Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Lung-derived soluble factors reprogram the triple negative breast cancer phosphoproteome to drive metastasis via activation of CD44 and Rho-family signaling

**Vasudeva Bhat**

vbhat@uwo.ca

Western University

**Urvi Patel**

Western University

**David Goodale**

London Health Sciences Centre Research Institute

**Carl Postenka**

London Health Sciences Centre Research Institute

**Matthew Piaseczny**

Western University

**Daniel Menyhart**

Western University

**Laszlo Gyenis**

Western University

**David W. Litchfield**

Western University

**Alison L. Allan**

Western University

**Research Article**

**Keywords:** Triple negative breast cancer, lung metastasis, lung-derived soluble factors, phosphoproteomics, Rho GTPase, CD44

**Posted Date:** November 19th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11474787/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

## Abstract

# Background

Triple negative breast cancer (TNBC) is an aggressive and highly metastatic subtype of breast cancer. Lung is one of the most common sites for TNBC metastasis, suggesting that the lung microenvironment provides a conducive microenvironment or “niche” for metastatic colonization of TNBC cells. There are limited therapeutic options for treating or preventing lung metastasis, resulting in significant patient morbidity and mortality. This highlights an urgent need to better understand the molecular signaling between the metastatic microenvironment and disseminated TNBC cells.

# Methods

To elucidate microenvironment-driven intracellular signaling events, we performed integrated quantitative global proteomics and phosphoproteomics on TNBC cells exposed to lung-conditioned media from either tumor naïve (tnLCM) or primary TNBC tumor bearing mice (tbLCM).

# Results

Global proteomic profiling revealed an enrichment of pathways associated with extracellular matrix (ECM) remodelling, MET signaling, and metabolic remodelling in TNBC cells exposed to tbLCM. Phosphoproteomic changes in tbLCM exposed TNBC cells resembled signatures associated with altered PI3K signaling, metastasis permissive phenotype and therapy resistance. Additionally, phosphoproteomic analysis demonstrated that tbLCM selectively induces phosphorylation of CD44 at S686 and enhanced components of Rho GTPase pathway. Furthermore, using previously identified lung-derived soluble factors, we generated ligand-receptor network, which identified a highly redundant, multi-ligand network of lung-derived factors secreted specifically in the presence of primary tumor that converged on the CD44 receptor, suggesting its activation in TNBC cells in the primed lung microenvironment and potentially driving metastatic colonization via Rho GTPase signaling axis.

# Conclusions

Together, these findings demonstrate that primed niche factors in the lung can reprogram the TNBC proteome and phosphoproteome to promote metastatic behavior. This is potentially achieved by activation of invasive CD44-Rho GTPase signaling, presenting a novel multi-nodal target for treatment of metastatic TNBC disease.

## Background

Despite advancements in early detection, prevention and therapy options, breast cancer remains the second leading cause of cancer-related death in Canada. Among the different breast cancer subtypes, triple negative breast cancer (TNBC) is the most aggressive and highly metastatic[1,2]. Patients with TNBC often display an enhanced propensity for lung metastasis, resulting in increased morbidity and mortality[3]. Current treatment options are non-curative in the metastatic setting, mainly due to a lack of understanding about how the metastatic microenvironment provides a conducive condition for survival and colonization of disseminated tumor cells (DTCs)[4,5]. Therefore, elucidation of the metastatic microenvironment and its interactions with the DTCs is imperative to design more effective therapeutic approaches to treat or prevent TNBC metastasis.

We previously demonstrated that TNBC cells with high aldehyde dehydrogenase activity ($ALDH^{high}$) and expression of CD44 ($CD44^{+}$), considered as a 'stem-like' phenotype, display enhanced metastatic potential and higher propensity to colonize the lungs[6]. Subsequently, we observed that the presence of a primary TNBC tumor 'primes' the lung microenvironment in preparation for metastatic colonization[7]. We further demonstrated that lung-derived soluble factors, including fibroblast growth factor 2 (FGF2) secreted in the presence of primary TNBC tumors can enhance the acquisition of stemness and drive lung colonization[8]. While FGF2 traditionally signals through canonical cascades such as the PI3K/AKT pathway, targeting these single signaling axes often fails in the clinical settings[9]. This suggests that the primed microenvironment orchestrates networks of alternate signals. However, the molecular events activated through the interaction between lung-derived soluble factors and the DTCs in regulating the survival, stemness and metastatic colonization of TNBC is poorly understood.

In this study, we observed that continuous FGF2 signaling contributed to the maintenance of metastatic behaviour in TNBC cells, and that inhibition of the FGF2/AKT signaling axis was compensated by potential activation of alternate signaling pathways. We next performed quantitative global proteomics and phosphoproteomics profiling of TNBC cells exposed to lung-derived soluble factors to further dissect the microenvironment cues triggering TNBC cell intrinsic changes. We observed that exposure to soluble factors from the primed lung microenvironment induces a proteomic shift towards extracellular matrix (ECM) remodeling, metabolic adaptation, and MET signaling pathway. Furthermore, at the post-translational level, the phosphoproteomic profile resembled a chemoresistant signature. Ligand-receptor modeling revealed that the primed metastatic niche may drive metastatic behaviour via a multi-ligand convergence on the CD44-Rho GTPase signaling axis. Overall, our findings provide crucial insights into the role of lung-derived soluble factors secreted in the presence of primary TNBC tumors in activating a coordinated signaling cascade that supports stemness and metastatic behaviour of breast cancer cells and offers novel insight about multi-modal therapeutic vulnerabilities.

## Materials and Methods

# Cell culture and reagents

Human MDA-MB-231 TNBC cells were obtained from American Type Culture Collection (ATCC, Manassas, VA, USA) and maintained in Dulbecco's modified Eagle's medium (DMEM)/F12 + 10% fetal bovine serum (FBS). Human SUM159 TNBC cells were obtained from Asterand Inc (Detroit, MI, USA) and maintained in HAMS:F12 + 5% FBS supplemented with 5µg/ml insulin, 1 µg/ml hydrocortisone and 1mM HEPES. All the cell lines were authenticated via third party testing (IDEXX BioAnalytic, Columbia MO, USA; 04/2024; 9 markers). Breast cancer cell lines stably expressing red fluorescent protein (MDA-MB-231-tdTomato and SUM159-mCherry) were maintained using puromycin as a selection marker.

## Generation of lung-conditioned media

Animal experiments were carried out in accordance with the Canadian Council of Animal Care under a protocol approved by the Western University's Animal Care Committee (protocol #2021-076). Lung-conditioned media was generated as previously described[8,10]. Mice bearing SUM159 tumors were sacrificed 4 weeks post mammary fat pad injections. Age-matched tumor naïve mice were used as controls. Tumor naïve and tumor-bearing mice were euthanized, and lung-conditioned media (LCM) were generated essentially as previously described[7,8,11].

## Immunoblotting

Human TNBC cells were exposed to base media control, tnLCM or tbLCM (SUM159) or vehicle control, recombinant mouse or human FGF2 (SUM159 and MDA-MB-231) for 72 hours. Total protein was isolated using 1X RIPA lysis buffer (Cat. No. 89900, Thero Fisher Scientific, USA) containing 1X protease inhibitor (Cat No. 78430, Thermo Scientific, USA) and 1X phosphatase inhibitor (Cat. No. 1862495, Thermo Scientific, USA). Lowry assays were performed to quantify total protein. Approximately 35 µg of protein per sample was subjected to SDS-polyacrylamide gel electrophoresis and separated proteins were transferred onto a 0.45 µm nitrocellulose membrane (Thermo Scientific, IL, USA). Membranes were incubated with blocking solution (5% BSA or skimmed milk) to minimize non-specific binding. The membranes were then incubated with target-specific antibodies (**Supplementary Table 1**) diluted in 5% blocking solution at room temperature or overnight, followed by horseradish peroxidase conjugated secondary antibodies for 1 hour (**Supplementary Table 1**). Amersham ECL primer detection reagent (GE Healthcare, WI, USA) was used to visualize chemiluminescence in BioRad ChemiDoc™ MP imaging system. Antibody re-probing was performed by stripping the membrane using a stripping buffer (1.5% glycine, 0.1% SDS, 1% Tween 20, pH set to 2.0). The densitometric analysis was preformed using Image Lab software (v6.1, Bio-Rad Laboratories, Inc).

## Colony and organoid forming assay

SUM159 and MDA-MB-231 TNBC cells were plated at a density of ~ 20 cells/cm$^2$ in 9.6 cm$^2$ cell culture dishes. The following day (16–20 hours after seeding), cells were treated with either DMSO (vehicle

control; 0.1%) or varying concentrations (0.01μM, 0.1μM, 1μM, 10μM or 100μM) of AKT inhibitor (AKT VII, Cat. no.124018, Sigma Aldrich, Darmstadt, Germany). Colonies formed after 7 days were fixed with methanol:acetone (1:1 v/v), stained with crystal violet and total number of colonies (defined as > 50 cells) were counted. For 3D organoid formation assay, 20,000 MDA-MB-231 and SUM159 cells were mixed in 50 μl of Culturex reduced growth factor basement membrane extract, type 2 (Bio-Techne, Canada), and cultured in organoid media as previously described[12]. After 24 hours of culturing, varying doses of AKT inhibitor (AKT VII, Isozyme-selective, Cat. No. 124018, Sigma Aldrich, Darmstadt, Germany) were added with DMSO as a control. Changes in organoid formation were assessed in real time using an IncuCyte live-imaging instrument (Sartorius Canada, Oakville, ON, Canada). Changes in total area of organoids ($\mu m^2$) at different time intervals in the presence of varying doses of AKT inhibitor were calculated using default settings.

## In vivo mouse experiments

All *in vivo* animal experiments were carried out using 6–8 weeks old female NOD/SCID mice. All animal procedures were carried out as per the recommendations of the Canadian Council on Animal Care, under a protocol approved by Western University's Animal Care Committee (protocol #2021-076). Human MDA-MB-231 TNBC cells pre-treated with 100 ng FGF2, or vehicle control for 3 hours were resuspended in sterile Hank's buffered saline solution at 5 x $10^6$ cells/ml. 200 μl (1 x $10^6$ cells) were delivered via tail vein injection (*n* = 6/group). Metastatic growth in the lung and extrapulmonary organs was allowed to develop for 6–7 weeks and mice were sacrificed to assess metastatic burden. All tissues were harvested and examined for visible macroscopic metastases. Tissues collected were fixed in 10% neutral-buffered formalin, paraffin embedded and sectioned (4μm). Lung sections were subjected to standard hematoxylin and eosin staining and evaluated by light microscope in a blinded fashion to identify and characterize regions of micrometastases (100–400μm) and macrometastases (> 400μm).

# Ex vivo pulmonary metastasis assay (PuMA)

Human MDA-MB-231 and SUM159 TNBC cells were resuspended in 1X sterile PBS at a concentration of 5 x $10^6$ cells/mL and 100 μL (5 x $10^5$ cells) of cell suspension was injected into the lateral tail vein of 6–8 week-old female nude (SUM159) and NOD/SCID (MDA-MB-231) mice (*n* = 4 mice/group). Fifteen minutes post injection, lungs were isolated and subjected to a *ex vivo* pulmonary metastasis assay (PuMA) as described previously[8,13,14]. Lung sections cultured in the *ex vivo* PuMA (days 7, 14 and 21 post injection) were treated with either DMSO control or varying doses of AKT inhibitor (2.5μM, 5.0μM or 10μM AKT inhibitor). Three lung sections for each mouse (*n* = 3) were immediately fixed in 10% buffered formalin phosphate + 25% sucrose (w/v) with overnight incubation. At the end of the experiment, lung sections from each time points were rinsed with PBS three times and imaged using an upright Nikon A1R confocal microscope at 10X objective, with a 591 nm emission laser with pin hole set to 10. Three separate lung sections were imaged per timepoint with three images taken per lung section. To assess growth and colonization of breast cancer cells, mean fluorescent area per field of view (FOV) for each

lung section ($\mu m^2$) was measured using ImageJ software (v2.16.0/1.54p, NIH, Bethesda, WA, USA). Data were normalized to 1000$\mu m^2$ at day 0 to account for variability in cellular delivery during tail-vein injection. The concentration of FGF2 secreted from unseeded lung tissue was measured using a mouse bFGF ELISA kit (RayBiotech, USA) according to the manufacturer's instructions.

## Label-free proteomics and phosphoproteomics analysis

For proteomic and phosphoproteomic analysis, SUM159 TNBC cells were exposed to base media (BM), LCM from tumor naïve (tnLCM) and tumor bearing mice (tbLCM) for 24 hours (*n* = 3 biological replicates). Cells were harvested, snap-frozen and shipped to the Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital (Toronto, ON) to perform label-free quantitative proteomics and phosphoproteomics. Cell pellets were lysed with 500 µL 5% SDS, 50 mM triethylammonium bicarbonate (TEAB) lysis buffer. Samples were sonicated at 30% amplitude for 15 seconds (5 sec on, 3 sec off for three cycles) with 1/8" microtip. Samples were centrifuged at 16,000 xg for 5 minutes. Supernatant was moved to a new tube and 1/2 of the clarified lysate was used for S-Trap. One half of protein material (250 µL) (in 5% SDS, 50mM TEAB) plus an additional 250 µL lysis buffer (500 µL total) was reduced with 20 mM DTT for 10 minutes at 95 ºC and alkylated with 40 mM iodoacetamide for 30 minutes in the dark. Samples were brought up to final concentration of 5% SDS and trifluoroacetic acid (TFA) was added to a final concentration of 0.9%. 3.3 mL of S-Trap protein binding buffer (90% methanol, 100 mM TEAB) was added to 550 µL of acidified lysate. The resulting mixture was passed through the micro column at 4000 xg. The midi column was washed 3 times with the S-Trap protein binding buffer. Each sample was digested with 30 µg of trypsin in 350 µL digestion buffer (50 mM TEAB) for 1hr at 47 ºC. Prior to elution, 500 µL of 50 mM TEAB pH 8 was added to the column. Peptides were eluted by centrifugation at 4000 xg. Peptides were eluted 2 more times with 500 µL 0.2% formic acid and 500 µL of 50% acetonitrile + 0.2% formic acid. Eluted peptides were dried down and stored at -40 ºC.

MagReSyn Ti-IMAC beads were used for phosphoenrichment. The protocol was followed as per manufacturer's instructions for 100 µg protein digest. Lactic acid (0.1 M) was used in place of glycolic acid (0.1 M) in the loading buffer. Two washes were performed with wash buffer 2. A second elution was also performed, as per manufacturer's instructions (MagReSyn TiIMAC PDF https://resynbio.com/ti-imac/). Peptides were cleaned up using Evotips as per manufacturer's instructions (https://www.evosep.com/wp-content/uploads/2020/08/Sample-loading-protocol_A6_v4_28.02_WEB.pdf), and eluted 2x with 30 µL of 66% acetonitrile/0.1% formic acid. Cleaned-up peptides were dried down and stored at -40 ºC. For mass spectrometry acquisition, samples were resuspended in 5% formic acid.

For data-independent acquisition (DIA) LC-MS/MS, 250 ng of digested peptides for total proteome or all of the Ti-IMAC enrichment were analyzed using a nano-HPLC (High-performance liquid chromatography) coupled to MS. Samples were loaded onto Evotip Pure as per the manufacturer's instructions. Peptides were eluted from the column (Cat. No. EV-1137, 15 cm x 150 µm with 1.5µm beads) with the 30SPD pre-

formed acetonitrile gradient generated by an Evosep One system and analyzed on a timsTOF Pro 2. The column toaster was set to 40 ºC, with a total DIA protocol of 44 minutes. The MS1 scan had a mass range of 100-1700Da in dia-PASEF. TIMS settings were accumulation and ramp time of 100ms, and within the mobility range (1/K0) of 0.6 to 1.6V·s/cm2. Cycle time 2 seconds. For MS2, 1 mobility window with 2 ramps was used for 32 mass windows, 29.8 Da wide with 5Da mass overlap. The mobility range was from 0.61/K0 to 1.451/K0 at a duty cycle of 100% and a ramp rate of 9.52Hz. 1 + ions were excluded from fragmentation using a polygonal filter. Auto calibration was off.

## Mass spectrometry data analysis

Spectronaut v19.0 directDIA + BGS Phospho PTM workflow was used to search the data with the Spectronaut generated human (Human_PDB_2023) and mouse (Mouse_PDB_2023) spectral library. Parameters for the search were default. Data normalization was performed through cross-run normalization for both proteomics and phosphoproteomics data. The protein, peptide, and PTM site reports have been log transformed prior to all analysis. A cutoff of 0.75 was used for site localization. Of 8,881 protein groups in the raw Spectronaut output, 7,582 were human only, 776 contained both human and mouse entry (shared), 472 were mouse-only, 47 were contaminants and 4 could not be mapped. About 9% of protein groups had ambiguous, multi-protein group assignments (**Supplementary Table 2**). In all analysis, mouse only proteins were excluded and human only + shared proteins were included. For all downstream analyses (pathway enrichment, Kinase substrate enrichment analysis, NicheNet, PTM SignalR and others), each protein group was represented by a single gene symbol, the first-listed gene of the group (in case of shared, human gene symbol was considered). Gene symbols with multiple entries were collapsed by taking the median.

Differential abundance of proteins and phosphosites was assessed by Spectronaut's unpaired, two-sided t-test, with multiple-testing correction (q-value). Significance was defined by $q < 0.05$, $|\log2FC| > 0.58$. For the proteome-level analysis in R, missing values were imputed and moderated two-sided t-statistics were computed using limma (linear model with a group-means design and contrast, empirical Bayer moderation), with p-value adjusted by the Benjamini-Hochberg (BH) method. Over-representation analysis (Reactome, KEGG) used a one-sided hypergeometric test vis clusterProfiler, with BH adjustment and all quantified proteins as the background. Gene-set enrichment of phosphosite-derived rankings was performed with fgsea (permutation-based, ranked by log2fold-change), and PTM-SEA was run with ssGSEA2 (1,000 permutations, minimum overlap of 5 sites). False discovery rate (FDR) < 0.25 and $|NES| > 1$ was used to obtain fgsea enriched pathway list. KSEA z-score were converted to one-sided p-values from the standard normal distribution and adjusted by BH. Survival associations were tested using Cox proportional-hazards regression (two-sided Wald test), with BH adjustment across the screen.

R Studio (v2026.01.2 + 418) was used to perform downstream analysis, pathways enrichment analysis and plot generation. Principal component analysis (PCA) was performed on log2-transformed data. Network and interaction diagrams were generated using circlize v0.4.18 (chord diagrams), ggalluvial

(v0.12.6; alluvial diagram), ggraph (v2.2.2; force-directed networks), and ComplexHeatmap (v2.28.0). Figures generated in R (v4.6.1) used ggplot2 package (v4.0.3).

# Pathway Enrichment Analysis

Over-representation analysis was performed using clusterProfiler v4.20.0 (enrichGO, ontology:Biological Process; enrichKEGG) and ReactomePA v1.56.0 (enrichPathway), against the detected proteins with Benjamini-Hochberg corrected (adjusted $p \leq 0.05$; $q \leq 0.2$). All tested protein sets were retained internally to allow inspection of non-significant results alongside significant ones; only adjusted $p \leq 0.05$ terms were plotted. Pathway analysis used gene symbols, converted to Entrez IDs via bitr(). Gene lists for enrichment used standard deduplication, unique(), retaining the first-encountered measurement.

Rank-based pathway analysis was additionally performed using fGSEA v1.38.0 on the complete ranked phosphosite list (ranked by maximum |log2fold-change| per protein across all detected sites, not restricted to significant sites), against a combined Reactome, Hallmark, and KEGG database. Pathways were classified as activated or inhibited by the sign of the normalized enrichment score (NES; FDR ≤ 0.25, |NES| ≥1).

# Epithelial-to-mesenchymal transition (EMT) scores

EMT scores were calculated for each sample by single-sample Gene Set Enrichment Analysis (ssGSEA, GSEA v2.6.3; ssgseaParam, minimum gene-set size 10). Protein abundance matrix, log2-transformed and median centred values per sample were used as input. Proteins quantified in fewer than 50% of samples were removed, and remaining missing values were imputed with the median of that protein across samples. Gene sets were used from MSigDB (msigdbr v26.1.1). The EMT score was calculated using ssGSEA score of the Hallmark of Epithelial-Mesenchymal Transition gene set alone, calculated on the same matrix.b Group differences were assessed by one-way ANOVA with Turkey's HSD post hoc test ($n$ = 3 per condition).

# Candidate upstream ligand prediction

Candidate upstream ligands were predicted using NicheNet[15], developed for transcriptomic data. Proteins significantly changed ($q \leq 0.05$, $|log2FC| \leq 0.58$) in each comparison were treated as genes of interest. Spectronaut protein groups were mapped to official HGNC gene symbols, using the first gene name listed for each protein group. Multiple protein groups mapped to same gene symbol were collapsed to one value per symbol by taking the median q-value and the median log2 fold change. All ligands in NicheNet's reference database were tested as candidates. Ligands with no informative target genes or fewer than 1000 target genes in the reference database in the reference model were excluded prior to ranking. Ligand activity was scored using the background-corrected area under the precision-recall curve (AUPR).

# Survival analysis

Significant proteins unique to tnLCM vs BM or tbLCM vs BM ($q \leq 0.05$, $|log2FC| \leq 0.58$) were screened in two public breast cancer cohorts, METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) and TCGA-BRCA (The Cancer Genome Atlas – Breast Invasive Carcinoma) accessed through cBioPortal using cBioPortalData (v2.18.2): METABRIC (brca_metabric) and TCGA-BRCA PanCancer Atlas 2018 (brca_tcga_pan_atlas_2018), accessed on July 18, 2026. Gene symbols were mapped to Entrez IDs, and patients with more than one sample, expression values were averaged. Patients were divided into high- and low-expression groups at the cohort median for each gene. Overall survival (OS) was analyzed in both cohorts. Recurrence-free survival (RFS) was used in METABRIC, and progression-free survival (PFS) was used in TCGA-BRCA. Associations were tested by univariable Cox proportional-hazards regression (survival v3.8-6), with hazard ratios (HR), 95% confidence intervals and two-sided Wald-test p-values were reported. Genes with fewer than 10 evaluable patients were excluded. P-values were adjusted by the BH method across all genes within each cohort and outcome ($q \leq 0.05$). Kaplan-Meier curves were drawn with survminer (v0.5.2), with p-value from the log-rank test.

# Concordance and integration

Significantly changed phosphosites ($q \leq 0.05$, $|log2FC| \leq 0.58$) were classified according to the differential abundance of their parent protein in the same comparison (tnLCM vs BM and tbLCM vs BM), with sites matched to proteins by gene symbol. A site was classified as protein-driven when the parent protein was significantly changed in the same direction. A site was classified as discordant when the parent protein was significantly changed in the opposite direction. A site was classified as signaling-driven when the parent protein was not significantly changed or was not detected in the proteome. As a complementary check, phosphosite intensities were normalized to their parent protein by subtracting the log2 protein abundance from the log2 phosphosite intensity within each sample. Sites without a quantified parent protein, or with fewer than three non-missing normalized values, were excluded. The normalized values were fitted with a robust linear model (msqrobLm; model: ~condition). The contrasts tnLCM vs BM, tbLCM vs BM and tnLCM vs tbLCM were tested, and p-values were adjusted by the BH method. For each originally significant site, the direction of change and the adjusted significance (adjusted $p < 0.05$) before and after normalization were compared.

# Kinase activity inference

Kinase-Substrate Enrichment Analysis (KSEA) was performed using KSEAapp[16], with NetworKIN-predicted kinase-substrate relationships (NetworKIN score cutoff = 5). Phosphosites were matched to the KSEA substrate database by gene symbol and phosphorylated residue number only (flanking sequences were not used). The gene symbol was the first gene name listed for the protein group, and the residue and phosphosite (e.g., S686) were taken from Spectronaut's site identifier, referenced to the UniProt accession. Uniqueness of every Gene+Site combination was confirmed (0 duplicates across all comparisons) to prevent site-level duplication. Spectronaut's PTM-level output preserved ambiguous phosphosite assignments as fully distinct entries, rather than collapsing to a single identifier. Kinases were classified as activated or inhibited by KSEA z-score sign. A reliability filter of ≥ 5 detected substrate and FDR ≤ 0.05 was applied for generating the primary kinase-activity plots. Kinase-substrate network

diagrams were built from the same KSEAapp substrate table using ggraph (v2.2.2; Fruchterman-Reingold force-directed layout, fixed random seed reproducibility).

# Post-translational modification signature enrichment analysis (PTM-SEA)

PTM-SEA analysis was performed using ssGSEA2 implementation (ssGSEA2, v1.0.0) against the PTMsigDB database (v2.0.0, human, UniProt-centric site identifiers), containing site-specific, directional signatures of kinase activity (KINASE), perturbation (PERT), pathway (PATH) and disease (DISEASE). All quantified phosphosites (serine, threonine and tyrosine, STY) in each comparison (tnLCM vs BM, tbLCM vs BM and tnLCM vs tbLCM) were used. Each site was represented by its residue and position and the gene symbol of its protein group, and the gene symbol was converted to the canonical reviewed UniProt accession to form PTMsigDB-style identifiers. The site-level log2 fold-change of each comparison was used as the input score, with one input column per comparison. PTM-SEA was run with rank-based weighing (weight = 1), the area-under-the-running-enrichment-score statistic, a minimum overlap of five sites per signature, and 1,000 permutations. Normalized enrichment scores (NES) and FDR were calculated for each signature within each comparision, and signatures with FDR < 0.25 were reported as significant. Results were summarized separately for each signature category, PATH, KINASE and PERT.

# Ligand-receptor-phosphosite network

Ligand-receptor networks were constructed using PTMSignal R (v0.0.10)[17]. Post-translation modification (PTM) data was integrated into ligand-receptor pathway inference. For each pairwise comparison (tnLCM vs BM, tbLCM vs BM and tnLCM vs tbLCM), protein and phosphosite intensity matrices were derived from protein and PTM outputs derived from Spectronaut (v19.0) analysis, restricting phosphosite data to residues annotated as “Phospho (STY)”. Phosphosite intensities were used as measured (site-level abundance) for differential analysis. Ligand-receptor pairs were extracted from the package’s built-in LRdb reference database[18], integrated with Reactome Neo4j-derived PTM reference database. Significant ligand-receptor interaction was Benjamini-Hochberg corrected; q ≤ 0.05 was used as the significance threshold throughout.

# Statistical analysis

All *in vitro* experiments were performed in triplicate ($n$ = 3) unless otherwise stated. The proteomics and phosphoproteomics analysis were performed on three biological replicates. *Ex vivo* experiments were carried out using multiple mice (PuMA: $n$ = 4 mice/group/cell line), LCM generation: $n$ = 4 mice/group, *in vivo* metastasis assay ($n$ = 6/group). Statistical analyses were carried out using GraphPad Prism 6.0 (San Diego, CA, USA). Data are presented as the mean ± standard deviation (SD). A two-way analysis of variance (ANOVA) was used to compare multiple means across different groups. In all case, p values of ≤ 0.05 were considered to be statistically significant.

## Results

# Continuous lung-derived FGF2 signaling is required to sustain TNBC metastatic behavior

We previously demonstrated that lung-derived soluble factors (including fibroblast growth factor 2 (FGF2)) secreted in the presence of a primary TNBC tumor promote the acquisition of a stem-like phenotype ($ALDH^{high}CD44^{+}$) in breast cancer cells[8]. Furthermore, blocking FGF2 ligand activity significantly impaired the growth and colonization of these cells in the lung[8], suggesting that microenvironmental FGF2 signaling drives TNBC stemness and metastatic behaviour. Interestingly, FGFR signaling has been shown to promote stemness via AKT phosphorylation and subsequent SOX2 upregulation in pancreatic cancer[19]. Since lung-derived soluble factors also increased SOX2 and OCT4 levels in our previous study[8], we investigated whether exposure to lung conditioned media (LCM) induces AKT phosphorylation in TNBC cells. TNBC cells were treated with LCM from tumor naïve (tnLCM) and TNBC tumor bearing (tbLCM) mice, using base media (BM) as control. TNBC cells exposed to tbLCM exhibited enhanced phosphorylation of AKT (Thr308) as compared to tnLCM and BM, which was detectable at 24 hours and retained up to 72 hours (Fig. 1A**&B**).

To determine if AKT activity is essential for TNBC metastatic behavior *in vitro*, we performed 2D colony-forming and 3D organoid-forming assays using varying doses of AKT1/2 inhibitor (AKT inhibitor VIII). We observed a dose dependent reduction in colony- and organoid-forming ability of TNBC cells under 2D and 3D culture conditions (**Supplementary Figure S1A-D**). To validate this observation in the context of microenvironmental stimulation, TNBC cells were exposed to tnLCM and tbLCM in the presence or absence of a AKT inhibitor. Interestingly, pharmacological inhibition of AKT significantly reduced the tbLCM-induced upregulated FGF2 expression in TNBC cells (Fig. 1C**&D**). Additionally, we observed significant increase in FGF2 expression 48 hours post tbLCM treatment (Fig. 1E). This suggests an autocrine feed-forward loop where microenvironmental factors trigger AKT phosphorylation to drive endogenous tumor cell FGF2 expression.

To further test the functional consequence of AKT phosphorylation in TNBC cells in a more physiologically relevant microenvironment, we utilized an *ex vivo* pulmonary metastasis assay (PuMA) model. SUM159 and MDA-MB-231 TNBC cells were injected into tail vein of 6–8 weeks old mice, and lung sections were cultured *ex vivo* with either DMSO (vehicle control) or varying dose of AKT inhibitor VIII. Despite the significant dose-dependent decrease in survival observed in the *in vitro* 2D and 3D assays, AKT inhibition had no significant effect on the growth and colonization of TNBC cells in the lung over a 21-day period (**Supplementary Figure S1E&F**). Additionally, in unseeded lung tissue sections, we observed a gradual increase in FGF2 secretion over the 21 days of *ex vivo* culture suggesting a continuous baseline secretion of FGF2 in the lung microenvironment (Fig. 1F). This exogenous FGF2 may be compensating for AKT inhibition, suggesting that continuous microenvironmental signals likely drive lung colonization.

To determine if transient exposure to FGF2 is sufficient to influence metastatic behavior, MDA-MB-231 cells pre-treated *in vitro* with recombinant human FGF2 (rhFGF2) protein or vehicle control (VC) were delivered to the lungs of 6–8 week NOD/SCID mice via tail vein injection (n = 6/treatment). After 6–7 weeks post injection, mice were sacrificed, and lungs were harvested to assess metastatic burden. No significant difference in the lung metastatic burden was observed (**Supplementary Figure S2A-C**), suggesting that transient FGF2 exposure is in itself insufficient to support stemness and metastatic behavior of TNBC cells. Instead, continuous and multiple cues from the primed lung microenvironment are likely required to support metastatic colonization of TNBC cells.

# Lung-derived soluble factors induce proteome-wide changes linked to steroid metabolism, MET signaling and ECM signaling

FGF2 relays intracellular signals through both canonical and non-canonical receptors, thereby activating multiple downstream signaling cascades, including the PI3K/AKT, MAPK/ERK, and PLC$\gamma$ pathways[20]. However, the possibility that additional lung-derived soluble factors work in concert to influence TNBC metastatic behaviour suggests that a broader signaling network may orchestrate these cell intrinsic changes. To investigate this, we performed quantitative label-free proteomics and phosphoproteomics on TNBC cells exposed to BM, tnLCM and tbLCM for 24 hours. This specific timepoint was selected because AKT phosphorylation was observed starting at 24 hours post exposure (Fig. 1A**&B**).

An unsupervised Principal Component Analysis (PCA) of the global proteome revealed that tnLCM and tbLCM treated cells exhibited proteomic profiles distinct from those treated with BM (**Supplementary Figure S3A**). Notably, tnLCM and tbLCM groups showed only modest separation from each other along the principal components. This suggests that during early metastatic events, the presence of a primary tumor doesn't induce a strong shift in global proteome relative to tnLCM. Instead, the distinct phenotypic effect of tbLCM may be mediated by subtle proteomic changes, or through alterations at the post translational level and pathway activation status.

To identify specific changes in the global proteome, we assessed differential protein abundant proteins between conditions were assessed in Spectronaut v21.1 using an unpaired, two-sided t-test with multiple-testing correction (q-value). Proteins with $q < 0.05$, $|Log2FC| > 0.58$) were considered significantly changed. Our global proteome analysis identified 7,327 proteins in tnLCM-treated cells compared to BM (tnLCM vs BM). Among them 260 proteins were significantly upregulated, and 412 proteins were significantly downregulated (Fig. 2A). In tbLCM-treated cells (tbLCM vs BM) 7,314 proteins were identified, of which 181 proteins were upregulated, and 463 proteins were downregulated (Fig. 2B). Direct comparison between the two LCM-treated groups (tnLCM vs tbLCM), identified 143 abundant proteins elevated in tnLCM-treated cells, and 89 abundant proteins in tbLCM-treated cells (Fig. 2C).

Pathway analysis was next performed to characterize the biological functions of these proteomic changes. Reactome pathway analysis showed no significant enriched pathways in tnLCM-treated cells

compared to BM. The tbLCM-treated cells showed significant enrichment for pathways involved in steroid metabolism, cholesterol biosynthesis, ECM remodelling and MET signaling (Fig. 2D**&E**). However, in tnLCM-treated cells, only complement and coagulation cascade was enriched in KEGG pathway analysis (**Supplementary Figure S3B**). Consistent with findings from Reactome pathway analysis, KEGG pathway analysis confirmed that steroid and terpenoid backbone biosynthesis was highly enriched in tbLCM treated TNBC cells (**Supplementary Figure S3C**). We also observed p53 signaling pathway to be significantly enriched (**Supplementary Figure S3C**). Additionally, assessment of epithelial-mesenchymal transition (EMT) scores revealed that, while exposure to both tnLCM and tbLCM showed an upward trend in the EMT score in TNBC cells relative to BM, only tbLCM exposed appear to significantly increase the score (**Supplementary Figure S3D**). Taken together these global proteomic analyses suggest that tbLCM exposure induces activation of metabolic pathways alongside structural adaptations, potentially providing a crucial microenvironmental link to stemness, EMT and metastasis in TNBC cells.

We next assessed the clinical relevance of uniquely expressed proteins following respective LCM exposure. To determine their association with patient outcomes, proteins unique to each condition were screened in two independent breast cancer cohorts, METABRIC and TCGA-BRCA via the cBioPortal database. Overall survival (OS) was tested in both cohorts, and recurrence-free survival (RFS) and progression-free survival (PFS) were assessed in METABRIC and TCGA-BRCA respectively. In tnLCM treated cells, higher expression of ALG3, CENPA and KIF23 were associated with poor overall survival (OS) and recurrence-free survival (RSF) (**Supplementary Figure S4A-F**). Additionally, higher expression of SRFBP1 was associated with poor OS in breast cancer patients (**Supplementary Figure S4G**). Lower expression of IRS1 (**Supplementary Figure S4H**) was associated with poor OS, and lower expression of MAPT (**Supplementary Figure S4I**) was associated with poor RFS. In tbLCM treated cells, lower expression of IDNK (**Supplementary Figure S4J&K**) was associated with poor OS and RFS. Additionally, lower expression of EVA1B and SPRY2 (**Supplementary Figure S4L&M**) was associated with poor OS, and lower expression of KLHL5, MORC2, NUDT18 and STAT3 (**Supplementary Figure S4N-Q**) was associated with reduced RFS. Taken together, these findings indicate that treatment with different LCM resulted in differential expression of proteins associated with poor OS and RFS.

**Computational modeling predicts selective enrichment of TGFβ1 and TNF driven signals in the primed metastatic niche**

Next, we sought to identify exogenous ligands responsible for initiating downstream signaling cascade within the lung microenvironment. To predict potential lung-derived ligands that drive these intracellular changes, we employed NicheNet[15], a computational method designed to infer upstream ligands based on downstream gene expression changes. Although NicheNet was originally developed and validated using transcriptomics data, we applied it here to our global proteomics dataset by considering protein significantly changed ($q < 0.05$, $|Log2FC| > 0.58$) as genes of interest. The use of this method can be biologically justified given that our proteomics profiling was conducted 24 hours post treatment, a timeframe sufficient for upstream ligand-receptor interactions to translate into altered downstream protein abundance.

Analysis of the NicheNet ligand activity scores revealed strong correlation (Pearson r = 0.93) between tnCLM and tbLCM groups across the entire predicted ligand set, suggesting a highly similar relative ranking of active ligands (**Supplementary Table 3**). We observed a distinct ligand regulatory profile between tnLCM and tbLCM relative to BM (Fig. 2F**&G**). In the tnLCM treatment group, growth factor signaling related ligands including EGF (r = 0.1120), FGF1 (r = 0.1024), OSM (r = 0.1002), and VEGFA (r = 0.1083) were predominant ligands predicted based on perturbation of downstream effector including PLAUR, TIMP1, CD55, and EREG (Fig. 2F). In contrast, in the tbLCM treatment condition, TNF (r = 0.1035), BMP2 (r = 0.0996), and TGF-β1 were predicted as upstream ligands, exhibiting high regulatory potential across a broad target spectrum, including upregulated matrix remodeling factors and extracellular protease such as SERPINE1, PLAU, ICAM1, and COL1A1 (Fig. 2G). Furthermore, TNF and IL1A were also predicted as significant ligands mapped to downstream effectors including TNFAIP2, TNFAIP3, JUNB, and SOD2. Notably, PDGFB (r = 1.1017) was specifically associated with downstream inflammatory target upregulation, including SDC4 and AKAP12, as well as cytoprotective enzymes, HMOX1 and SOD2 (Fig. 2G). Taken together, the computational modeling results suggest that while tnLCM exposure results in EGF-driven signaling, tbLCM exposure results in highly specific phenotypic changes driven by TNF, PDGFB and TGF-β dominant signaling that promotes pro-metastatic phenotype in TNBC cells.

# Phosphoproteomics analysis reveals secretome-specific reprogramming independent of protein abundance changes

To further elucidate the functional consequences of cellular reprogramming induced by tbLCM treatment in TNBC cells, we also assessed phosphoproteomic changes. Consistent with our global proteomics findings, unsupervised PCA of the phosphoproteome showed a clear separation of LCM-treated samples from BM-treated controls after 24 hours (**Supplementary Figure S5A**). However, tnLCM and tbLCM-treated samples exhibited a heterogenous, and subtle phosphorylation shift rather than a very distinct separation (**Supplementary Figure S5A**).

Compared to BM, tnLCM treatment resulted in 4,802 altered phosphosites, while tbLCM treated resulted in 4,411 phosphosites (Fig. 3A). Among these, 1,439 proteins sharing the same post-translational modifications (PTMs) were common to both tnLCM and tbLCM treatment (Fig. 3B). Whereas 53 common proteins displayed distinct, treatment-specific phosphosites, there were 398 and 324 phosphorylated proteins unique to tnLCM vs BM and tbLCM vs BM comparisons, respectively (Fig. 3B). To determine whether phosphosite-level changes reflected underlying changes in protein abundance or independent, signaling-driven events, we classified each significantly altered phosphosites ($q < 0.05$, $|\log2FC| > 0.58$) by comparing its fold-change to that of its parent protein in the same comparison. Phosphosites were classified as protein-driven (concordant with a significant change in parent protein abundance), signaling-driven (significant phosphorylation change with no corresponding change in protein abundance), or discordant (phosphosite and protein changes in opposite directions). Across the tnLCM vs BM and tbLCM vs BM comparisons, 88.6% and 91.5% phosphosites, respectively, were

classified as signaling-driven as compared to 7.1% and 6.3%, respectively, being protein-driven (Fig. 3C**&D**). To confirm this observation, an independent statistical approach was used. We additionally applied protein-abundance correction (log2 phosphosites – log2 protein, per sample) followed by robust linear modeling. Across all originally significant phosphosites with a quantified parent protein, 93.9% and 94.5% in tnLCM vs BM and tbLCM vs BM comparisons, respectively, preserved their original direction of change after correction. However, due to small sample size ($n$ = 3), a smaller subset of sites was retained (4.3% and 2.2% in tnLCM vs BM and tbLCM vb BM comparisons, respectively) as signaling-driven (**Supplementary Figure S5B**).

Comparative analysis of phosphosites significant in both conditions ($q < 0.05$, $|log2FC| > 0.58$) identified a subset of phosphosites exhibiting opposite directional changes between tbLCM and tnLCM-treated cells. Specifically, key regulators of Rho GTPase effector network, cytoskeletal dynamics, mitogenic signaling and cellular metabolism were uniquely showed increased phosphorylation in tbLCM-treated cells while being unphosphorylated in tnLCM-treated cells (Fig. 3D, **Supplementary Figure S5C**). These included MARCKS (S167), LIMCH1 (S231), MAPK1/ERK2 (T190), and the key metabolic enzyme PDHA1 (S300). Conversely, proteins such as MAP4 (T282), BAG3 (S386), and PIEZO1 (S1396) showed inverse pattern of phosphosites being downregulated in tbLCM and upregulated in tnLCM-treated cells (Fig. 3D).

Notably, tbLCM treatment selectively increased phosphorylation of Radixin (RDX; T564), a member of the Ezrin-Radixin-Moesin (ERM) protein family, which functions as the primary linker tethering the CD44 cytoplasmic tail to the actin cytoskeleton upon phosphorylation[21]. Concurrently, we observed increased phosphorylation of Rho GTPase effector, CDC42EP4 (T23) as well as actin-crosslinking Src activator AFAP1 (T689). Furthermore, proteins related to actomycin contractility, including MYL12B (T19/T20), JCAD (S696) and vinculin (VCL; S288, S290) was selectively observed in tbLCM-treated cells. Similarly, core RNA processing and nuclear matrix factors such as SRRM1 (S754, S756) and LARP7 (S99, S300) exhibited decreased phosphorylation in response to tbLCM treatment (**Supplementary Figure S5C**). These results suggest an activation of a pro-invasive and pro-metastatic signatures, particularly the potential activation of a CD44-ERM-Rho GTPase effector cascade in cells exposed to soluble factors from primed lung microenvironment. In contrast, tnLCM-treated cells were uniquely enriched for phosphosites associated with cytoskeletal stabilization and organelle homeostasis. Phosphosites displaying increased phosphorylation included TBC1D108 (S658), SERBP1 (S394), CLASP1 (S636), and ANLN (S72) (**Supplementary Figure S5D**). Conversely, dynein adaptors and mitochondrial fission regulators including, BICD2 (S318) and DNM1L/DRP1 (S607) displayed decreased phosphorylation in tnLCM-treated cells. Taken together, these results suggest microenvironmental-specific reprogramming of TNBC cells.

To identify perturbed pathways, we performed pathway enrichment analysis. We only used proteins that had a phosphosite identified in each comparison. We found that tnLCM exposure (compared to BM) induced mitotic and homeostatic signatures (**Supplementary Figure S5E**) while sharing some baseline cellular response pathways with tbLCM treated condition (Fig. 3E). There were distinct enrichment clusters related to nuclear pore complex disassembly, SUMOylation, and transcriptional regulation.

Furthermore, activation of ATR response to replication stress was negatively enriched (**Supplementary Figure S5E**). These results suggest that tnLCM exposure maintains homeostatic cell-cycle checks. In tbLCM-treated cells, we observed a pro-inflammatory and stress signature and a unique and predominant enrichment of the TNFA signaling pathway (Fig. 3E). Additionally, we observed a positive enrichment of RNA processing and stress responses related pathways (Fig. 3E). Together, the pathway enrichment analysis suggested that lung-derived soluble factors in the primed microenvironment, upregulate survival, transcription, and structural remodelling pathways.

Since kinases play a critical role in regulating PTM dynamics and downstream signaling cascade, we analysed upstream kinase activities based on the phosphorylation profile of the substrates present in our phosphoproteomic data. Kinase substrate enrichment analysis (KSEA) of the tnLCM treated samples (compared to BM), identified CDK1 (Cyclin-dependent kinase1) as a kinase with a positive KSEA score (Fig. 3F). In tbLCM-treated samples, CDK1 kinases exhibited a similar increase in KSEA score observed in the tnLCM condition (Fig. 3G). We also found CSNK2A1 (Casein kinase 2 alpha 1) to be negatively enriched in both comparisons (Fig. 3F**&G**). However, TTK (Threonine tyrosine kinase) was uniquely enriched in the tbLCM group, with a positive KSEA score (Fig. 3G). Direct comparison of KSEA profiles between tnLCM and tbLCM treatment groups showed ROCK1 (Rho associated coiled-coil containing protein kinase 1) kinase to be negatively enriched in upstream kinase activity when evaluated globally (Fig. 3H).

To gain further insight into these signaling cascade, we performed kinase-substrate network mapping. In tnLCM-treated cells, mapping of upstream kinases CDK1 and CSNK2A1 revealed a network with multiple significantly altered substrates (Fig. 4A). Both CDK1 and CSNK2A1 networks were linked to multiple substrates, and they were interconnected through LIG1 (DNA ligase 1) substrate (Fig. 4A). The CDK1 network was linked to substrates involved in cell cycle and signal transduction, including ANAPC1, PTPN1, CALD1, RANBP2, PI3KC2A, SQSTM1, HMGA1, FLNA, PKN2, PI4KB, BRCA1, MAP2K1 and SIRT2 (Fig. 4A). The CSNK2A1 network was linked to substrates involved in cell cycle, DNA repair and cytoskeletal remodelling, including EIF3J, ABCF1, HDAC1, HDAC2, JUN, MDC1, TP53, MCM2, IGF2R, MYL12B, EIF2S2, CAPZA1, and NPM1 (Fig. 4A).

Kinase-substrate network mapping of the tbLCM treatment condition (compared to BM) revealed three distinct multi-hub phosphoregulatory networks (Fig. 4B). While common CDK1 and CSNK2A1 substrates were identified in the tbCLM treatment condition (similar to those observed in tnLCM condition) we found that substrates involved in cell division, cholesterol biosynthesis, and protein folding including TPR, NCOR1, HMGCS1, SSR1, UIMC1, and YAP1 were unique to CDK1 kinase. Substrates involved in cell division including, PPP1R2, MYL12B, and CTDP1 were unique to CSNK2A1 kinase. The TTK kinase network consisted of cell division associated substrates BUB1B and MKI67 (Fig. 4B). Overall, compared to the tnLCM network, the tbLCM induced network appeared to be more interconnected, suggesting highly coordinated reprogramming of the phosphoproteome in TNBC cells in response to lung-derived factors induced in the presence of a primary tumor.

Finally, because we had earlier identified increased phosphorylation of AKT in TNBC cells exposed to tbLCM (Fig. 1A**&B**), we sought to identify specific substrates potentially phosphorylated by AKT kinase. We used AKT1/2/3 as a seed protein for kinase-substrate network generation for both tnLCM and tbLCM conditions. In the tnLCM condition, the AKT1 network consisted of unique substrates including FLN4, MDM4, and SLC9A1 (**Supplementary Figure S6A**). In the tbLCM condition, the network consisted of unique substrates including, CELF1, PFKFB2, NR3C1, AKT1S1, BAD, ZYX, BRAF and PEA15 (**Supplementary Figure S6B**). Substrates such as PALLD, GSK3B, IRS1, ZFP36L1, HSPB1, PIKFYVE, LMNA, and TBC1D4 were commonly present in the kinase-substrate network (**Supplementary Fig. 5A&B**).

While KSEA infers kinase activity from the aggregate behaviour of a kinase's known substrates, it fails to investigate whether the changes in phosphoproteome resemble a previously characterized signatures of specific perturbation in pathways or kinases. To complement KSEA and to test the specificity of individual phosphosite-level changes, we additionally performed PTM Signature Enrichment Analysis (PTM-SEA). Among the kinases identified as significantly altered by KSEA, two significant kinases identified in KSEA, CDK1 and CSNK2A1 were independently corroborated by PTM-SEA (Fig. 4C**&D**). While CDK1, MAPK1, MAPK3, MAPK8, MAPK9, GSK3B, PRKDC, and CSNK2A1 kinases were commonly activated in both tnLCM and tbLCM condition groups relative to BM, MSK1 (activated), SGK2 (activated), PRKCA (activated), PRKCI (activated), DYRK2 (inhibited) and PRKCH (inhibited) activity was specific to tbLCM (Fig. 4D). These results suggested a selective modulation of kinases by tbLCM.

Next, we applied PTM-SEA perturbation (PERT) signatures of PTMsigDB to correlate the global phosphorylation changes following tnLCM and tbLCM exposure to known signaling perturbations. In both tnLCM vs BM and tbLCM vs BM comparisons, phosphorylation profiles resembled signatures of kinase-inhibitor treatment (Fig. 4E**&F**). Kinase inhibitor (VX-970, PS-1145 and IPI145) signatures were negatively enriched in both conditions, suggesting activation of kinase pathway particularly PI3K pathway (negative enrichment of PI3K inhibitor IPI145) (Fig. 4E**&F**). Signatures such as the MEK inhibition by PD98059, p38 MAPK inhibition by losmapimod, phosphodiesterase 3 inhibition by Milrinone and IGF1 stimulation were positively enriched in tnLCM vs BM comparison and signatures such as EGFR inhibition (gefitinib), pan protein tyrosine phosphatase inhibition (vanadate), and transcriptional modifiers (EPZ-5687 and Lenalidomide) were negatively enriched (Fig. 4E). In tbLCM vs BM comparison, while the EGF treatment signature was positively enriched, signatures related to Taxol treatment, tyrosine kinase inhibition (TG101348, Lapatinib, Vemurafenib, Dactolisib) and resveratrol treatment were negatively enriched (Fig. 4F). Together, these signatures indicate a phosphorylation program resembling altered PI3K-related signaling in both conditions, with increased chemoresistant and active kinase signatures in tbLCM-treated cells.

# Lung-derived soluble factors from primed lung displays enhanced Rho GTPase signaling

Our earlier analysis of global phosphoproteome changes revealed that while majority of predicted intracellular signaling pathways were shared between tnLCM and tbLCM treatment groups, a smaller subset is enhanced in tbLCM-treated cells. To further evaluate this observation, we first used protein-containing significant phosphosites ($q < 0.05$, $|\log2FC| < 0.58$) in two groups and performed pathways analysis separately. Reactome pathway analysis revealed predominant enrichment of signaling pathways associated with Rho GTPase in both tnLCM and tbLCM comparisons (Fig. 5A).

To further investigate components of the Rho GTPase pathway altered in the presence of tbLCM, we examined phosphosite occupancy across different components of the Rho GTPases pathway (using a curated gene list compiled from Gene Ontology Rho-GTPase signaling terms) including guanine nucleotide dissociation inhibitors (GDIs), guanine nucleotide exchange factors (GEFs), GTPase-activating proteins (GAPs), effectors and cytoskeletal regulators, and the FAK-adhesion axis (Fig. 5B-G, **Supplementary Table 4**). While the majority of phosphosite changes were shared between the tnLCM and tbLCM treatment conditions, a smaller set of phosphosites changed significantly only in tbLCM-treated cells. Both phosphosites on myosin regulatory light chain 12B (MYL12B; T19 and S20) were significantly increased tbLCM-treated cells but remained unchanged in tnLCM-treated cells (Fig. 5C). Additionally, phosphosites on AFAP1 (T689), CDC42EP4 (T23), ROCK (S1374), and ZYX (S142) (Fig. 5D) were increased selectively in tbLCM-treated cells. GEF and GAP components displayed increased phosphorylation of AKAP13 (S2398 and S1282), ALSIN/ALS2 (T510), ARHGEF12 (S1327 and S599), TNS3 (S850), MYO9B (S1261 and T1271) in tbLCM-treated cells (Fig. 5E**&F**). Among FAK-adhesion axis components, we found that PTK2, (FAK) displayed increased phosphorylation (S910) in tbLCM-treated cells, however, its parent protein levels were decreased (Fig. 5D). We also observed a decrease in the phosphorylation of ARHGDIA (S24) (Fig. 5B), MAP4 (S859), SHROOM3 (S439), PARD3B (S749), MPRIP (S977) (Fig. 5C), vinculin (VCL; S288/S290), TLN1 (T1263) (Fig. 5D), ARAP3 (S1380), ARHGAP32 (S1585) (Fig. 5E), AKAP13 (S1559) (Fig. 5F), DENND4C (S1589), SIPA1L1 (S1116) (Fig. 5G), and MYH9 (S841) selectively in tbLCM-treated cells. While phosphorylation of CD44 at S697/S706 was decreased in both tnLCM- and tbLCM-treated cells, increased phosphorylation at S686 ($q = 0.056$ vs BM) was selectively observed in tbLCM-treated cells. Taken together, our analysis suggests that while both tnLCM and tbLCM activate Rho pathway signaling, tbLCM additionally enhanced the phosphorylation of selected actomyosin and adhesion regulators.

Since these signaling cascades require upstream microenvironmental stimulation, we sought to integrate these receptor profiles with our previously identified secretome of the ‘primed’ lung, where we identified unique soluble ligands secreted in the presence of the primary TNBC tumor[7,8]. To identify matching and functionally active receptors in our global phosphoproteomics datasets following tbLCM treatment, we utilized the LIANA (**LI**gand-receptor **AN**alysis fr**A**mework) computational package[22], which enables the prediction of both canonical and non-canonical receptor-ligand interactions. Through this analysis, we found that the 5 ligands identified in the primed lung in our earlier findings paired with receptors present in our global phosphoproteomics dataset (Fig. 6A). Notably, several of these ligands including dkk4, vegfa, thsd1, mmp3 and fgf2 were predicted to bind to two or more receptors, suggesting

activation of highly redundant signaling cascade (Fig. 6A). To determine whether the phosphorylation state of these receptors was modulated as a direct consequence of ligand binding, we generated an integrated ligand-receptor-effector signaling network (Fig. 6A**&B**). This network mapping revealed several key signaling hubs. The ligand tmeff1 converged on the ERBB2 receptor (decreased phosphorylation at threonine 1060) and the EGFR receptor (decreased phosphorylation at serine 1166). Additionally, metalloproteinase 3 (mmp3) similarly converged on ITGB3 and ITGAV integrin complexes (Fig. 6A**&B**).

Notably, we observed that the pro-metastatic ligands vegfa, mmp3, fgf2 all converged on the stemness-associated receptor CD44 (Fig. 6B). At the post-translational level, CD44 was found to have increased phosphorylation at serine 686 ($q = 0.056$ vs BM; $q = 0.044$ direct comparison with tnLCM) following tbLCM exposure only and concurrently decreased phosphorylation at serine 697 sites (common to both tnLCM and tbLCM groups). This suggests that multiple ligands in the primed lung may interact with CD44 to trigger its activation. In turn, this localized receptor activation likely acts as a major upstream trigger for downstream Rho GTPase signaling, providing ECM remodelling and stemness-related programming required to sustain TNBC colonization in the lung[23]. Taken together, our global proteomics, phosphoproteomics, and ligand-receptor networks analysis suggests that exposure to primed lung microenvironment activates a highly coordinated, multi-ligand signaling program designed to sustain TNBC stemness and metastatic behaviour.

# Discussion

Current TNBC therapies often fail in the metastatic setting, resulting in patient morbidity and mortality[24,25]. Although metastasis is the primary cause of cancer death, colonization of distant organ sites remains a rate-limiting process[26] and depends on a conducive microenvironment where disseminated tumor cells (DTCs) can thrive and eventually result in clinically impactful lesions. Lung metastasis is commonly associated with TNBC[24,27], suggesting that the lung microenvironment is enriched in supportive factors needed for survival and growth of breast cancer cells. Growing evidence supports a bidirectional relationship between DTCs and the metastatic microenvironment[28–30], where the metastatic microenvironment is capable of inducing tumor cell intrinsic changes that in turn contribute to continued evolution of a metastasis-promoting microenvironment. Therefore, dissecting the cellular and molecular interactions between these entities is critical to identifying key drivers of TNBC metastasis.

In this study, we demonstrated that TNBC cells rely on continuous FGF2 driven signaling to sustain metastatic behavior. Exogenous microenvironmental cues induce AKT phosphorylation, driving endogenous FGF2 expression in tumor cells, suggesting a potential feed-forward signaling mechanism. This feed-forward/autocrine loop could make TNBC cells self-sufficient over time, resulting in co-evolution and progression of a single DTC to micrometastases and macrometastases. Interestingly, our data also showed that FGF2 secretion in unseeded lung is a baseline biological state, suggesting that in

the primed lung, other factors in addition to FGF2 may orchestrate a multi-modal signal network required for metastatic colonisation. Furthermore, while AKT phosphorylation correlated with TNBC survival during *in vitro* colony/organoid formation, *ex vivo* functional assays showed that additional microenvironmental factors compensate for the loss of AKT signaling, underscoring the need to target niche factors alongside tumor intrinsic pathways.

Our global proteomics analysis revealed that the primed lung microenvironment actively rewires the TNBC proteome towards pro-metastatic potential. While the tumor naïve (tnLCM) and TNBC tumor bearing (tbLCM) lung-conditioned media both altered the proteome, tbLCM uniquely triggers ECM remodelling, MET signaling, and metabolic programming. Similarly, metabolic pathways such as steroid and cholesterol biosynthesis were significantly enriched in tbLCM treated cells. Emerging evidence suggests that cancer stem cells heavily rely on cholesterol metabolism and their metabolic derivatives can activate several evolutionary conserved signaling pathways[31]. This suggests that tbLCM-induced metabolic adaptations towards cholesterol biosynthesis may support survival of breast cancer stem cells in the lung. This also provides a strong rationale for targeting microenvironmental factors along with tumor cell intrinsic signaling to prevent metastatic growth and colonization.

Unsupervised PCA showed modest proteome separation between tnLCM and tbLCM treatments. In this study we used tbLCM that was generated 4-weeks post mammary fat pad injection to capture early changes in the lung microenvironment prior to tumor seeding in the lung. This modest change may indicate the early transition from a physiological 'normal' lung microenvironment to a potentially 'primed' metastatic niche. NicheNet[15] modeling identified potential ligand-receptor interactions that may be a characteristic of the global proteomic shift in TNBC cells exposed to lung conditioned media. We identified TNF, PDGFB and TGFβ1 ligands as potential upstream ligands present in the primed microenvironment. TGFβ1 is known to play multifaceted roles, including the maintenance of stemness and metastasis in breast cancer[32]. In contrast to TGFβ1, TNF exhibits dual roles, encompassing tumor promoting/pro-metastatic and antitumor effects[33–35]. Additionally, accumulating evidence also suggests a role for PDGFB in promoting breast cancer growth, angiogenesis and metastasis[36]. Notably, all three ligands, TNF, PDGFB and TGFβ1 are associated with EMT[37–39]. Consistent with these findings, our global proteomics analysis reveals a significant increase in the EMT score of TNBC cells exposed to tbLCM. Taken together, these results suggest that the primed lung microenvironment may secrete unique ligands to promote EMT, stemness and metastatic behavior in TNBC cells.

The phosphoproteomics changes observed following LCM treatment resembled signatures consistent with altered PI3K signaling (negative enrichment of IPI145 inhibitor signature). This further confirmed our early observation whereby inhibition of AKT did not affect metastatic colonization, potentially due to activation compensatory receptor tyrosine kinase signaling. Furthermore, tbLCM-specific phosphoproteome changes resembled signatures opposite to resvertarol treatment, which suppresses cell proliferation, EMT, metastasis and induces apoptosis in cancer cells[40]. Negative enrichment of resveratrol treatment signatures in tbLCM exposed TNBC cells suggests a pro-metastatic phenotype.

Additionally, resveratrol has been shown to directly inactive the RHO GTPase family member, RhoA in breast cancer cells[41]. Notably, the phosphoproteome of tbLCM-treated cells resemble a Taxol resistant signature. This suggests that the primed microenvironment induces early subtle changes that may make TNBC cells resistant to conventional chemotherapy agents. Rho GTPases are involved in coordinating actin dynamics, focal adhesion assembly (FAK), cellular motility and stemness[42]. Analysis of proteins with significant phosphosites revealed shared enrichment of Rho GTPase signaling in both tnLCM vsBM and tbLCM vs BM comparisons. However, tbLCM induced phosphorylation of unique proteins associated with the Rho GTPase pathway, suggesting a critical role of primed microenvironment in regulating this signaling cascade. Overall, our findings that tbLCM exposure may enhance Rho GTPase signaling suggests a key role in inducing stemness and metastasis in TNBC cells.

Previously we had identified lung-derived soluble factors secreted specifically in the presence of primary TNBC tumors[7]. When we modelled ligand-receptor networks by comparing identified ligands with the phosphoproteomics data, we found that ligands including fgf2, mmp3, and vegfa in the tbLCM converged on CD44 receptor. CD44 is a known breast cancer stem cell marker[43], documented to transduce FGF2 signaling[44]. CD44 interaction with its canonical ligand hyaluronan (HA) initiates downstream signaling including Rho GTPase[45,46]. CD44 signaling plays a key role in proliferation, cell motility, invasiveness and therapy resistance[47]. Interestingly, in our phosphoproteomic analysis, we identified increased phosphorylation of CD44 at Serine 686 ($q = 0.056$), which was specific to tbLCM group. While phosphorylation of CD44 (Serine 323 and 325) is key for HA binding and signal transduction[47], the specific role of Serine 686 has not been investigated. Our analysis strongly suggests a coordinated multi-ligand (fgf2, mmp3, vegfa) activation of CD44 and subsequent Rho GTPase signaling cascade (CD44-FAK-Rho/ROCK axis), which may be critical to support stemness and metastatic behaviour of TNBC cells. While this prediction is correlative, direct experimental validation, such as receptor blockade or ligand neutralization, would be required to distinguish between these possibilities.

# Conclusion

Taken together, our findings shows that the primed lung microenvironment supports TNBC progression. Through global proteomics and phosphoproteomics profiling, we uncovered that lung-derived soluble factors in the primed lung triggers a complex and highly coordinated signaling network encompassing PI3K signaling, and that Rho GTPase signaling and multi-ligand activation of CD44 signaling axis, which may collectively drive stemness and metastatic behavior. While tumor-intrinsic therapies such as AKT inhibition show potent *in vitro* efficacy, our *ex vivo* assays reveal that microenvironmental factors can compensate for AKT inhibition. This underscores the critical need to target both tumor-intrinsic pathways and microenvironmental niche factors as a combination treatment approach. Finally, this study provides a strong foundation for future functional validation of ligand-receptor networks critical for induction of stemness and metastatic behavior in TNBC cells. Additionally, the study provides an opportunity for testing combinatorial therapeutic strategies that pair microenvironmental niche targeting

agents with standard-of-care TNBC treatments to evaluate their efficacy in preventing metastatic outgrowth in preclinical models.

## Declarations

**Ethical Approval and Consent to participate:**

The Western University's Animal Care Committee (protocol #2021-076) approved all animal studies. No additional approvals were required.

**Consent for publication:**

All the authors read and approved the final version of the manuscript for submission.

**Competing interest:**

The authors declare no competing interest.


# Funding:

This work was funded by Verspeeten Family Cancer Centre Catalyst Grant (# LHR F8319), and donor funding from Richard and Susanne Shaftoe and Frank Vita through the London Health Sciences Foundation. V.B., U.P., and A.L.A received support from Breast Cancer Canada.


# Author Contribution

V.B. and A.L.A conceived and planned the experiments. V.B., U.P., D.G. and M.P. performed the experiments. C.P. contributed to histological analysis. V.B., U.P. and M.P contributed to data acquisition and analysis. V.B., D.M., L.G., D.W.L., and A.L.A. contributed to data interpretation. V.B. prepared the first draft of the manuscript. U.P., D.G., C.P., D.M., L.G., D.W.L., and A.L.A. reviewed and edited the manuscript. All the authors read and approved the final manuscript.


# Acknowledgement

We would like to acknowledge Dr. Christopher L. Pin for providing access to the IncuCyte instrument at the Organoid Facility (part of the Baker Centre for Pancreatic Cancer), at the Verspeeten Family Cancer Centre (London, ON), and Ms. Courtney Brooks for her technical assistance with the instrument. We would also like to thank Ms. Cassandra Wong of the Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital (Toronto, ON) for her expertise with data analysis.

# Data Availability

The Spectronaut output data used for all downstream analysis is submitted as Supplementary Table 2.

# References

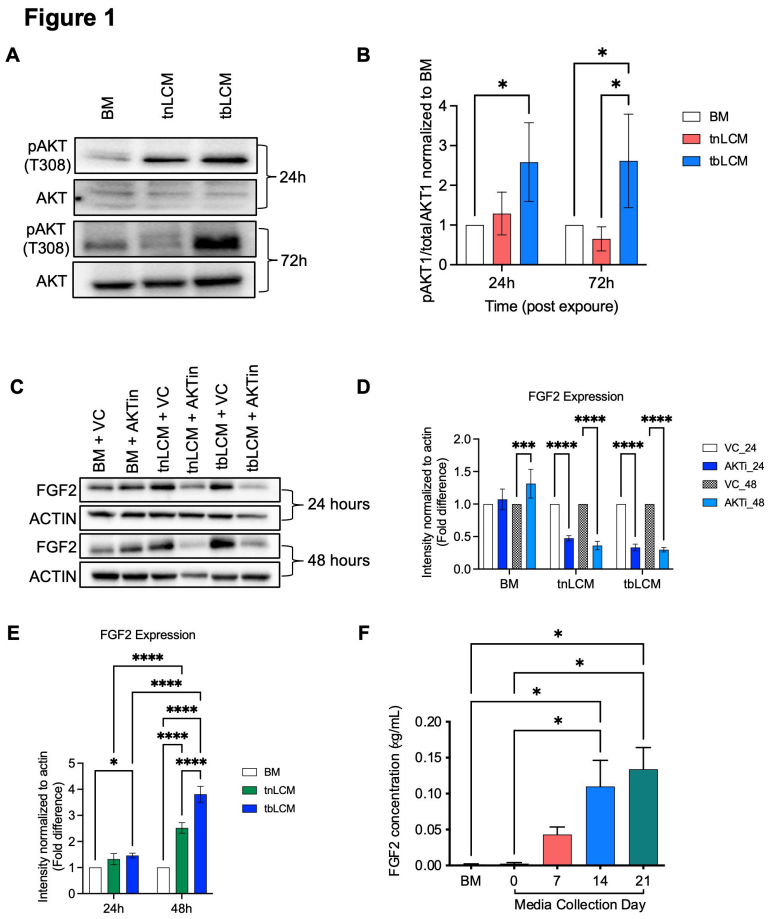

1. Kennecke H, et al. Metastatic behavior of breast cancer subtypes. J Clin Oncol. 2010;28:3271–7. https://doi.org/10.1200/jco.2009.25.9820.
2. Anders C, Carey LA. Understanding and treating triple-negative breast cancer. Oncol (Williston Park). 2008;22:1233–9. discussion 1239–1240.
3. Wu Q, et al. Breast cancer subtypes predict the preferential site of distant metastases: a SEER based study. Oncotarget. 2017;8:27990–6. https://doi.org/10.18632/oncotarget.15856.
4. Loibl S, Poortmans P, Morrow M, Denkert C, Curigliano G. Breast cancer. Lancet. 2021;397:1750–69. https://doi.org/10.1016/s0140-6736(20)32381-3.
5. Ji X, et al. Chemoresistance mechanisms of breast cancer and their countermeasures. Biomed Pharmacother. 2019;114:108800. https://doi.org/10.1016/j.biopha.2019.108800.
6. Croker AK, et al. High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability. J Cell Mol Med. 2009;13:2236–52. https://doi.org/10.1111/j.1582-4934.2008.00455.x.
7. Medeiros B, et al. Triple-Negative Primary Breast Tumors Induce Supportive Premetastatic Changes in the Extracellular Matrix and Soluble Components of the Lung Microenvironment. Cancers (Basel). 2020;12. https://doi.org/10.3390/cancers12010172.
8. Bhat V, et al. Lung-derived soluble factors support stemness/plasticity and metastatic behaviour of breast cancer cells via the FGF2-DACH1 axis. Clin Exp Metastasis. 2024;41:717–31. https://doi.org/10.1007/s10585-024-10284-4.
9. Chen J, Wang Q, Wu H, Huang X, Cao C. Therapies targeting triple-negative breast cancer: a perspective on anti-FGFR. Front Oncol. 2024;14:1415820. https://doi.org/10.3389/fonc.2024.1415820.
10. Hedley BD, et al. Recombinant human erythropoietin in combination with chemotherapy increases breast cancer metastasis in preclinical mouse models. Clin cancer research: official J Am Association Cancer Res. 2011;17:6151–62. https://doi.org/10.1158/1078-0432.Ccr-10-3298.
11. Chu JE, et al. Lung-derived factors mediate breast cancer cell migration through CD44 receptor-ligand interactions in a novel ex vivo system for analysis of organ-specific soluble proteins. Neoplasia (New York N Y). 2014;16:180–91. https://doi.org/10.1593/neo.132076.
12. Sachs N, et al. A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. Cell. 2018;172:373–e386310. https://doi.org/10.1016/j.cell.2017.11.010.
13. Mendoza A, et al. Modeling metastasis biology and therapy in real time in the mouse lung. J Clin Investig. 2010;120:2979–88. https://doi.org/10.1172/jci40252.

14. Khan SU, Xia Y, Goodale D, Schoettle G, Allan AL. Lung-Derived Selectins Enhance Metastatic Behavior of Triple Negative Breast Cancer Cells. Biomedicines. 2021;9. https://doi.org/10.3390/biomedicines9111580.
15. Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. Nat Methods. 2020;17:159–62. https://doi.org/10.1038/s41592-019-0667-5.
16. Wiredja DD, Koyutürk M, Chance MR. The KSEA App: a web-based tool for kinase activity inference from quantitative phosphoproteomics. Bioinformatics. 2017;33:3489–91. https://doi.org/10.1093/bioinformatics/btx415.
17. Giroux P, Maillard M, Colinge J. Post-Transcriptional Modification Integration for Ligand-Receptor Cellular Network Inference. Mol Cell Proteom. 2026;25:101493. https://doi.org/10.1016/j.mcpro.2025.101493.
18. Cabello-Aguilar S, et al. SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. Nucleic Acids Res. 2020;48:e55. https://doi.org/10.1093/nar/gkaa183.
19. Quan MY, et al. An FGFR/AKT/SOX2 Signaling Axis Controls Pancreatic Cancer Stemness. Front Cell Dev Biol. 2020;8:287. https://doi.org/10.3389/fcell.2020.00287.
20. Akl MR, et al. Molecular and clinical significance of fibroblast growth factor 2 (FGF2 /bFGF) in malignancies of solid and hematological cancers for personalized therapies. Oncotarget. 2016;7:44735–62. https://doi.org/10.18632/oncotarget.8203.
21. Mori T, et al. Structural basis for CD44 recognition by ERM proteins. J Biol Chem. 2008;283:29602–12. https://doi.org/10.1074/jbc.M803606200.
22. Dimitrov D, et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. Nat Commun. 2022;13:3224. https://doi.org/10.1038/s41467-022-30755-0.
23. Crosas-Molist E, et al. Rho GTPase signaling in cancer progression and dissemination. Physiol Rev. 2022;102:455–510. https://doi.org/10.1152/physrev.00045.2020.
24. Xiao W, et al. Risk factors and survival outcomes in patients with breast cancer and lung metastasis: a population-based study. Cancer Med. 2018;7:922–30. https://doi.org/10.1002/cam4.1370.
25. Chen S, et al. Prognostic factors and survival outcomes according to tumor subtype in patients with breast cancer lung metastases. PeerJ. 2019;7:e8298. https://doi.org/10.7717/peerj.8298.
26. Luzzi KJ, et al. Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases. Am J Pathol. 1998;153:865–73. https://doi.org/10.1016/s0002-9440(10)65628-3.
27. Rashid OM, Takabe K. The evolution of the role of surgery in the management of breast cancer lung metastasis. J Thorac Dis. 2012;4:420–4. https://doi.org/10.3978/j.issn.2072-1439.2012.07.16.
28. Baumann Z, Auf der Maur P, Bentires-Alj M. Feed-forward loops between metastatic cancer cells and their microenvironment-the stage of escalation. EMBO Mol Med. 2022;14:e14283. https://doi.org/10.15252/emmm.202114283.

29. Izraely S, Witz IP. Site-specific metastasis: A cooperation between cancer cells and the metastatic microenvironment. Int J Cancer. 2021;148:1308–22. https://doi.org/10.1002/ijc.33247.
30. Vasilaki D, Bakopoulou A, Tsouknidas A, Johnstone E, Michalakis K. Biophysical interactions between components of the tumor microenvironment promote metastasis. Biophys Rev. 2021;13:339–57. https://doi.org/10.1007/s12551-021-00811-y.
31. Lee H, et al. Lipid metabolism in cancer stem cells: reprogramming, mechanisms, crosstalk, and therapeutic approaches. Cell Oncol (Dordr). 2025;48:1181–201. https://doi.org/10.1007/s13402-025-01081-6.
32. Xie F, Ling L, van Dam H, Zhou F, Zhang L. TGF-β signaling in cancer metastasis. Acta Biochim Biophys Sin (Shanghai). 2018;50:121–32. https://doi.org/10.1093/abbs/gmx123.
33. Yu L, Mu Y, Sa N, Wang H, Xu W. Tumor necrosis factor α induces epithelial-mesenchymal transition and promotes metastasis via NF-κB signaling pathway-mediated TWIST expression in hypopharyngeal cancer. Oncol Rep. 2014;31:321–7. https://doi.org/10.3892/or.2013.2841.
34. Ababneh O, Nishizaki D, Kato S, Kurzrock R. Tumor necrosis factor superfamily signaling: life and death in cancer. Cancer Metastasis Rev. 2024;43:1137–63. https://doi.org/10.1007/s10555-024-10206-6.
35. Laha D, Grant R, Mishra P, Nilubol N. The Role of Tumor Necrosis Factor in Manipulating the Immunological Response of Tumor Microenvironment. Front Immunol. 2021;12:656908. https://doi.org/10.3389/fimmu.2021.656908.
36. Thies KA, et al. Stromal Platelet-Derived Growth Factor Receptor-β Signaling Promotes Breast Cancer Metastasis in the Brain. Cancer Res. 2021;81:606–18. https://doi.org/10.1158/0008-5472.Can-19-3731.
37. Katsuno Y, Derynck R. Epithelial plasticity, epithelial-mesenchymal transition, and the TGF-β family. Dev Cell. 2021;56:726–46. https://doi.org/10.1016/j.devcel.2021.02.028.
38. Li CW, et al. Epithelial-mesenchymal transition induced by TNF-α requires NF-κB-mediated transcriptional upregulation of Twist1. Cancer Res. 2012;72:1290–300. https://doi.org/10.1158/0008-5472.Can-11-3123.
39. Jechlinger M, et al. Autocrine PDGFR signaling promotes mammary cancer metastasis. J Clin Invest. 2006;116:1561–70. https://doi.org/10.1172/jci24652.
40. Ren B, et al. Resveratrol for cancer therapy: Challenges and future perspectives. Cancer Lett. 2021;515:63–72. https://doi.org/10.1016/j.canlet.2021.05.001.
41. Kim YN, et al. Resveratrol suppresses breast cancer cell invasion by inactivating a RhoA/YAP signaling axis. Exp Mol Med. 2017;49:e296. https://doi.org/10.1038/emm.2016.151.
42. Zhang Z, Liu M, Zheng Y. Role of Rho GTPases in stem cell regulation. Biochem Soc Trans. 2021;49:2941–55. https://doi.org/10.1042/bst20211071.
43. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A. 2003;100:3983–8. https://doi.org/10.1073/pnas.0530291100.

44. Shimabukuro Y, et al. Fibroblast growth factor-2 stimulates directed migration of periodontal ligament cells via PI3K/AKT signaling and CD44/hyaluronan interaction. J Cell Physiol. 2011;226:809–21. https://doi.org/10.1002/jcp.22406.
45. Bourguignon LY. Hyaluronan-mediated CD44 activation of RhoGTPase signaling and cytoskeleton function promotes tumor progression. Semin Cancer Biol. 2008;18:251–9. https://doi.org/10.1016/j.semcancer.2008.03.007.
46. Ouhtit A, Rizeq B, Saleh HA, Rahman MM, Zayed H. Novel CD44-downstream signaling pathways mediating breast tumor invasion. Int J Biol Sci. 2018;14:1782–90. https://doi.org/10.7150/ijbs.23586.
47. Guo Q, Yang C, Gao F. The state of CD44 activation in cancer progression and therapeutic targeting. Febs j. 2022;289:7970–86. https://doi.org/10.1111/febs.16179.

# Figures

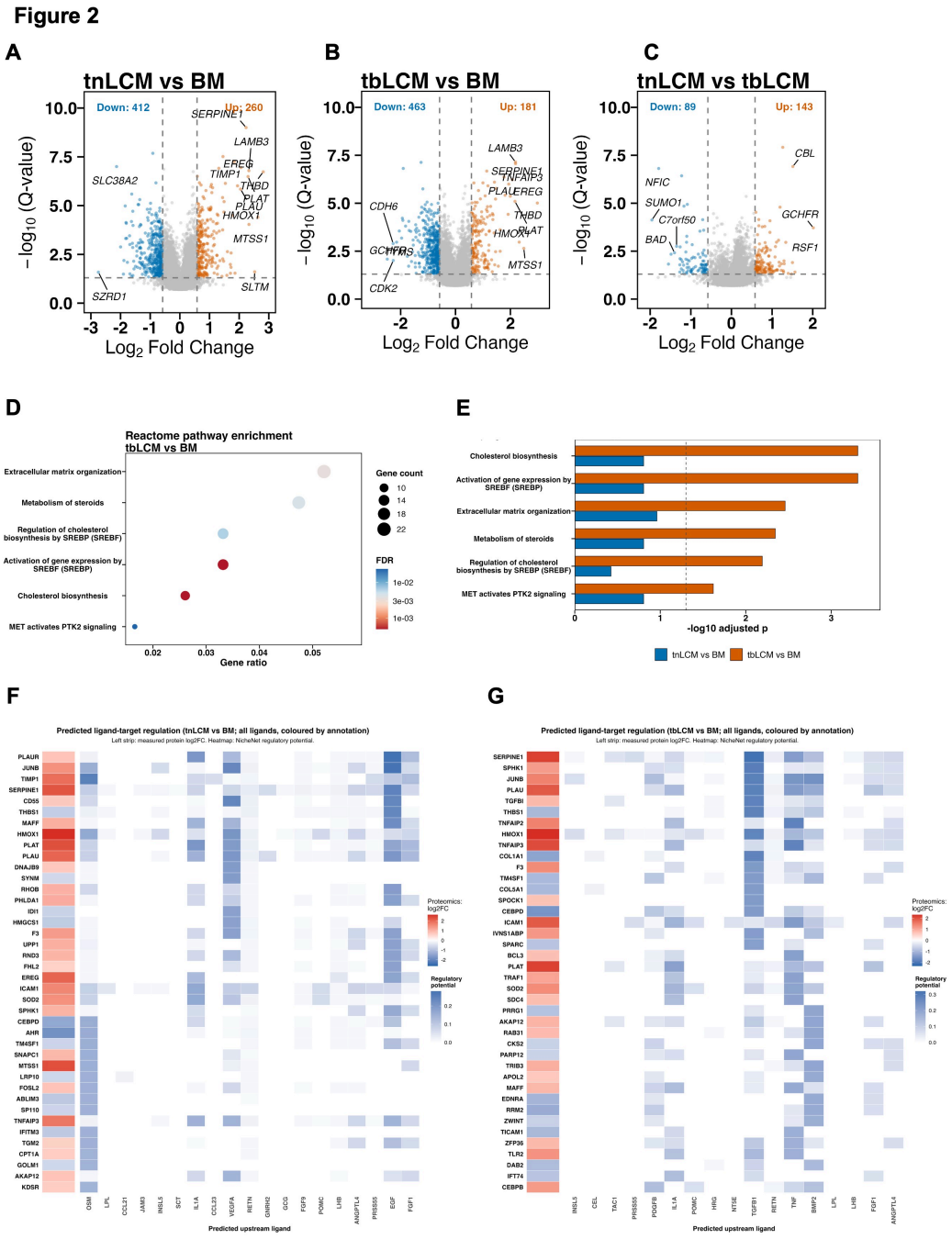



## Figure 1

***Exogenous FGF2 signaling triggers autocrine signaling in TNBC cells.***

*SUM159 TNBC cells were exposed to lung conditioned media from tumor naïve (tnLCM) and SUM159 tumor bearing (tbLCM) mice, with base media (BM) as control. TNBC cells were harvested at different time points, lysed and immunoblotting was performed to assess AKT phosphorylation (Thr308).* ***(A)****Representative*

*immunoblot images showing time-dependent increase in AKT phosphorylation in SUM159 TNBC cells exposed to BM, tnLCM and tbLCM.* ***(B)****Densitometric analysis of the ratio between phosphorylated (Thr308) AKT and total AKT levels at different time points.* ***(C)*** *Representative immunoblots showing FGF2 levels in SUM159 cells exposed to BM, tnLCM and tbLCM in the presence or absence of AKT inhibitor.* ***(D)*** *Densitometric analysis of FGF2 protein levels in SUM159 cells exposed to different lung conditioned media.* ***(E)*** *Densitometric analysis shows significant increase in FGF2 protein levels 48 hours post tbLCM treatment compared to BM and tnLCM. Data are presented as the mean +/- SD (n=3).* ***(F)*** *FGF2 secretion from the lung section cultured ex vivo in the PuMA assay was quantified via ELISA at days 0, 7, 14 and 21. p-values = *≤0.05; **≤0.005; ***≤0.0005; ****≤0.0001.*



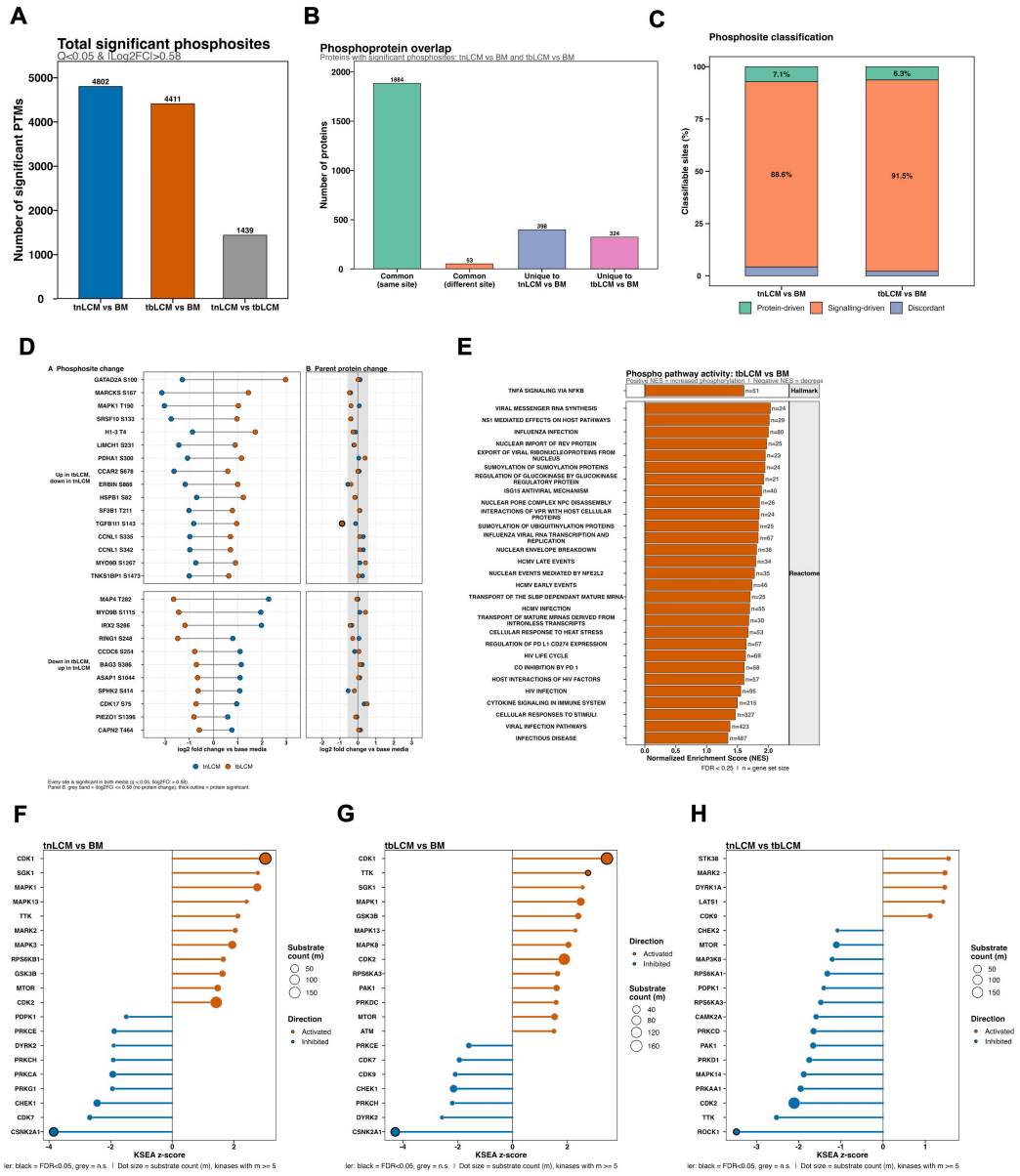

## Figure 2

### *Global proteome dynamics reveal differential expression of proteins in TNBC cells exposed to lung-derived soluble factors*

*SUM159 cells were exposed to BM, tnLCM, and tbLCM for 24 hours and quantitative global proteomics was performed. Global proteomic analysis identified 7499 and 7500 proteins in tnLCM and tbLCM treated cells,*

*respectively. A threshold of q<0.05, |log2FC| >0.58 was employed to identify list of significant proteins. Volcano plots showing differentially expressed proteins in* ***(A)*** *tnLCM-treated cells compared to BM,* ***(B)****tbLCM-treated cells compared to BM and* ***(C)*** *direct comparison between tnLCM and tbLCM-treated cells.* ***(D&E)*** *Reactome pathway analysis showing significantly enriched pathways in TNBC cells exposed to* ***(D)****tnLCM.* ***(E)*** *Bar graph shows tbLCM specific enriched pathways relative to tnLCM. All quantified proteins (n=7,470) in all conditions were used as background in the analysis. Differentially regulated target proteins were used to predict upstream ligand in different lung conditioned media using NicheNet computational method.* ***(F&G)*** *Heatmap shows presence of candidate ligands (x-axis) in different lung conditioned media and their corresponding downstream target proteins differentially expressed in our proteomics dataset (y-axis).*

Figure 3



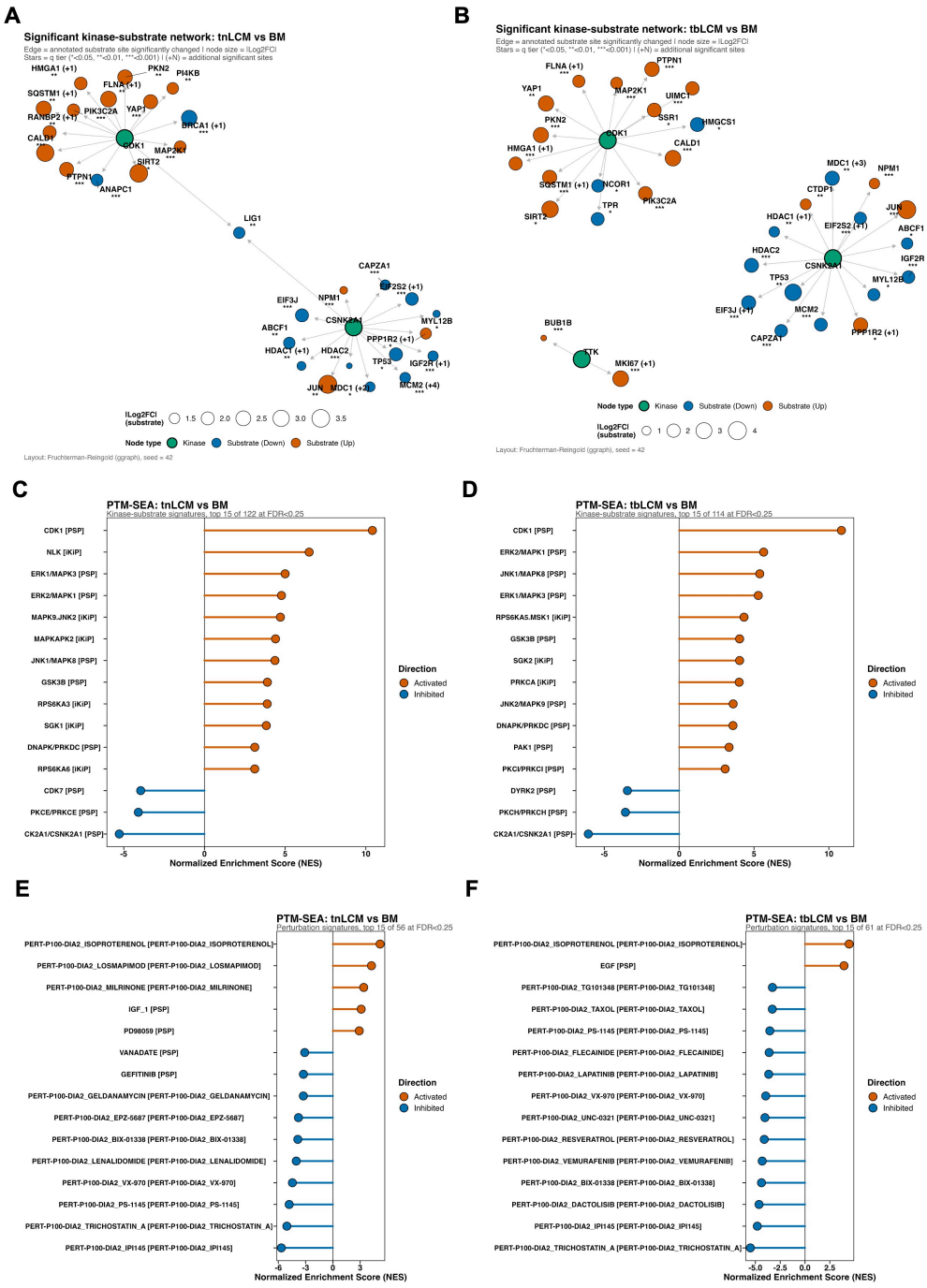

## Figure 3

### *Microenvironment specific phosphoproteome dynamics regulates multiple signaling cascades*

*Phosphoproteomics analysis was performed on SUM159 TNBC cells exposed to BM, tnLCM and tbLCM to identify differentially phosphorylated proteins.* ***(A)*** *Bar graphs show the number of phosphosites per comparison (tnLCMvsBM=5,031; tbLCMvsBM=4,600 and tnLCMvstbLCM=1,459).* ***(B)*** *Bar graphs show the*

*number of common and unique phosphosite under different treatment conditions.* ***(C)*** *Significantly altered phosphosites were classified into three groups in the same comparison based on the log2fold-change value of parent protein. Protein-drive: the parent protein was significantly changed in the opposite direction. Signaling-driven: the parent protein was not significantly changed or was not detected.* ***(C)****In tnCLM vs BM and tbLCM vs BM comparison, 88.6% and 91.5% of phosphosite-level changes were signaling-driven.* ***(D)*** *Left panel, lollipop plot shows phosphosites changes (q<0.05, |log2FC| >0.58) in shared proteins in tnLCM and tbLCM treatment groups relative to BM. Right panel shows log2FC of parent proteins. The grey band corresponds to |log2FC|≤0.58.* ***(E)*** *Fast gene set enrichment analysis (fGSEA) shows significantly enriched signaling pathways in tbLCM-treated cells. Background list used – tnLCMvsBM = 3,723 proteins and tbLCMvsBM = 3,627 proteins.* ***(F-H)*** *Kinase Substrate Enrichment Analysis (KSEA) identified potential upstream kinases based on the substrate phosphorylation status in* ***(D)*** *tnLCMvsBM* ***(E)*** *tbLCMvsBM and* ***(F)****tnLCMvstbLCM comparisons.*

**Figure 4**



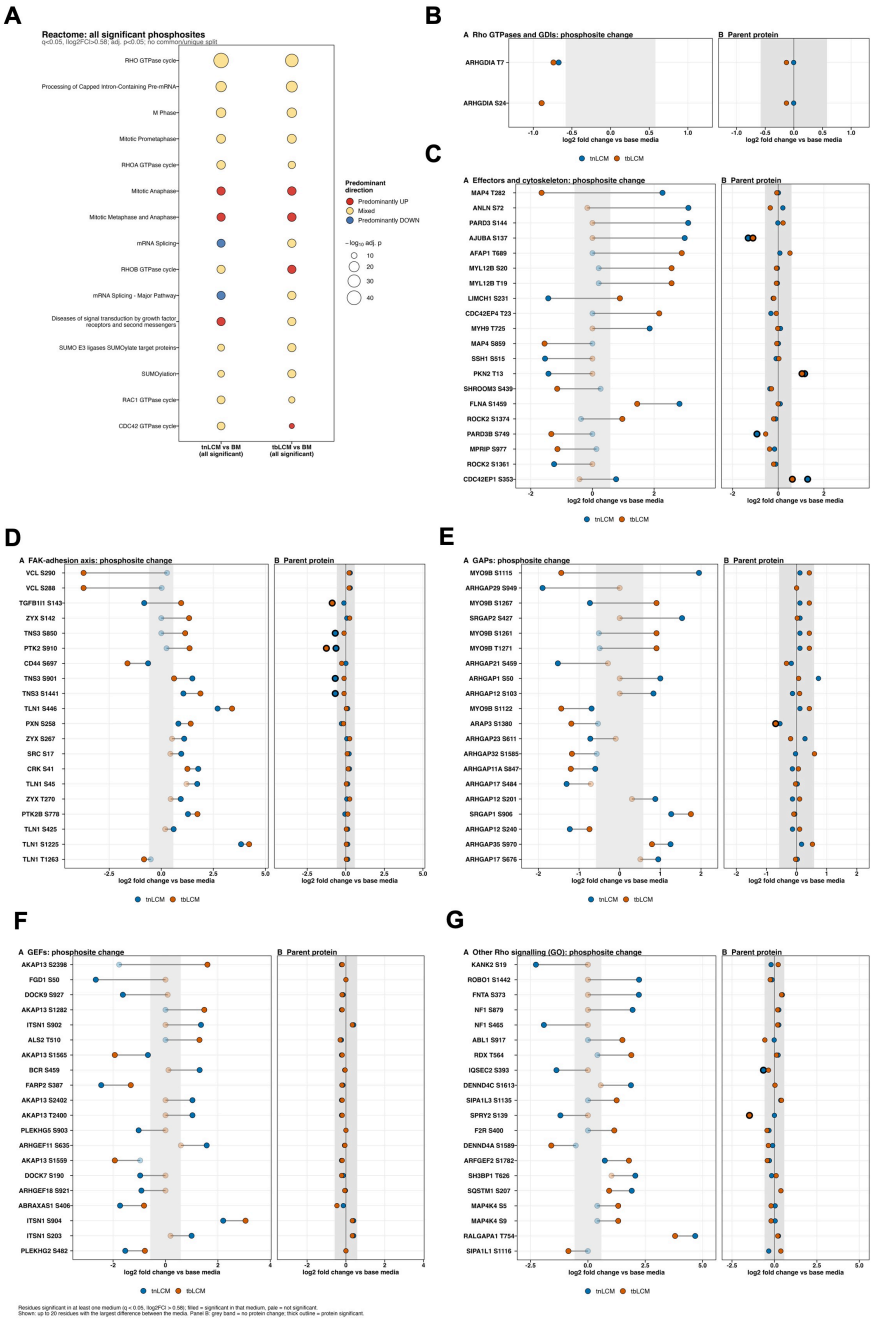

## Figure 4

***Kinase-substrate network reveals microenvironment driven distinct substrate profiles.***

*Using predicted upstream kinases as hub and substrate phosphorylation status in our datasets, kinase-substrate networks was generated.* ***(A)*** *The kinase-substrate network of the tnLCM group shows CDK1 and CSNK2A1 as central hubs with respective substrates.* ***(B)*** *In the tbLCM group, kinase-substrate networks*

*were distinct and included CDK1, TTK, and CSNK2A1.* ***(C-F)*** *PTM-SEA was performed to identify significantly enriched kinases and perturbed phosphorylation signatures. Lollipop plot showing top 15 kinase-substrate signatures ranked by false discovery rate (FDR<0.25) and the normalized enrichment score (NES) in* ***(C)*** *tnLCM vs BM and* ***(D)****tbLCM vs BM.* ***(E&F)*** *Lollipop plot showing the top 15 perturbation signatures ranked by FDR<0.25 and NES in (E) tnLCM vs BM and (F) tbLCM vs BM. Orange: positive NES, Blue: negative NES score.*

**Figure 5**

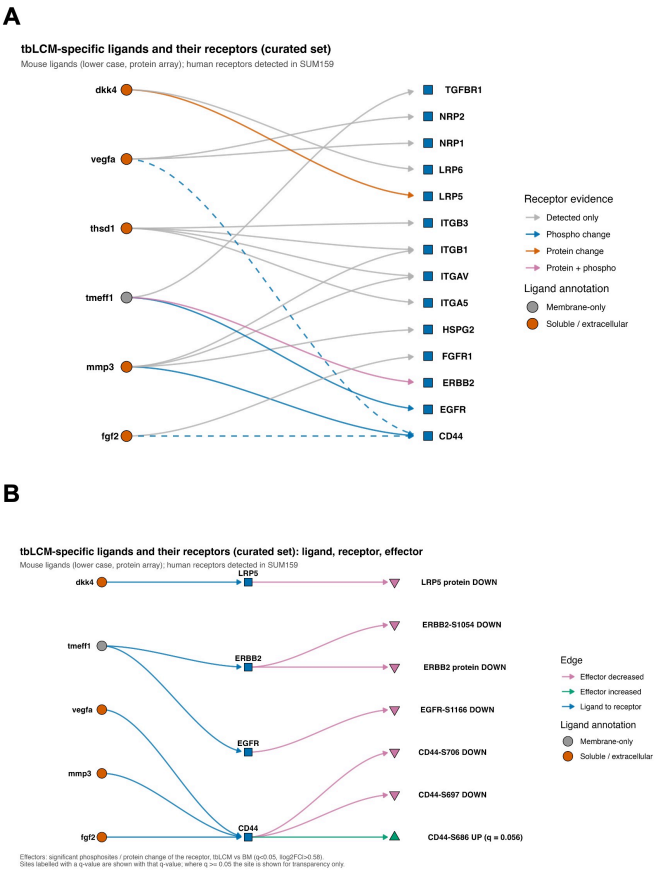


## Figure 5

***Microenvironment-specific signals activated Rho-GTPase signaling***

*Proteins with all significant phosphosites identified in each comparison were separately used to perform pathway analysis.* ***(A)****Reactome pathway enrichment analysis revealed enrichment of Rho GTPase pathways shared between tnLCM vs BM and tbLCM vs BM comparisons. Phosphorylation of Rho GTPAse pathway components were assessed in both tnLCM vs BM and tbLCM vs BM comparison* ***(B-G)****. Left panel, lollipop plots show log2 fold change of significant phosphosites (q<0.05, |log2FC| >0.58) on proteins related to Rho GTPase signaling pathway in different treatment conditions ranked by the difference between the two treatments,* ***(B)*** *Rho GTPase and GDIs,* ***(C)****effectors and cytoskeleton,* ***(D)*** *FAK-adhesion axis,* ***(E)*** *GAPs,* ***(F)****GEFs and* ***(G)*** *other Rho signaling components. Right panel shows log2FC of parent proteins. The grey band corresponds to |log2FC|≤0.58.*

## Figure 6




## Figure 6

***A multi-ligand and receptor interaction induces CD44 signaling in TNBC cells.***

***(A-B)*** *Soluble factors uniquely secreted in the presence of primary TNBC tumor compared to our proteomics* ***(A)*** *and phosphoproteomics data* ***(B)*** *shows multiple ligands converging at CD44 and its*

*corresponding phosphosites. Dashed line in **(A)** corresponds to non-canonical pairs, square = receptor, circle = ligand.*

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupFigS1.jpeg
- SupFigS2.jpeg
- SupFigS3.jpeg
- SupFigS4.jpeg
- SupFigS5.jpeg
- SupFigS6.jpeg
- SupplementaryTable1Reagents.xlsx
- SupplementaryTable2SpectronautAnalysis.xlsx
- SupplementaryTable3NicheNet.xlsx
- SupplementaryTable4RhoGTPasepathwaycomponent.xlsx