastasize to the bone (65). Interestingly, CRTAP has been implicated in the activation of Procollagen lysyl hydroxylase (PLOD) activities (66). Procollagen lysyl hydroxylase 2 (PLOD2) is a crucial player in the formation of fibrillar collagen in breast tissue and has been reported to regulate TGF-beta activity as well as enhance tumor stiffness (67), thereby promoting mesenchymal cancer cell phenotypes and stemness (68). PLOD2 upregulation has also been shown to correlate with tumors with poor outcome characterized by an extensive immune cell infiltration, in particular eosinophils, a distinct MBC phenotype as discussed below (69). Additionally, PLOD2, through

interactions with alpha-ketoglutarate, can generate succinate which can subsequently suppress TET proteins and disrupt DNA demethylation. Another protein of interest was insulin-like growth factor binding protein-3 (IGFBP3), a protein known to stimulate survival and autophagy in breast cancer cells experiencing glucose deprivation and hypoxia (70), two common conditions in the tumor microenvironment. The collagen triple helix repeat-containing 1 (CTHRC1), a pro-metastatic protein, was found to have a positive correlation with the abundance of MMP7 and MMP9, thereby promoting tumor migration, motility, and invasion, particularly in non-small cell lung cancer (71). Furthermore, complement C1r subcomponent-like protein (C1RL), as part of the immune response, is activated during chronic inflammation (72).

The 'green' WGCNA module (Fig. 6*D*) was comprised of a diverse array of proteins with the highest upregulation in MBC. These proteins included the SH3 domain binding glutamate-rich protein like 3 (SH3BGRL3), which is involved in cell growth and apoptosis signaling pathways (73), and the quaking protein, KH domain RNA binding (QKI), which is crucial for post-transcriptional mRNA processing and cell differentiation (74). Pentraxin 3 (PTX3) is part of the innate immune system and is known to influence several aspects of cancer biology, including tumorigenesis and breast cancer metastasis to bone (75). The procollagen-lysine 2-oxoglutarate 5-dioxygenase 3 (PLOD3) is involved in collagen formation, a key component of the ECM, and is associated with tumorigenesis and metastasis (76). Cathepsin B (CTSB) is a cysteine protease involved in protein breakdown associated with various pathological states such as cancer and inflammation (77), and carboxypeptidase, vitellogenic-like (CPVL), contributes to protein processing and degradation. Of note, CPVL has been shown to

facilitate resistance to CDK4/6 inhibitors in breast cancer (78). Furthermore, the simultaneous upregulation of CPVL, fatty acid binding protein 3 (FABP3), integrin subunit alpha 5 (ITGA5), orosomucoid 1 (ORM1), ORM2, ribosomal protein S19 (RPS19), and SERPINE2 and downregulation of transmembrane protein 205 (TMEM205) evoked heightened immunosuppression in MBC through polarization of macrophages toward a pro-tumorigenic M2 phenotype and/or through proliferation of regulatory T cells. Finally, the actin beta like 2 (ACTBL2) protein maintains cell structural integrity and facilitates processes such as motility and division (79).

Metaplastic Breast Cancers (MBC) Display a Unique Tissue Landscape Typifying their Lineage Promiscuity

Strikingly, we discovered that MBC featured twice as many unique differentially changing signature proteins, in contrast to the other cancer subtypes that also showed unique signatures but to a much lesser extent, this was true even for TNBC. Overall, MBC featured 181 unique proteins that were upregulated and 168 unique proteins that were downregulated in MBC specifically (Fig. 2*F*). The identification of a large fraction of unique significant changes in protein levels in MBC (Supplemental Table S3-*B*) was anticipated (and now confirmed) based on the histological heterogeneity and the emergence of ectopic tissues or metaplasia (bone, muscle, etc.) in these rare, poorly differentiated, and particularly aggressive breast cancers (Supplemental Fig. S1). This unique plasticity in the breast for MBC was illustrated by the upregulation of previously defined ‘stemness’ proteins, such as adenosine monophosphate deaminase 2

(AMPD2), cluster of differentiation 109 (CD109), phosphoribosyl pyrophosphate amidotransferase (PPAT), and phosphoserine aminotransferase 1 (PSAT1), as well as downregulation of proteins, such as ATPase H+ transporting V0 subunit a1 (ATP6V0A1), a phenotype previously observed upon abnormal brain differentiation associated with lysosomal dysfunction(80) (Supplemental Fig. S3*C*). The suggested 'stemness' signature of MBCs was further exemplified by unique upregulation of proteins typically expressed in other tissues but not usually in breast tissue. Examples of such proteins included proteins reminiscent of osteocytes, i.e. bone-producing cells, such as osteolectin (CLEC11A). The upregulation of bone formation-associated proteins, specifically in MBC, aligned with its known propensity for ectopic bone (and cartilage) formation in the breast (81, 82) and bone dissemination (65). Additionally, MBC showed high keratin 17 (KRT17) level and anterior gradient 3 (AGR3) level, two phenotypes previously reported to be more frequent in TNBC and to be associated with worse outcome (83). This latter feature was likely an illustration of a squamous subtype of MBC. The striking upregulation of branched-chain-amino-acid aminotransferase (BCAT1) only observed in MBC (Supplemental Table S3*B*) further illustrated the complex cross-talk between epithelial cells and fibroblasts contributing to the acquisition of a CAF phenotype (84) and of a microenvironment favoring EMT and stemness (85). Finally, the extensive downregulation of selenium-binding protein SELENBP1 in MBC may also reflect a loss of cell differentiation. Previous reports demonstrate induction of SELENBP1 in colon cancer cells triggers colonocyte differentiation (86). A heightened EMT profile, characteristic of tissues with high plasticity, and previously reported as a major phenotype in MBC (5), could be validated in our FFPE specimen analysis.

Indeed, 290 of the 1153 'EMTome' proteins (87) were captured in our proteomic analysis. Seventeen proteins were upregulated and 13 proteins downregulated specifically in MBC. The most consistent of these MBC-specific changes included upregulation of CD44, NDRG1, SERPINE2 and TGFBI and downregulation of AP1M2, SELENBP1 and STAT5A. The detection of such a wide array of unexpected proteins in breast, unique to MBC, highlighted the exquisite adaptability of cancer cells in a quite unique tumor microenvironment.

MBC Display a Unique Tissue Landscape Characterized by an Extensive Deposition of Proteins Reflecting Chronic Inflammation and Involvement of Coagulation

Besides upregulated proteins involved in cell proliferation, cell survival and cell migration/invasiveness biological processes, the MBC-specific signature identified many upregulated proteins involved in chronic inflammation, fibrosis, and immune response (Supplemental Fig. S7, *A* and *B*). Thus, we observed upregulation of a panel of proteins indicative of chronic inflammation (Supplemental Table S3*B*): serum amyloid A-4 protein (SAA4), alpha-2-HS-glycoprotein (AHSG), tumor necrosis factor alpha-induced protein 8 (TNFAIP8), transthyretin (TTR), typically expressed in liver, choroid plexus, pancreas and kidney, leukosialin (SPN), typically expressed in lung, skin and bone marrow, plasminogen (PLG) typically expressed in liver and kidney; complement proteins (C4BPA, C4BPB and C8B) typically expressed in liver, lung and/or, testis; and afamin (AFM) typically expressed in liver, testis and kidney. Upregulation of afamin (AFM) in MBC strongly supported stabilization of Wnt ligands and as a result heightened Wnt

signaling through this unique mechanism. The upregulation of transforming growth factor-beta-induced protein ig-h3 (TGFBI) also reflected engagement of a key node in fibrosis pathways. Additional markers typifying (chronic) inflammation included upregulation of agrin (AGRN), interleukin 18 (IL18), protein arginine methyltransferase 1 (PRMT1), SERPINA3, SERPINA4, SERPINB8, SERPINB9, SERPIND1, SERPINF2, and Thimet oligopeptidase 1 (THOP1), as well as downregulation of repressors of inflammation and fibrosis, such as mediator complex subunit 23 (MED23). It was intriguing to observe that MBC-specific protein signatures were indicative of activation of complement (C) or coagulation factors (F), as illustrated by upregulation of C7, CFI, F2, F9, F12, F13A1, kininogen 1 (KNG1), plasminogen (PLG), a biological function previously reported as enriched in MBC (5). It has recently emerged that activation of this coagulation cascade is not only leading to platelet, neutrophil and endothelial cell activation but also promotes cancer cell growth and invasion (88).

A significant number of the protein level changes observed may be attributable to increased MYC activity. Indeed, a myc signature has been identified in MBC (5) but also in (metaplastic) lung squamous cell carcinoma but not in (non-metaplastic) lung adenocarcinoma (89). A systematic search of the proteins captured in our proteomic analysis among a 198 'HALLMARK MYC TARGET' published database (90) identified 122 'hallmark myc target' proteins. Five of these proteins were upregulated (EIF3I, HNRNPA2B1, PPIA, PSMA1 and UBA2) and 1 protein downregulated (CSTF2) specifically in MBC. The most consistent MBC-specific change was the upregulation of peptidylprolyl Isomerase A (PPIA).

Breast Tissue Injury and Aging leads to Dysregulation of Stromal Proteins and Cancer Progression

Cancer progression and tumorigenesis is facilitated and perhaps even driven by disruptions to the ECM that lead to the unrestrained release of soluble factors and changes in mechanotransduction, thus creating a chronically inflamed microenvironment that favors cancer progression and increased senescence burden (Fig. 4*A*). A graphical display of MBC vs disease-free significant protein changes was visualized using a map of proteins localized to organelles (Supplemental Fig. S7*C*) (24, 26). This display indicated significant changes in levels of proteins involved in regulation of mitochondria, peroxisome, endoplasmic reticulum, and proteasome function. A predominant downregulation of mitochondrial and peroxisomal proteins would support reduced energy production (91, 92), while upregulation of endoplasmic reticulum and proteasome proteins would correlate with increased signaling and demand to degrade proteins as the cells enter a stress mode associated with tumorigenesis (93, 94). These changes, specifically observed for MBC, highlighted characteristics of a disrupted tissue microenvironment in metaplastic cancers that might contribute to the unique lineage promiscuity and cell plasticity, and importantly, to the heightened aggressiveness of these rare cancers. In our investigation of the coping mechanisms of cancer cells amidst various neoplastic stresses, we observed notable alterations in endoplasmic reticulum (ER) proteostasis (Fig. 4*B* and Supplemental Fig. S5*B*). Proteostasis relies on regulated protein translation, proper protein folding by molecular chaperones, and activation of protein degradation pathways such as autophagy and proteasome-

mediated degradation. The fine tuning of these mechanisms in response to homeostatic variations or stress conditions is pivotal to the maintenance of an appropriately folded proteome. This tuning relies on proper and timely engagement of cellular stress response pathways - heat shock response (HSR) as well as mitochondrial and ER unfolded protein responses (UPRs) (95). This is a dual-edged phenomenon that can promote cell adaptation and survival under moderate stress but initiate instead cell suicide under persistent stress (93). The upregulation of various ER proteins in breast cancer subtypes (Fig. 4*B*) compared to disease-free tissue underscores the importance of ER proteostasis in neoplastic stress responses, highlighting the importance of ER-based processes in both the promotion and suppression of tumorigenesis. The upregulation of flap endonuclease 1 (FEN1), in conjunction with the dramatic downregulation of double strand break repair protein (RAD50, the most downregulated protein in the MBC dataset; -6.15) and of inosine triphosphatase (ITPA), further illustrated profound defects in DNA damage repair expected in stressed tissues upon injury or aging.(96)

DISCUSSION

The generation of these in-depth comparative proteomic profiles of five major human breast cancer subtypes and disease-free human breasts was the result of the use of modern mass spectrometry technology in combination with the distinctly optimized FFPE sample processing through extensive sample deparaffinization and de-lipidation, that became necessary due to the high adiposity of breast tissues. The optimization of

such efficient quantitative proteomic workflows for FFPE cancer tissues opens access to almost 'limitless' archived FFPE tissues from biobanks. This becomes particularly relevant when conducting longitudinal studies using retrospective cohorts with long term patient follow-up to assess, for example, disease progression or relapse for risk stratification purposes or response to therapy. This rich dataset identified both common and unique protein signatures underlying general and specific aspects of tumor pathogenesis across the breast cancer subtype spectrum. The fact that these FFPE samples were not enriched for specific cellular or acellular (ECM) tissue components allowed us to query a wide range of biological phenotypes. Furthermore, while some of our observations confirmed and extended some epithelial phenotypes, they also extended our insights into key stromal phenotypes, both in terms of cellular and acellular tissue components, i.e. stromal cells and ECM, respectively.

Thus, we have delineated across all breast cancer subtypes conserved upregulation and downregulation of proteins that support: (1) increased proliferation typified by higher levels of proteins involved in transcription and translation, (2) dramatic switches in lipid and sugar metabolism as well as mitochondrial function, and (3) extensive tissue remodeling associated with stress response and differentiation. Notably, although some of these changes were shared among all subtypes, a significant number of them showed a gradual increase from more differentiated breast cancers (luminal A and B) to less differentiated breast cancers (TNBC and MBC). This observation afforded us with the ability to highlight malignant events whose extent of dysregulation seemed to correlate with tumor aggressiveness. A more granular analysis of the various breast

cancer subtypes identified protein signatures unique to each subtype. The least differentiated breast cancer, MBC, displayed several striking phenotypes.

First, in MBC, we observed a dramatic remodeling of ECM illustrated by an extensive loss of ECM proteins: downregulation of collagens, in particular collagens 4, 6, 15 and 18, as well as small leucine rich protein decorin. A loss of desmosome proteins (desmocollins, desmogleins and desmoplakin) illustrated loss of cell-cell contacts that was going hand in hand with an extensive EMT signature. Indeed, just as striking was the dramatic upregulation of ECM proteins known to foster a pro-tumorigenic environment including periostin, fibronectin, versican and tenascin (97-99). The upregulation of fibromodulin, specifically restricted to MBC, was noteworthy as its upregulation has been linked to breast cancer metastasis (100). Furthermore, this phenotype is in accordance with increased Wnt-beta catenin signaling and is a mark of chronically inflamed tissue. The upregulation of fibromodulin can be inhibited by anti-inflammatory drugs such as aspirin. Notably, we observed a dramatic and specific upregulation of pivotal ECM-modifying proteins, such as MMP9, a protease widely implicated in tumorigenesis, as well as neutrophil collagenase MMP8. The dramatic upregulation of such ECM-proteolytic enzymes specific to MBC had far reaching consequences, not only leading to an extensive loss of ECM architecture but also more broadly contributing to a loss of tissue homeostasis and to the acquisition of a particularly aggressive and immunosuppressed tumor microenvironment. The upregulation of PLOD proteins, maximal for PLOD1 and PLOD3 and unique for PLOD2 in MBC, and the ensuing extensive collagenolysis have been recently related to promoting desmoplastic cancers with very poor patient outcome (69). The upregulation

of BCAT1 was another example of an epithelium-stroma cross-talk leading to acquisition of a CAF phenotype associated with cancers with heightened EMT and cancer cell stemness (84, 85).

Our analysis identified profound changes affecting multiple stromal tissue components. For example, the loss of ADH1B illustrated a shift in fibroblast identity towards a CAF phenotype (33). The loss of CD36, previously reported by us across all breast cancer subtypes (34), reflected a shift in vasculature identity characterized by a loss of physiological vasculature architecture and identified by the detachment of a recently characterized capillary population involving pericytes (36). The upregulation of myosin 9 (MYH9) and the myofibroblast-specific myosin 11 (MYH11) further exemplified the acquisition of a highly EMT phenotype in breast cancers (101). Lastly, the loss of several mast cell-specific proteins (chymase (CMA1) and carboxypeptidase A (CPA3)) further documented these dramatic changes in tissue ecology. These relationships unveil the too long ignored proclivity of an altered stroma to dictate epithelial tumor cell fate. We document below the remarkable changes observed in immune tissue landscape.

The importance of the immune (stromal) environment first indicated by Finak et al. (8), has revolutionized classification of tumors and associated patient outcome. Highlighting the importance of this tissue ecology aspect, our second observation was that MBC was characterized by a dramatically immunosuppressed tumor environment. Indeed, while alterations in immune response were seen across all 5 subtypes of breast cancer, this phenotype exhibited the most variation that was unique to each subtype. The concomitant upregulation of EPX, RNASE3, DEFA1, MPO, PRTN3, ELANE and

ITGAM, identified through proteomic analysis (Supplemental Table S3-*B*) and WGCNA (Fig. 6*D*), documented recruitment of eosinophils and neutrophils, as a phenotype dramatically exemplified in the context of the MBC tissue microenvironment (102, 103). Due to its unique mode of action relying on the use of hydrogen peroxide to form different hypohalous acids leading to the production of reactive oxygen and reactive nitrogen species, uncontrolled or sustained release of MPO would exaggerate inflammation and contribute to oxidative stress-associated tissue damage even in absence of inflammation (104). We have previously reported upregulation of such biomarkers in other human CIACs, i.e., esophageal (10) and lung (14). It therefore appears that recruitment of specialized immune cells is a prominent phenotype of CIACs shared across tissue types, i.e., metaplastic breast, lung, and esophageal cancers. We have also recently identified eosinophil and neutrophil signatures (EPX and MPO upregulation) in a mouse model of colitis, highlighting the importance of such biological changes as early events in chronic inflammation-driven tumorigenesis (105). Previous findings in oral cancer models directly linked tumor-associated neutrophils to metastatic spread (63). Furthermore, neutrophils have been shown to play a key role in maintenance of metaplastic phenotypes in a murine model of lung metaplastic carcinoma, i.e. lung squamous cell carcinoma (9). Relevant to this study, MPO-derived oxidants have been involved in inducing premature cell senescence (106). Additional MBC-specific proteomic signatures further evoked heightened immunosuppression through polarization of macrophages toward a pro-tumorigenic M2 phenotype (typified by a loss of CD36 level) and/or through proliferation of regulatory T cells. Of particular novelty and introducing a new concept, recent evidence has pointed to the activation of

neutrophil extracellular traps by SASP factors, which further promotes cancer cell proliferation, invasion, and metastasis in contrast to the traditional view of neutrophils killing pathological cells (107, 108). Our data further supported this hypothesis as SASP and neutrophilic proteins that give rise to neutrophil extracellular traps showed the highest abundance in MBC.

Third, another major trend was the more extensive deposition of proteins reminiscent of chronic inflammation, a key driver of tumorigenesis for metaplastic cancers. Indeed, we observed upregulation of a large panel of proteins indicative of chronic inflammation (AGRN, AHSG, IL18, PRMT1, SAA4, SPN, TNFAIP8, TTR) and complement proteins (C4BPA, C4BPB and C8B). The upregulation of some of these proteins suggested dysregulated signaling: sustained Wnt signaling associated with AFM upregulation or TGF-beta signaling associated with upregulation of TGFBI. A striking observation was the upregulation of serpins. We extensively discussed the upregulation of SERPINH1, but other serpins were upregulated as well. This pro-inflammatory signature was exemplified by the downregulation of repressors of inflammation and fibrosis (MED23). An intriguing observation was a MBC-specific protein signature indicative of deposition of proteins involved in activation of complement (C7, CFI) or coagulation factors (F2, F12), kininogen 1 (KNG1) and plasminogen (PLG). As discussed above, activation of this coagulation cascade is not only leading to platelet, neutrophil and endothelial cell activation but also contributes to cancer cell growth and invasion (88).

Lastly, our study extended the recent concept that (breast) cancer evolves as an accelerated process of tissue aging (109, 110). Such ‘biological’ aging likely results from

a response to stress exemplified by chronic inflammation driving tumorigenesis in the case of metaplastic cancers. Aging-associated alterations in the ECM architecture and remodeling include loss of basement membrane components, increased collagen cross-linking and stiffness. We detected a dramatic upregulation of a large number of proteins associated with a senescence (SASP) phenotype (45). The risk and incidence of breast cancer increases with age (1), as does the accumulation of senescent cells, i.e., of cells that are in a state of permanent cell cycle arrest (111). Senescent cells are known to secrete inflammatory molecules and cytokines into the ECM (45, 112), where these senescent associated secretory phenotype (SASP) factors promote increased tissue stiffness and a tumor-prone microenvironment (112). In our study, activation of senescence and inflammatory pathways was evident across breast cancer subtypes through conserved upregulation of numerous SASP factors (Fig. 4*C*). Further extending a stress response signature in breast cancers, we observed notable alterations in endoplasmic reticulum (ER) proteostasis and an upregulation of many damage-associated molecular patterns (DAMPs) including ECM proteins (VCAN, TNC) but also several heat shock proteins (HSPA1A, HSPA5, HSPA8, HSP90AA1, HSP90AB1 and HSP90B1). Upregulation of one of these DAMPs, CD44, along with SASP-Associated candidate, SERPINF2 was only observed in MBC. The consistent overexpression of proteins like SERPINH1, POSTN, TNC, and FN1 across breast cancer subtypes, also implicated as pro-tumorigenic drivers in LSCC (14), implied the common engagement of local and systemic SASP programs as one of the programs promoting cancer progression. Relevant to this discussion, engagement of such proteins influences not only immune function but also ECM processing (113).

The resulting cell secretome likely acts both locally to remodel the ECM and distally to spread inflammatory signals, thus contributing to sustained matrix stiffening, chronic inflammation and more broadly a tumor-supportive microenvironment. The role of ECM alterations is also closely related with aging, as ECM stiffening with age presents additional prevalence for cancer invasion (39). Changes in tissue (stromal) architecture can already be identified at a pre-malignant stage and the extent of these changes can be used as a tool for risk stratification for developing subsequent breast cancer (36). Interestingly, in many breast cancers, epithelial cells show a striking loss of lineage fidelity. Thus, luminal epithelial cells start to lose their lineage fidelity, and begin to express markers typically associated with myoepithelial cells (110). This fidelity to cellular identity and function is compromised during the aging process and cancer development, as DNA damage incites inflammation and disruption of tissue homeostasis (109). Inflammatory processes gradually undermine the homeostatic cues maintaining cellular identity, leading to loss of cellular differentiation and further disarray in cellular function.

Our findings provide a foundation to further elucidate mechanistic events linking loss of cell identity and function observed in cancer to chronic inflammation, stress response and accelerated (chronological) aging (109). Knowledge of the molecular and pathological differences between the various breast cancer subtypes studied here, not only focusing on the epithelial component but also taking into consideration cellular and acellular stromal components, is critical for the development of therapies customized to each breast cancer subtype. However, some limitations should be noted. Additional patient cohorts will be needed to further validate our findings. Interestingly, despite the

relatively small human cohort size, dramatic protein signatures emerged throughout all cancer subtypes – common and unique – with remarkable statistical significance for many cancer-relevant proteins and pathways. Moving forward, high-priority foci for future work include: i) immunohistochemistry-based validation of differential levels of subtype protein markers in larger/independent cohorts, ii) elucidation of mechanisms via functional studies, and iii) exploration of biomarker applications to design therapeutic regimens targeting some of the pathways highlighted in this study. The development of subtype-specific protein signatures into clinically viable tests could enable better prognostication and therapy selection. Further exploration of therapeutics designed to target conserved metabolism/ECM factors or unique SASP/metastasis modules may improve treatment responses. With further mechanistic clarification and clinical validation, our proteomic findings may ultimately translate into personalized management strategies tailored to each patient's tumor profile.

DATA AVAILABILITY

The raw data and complete mass spectrometry data sets have been uploaded to the Mass Spectrometry Interactive Virtual Environment (MassIVE) repository, which is maintained by the Center for Computational Mass Spectrometry at the University of California San Diego. These data can be accessed and downloaded using the following link: https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?task=bfbb9bbd82244bd0810b6c5809db91ca (MassIVE ID number: MSV000094056; ProteomeXchange ID: PXD049314). The R scripts used in this study are available through our GitHub repository

(https://github.com/JoBBurt/Human-FFPE-Breast-Cancer-Subtypes) and can be cited using the DOI: https://doi.org/10.5281/zenodo.10651533. [Note to the reviewers: To access the data repository MassIVE (UCSD) for MS data, please use: Username: MSV000094056_reviewer; Password: winter].

Supplemental data – This article contains supplemental data.

Acknowledgments – FFPE tissue blocks used for proteomic and mIHC analyses were obtained from the Cooperative Human Tissue Network Western and Mid-Atlantic divisions.

Funding and additional information – This work was supported by National Institute of Aging (NIA) Award U01 AG060906 (to BS), the Anna Fuller Foundation (to BS), National Cancer Institute (NCI) Award R50 CA211543 (to PG), and NCI Outstanding Investigator Award R35 CA197694 (to TDT). The authors acknowledge the very generous support from SCIEX for the ZenoTOF 7600 system and a Waters M-class HPLC system at the Buck Institute.

Author contributions – J.B.B., B.S., P.G. and T.D.T. manuscript writing; J.B.B., D.P., R.B., C.C.-T. and J.A.C. methodology; J.B.B. and J.B. investigation; J.B.B., J.B. and B.S., formal analysis; J.B.B., B.S. and P.G. data curation; P.G., B.S., and T.D.T conceptualization; B.S. and T.D.T. supervision; C.L.H, B.S. and T.D.T. resources; P.G. project administration; B.S., P.G. and T.D.T. funding acquisition.

Conflicts of interest – Dr. Birgit Schilling is an editorial board member for this journal and was not involved in the editorial review or the decision to publish this article. Dr. Christie L Hunter previously was an employee of SCIEX.

Abbreviations - The abbreviations used are: DIA, data-independent acquisition; ECM, extracellular matrix; ER, Estrogen Receptor; FDR, false discovery rate; FFPE, Formalin-Fixed Paraffin-Embedded; GO, gene ontology; Her2, human epidermal growth factor receptor 2; LumA, Luminal A; LumB, Luminal B; MBC, metaplastic breast cancers; MS/MS, tandem MS; mIHC, Multiplex Immunohistochemistry; PR, Progesterone Receptor; RM, reduction mammoplasty; SASP, Senescence-Associated Secretory Phenotype; TNBC, triple negative breast cancer

FIGURE LEGENDS

FIG. 1. Breast cancer subtypes and workflow for proteomic analysis of human breast tissue FFPE specimens. *A*, Breast cancer subtypes exhibit a continuum from

well differentiated to poorly differentiated phenotype, *adapted from Lim et. al. Nat. Med. 2009 (2).* The 5-year survival rates are 55% for metaplastic breast cancer (MBC), 77% for triple negative breast cancer (TNBC), 83% for HER2 positive breast cancer (HER2), 92% for luminal B breast cancer (LumB), and 89% for luminal A breast cancer (LumA) (1). *B*, Showcase of one hematoxylin and eosin stained FFPE tissue for representative sample subtypes, including a reduction mammoplasty (RM) control. *C*, A workflow for processing and acquiring proteomic data from human breast FFPE tissue is shown. Briefly, 42 FFPE samples were obtained with known pathologies. Tissue scrolls were collected and deparaffinized before protein extraction. Folch extracted proteins were digested with trypsin and deglycosylated with PNGase F.

FIG. 2. Breast Cancer Subtype Proteomes reflect dramatic tissue remodeling. Volcano plots of the 5,858 protein groups identified with 2 or more unique peptides show proteome remodeling for comparisons to disease-free tissue. Significantly changing proteins are indicated in red (up-regulated) or blue (down-regulated): significantly altered proteins were selected with strict q-values ≤ 0.001 and absolute fold changes at |Log2(FC)| ≥ 0.58; non-significantly changing proteins are indicated in grey. *A*, Luminal A (ER+PR+Her2-) breast cancers showed 1,805 proteins with significantly altered levels: with 1,010 downregulated proteins (blue), and 795 upregulated proteins (red). Of these significantly changing proteins 244 proteins (13.5%) only changed significantly in the luminal A subtype, but not in the other subtypes. *B*, Luminal B (ER+PR+Her2+) breast cancers showed 1,548 significantly altered proteins with 730 downregulated proteins (blue), and 818 upregulated proteins (red). Of these significantly

changing proteins 169 proteins (10.9%) only changed significantly in the luminal B subtype, but not in the other subtypes. *C*, Her2+ (ER-PR-Her2+) breast cancers showed 1,851 significantly altered proteins (including 148 unique proteins; 8%) with 961 downregulated proteins (blue), and 890 upregulated proteins (red). Of these significantly changing proteins 148 proteins (8.0%) only changed significantly in the Her2+ subtype, but not in the other subtypes. *D*, TNBCs (ER-PR-Her2-) showed 1,861 significantly altered proteins with 996 downregulated proteins (blue), and 867 upregulated proteins (red). Of these significantly changing proteins 166 proteins (8.9%) only changed significantly in the TNBC subtype, but not in the other subtypes. *E*, MBCs showed 1,624 significantly altered proteins (including 349 unique proteins; 21.5%) with 832 downregulated proteins (blue), and 792 upregulated proteins (red). Of these significantly changing proteins 349 proteins (21.5%) only changed significantly in the MBC subtype, but not in the other subtypes. *F*, Venn diagram showing the overlap of 576 significantly altered protein groups that changed in comparisons of all subtypes vs control, and significantly altered protein groups that are unique to each cancer subtype (breast cancer subtype vs. disease-free). *G*, Dotplot of the upregulated KEGG biological processes with a q-value ≤ 0.01 and term level ≥ 5 in all breast cancer subtype comparisons with disease-free tissue.

FIG. 3. Breast cancers display alterations in tissue microenvironment characterized by quantitative and qualitative changes in ECM proteins*. A*, Heatmap of the significantly altered small leucine-rich proteins. *B*, Heatmap of the significantly altered desmosome proteins. *C*, Heatmap of significantly altered basement

membrane proteins. The fold-change scale is displayed for all heatmaps where the positive $\log_2$(FC) values are shown in red and the negative values in blue.

FIG. 4. Breast Cancer Subtypes Show an Upregulation of SASP Markers Compared to Disease-Free Samples. *A*, biological overview of the transition from normal tissue to disrupted tissue with increased markers of stress and cellular senescence due to acute inflammation to invasive carcinoma as acute inflammation becomes chronic inflammation. *Adapted from Luthold et al. Cancers, 2022* (114). *B*, Heatmap showing upregulation of stress-response markers from the endoplasmic reticulum. *C*, Heatmap showing higher expression of selected SASP markers in breast cancer subtypes with the worse 5-year survival. Note that all SASP markers are upregulated in the secretome of irradiated IMR 90 cells. Proteins shown have a q-value ≤ 0.001, positive fold change values are displayed in red and negative fold change values are displayed in blue.

FIG. 5. Stromal signature characterized by upregulation of SERPINH1 is exemplified in the least differentiated breast cancers. *A*, Tissue sections from disease-free reduction mammoplasties (RMs) and five breast cancer subtypes (Luminal A (LumA), Luminal B (LumB), Her2-positive (Her2), triple negative (TNBC) and metaplastic (MBC) were sequentially immunostained as described in Methods for collagen chaperone *SERPINH1* detected with FITC and the stromal marker *vimentin* detected with Cy3. Stained slides were counterstained with DAPI and mounted prior to imaging. Representative images (at 20 x magnification) for one specimen of each of the

6 tissue specimen groups described above are shown. Scale bar: 0.1mm. *B*, Quantitative analysis of chaperone SERPINH1. Images of immunostained breast tissue sections were subjected to quantitative analysis (Table S5) as described in Methods. Numbers of SERPINH1+VIM+ cells (fibroblasts) per mm2 of analyzed tissue specimen are displayed for disease-free breast (RM; n=13) and each breast cancer subtype (LumA; n=20, LumB; n=9, Her2; n=8, TNBC; n=11 and MBC; n=7). Degree of significance for each pair comparison (**, ****) are displayed at the top of the panel. P-values: LumA vs. RM: <0.0001; LumB vs RM: <0.0001; Her2 vs. RM: <0.0001; TNBC vs. RM: <0.0001; MBC vs. RM: p=0.0018. Parameters used to generate the graph are shown in Supplemental Table S5.

FIG. 6. Protein modules of interest identified by Weighted Gene Correlation Network Analysis (WGCNA) define breast cancer subtype-specific proteomic signatures. *A*, WGCNA correlation analysis of select protein modules to breast cancer subtype is displayed. Heatmap of selected significantly altered protein groups from modules defined in WGCNA that are significantly different for *B*, Luminal A and B; *C*, HER2-positive; and *D*, TNBC and Metaplastic breast cancers are shown. The pale turquoise WGCNA module proteins are upregulated signatures for LumA and LumB. The light cyan and dark orange WGCNA module proteins are upregulated in HER2. The WGCNA modules white, royal blue, and green proteins are upregulated in MBC.

FIG. 7. Stromal signatures characterized by upregulation of Myeloperoxidase and RNAse3 are exemplified in the least differentiated breast cancers. Tissue sections

from disease-free reduction mammoplasties (RMs) and five breast cancer subtypes (Luminal A (LumA), Luminal B (LumB), Her2-positive (Her2), triple negative (TNBC) and metaplastic (MBC)) were sequentially immunostained as described in Methods for *A*, neutrophil-enriched *myeloperoxidase* (MPO) detected with FITC, the myeloid marker *ITGAM/CD11b* detected with Cy3 and the immune cell marker CD45 detected with Cy5; or *B*, RNAse3/eosinophil-specific *Eosinophil cationic Protein* detected with FITC and the myeloid marker *ITGAM/CD11b* detected with Cy3. Stained slides were counterstained with DAPI and mounted prior to imaging. Representative images (at 20 x magnification) for one specimen of each of the 6 tissue specimen groups described above are shown. White arrows point to some of the numerous MPO-expressing neutrophils (panel *A*) or RNAse3-expressing eosinophils (panel *B*) in MBC. Scale bar: 0.1mm. C) Quantitative analysis of neutrophil myeloperoxidase (MPO) and eosinophil cationic protein (RNAse3). Images of immunostained breast tissue sections were subjected to quantitative analysis (Supplemental Table S5) as described in Experimental Procedures. Left panel: Numbers of MPO+ITGAM+ cells (neutrophils) per mm2 of analyzed tissue specimen are displayed for disease-free breast (RM; n=13) and each breast cancer subtype (LumA; n=20, LumB; n=9, Her2; n=8, TNBC; n=11 and MBC; n=8). Degrees of significance for each pair comparison (*, **) are displayed at the top of the panel. P-values: LumA vs. RM: p=0.0434; LumB vs RM: p=0.0203; Her2 vs. RM: p=0.0033; TNBC vs. RM: p=0.068; MBC vs. RM: p=0.0025. Right panel: Same as above for RNASE3+ITGAM+ cells (eosinophils). Same number of specimens as above except for TNBC; n=9. Degrees of significance for each pair comparison (*, **, ***) are displayed at the top of the panel. P-values: LumA vs. RM: p=0.0227; LumB vs RM:

p=0.047; Her2 vs. RM: p=0.0022; TNBC vs. RM: p=0.023; MBC vs. RM: p=0.0008. Parameters used to generate graphs are shown in Supplemental Table S5.

REFERENCES

1. Society, A. C. (2019) Breast Cancer Facts & Figures 2019-2020. American Cencer Society, Inc., Atlanta

2. Lim, E., Vaillant, F., Wu, D., Forrest, N. C., Pal, B., Hart, A. H., Asselin-Labat, M.-L., Gyorki, D. E., Ward, T., Partanen, A., Feleppa, F., Huschtscha, L. I., Thorne, H. J., kConFab, Fox, S. B., Yan, M., French, J. D., Brown, M. A., Smyth, G. K., Visvader, J. E., and Lindeman, G. J. (2009) Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nature Medicine* 15, 907-913

3. McCart Reed, A. E., Kalaw, E. M., and Lakhani, S. R. (2021) An Update on the Molecular Pathology of Metaplastic Breast Cancer. *Breast Cancer: Targets and Therapy* 13, 161-170

4. Krings, G., and Chen, Y. Y. (2018) Genomic profiling of metaplastic breast carcinomas reveals genetic heterogeneity and relationship to ductal carcinoma. *Mod Pathol* 31, 1661-1674

5. Djomehri, S. I., Gonzalez, M. E., da Veiga Leprevost, F., Tekula, S. R., Chang, H.-Y., White, M. J., Cimino-Mathews, A., Burman, B., Basrur, V., Argani, P., Nesvizhskii, A. I., and Kleer, C. G. (2020) Quantitative proteomic landscape of metaplastic breast carcinoma pathological subtypes and their relationship to triple-negative tumors. *Nat Commun* 11, 1723

6. Kalaw, E., Lim, M., Kutasovic, J. R., Sokolova, A., Taege, L., Johnstone, K., Bennett, J., Saunus, J. M., Niland, C., Ferguson, K., Gresshoff, I., Bettington, M.,

Pathmanathan, N., Tse, G. M., Papadimos, D., Pathmanathan, R., Harris, G., Yamaguchi, R., Tan, P. H., Fox, S., O'Toole, S. A., Simpson, P. T., Lakhani, S. R., and McCart Reed, A. E. (2020) Metaplastic breast cancers frequently express immune checkpoint markers FOXP3 and PD-L1. *Br J Cancer* 123, 1665-1672

7. Kumar, B., Khatpe, A. S., Guanglong, J., Batic, K., Bhat-Nakshatri, P., Granatir, M. M., Addison, R. J., Szymanski, M., Baldridge, L. A., Temm, C. J., Sandusky, G., Althouse, S. K., Cote, M. L., Miller, K. D., Storniolo, A. M., and Nakshatri, H. (2023) Stromal heterogeneity may explain increased incidence of metaplastic breast cancer in women of African descent. *Nat Commun* 14, 5683

8. Finak, G., Bertos, N., Pepin, F., Sadekova, S., Souleimanova, M., Zhao, H., Chen, H., Omeroglu, G., Meterissian, S., Omeroglu, A., Hallett, M., and Park, M. (2008) Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med* 14, 518-527

9. Mollaoglu, G., Jones, A., Wait, S. J., Mukhopadhyay, A., Jeong, S., Arya, R., Camolotto, S. A., Mosbruger, T. L., Stubben, C. J., Conley, C. J., Bhutkar, A., Vahrenkamp, J. M., Berrett, K. C., Cessna, M. H., Lane, T. E., Witt, B. L., Salama, M. E., Gertz, J., Jones, K. B., Snyder, E. L., and Oliver, T. G. (2018) The Lineage-Defining Transcription Factors SOX2 and NKX2-1 Determine Lung Cancer Cell Fate and Shape the Tumor Immune Microenvironment. *Immunity* 49, 764-779 e769

10. Strasser, M. K., Gibbs, D. L., Gascard, P., Bons, J., Hickey, J. W., Schurch, C. M., Tan, Y., Black, S., Chu, P., Ozkan, A., Basisty, N., Sangwan, V., Rose, J., Shah, S., Camilleri-Broet, S., Fiset, P. O., Bertos, N., Berube, J., Djambazian, H., Li, R.,

Oikonomopoulos, S., Fels-Elliott, D. R., Vernovsky, S., Shimshoni, E., Collyar, D., Russell, A., Ragoussis, I., Stachler, M., Goldenring, J. R., McDonald, S., Ingber, D. E., Schilling, B., Nolan, G. P., Tlsty, T. D., Huang, S., and Ferri, L. E. (2023) Concerted epithelial and stromal changes during progression of Barrett's Esophagus to invasive adenocarcinoma exposed by multi-scale, multi-omics analysis. *bioRxiv*

11. Lee, S. H., Contreras Panta, E. W., Gibbs, D., Won, Y., Min, J., Zhang, C., Roland, J. T., Hong, S. H., Sohn, Y., Krystofiak, E., Jang, B., Ferri, L., Sangwan, V., Ragoussis, J., Camilleri-Broet, S., Caruso, J., Chen-Tanyolac, C., Strasser, M., Gascard, P., Tlsty, T. D., Huang, S., Choi, E., and Goldenring, J. R. (2023) Apposition of Fibroblasts With Metaplastic Gastric Cells Promotes Dysplastic Transition. *Gastroenterology* 165, 374-390

12. Gillet, L. C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., Bonner, R., and Aebersold, R. (2012) Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Molecular & Cellular Proteomics* 11, 6

13. Collins, B. C., Hunter, C. L., Liu, Y., Schilling, B., Rosenberger, G., Bader, S. L., Chan, D. W., Gibson, B. W., Gingras, A.-C., Held, J. M., Hirayama-Kurogi, M., Hou, G., Krisp, C., Larsen, B., Lin, L., Liu, S., Molloy, M. P., Moritz, R. L., Ohtsuki, S., Schlapbach, R., Selevsek, N., Thomas, S. N., Tzeng, S.-C., Zhang, H., and Aebersold, R. (2017) Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nat Commun* 8, 291

14. Bons, J., Pan, D., Shah, S., Bai, R., Chen-Tanyolac, C., Wang, X., Elliott, D. R. F., Urisman, A., O'Broin, A., Basisty, N., Rose, J., Sangwan, V., Camilleri-Broët, S., Tankel, J., Gascard, P., Ferri, L., Tlsty, T. D., and Schilling, B. (2023) Data-independent acquisition and quantification of extracellular matrix from human lung in chronic inflammation-associated carcinomas. *Proteomics* 23, 2200021

15. Burton, J. B., Silva-Barbosa, A., Bons, J., Rose, J., Pfister, K., Simona, F., Gandhi, T., Reiter, L., Bernhardt, O., Hunter, C. L., Goetzman, E. S., Sims-Lucas, S., and Schilling, B. (2023) Substantial downregulation of mitochondrial and peroxisomal proteins during acute kidney injury revealed by data-independent acquisition proteomics. *Proteomics*, e2300162

16. Lee, M. Y., Bedia, J. S., Bhate, S. S., Barlow, G. L., Phillips, D., Fantl, W. J., Nolan, G. P., and Schurch, C. M. (2022) CellSeg: a robust, pre-trained nucleus segmentation and pixel quantification software for highly multiplexed fluorescence images. *BMC Bioinformatics* 23, 46

17. Geoui, T., Urlaub, H., Plessmann, U., and Porschewski, P. (2010) Extraction of proteins from formalin-fixed, paraffin-embedded tissue using the Qproteome extraction technique and preparation of tryptic peptides for liquid chromatography/mass spectrometry analysis. *Curr Protoc Mol Biol* Chapter 10, Unit 10.27.11-12

18. Escher, C., Reiter, L., MacLean, B., Ossola, R., Herzog, F., Chilton, J., MacCoss, M. J., and Rinner, O. (2012) Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* 12, 1111-1121

19. Burton, J. B., Silva-Barbosa, A., Bons, J., Rose, J., Pfister, K., Simona, F., Gandhi, T., Reiter, L., Bernhardt, O., Hunter, C. L., Goetzman, E. S., Sims-Lucas, S., and Schilling, B. Substantial downregulation of mitochondrial and peroxisomal proteins during acute kidney injury revealed by data-independent acquisition proteomics. *Proteomics* 23, 2300162

20. Rosenberger, G., Koh, C. C., Guo, T., Röst, H. L., Kouvonen, P., Collins, B. C., Heusel, M., Liu, Y., Caron, E., Vichalkovski, A., Faini, M., Schubert, O. T., Faridi, P., Ebhardt, H. A., Matondo, M., Lam, H., Bader, S. L., Campbell, D. S., Deutsch, E. W., Moritz, R. L., Tate, S., and Aebersold, R. (2014) A repository of assays to quantify 10,000 human proteins by SWATH-MS. *Sci Data* 1, 140031

21. Burger, T. (2018) Gentle introduction to the statistical foundations of false discovery rate in quantitative proteomics. *Journal of proteome research* 17, 12-22

22. Dennis, G., Jr., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., and Lempicki, R. A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4, P3

23. Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*, Springer

24. Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., Alm, T., Asplund, A., Björk, L., Breckels, L. M., Bäckström, A., Danielsson, F., Fagerberg, L., Fall, J., Gatto, L., Gnann, C., Hober, S., Hjelmare, M., Johansson, F., Lee, S., Lindskog, C., Mulder, J., Mulvey, C. M., Nilsson, P., Oksvold, P., Rockberg, J., Schutten, R., Schwenk, J. M., Sivertsson, Å., Sjöstedt, E., Skogs, M., Stadler, C., Sullivan, D. P., Tegel, H., Winsnes, C., Zhang, C., Zwahlen, M., Mardinoglu, A.,

Pontén, F., von Feilitzen, K., Lilley, K. S., Uhlén, M., and Lundberg, E. (2017) A subcellular map of the human proteome. *Science* 356

25. Gatto, L., Breckels, L. M., Wieczorek, S., Burger, T., and Lilley, K. S. (2014) Mass-spectrometry-based spatial proteomics data analysis using pRoloc and pRolocdata. *Bioinformatics* 30, 1322-1324

26. Burton, J. B., Carruthers, N. J., Hou, Z., Matherly, L. H., and Stemmer, P. M. (2022) Pattern Analysis of Organellar Maps for Interpretation of Proteomic Data. *Proteomes* 10, 18

27. Langfelder, P., and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559

28. Schilling, B., Gibson, B. W., and Hunter, C. L. (2017) Generation of High-Quality SWATH(®) Acquisition Data for Label-free Quantitative Proteomics Studies Using TripleTOF(®) Mass Spectrometers. *Methods Mol Biol* 1550, 223-233

29. Friedrich, C., Schallenberg, S., Kirchner, M., Ziehm, M., Niquet, S., Haji, M., Beier, C., Neudecker, J., Klauschen, F., and Mertins, P. (2021) Comprehensive micro-scaled proteome and phosphoproteome characterization of archived retrospective cancer repositories. *Nat Commun* 12, 3576

30. Carr, S., Aebersold, R., Baldwin, M., Burlingame, A., Clauser, K., and Nesvizhskii, A. (2004) The Need for Guidelines in Publication of Peptide and Protein Identification Data: Working Group On Publication Guidelines For Peptide And Protein Identification Data *. *Molecular & Cellular Proteomics* 3, 531-533

31. Feng, Y., Xiong, Y., Qiao, T., Li, X., Jia, L., and Han, Y. (2018) Lactate dehydrogenase A: A key player in carcinogenesis and potential target in cancer therapy. *Cancer Med* 7, 6124-6136

32. Lang, B. J., Guerrero-Gimenez, M. E., Prince, T. L., Ackerman, A., Bonorino, C., and Calderwood, S. K. (2019) Heat Shock Proteins Are Essential Components in Transformation and Tumor Progression: Cancer Cell Intrinsic Pathways and Beyond. *Int J Mol Sci* 20, 18: 4507

33. Villeger, R., Chulkina, M., Mifflin, R. C., Markov, N. S., Trieu, J., Sinha, M., Johnson, P., Saada, J. I., Adegboyega, P. A., Luxon, B. A., Beswick, E. J., Powell, D. W., and Pinchuk, I. V. (2023) Loss of alcohol dehydrogenase 1B in cancer-associated fibroblasts: contribution to the increase of tumor-promoting IL-6 in colon cancer. *Br J Cancer* 128, 537-548

34. DeFilippis, R. A., Chang, H., Dumont, N., Rabban, J. T., Chen, Y. Y., Fontenay, G. V., Berman, H. K., Gauthier, M. L., Zhao, J., Hu, D., Marx, J. J., Tjoe, J. A., Ziv, E., Febbraio, M., Kerlikowske, K., Parvin, B., and Tlsty, T. D. (2012) CD36 repression activates a multicellular stromal program shared by high mammographic density and tumor tissues. *Cancer Discov* 2, 826-839

35. DeFilippis, R. A., Fordyce, C., Patten, K., Chang, H., Zhao, J., Fontenay, G. V., Kerlikowske, K., Parvin, B., and Tlsty, T. D. (2014) Stress signaling from human mammary epithelial cells contributes to phenotypes of mammographic density. *Cancer Res* 74, 5032-5044

36. Caruso, J. A., Wang, X., Murrow, L. M., Rodriguez, C. I., Chen-Tanyolac, C., Vu, L., Chen, Y. Y., Gascard, P., Gartner, Z. J., Kerlikowske, K., and Tlsty, T. D. (2023) Loss of PPARgamma activity characterizes early protumorigenic stromal reprogramming and dictates the therapeutic window of opportunity. *Proc Natl Acad Sci U S A* 120, e2303774120

37. Corn, K. C., Windham, M. A., and Rafat, M. (2020) Lipids in the tumor microenvironment: From cancer progression to treatment. *Prog Lipid Res* 80, 101055

38. Saraswathibhatla, A., Indana, D., and Chaudhuri, O. (2023) Cell-extracellular matrix mechanotransduction in 3D. *Nat Rev Mol Cell Biol* 24, 495-516

39. Lepucki, A., Orlińska, K., Mielczarek-Palacz, A., Kabut, J., Olczyk, P., and Komosińska-Vassev, K. (2022) The Role of Extracellular Matrix Proteins in Breast Cancer. *Journal of Clinical Medicine* 11, 1250

40. Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K. J., and Werb, Z. (2020) Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun* 11, 5120

41. Shao, X., Gomez, C. D., Kapoor, N., Considine, J. M., Grams, C., Gao, Y. T., and Naba, A. (2023) MatrisomeDB 2.0: 2023 updates to the ECM-protein knowledge database. *Nucleic Acids Res* 51, D1519-d1530

42. Jayadev, R., Morais, M., Ellingford, J. M., Srinivasan, S., Naylor, R. W., Lawless, C., Li, A. S., Ingham, J. F., Hastie, E., Chi, Q., Fresquet, M., Koudis, N. M., Thomas, H. B., O'Keefe, R. T., Williams, E., Adamson, A., Stuart, H. M., Banka, S., Smedley, D.,

Sherwood, D. R., and Lennon, R. (2022) A basement membrane discovery pipeline uncovers network complexity, regulators, and human disease associations. *Sci Adv* 8, eabn2265

43. Rujchanarong, D., Lefler, J., Saunders, J. E., Pippin, S., Spruill, L., Bethard, J. R., Ball, L. E., Mehta, A. S., Drake, R. R., Ostrowski, M. C., and Angel, P. M. (2021) Defining the Tumor Microenvironment by Integration of Immunohistochemistry and Extracellular Matrix Targeted Imaging Mass Spectrometry. *Cancers* 13, 4419

44. Su, H., Yang, F., Fu, R., Trinh, B., Sun, N., Liu, J., Kumar, A., Baglieri, J., Siruno, J., Le, M., Li, Y., Dozier, S., Nair, A., Filliol, A., Sinchai, N., Rosenthal, S. B., Santini, J., Metallo, C. M., Molina, A., Schwabe, R. F., Lowy, A. M., Brenner, D., Sun, B., and Karin, M. (2022) Collagenolysis-dependent DDR1 signalling dictates pancreatic cancer outcome. *Nature* 610, 366-372

45. Basisty, N., Kale, A., Jeon, O. H., Kuehnemann, C., Payne, T., Rao, C., Holtz, A., Shah, S., Sharma, V., Ferrucci, L., Campisi, J., and Schilling, B. (2020) A proteomic atlas of senescence-associated secretomes for aging biomarker development. *PLOS Biology* 18, e3000599

46. Fujihara, M., Shien, T., Shien, K., Suzawa, K., Takeda, T., Zhu, Y., Mamori, T., Otani, Y., Yoshioka, R., Uno, M., Suzuki, Y., Abe, Y., Hatono, M., Tsukioki, T., Takahashi, Y., Kochi, M., Iwamoto, T., Taira, N., Doihara, H., and Toyooka, S. (2021) YES1 as a Therapeutic Target for HER2-Positive Breast Cancer after Trastuzumab and Trastuzumab-Emtansine (T-DM1) Resistance Development. *Int J Mol Sci* 22

47. Song, F. f., Xia, L. l., Ji, P., Tang, Y. b., Huang, Z. m., Zhu, L., Zhang, J., Wang, J. q., Zhao, G. p., Ge, H. l., Zhang, Y., and Wang, Y. (2015) Human dCTP pyrophosphatase 1 promotes breast cancer cell growth and stemness through the modulation on 5-methyl-dCTP metabolism and global hypomethylation. *Oncogenesis* 4, e159-e159

48. Hu, Y. H., Lu, Y. X., Zhang, Z. Y., Zhang, J. M., Zhang, W. J., Zheng, L., Lin, W. H., Zhang, W., and Li, X. N. (2019) SSH3 facilitates colorectal cancer cell invasion and metastasis by affecting signaling cascades involving LIMK1/Rac1. *Am J Cancer Res* 9, 1061-1073

49. Kunji, E. R. S., King, M. S., Ruprecht, J. J., and Thangaratnarajah, C. (2020) The SLC25 Carrier Family: Important Transport Proteins in Mitochondrial Physiology and Pathology. *Physiology (Bethesda)* 35, 302-327

50. Huo, Q., Wang, J., and Xie, N. (2023) High HSPB1 expression predicts poor clinical outcomes and correlates with breast cancer metastasis. *BMC Cancer* 23, 501

51. Su, X., Li, H., Chen, S., and Qin, C. (2021) Study on the Prognostic Values of Dynactin Genes in Low-Grade Glioma. *Technol Cancer Res Treat* 20, 15330338211010143

52. Sun, H., Zhou, R., Zheng, Y., Wen, Z., Zhang, D., Zeng, D., Wu, J., Huang, Z., Rong, X., Huang, N., Sun, L., Bin, J., Liao, Y., Shi, M., and Liao, W. (2021) CRIP1 cooperates with BRCA2 to drive the nuclear enrichment of RAD51 and to facilitate homologous repair upon DNA damage induced by chemotherapy. *Oncogene* 40, 5342-5355

53. Zhou, Y., and Luo, G. (2020) Apolipoproteins, as the carrier proteins for lipids, are involved in the development of breast cancer. *Clin Transl Oncol* 22, 1952-1962

54. Ren, L., Yi, J., Li, W., Zheng, X., Liu, J., Wang, J., and Du, G. (2019) Apolipoproteins and cancer. *Cancer Med* 8, 7032-7043

55. Wang, D., Li, J., Cai, F., Xu, Z., Li, L., Zhu, H., Liu, W., Xu, Q., Cao, J., Sun, J., and Tang, J. (2019) Overexpression of MAPT-AS1 is associated with better patient survival in breast cancer. *Biochem Cell Biol* 97, 158-164

56. Zhang, T., Yu, S., and Zhao, S. (2021) ANXA9 as a novel prognostic biomarker associated with immune infiltrates in gastric cancer. *PeerJ* 9, e12605

57. Rochette, L., Meloux, A., Zeller, M., Malka, G., Cottin, Y., and Vergely, C. (2020) Mitochondrial SLC25 Carriers: Novel Targets for Cancer Therapy. *Molecules* 25

58. Jansen, M. P., Sas, L., Sieuwerts, A. M., Van Cauwenberghe, C., Ramirez-Ardila, D., Look, M., Ruigrok-Ritstier, K., Finetti, P., Bertucci, F., Timmermans, M. M., van Deurzen, C. H., Martens, J. W., Simon, I., Roepman, P., Linn, S. C., van Dam, P., Kok, M., Lardon, F., Vermeulen, P. B., Foekens, J. A., Dirix, L., Berns, E. M., and Van Laere, S. (2015) Decreased expression of ABAT and STC2 hallmarks ER-positive inflammatory breast cancer and endocrine therapy resistance in advanced disease. *Mol Oncol* 9, 1218-1233

59. El Khoury, W., and Nasr, Z. (2021) Deregulation of ribosomal proteins in human cancers. *Biosci Rep* 41

60. Zhao, H. B., Zhang, X. F., Wang, H. B., and Zhang, M. Z. (2017) Migration and invasion enhancer 1 (MIEN1) is overexpressed in breast cancer and is a potential new therapeutic molecular target. *Genet Mol Res* 16

61. Bivin, W. W., Yergiyev, O., Bunker, M. L., Silverman, J. F., and Krishnamurti, U. (2017) GRB7 Expression and Correlation With HER2 Amplification in Invasive Breast Carcinoma. *Appl Immunohistochem Mol Morphol* 25, 553-558

62. Pillai, S., Mahmud, I., Mahar, R., Griffith, C., Langsen, M., Nguyen, J., Wojtkowiak, J. W., Swietach, P., Gatenby, R. A., Bui, M. M., Merritt, M. E., McDonald, P., Garrett, T. J., and Gillies, R. J. (2022) Lipogenesis mediated by OGR1 regulates metabolic adaptation to acid stress in cancer cells via autophagy. *Cell Rep* 39, 110796

63. Xiong, S., Dong, L., and Cheng, L. (2021) Neutrophils in cancer carcinogenesis and metastasis. *Journal of Hematology & Oncology* 14, 173

64. Aloe, C., Wang, H., Vlahos, R., Irving, L., Steinfort, D., and Bozinovski, S. (2021) Emerging and multifaceted role of neutrophils in lung cancer. *Transl Lung Cancer Res* 10, 2806-2818

65. Zabkiewicz, C., Resaul, J., Hargest, R., Jiang, W. G., and Ye, L. (2017) Bone morphogenetic proteins, breast cancer, and bone metastases: striking the right balance. *Endocr Relat Cancer* 24, R349-r366

66. Afroz, S., Giddaluru, J., Abbas, M. M., and Khan, N. (2016) Transcriptome meta-analysis reveals a dysregulation in extra cellular matrix and cell junction associated gene signatures during Dengue virus infection. *Sci Rep* 6, 33752

67. Chen, Y., Guo, H., Terajima, M., Banerjee, P., Liu, X., Yu, J., Momin, A. A., Katayama, H., Hanash, S. M., Burns, A. R., Fields, G. B., Yamauchi, M., and Kurie, J. M. (2016) Lysyl Hydroxylase 2 Is Secreted by Tumor Cells and Can Modify Collagen in the Extracellular Space *. *Journal of Biological Chemistry* 291, 25799-25808

68. Nallanthighal, S., Heiserman, J. P., and Cheon, D. J. (2019) The Role of the Extracellular Matrix in Cancer Stemness. *Front Cell Dev Biol* 7, 86

69. Tong, Y., Qi, Y., Xiong, G., Li, J., Scott, T. L., Chen, J., He, D., Li, L., Wang, C., Lane, A. N., and Xu, R. (2023) The PLOD2/succinate axis regulates the epithelial-mesenchymal plasticity and cancer cell stemness. *Proc Natl Acad Sci U S A* 120, e2214942120

70. Grkovic, S., O'Reilly, V. C., Han, S., Hong, M., Baxter, R. C., and Firth, S. M. (2013) IGFBP-3 binds GRP78, stimulates autophagy and promotes the survival of breast cancer cells exposed to adverse microenvironments. *Oncogene* 32, 2412-2420

71. He, W., Zhang, H., Wang, Y., Zhou, Y., Luo, Y., Cui, Y., Jiang, N., Jiang, W., Wang, H., Xu, D., Li, S., Wang, Z., Chen, Y., Sun, Y., Zhang, Y., Tseng, H.-R., Zou, X., Wang, L., and Ke, Z. (2018) CTHRC1 induces non-small cell lung cancer (NSCLC) invasion through upregulating MMP-7/MMP-9. *BMC Cancer* 18, 400

72. Keller, L. E., Fortier, L. A., Lattermann, C., Hunt, E. R., Zhang, S., Fu, Q., and Jacobs, C. A. (2023) Complement system dysregulation in synovial fluid from patients with persistent inflammation following anterior cruciate ligament reconstruction surgery. *Journal of Cartilage & Joint Preservation*, 100114

73. Li, H., Zhang, M., Wei, Y., Haider, F., Lin, Y., Guan, W., Liu, Y., Zhang, S., Yuan, R., Yang, X., Yang, S., and Wang, H. (2020) SH3BGRL confers innate drug resistance in breast cancer by stabilizing HER2 activation on cell membrane. *Journal of Experimental & Clinical Cancer Research* 39, 81

74. Cao, Y., Chu, C., Li, X., Gu, S., Zou, Q., and Jin, Y. (2021) RNA-binding protein QKI suppresses breast cancer via RASA1/MAPK signaling pathway. *Ann Transl Med* 9, 104

75. Choi, B., Lee, E. J., Song, D. H., Yoon, S. C., Chung, Y. H., Jang, Y., Kim, S. M., Song, Y., Kang, S. W., Yoon, S. Y., and Chang, E. J. (2014) Elevated Pentraxin 3 in bone metastatic breast cancer is correlated with osteolytic function. *Oncotarget* 5, 481-492

76. Shi, J., Bao, M., Wang, W., Wu, X., Li, Y., Zhao, C., and Liu, W. (2021) Integrated Profiling Identifies PLOD3 as a Potential Prognostic and Immunotherapy Relevant Biomarker in Colorectal Cancer. *Front Immunol* 12, 722807

77. Nouh, M. A., Mohamed, M. M., El-Shinawi, M., Shaalan, M. A., Cavallo-Medved, D., Khaled, H. M., and Sloane, B. F. (2011) Cathepsin B: a potential prognostic marker for inflammatory breast cancer. *J Transl Med* 9, 1

78. Zhu, X., Xu, X., Zhang, L., and Yang, X. (2023) Carboxypeptidase vitellogenic like facilitates resistance to CDK4/6 inhibitors in breast cancer. *Thorac Cancer* 14, 983-991

79. Malek, N., Michrowska, A., Mazurkiewicz, E., Mrówczyńska, E., Mackiewicz, P., and Mazur, A. J. (2021) The origin of the expressed retrotransposed gene ACTBL2 and

its influence on human melanoma cells' motility and focal adhesion formation. *Scientific Reports* 11, 1: 3329

80. Aoto, K., Kato, M., Akita, T., Nakashima, M., Mutoh, H., Akasaka, N., Tohyama, J., Nomura, Y., Hoshino, K., Ago, Y., Tanaka, R., Epstein, O., Ben-Haim, R., Heyman, E., Miyazaki, T., Belal, H., Takabayashi, S., Ohba, C., Takata, A., Mizuguchi, T., Miyatake, S., Miyake, N., Fukuda, A., Matsumoto, N., and Saitsu, H. (2021) ATP6V0A1 encoding the a1-subunit of the V0 domain of vacuolar H(+)-ATPases is essential for brain development in humans and mice. *Nat Commun* 12, 1: 2107

81. Chhieng, C., Cranor, M., Lesser, M. E., and Rosen, P. P. (1998) Metaplastic carcinoma of the breast with osteocartilaginous heterologous elements. *Am J Surg Pathol* 22, 188-194

82. Reddy, T. P., Rosato, R. R., Li, X., Moulder, S., Piwnica-Worms, H., and Chang, J. C. (2020) A comprehensive overview of metaplastic breast cancer: clinical features and molecular aberrations. *Breast Cancer Res* 22, 1: 121

83. Merkin, R. D., Vanner, E. A., Romeiser, J. L., Shroyer, A. L. W., Escobar-Hoyos, L. F., Li, J., Powers, R. S., Burke, S., and Shroyer, K. R. (2017) Keratin 17 is overexpressed and predicts poor survival in estrogen receptor-negative/human epidermal growth factor receptor-2-negative breast cancer. *Hum Pathol* 62, 23-32

84. Zhu, Z., Achreja, A., Meurs, N., Animasahun, O., Owen, S., Mittal, A., Parikh, P., Lo, T. W., Franco-Barraza, J., Shi, J., Gunchick, V., Sherman, M. H., Cukierman, E., Pickering, A. M., Maitra, A., Sahai, V., Morgan, M. A., Nagrath, S., Lawrence, T. S.,

and Nagrath, D. (2020) Tumour-reprogrammed stromal BCAT1 fuels branched-chain ketoacid dependency in stromal-rich PDAC tumours. *Nat Metab* 2, 775-792

85. Barata, T., Duarte, I., and Futschik, M. E. (2023) Integration of Stemness Gene Signatures Reveals Core Functional Modules of Stem Cells and Potential Novel Stemness Genes. *Genes (Basel)* 14, 3:745

86. Scheller, A. S., Philipp, T. M., Klotz, L. O., and Steinbrenner, H. (2022) Altered Capacity for H(2)S Production during the Spontaneous Differentiation of Caco-2 Cells to Colonocytes Due to Reciprocal Regulation of CBS and SELENBP1. *Antioxidants (Basel)* 11, 10: 1957

87. Vasaikar, S. V., Deshmukh, A. P., den Hollander, P., Addanki, S., Kuburich, N. A., Kudaravalli, S., Joseph, R., Chang, J. T., Soundararajan, R., and Mani, S. A. (2021) EMTome: a resource for pan-cancer analysis of epithelial-mesenchymal transition genes and signatures. *Br J Cancer* 124, 259-269

88. Hill, C. N., Hernández-Cáceres, M. P., Asencio, C., Torres, B., Solis, B., and Owen, G. I. (2020) Deciphering the Role of the Coagulation Cascade and Autophagy in Cancer-Related Thrombosis and Metastasis. *Front Oncol* 10: 605314

89. Sellers, K., Allen, T. D., Bousamra, M., 2nd, Tan, J., Mendez-Lucas, A., Lin, W., Bah, N., Chernyavskaya, Y., MacRae, J. I., Higashi, R. M., Lane, A. N., Fan, T. W., and Yuneva, M. O. (2019) Metabolic reprogramming and Notch activity distinguish between non-small cell lung cancer subtypes. *Br J Cancer* 121, 51-64

90. Schulze, A., Oshi, M., Endo, I., and Takabe, K. (2020) MYC Targets Scores Are Associated with Cancer Aggressiveness and Poor Survival in ER-Positive Primary and Metastatic Breast Cancer. *Int J Mol Sci* 21, 21: 8127

91. Avagliano, A., Ruocco, M. R., Aliotta, F., Belviso, I., Accurso, A., Masone, S., Montagnani, S., and Arcucci, A. (2019) Mitochondrial Flexibility of Breast Cancers: A Growth Advantage and a Therapeutic Opportunity. *Cells* 8, 5: 401

92. Kim, J. A. (2020) Peroxisome Metabolism in Cancer. *Cells* 9, 7: 1692

93. Oakes, S. A. (2020) Endoplasmic Reticulum Stress Signaling in Cancer Cells. *Am J Pathol* 190, 934-946

94. Morozov, A. V., and Karpov, V. L. (2019) Proteasomes and Several Aspects of Their Heterogeneity Relevant to Cancer. *Front Oncol* 9: 761

95. Taylor, R. C., Berendzen, K. M., and Dillin, A. (2014) Systemic stress signalling: understanding the cell non-autonomous control of proteostasis. *Nat Rev Mol Cell Biol* 15, 211-217

96. Schumacher, B., Pothof, J., Vijg, J., and Hoeijmakers, J. H. J. (2021) The central role of DNA damage in the ageing process. *Nature* 592, 695-703

97. Dorafshan, S., Razmi, M., Safaei, S., Gentilin, E., Madjd, Z., and Ghods, R. (2022) Periostin: biology and function in cancer. *Cancer Cell Int* 22: 315

98. Rick, J. W., Chandra, A., Dalle Ore, C., Nguyen, A. T., Yagnik, G., and Aghi, M. K. (2019) Fibronectin in malignancy: Cancer-specific alterations, protumoral effects, and therapeutic implications. *Semin Oncol* 46, 284-290

99. Yilmaz, A., Loustau, T., Salomé, N., Poilil Surendran, S., Li, C., Tucker, R. P., Izzi, V., Lamba, R., Koch, M., and Orend, G. (2022) Advances on the roles of tenascin-C in cancer. *J Cell Sci* 135: 18

100. Khan, F. U., Owusu-Tieku, N. Y. G., Dai, X., Liu, K., Wu, Y., Tsai, H. I., Chen, H., Sun, C., and Huang, L. (2019) Wnt/beta-Catenin Pathway-Regulated Fibromodulin Expression Is Crucial for Breast Cancer Metastasis and Inhibited by Aspirin. *Front Pharmacol* 10, 1308

101. Wang, B., Qi, X., Liu, J., Zhou, R., Lin, C., Shangguan, J., Zhang, Z., Zhao, L., and Li, G. (2019) MYH9 Promotes Growth and Metastasis via Activation of MAPK/AKT Signaling in Colorectal Cancer. *J Cancer* 10, 874-884

102. Ye, L., Wang, H., Li, H., Liu, H., Lv, T., Song, Y., and Zhang, F. (2019) Eosinophil peroxidase over-expression predicts the clinical outcome of patients with primary lung adenocarcinoma. *J Cancer* 10, 1032-1038

103. Zeindler, J., Angehrn, F., Droeser, R., Däster, S., Piscuoglio, S., Ng, C. K. Y., Kilic, E., Mechera, R., Meili, S., Isaak, A., Weber, W. P., Muenst, S., and Soysal, S. D. (2019) Infiltration by myeloperoxidase-positive neutrophils is an independent prognostic factor in breast cancer. *Breast Cancer Res Treat* 177, 581-589

104. Khan, A. A., Alsahli, M. A., and Rahmani, A. H. (2018) Myeloperoxidase as an Active Disease Biomarker: Recent Biochemical and Pathological Perspectives. *Med Sci (Basel)* 6: 33

105. Moutin, E. B., Bons, J., Giavara, G., Lourenco, F., Pan, D., Burton, J. B., Shah, S., Colombe, M., Gascard, P., Tlsty, T., Schilling, B., and Winton, D. J. (2024)

Extracellular Matrix Orchestration of Tissue Remodeling in the Chronically Inflamed Mouse Colon. *Cell Mol Gastroenterol Hepatol* 17, 4, 639-656

106. Liu, W. Q., Zhang, Y. Z., Wu, Y., Zhang, J. J., Li, T. B., Jiang, T., Xiong, X. M., Luo, X. J., Ma, Q. L., and Peng, J. (2015) Myeloperoxidase-derived hypochlorous acid promotes ox-LDL-induced senescence of endothelial cells through a mechanism involving beta-catenin signaling in hyperlipidemia. *Biochem Biophys Res Commun* 467, 859-865

107. Sen, B., Aggarwal, S., Nath, R., Sehgal, R., Singh, R., Agrawal, K., Shashidhara, A. N., Rastogi, A., Bajpai, M., Pamecha, V., Trehanpati, N., and Ramakrishna, G. (2022) Secretome of senescent hepatoma cells modulate immune cell fate by macrophage polarization and neutrophil extracellular traps formation. *Medical Oncology* 39, 134

108. Zheng, C., Xu, X., Wu, M., Xue, L., Zhu, J., Xia, H., Ding, S., Fu, S., Wang, X., Wang, Y., He, G., Liu, X., and Deng, X. (2023) Neutrophils in triple-negative breast cancer: an underestimated player with increasingly recognized importance. *Breast Cancer Research* 25: 88

109. Caruso, J. A., and Tlsty, T. D. (2021) Remaining true to one's identity. *Nature Aging* 1, 757-759

110. Shalabi, S. F., Miyano, M., Sayaman, R. W., Lopez, J. C., Jokela, T. A., Todhunter, M. E., Hinz, S., Garbe, J. C., Stampfer, M. R., Kessenbrock, K., Seewaldt, V. E., and LaBarge, M. A. (2021) Evidence for accelerated aging in

mammary epithelia of women carrying germline BRCA1 or BRCA2 mutations. *Nature Aging* 1, 838-849

111. Campisi, J., and d'Adda di Fagagna, F. (2007) Cellular senescence: when bad things happen to good cells. *Nature Reviews Molecular Cell Biology* 8, 729-740

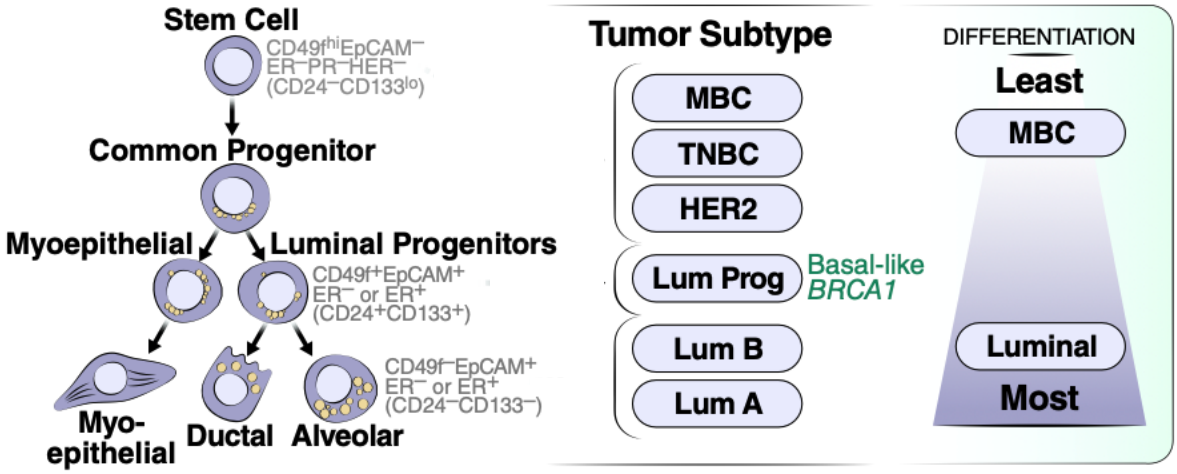
112. Coppé, J.-P., Desprez, P.-Y., Krtolica, A., and Campisi, J. (2010) The senescence-associated secretory phenotype: the dark side of tumor suppression. *Annual Review of Pathology* 5, 99-118

113. Marozzi, M., Parnigoni, A., Negri, A., Viola, M., Vigetti, D., Passi, A., Karousou, E., and Rizzi, F. (2021) Inflammation, Extracellular Matrix Remodeling, and Proteostasis in Tumor Microenvironment. *Int J Mol Sci* 22, 15: 8102

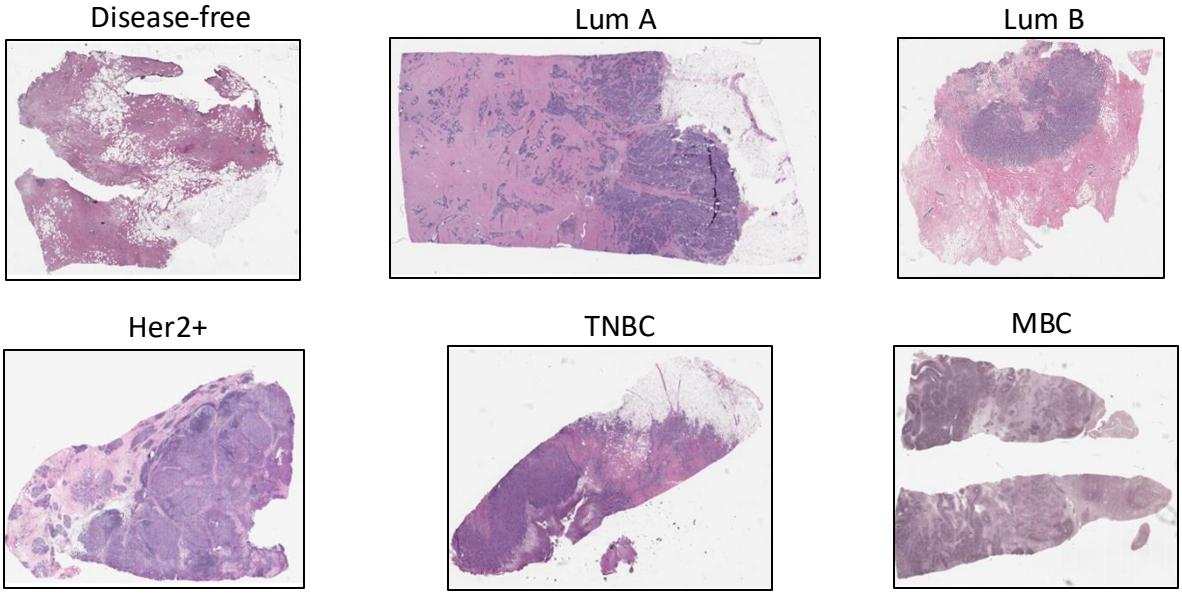
114. Luthold, C., Hallal, T., Labbé, D. P., and Bordeleau, F. (2022) The Extracellular Matrix Stiffening: A Trigger of Prostate Cancer Progression and Castration Resistance? *Cancers 14: 2887*

Figure 1

A Breast Tumor Subtype Hierarchy



B Histological Examples of Disease-free Breast and Breast Cancer Subtypes



C Workflow

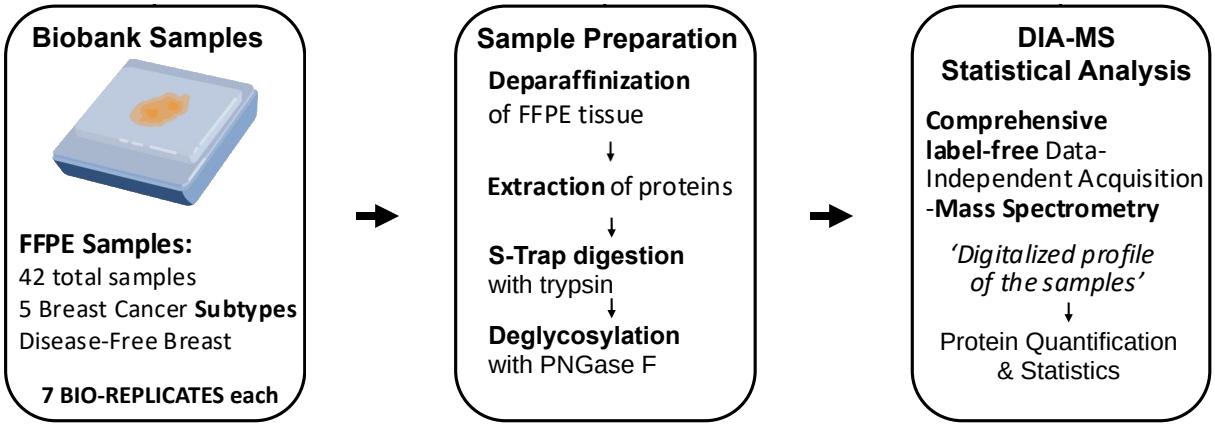


Figure 2

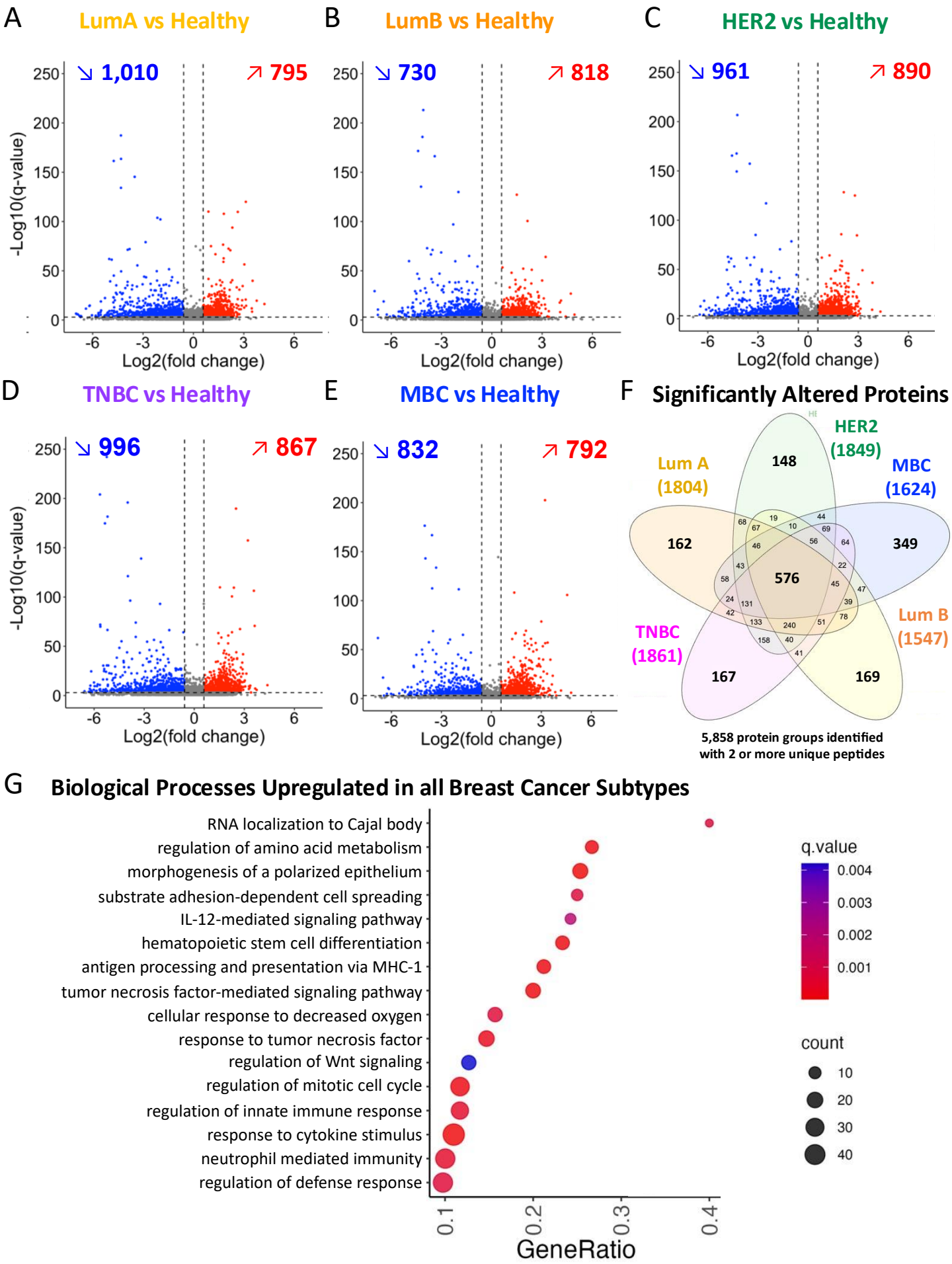

A
LumA vs Healthy
↘ 1,010
↗ 795
B
LumB vs Healthy
↘ 730
↗ 818
C
HER2 vs Healthy
↘ 961
↗ 890
D
TNBC vs Healthy
↘ 996
↗ 867
E
MBC vs Healthy
↘ 832
↗ 792
-Log10(q-value)
Log2(fold change)
F
Significantly Altered Proteins
Lum A (1804)
HER2 (1849)
MBC (1624)
TNBC (1861)
Lum B (1547)
162
148
349
576
167
169
5,858 protein groups identified with 2 or more unique peptides
G
Biological Processes Upregulated in all Breast Cancer Subtypes
RNA localization to Cajal body
regulation of amino acid metabolism
morphogenesis of a polarized epithelium
substrate adhesion-dependent cell spreading
IL-12-mediated signaling pathway
hematopoietic stem cell differentiation
antigen processing and presentation via MHC-1
tumor necrosis factor-mediated signaling pathway
cellular response to decreased oxygen
response to tumor necrosis factor
regulation of Wnt signaling
regulation of mitotic cell cycle
regulation of innate immune response
response to cytokine stimulus
neutrophil mediated immunity
regulation of defense response
GeneRatio
q.value
0.004
0.003
0.002
0.001
count
10
20
30
40

Figure 3

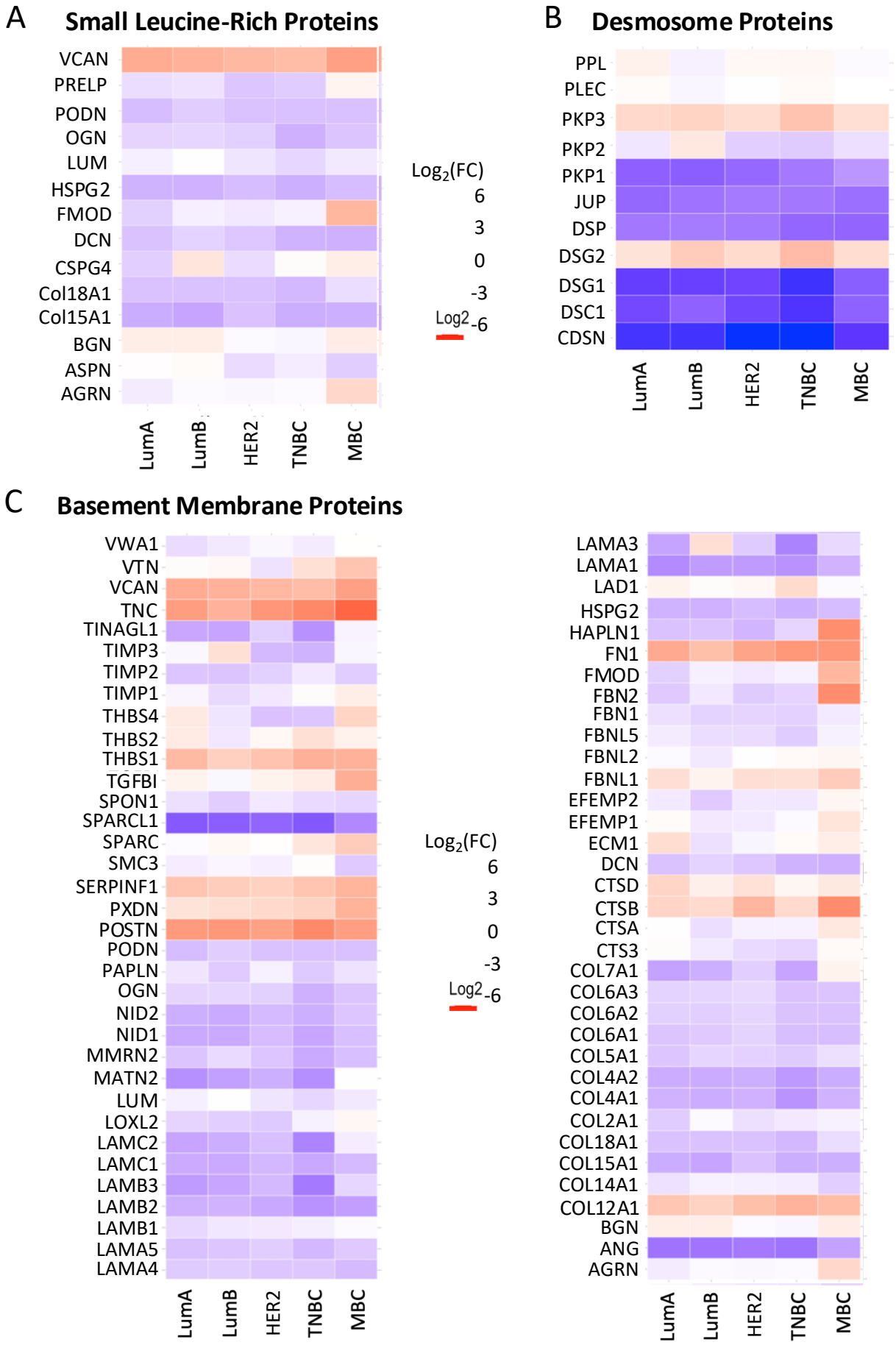
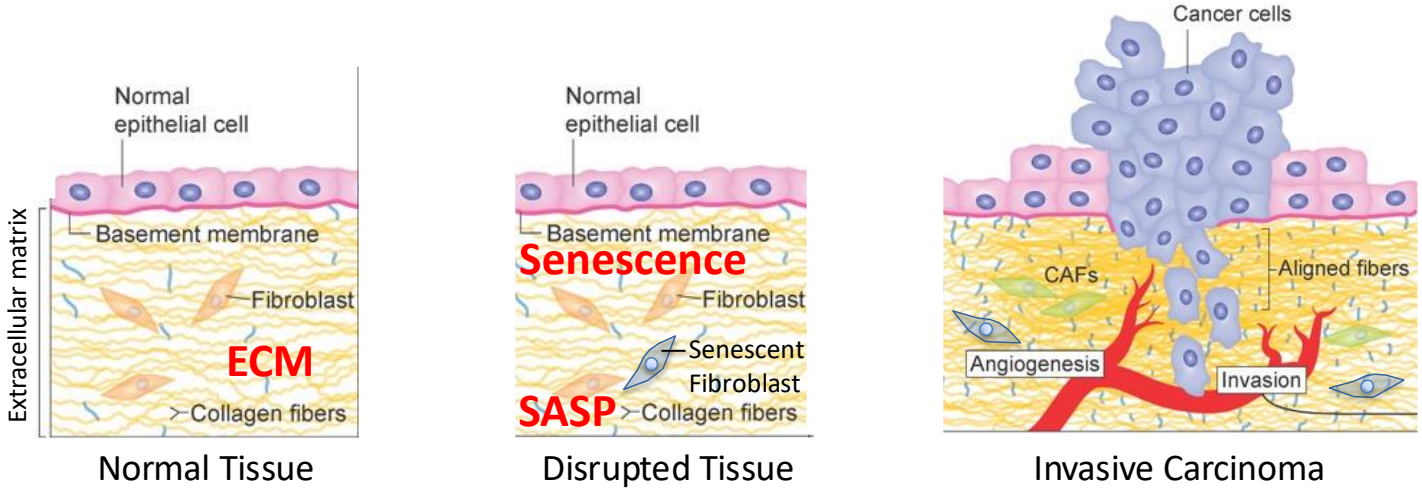
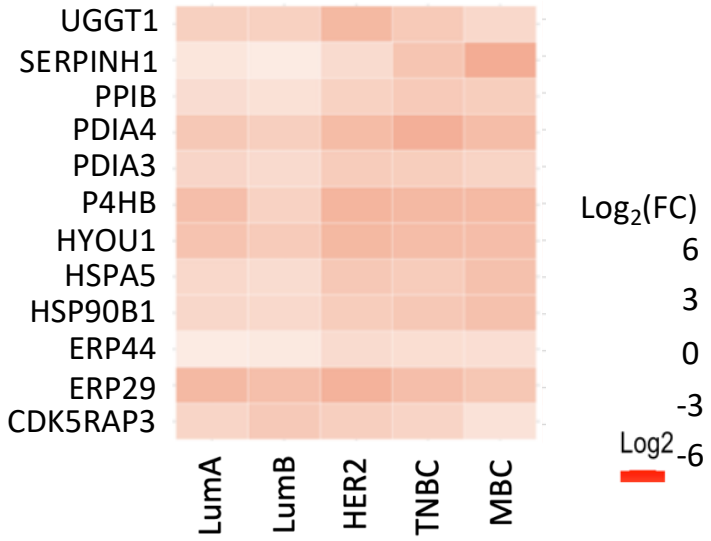

A
Small Leucine-Rich Proteins
VCAN
PRELP
PODN
OGN
LUM
HSPG2
FMOD
DCN
CSPG4
Col18A1
Col15A1
BGN
ASPN
AGRN
LumA
LumB
HER2
TNBC
MBC
Log2(FC)
6
3
0
-3
Log2 -6
B
Desmosome Proteins
PPL
PLEC
PKP3
PKP2
PKP1
JUP
DSP
DSG2
DSG1
DSC1
CDSN
LumA
LumB
HER2
TNBC
MBC
C
Basement Membrane Proteins
VWA1
VTN
VCAN
TNC
TINAGL1
TIMP3
TIMP2
TIMP1
THBS4
THBS2
THBS1
TGFBI
SPON1
SPARCL1
SPARC
SMC3
SERPINF1
PXDN
POSTN
PODN
PAPLN
OGN
NID2
NID1
MMRN2
MATN2
LUM
LOXL2
LAMC2
LAMC1
LAMB3
LAMB2
LAMB1
LAMA5
LAMA4
LumA
LumB
HER2
TNBC
MBC
Log2(FC)
6
3
0
-3
Log2 -6
LAMA3
LAMA1
LAD1
HSPG2
HAPLN1
FN1
FMOD
FBN2
FBN1
FBNL5
FBNL2
FBNL1
EFEMP2
EFEMP1
ECM1
DCN
CTSD
CTSB
CTSA
CTS3
COL7A1
COL6A3
COL6A2
COL6A1
COL5A1
COL4A2
COL4A1
COL2A1
COL18A1
COL15A1
COL14A1
COL12A1
BGN
ANG
AGRN
LumA
LumB
HER2
TNBC
MBC

Figure 4

A Chronic Inflammation Associated Cancer, ECM, and Senescence in Aging and Cancer



B Endoplasmic Reticulum Stress-Response



C Highly Regulated SASP Markers

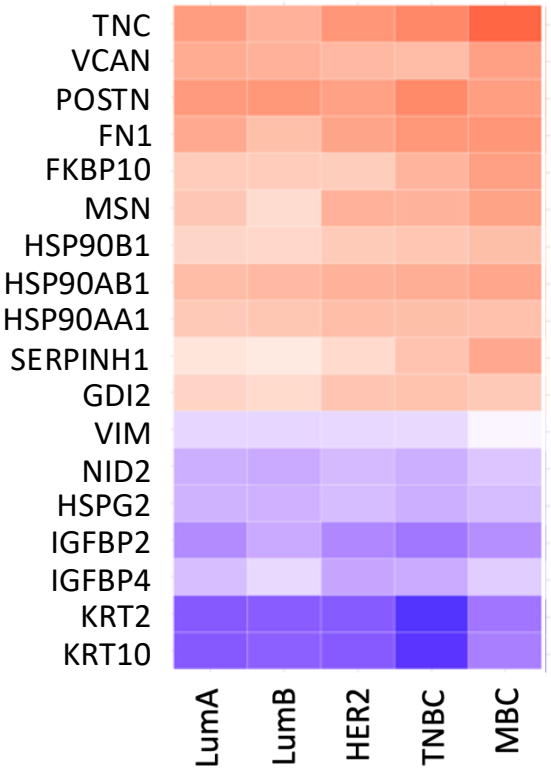
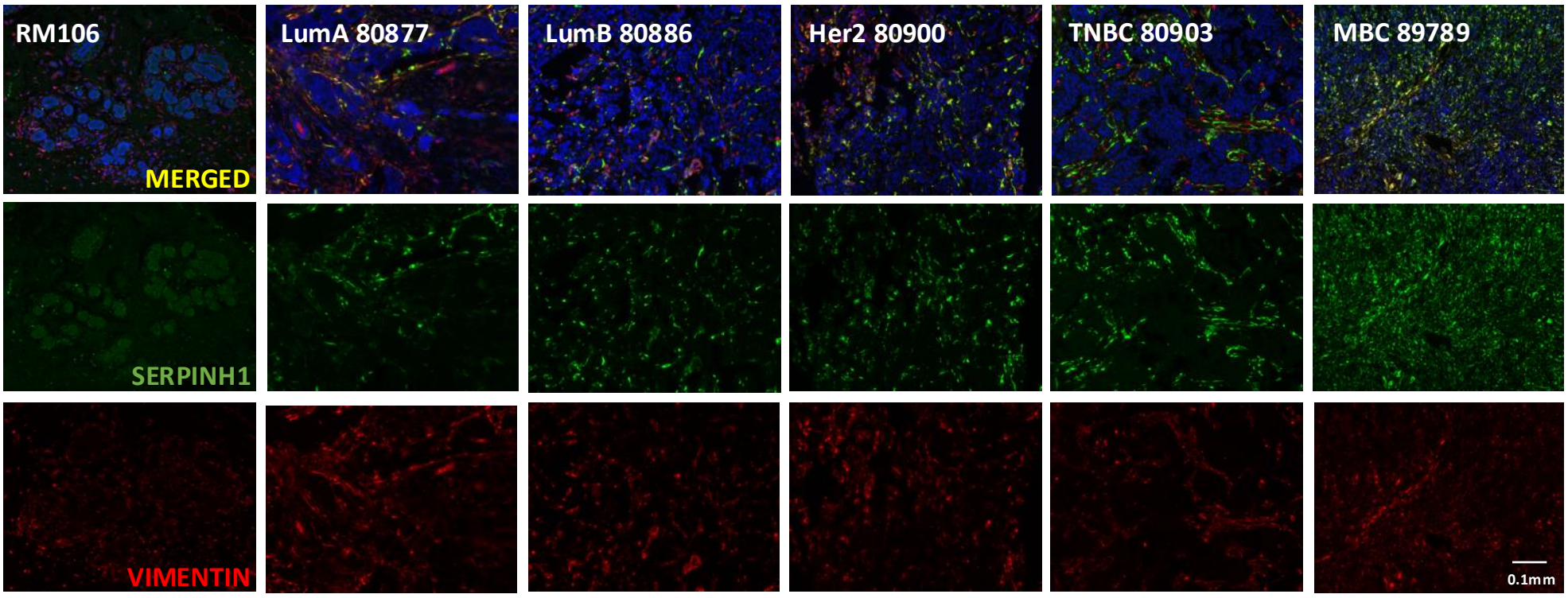


Figure 5

A SERPINH1 MODULE



B

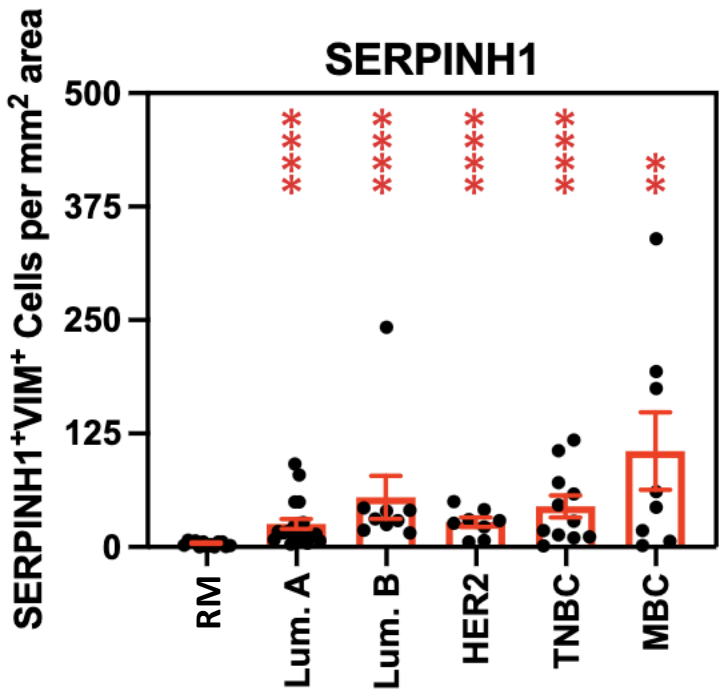


Figure 6

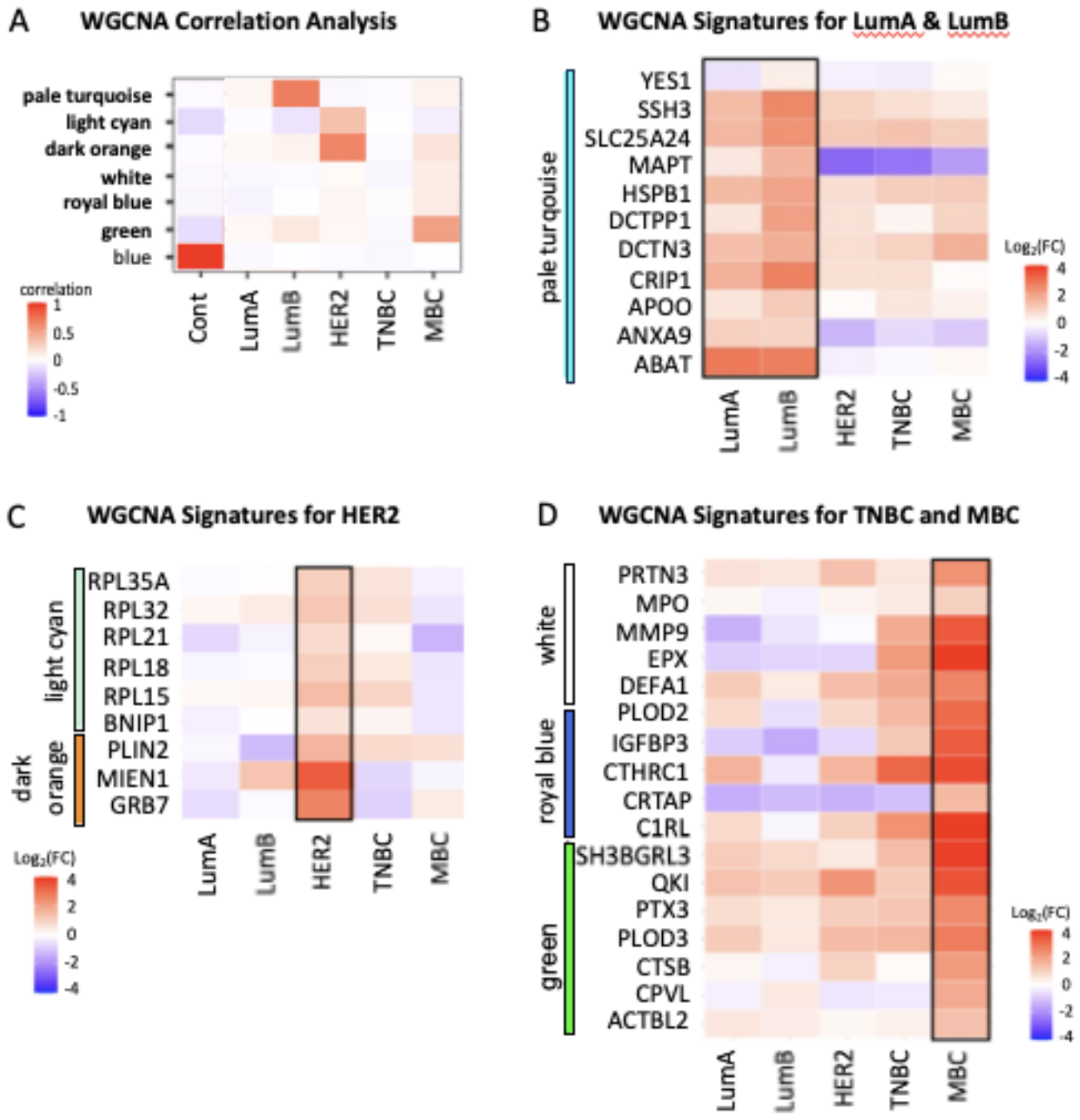
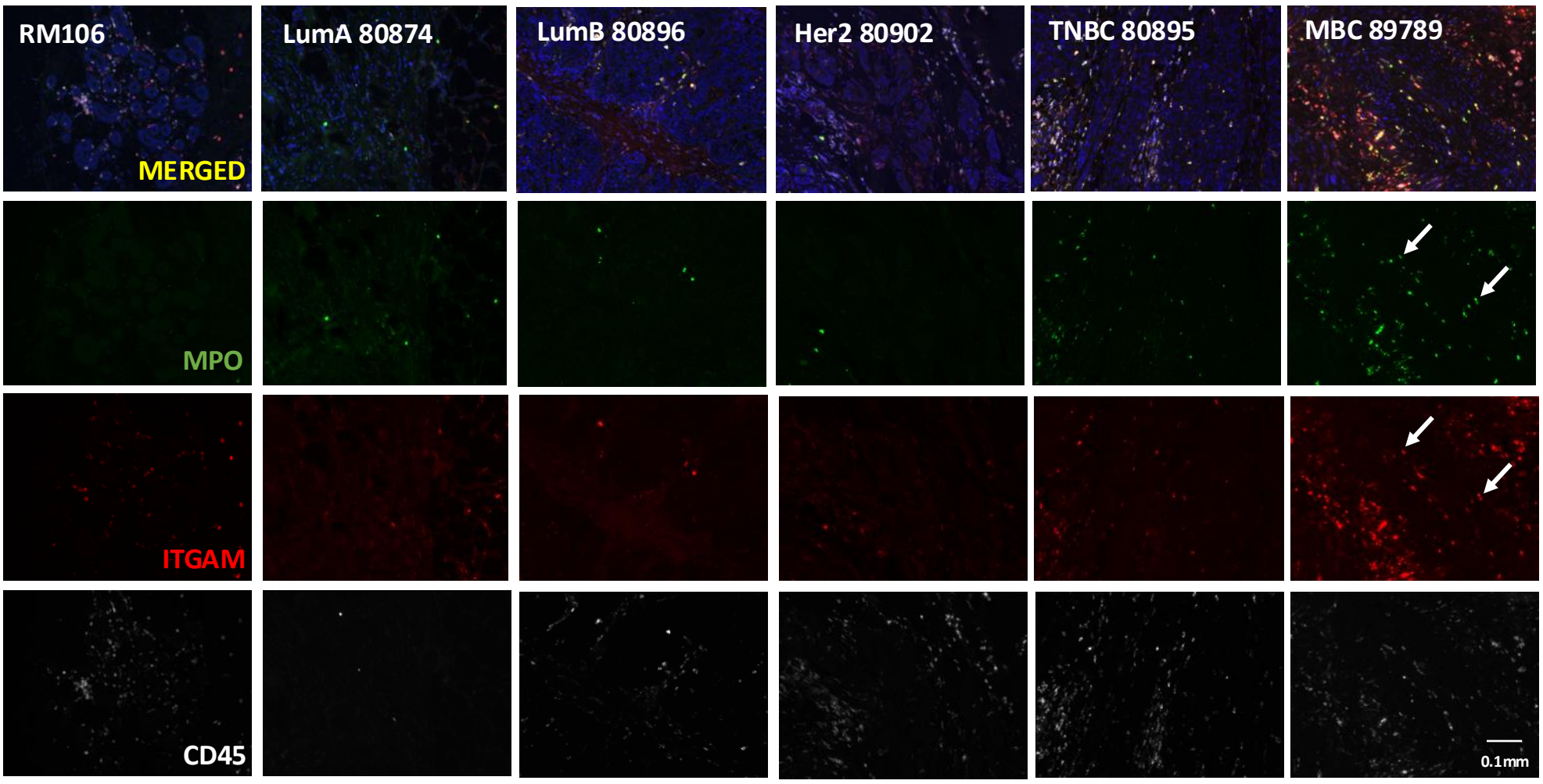
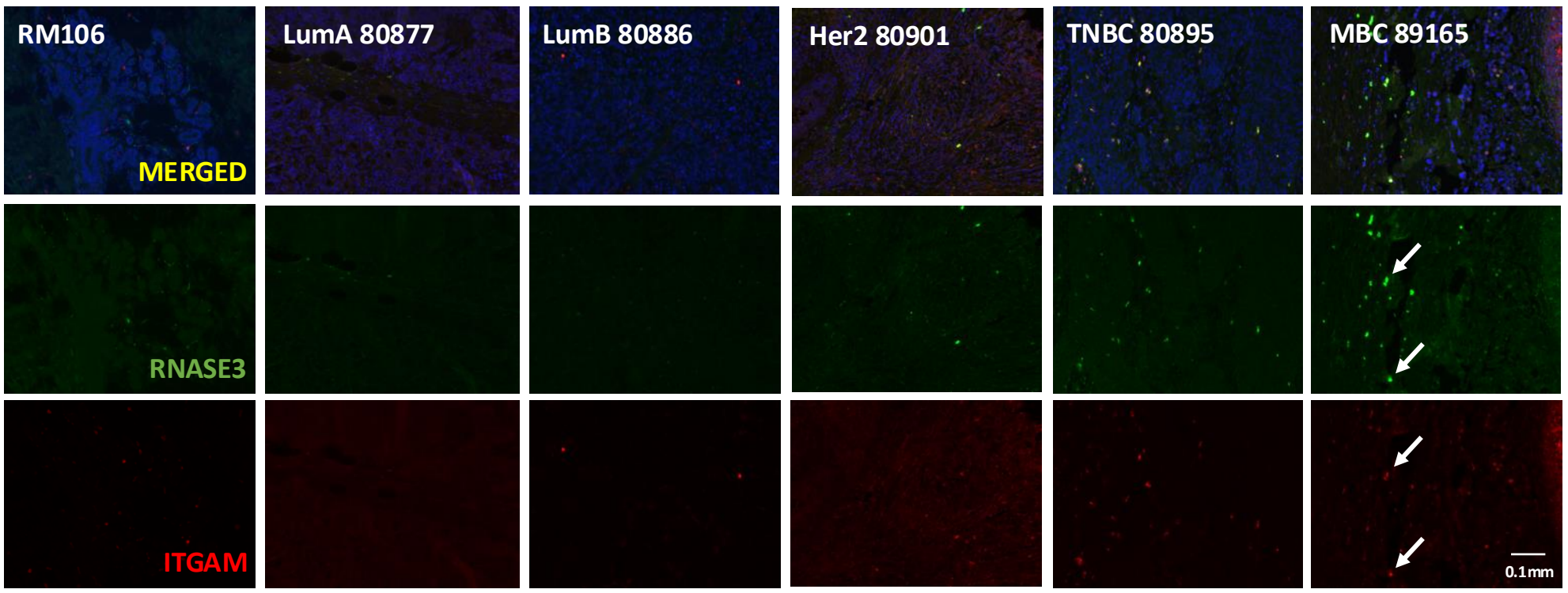
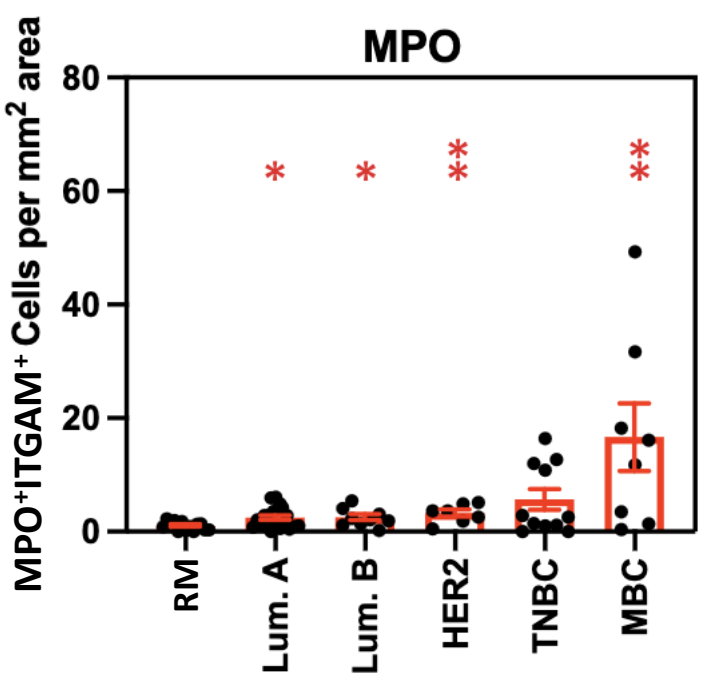
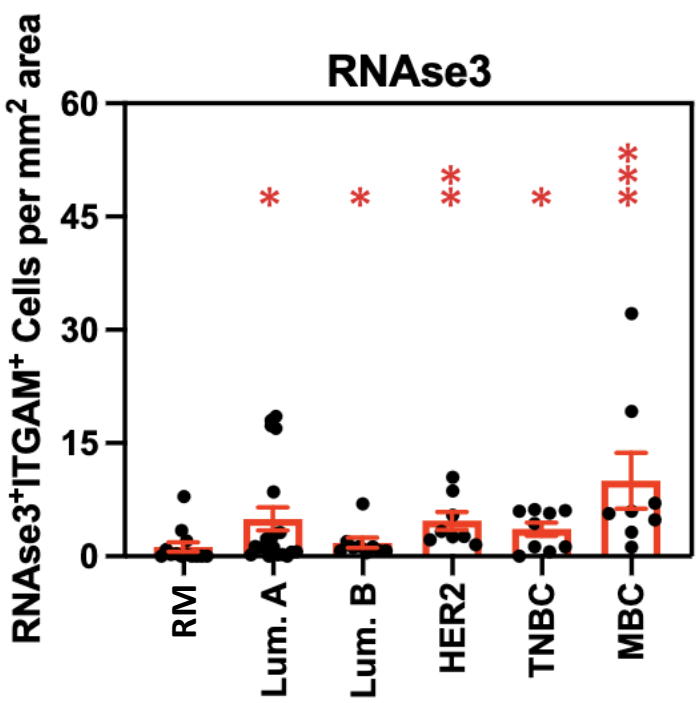

A
WGCNA Correlation Analysis
pale turquoise
light cyan
dark orange
white
royal blue
green
blue
Cont
LumA
LumB
HER2
TNBC
MBC
correlation
1
0.5
0
-0.5
-1
B
WGCNA Signatures for LumA & LumB
pale turqouise
YES1
SSH3
SLC25A24
MAPT
HSPB1
DCTPP1
DCTN3
CRIP1
APOO
ANXA9
ABAT
LumA
LumB
HER2
TNBC
MBC
Log2(FC)
4
2
0
-2
-4
C
WGCNA Signatures for HER2
light cyan
dark orange
RPL35A
RPL32
RPL21
RPL18
RPL15
BNIP1
PLIN2
MIEN1
GRB7
LumA
LumB
HER2
TNBC
MBC
Log2(FC)
4
2
0
-2
-4
D
WGCNA Signatures for TNBC and MBC
white
royal blue
green
PRTN3
MPO
MMP9
EPX
DEFA1
PLOD2
IGFBP3
CTHRC1
CRTAP
C1RL
SH3BGRL3
QKI
PTX3
PLOD3
CTSB
CPVL
ACTBL2
LumA
LumB
HER2
TNBC
MBC
Log2(FC)
4
2
0
-2
-4

Figure 7

A MPO MODULE



B **RNASE3 MODULE**



C