

Breast Cancer Cell Secretome Analysis to Decipher miRNA Tumor Biology and Discover Potential Biomarkers

Riley Feser

Brandon University

Reid Opperman

Brandon University

Braydon Nault

Brandon University

Sujit Maiti

Brandon University

Vincent Chen

Brandon University

Mousumi Majumder (✉ majumderm@brandonu.ca)

Brandon University

Article

Keywords: Breast Cancer, MicroRNA, miR526b, miR655, Secretome, Biomarker, Tumor Microenvironment (TME), Estrogen Receptor-Positive (ER+), Luminal A, Bioinformatics, Systems Biology

Posted Date: February 7th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1328838/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Breast Cancer Cell Secretome Analysis to Decipher miRNA Tumor Biology and Discover Potential Biomarkers

Riley Feser¹, Reid M Opperman^{1†}, Braydon Nault^{1†}, Sujit Maiti¹, Vincent C Chen², and Mousumi Majumder^{1*}

¹Department of Biology, Brandon University, 3rd Floor, John R. Brodie Science Centre, 270 18th Street, Brandon, MB, Canada, R7A 6A9

²Department of Chemistry, Brandon University, 4th Floor, John R. Brodie Science Centre, 270 18th Street, Brandon, MB, Canada, R7A 6A9

*Correspondence: majumderm@brandonu.ca, +1(204) 727-7324

†These authors contributed equally to this work

Abstract

MicroRNA (miRNA/miR) miR526b and miR655 overexpressed tumor cell-free secretions promote breast cancer phenotypes in the tumor microenvironment (TME). However, the mechanisms of miRNA regulating TME have never been investigated. With mass spectrometry analysis of MCF7-miRNA-overexpressed versus miRNA-low MCF7-Mock tumor cell secretomes, we identified 34 novel secretory proteins coded by eight genes *YWHAB*, *TXNDC12*, *MYL6B*, *SFN*, *FNI*, *PSMB6*, *PRDX4*, and *PEA15* those are differentially regulated. We used bioinformatic tools and systems biology approaches to identify these markers' role in breast cancer. Gene ontology analysis showed that the top functions are related to apoptosis, oxidative stress, membrane transport, and motility, supporting miRNA-induced phenotypes. These secretory markers expression is high in breast tumors, and a strong positive correlation exists between upregulated markers' mRNA expressions with miRNA cluster expression in luminal A breast tumors. Gene expression of secretome markers is higher in tumor tissues compared to normal samples, and immunohistochemistry data supported gene expression data. Moreover, both up and downregulated marker expressions are associated with breast cancer patient survival. miRNA regulates these marker protein expressions by targeting transcription factors of these genes. Premature miRNA (pri-miR526b and pri-miR655) are established breast cancer blood biomarkers. Here we report novel secretory markers upregulated by miR526b and miR655 (*YWHAB*, *MYL6B*, *PSMB6*, and *PEA15*) are significantly upregulated in breast cancer patients' plasma and are potential breast cancer biomarkers.

Keywords: Breast Cancer, MicroRNA, miR526b, miR655, Secretome, Biomarker, Tumor Microenvironment (TME), Estrogen Receptor-Positive (ER+), Luminal A, Bioinformatics, Systems Biology

Introduction

In 2021, breast cancer surpassed lung cancer as the most diagnosed cancer worldwide, accounting for 11.7% of all cancer cases¹. In addition, the number of breast cancer-related deaths continues to rise globally, as there was a 9.3% increase from 2018 to 2020^{1,2}. A way to cope with the increase in breast cancer cases and deaths is by incorporating accessible, non-invasive, cost-effective methods to regularly screen for breast cancer alongside current procedures, as breast cancer patient survival is 99% when detected in the localized stage³.

The complexity of breast cancer has led to the development of specialized treatment regimens that are dependent on tumor stage, grade, and the presence or absence of hormone receptors (HR; estrogen receptor (ER) and progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2)). The majority of diagnosed breast tumors (73%) are subtype luminal A (HR+/HER2-) followed by 12% that are triple-negative (HR-/HER2-), 11% are luminal B (HR+/HER2+), and 4% are HER2-enriched (HR-/HER2+)⁴. Improvement in patient survival is possible with precision medication and personalized therapy⁵. However, traditional mammogram screening for breast cancer detection starts at age 50 in most countries⁶. With increased breast cancer incidences in younger populations, finding early detection biomarkers is a significant focus of breast cancer research. Blood-based biomarkers can serve as sensitive and specific detection tools to diagnose breast cancer in the early stages.

Cells secrete various molecules such as growth factors, ligands, and metabolites to participate and communicate with neighboring cells. The composition of these secretions changes during oncogenic transformation, establishing a specific tumor microenvironment (TME) that may determine the fate of metastasis⁷. The secretome is the collection of proteins secreted by cells into extracellular space. It makes up a substantial amount (13-20%) of the human proteome and plays essential roles in cell migration, cell signaling, and cell-cell communication⁸. In addition, secretory proteins regulate numerous cancer phenotypes, and bodily fluids near tumors have more secreted proteins, allowing secreted proteins to be detected in blood^{8,9}. Hence, the secretome has diverse regulatory mechanisms in cancer and carcinogenesis.

miRNAs are short non-coding RNAs that regulate post-transcriptional gene expression. In breast cancer, miRNA dysregulation is associated with nearly all aspects of cancer initiation and progression¹⁰. The overexpression of cyclooxygenase-2 (COX-2), an inflammatory enzyme, in poorly metastatic luminal A breast cancer cell line MCF7 upregulates two oncogenic miRNAs, miR526b and miR655^{11,12}. Overexpression of miR526b and miR655 in poorly metastatic luminal A breast tumor cells MCF7 and SKBR3 promoted aggressive breast cancer phenotypes such as epithelial to mesenchymal transition (EMT), cell migration, invasion, induction of cancer stem cells (CSCs), tumor-associated angiogenesis, and lymphangiogenesis, oxidative stress and hypoxic responses¹¹⁻¹⁵. Furthermore, miRNA-overexpressed cell secretions and metabolites induced oxidative stress, tumor-associated angiogenesis, and lymphangiogenesis in poorly metastatic MCF7 cells and enhanced angiogenic properties of primary endothelial

HUVECs^{13–15}. Both miRNA expressions in human breast tumors were significantly higher than disease-free control tissues, and high miR526b and miR655 expression are associated with poor breast cancer patient survival^{11,12}. In patient blood plasma, both pri-miR526b and pri-miR655 can be detected, and pri-miR526b can detect breast cancer as early as stage I and is a sensitive blood-based breast cancer biomarker¹⁶. Investigating these oncogenic miRNAs' effects on secretome composition will identify novel miRNA regulations impacting the TME. Thus, a deeper understanding of oncogenic miR526b- and miR655-high tumor secretions may help better understand miRNA tumor biology and identify novel biomarkers. Here, we performed a systematic analysis of MCF7-Mock, MCF7-miR526b, and MCF7-miR655 secretomes to identify differentially regulated proteins expressed in miRNA-overexpressing tumor secretomes. We aim to discover novel secretory proteins regulating miRNA mechanisms in the TME and potential blood-based breast cancer biomarkers.

Results

Proteins differentially expressed in MCF7-miR526b and MCF7-miR655 cell secretions

We systematically analyzed MCF7-miR526b and MCF7-miR655 cell secretomes and compared them to MCF7-Mock (Figure 1A). Nano-high-performance liquid chromatography-mass spectrometry (nanoHPLC-MS) determined a total of 1535 secreted proteins coded by 442 genes differentially expressed in both MCF7-miRNA-high cell secretomes compared to MCF7-Mock (Supplementary Data 1). A gene ontology (GO) analysis of differential secretions was completed to identify statistically significant GO enrichment functions. The 1535 proteins were involved in 426 biological processes, 145 molecular functions, and 458 pathways, which can be found across 168 different cellular components with 58.1% of the proteins mapped to the extracellular region (Supplementary Data 2). Consistent with secretome analysis, the seven most significant cellular components were within extracellular compartments, including extracellular exosome ($p = 3.34 \times 10^{-111}$), extracellular membrane-bounded organelle ($p = 6.95 \times 10^{-111}$), extracellular organelle ($p = 9.27 \times 10^{-111}$), extracellular vesicle ($p = 1.15 \times 10^{-110}$), vesicle ($p = 1.3 \times 10^{-79}$), extracellular space ($p = 1.05 \times 10^{-77}$), and extracellular region ($p = 8.37 \times 10^{-64}$) (Figure 1B). So, major functions regulated by these markers are extracellular.

The most significant biological process, organelle organization ($p = 1.19 \times 10^{-27}$), allows proper organization of the cytoskeleton, Golgi apparatus, vacuoles, and extracellular vesicles (Figure 1C). Another important biological process, cellular process ($p = 1.79 \times 10^{-25}$), regulates cell export, cell adhesion, cell-cell fusion, and cell communication. These processes are crucial to ensure proper protein secretion and transportation by cells. Eight of the top 10 most significant molecular functions involve the binding of molecules. More specifically, seven functions had binding of molecules that include heterocyclic compounds ($p = 2.41 \times 10^{-50}$) and organic cyclic compounds ($p = 1.48 \times 10^{-49}$), RNA ($p = 3.55 \times 10^{-47}$), nucleic acid ($p = 7.42 \times 10^{-30}$), unfolded protein ($p = 1.19 \times 10^{-24}$), cadherin ($p = 8.38 \times 10^{-23}$) and nucleoside phosphate

($P = 1.24 \times 10^{-21}$) (Figure 1D). This indicates these proteins have major roles in macro-molecule biogenesis. The most significantly enriched pathway is cellular response to stress ($p = 1.78 \times 10^{-53}$) (Figure 1E), which directly relates to the reported roles of miR526b and miR655 to promote oxidative stress in breast cancer¹³. The following most enriched pathway was cellular response to external stimuli ($p = 2.96 \times 10^{-53}$). This result supports our earlier report that miRNA secrete angiogenesis and lymphangiogenesis stimulating factors which alter cell migration and tube formation capacity of endothelial cells¹⁴.

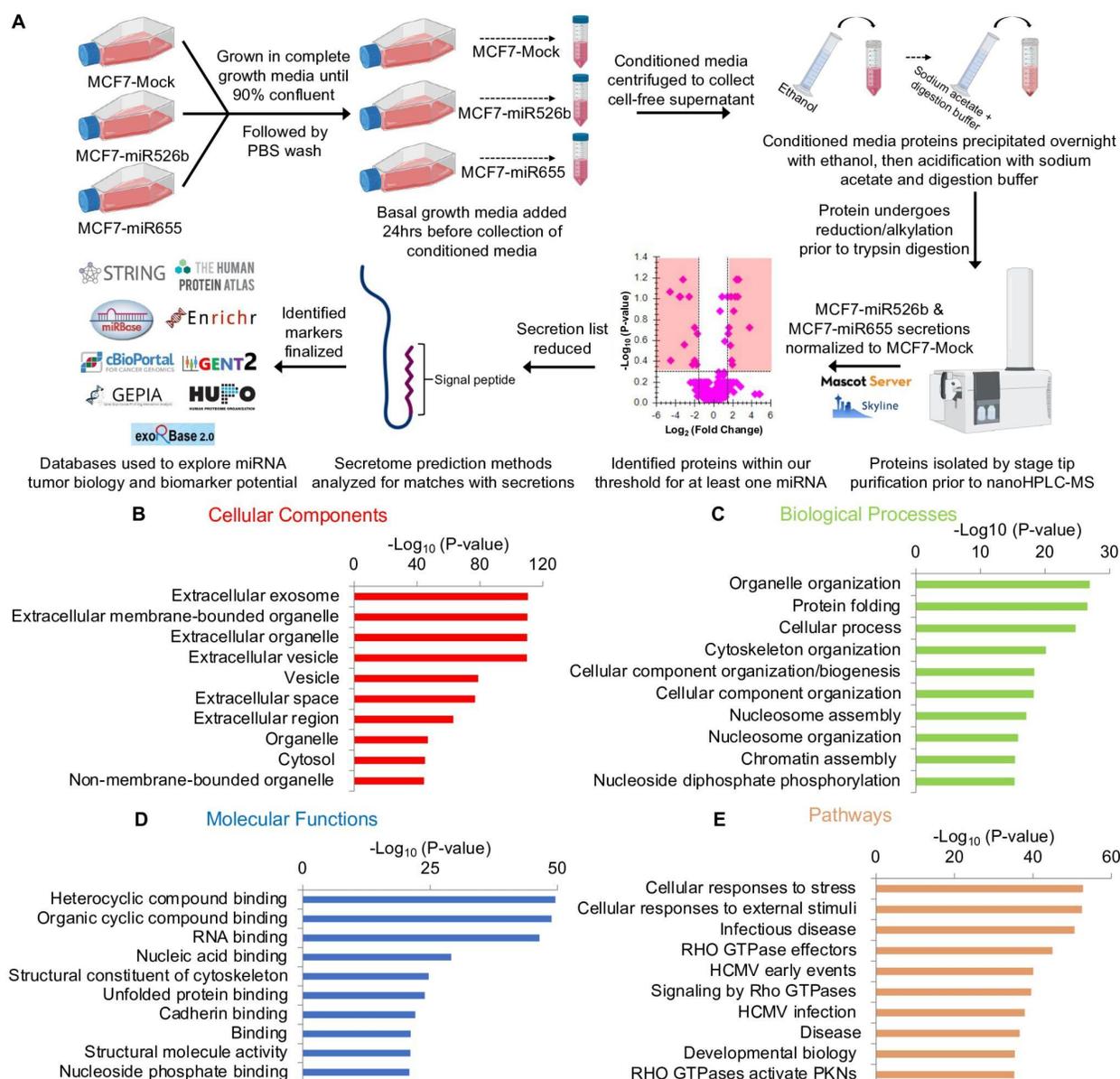


Figure 1. Schematic overview of the project and the top ten most enriched GOs and pathways of differential miRNA-high secretions. (A) Summary of the miRNA-high secretome project. Top ten most significant (B) cellular components, (C) biological processes, (D) molecular functions,

and (E) pathways of differentially expressed secretions in MCF7-miR526b and MCF7-miR655 secretomes.

Data curation and threshold determination

We established a $>1.5 / <-1.5 \log_2$ fold change and $>0.3 -\log_{10}$ p-value as our threshold, following previous publications^{17,18}. We found 96 proteins coded by 32 genes in MCF7-miR526b secretome (Figure 2A) and 95 proteins coded by 29 genes in MCF7-miR655 secretome (Figure 2B) that were within our threshold. The total proteins we identified in both miRNA secretomes are 136 proteins, coded by 39 genes, which is approximately the top 92nd percentile of proteins and genes (Figure 2C). Of the 39 protein-coding genes, 13 were upregulated, and 26 downregulated in both miRNA-secretomes (Figure 2C). Interestingly, many protein-coding genes came from the same protein domains, including 13 from the c-terminus of histone H2A, five from Septin, four from 14-3-3, and three from thioredoxin, a detailed list is provided in Figure S1A. Most protein domains were grouped on the agglomerative hierarchical clustering heatmap, including the c-terminus of histone H2A, septin, and 14-3-3. The total list is provided in Figure S1B. Next, differential secretory proteins within our threshold were analyzed to ensure they were found in the breast-specific proteome. There are 19670 human protein-coding genes, and 14227 are listed in the breast-specific proteome. Of our 39 genes, 33 were found in the breast-specific proteome. The six genes not in the breast-specific proteome (*H2AC1*, *H2AC12*, *H2AC14*, *H2AC18*, *OBSCN*, *SEPT14*) were excluded from further analysis.

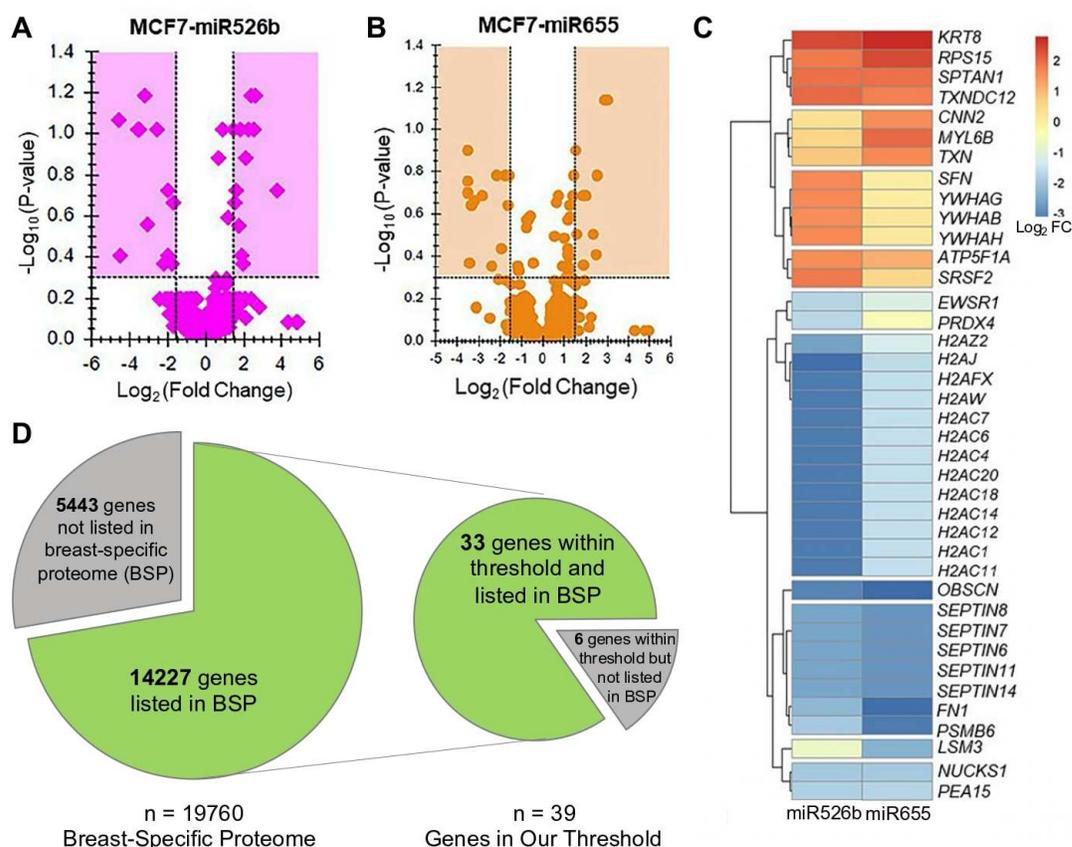


Figure 2. Secretions from MCF7-miR526b and MCF7-miR655 within our threshold and a part of the breast-specific proteome. Volcano plot highlighting proteins within our threshold in (A) MCF7-miR526b and (B) MCF7-miR655. (C) Heatmap showing log₂ fold change values of the total protein-coding genes found from MCF7-miR526b and MCF7-655 volcano plots. (D) Genes within our threshold that are a part of the breast-specific proteome.

Validation of differential secretory proteins using secretome prediction methods

The 33 genes within the breast-specific proteome (BSP) were analyzed by five classical (signal peptide) secretome prediction methods (HPA, Majority Decision-Based Method for Secreted proteins (MDSEC), Phobius, SignalP, SPOCTOPUS), and we also explored one non-classical (no signal peptide) secretome prediction method, SecretomeP. This determined which of the differentially regulated secretory proteins in miRNA-high cells are already known secretory proteins, the total list is provided in Supplementary Data 3. We identified six protein-coding genes that follow the classical secretory pathway, thioredoxin domain containing 12 (*TXNDC12*), myosin light chain 6B (*MYL6B*), stratifin (*SFN*), fibronectin 1 (*FNI*), proteasome 20S subunit beta 6 (*PSMB6*), and peroxiredoxin 4 (*PRDX4*) and two protein-coding genes that follow the non-classical secretory pathway, tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein beta (*YWHAB*) and proliferation and apoptosis adaptor protein 15 (*PEA15*) (Figure 3A). Secretome markers were considered classically secreted if found in

both classical and non-classical secretion methods¹⁹. Of the eight identified secretome markers, four were upregulated (*YWHAB*, *TXNDC12*, *MYL6B*, *SFN*), and four were downregulated (*FN1*, *PSMB6*, *PRDX4*, *PEA15*) by miR526b or miR655 (Figure 3B).

Overall, our data curation pipeline identified eight secretome markers from the beginning 1535 secreted proteins, as summarized in Figure 3C. These eight secretome markers present in both miRNA-secretomes will be further investigated through different systems biology approaches, examining their relationships with miR526b and miR655 and evaluating their blood-based biomarker potential.

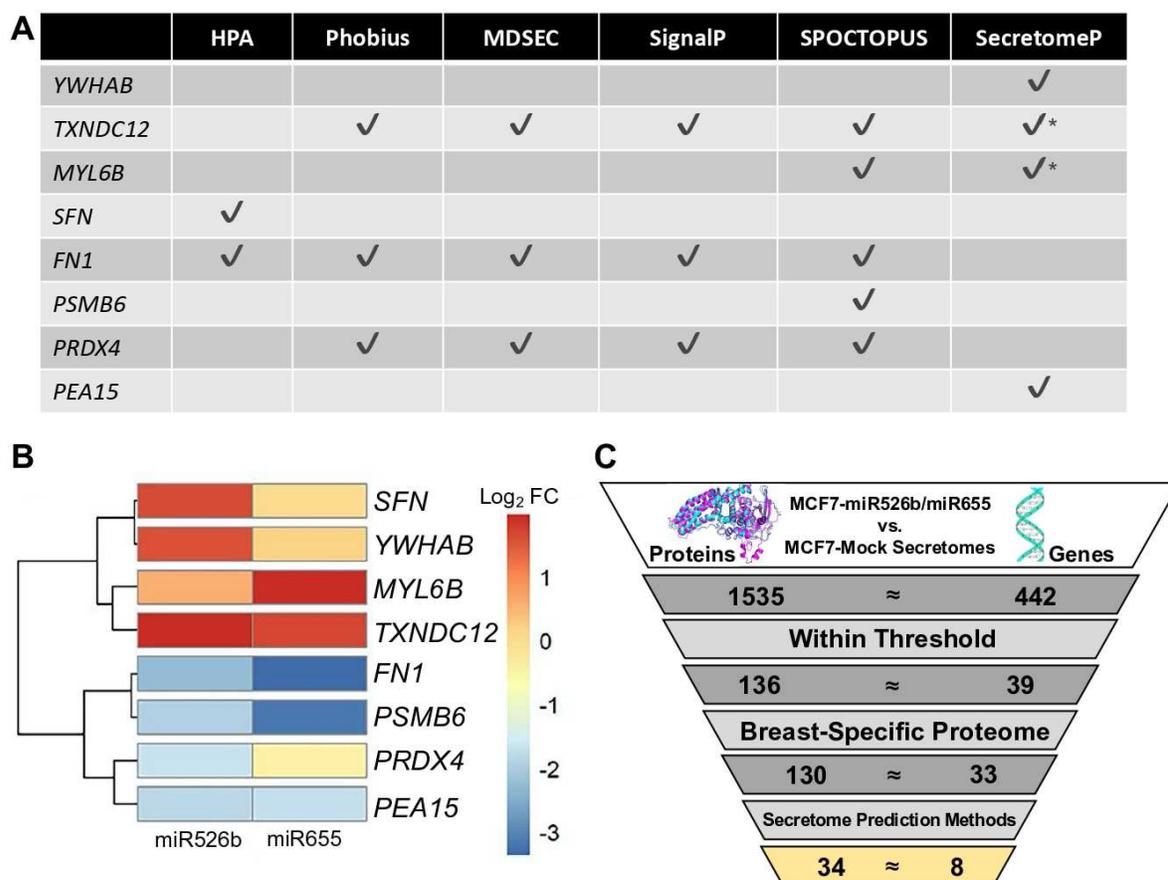


Figure 3. Secretions identified by secretome prediction methods and overview of data curation pipeline. (A) Overview of which protein-coding genes were identified by which secretome prediction method. *= Secretome markers were considered classically secreted if found in both classical and non-classical secretion methods. (B) Heatmap showing the log₂ fold change of the eight selected markers. (C) Each step of the data curation process highlights how the final list of secretome markers was achieved.

Gene ontology analysis of secretome markers

First, we determined the secretome markers' general functions and individual cellular components, biological processes, and molecular functions with GO analysis (Figure 4A and

Table S1). Interestingly, many of the secretome markers' functions and GO enrichments overlapped with miR526b and miR655 induced breast cancer phenotypes. For example, *PRDX4* and *TXNDC12* have contrasting roles in oxidative stress. *TXNDC12* was found upregulated in both miRNA secretomes. It is a negative regulator of the endoplasmic reticulum stress-induced intrinsic apoptotic signaling pathways, allowing cells to survive while stressed. *PRDX4* is an antioxidant enzyme that neutralizes oxygen species and protects cells against oxidative stress. It is a key protein that regulates cell redox homeostasis. Since miRNA overexpression has been shown to increase oxidative stress in breast cancer, decreased expression of *PRDX4* and increased expression of *TXNDC12* in miRNA-secretomes shows the complexity of the breast cancer TME and the tumor regulatory roles of miRNA. *FNI* modulates angiogenesis and is downregulated in miRNA secretomes. Since miRNAs promote angiogenesis, *FNI* expression in miRNA-high cells might be regulated by miRNA targeting transcription factors of *FNI*.

Next, the eight secretome markers were analyzed further to find any significantly enriched function. This found four cellular components (Figure 4B), four biological processes (Figure 4C), one molecular function, which is phosphoprotein binding (Figure 4D), and 30 different pathways (Figure S2) linked to these markers. The top ten pathways are shown in Figure 4E. Interestingly, all cellular components were related to the extracellular region, further confirming that these secretome markers are secreted by miRNA-overexpressed breast tumor cells, suggesting they might play a role in TME regulation. Additionally, three biological processes were related to apoptotic regulation, and one was involved in redox reaction, and both processes are altered by miR526b and miR655 in breast cancer. The most significantly enriched pathway was FLT3 signaling. This pathway is involved in the differentiation, proliferation, and survival of dendritic and hematopoietic progenitor cells²⁰. When adapter and signaling molecules bind with FLT3's active receptor, activation of downstream pathways such as PI3K/Akt and MAPK cascades occur. MAPK signaling is also a significant pathway listed. This pathway responds to various extracellular stimuli to activate intracellular processes such as gene expression, metabolism, proliferation, and apoptosis²¹. The RAF/MAP kinase cascade is highly mutated in cancer, RAS mutations are found in ~30% of all human cancers, and the most active activator of this pathway, BRAF, is mutated in ~7% of cancers, including breast cancer²². FOXO transcription factors (FOXO1, FOXO3, FOXO4) bind to 14-3-3 proteins, allowing their retention in the cytosol²¹, and 14-3-3 proteins are upregulated in the secretome. Activation of BAD and translocation to mitochondria is sequestered by 14-3-3 proteins after Akt1 phosphorylation²¹. Most of these pathways are known regulatory functions of miR526b and miR655 involvement.

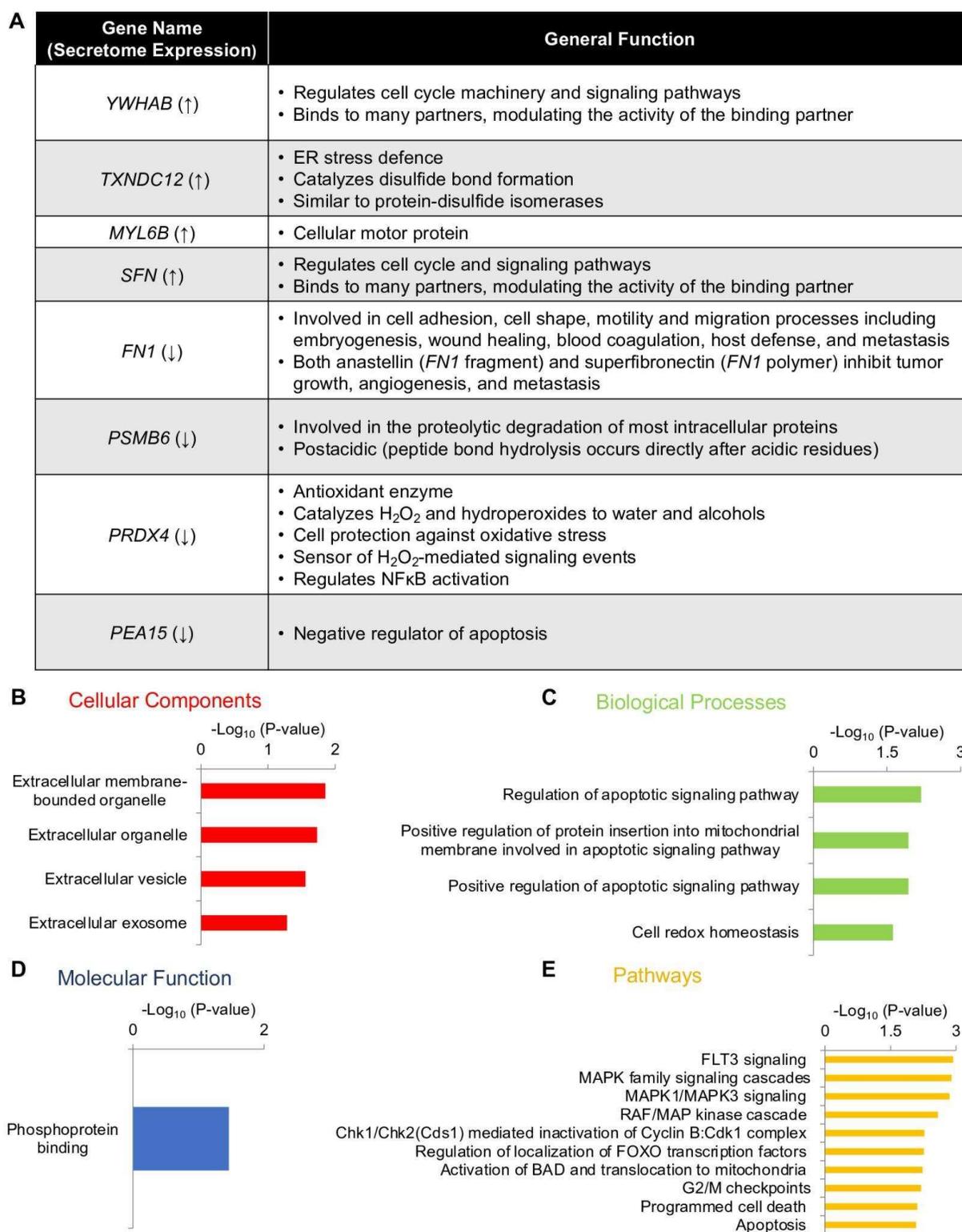
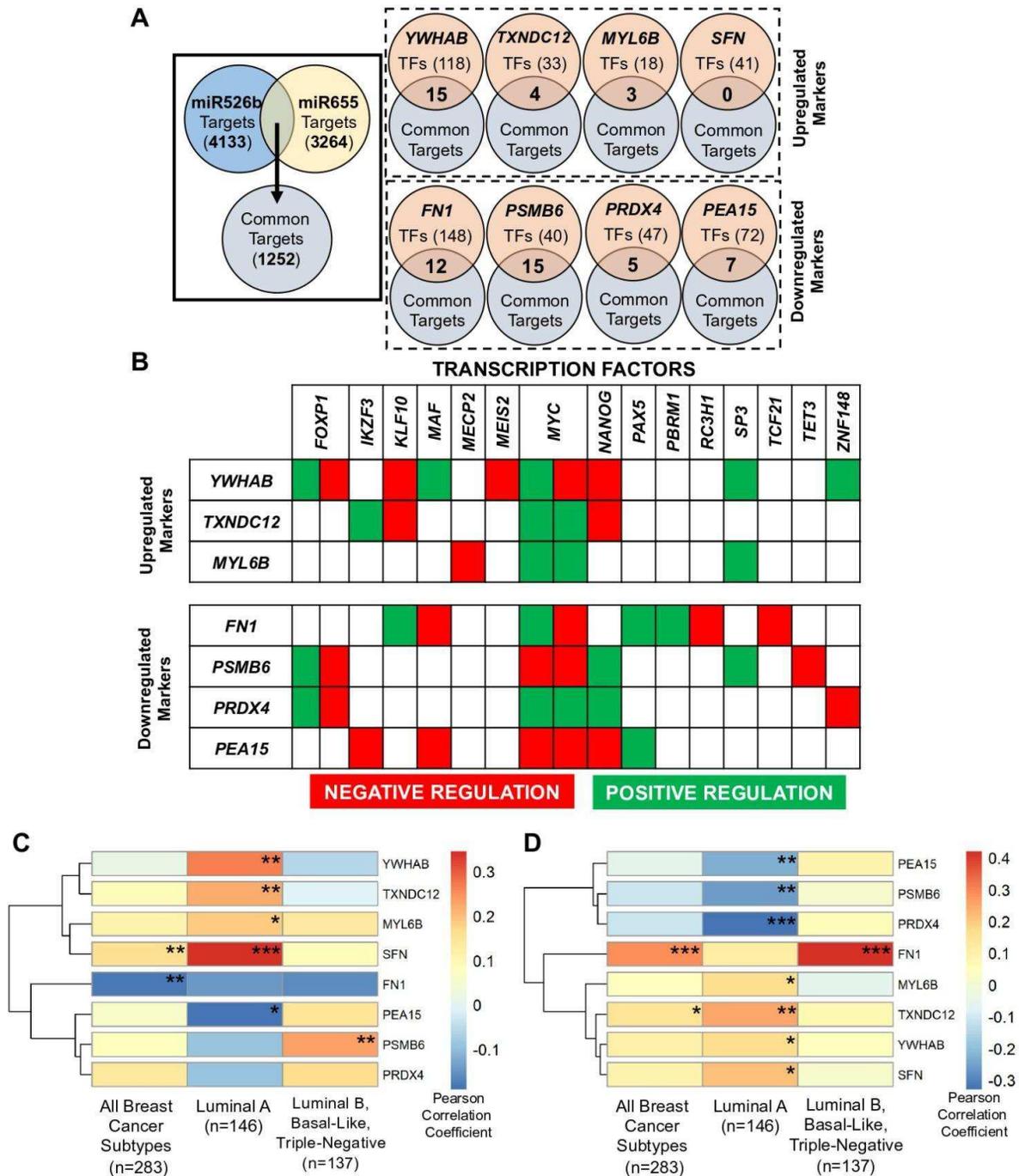


Figure 4. Overview of our identified secretome markers. (A) Secretome markers general

functions. Shared significantly enriched (B) cellular components, (C) biological processes, (D) molecular component GOs, and (E) pathways of secretome markers.

In silico analysis of miRNA regulating secretome markers gene expressions

To establish a regulatory connection between miRNA and secretome markers, we cross-referenced known transcription factors (TFs) of the eight markers with predicted targets of both miRNAs. Altogether, miR526b showed 4133 predicted targets, and miR655 had 3264 predicted targets, of which 1252 were predicted common targets of both miRNAs (Figure 5A). *TXNDC12* was an indirect target of miR526b and a predicted direct target of miR655. However, *TXNDC12* was found upregulated in both miRNA secretomes. Therefore, we found TFs that negatively regulate *TXNDC12* and are common targets of both miRNAs (Figure 5B). miRNA downregulates negative regulator TFs *NANOG* and *KLF10*, which upregulates *TXNDC12* expression. Upregulated secretome marker *MYL6B* has two positive regulation TFs, *MYC* and *SP3*, and one negative regulation TF *MECP2*. Since *MYL6B* is upregulated in both miRNAs' secretomes, we predict that *MECP2* is completely downregulated by miRNAs, leading to *MYL6B*'s increased secretome expression. Alternatively, downregulated secretome marker *PRDX4* has three positive regulation TFs, *FOXP1*, *MYC*, and *NANOG*, and two negative regulation TFs, *FOXP1* and *ZNF148*. Since *PRDX4* is downregulated in miRNA-high secretomes, we anticipate that the three positive regulation TFs are targeted by miRNAs resulting in *PRDX4* downregulation. We could not find TFs of *SFN* that are common miRNA targets. However, seven TFs have predicted miR526b targets (Figure S3A), and six have predicted targets of miR655 (Figure S3B). Therefore, in miR526b's secretome, *SFN* could be upregulated through one or more of its negative regulation TFs *ETS2*, *SOX4*, *THRB*, *ESR1*, and *POU5F1* downregulation. In miR655's secretome, *SFN* could be upregulated via targeting negative regulator TFs *THRA*, *SRF*, and *HNF4A*. Interestingly, two negative regulators of *SFN*, *THRB*, and *THRA* (thyroid hormone receptors) are targets of miR526b and miR655, respectively.



heatmap for secretome markers in breast cancer subtypes for (C) miR526b and (D) miR655. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Secretome markers correlation with miRNAs cluster expressions in breast cancer

To determine a correlation between miR526b and miR655 with secretome markers in breast tumors, miRNA cluster expressions and secretome markers mRNA expression data were extracted, and the Pearson correlation coefficient was measured. For all breast cancer subtypes, there were a total of 283 tumor tissue samples available for both miRNA clusters and secretome marker mRNA expression (Figure 5C-D). In non-stratified samples, upregulated secretome markers *SFN* and *TXNDC12* showed significant positive correlations with miR526b and miR655 clusters, respectively. At the same time, downregulated secretome marker *FNI* had a significant negative correlation to miR526b's cluster. Tumor samples were then stratified into luminal A or luminal B, HER2-enriched, and triple-negative tumor subtypes.

In luminal A samples, miRNA cluster expressions showed a stronger correlation to secretome markers in luminal A breast cancer for all markers in miR655's cluster, and seven correlations were significant (Figure 5C). This is denoted by the increased Pearson correlation coefficient in total tumor samples compared to luminal A samples for upregulated markers and decreased Pearson correlation coefficient between the two categories for downregulated markers. In miR526b's cluster, all markers except *FNI* had improved correlations in luminal A breast cancer, and five markers correlation coefficients were significant (Figure 5D). When comparing total samples to luminal B, HER2-enriched, and triple-negative subtypes, only one enhanced correlation was found, as *FNI* had a stronger negative correlation with miR526b's cluster compared to all samples, but this was not significant.

Secretome markers gene expression in breast cancer cell lines

Gene expression of secretome markers in poorly metastatic ER-positive MCF7, T47D, HER2 positive SKBR3, and highly metastatic triple-negative HS578T and MDA-MB-231 breast cancer cell lines was compared to immortalized mammary epithelial cell line MCF710A. Beginning with upregulated secretome markers, *YWHAB* showed significant upregulation in both poorly metastatic T47D, SKBR3 cell lines, and highly aggressive MDA-MB-231 cells (Figure 6A); *TXNDC12* was significantly upregulated in poorly metastatic T47D and SKBR3 cells (Figure 6B); *MYL6B* showed significant upregulation in T47D cells (Figure 6B), and *SFN* was significantly downregulated in all breast cancer cell lines (Figure 6D). So, all upregulated markers except *SFN* validated secretome results.

Among downregulated secretome markers, *FNI* was significantly downregulated in MCF7 and T47D cells and was significantly upregulated in HS578T cells (Figure 6E); *PSMB6* was high in all cell lines but was significantly upregulated in MDA-MB-231 cells (Figure 6F); *PRDX4* showed significant upregulation only in aggressive cell lines HS578T and MDA-MB-231 (Figure 6G); and *PEA15* is significantly downregulated only in MCF7 cells, while was

significantly upregulated in all other cell lines (Figure 6H). Secretome is protein data, and these are gene expression data. However, we found *YWHAB*, *TXNDC12*, *MYL6B*, *FN1*, and *PEA15* marker expression is similar in secretome and cell lines. miRNAs play a major role in gene expression and regulation; hence these observations need further validation in miRNA upregulated cells.

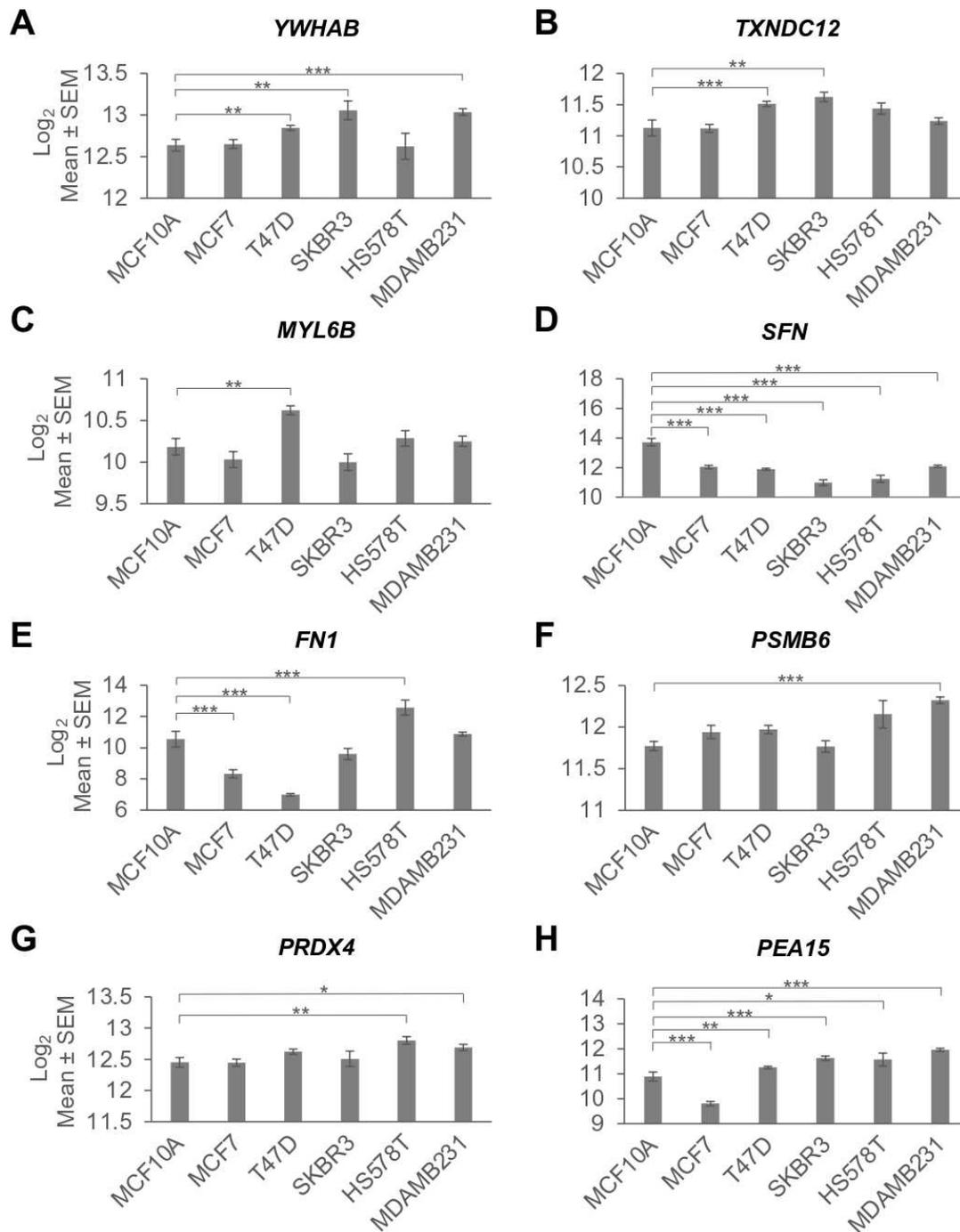


Figure 6. Secretome markers gene expression in poorly metastatic and highly metastatic breast cancer cell lines. Gene expression in breast cancer cell lines for secretome marker (A) *YWHAB*, (B) *TXNDC12*, (C) *MYL6B*, (D) *SFN*, (E) *FNI*, (F) *PSMB6*, (G) *PRDX4*, and (H) *PEA15*.

Translational validation of secretome markers in breast cancer: Gene expression analysis in breast cancer tissue and blood

Next, the eight secretome marker expressions were analyzed in breast tumor tissue (n=1085) and normal tissues (n=291). All markers were upregulated in breast cancer tumor tissues compared to control (Figure 7A). This supports upregulated secretome markers *YWHAB*, *TXNDC12*, *MYL6B*, and *SFN*. However, downregulated secretome markers *FNI*, *PSMB6*, *PRDX4*, and *PEA15* were also upregulated at the mRNA level in breast tumors. The observed downregulation of secretome markers in miRNA-high secretomes could be due to miRNA targeting the positive regulator TFs of the identified markers as shown in miRNA target analysis.

Using the Human Proteome Organizations (HUPO) Human Proteome and Plasma Proteome Projects data, we identified that all secretome markers could be found in human blood at the highest level of protein evidence (protein level) and the highest level of certainty (canonical) (Table S2). Additionally, secretome marker mRNA expression was explored in the blood exosomes of breast cancer patients (n=140) and healthy controls (n=118) (Figure 7B-I). In breast cancer patients' blood, *YWHAB*, *MYL6B*, *PSMB6*, and *PEA15* expressions were significantly higher, and *SFN* and *PRDX4* expressions were also higher but not significant compared to control blood samples. On the other hand, *TXNDC12* and *FNI* showed no difference in mRNA expression in blood plasma.

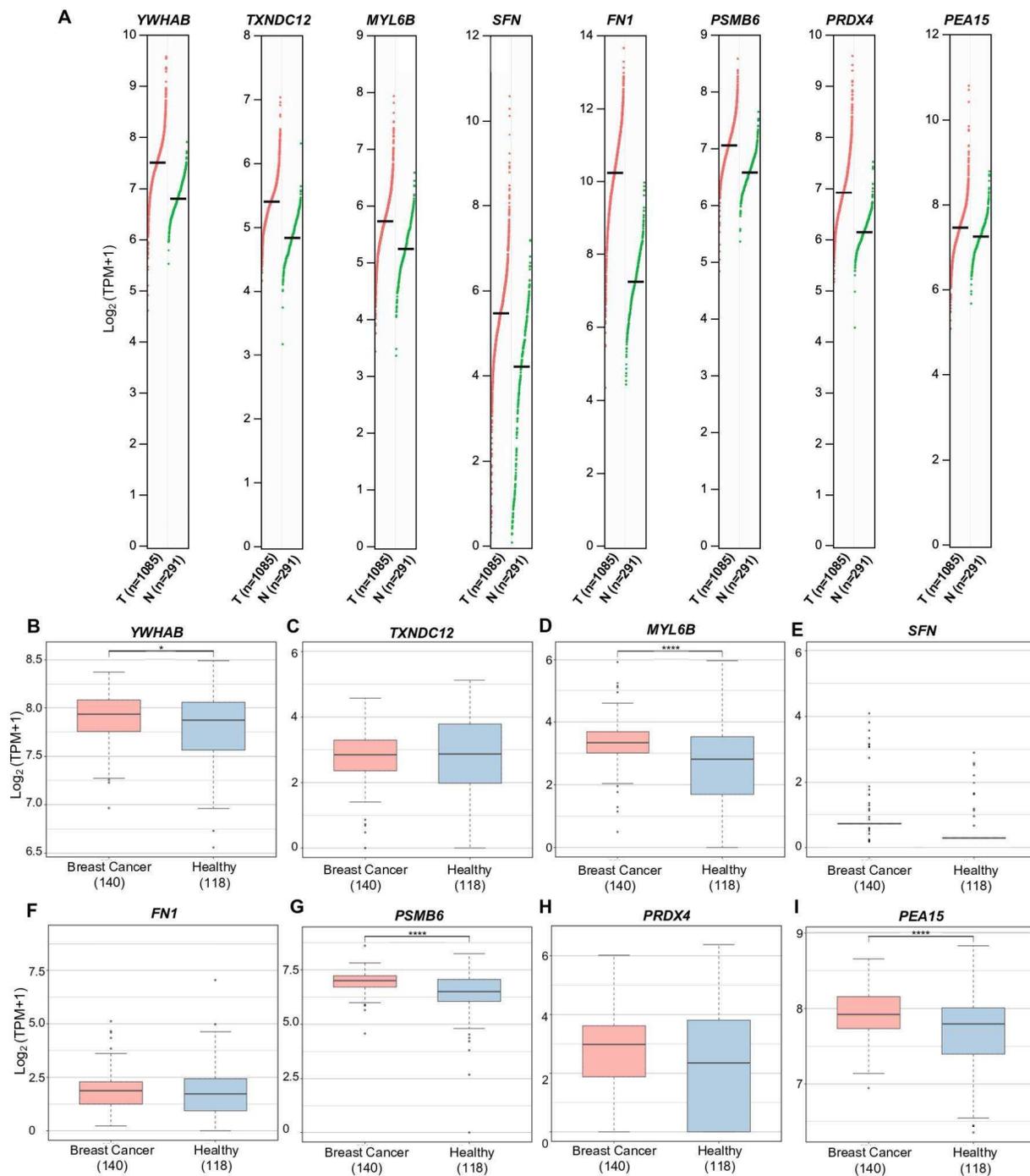


Figure 7. Secretome markers gene expression breast cancer tissue and blood exosomes. (A) Secretome marker gene expression in breast cancer tumor tissue. (T): Breast cancer tumor samples. (N): Normal breast samples. The black line represents the median expression value, and each dot is a patient sample. Secretome marker mRNA expression in breast cancer patient blood exosomes for (B) *YWHAB*, (C) *TXNDC12*, (D) *MYL6B*, (E) *SFN*, (F) *FN1*, (G) *PSMB6*, (H) *PRDX4*, and (I) *PEA15*. * $p < 0.05$, **** $p < 0.0001$.

Immunohistochemistry analysis of identified markers in breast tumor and normal tissue

Immunohistochemistry staining data of breast cancer tumors and normal tissues were examined to determine secretome markers protein expression and localization. SFN had five of 12 breast tumor tissues showing high or low positive intensity staining, compared to normal tissue not showing any expression (Figure 8A). YWHAB had all breast tumor samples stained positive at high or medium levels compared to YWHAB found only at medium expression in normal tissue (Figure 8B). MYL6B showed the same expression in both tumor and control tissue (Figure 8C). TXNDC12 had no immunohistochemistry data. PRDX4 had eight of 11 samples showing low staining in tumor tissue, compared to normal samples, all having medium intensity (Figure 8D). FN1 had most samples (eight of 12 samples) showing no expression in breast tumor tissue, and normal tissues showed low-intensity staining (Figure 8E). No breast cancer tissues showed PEA15 expression, but all three control samples showed low PEA15 expression (Figure 8F). PSMB6 in both breast and normal tissues showed medium or high-intensity staining (Figure 8G). Overall, SFN, YWHAB, PRDX4, FN1, and PEA15 immunohistochemistry data reflected secretome expression.

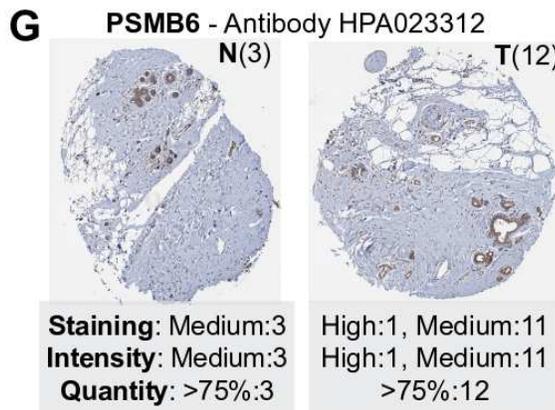
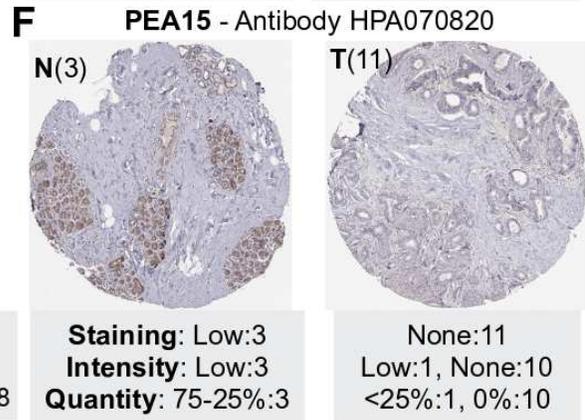
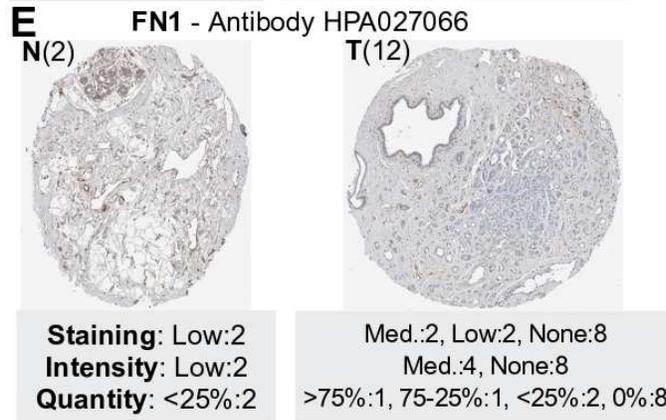
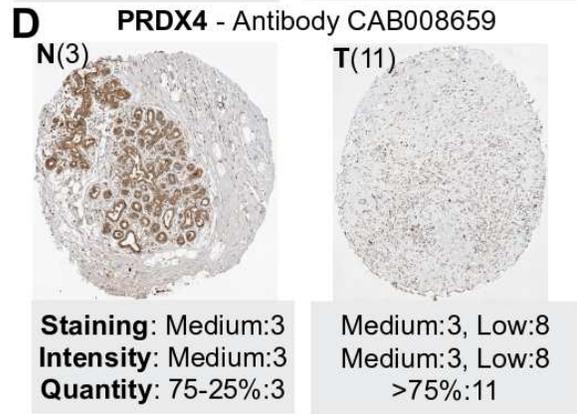
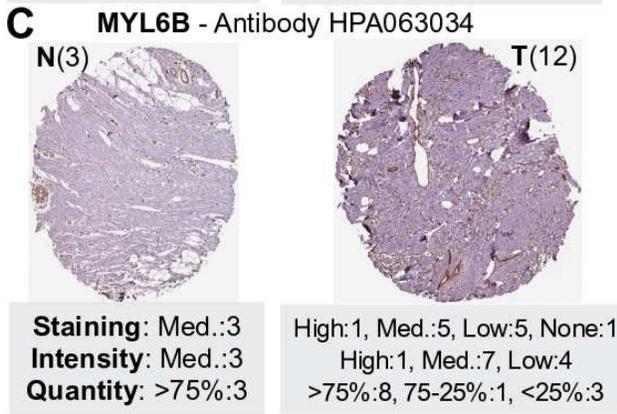
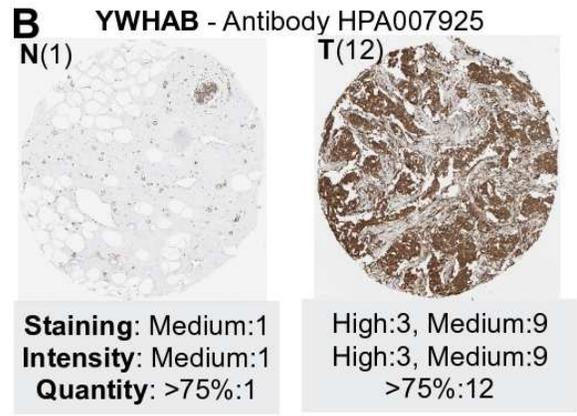
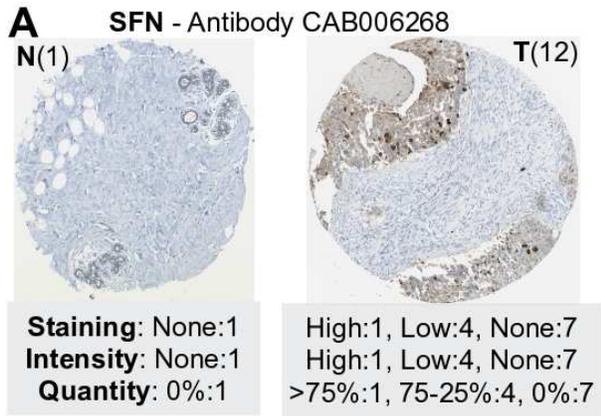


Figure 8. Secretome markers immunohistochemical staining. Immunohistochemistry staining data for secretome marker (A) SFN, (B) YWHAB, (C) MYL6B, (D) PRDX4, (E) FN1, (F) PEA15, and (G) PSMB6. TXNDC12 has no immunohistochemistry data. The brown staining indicates where an antibody has been bound to its corresponding antigen, and blue represents counterstaining.

Roles of secretome markers in breast cancer patient survival

We have conducted Kaplan-Meier survival plot analysis for secretome markers in non-stratified (all stages I-IV) and in stratified early (stages I & II) and late (stages III & IV) tumor stages to examine if secretome markers expression is associated with poor patient survival (Figure 9). High *YWHAB* expression showed reduced survival in non-stratified samples, but that was not significant ($p = 0.092$) (Figure 9A). However, *YWHAB* significantly reduced survival in early-stage tumors ($p = 0.020$) (Figure 9B). But in late-stage tumors, *YWHAB* expression was not associated with patient survival (Figure 9C). *TXNDC12* expression was not correlated with breast cancer patient survival in non-stratified samples ($p = 0.95$) (Figure 9D), and in the early stages, low *TXNDC12* expression showed slightly reduced survival, but that was not significant ($p = 0.23$) (Figure 9E). But in late-stage tumors, high *TXNDC12* expression is significantly associated with poor patient survival ($p = 0.00087$) (Figure 9F). Low *MYL6B* expression showed reduced survival in non-stratified samples but was not significant ($p = 0.09$) (Figure 9G); however, low *MYL6B* expression significantly reduced survival in early stages ($p = 0.044$) (Figure 9H). Alternatively, high *MYL6B* expression led to slightly decreased survival in late stages, but that was not significant ($p = 0.58$) (Figure 9I). Decreased *SFN* expression led to poor patient survival in non-stratified samples ($p = 0.056$) (Figure 9J) and significantly reduced survival in early-stages ($p = 0.018$) (Figure 9K). On the other hand, in late stages, high *SFN* expression led to significantly reduced survival ($p = 0.0039$) (Figure 9L). High *FN1* expression led to slightly decreased survival in non-stratified samples ($p = 0.21$) (Figure 9M). In the early tumor stages, *FN1* was not associated with breast cancer patient survival (Figure 9N). Increased *FN1* expression showed marginally reduced survival in late stages but not significantly (Figure 9O). *PSMB6* expression was not connected to patient survival in non-stratified samples ($p = 0.94$) (Figure 9P). However, low *PSMB6* expression led to slightly decreased survival in early stages ($p = 0.16$) (Figure 9Q), and high *PSMB6* expression showed moderately reduced survival in late stages ($p = 0.21$) (Figure 9R). High *PRDX4* expression led to significantly decreased survival in non-stratified samples ($p = 0.018$) (Figure 9S). Low *PRDX4* expression led to reduced survival in early-stage tumors but was not significant ($p = 0.72$) (Figure 9T). In late stages, high *PRDX4* expression showed significantly decreased survival ($p = 0.00049$) (Figure 9U). *PEA15* had no survival data available. Overall, the most important relationships in breast cancer patient survival were found with upregulated secretome marker expressions.

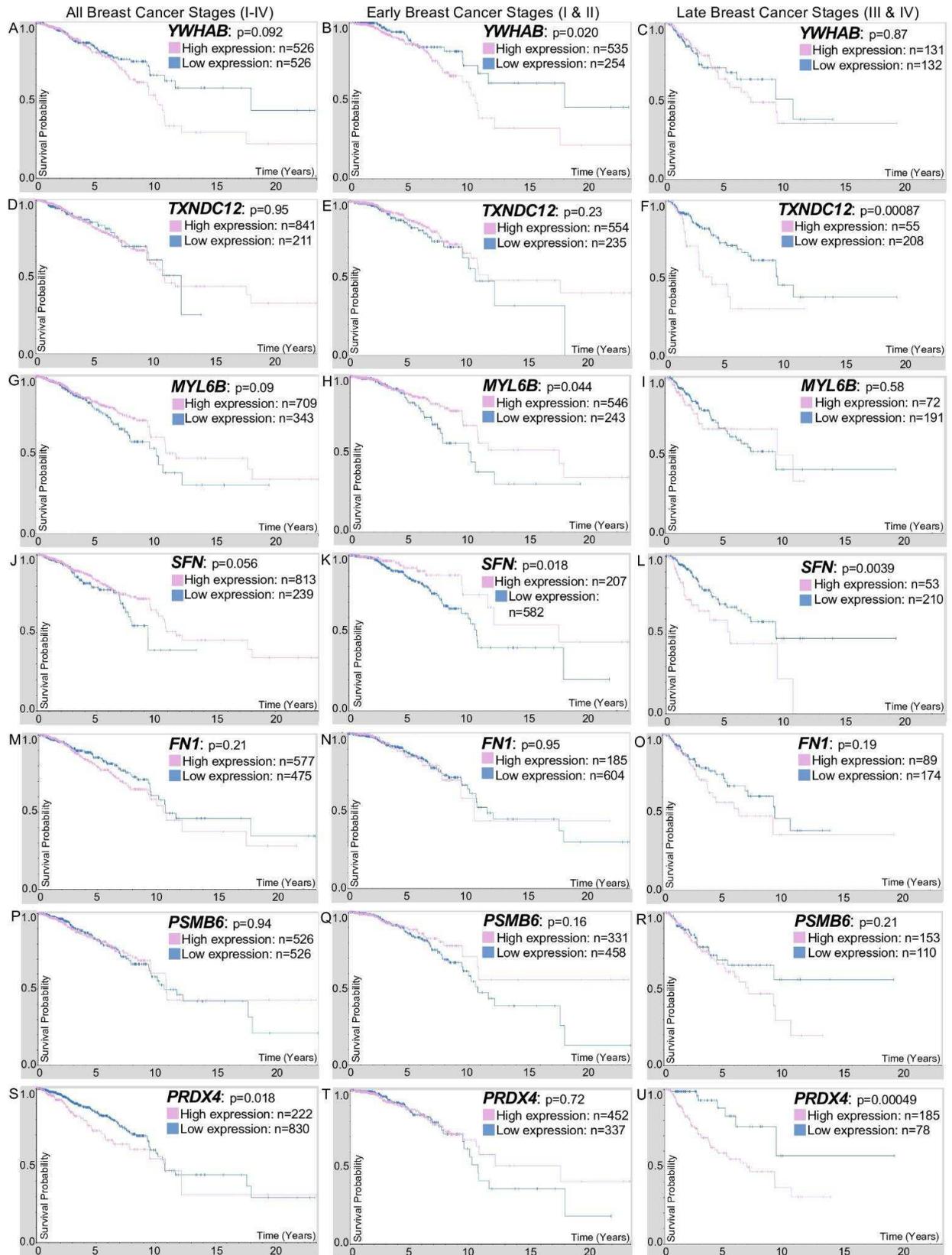


Figure 9. Kaplan-Meier survival plots for secretome markers in all (non-stratified), early (stages I-II), or late (III-IV) stage breast cancer. Survival data presented for secretome marker (A-C) *YWHAB*, (D-F) *TXNDC12*, (G-I) *MYL6B*, (J-L) *SFN*, (M-O) *FNI*, (P-R) *PSMB6*, and (S-U) *PRDX4*. *PEA15* had no data. The x-axis shows time (years), and the y-axis represents survival probability where 1.0 equals 100%.

Discussion

miR526b and miR655 are oncogenic miRNAs that promote aggressive breast cancer phenotypes and alter cells within TME. We have identified that cell-free miRNA and miRNA-high tumor cell secretory proteins change the phenotypes of cells present in TME. Analysis of miR526b and miR655 cell secretomes in ER-positive breast cancer cell lines might decipher the mechanisms of miRNA regulating TME and identify biomarker candidates. Extensive secretome analysis can impose difficulties, as many extracellular proteins are signaling molecules found at low levels hence might not be selected due to higher threshold¹⁷. However, our platform of combining nano-high-performance liquid chromatography with large sensitivity mass spectrometry ensured in-depth, sensitive secretome analysis. After systematic data curation, we identified four upregulated (*YWHAB*, *TXNDC12*, *MYL6B*, *SFN*) and four downregulated (*FNI*, *PSMB6*, *PRDX4*, *PEA15*) markers in miRNA-overexpressed tumor secretomes.

miR526b and miR655 have been shown to induce oxidative stress by overproduction of ROS, and ROS levels were further enhanced during hypoxia in miRNA overexpressed cells^{13,15}. These miRNA-induced functions are also regulated by secretory marker expression by tumor cells. ROS production causes DNA damage so often triggers apoptosis, and in growing aggressive tumors, hypoxia influences the apoptotic pathways. Here we identified that the most enriched biological processes regulated by the eight secretome markers were related to apoptosis regulation and cell redox homeostasis. In miRNA-high breast cancer cells, instead of induction of apoptosis, hypoxia promotes oxidative stress, cell migration, tube formation¹⁵. Secretome markers regulate ROS levels, apoptosis, and hypoxic response, supporting miR526b and miR655 regulating these processes in breast cancer.

Strong relationships between upregulated secretome markers and breast cancer were identified individually when analyzing each secretome marker. *YWHAB* is one of seven 14-3-3 family proteins, and all family members, except *SFN*, have oncogenic properties in epithelial carcinomas and are known to potentiate tumor growth and progression²³. In luminal A breast cancer, *YWHAB* overexpression led to a worse breast cancer prognosis²⁴ and showed higher expressions in non-triple-negative breast cancer tissues and cell lines²⁵. In our study, *YWHAB* was upregulated in both miRNA secretomes by miRNA targeting negative regulator TFs of *YWHAB* and showed strong positive correlations to miRNA cluster expression in luminal A breast cancer. In addition, *YWHAB* was significantly upregulated in luminal A breast cancer cell T47D, HER2-enriched SKBR3, and triple-negative MDA-MB-231 cells. *YWHAB* was also significantly upregulated in breast cancer blood samples and tissues and showed high expression

in immunohistochemistry staining in breast tumors compared to controls. High *YWHAB* expression could significantly predict a worse prognosis in early-stage breast cancer. Overall, *YWHAB* promotes oncogenic breast cancer functions and shows the potential to serve as a biomarker for luminal A breast cancer.

TXNDC12 inhibits cell death, regulates intrinsic apoptotic signaling pathways, and modulates oxidoreductase activity during oxidative stress²⁶. Upregulation of *TXNDC12* inhibits apoptosis by endoplasmic reticulum stress-inducing agents and promotes EMT and metastasis in many epithelial cancers^{27,28}. However, *TXNDC12* has not been studied in breast cancer. Two of its gene family members, *ARG2* and *ARG3*, are known serum-based breast cancer biomarkers²⁹. In our study, *TXNDC12* was upregulated in both miRNA secretomes; however, *TXNDC12* is a predicted target of miR655. Negative regulator TFs of *TXNDC12*, *NANOG*, and *KLF10*, are targets of both miR526b and miR655, which may explain *TXNDC12* upregulation in miRNA secretomes. In our study, *TXNDC12* had significant positive correlations with both miRNA clusters in luminal A breast cancer and was also significantly higher in T47D and SKBR3 cell lines. *TXNDC12* was highly expressed in breast cancer tissue, and high *TXNDC12* expression leads to significantly reduced survival in late-stage breast cancer. Thus, *TXNDC12* shows potential as a prognostic biomarker for metastatic breast cancer. Since both miRNAs and *TXNDC12* promote EMT, *TXNDC12* likely collaborates with miRNAs to promote breast cancer cell migration, invasion, and metastasis.

MYL6B is an essential light chain subunit for myosin motor proteins and regulates cell mobility functions³⁰. Although no studies have investigated *MYL6B* in breast cancer, in other epithelial cancers, such as rectal adenocarcinoma, *MYL6B* overexpression promoted EMT, and high *MYL6B* expression was a predictor of poor patient survival³¹. In our study, *MYL6B* is upregulated in miRNA secretions potentially through both miRNAs targeting *MYL6B*'s negative regulator TF *MECP2*. Since miRNA secretions enhance mesenchymal phenotypes in primary human endothelial and MCF7 cells¹⁴, *MYL6B* could contribute towards miRNA regulating EMT. Furthermore, *MYL6B* had a significant positive correlation with miRNA cluster expression in luminal A breast cancer. Additionally, *MYL6B* has significantly increased expression in luminal A T47D cells and in the blood and tissue of breast cancer patients; hence, *MYL6B* can serve as a blood biomarker. However, *MYL6B* mRNA expression is not associated with breast cancer patient survival. So, *MYL6B* might be a diagnostic biomarker to add to the list along with pri-miRNAs.

SFN is another member of the 14-3-3 protein family and is the only member with a tumor suppressive role³². Many studies highlight the importance of the loss of *SFN* expression in breast cancer development³³. In this study, *SFN* was upregulated in both miRNA secretomes. However, there were no common *SFN* TFs regulated by both miRNAs. However, the beta and alpha thyroid hormone receptors (*THRB* and *THRA*), two paralogous genes of the same family, are negative regulators of *SFN* and are targets of miR526b and miR655, respectively. Therefore, *SFN* could be upregulated in miRNA secretomes through these negative regulator TFs.

Furthermore, *SFN* expression in human tumor tissues could significantly differentiate between early- and late-stage breast cancer, as higher *SFN* expression significantly decreased breast cancer patient survival in late stages, while lower *SFN* expression significantly reduced survival in early stages. This might be due to *SFN* having a tumor suppressor role, and loss of *SFN* is required to initiate metastasis, and at a progressive stage, mutated *SFN* behaves more like an oncogene. Additionally, *SFN* expression correlates with miRNA cluster expression, specifically in the luminal A tumor subtype, and is highly expressed in breast cancer tissue and blood. miRNA-overexpression in MCF7 cells promotes aggressive tumor phenotypes, and high *SFN* expression is significantly associated with poor patient survival in progressive tumor stages; hence *SFN* has more oncogenic functions in advanced stages of breast cancer.

The TF *MYC* is a regulator of seven secretome markers except *SFN*. *MYC* had been deemed both a positive and negative regulator of *YWHAB* and *FNI*. *MYC* regulates ~15% of human genes, and in breast cancer, *MYC* target genes participate in the CSC phenotype, angiogenesis, cell growth, and transformation³⁴. We have previously observed that miRNA enhanced CSC phenotypes in breast cancer^{11,12}, and *MYC* expression in MCF7-miR655 cells is marginally upregulated¹²; however, *MYC* is a miR655 target. We speculate that if *MYC* is a miR655 target, its expression may be compensated for by other downstream effects. Thus, the underlying relationship between *MYC* in the secretome and the CSC phenotype of miRNA-high breast tumor cells requires further investigation.

All secretome markers are deemed prognostic markers in at least one other epithelial cancer and can be identified in the blood³⁵. Downregulated secretome markers *PSMB6*, *PRDX4*, and *PEA15* showed significant negative correlations with miRNA clusters expression in the luminal A subtype tumors. However, they were upregulated in breast cancer tumor tissue, *PSMB6* and *PEA15* were significantly upregulated in the blood of breast cancer patients, and high expression of *PRDX4* had significantly reduced survival in late-stage breast cancer. *FNI* mRNA was downregulated in all luminal breast cancer cell lines, supporting secretome data. And FN1, PRDX4, and PEA15 immunohistochemistry staining showed lower expression in tumoral samples than the normal controls, indicating these markers might have tumor-suppressor-like functions. The observed downregulation of these four markers in miRNA-secretomes could be related to miRNA epigenetically regulating these marker expressions in breast cancer. Upregulated secretome markers *YWHAB*, *TXNDC12*, *MYL6B*, and *SFN* might be more beneficial in establishing a battery of breast cancer biomarkers alongside pri-miR526b and pri-miR655 since their translational bioinformatic analysis supports our secretome findings.

To our knowledge, this is the first time secretory proteins from miR526b and miR655 overexpressing ER-positive cells have been identified and investigated to know the roles of miRNA in TME regulation. Upregulated secretome markers might increase the sensitivity and specificity of breast cancer early detection in combination with pri-miRNAs, allowing us to develop a battery of blood-based breast cancer biomarkers for luminal A or ER-positive breast cancers. Although we used only ER-positive breast cancer cells for our secretome analysis, we

validated secretome markers in various breast cancer cell lines and tumor tissues. In addition, the majority of secretome marker expressions were correlated with miRNA cluster expressions in luminal A breast cancer, strengthening our hypothesis that upregulated secretome markers might be potential blood-based biomarkers for ER-positive breast cancer. Many low abundances but vital secretory proteins were excluded during data curation to find common secretory proteins present in both miRNA-secretomes. In the future, each miRNA secretome will be investigated separately to identify specific miRNA secretory proteins regulating individual miRNA TMEs. Additionally, we will further investigate the roles of upregulated secretome markers in breast cancer progression and metastasis. Some of the identified markers are showing potential as drug targets. Novel secretory proteins identified in miRNA-high tumor secretions will be tested in breast cancer patient plasma and tissue to test their blood-based breast cancer biomarker potential.

Materials and Methods

Collection of conditioned media

MCF7-Mock, MCF7-miR526b, and MCF7-miR655 cell lines were grown in complete RPMI 1640 media (Gibco, ON, Canada) with 10% fetal bovine serum (VWR, ON, Canada) and 1% Pen-Strep. Stable miRNA-overexpressed cells and mock (empty vector-transfected) cells received Geneticin (G418) at a concentration of 200 μ g/ml (Biobasic, ON, Canada), and all cell lines were maintained in cell culture at 37°C, 5% CO₂. Once cell confluency reached 90%, all cells were washed with 1x phosphate-buffered saline (PBS) (Gibco, ON, Canada) to remove any trace of complete media. All cells were followed by being serum-starved with basal media for 16 hours before collection of the cell secretion in the media. The conditioned media, which contains all the secretory proteins and cell metabolites, was collected from all cell lines.

Preparation of proteins to be analyzed by nanoHPLC-MS

Conditioned media proteins were precipitated overnight with 35% ethanol, followed by acidification with sodium acetate and the addition of a digestion buffer (1% sodium deoxycholate and 50mM NH₄HCO₃). Precipitated proteins were quantified by BCA Protein Assay Kit (Pierce, Rockford, IL) with at least 100 μ g protein used per cell line. The protein underwent reduction and alkylation (dithiothreitol and iodoacetamide) before digestion with trypsin. Peptides were isolated by stage tip purification before analysis by nanoHPLC-MS (Agilent 6530 Accurate-Mass Q-TOF LC/MS, Santa Clara, CA). In a single run, we had two experimental replicates of each sample. So, with four biological replicates, we generated n=8 sets of data for each cell line.

Identification of secreted protein IDs

NanoHPLC-MS used Mascot Server (version 2.6) to identify peptides³⁶. A full scan of peptides was quantified by MS1 filtering, extracted ion chromatogram, and verified by spectral matching (Uniprot human protein reference data file) and amino acid database search (<1% false discovery rate (FDR)). Mass spectrometry raw data files and peptide masses were analyzed using Skyline (version 20.1.0.155), which allowed us to acquire a list of protein IDs³⁷. MCF7-miR526b and MCF7-miR655 cells (case) were normalized to MCF7-Mock (control), and Skyline gave their protein IDs with corresponding fold changes and p-values.

GO analysis of all differentially expressed miRNA-high proteins

All protein IDs differentially expressed between miRNA-high versus miRNA-low cells were entered into The Gene Ontology Resource for Homo sapiens (release 2021-09-01), and data was extracted^{38,39}.

Threshold determination and data curation

We extracted the volcano plots for differentially expressed secreted protein IDs for MCF7-miR526b and MCF7-miR655 from Skyline (version 20.1.0.155)³⁷. First, we converted Skyline's original fold change, and p-value data of differential secreted protein IDs to \log_2 fold change and $-\log_{10}$ p-value, respectively, since Skyline gave volcano plots with this output and variables. Next, we submitted protein IDs to Uniprot and extracted protein names, primary gene names, and synonyms⁴⁰. One Skyline protein ID was unmapped (no peptide ID found), and 13 Skyline protein IDs had no gene names in Uniprot. Therefore, these Skyline protein IDs were excluded from our study. Gene names were used for all differentially secreted protein IDs. If a protein ID corresponded to the same gene name, and one or both \log_2 fold change and $-\log_{10}$ p-value differed, we took the average of all IDs with the same gene name. The \log_2 fold change and $-\log_{10}$ p-value of gene names were further analyzed to identify proteins that abided by our threshold in at least one miRNA-secretome. We had no statistically significant protein IDs ($p < 0.05$), so we considered the top 92nd percentile of data, which roughly translates to 0.3 $-\log_{10}$ p-value.

Generation of heatmaps

We used R-studio version 1.4.1103 and R version 4.0.3 to make gene heatmaps. We used the gene name and corresponding mean \log_2 fold change values for MCF7-miR526b and MCF7-miR655-secretomes as an input file for R to make heatmaps. Heatmap functions with default arguments in R were used to produce the agglomerative hierarchical clustered heatmap.

Breast-specific proteome

The 14227 human protein-coding genes within the breast-specific proteome were extracted from the HPA (version 20.1)³⁵ and compared to genes within our secretome threshold.

Secretome prediction methods

We downloaded the classical secretome prediction method data from the HPA (version 20.1)³⁵ for HPA, MDSEC, Phobius, SignalP, and SPOCTOPUS and compared genes within each method to our list of genes that followed our threshold.

For the non-classical secretome prediction method, SecretomeP, FASTA sequences of our threshold protein IDs were obtained via Uniprot and submitted in SecretomeP 2.0 (December 2020) for mammalian sequences^{19,40}. Genes were considered non-classically secreted following the previously established guideline of a neural network score >0.6 and odds >3 ^{19,41}. Secretome markers were considered classically secreted if found in both classical and non-classical secretion methods¹⁹.

Secretome marker functions and GO analysis

Secretome marker general functions were obtained by using www.GeneCards.org version 5.6.0 Build 515⁴². Individual GO functions of each secretome marker were obtained through Uniprot (Last modified: February 2, 2021)⁴⁰ and QuickGo (GO version 2021-11-08)²⁶. Shared GO of the eight secretome markers was found by analyzing all protein IDs as one quarry into the Gene Ontology Resource for Homo sapiens (release 2021-09-01)^{38,39}. This obtained cellular component and Reactome pathway results. Additionally, secretome markers were analyzed with STRING database (version 11.0b)⁴³ to identify biological processes and molecular functions GO.

miRNA target genes and TF analysis

miR526b (hsa-mir-526b) and miR655 (hsa-mir-655) targets were downloaded for both mature five-prime sequences from TargetScanVert (Release 7.1)⁴⁴. Only *TXNDC12* was found to be a predicted direct target of miR655. Therefore, a combined list was created, which included common targets of both miRNAs and Enrichr, then identified the TFs for our secretome markers⁴⁵. Each marker and their TFs were matched against miR526b and miR655 common targets. If the miRNA target gene matches the TF of the gene, and if that TF is a positive regulator of the gene, then gene expression will be downregulated. In contrast, if miRNA targets the negative regulator TF of a gene, gene expression will be upregulated.

miRNA cluster and secretome marker correlations

miRNAs within miR526b and miR655 clusters were identified through miRBase (release 22.1)⁴⁶. miR526b's cluster had expression data for two miRNAs, miR516a-1, and miR516a-2. miR655's cluster had expression data for 16 miRNAs within the cluster: miR134, miR154, miR369, miR376c, miR379, miR381, miR382, miR409, miR410, miR411, miR487b, miR495, miR539, miR655, miR758, and miR889. We took the mean of the miRNAs expression within each miRNA cluster and presented this as miR526b and miR655 cluster expression. All 19 breast cancer studies were included, and secretome markers and miRNAs within miR526b and miR655 clusters were submitted with all options default on cBioPortal (v 3.7.15)⁴⁷. The "mRNA

expression (microarray)” and “microRNA expression” data were downloaded in tab-delimited format. Pearson correlation coefficient (R) was measured between each secretome marker and miR526b or miR655 cluster expression in breast cancer tumor subtypes.

Secretome markers gene expression in breast cancer cell lines

All secretome marker gene expressions in breast cancer cell lines were obtained from the GENT2 database [GPL570 platform (HG-U133_Plus_2)]⁴⁸. Data was analyzed using Microsoft Excel.

Secretome markers mRNA expression in breast cancer tissue

mRNA expression of secretome markers in breast cancer tumor tissue versus controls was obtained from GEPIA 2 (2018 version)⁴⁹, with all parameters set as default.

Secretome markers in blood exosomes

We used the HUPO website (www.hupo.org) to retrieve the “Human Plasma Proteome Project.” We were redirected to PeptideAtlas (build: Plasma Non-Glyco 2017-04 - Mapping 2017-05-12)⁵⁰, and secretome markers’ presence in human plasma was extracted. Secretome marker expression in breast cancer blood exosomes versus healthy controls was retrieved from ExoRBase2.0 (version 2.0)⁵¹.

Secretome markers Immunohistochemistry

Normal and breast cancer immunohistochemistry staining data were obtained from the HPA (version 20.1)³⁵. For each normal sample, adipocyte, glandular cell, and myoepithelial cell parameters were combined (staining, intensity, and quantity), and the median intensity value was considered. *TXNDC12* had no immunohistochemistry data available in the HPA. For most secretome markers, data for several antibodies were available. We selected the antibody that gave signals for most tumor samples, and signal intensity was evident.

Secretome markers Kaplan-Meier survival analysis

Kaplan-Meier survival data were retrieved from the HPA (version 20.1)³⁵. Non-stratified (all stages), early stages (stages I & II) included stages I, IA, IB, II, IIA, IIB. Late stages (stages III & IV) included stages III, IIIA, IIIB, IIIC, IV. The cut-off values used for each secretome marker shown in order of all early and late stages are as follows: *YWHAB* (107.89, 95.57, and 109.48), *TXNDC12* (15.97, 17.67, and 22.97), *MYL6B* (7.01, 6.76, and 11.15), *SFN* (17.48, 49.56, and 57.21), *FNI* (192.40, 425, and 309), *PSMB6* (50.47, 54.28, and 47.40) and *PRDX4* (62.87, 37.28, and 31.98). *PEAI5* had no data.

Statistical analysis

Skyline calculated false discovery rate (FDR) corrected p-values to identify differently secreted miRNA-high proteins using the MSstats R package (version 3.13.6)³⁷. GO statistics from The Gene Ontology Resource and STRING database used FDR corrected p-values^{38,39,43}. These p-values were converted to $-\log_{10}$ p-value in Microsoft Excel to be easily graphed. miRNA cluster and secretome marker correlation were measured using Pearson Correlation coefficient, and gene expression in cell lines was compared using unpaired, two-tailed *t*-tests in GraphPad Prism (Version 9.2.0). Gene expression in normal and breast cancer tissue was measured using the absolute value of fold change cutoff of one and q-value cutoff of 0.01 (ANOVA)⁴⁹. Gene expression in blood exosomes between two groups was compared using unpaired, two-tailed *t*-test⁵¹. Survival analysis of secretome markers p-value was determined using log-rank p-value for Kaplan-Meier survival plots showing results from analysis of correlation between mRNA expression level and patient survival³⁵.

Data availability

Breast-specific proteome data was extracted from the HPA (version 20.1)³⁵. Classical secretome prediction method data was extracted from the HPA (version 20.1). Non-classical secretome prediction method data was completed in SecretomeP 2.0 (December 2020)¹⁹. miR526b and miR655 targets were identified through TargetScanVert (Release 7.1)⁴⁴. Enrichr identified the TFs of secretome markers⁴⁵. miR526b and miR655 clusters were identified through miRBase (release 22.1)⁴⁶. miRNA cluster expression and secretome marker mRNA expression were obtained through cBioPortal (v 3.7.15)⁴⁷. Secretome marker gene expressions in breast cancer cell lines were obtained via GENT2 [GPL570 platform (HG-U133_Plus_2)]⁴⁸. mRNA expression of secretome markers in breast tumors and controls was collected from GEPIA 2 (2018 version)⁴⁹. Secretome marker expression in breast cancer blood exosomes and healthy controls were collected from ExoRBase2.0 (version 2.0)⁵¹. Breast cancer and normal immunohistochemistry staining data were obtained from the HPA (version 20.1)³⁵. Kaplan-Meier survival data were found from the HPA (version 20.1)³⁵. Computer code and all other data supporting the findings of this study are available from the corresponding author upon request.

References

1. Sung, R. *et al.* Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* <http://gco.iarc.fr/today/home> (2021).
2. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **68**, 394–424 (2018).
3. SEER Explorer. SEER Explorer: An interactive website for SEER cancer statistics. Surveillance Research Program, National Cancer Institute. <https://seer.cancer.gov/explorer/> (2021).
4. American Cancer Society, A. C. S. Breast Cancer Facts & Figures 2019-2020. *American Cancer Society* <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2019-2020.pdf> (2019).
5. Krzyszczyk, P. *et al.* The growing role of precision and personalized medicine for cancer treatment. *Technology* **6**, 79–100 (2018).
6. Ebell, M. H., Thai, T. N. & Royalty, K. J. Cancer screening recommendations: an international comparison of high income countries. *Public Health Rev.* **39**, (2018).
7. Giraldo, N. A. *et al.* The clinical role of the TME in solid cancer. *Br. J. Cancer* **120**, 45–53 (2019).
8. Mukherjee, P. & Mani, S. Methodologies to decipher the cell secretome. *Biochim. Biophys. Acta* **1834**, 2226–2232 (2013).
9. Pavlou, M. P. & Diamandis, E. P. The cancer cell secretome: A good source for discovering biomarkers? *J. Proteomics* **73**, 1896–1906 (2010).
10. MicroRNAs: The Master Regulators of the Breast Cancer Tumor Microenvironment. in *Handbook of Oxidative Stress in Cancer: Therapeutic Aspects* (Springer-Nature).
11. Majumder, M., Landman, E., Liu, L., Hess, D. & Lala, P. K. COX-2 Elevates Oncogenic miR-526b in Breast Cancer by EP4 Activation. *Mol. Cancer Res. MCR* **13**, 1022–1033 (2015).
12. Majumder, M. *et al.* COX-2 induces oncogenic micro RNA miR655 in human breast cancer. *Sci. Rep.* **8**, 327 (2018).
13. Shin, B. *et al.* miR526b and miR655 Induce Oxidative Stress in Breast Cancer. *Int. J. Mol. Sci.* **20**, 4039 (2019).
14. Hunter, S., Nault, B., Ugwuagbo, K. C., Maiti, S. & Majumder, M. Mir526b and Mir655 Promote Tumour Associated Angiogenesis and Lymphangiogenesis in Breast Cancer. *Cancers* **11**, (2019).
15. Gervin, E. *et al.* Chemically Induced Hypoxia Enhances miRNA Functions in Breast Cancer. *Cancers* **12**, 2008 (2020).
16. Majumder, M., Ugwuagbo, K. C., Maiti, S., Lala, P. K. & Brackstone, M. Pri-miR526b and Pri-miR655 Are Potential Blood Biomarkers for Breast Cancer. *Cancers* **13**, 3838 (2021).
17. Wang, X.-D. *et al.* Subtype-specific secretomic characterization of pulmonary

- neuroendocrine tumor cells. *Nat. Commun.* **10**, 3201 (2019).
18. Kammers, K., Cole, R. N., Tiengwe, C. & Ruczinski, I. Detecting significant changes in protein abundance. *EuPA Open Proteomics* **7**, 11–19 (2015).
 19. Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G. & Brunak, S. Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel. PEDS* **17**, 349–356 (2004).
 20. Grafone, T., Palmisano, M., Nicci, C. & Storti, S. An overview on the role of FLT3-tyrosine kinase receptor in acute myeloid leukemia: biology and treatment. *Oncol. Rev.* **6**, e8 (2012).
 21. Griss, J. *et al.* ReactomeGSA - Efficient Multi-Omics Comparative Pathway Analysis. *Mol. Cell. Proteomics MCP* **19**, 2115–2125 (2020).
 22. Cantwell-Dorris, E. R., O’Leary, J. J. & Sheils, O. M. BRAFV600E: Implications for Carcinogenesis and Molecular Therapy. *Mol. Cancer Ther.* **10**, 385–394 (2011).
 23. Tang, Y., Lv, P., Sun, Z., Han, L. & Zhou, W. 14-3-3 β Promotes Migration and Invasion of Human Hepatocellular Carcinoma Cells by Modulating Expression of MMP2 and MMP9 through PI3K/Akt/NF- κ B Pathway. *PloS One* **11**, e0146070 (2016).
 24. Zhu, Q.-N., Renaud, H. & Guo, Y. Bioinformatics-based identification of miR-542-5p as a predictive biomarker in breast cancer therapy. *Hereditas* **155**, 17 (2018).
 25. Boac, B. M. *et al.* Expression of the BAD pathway is a marker of triple-negative status and poor outcome. *Sci. Rep.* **9**, 17496 (2019).
 26. Binns, D. *et al.* QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics* **25**, 3045–3046 (2009).
 27. Jeong, W., Lee, D.-Y., Park, S. & Rhee, S. G. ERp16, an endoplasmic reticulum-resident thiol-disulfide oxidoreductase: biochemical properties and role in apoptosis induced by endoplasmic reticulum stress. *J. Biol. Chem.* **283**, 25557–25566 (2008).
 28. Yuan, K. *et al.* TXNDC12 promotes EMT and metastasis of hepatocellular carcinoma cells via activation of β -catenin. *Cell Death Differ.* **27**, 1355–1368 (2020).
 29. Garczyk, S. *et al.* AGR3 in Breast Cancer: Prognostic Impact and Suitable Serum-Based Biomarker for Early Cancer Detection. *PLOS ONE* **10**, e0122106 (2015).
 30. Heissler, S. M. & Sellers, J. R. Myosin light chains: Teaching old dogs new tricks. *BioArchitecture* **4**, 169–188 (2014).
 31. Li, J.-L., Wang, Z.-Q. & Sun, X.-L. MYL6B drives the capabilities of proliferation, invasion, and migration in rectal adenocarcinoma through the EMT process. *Open Life Sci.* **15**, 522–531 (2020).
 32. Kaplan, A., Bueno, M. & Fournier, A. E. Extracellular functions of 14-3-3 adaptor proteins. *Cell. Signal.* **31**, 26–30 (2017).
 33. Simpson, P. T. *et al.* Distribution and significance of 14-3-3 σ , a novel myoepithelial marker, in normal, benign, and malignant breast tissue. *J. Pathol.* **202**, 274–285 (2004).
 34. Xu, J., Chen, Y. & Olopade, O. I. MYC and Breast Cancer. *Genes Cancer* **1**, 629–640

- (2010).
35. Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
 36. Perkins, D. N., Pappin, D. J. C., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *ELECTROPHORESIS* **20**, 3551–3567 (1999).
 37. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinforma. Oxf. Engl.* **26**, 966–968 (2010).
 38. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
 39. Gene Ontology Consortium. The Gene Ontology resource: enriching a GOLD mine. *Nucleic Acids Res.* **49**, D325–D334 (2021).
 40. Consortium, T. U. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).
 41. Lin, B. *et al.* Evidence for the presence of disease-perturbed networks in prostate cancer cells by genomic and proteomic analyses: a systems approach to disease. *Cancer Res.* **65**, 3081–3091 (2005).
 42. Stelzer, G. *et al.* The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinforma.* **54**, 1.30.1-1.30.33 (2016).
 43. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
 44. Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **4**, (2015).
 45. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90-97 (2016).
 46. Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **47**, D155–D162 (2019).
 47. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
 48. Park, S.-J., Yoon, B.-H., Kim, S.-K. & Kim, S.-Y. GENT2: an updated gene expression database for normal and tumor tissues. *BMC Med. Genomics* **12**, 101 (2019).
 49. Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **47**, W556–W560 (2019).
 50. Desiere, F. *et al.* The PeptideAtlas project. *Nucleic Acids Res.* **34**, D655-658 (2006).
 51. Li, S. *et al.* exoRBase: a database of circRNA, lncRNA and mRNA in human blood exosomes. *Nucleic Acids Res.* **46**, (2017).

Acknowledgments

We would like to thank Kingsley Chukwunonso Ugwuagbo for helping us prepare conditioned media for Mass-Spec analysis. This work is funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant and the Brandon University Research Committee (BURC) Grant to M.M. M.M. is supported by the Canada Research Chair Program (CRCP) and Canada Foundation for Innovation (CFI). R.F. and B.N. have received Research Manitoba MSc scholarship. R.M.O. received CIHR MSc scholarship.

Author Contributions

R.F. contributed manuscript writing, data analysis, bioinformatics, figures and image data processing, and manuscript editing. R.M.O. contributed manuscript writing and editing. B.N. contributed data analysis, bioinformatics, figures and image data processing, and manuscript editing. S.M. contributed data analysis, bioinformatics, figures and image data processing, and manuscript editing. V.C.C. contributed to running nanoHPLC-MS samples. M.M. contributed concept, project design, funding acquisition, supervision, and manuscript writing and editing. All authors have read and agreed to the published version of the manuscript.

Competing Interests

The authors declare no conflict of interest.

Supplementary Files

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

- [SupplementaryData1.xlsx](#)
- [SupplementaryData2.xlsx](#)
- [SupplementaryData3.xlsx](#)
- [SupplementaryData.pdf](#)
- [DescriptionofAdditionalSupplementaryFiles.pdf](#)