# *DOCTORAL THESIS*

UNIVERSIDAD DE GRANADA

**DOCTORATE PROGRAMME OF BIOMEDICINE**

**EL PAPEL DE LOS MICROARNs EN EL DIAGNÓSTICO Y PRONÓSTICO DEL CÁNCER DE MAMA**

---

**THE ROLE OF MICRORNAs IN THE DIAGNOSIS AND PROGNOSIS OF BREAST CANCER**


**Coral González Martínez**

Thesis supervisors:

José Antonio Lorente Acosta

Francisco Gabriel Ortega Sánchez

Genyo

CENTRO PFIZER-UNIVERSIDAD DE GRANADA-JUNTA DE ANDALUCÍA
DE GENÓMICA E INVESTIGACIÓN ONCOLÓGICA

Editor: Universidad de Granada. Tesis Doctorales
Autor: Coral González Martínez
ISBN: 978-84-1195-811-0
URI: https://hdl.handle.net/10481/104891

*Estoy dispuesto a seguir luchando, pero necesito saber qué es lo que estoy luchando*

*Siddhartha Mukherjee, "El Emperador de todos los males"*

Gracias a mis directores, Francisco Gabriel Ortega y José Antonio Lorente, y a la IP de mi laboratorio, Mª José Serrano. Por ser guía en este camino y luz en la oscuridad. Sin vosotros, esta tesis no habría sido posible.

Gracias al Dr. Pieter Vader y a su laboratorio, por acogerme y hacerme sentir como en casa. Con vosotros, la ciencia y la vida se hace fácil, dos cosas bastante difíciles de lograr.

Estoy infinitamente agradecida de haber podido realizar esta tesis doctoral en la pequeña gran familia del centro de investigación GENyO, donde cada saludo por los pasillos te hace los días oscuros un poco más luminosos. Gracias al laboratorio vecino, el 4, a Mariadel, siempre con una sonrisa cómplice a través del cristal y dispuesta a prestarte un anticuerpo. A Alba, Rita y Manzaneque, por ese tránsito entre laboratorios y conversaciones que hacen los días amables. A todos los miembros del laboratorio 6 y 7. A la sala de la primera planta, espacio de trabajo – no trabajo, pero de complicidad, sin duda. Mari Carmen, eres simplemente maravillosa. Tere, has pasado de todo y por todo, hablar contigo siempre es un bálsamo. Josema, tenemos el peso de ser los Capricornio de GENyO. Victoria, Heavy, Joan, Gonzalo, Mónica, Peris, no hemos hablado mucho, pero os siento muy cercanos, el milagro de la sala y de compartir penas.

Gracias a Diego y a Alba, por ser mis padrinos en el bautismo de la ciencia.

Gracias al Laboratorio 5, a Valeria y a Juanfry por estar dispuestos a compartir su conocimiento, a Pili, siempre con un apaño para cada problema, a Abel, con tu ser y tus bromas haciendo la poyata un lugar menos hostil, a Jorge, con tu gabinete psicológico y la sabiduría de 100

posdoctorales. A Laura, por tu ansia de aprender y tu bondad, a Fabián, mi sucesor en vesículas extracelulares, tienes el sol de Málaga en tu pecho.

Gracias a ti, Inés. Tu compañía, emotional support y amistad durante estos años ha sido tan profunda, que estoy convencida de que vamos a estar la una en la vida de la otra para siempre. Gracias, gracias, gracias.

Gracias a todos los TFGs y TFMs con los que he coincidido en espacio-tiempo, habéis sido una brisa de aire fresco en un ambiente a veces estanco. Gracias a ti, Chiara, por demostrarme que la amistad que se crea entre pipetas se puede mantener en la distancia.

Gracias al Depto. de Medicina Legal, Toxicología y Antropología física, por ser mi respaldo en la Universidad. Gracias, Miguel, por abrirme nuevos caminos apasionantes en la investigación y por los cafés acompañados de una buena conversación.

Gracias a todos los amigos que conocí en la Universidad y que aún siguen en mi día a día: Irene, Marga, Laura, Juanan, Juanma, Mariem, Mónica y Guty. A día de hoy, sois mi núcleo. Gracias por estar siempre ahí, transitar una tesis es mucho más fácil sabiendo que tienes una red debajo por si te caes.

Gracias a los amigos del otro lado del charco (Melilla), hacéis que volver a mi tierra en los descansos de laboratorio sea volver a casa. Gracias en particular a ti, Sara.

Gracias Cathaysa, por aparecer en mitad de esta etapa de mi vida y quedarte. Por celebrar conmigo mis victorias y ser la mano que me ayuda a salir cada vez que toco fondo. Por tu apoyo, amor y comprensión.

Gracias infinitas a mi familia. A mis padres, María del Mar y Román, por apoyarme en todo lo que habéis podido y más, sin vosotros no sería quien soy. Gracias a mi hermano Romi, por enseñarme a quitarle hierro a la vida. Gracias a mi abuela, cada minuto contigo vale oro.

Gracias a todo el personal clínico, por la gran labor que desempeñáis día a día.

Gracias a todas las pacientes oncológicas que habéis participado en este estudio con la ilusión y la esperanza de un sistema sanitario mejor, a las que seguís luchando y a las que no, esto es por vosotras.

# TABLE OF CONTENTS

## ABSTRACT

Breast cancer is the most frequently diagnosed cancer worldwide in women and the second most common overall, also being the leading cause of cancer mortality in women[1]. In recent years, significant advances have been made in the diagnosis and treatment of this disease, with two key points for effective breast cancer management: 1) early-stage diagnosis and 2) providing timely treatment following diagnosis[2]. Five-year survival rates are 99% when the cancer is localized and 86% when breast cancer is loco-regional. However, five-year survival falls to 31% when the patient presents with distal metastasis to other organs[3].

Among breast cancer screening and diagnostic techniques, mammography is currently the gold standard. However, when an abnormality is detected, a breast biopsy is still necessary to confirm the diagnosis, an invasive procedure that is not recommended for all cancer patients[4]. Similarly, invasive is the early detection of lymph node metastasis (LNM) through sentinel lymph node biopsy. This technique is also limited to patients undergoing initial surgery or neo-adjuvant chemotherapy[5].

These limitations have highlighted the need for new diagnostic and prognostic markers in breast cancer. In this context, microRNAs present themselves as important gene regulators in breast cancer, whose expression variation reflects tumor activity. This doctoral thesis proposes microRNAs as potential markers, evaluating their role at two different stages of the disease: 1) in early-stage breast cancer and 2) in breast cancer with lymph node metastasis. This research was approached from two angles: first, a bioinformatic approach, which included an exhaustive review of all existing literature on the role of microRNAs in LNM, to

then confront this data in a meta-analysis and offer a resulting biomarker from the combination of all studies analyzed. Second, a sequencing of microRNAs encapsulated in extracellular vesicles, thanks to their protection against RNAases and pathological status information, was conducted from two cohorts of breast cancer patients, one in early stages and the other with metastasis to distal organs, as well as a cohort of healthy donors, to analyze expression changes between cancer and non-cancer states, and in disease progression.

The results obtained in the meta-analysis showed a total of two microRNAs differentially expressed between patients with localized early-stage breast cancer and patients with breast cancer with LNM: miR-34a and miR-155. miR-34a, which was down-regulated compared to localized cancer, is a tumor suppressor in the p53 network, while miR-155, overexpressed compared to localized stages, is an oncogene widely proven to promote breast cancer. This analysis provides a potential diagnostic signature of lymph node metastasis in breast cancer, pending future experimental validation.

On the other hand, the sequencing of microRNAs transported in extracellular vesicles yielded a total of three microRNAs that met the criteria for significant differential expression between groups: miR-320b, miR-423-5p, and miR-141-3p. Bioinformatic analysis showed that these microRNAs were involved in other hormone-related cancers, such as prostate and endometrial cancer. It was also found that silencing the target genes of these microRNAs affected key breast pathways, such as proteoglycan production, progesterone-mediated oocyte maturation, and ovarian steroidogenesis.

Similarly, validation of these microRNAs in an independent cohort from the first one, comprising early-stage and metastatic breast cancer patients, as well as healthy donors, showed a variation in differential expression between groups for miR-423-5p. In a second marker validation between healthy women and women with early-stage breast cancer, this microRNA showed large expression differences, with high levels also correlated with the presence of breast cancer.

The location of miR-423-5p in extracellular vesicles was also studied, as well as its presence or absence in EpCAM+ vesicles, hypothesized to be tumor-derived. In cancer patients, miR-423 was found both inside and outside the overall EVs, while in healthy women it was found mainly inside the EVs. After pulling down EVs employing anti-EpCAM immune-magnetic precipitation, this miRNA was also observed in association with EpCAM+ vesicles.

Overall, these results suggest that EVs associated miR-423-5p could be a good candidate for the identification of breast cancer patients from healthy women. Additionally, its increased in extracellular vesicles make it a potential clinical tool for evaluating disease progression, as its study is minimally invasive and can be repeated over time. However, further studies in large cohorts are needed to confirm its clinical utility and additional studies in patients with different stages of breast cancer and breast symptoms and pathologies.

# RESUMEN

El cáncer de mama es el tipo de cáncer más diagnosticado a nivel mundial en mujeres y el segundo en general, siendo además la primera causa por mortalidad de cáncer en mujeres[1].

En los últimos años se han realizado grandes avances en el diagnóstico y tratamiento de esta enfermedad, con dos puntos clave para el manejo del cáncer de mama efectivo: 1. El diagnóstico en estadios tempranos y 2. proporcionar un tratamiento oportuno tras el diagnóstico[2]. Los datos de supervivencia a los 5 años cuando el cáncer es localizado son del 99% y del 86% cuando el cáncer de mama es loco-regional. Sin embargo, la supervivencia a los 5 años cae hasta el 31% cuando la paciente presenta metástasis distal a otros órganos[3].

Entre las técnicas de *screening* y diagnóstico del cáncer de mama, la mamografía es actualmente el *gold standard*. Sin embargo, cuando una anomalía es detectada, sigue siendo necesario realizar una biopsia de la mama para certificar el diagnóstico, un procedimiento invasivo y no recomendado para todas las pacientes con cáncer[4]. Igualmente invasiva es la detección temprana de metástasis en nódulos linfáticos (MNL), mediante biopsia del ganglio linfático centinela. Esta técnica está además limitada a pacientes sometidas a una cirugía inicial o con quimioterapia neoadyuvante[5].

Estos inconvenientes han puesto de manifiesto la necesidad de encontrar nuevos marcadores tanto diagnósticos como pronósticos en el cáncer de mama. En este sentido, los microARNs se presentan como importantes reguladores génicos en el cáncer de mama, cuya variación en la expresión actúa como reflejo de la actividad tumoral. Esta tesis doctoral propone los microARNs como potenciales marcadores, evaluándose su

papel en dos puntos distintos de la enfermedad: 1. en el cáncer de mama en estadios tempranos y 2. En el cáncer de mama con metástasis en los nódulos linfáticos. Para ello, esta investigación se abordó desde dos aproximaciones: Una primera aproximación bioinformática, en la que se ejecutó una revisión exhaustiva de toda la literatura existente sobre el papel que juegan los microARNs en la MNL para, posteriormente, confrontar esos datos en un metaánalisis y ofrecer un biomarcador resultante de la conjunción de todos los estudios analizados. En segundo lugar, se realizó una secuenciación de los microARNs encapsulados en vesículas extracelulares, gracias a la protección que proporcionan frente a la degradación por ARNasas y su información sobre el estado patológico del organismo, de dos cohortes de pacientes con cáncer de mama, una en estadios tempranos y otra con metástasis en órganos distales, así como una cohorte de donantes sanas, con el fin de analizar los cambios de expresión entre cáncer-no cáncer, y en la progresión de la enfermedad.

Los resultados obtenidos en el metaanálisis mostraron un total de dos microARNs diferencialmente expresados entre pacientes con cáncer de mama temprano localizado y pacientes con cáncer de mama con MNL: miR-34a y miR-155. MiR-34a, que resultó regulado a la baja en comparación con el cáncer localizado, es un supresor tumoral perteneciente a la red de p53 mientras que miR-155, sobre-expresado en comparación con estadios localizados, es un oncogén cuya relación con el cáncer de mama como promotor de la enfermedad está ampliamente demostrada. Este análisis ofrece una potencial firma diagnóstica de la metástasis en nódulos linfáticos en cáncer de mama, pendiente de futuras validaciones experimentales.

Por otro lado, la secuenciación de los microARNs transportados en vesículas extracelulares resultó en un total de tres microARNs que cumplieron los criterios de expresión diferencial significativa entre grupos: miR-320b, miR-423-5p y miR-141-3p. Mediante análisis bioinformáticos, se vio que estos microARNs estaban involucrados en otros cánceres de tipo hormonal, como el cáncer de próstata y el cáncer de endometrio. También se comprobó que el silenciamiento de los genes dianas de estos microARNs afectaba en la mama a rutas importantes como la producción de proteoglicanos, la maduración de los oocitos mediante la progesterona, o la esteroidogenesis ovárica.

Así mismo, la validación de estos microARNs en una cohorte independiente a la primera de pacientes con cáncer de mama temprano y metastásico, así como donadoras sanas, mostró una variación de la expresión diferencial entre grupos del miR-423-5p. En una segunda validación entre mujeres sanas y mujeres con cáncer de mama temprano, este microARN presentó grandes diferencias de expresión, estando además correlacionado sus niveles altos con la presencia de cáncer de mama.

También se estudió la localización de miR-423-5p en las vesículas extracelulares en general, así como su presencia o ausencia en vesículas EpCAM+, hipotetizadas como tumorales. Así, se vio que, en pacientes con cáncer, miR-423 estaba tanto dentro como fuera de la vesícula en general, probablemente resultado de los procesos de vertido y de inflamación durante el cáncer, mientras que en mujeres sanas se encontró principalmente en el interior de la vesícula. Igualmente, este miRNA estuvo presente en vesículas EpCAM+.

En conjunto, estos resultados sugieren que miR-423-5p podría ser un buen candidato para diferenciar pacientes con cáncer de mama de mujeres sanas. Además, su incremento de expresión y su presencia tanto en vesículas extracelulares como circulante, lo hacen una potencial herramienta clínica para evaluar la evolución de la enfermedad, siendo su estudio poco invasivo y repetible a lo largo del tiempo. Sin embargo, más estudios son necesarios para avalar su utilidad clínica, sobre todo en pacientes con cáncer de mama metastásico.

**Bibliography / Bibliografía**


1. World Health Organization. Breast Cancer. 2024
2. Bhushan A, Gonsalves A, Menon JU. Current state of breast cancer diagnosis, treatment, and theranostics. Pharmaceutics. 2021;13(5):1–24
3. American Cancer Society. Understanding a Breast Cancer Diagnosis, Atlanta. Https://WwwCancerOrg/Cancer/Breast-Cancer/About/Types-of-Breast-CancerHtml#References. 2016;1–36.
4. Løberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. Breast Cancer Res. 2015;17(1):1–12.
5. Trialists TS. The UK Standardisation of Breast Radiotherapy (START) Trial A of radiotherapy hypofractionation for treatment of early breast cancer: a randomised trial. Lancet Oncol. 2008;9(4):331–41

# ABBREVIATIONS

AC: Adjuvant Chemotherapy

Ago: Argonaute

AI: Aromatase Inhibitor

AJCC: American Joint Committee on Cancer

ALN: Axillary Lymph Nodes

ALND: Axillary Lymph Node Dissection

ARE: AU-rich Element

ATCC: American Type Culture Collection

AUC: Area Under the Curve

BC: Breast Cancer

BCS: Breast-Conserving Surgery

BMI: Body Mass Index

IDC: Invasive Ductal Carcinoma

IHC: Immunohistochemical

ILC: - Invasive lobular carcinoma

ILV: Intraluminal Vesicle

LCIS: - Lobular Carcinoma In Situ

LNM: Lymph Node Metastasis

MHCI: Major Histocompatibility Complex Class I

miRISC: miRNA-induced silencing complex

miRNAs: microRNAs

MRD: Minimal Residual Disease

MRI: Magnetic Resonance Imaging

MV: Microvesicle

BRCA1: Breast Cancer Susceptibility Gene 1

BRCA2: Breast Cancer Susceptibility Gene 1

cDNA: complementary DNA

CTCs: Circulating Tumor Cells

DCIS: Ductal Carcinoma In Situ

DGCR8: DiGeorge critical region 8

dnMBC: De Novo Metastasis Breast Cancer

EGFR1: Epidermal Growth Factor Receptor Type 1

EMT: Epithelial – Mesenchymal Transition

ER: Estrogen Receptor

ESCRT: Endosomal Sorting Complex Required for Transport

MVB: Multivesicular Bodies

NAC: Neo-adjuvant Chemotherapy

NGS: Next Generation Sequencing

NTA: Nanoparticle Tracking Analysis

OS: Overall Survival

PET/CT: Positron Emission Tomography/Computed Tomography

PR: Progesterone Receptor

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis

QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies

RISC: RNA-induced silencing complex

RT: Radiotherapy

EV: Extracellular Vesicle

FC: Fold Change

FDR: False Discovery Rate

FFPE: Formalin-Fixed Paraffin-Embedded

GEP/S: Gene Expression Profiling/Signatures

GLOBOCAN: Global Cancer Observatory

HDI: Human Development Index

HER2: Human Epidermal Growth Factor 2

HMW: High Molecular Weight

HR: Hazard Ratio

SEC: Size Exclusion Chromatography

SEER: Surveillance, Epidemiology and End Results

SLNB: Sentinel Lymph Node Biopsy

TEM: Transmission Electron Microscopy

TME: Tumor Microenvironment

TNBC: Triple Negative Breast Cancer

TNM: Tumor, Node and Metastasis

UTR: 3' Untranslated Region

WHO: World Health Organisation

# CHAPTER I. INTRODUCTION

## 1. BREAST CANCER

According to the World Health Organisation (WHO), breast cancer (BC) is a disease caused by an uncontrolled growth of cells located in the breast that eventually forms a tumor in the area of the breast[1].

### 1.1 EPIDEMIOLOGY

Based on Global Cancer Observatory (GLOBOCAN) estimates for 2022, breast cancer is the most commonly diagnosed cancer in women worldwide (23.8%) and the second most common cancer in both sexes (11.5%) after lung cancer (Figure 1a). The regions with the highest recorded incidence of breast cancer are Asia with 985, 817 cases (42.9%) and Europe with 557,532 cases (24.3%)[2]. This increase in the incidence rate in Human Development Index (HDI) countries may be due to a higher prevalence of numerous reproductive and lifestyle-related risk factors, such as older age at first birth, fewer children, oral contraceptives, alcohol consumption and physical inactivity[3].

Similarly, breast cancer is the leading cause of cancer mortality in women globally, with 666,103 deaths in 2022 (15.4%) (Figure 1b). The continent with the highest breast cancer mortality rate is Asia, accounting for 315,309 deaths (47.3%), with Japan and the Republic of Korea being the countries with the highest mortality rates [4]. Similarly, many countries in sub-Saharan Africa have high mortality rates [5], so it appears that geographical and temporal variations in breast cancer mortality are related to the level of health service coverage[6].

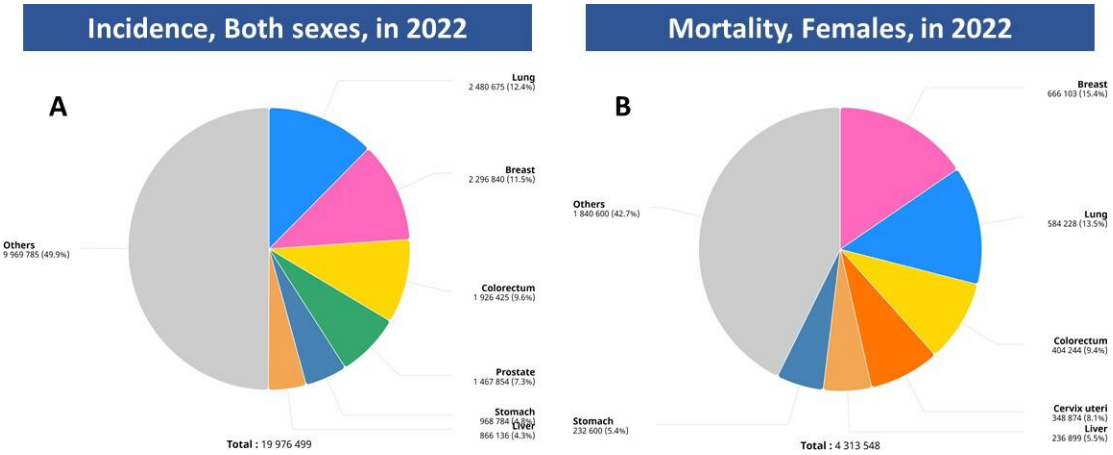



Figure 1. Estimated incidence in 2022 in all cancer types worldwide, both sexes (A), and mortality across cancer types worldwide in females (B). Adapted from *The Global Cancer Observatory: Cancer Today.*

Despite high incidence and mortality rates, according to American Cancer Society data for women diagnosed with breast cancer between 2013 and 2019, the 5-year survival rate when the cancer is localized is 99% and 86% for loco-regional breast cancer. However, the 5-year survival rate decreases to 31% for breast cancer with distal metastases [7].

## 1.2 ETIOLOGY AND RISK FACTORS

The etiology of breast cancer is attributed to a complex interaction between various modifiable and non-modifiable factors. This etiology is determined by genetic, environmental, hormonal and hereditary elements that contribute to the development of this disease.

### Non – modifiable risk factors

One of the most important non-modifiable risk factors for the development of breast cancer is gender. It is estimated that women are 100 times more likely to develop breast cancer than men[8], mainly due to

high estrogen and progesterone stimulation [9]. In women, earlier age at menarche and later age at menopause have been associated with an increased risk of breast cancer, attributed to increased lifetime exposure to endogenous estrogens[10]. Also, higher levels of estrogen in premenopausal and postmenopausal women have been associated with an increased risk of breast cancer [11].

Another relevant non-modifiable risk factor is age. The Surveillance, Epidemiology, and End Results (SEER) database shows that the probability of a woman developing breast cancer between the ages of 50 and 59 is 2.4%, 3.5% between the ages of 60 and 69, and 7.0% after the age of 70 [12]. Finally, genetic predisposition also plays an important role in the development of breast cancer. About 5-10% of all breast cancers are hereditary [13], and it is estimated that a woman with a first-degree relative with breast cancer is twice as likely to develop breast cancer as a woman with no family history of breast cancer [14]. Among the most important genes for the development of this disease are breast cancer susceptibility gene 1 (*BRCA1*) and breast cancer susceptibility gene 2 (*BRCA2*). 55%-72% of women who inherit a damaged *BRCA1* variant and 45%-69% of women who inherit a damaged *BRCA2* variant will develop breast cancer by the age of 70-80 years [15,16]. Among others, the DNA repair genes *ATM* and *CHEK2* and the tumor suppressor gene *TP53*, whose loss of function is associated with an increased overall risk of breast cancer, also play an important role in the development of breast cancer[17].

### Modifiable risk factors

There are lifestyle-associated risk factors that may cause genetic and epigenetic changes leading to the development of breast cancer, such as

physical activity[18], diet[19], obesity[20], alcohol consumption[21], smoking, nulliparity, breastfeeding, or the use of hormone replacement therapy[22].

Increased physical activity has been shown to be associated with a lower risk of breast cancer[18]. Similarly, in general, a diet high in saturated fats, processed foods and red meat has been associated with an increased risk of breast cancer, while a high intake of fruits and vegetables is linked to a lower risk[23]. However, there is contradictory data on this issue[23–25], so more studies are needed to reach a conclusion. On the other hand, a higher Body Mass Index (BMI) is associated with an acceleration in genetic ageing and an inflammatory profile that promotes tumor development, through different genetic alterations based on hormone receptor status[26]. Both moderate and heavy alcohol consumption is associated with BC possibly related to increased levels of estrogen in the blood[27]. In postmenopausal women, those who drank alcohol were found to have higher estrogen levels compared to those who did not[28]. Active smoking, especially before the first birth, is associated with a modest increase in the risk of BC[29].

Both multiparity and breastfeeding have a protective effect against breast cancer[30,31]. While the protective effect of parity has been seen especially in hormone receptor positive BC [32], the protective effect of breastfeeding is against hormone receptor negative breast cancers, which are more frequent in younger women and usually have a worse prognosis [30,33]. However, more studies are needed to corroborate the influence of hormone receptors as well as menopausal status. On the other hand, in postmenopausal women, combined estrogen and progesterone hormone replacement therapy has an increased risk of breast cancer when used long-term (> 10 years), but does not significantly increase the risk when

used for a short period of time (<= 4 years), although it does hinder mammographic detection of breast cancer[34].

## 1.3 HISTOPATHOLOGY AND MOLECULAR TYPES

Breast cancer comprises a complex and heterogeneous group of diverse tumors associated with different histological patterns and molecular characteristics [35].

### 1.3.1 Histopathology

Since the vast majority of breast tumors originate in epithelial cells, they are called adenocarcinomas (except less common tumors such as inflammatory breast cancer, sarcomas, lymphomas...). For the histological study of breast cancer it is important to identify tumor from ducts or lobules (ductal or lobular), and whether it is limited to the epithelia or has invaded the surrounding stroma (in situ or invasive) [36]. While invasive carcinomas constitute the 70-85%, carcinomas in situ represent about 15-30% of breast biopsies[37].

- **Ductal carcinoma in situ**: It accounts for approximately 83% of in situ breast cancer cases[38]. The neoplastic proliferation is originated and limited to the ducts. Depending on the degree of its evolution it is classified as low, medium or high grade. The tendency for invasive breast cancer to develop depends directly on the grade of the tumor. It can occur in five architectural subtypes: comedo, solid, cribriform, papillary and micropapillary. There is a minority of rare morphological variations, including neuroendocrine differentiation, apocrine metaplastic cells and squamous cell carcinoma in situ[39].

- **Lobular carcinoma in situ (LCIS):** It represents about 12% of in situ breast cancer diagnoses[38]. The cancer begins in the lobules of the breast, with proliferation of small, poorly cohesive cells, which can lead to invasive carcinoma in 25-35% of cases. Typical markers to differentiate LCIS are the absence of E- cadherin and β-catenin expression (whereas DCIS is positive for both) and positivity for high molecular weight keratin (HMW)[40].
- **Invasive ductal carcinoma:** Represents the 80% of invasive carcinomas[41]. Ductal neoplastic proliferation occurs together with stromal invasion, in the presence or absence of DCIS. IDC is further classified into different morphological subtypes based on cell type, number, location and type of secretion and immunohistochemical profile, among others[42]. However, 75% of CDIs do not have enough common features to be classified into specific morphological subtypes, and are referred to as non-special type (NST) CDIs. [43].
- **Invasive lobular carcinoma (ILC):** It constitutes 5-20% of invasive breast carcinomas and usually affects older women previously with IBC[41]. Like LCIS, they are small, round, poorly cohesive cells with a unilateral stromal infiltration pattern. They also have E-cadherin silenced by mutation, loss of heterozygosity or methylation [44]. The incidence of ILC is increasing over time, particularly in postmenopausal women, which has been associated with hormone replacement therapy [45].

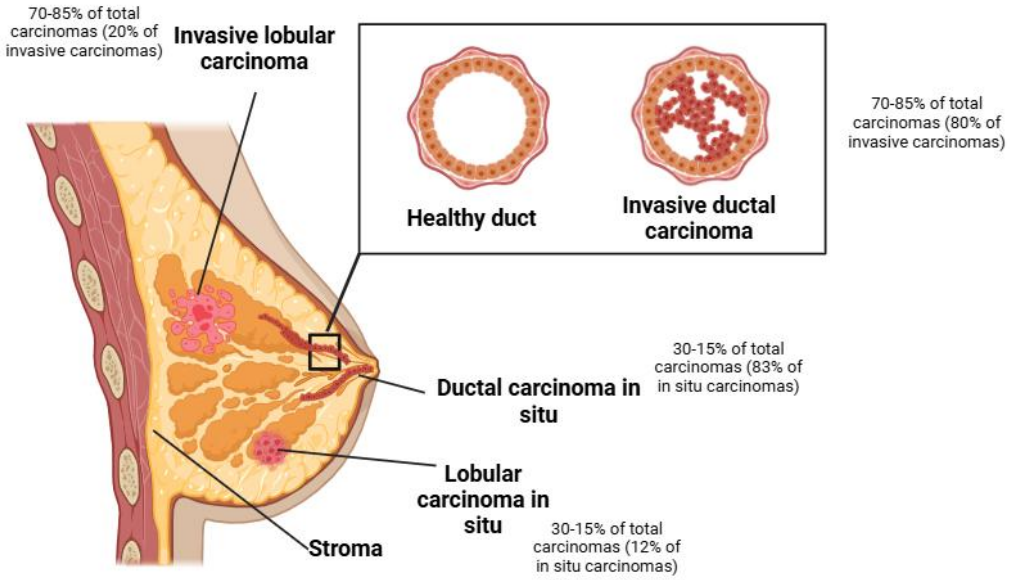



Figure 2. Main histopathological breast cancer types and their breast location versus a healthy duct. Created with Biorender.com

### 1.3.2 Molecular types

Breast cancer is a heterogeneous disease, and the classification of its multiple subtypes has evolved over time. Currently, the most common classification of breast cancer is based on the immunohistochemical expression of the following hormone receptors: 1. Estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor 2 (HER2) (Figure 3)[46]. This molecular classification helps determine which patients might benefit from targeted therapy, such as hormone therapy or anti-HER2 therapy [47]. According to the presence or absence of these receptors along with other clinical features, the tumor is classified into the following four subtypes:

- **Luminal A:** Luminal A tumors have ER and/or PR and lack HER2. They have low expression of the cell proliferation marker ki67 (< 20%) and low histological grade. They are the tumors with the lowest incidence of relapse (3,7%)[48], high survival rates and the best prognosis. Patients with luminal A breast cancer have a good response to hormonal therapy, but a more limited benefit from chemotherapy[49].
- **Luminal B:** They are ER-positive tumors and may have absent RP. They constitute 10-20% of luminal tumors. They have high ki67 expression (> 20%) and medium/high histological grade. They have a worse prognosis compared to luminal A tumors due to their rapid tumor growth[50], and benefit from hormone therapy in combination with chemotherapy[51].
- **HER2:** They are characterised by high HER2 expression and the absence of ER and PR. They account for 10-15% of all breast cancers. Within HER2 tumors, two groups can be distinguished: luminal HER2 (E+, PR+, HER2+ and Ki-67:15-30%) and HER2-enriched (HER2+, E-, PR-, Ki-67>30%)[52]. In general, they are faster growing than the luminal subtypes and are more aggressive, although their prognosis has greatly improved with the introduction of HER2-targeted therapies (trastuzumab, pertuzumab or tyrosine kinase inhibitors, among others) and they have a high response to chemotherapy[53].
- **Triple negative breast cancer (TNBC):** TNBCs do not have any hormone receptors (ER-, PR- and HER2-). They constitute 20% of all breast cancers, and 80% of tumors with mutated *BRCA1/BRCA2* belong to this group[54]. Most TNBC manifests as invasive ductal carcinoma of no special type [55]. It is a tumor

characterised by its aggressiveness, high rate of cell proliferation and alterations in DNA repair genes. Immunohistochemically, it is divided into basal (expression of cytokeratins 5/6 and human epidermal growth factor receptor type 1 (EGFR1)) and non-basal, with absence of cytokeratins 5/6[46].

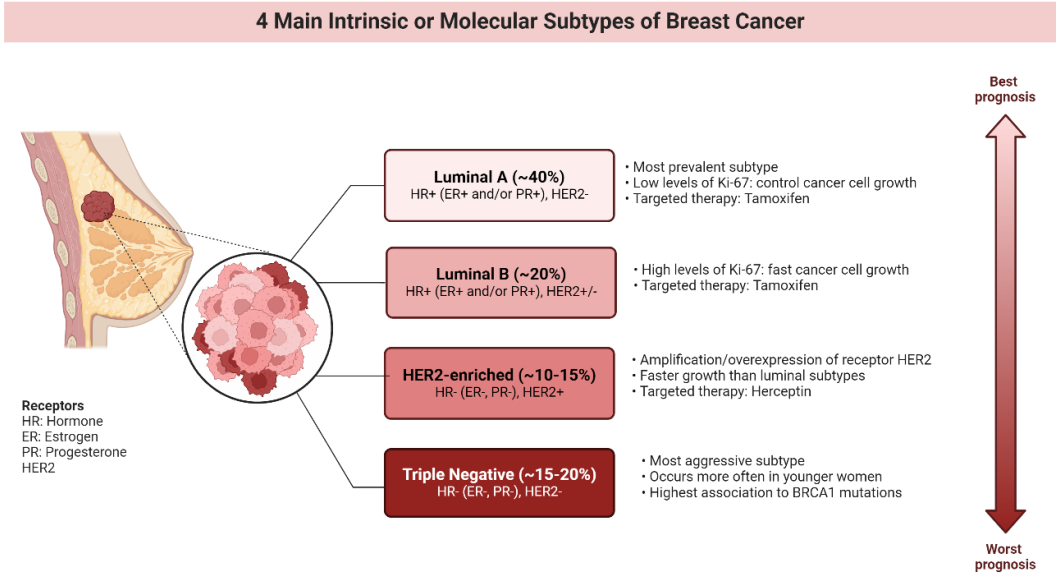



Figure 3. Molecular subtypes of breast cancer and their main characteristics. Adapted template from Biorender.com

## 1.4 CARCINOGENESIS

Breast carcinogenesis refers to the process by which normal breast cells transform into cancerous cells. As we have seen, this process is driven by a combination of genetic, hormonal and environmental factors. Although risk factors have been identified, breast carcinogenesis is not well understood. It cannot be considered as a fixed-step pattern of genetic progression, but as the result of the accumulation of several major and minor genetic events in a rather random order[56].

In general, sporadic breast carcinogenesis shares similarities with other epithelial malignancies. One of the earliest events in breast carcinogenesis is the inactivation of tumor suppressor genes through DNA methylation[57]. Additionally, loss of heterozygosity may occur early, leading to the inactivation of crucial genes (like *BRCA1*, *BRCA2*, *ATM*, and *CHEK2* in familiar cancers) which impacts DNA repair and allows the accumulation of further genetic changes[58].

Estrogen and progesterone receptors are highly expressed in nearly all pre-invasive breast lesions, stimulating cell proliferation[59]. Similarly, many growth factor receptors, such as HER2 and EGFR, are frequently overexpressed. Also, several cell-cycle control proteins play a crucial role in this process. Cyclin D1 may be amplified or induced in an ER-dependent manner, while cyclin E is often overexpressed. Cyclin-dependent kinases like CDK4 can also be overexpressed [60] and inactivating mutations in tumor suppressor genes like p53 promote the outgrowth of the tumor[61]. Disruptions in cell cycle control and apoptosis-signaling pathways create an imbalance between proliferation and apoptosis, leading to the growth of the tumor. Increased and uncontrolled proliferation generate hypoxia that aberrant angiogenesis, often forming a network of vessels around malignant ducts[62].

## 1.5 BREAST CANCER DISSEMINATION AND METASTASIS

In the advanced pre-invasive stages, invasion-related genes are activated (as *TNC*, *JAG2* or *EREG*), enabling cells to degrade the basement membrane and extracellular matrix, facilitating invasion into the surrounding stroma[63]. This invasion provides access to lymphatic and blood vessels, allowing cells to enter the lymphatics and bloodstream, leading to loco-regional and distant metastases. For distant metastasis,

circulating tumor cells (CTCs) must home to distant sites, adhere to the endothelium, invade local tissue, and establish an optimal microenvironment to escape dormancy and develop into clinically manifest metastases[64].

Metastasis is the complex process by which a primary tumor develops into a secondary tumor at a distant site. It is a defining characteristic of cancer and often results in treatment failure, causing the death of many patients. This multi-step process begins with local invasion, followed by intravasation, migration through blood or lymphatic vessels, and concludes with extravasation and colonization of distant organs[65]. It is characterized by epithelial– mesenchymal transition (EMT), epithelial cell polarity and cohesion[66]. Inflammatory cytokines, especially IL-6, play a vital role by activating the JAK/STAT3 pathway, which promotes EMT via estrogen receptor α (ERα). Also, immune cells and the tumor microenvironment (TME) play an important role in breast cancer metastasis. Particularly tumor-associated macrophages activated by IL-4 from CD4+ T cells secretes factors that enhance metastasis by increasing cell adhesion to the extracellular matrix. Nowadays, all these components are being studied as potential targets for treating breast cancer patients[67].

### 1.5.1 LOCO-REGIONAL METASTASIS

There is an intermediate stage between localized breast cancer and breast cancer with metastases to other organs known as regional breast cancer or breast cancer with lymph node metastasis (LNM). Clinicians distinguish this type of metastasis from distant metastasis during staging, which differentiates stage III from stage IV in breast cancer patients.

Biologically, here the tumor does not migrate to a distant organ but remains around the breast and colonizes the mammary lymph nodes. The 10-year survival rate for regional breast cancer is 60%[68], and approximately 27% of women with breast cancer present with LNM at diagnosis[69]. Additionally, only 20–30% of patients with LNM remain free of distant metastases in the future[70].

Early detection of LNM is crucial as it significantly impacts clinical management, treatment, and prognosis[71]. Radiologists play a key role in the preoperative diagnosis of abnormal lymph nodes using mammography, magnetic resonance imaging, ultrasonography, and other techniques[72]. However, imaging screening methods sometimes fall short in accurately staging patients. Currently, the best method for preoperative patient staging is Sentinel Lymph Node Biopsy (SLNB)[73]. Nevertheless, SLNB is limited to patients undergoing initial surgery or neo-adjuvant chemotherapy and is highly invasive. Moreover, SLNB has shown more than 10% false negatives in patients with LNM after preoperative systemic therapy[74].

### 1.5.2 DISTAL METASTASIS

In breast cancer, distal metastasis is responsible for nearly all deaths[75]. The English surgeon Stephen Paget proposed the “Seed and soil” hypothesis in 1889. It suggests that tumor cells are like seeds dispersed from a plant and only those landing on compatible and fertile soil will grow. According to this theory, several factors contribute to pre-metastatic niche and the metastasis: extrinsic factors, such as tumor-secreted elements via extracellular vesicles (EVs) and cytokines, modulate the extracellular matrix of both primary and secondary sites.

Also, intrinsic factors like the EMT and autophagy mechanisms, are crucial for CTCs survival and colonization[76] (Figure 4).

Thus, breast cancer exhibits distinct patterns of organ-specific metastasis (organotropisms) depending on its molecular subtype. For example, the HR+/HER2- subtype has a high risk of bone metastasis [77]. The HER2+ subtype is significantly associated with increased rates of liver metastases[78]. Meanwhile, triple-negative breast cancers are more likely to metastasize to the brain and lungs [79,80].

BC metastases can be classified as primary tumor relapse or as *de novo* metastasis, which present molecular and clinic-pathological differences:

- *De novo* metastasis breast cancer (dnMBC): encompasses a subpopulation of breast cancer patients who present metastasis or stage IV at diagnosis[81]. Approximately 6% of new breast cancers presents metastases at the time of diagnosis[82]. Clinically, de novo metastasis present increased hormone receptor-positive status, increased lymph node involvement and better survival outcomes relative to metastasis at relapse, perhaps due to the treatment clonal selection[83].De novo metastasis tumors are more likely to be HR+ and HER2+, increasing the therapeutic possibilities as endocrine therapy (tamoxifen/anastrazole) or epidermal growth factor-targeted therapy (trastuzumab and lapatinib). In terms of genome alterations, dnMBC had 4-fold higher *PTEN* mutations and poorer survival with *ABL2* and *GATA3* alterations than metastasis at relapse. In terms of expression, dnMBCs down-regulated TNFa signaling, IL-17, and chemotaxis, while they up-

regulated steroid biosynthesis, cell migration, and cell adhesion[83].

- Metastasis at relapse: refers to the return of cancer after a period of remission or successful treatment. 25-30% of patients develop disease recurrence and die from disease dissemination[84]. The risk of metastatic relapse depends on various factors, such as the stage of cancer at diagnosis, the aggressiveness of the cancer cells (highly related to the molecular characteristics), and the effectiveness of the initial treatment. Metastases at relapse BC are more likely to be basal in the molecular subtype, without chances of target therapy[85].

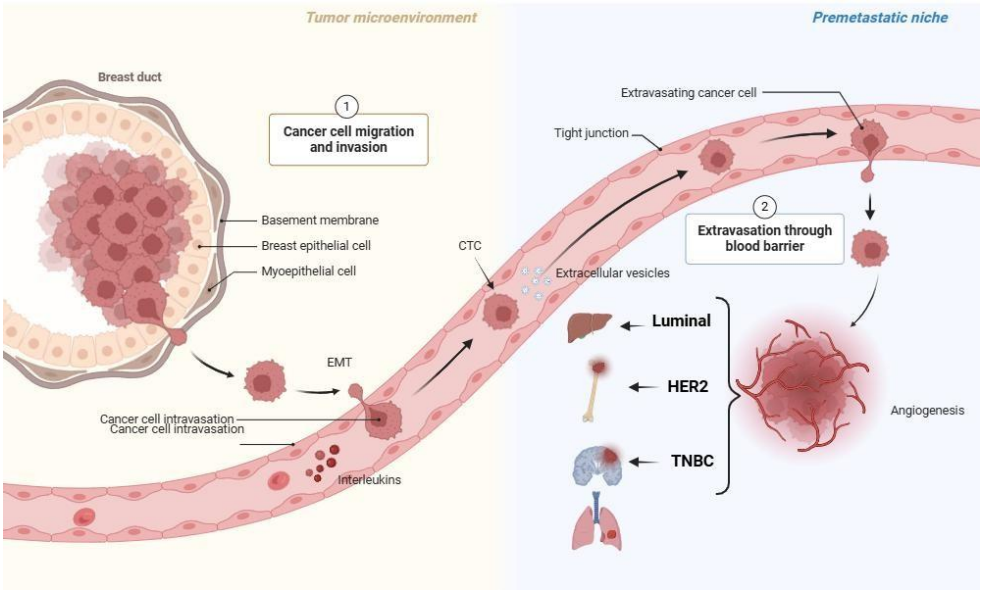



Figure 4. Cancer cell intravasation and migration through the blood flow to form a pre-metastatic niche in a distal organ. Adapted from Biorender.com

## 1.6 TUMORAL HETEROGENEITY

Cancer is a dynamic and constantly evolving process. From the initial genetic diversity during tumorigenesis to clonal selection and expansion, the tumor is modulated by the microenvironment, leading to tumor heterogeneity, which is one of the hallmarks of malignancy. Intertumor heterogeneity is observed in breast carcinomas across different individuals, while intratumor heterogeneity arises from the presence of diverse cell populations within a single tumor [86].

The cellular heterogeneity of breast cancer was acknowledged as far back as the nineteenth century[87]. However, its significance in clinical settings was only recognized 30 years ago with the advent of estrogen receptor (ER) testing[88]. Differences in ER expression between various tumors, and among different cell populations within the same tumor, were identified as reasons for variations in clinical behavior and treatment responses[89].

Several mechanisms have been described to cause tumor heterogeneity:

**Clinical and Histopathologic Heterogeneity:** Standard breast cancer treatment is tailored to the tumor's characteristics, such as clinical stage, histopathologic features and biomarker profile, which significantly impact patient survival and largely explain the differences in clinical outcomes among breast cancer patients [90].

Morphologic intratumor heterogeneity can appear as variability within different regions of a tumor (spatial heterogeneity) or as changes occurring over time (temporal heterogeneity)[86]. Spatial heterogeneity is commonly seen in surgical pathology practices within a single tumor and can also be observed between a primary breast carcinoma and

synchronous lymph node metastases, as well as among synchronous metastases from different sites. On the other side, temporal heterogeneity includes the tumor's evolution over time or in response to treatment, the emergence of asynchronous metastatic disease, and the transition from in situ to invasive carcinoma [91,92].

**Biomarker Heterogeneity**: Although 70–80% of ER-positive breast cancers also express PR (ER+/PR+), there are cases where tumors are ER+/PR− or, less commonly, ER−/PR+. Depend on that, hormonal treatment responses can vary, with the highest response rate (around 60%) observed in ER+/PR+ tumors, and lower rates in ER+/PR− and ER−/PR+ tumors[93]. The HER2 oncoprotein is overexpressed in approximately 15–20% of primary breast carcinomas[94]. Biomarker expression within a single tumor can be highly variable, causing challenges in interpretation and leading to inconsistent results in small biopsies. The percentage of ER/PR-expressing cells in individual tumors ranges from 1% to 100%, with higher expression levels correlating with better response to endocrine therapy[95]. Nonetheless, this approach does not fully account for intratumor heterogeneity, which can diminish the clinical relevance of classifying tumors with uneven ER expression as ER-positive. Likewise, HER2 gene amplification can be very heterogeneous, affecting disease-free survival[96].

**Genetic Heterogeneity:** While gene expression patterns hold promise for predicting chemotherapy response and recurrence risk, the classification of breast cancer based on gene expression is hindered by both clinical and molecular heterogeneity. Even among patients with the same molecular subtype receiving identical treatments, outcomes can differ, and therapy resistance may emerge[97].

Breast cancer can show significant intratumor heterogeneity in terms of chromosomal and genomic changes[98,99], impacting various biological processes[100]. Within tumors, different cell clones may either segregate into distinct regions or intermingle within the same area[101]. Additionally, DNA methylation can influence the activity of tumor suppressor genes, as well as the expression of ER/PR/HER2 receptors[102].

There are different mechanisms that could explain breast cancer heterogeneity, like the differentiation state of the cell of origin, within the resulting tumor phenotype may not consistently represent the original cell type [103]. Other mechanisms that could ilustrate it are the cell plasticity, principally enabled by cancer stem cells[104], the influence of the tumor microenvironment[105], and the genetic evolution of the tumor, driven by Darwinian selection of the most adaptable cells and genetic instability[106]. Furthermore, treatment process can accelerate the clonal evolution of the tumor, selecting cells with mutations that confer growth and resistance advantages[107]. This heterogeneity and complexity of breast cancer highlights the need to evolve towards personalised and precision medicine, with therapies tailored to each patient.

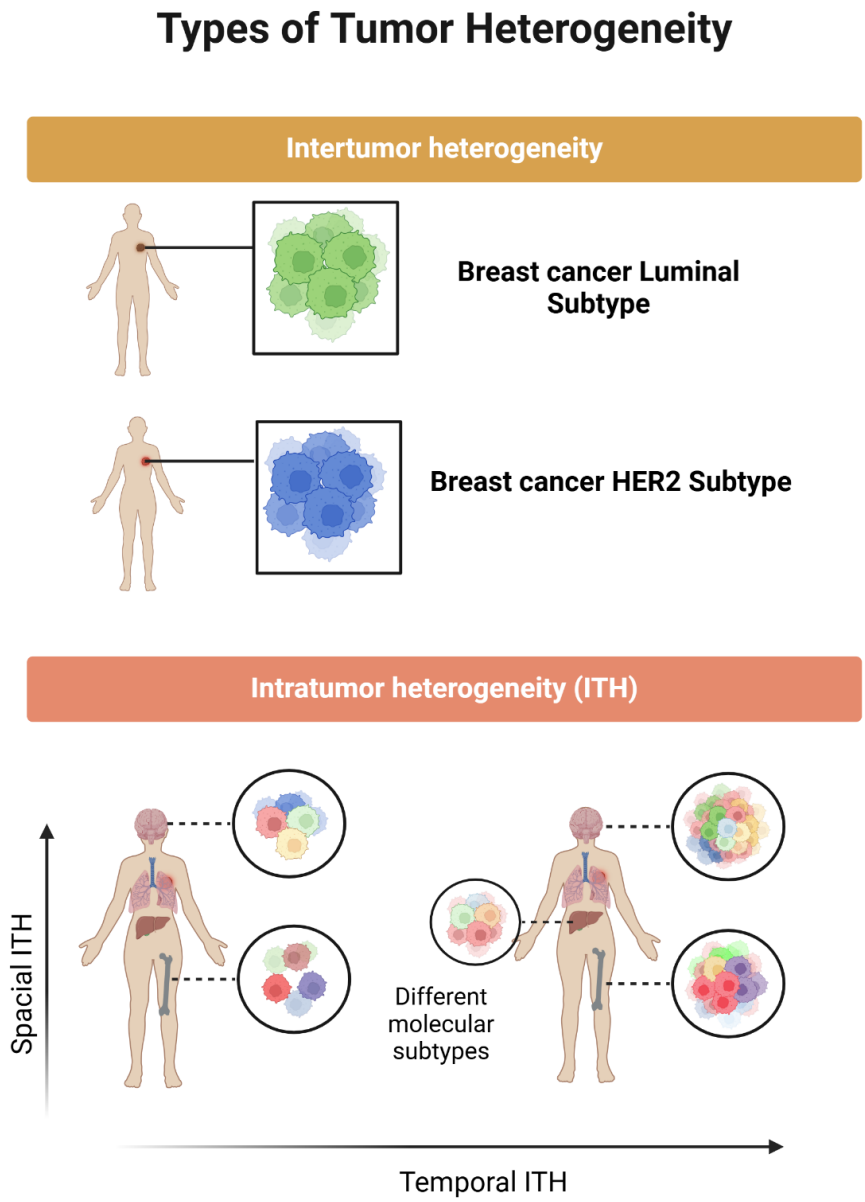



Figure 5. Tumor spatial and temporal heterogeneity, inter and intratumoral. Inspired in Gilson et al., [108] and created with Biorender.com

## 1.7 CLINICAL MANAGEMENT OF BREAST CANCER

Early diagnosis of breast cancer is key to effective treatment and a positive prognosis. Patients with tumors classified as T1 at the moment of diagnosis have a 10-year survival rate of approximately 85%, whereas T3 tumors, often due to delayed diagnosis, have a 10-year survival rate of less than 60%[109]. Currently, the two main focuses for effective breast cancer management are: (i) diagnosing breast cancer at its earliest stages and (ii) providing timely treatment after diagnosis[110].

### 1.7.1 SCREENING AND DIAGNOSIS

Current screening and detection methods in clinical practice include Breast Imaging Modalities (such as mammography, ultrasonography and magnetic resonance imaging, among others) and Breast Physical Examination. After the detection of an anomaly by imaging techniques, biopsies are required for the breast cancer diagnosis[111].

Between imaging techniques, mammography is the current gold standard for BC screening. A significant drawback of traditional mammography is over-diagnosis, where harmless tissue irregularities or nonaggressive tumors are mistakenly identified, resulting in unwarranted procedures and treatments [112]. Furthermore, the sensitivity of mammograms varies with age and breast tissue density [113]. Ultrasonography is utilized in breast cancer diagnosis to differentiate between fluid-filled cysts and solid tumors [109,114]. However, ultrasonography has a low detection rate for calcifications, lower specificity compared to mammography and requires a highly trained technician to perform the test [115]. On the other hand, Magnetic Resonance Imaging (MRI) can detect breast cancers with an increased sensitivity of 58% in high-risk women compared to

mammography alone [116], though it has high rates of false positives and significant examination costs. Moreover, MRI is contraindicated in patients with implantable electronic devices (e.g. peacemakers) or tissue expanders[117,118]. Lastly, there is Positron Emission Tomography/Computed Tomography (PET/CT), which the main advantage is its combination of anatomical and functional imaging, but the National Comprehensive Cancer Network 2020 guidelines do not suggest PET/CT scanning in patients with clinical stage I or II and operable stage III breast cancer [119].

When an anomaly is detected in breast tissue through imaging techniques, biopsies are necessary for an accurate diagnosis. Breast biopsy is an invasive procedure which involves removing abnormal breast fluid or tissue for cytological, histological, and molecular analysis. It is recommended only in suspected cancer cases based on the BI-RADS lexicon scale used by radiologists[120]. Despite its invasiveness and unsuitability for cancer patients, breast biopsy remains the gold standard for confirming whether a tumor is benign or malignant [121].

### 1.7.2 GRADING AND STAGING

After BC diagnosis, grading and staging of the tumor are performed in order to apply the most appropriate treatment for the patient.

Grading assesses the appearance of cancer cells compared to healthy cells and predicts their growth and spread rate. Breast tumors can be categorized as low-grade, with uniform, slow-growing cells; intermediate-grade, with larger, variably shaped cells that grow faster than normal; and high-grade, with rapidly growing cells of diverse sizes

and shapes. Higher grade breast cancers are generally more aggressive[111].

For staging, the American Joint Committee on Cancer (AJCC) Cancer Staging Manual uses the Tumor, Node, and Metastasis (TNM) system to describe the size of the tumor (T), the status of regional the lymph nodes (N) and distal metastasis status (M), respectively. In the AJCC Cancer Staging Manual 8th Edition [122] stage 0 represents DCIS or Paget disease without nodal involvement or distant metastasis. Stage I is divided into IA and IB. IA includes tumors ≤20mm in size. IB includes cases with no evidence of the primary tumor or a tumor size of ≤20mm, with micrometastasis (nodal involvement) of 0.2-2mm. Stage II is also split into IIA and IIB. IIA involves tumors ≤20mm (or without evidence of the primary tumor) with ipsilateral level I or II axillary nodes, and tumors 20-50mm without lymph node involvement. IIB encompasses tumors 20-50mm with level I or II axillary lymph node involvement, or tumors >50mm without lymph node metastases. Stage III is subdivided into IIIA, IIIB, and IIIC. IIIA includes a range of tumor sizes from no evidence of the primary tumor to tumors >50mm. IIIB comprises tumors of any size with direct extension to the chest wall or skin invasion by malignant cells, potentially with the same lymph node involvement as IIIA. IIIC includes tumors of any size with involvement of ≥10 lymph nodes. Stage IV includes any tumor size and lymph node involvement but requires evidence of distant metastasis (Table 1) [111,122,123].

| Stage | Tumor | Node | Metastasis |
|---|---|---|---|
| 0 | Tis | N0 | M0 |
| IA | T1 | N0 | M0 |
| IB | T0 | N1mi | M0 |
| | T1 | N1mi | M0 |
| IIA | T0 | N1 | M0 |
| | T1 | N1 | M0 |
| | T2 | N0 | M0 |
| IIB | T2 | N1 | M0 |
| | T3 | N0 | M0 |
| IIIA | T0 | N2 | M0 |
| | T1 | N2 | M0 |
| | T2 | N2 | M0 |
| | T3 | N1 | M0 |
| | T3 | N2 | M0 |
| IIIB | T4 | N0 | M0 |
| | T4 | N1 | M0 |
| | T4 | N2 | M0 |
| IIIC | AnyT | N3 | M0 |
| IV | AnyT | AnyN | M1 |

Tis = *in situ*, mi = micrometasis

Table 1. TNM. Breast Cancer Staging AJCC Cancer Staging Manual 8th Edition[122].

### 1.7.3 GENETIC TESTS

Gene Expression Profiling/Signatures (GEP/S) is a valuable tool for BC prognosis and management, as it identifies aggressiveness differences among tumors with the same anatomical staging, immunohistochemical (IHC) markers or genetic predisposition[124]. For instance, GEP/S can aid in therapeutic decision-making for low-grade breast cancer that may become aggressive and resistant to chemotherapy[125]. Most of the genes used for GEP/S are different, making each signature unique and not interchangeable. Here we describe two of the most clinically used: MammaPrint® and Oncotype DX®.

**MammaPrint®** evaluates the risk of breast cancer recurrence in patients with an early BC, ER+ and LN+ (1-3 nodes). It analyzes the activity of 70 genes within a breast cancer tumor to predict the likelihood of cancer returning after treatment. This test helps in guiding treatment decisions, particularly in determining whether chemotherapy would be beneficial for patients with early-stage breast cancer[126].

**Oncotype DX®** assesses the risk of breast cancer recurrence and to help guide treatment decisions, particularly for patients with early-stage ER+, HER2- and LN- or LN+ BC. The test analyzes the expression of 21 genes in the tumor sample to produce a Recurrence Score® between 0 and 100. This score indicates the likelihood of cancer returning within 10 years of the initial diagnosis and helps determine the potential benefit of chemotherapy in addition to hormone therapy[127].

The AJCC has introduced "Prognostic Staging" in its latest Cancer Staging Manual, combining TNM staging with tumor grade, receptor status, and genomic tests for a more accurate prognosis. For example, a tumor previously staged as IIIa could be reclassified as Ib if the patient has a low Oncotype DX® score (<11)[124]. This update allows for a more comprehensive understanding of prognosis and enables more tailored treatment decisions.

### 1.7.4 TREATMENT

Ideally, for BC treatment, tumor grading and staging is combined with hormone receptor status and genomic testing, allowing a more complete cancer management. Specific treatments encompass surgical resection (when possible), neo-adjuvant and adjuvant treatments, such a chemotherapy, endocrine therapy, radiotherapy, targeted therapies and immunotherapy.

## SURGERY

Breast-conserving surgery (BCS) and mastectomy, whether or not followed by immediate reconstruction, are established methods for managing early invasive breast cancer. In cases of non-metastatic breast cancer, the primary treatment is surgical[128]. However, for metastatic breast cancer, systemic therapy is preferred, with surgery reserved for palliative purposes[129]

Axillary management depends on the status of the axillary lymph nodes (ALNs) at diagnosis and the use of neo-adjuvant therapy. Generally, all newly diagnosed invasive breast cancer patients with a clinically negative axilla, except older patients and those with significant comorbidities, should have axillary staging via SLNB [130]. Completion axillary lymph node dissection (ALND) is necessary for patients with three or more positive sentinel lymph nodes or those with matted nodes detected during surgery[131]. For patients with clinically node-negative (cN0) breast cancer, the approach to axillary management is debated. SLNB may be adequate for most cN0 patients, while additional axillary radiation is suggested only for specific patients, such as those with three involved sentinel lymph nodes[128,131,132].

## NEO-ADJUVANT AND ADJUVANT CHEMOTERAPHY

Neo-adjuvant chemotherapy (NAC) was originally employed to make locally advanced, inoperable breast cancer suitable for surgery. More recently, NAC has been applied to operable tumors to shrink the disease in the breast and axilla, thereby facilitating breast-conserving surgery and sometimes eliminating the need for ALND[130]. NAC recommended for patients with a large tumor, and aggressive molecular subtypes, such as triple-negative and HER2[130], and has been demonstrated that it

significantly decreases the incidence of axillary metastases in women who are clinically node-negative. Also, the oncologic safety of NAC has been tested [133,134].

Adjuvant chemotherapy (AC) administration has been shown to reduce the risk of recurrence by approximately 30% in early breast cancer cases [135]. Multigene assays and molecular typing tools, as Oncotype Dx, are valuable implements for identifying patients who would benefit most from chemotherapy, particularly in cases of node-negative ER-positive, HER2-positive or TNBC diseases[136,137]. The current preferred NAC and AC regimen involves taxane with or without anthracycline, administered sequentially or in combination[138]. While the use of anthracyclines remains contentious, it appears to be crucial in high-risk patients, such as those with triple-negative and HER2-positive subtypes[139].

### RADIOTHERAPY

Radiotherapy (RT) is the standard of care for patients undergoing breast cancer surgery[140,141] or those with unresectable tumors and metastatic tumors as a symptom palliative[142]. For patients receiving a mastectomy, the decision to administer radiation is typically based on the number of involved ALNs, particularly in cases where four or more ALNs are affected[128].

For patients with one to three ALNs, recent studies indicate that there is no additional survival benefit in the context of systemic treatments[143]. Therefore, it is crucial to identify high-risk patients for postmastectomy radiation therapy, such as those of younger age, and with a higher tumor burden in the breast and axilla[128]. The last edition of AJCC pathological prognostic staging system incorporates molecular markers (ER, PR, HER2 status and tumor grade) to guide RT decisions for patients with

N1 breast cancer[144].

### ENDOCRINE THERAPY

Endocrine therapy is considered the standard adjuvant treatment for patients with hormone receptor-positive BC for a duration of 5–10 years. Its effectiveness directly linked to the expression of hormone receptors[145]. For patients at high risk of relapse, ovarian suppression drugs (as triptorelin) combined with either tamoxifen or an aromatase inhibitor (AI) have shown improved disease-free survival, though they come with higher toxicity compared to tamoxifen alone[146]. An example is the MA.17 trial which demonstrated that after 5 years of tamoxifen, and additional 5 years of AIs can reduce the relative risk of recurrence by 40%, and a 34% reduction in recurrence with 10 years of AIs[147].

Tamoxifen is used as an estrogen receptor antagonist for both premenopausal and postmenopausal women. AIs are only used in postmenopausal women, and are generally preferred over tamoxifen, but can also be used sequentially with tamoxifen[148].

### TARGETED THERAPY

The treatment paradigm and prognosis of HER2 BC changed with the discovery of anti-HER2 targeted therapy. The first anti-HER2 targeted drug discovered was trastuzumab, which has been widely used for HER2 BC diseases[128]. Currently, patients with stage I HER2 BC typically receive a combination of paclitaxel and trastuzumab. Until the approval of pertuzumab (HER2 dimerization inhibitor) by the United States Food

and Drug Administration in 2013, patients with stage II-III HER2 BC were treated with regimens that included trastuzumab added to either anthracycline - taxane[149]. Recent studies have shown that adding pertuzumab, to trastuzumab in the neo-adjuvant setting improves the pathologic complete response rate. Using dual-HER2 agents in the neo-adjuvant setting has now become the standard of care for patients with stage II-III HER2-positive breast cancer[150,151].

### IMMUNOTHERAPY

BC hinders the ability of activated T cells to fight tumor cells due to the realize of inhibitory factors and the interactions between PD-1, LAG-3, TIGIT, CTLA-4, and their ligands. This activity leads to T-cell exhaustion, decreasing their anti-tumoral activity and therapy efficiency. As a result, using immune checkpoint blockade as an anti-tumor treatment has shown limited effectiveness as a single-agent therapy in advanced breast cancer[152]. Although the results of a Phase 3 trial encouraging for the use of neo-adjuvant immune checkpoint inhibitors therapy with atezolizumab (against PD-L1) plus chemotherapy, longer follow-up is needed to confirm the long-term efficacy[153].

### GENE THERAPY

Gene therapy consists in delivering genetic material into target cells via a vector, in order to edit genes and alter their product, with the goal of treating cancers. For this, many strategies are employed, like DNA or RNA vaccination, targeting transcription factors, gene editing, microRNAs, etc[154].

A Phase I clinical trial assessed the efficacy and safety of genetic prodrug activation therapy targeting the human HER-2 gene promoter[155].

Besides, the use of microRNA in anti-cancer therapy has also shown promising results in inhibiting BC proliferation and development. Phase I of MRX34 is one of the first on miRNA replacement agent (miR-34a) and is now entering clinical trials[156].

## 1.8 DIFFICULTIES IN BREAST CANCER MANAGEMENT

The primary challenge in breast cancer is achieving early diagnosis, as it can improve the 5-year survival rate to 95%[109]. Mammograms and other imaging techniques, while useful, are not entirely accurate, necessitating a solid biopsy for confirmation[120]. This invasive procedure is often not suitable for breast cancer patients. Early detection of LNM is equally critical. Only 20-30% of LNM-positive patients remain free of distant metastases over time, significantly impacting clinical management, treatment, and prognosis[73]. Currently, SLNB is the most effective method for preoperative staging. However, SLNB is restricted to patients undergoing initial surgery or neo-adjuvant chemotherapy, is invasive, and has a notable false negative rate (over 10%) in patients with LNM after preoperative systemic therapy[74].

In the clinical practice, there are currentlykey biomarkers for early detection and prognostication in breast cancer including Ki67, hormone receptors (ER and PR), and HER2. There are also circulating tumor biomarkers as tumor antigens (CEA, CA15-3, CA-125, etc)[157].These biomarkers facilitate the molecular classification of tumors, aiding in management and treatment decisions. However, Ki67 expression can vary among different ethnic groups, requiring careful interpretation in the context of racial and ethnic cancer heterogeneity[158]. CEA and CA15-3 biomarkers have several limitations in the early diagnosis, as low sensitivity, poor specificity and limited utility in localized disease[159].

Additionally, intratumoral heterogeneity with hormonal receptors and HER2 receptor remains a challenge across different molecular subtypes[160,161].

## 2. LIQUID BIOPSY

Tissue biopsy is the gold standard in breast cancer diagnosis, together with image screening techniques. But solid tumor biopsy has many drawbacks as the derived from tumor heterogeneity, the high challenge to detect minimal residual disease (MRD) or its invasive nature, which often require surgical procedures to extract tissue samples and prevent patient follow-up, being potentially harmful to the health of certain patients[162].

In this scenario, liquid biopsy arises as a minimally invasive tool that can reduce bias due to the spatial heterogeneity of tumors, allows for real-time monitoring of cancer progression and don't require specialized equipment, which make easier to integrate liquid biopsy into routine clinical practice[163,164]. Liquid biopsy can be defined as a non – invasive tool that analyze molecular components in different body fluids, mainly blood. The implementation of liquid biopsy in clinical practice allows for real-time monitoring of cancer progression and treatment response. Also, it can provide a more comprehensive picture of the tumor's genetic landscape, that holds the promise for BC early detection and monitoring for recurrence[165].

There are many different components that are being study as liquid biopsy tool (cell free DNA, CTCs, extracellular vesicles, proteins…). In this thesis, because of their importance in the regulation of the tumor genetic landscape, we decided to focus on microRNAs

relevance in BC diagnosis and metastasis and we study the clinical relevance of them when transported by extracellular vesicles.

## 2.1 EXTRACELLULAR VESICLES

EVs are small membrane-derived particles secreted by all cell types, which can be classified according to their mechanism of biogenesis (e.g., exosomes, microvesicles, and apoptotic bodies) and size (e.g., small EVs). and large EVs) [166]. The primary subgroup consists of vesicles ranging from 150 nm to 1-2 μm in size, commonly called microvesicles (MVs), while the second subgroup includes vesicles that are 20 to 150 nm in diameter, known as exosomes. However, their classification is a topic of continuous debate among the scientific community of EVs, due to their complexity and heterogeneity in isolation and characterization techniques [167].

Initially, EVs were thought to be merely a way for cells to discard waste and unwanted substances or to be remnants of apoptotic cells. Over the years, tumor-derived EVs have become increasingly important in cancer development and tumor progression [167,168]. This is due to its important role in intercellular communication, transmitting its cargo from a sending cell to a receiving cell. In pathophysiological conditions, this cargo formed by proteins, DNA, microRNAs... is altered, being a faithful reflection of the disease and can compromise the functionality of the recipient cell[169]. Furthermore, it has also been seen that they play an important role in metastasis, favouring pre-metastatic niches where disseminated tumor cells grow, as well as the appearance of resistance to therapies [170,171].

### 2.1.1 BIOGENESIS AND RELEASE

MVs formation begin with the outward budding of the plasma membrane, while exosomes are derived from the endosomal pathway that starts with the formation of early endosomes. As these early endosomes mature, they transform into multivesicular bodies (MVBs), which are filled with intraluminal vesicles (ILVs) created by the inward budding of the endosomal membrane[172]. Within these MVBs, a meticulous sorting process occurs, directing specific proteins, lipids, and nucleic acids into the ILVs. This sorting can be mediated by the Endosomal Sorting Complex Required for Transport (ESCRT) machinery, a group of proteins adept at directing traffic within the cell. Alternatively, this can happen through ESCRT-independent mechanisms involving tetraspanins, lipids, and other sorting proteins. Once the MVB is packed with ILVs, it has two potential fates: it can fuse with lysosomes, leading to the degradation of its contents, or it can fuse with the plasma membrane, releasing the ILVs as exosomes into the extracellular environment[172] (Figure 6).

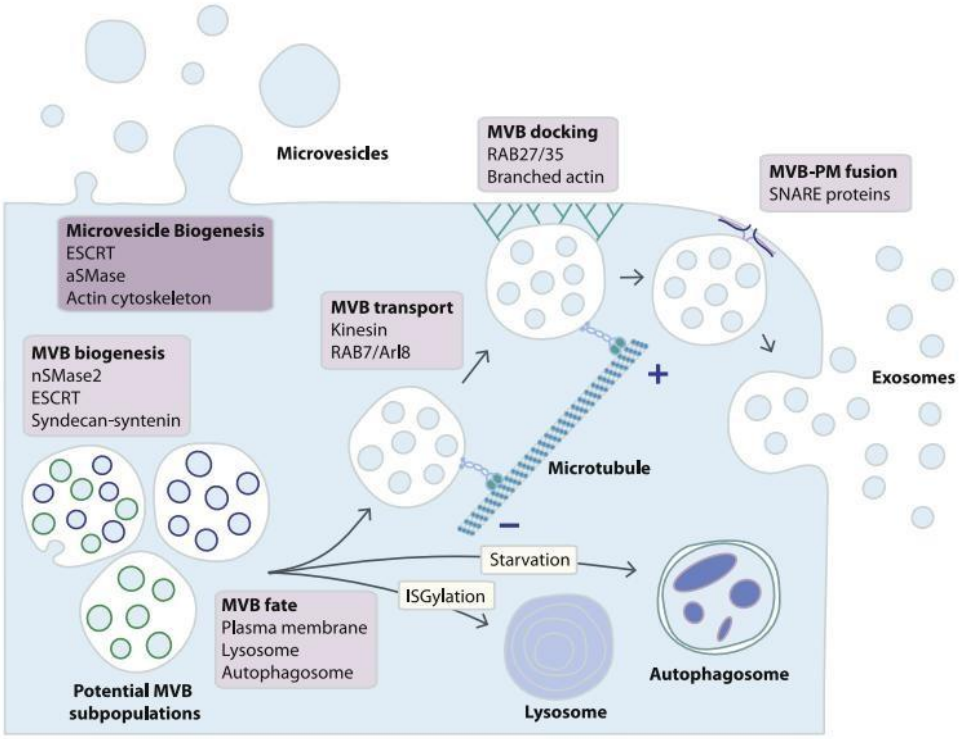



Figure 6. Biogenesis and function of extracellular vesicles in cancer[172].

When the EV reaches its target cell, the contents of the vesicle are released into the cytoplasm of the target cell by fusion of its membrane with the cell plasma membrane[173]. However, recently, the existence of a corona surrounding the EV has been discovered, which composition is influenced by the biofluid where the vesicle is located[174]. In blood, the EV corona is formed by coagulation factors, immunoglobulins, cytosines, complement proteins, enzymes, DNA and RNA[175]. Likewise, it has been seen that this corona and its composition influence the tissue biodistribution of the EVs in the body, and it is suspected that it may have an important role in the biology of cancer and its progression[176].

In breast cancer cells, EV biogenesis is often up-regulated, leading to increased release of EVs. This upregulation can be driven by oncogenic signaling pathways and environmental factors such as hypoxia. The EVs

from cancer cells carry distinct molecular cargo that can influence tumor progression, metastasis, and immune evasion.

### 2.1.2 MOLECULAR COMPOSITION AND CARGO

EVs are composed of proteins, lipids, and nucleic acids that typically mirror the originating cell's contents. Frequently detected proteins include Alix, Tsg101, and Hsp70, which are linked to the endosomal pathway and EV formation; tetraspanins like CD9 or CD63; and proteins involved in antigen presentation, such as Major Histocompatibility Complex Class I (MHCI) and MHCII[177].

The functional properties of EVs in the tumor microenvironment are largely determined by their cargo and the dynamics of their release and uptake. Proteins are incorporated into EVs by interacting with the EV biogenesis machinery. Membrane proteins, particularly those associated with tetraspanins, are sorted into EVs either through direct interaction or by entrapment in tetraspanin-enriched microdomains. Tumor exosomes carry mediators of tumorigenesis, such as oncoproteins or growth factors[172].

In 2006, it was first proposed that RNA could be transferred horizontally between donor and recipient cells through EVs. EVs can also carry nucleic acids, especially small RNAs like microRNAs, which are found in high concentrations within EVs. MicroRNAs as cargo in EVs play a pivotal role in regulating gene expression and mediating intercellular communication. Their selective packaging into EVs, impact on recipient cells, and potential as therapeutic and diagnostic tools highlight their importance in both normal physiology and disease states, particularly cancer[169].

### 2.1.3 EV CELLULAR TARGETING

Tumor-derived EVs are released into the blood, where they mix with EVs released by other healthy tissues and EVs from immune cells[178]. The process by which EVs are taken up is not fully understood yet. EVs can be adhered and fused to the cell surface or be engulfed into endosomes. In both scenarios, EVs release their contents into the cell cytoplasm through membrane fusion. The specificity of exosomes to distinct cell types is influenced by proteins enriched in the exosome membrane and the target cell's plasma membrane, such as tetraspanins, proteoglycans or integrins, through mutual recognition[179].

Therefore, studies that analyze the cargo of EVs in the plasma of cancer patients as potential biomarkers face the difficulty of demonstrating that their EVs are actually derived from the tumor and not from other healthy cells [180].

Today, to elucidate the tumoral origin of the EVs is a technical challenge that has been attempted to be addressed from different approaches. This is the case of proteomic[181] which has only been tested in EVs derived from tumor tissue cells, with tissue markers that may not be present in EVs [180] or nano-flow cytometry, suitable for lower throughput assays [182]. Another approach is the positive selection of EpCAM+ EVs. EpCAM is an epithelial biomarker absent in immune cells. While not all EpCAM+ EVs originate from the tumor, they are highly likely to do so. A further current problem is to find out the origin of EVs derived from CTCs, which has remained unsolved to date.

### 2.1.4 ISOLATION AND CHARACTERIZATION

EVs isolation is a critical step in understanding their functions and unlocking their potential applications. Among the various techniques available, ultracentrifugation is the most used, employed in 81% of the studies which employ EV isolation. This method separates EVs by their density or precipitation´s coefficient through a series of centrifugation steps. Another effective approach is size exclusion chromatography (SEC), which sorts EVs based on size using porous beads. Precipitation methods utilize agents like polyethylene glycol to draw EVs out of solution. Immuno-affinity capture leverages antibodies that target specific surface markers on EVs, enabling precise isolation. Additionally, microfluidic technologies, which use miniaturized devices, isolate EVs based on size, charge, and other physical properties. Lastly, for a quick and straightforward isolation, commercial kits offer a convenient solution. Each isolation method has a different EV recovery efficiency and purity. Each method has pro and contras, depending on what the EVs are needed for, one technology or another will be appropriate (Table 2).

| Isolation Method | Isolation Principle | Advantages/Limitations |
|---|---|---|
| Differential centrifugation | EV separation based on particle density, size and shape | - Commonly used; standardized; vesicle enrichment as pellet; EV subtypes isolation by density gradient centrifugation<br>- Vesicle aggregation; protein and soluble factors contamination; low recovery; laborious |
| Polymer-based precipitation | EV precipitation using polymers altering solubility | - Easy and inexpensive; high yield; effective with small amount of starting material; preservation of bioactivity<br>- Co-precipitation of protein contaminants and polymeric materials; not suitable for large scale studies; long incubation times |
| Size-exclusion chromatography (SEC) | EV isolation by gel filtration chromatography based on size | - Inexpensive; reproducible; high yield and purity; preservation of integrity and activity.<br>- Specific equipment; not suitable for large scale studies; long run times. |
| Immunoaffinity capture-based techniques | EV immuno-purification using magnetic beads conjugated with antibodies direct toward specific EV surface markers | - Sensitivity; specificity; high purity; EV subtypes isolation.<br>- Expensive; antibody cross-reactivity; low yield |

Table 2. Different methods for EV isolation and their advantages/limitations[183].

Once isolated, EVs must be thoroughly characterized to ensure their effective use in research and therapeutic applications. Several methods are employed for this purpose. Nanoparticle tracking analysis (NTA) measures the size and concentration of EVs, while various microscopy techniques, as transmission electron microscopy (TEM), provide detailed visual insights (Figure 7).

EVs hold significant promise as biomarkers for non-invasive disease diagnosis, since their presence and composition in different bodily fluids can provide valuable diagnostic information.

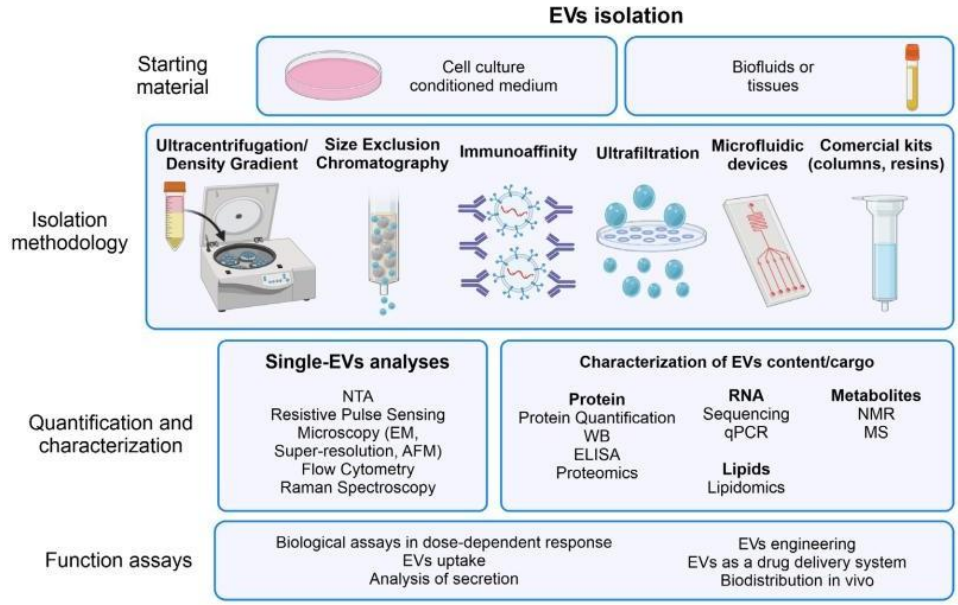



Figure 7. Isolation and characterization of EVs, highlighting the most commonly used methods at each stage[184].

## 2.2 MICRORNAs

MicroRNAs (miRNAs) are small non-coding RNAs of approximately 22-25 nucleotides in length that play a crucial role in regulating gene expression. They are predicted to regulate 60% of all human genes by binding to specific target sites, demonstrating both oncogenic and tumor-suppressive functions[185].

The involvement of miRNAs in human cancer was first highlighted by studies on B-cell chronic lymphocytic leukemia. Researchers discovered that the frequently deleted chromosome region 13q14 in these cancer cells contained two miRNA genes, miR-15a and miR-16-1. These miRNAs act as tumor suppressors by inducing apoptosis through the repression of the anti-apoptotic protein Bcl-2[186].

The dysregulation of miRNAs in cancer can occur through various mechanisms, including chromosomal abnormalities, transcriptional control changes, epigenetic modifications, and defects in the miRNA biogenesis machinery. miRNAs influence the hallmarks of cancer, such as sustaining proliferative signaling, evading growth suppressors, resisting cell death, activating invasion and metastasis, and inducing angiogenesis.

The potential of miRNAs as biomarkers for cancer diagnosis and prognosis has generated significant interest. miRNA profiling and deep sequencing have provided direct evidence of their dysregulation in various cancers, and their expression signatures can be used for tumor classification.

### 2.2.1 MICRORNAs BIOGENESIS

In humans, miRNAs are transcribed as large primary transcripts (pri-miRNAs) and processed into precursor miRNAs (pre-miRNAs) before being further cleaved into mature miRNAs. First, miRNAs undergo initial processing in the nucleus by RNA polymerase II, producing pri-miRNA transcripts that feature a 5' 7-methylguanosine cap and a 3' polyadenine tail. These pri-miRNAs are then processed by the enzyme Drosha, an RNase III family member, which in complex with the double-stranded RNA-binding protein DiGeorge critical region 8 (DGCR8), cleaves the hairpin to produce pre-miRNAs, with – 65 nucleotides of length. Pre-miRNAs are subsequently transported to the cytoplasm by exportin-5 and Ran-GPT, where another RNase III enzyme, Dicer, takes over. Dicer trims the pre-miRNAs into ≈22 nucleotide double-stranded miRNA[187,188].

I one hand, we have the guide strand, which is incorporated into the RNA-induced silencing complex (RISC) with Argonaute (Ago) proteins which guides the miRNA to its target mRNA. This process results in either translational repression or mRNA degradation, depending on the complementarity between the miRNA and the target mRNA. On the other hand, we have the passenger strand, which is typically degraded. The selection of the guide strand is influenced by the thermodynamic stability at the 5' end of the miRNA duplex, typically favoring the strand with lower 5' stability or a 5' uracil[188] (Figure 8).

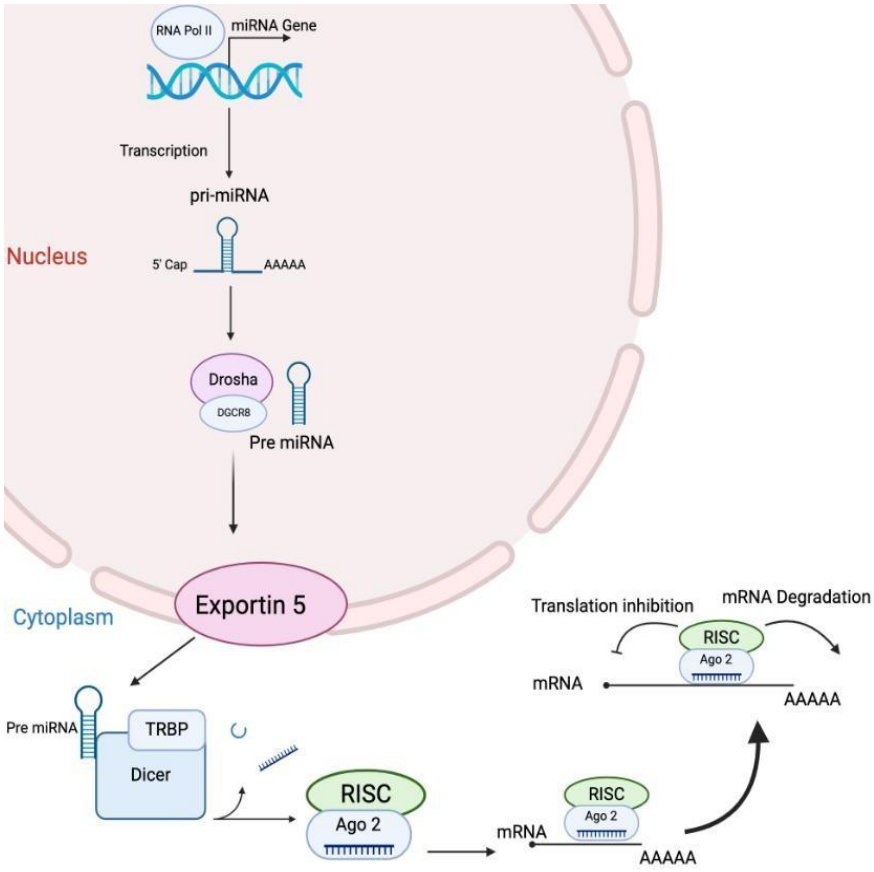



Figure 8. microRNA biogenesis. Template obtained from Biorender.com

### 2.2.2 MICRORNAs FUNCTION

miRNAs are vital regulators of gene expression with profound implications for human cancer. Their discovery and subsequent research have unveiled a complex landscape of regulatory networks and interactions.

miRNAs typically regulate gene expression by binding to specific sequences in the 3' untranslated region (UTR) of target mRNAs, leading to translational repression and mRNA degradation. miRNAs can also interact with other regions such as the 5' UTR and coding sequences, as well as promoter regions, to exert their regulatory effects. The degree of complementarity between the two sequences determines the fate of the mRNA. In mammals, there are four Ago proteins (Ago1-4), with Ago2 being the most prevalent. Ago2 is uniquely capable of degrading mRNA when the guide miRNA and mRNA sequences are perfectly complementary. When the miRNA-mRNA complementarity is full, Ago2 endonuclease activity is enhanced and the mRNA is cleaved[189,190].

The binding of miRNAs to their target mRNAs usually involves the formation of a miRNA-induced silencing complex (miRISC). This complex is recruited to the mRNA by the GW182 family of proteins, which provide scaffolding to bring in other effector proteins like poly(A)-deadenylase complexes. These complexes initiate and complete the deadenylation of the target mRNA, respectively. Following deadenylation, the mRNA undergoes decapping and subsequent degradation by the exoribonuclease XRN1 or inside the exosome by catalytic nucleases[190].

Interestingly, while most studies focus on the repressive role of miRNAs, there is evidence that miRNAs can also activate gene expression under certain conditions. For example, in serum-starved cells, miRNAs such as let-7 can associate with AGO2 and FXR1 to activate translation during cell cycle arrest. This miRNA-mediated activation typically involves binding to AU-rich elements (AREs) at the 3' UTR or the 5' UTR of target mRNAs[191].

Additionally, miRNAs can regulate gene expression within the nucleus. Through interactions with proteins like Importin-8 or Exportin-1, AGO2 shuttles between the nucleus and cytoplasm. In the nucleus, miRISC can regulate transcriptional rates and post-transcriptional levels of mRNA, often associating with euchromatin at actively transcribed gene loci. The exact mechanisms of nuclear miRNA functions remain an area of ongoing research.

### 2.2.3 MICRORNAs SUBCELLULAR COMPARTMENTALIZATION

MiRISC and target mRNA have been found in various subcellular compartments, such as the rough endoplasmic reticulum, processing (P)-bodies, the trans-Golgi network, early and late endosomes, multivesicular bodies, lysosomes, mitochondria and the nucleus[190] (Figure 9).

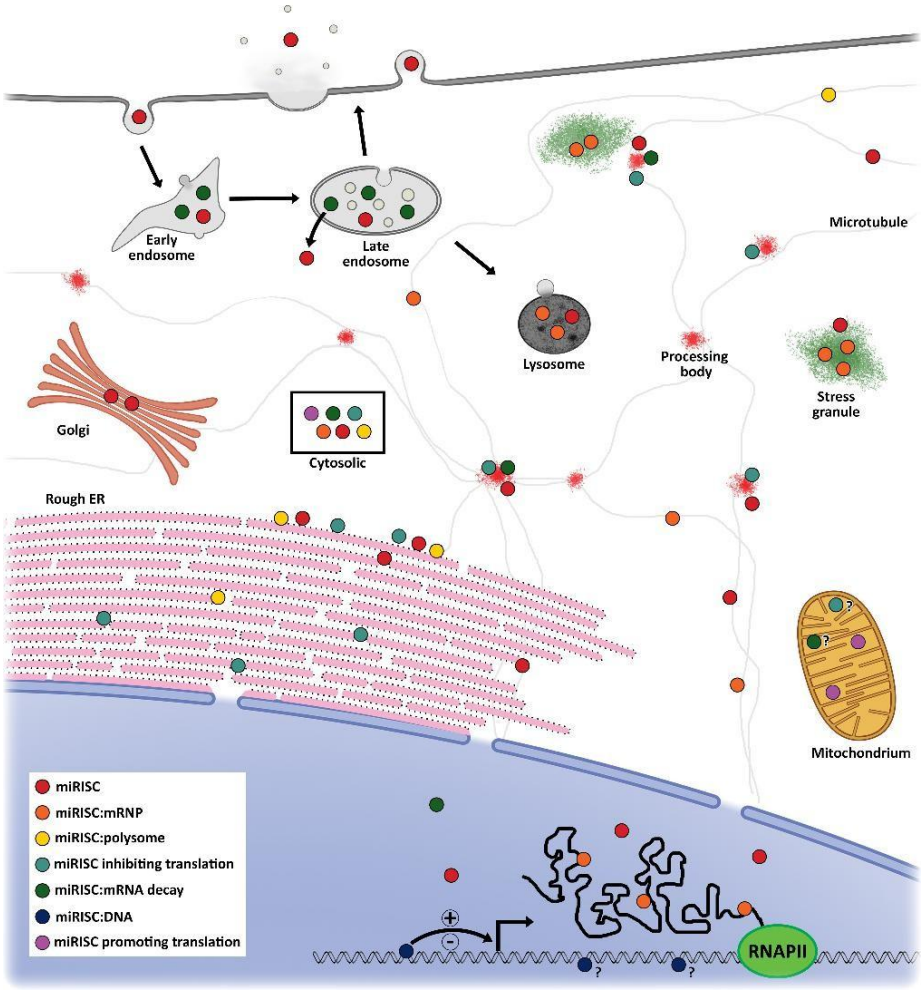



Figure 9. Proposed model by O'Brien et al., [190] of miRNA location and function.

Also, there are extracellular miRNAs in biological fluids such as plasma, serum, cerebrospinal fluid, saliva, breast milk, urine, and more. They exist in two forms: within vesicles like exosomes and microvesicles or associated with proteins, predominantly AGO2[192,193]. Their stability in extracellular environments is remarkable, enduring conditions like boiling and multiple freeze-thaw cycles[194].

The secretion and uptake of miRNAs are believed to be regulated processes rather than mere by-products of cellular activities.

Extracellular miRNAs may function as autocrine, paracrine, and endocrine regulators, having hormone-like activities and playing significant roles in intercellular communication[195].

## 3. BIG DATA AND CANCER

In today's world, as we enter the era of Industrial Revolution 4.0, the term "Big Data" has gained significant popularity. It symbolizes the massive volumes of data generated daily by individuals, organizations, and devices globally. Big Data is characterized by its large volume, high velocity, and wide variety, necessitating specialized technologies and analytical methods to transform it into valuable insights[196].

Applied to science, Big Data holds the promise of providing highly efficient new methods for planning, evaluating, and disseminating research. This field combines computational, algorithmic, statistical, and mathematical techniques to extract knowledge from extensive datasets. Researchers are utilizing the ability to link and cross-reference data from diverse sources to enhance the accuracy and predictive power of scientific findings and to identify future research directions, with cancer research being a prime example[197].

Traditionally, cancer research has focused on the molecular and clinical investigation of specific genes and pathways involved in cancer development and progression. However, the advent of high-throughput sequencing and other advanced technologies has led to a data explosion, ushering in the era of 'Big Data.' This has resulted in vast amounts of omics data, including genomics, transcriptomics, proteomics, and metabolomics. The integration of this information facilitates the

identification of novel biomarkers for early cancer detection, prognosis, and treatment response monitoring[198].

Despite these significant advancements, integrating Big Data into cancer research faces several challenges, such as ensuring access to high-quality, annotated datasets and effectively integrating diverse data into cohesive models. Moreover, the structure of science and the scientific publication system has led to the well-known "publish or perish" paradigm. Scientists are pressured to publish numerous articles in a short time to advance their careers, prioritizing quantity over quality[199]. This rapid publication pace has reduced the time scientists spend reading the available literature in their field, and cancer research is not an exemption. As a result, there is considerable experimental redundancy on the same topics that do not contribute new knowledge, leading to many low-quality articles.

## 3.1 SYSTEMATIC REVIEWS AND META-ANALYSES

Systematic reviews and meta-analyses are effective ways to review the literature published by colleagues in your field while producing high-quality articles. Systematic reviews represent a specific type of research where the units of analysis are original primary studies. They are essential tools for synthesizing available scientific information, increasing the validity of primary study conclusions, and identifying future research areas. A meta-analysis is a type of systematic review that uses statistical methods to combine the results of two or more studies. Meta-analyses do not provide a simple arithmetic average of the results but a weighted average, giving greater weight to studies with a larger information load, such as those that are larger or have a higher number of cases.

Meta-analyses, as an extension of systematic reviews, utilize statistical methods to combine data from multiple studies, yielding more robust and generalizable conclusions. By giving greater weight to larger and more reliable studies, meta-analyses provide a nuanced understanding of research outcomes, reducing the likelihood of false positives and negatives that might arise from smaller, isolated studies[200].

In the context of cancer research, these methodologies facilitate the identification of novel biomarkers, improve the precision of diagnostic and prognostic tools, and enhance the evaluation of treatment efficacy and safety. They help to streamline the vast amount of data generated in the era of Big Data, making it more manageable and interpretable. Consequently, systematic reviews and meta-analyses are invaluable for advancing scientific knowledge, informing clinical practice, and ultimately improving cancer patient outcomes.

# CHAPTER II. HYPOTHESIS

Breast cancer is the most commonly diagnosed cancer in women worldwide and the leading cause of cancer-related mortality among women globally. While screening techniques and imaging diagnostics have advanced significantly, with mammograms being the gold standard, they still present notable limitations such as over-diagnosis and varying sensitivity based on age and breast tissue density. These limitations underscore the need for new, complementary diagnostic tools.

Early detection is critical in breast cancer, with the 5-year survival rate exceeding 95% for localized cases but falling below 25% after metastasis. Despite widespread use of mammograms, the incidence of metastasis at diagnosis has not decreased. De novo metastatic breast cancer accounts for approximately 6–10% of all breast cancers and about 30% of metastatic breast cancers, with incidence rising over time. Early detection of LNM is vital for effective clinical management, treatment, and prognosis. However, current imaging techniques and the highly invasive SLNB often fall short in accurate preoperative staging and have significant limitations, including a false-negative rate exceeding 10% after preoperative systemic therapy.

Recent advancements in Next Generation Sequencing (NGS) have revealed substantial molecular differences between breast cancer subtypes and highlighted the therapeutic potential of intratumoral heterogeneity. miRNAs (microRNAs), which regulate gene expression rapidly and broadly, play a crucial role in cancer processes and phenotypical changes in tumor cells. These miRNAs, encapsulated in EVs circulating in the blood, play an important role in intercellular communication, and represent potential liquid biopsy biomarkers for early breast cancer diagnosis.

Taking these data into account, this doctoral thesis is based on two hypotheses:

1. The integration of all the data from all the investigations performed to date on the differential expression of miRNAs in NML by means of a meta-analysis will allow to identify specific genetic and miRNA profiles associated with disease progression, with a very reliable statistical power.

2. An experimental approach by sequencing EV-derived microRNAs in breast cancer patients at various stages will allow us to recognize the microRNAs that differentiate in the early stages of breast cancer, as well as those that differentiate when distal metastases occur. This comprehensive analysis will offer a less invasive, complementary diagnostic tool to mammography and SLNB and contribute to precision medicine, enhancing our understanding and management of breast cancer.

# CHAPTER III. OBJECTIVES

The are two main objectives of this doctoral thesis divided into more specific objectives.

1. To conduct a meta-analysis to identify microRNAs on all types of biological samples (tissue, serum, plasma...) involved in the diagnosis of LNM, collecting and statistically analyzing all relevant experimental studies to date.
    a. To analyze the differential expression of the microRNAs between LNM and early localized and metastatic stages in breast cancer.
    b. To pull all the fold change values find in experimental studies in LNM breast cancer in order to offer a high valuable statistic result.
    c. To provide a complementary, less invasive diagnostic tool to SLNB and mammography.

2. To use NGS techniques to sequence microRNAs transported by extracellular vesicles in two cohorts of breast cancer patients (early and metastatic stages) and a cohort of healthy donors.
    a. To identify specific EVs associated miRNA profiles of each tumoral stage and their associated target genes and molecular pathways.
    b. To analyze the clinical value of EV-derived microRNAs in each stage and their relation with clinic-pathological features.
    c. To evaluate the diagnostic and/or prognostic role of the expression of specific EV-derived miRNAs.
    d. To validate these results in an independent cohort, in order to improve the clinical value.

e. To describe the miRNA location in the EV and the origin of the EV.

# CHAPTER IV. MATERIAL AND METHODS

## 1. SYSTEMATIC REVIEW AND META-ANALYSIS

The goals of our first objective are twofold: first, to review the current literature to identify microRNAs that are specifically dysregulated in LNM among BC patients, and second, to evaluate the diagnostic and prognostic value of these microRNAs in the development of LNM and their role in predicting distal metastasis. To our knowledge, this is the first meta-analysis exclusively focused on microRNAs involved in the diagnosis of LNM in BC patients.

### 1.1 INCLUSION AND EXCLUSION CRITERIA

Studies were included in the systematic review if breast cancer patients were studied for differential expression of their microRNAs in loco-regional breast cancer or with lymph node metastases. Only studies in which the location of the patient's tumor was explicit were included, excluding studies that focused on the role of microRNAs in localized breast cancer or those with distal metastases. Likewise, studies in which this information was provided in an ambiguous manner were excluded.

Inclusion and exclusion criteria were the following:

*Inclusion Criteria*

1. Original research based on patient samples

2. Report outcomes of the role of miRNAs at the diagnosis or prognosis of the BC patients with LNM

3. measured miRNA expression levels in patient samples with LNM

*Exclusion Criteria*

1. Studies based on cell lines, databases, or animals

2. Meta-analysis, systematic reviews, and reviews

3. Studies based on the expression of microRNAs in localized breast cancer or distal metastatic breast cancer

4. Studies based on microRNAs as biomarkers in response to therapies

5. Studies assessing the effect of miRNA dysregulation through another gene

6. Other organ than breast and

7. Manuscripts retracted or published in languages other than English.

## 1.2 SYSTEMATIC SEARCH STRATEGY

The protocol of this review was registered in the international database of prospective registered systematic reviews (PROSPERO 2024, CRD42024534072). The following medical and health professionals' databases (PubMed, Web of Science, and SCOPUS) were searched until 26 March 2024:

- PubMed is a free search engine accessing primarily the MEDLINE database of references and abstracts on life sciences and biomedical topics. It is maintained by the United States National Library of Medicine at the National Institutes of Health.
- Web of Science is a comprehensive research platform owned by Clarivate Analytics that provides access to multiple databases containing reference and citation data from academic journals, conference proceedings, and other scholarly documents across various disciplines. Originally developed by the Institute for Scientific Information, Web of Science has evolved to include a

wide range of databases that support scientific and scholarly research.

- Scopus is a large, multidisciplinary abstract and citation database of peer-reviewed literature, including scientific journals, books, and conference proceedings, launched by the academic publisher Elsevier in 2004. It covers research in the fields of science, technology, medicine, social sciences, and arts and humanities

A systematic review and a meta-analysis were performed using these databases, following the guidelines in Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA)[201].

The search algorithm utilized was: breast AND (microRNA OR miRNA OR miR OR microRNAs OR miRNAs OR miRs) AND (lymph node metastasis). No filters or time restrictions were applied to ensure comprehensive coverage of relevant studies. Additionally, conference papers and book chapters from SCOPUS were included as grey literature to minimize search bias. The retrieved records were subsequently imported into an Excel file for further analysis.

### 1.3 DATA EXTRACTION

After removing duplicates, the titles and abstracts of the remaining articles were reviewed in the first screening phase to determine their alignment with the inclusion criteria. A second screening was then conducted based on both inclusion and exclusion criteria to assess the articles' eligibility (Supplementary Table I). The primary researcher performed data extraction, which was subsequently verified for completeness and accuracy by the primary supervisor. For data that was difficult to interpret or missing, the authors were contacted directly via

email to obtain the necessary information. The data extraction of the remain articles included the name of the microRNA/s, the breast cancer molecular subtype, the number of patients (with LNM) and control cases (patients with localized BC), the type of sample used, the method of analysis, state of validation, patient treatment and the direction of the regulation.

### 1.4 STUDY QUALITY ASSESSMENT

When conducting a meta-analysis, it is important that the included research is of good methodological quality to ensure the integrity of the meta-analysis. The systematic review quality and susceptibility to bias of all studies included in the review were evaluated by using the 27-item checklist PRISMA 2020 statement [201].

The Quality Assessment of Diagnostic Accuracy Studies -2 (QUADAS-2) criteria was employed as a tool to assess the quality of the studies included in the meta-analysis. QUADAS-2 is a tool designed to assess the quality of diagnostic accuracy studies in systematic reviews and meta-analyses. Developed as a revision of the original QUADAS tool, it aims to improve the evaluation of the risk of bias and applicability of studies assessing diagnostic tests. The tool consists of four main domains, each evaluated for risk of bias and applicability concerns:

- Patient Selection: This domain assesses how patients were selected for the study, including whether a random or consecutive sampling method was used and if there were any inappropriate exclusions.
- Index Test: This domain evaluates the conduct and interpretation of the diagnostic test being studied, checking if the test results

were interpreted without knowledge of the reference standard results.

- Reference Standard: This domain examines the reference standard used to confirm the diagnosis, ensuring it is likely to accurately classify the target condition.
- Flow and Timing: This domain looks at the patient flow through the study, including any patients who did not receive the index test or reference standard and the timing of these tests relative to each other.

## 1.5 META-ANALYSIS

### 1.5.1 Model

A meta-analysis can be conceptualized in two ways: using fixed-effect models or random-effects models [202]. The fixed-effect model assumes that all studies are estimating the same underlying effect size. It is typically used when the studies are very similar in terms of participants, interventions, and outcomes. On the other hand, the random-effects model assumes that the effect sizes vary between studies due to differences in study populations, interventions, or other factors. It is more appropriate when there is significant heterogeneity among the studies included in the meta-analysis. For this study, due to the enormous number of variables that can influence the expression of microRNAs and the varied array of methodologies employed for measure them, we decided to use the random-effects model and normalise the heterogeneity found. Furthermore, the results drawn from a random-effects model are generally more flexible and can be extended to a broader population of studies beyond the sample. A crucial parameter for a meta-analysis is the

effect size, representing a quantitative measure that reflects the magnitude of the relationship or difference between groups being studied. It provides a standardized way to compare results across different studies, allowing researchers to synthesize findings and draw broader conclusions.

For our meta-analysis, microRNAs appearing in fewer than three independent studies were excluded. Additionally, only miRNAs with directly provided effect sizes (sensitivity, specificity or fold change) and sample sizes (number of patients and controls) were included. Fold change (FC) was the only effect size consistently reported across the publications that met the inclusion criteria, and thus, it was selected as the effect size for the meta-analysis.

All fold change values were standardized to a consistent scale: no fold change was set to one, with downregulation ranging from 0 to 1. This was done because some publications presented no fold change as zero, assigning negative values for downregulation, while others used one as the no fold change value. To address the FC asymmetry caused by differences in reference group selection across studies, we used the log2(FC) of the standardized values as the primary endpoint [203].

### 1.5.2 Heterogeneity assessment

As part of conducting a M-A there is an underlying aim to attempt to measure and control for heterogeneity. Heterogeneity bias was assessed using Cochran's Q test, with the primary purpose of determine whether the observed variability in effect sizes across studies is greater than what would be expected by chance alone. A $p < 0.10$ suggests significant heterogeneity. We also employed $I_2$ statistic to quantify the percentage

of total variation across studies that is due to heterogeneity rather than chance[204], with $I_2 > 0.25$ indicating significant heterogeneity.

### 1.5.3 Publication bias

Publication bias refers to the tendency for studies with statistically significant or positive results to be more likely to be published than studies with non-significant or negative results. This can lead to an overestimation of the true effect size in the meta-analysis. Several authors have proposed various strategies for addressing publication bias over the long term, along with statistical techniques for identifying and mitigating it. To assess publication bias here we used Eager's, a regression-based test that assesses funnel plot asymmetry, and Begg's tests, a rank correlation test that also assesses funnel plot asymmetry.

## 2. IDENTIFICATION OF MICRONAs FOR BREAST CANCER DIAGNOSIS

### 2.1 STUDY POPULATION

A prospective observational study was conducted in two independent cohorts of BC patients: (i) newly diagnosed patients with early BC (with or without loco-regional metastasis) and with an indication for surgery and (ii) patients with advanced BC and distal metastasis. A population of sex-age matched healthy donors (HD) was included. The Ethical Committee of San Cecilio University Hospital (Granada) approved the study protocol, which was conducted following the Declaration of Helsinki. All cancer patients and healthy volunteers signed written informed consent before participating.

For the first cohort, the main inclusion criteria were patients older than 18 years, stage I-III, ECOG Performance Status ≤2 and having performed molecular pathological analysis of the tumor's hormone receptors. For the second cohort the inclusion criteria were patients older than 18 years, stage IV and having performed molecular pathological analysis of the tumor's hormone receptors. In both cohorts the exclusion criteria were having had or simultaneously having another type of tumor or other benign breast pathologies.

### 2.2 STUDY DESIGN

For the analysis of these circulating microRNAs, the study was divided into three phases: marker discovery, marker validation and blind validation. For this, the two cohorts of recruited BC patients and healthy donors were distributed as shown in Figure 10.

## 2.3 SAMPLE COLLECTION

Samples of BC (localized and metastatic BC) patients consisted in 10 mL of peripheral blood collected in EDTA Vacutainer tubes at the moment of diagnosis, stored at room temperature, and processed at GENYO Centre (Granada) within 4 hours. The clinical outcomes of the patients enrolled in the study were collected by the oncologist at the San Cecilio University Hospital (Granada). The total volume of blood was used for plasma extraction, from which circulating EV-derived microRNAs were extracted. The clinical-pathological characteristics of the patients involved in each step are summarized in Table 3 (early BC patients) and Table 4 (metastatic BC patients). Due to all the patients included in the study are women, individual sex is not specifying on the tables.

| | **Characteristics** | **Discovery (N=11)** | **PCR Validation (N =20)** | **Marker validation (N=100)** | **P value** |
|---|---|---|---|---|---|
| **EARLY BREAST CANCER** | Age | Mean ± SD | 53±12.25 | 53±9.45 | 53±8.09 | |
| | | <45 | 3 (27.3%) | 5 (25%) | 14 (14%) | 0.97 |
| | | 45-64 | 5 (45.5%) | 12 (60%) | 74 (74%) | |
| | | >65 | 3 (27.3%) | 3 (15%) | 12 (12%) | |

| | | | | | |
|---|---|---|---|---|---|
| Menopause | Yes | 5 (45.5%) | 14 (70%) | 64 (64%) | 0.693 |
| | No | 6 (54.5%) | 6 (30%) | 36 (36%) | |
| cTNM | T1N0M0 | 2 (18.2%) | 16 (80%) | 6 (6%) | 0.001 |
| | T1N1Mx | 3 (27.3%) | 0 | 1 (1%) | |
| | T2N0M0 | 1 (9.1%) | 2 (10%) | 16 (16%) | |
| | T2N1M0 | 0 | 0 | 26 (26%) | |
| | T2N1Mx | 0 | 0 | 50 (50%) | |
| | T2N2M0 | 1 (9.1%) | 0 | 1 (1%) | |
| | T3N1M0 | 4 (36.4%) | 0 | 0 | |
| | T3N0Mx | 0 | 1 (5%) | 0 | |
| | T4c | 0 | 1 (5%) | 0 | |
| Tumor size | Mean ± SD | 41±23.19 | 18±20.03 | 18±16.31 | |
| | < 10 | 0 | 5 (25%) | 19 (19%) | 0.00 |
| | 10 - 20 | 1 (9.1%) | 10 (50%) | 59 (59%) | |
| | 20 - 50 | 8 (72.7%) | 3 (15%) | 18 (18%) | |
| | > 50 | 2 (18.2%) | 2 (10%) | 4 (4%) | |

| | | | | | |
|---|---|---|---|---|---|
| Tumor stage | I | 4 (36.4%) | 12 (60%) | 50 (50%) | 0.35 |
| | II | 5 (45.5%) | 6 (30%) | 42 (42%) | |
| | III | 2 (18.2%) | 2 (10%) | 8 (8%) | |
| Perou Classification | Luminal A | 2 (18.2%) | 4 (20%) | 35 (35%) | 0.03 |
| | Luminal B | 4 (36.4%) | 16 (80%) | 62 (62%) | |
| | HER2 | 3 (27.3%) | 0 | 1 (1%) | |
| | Triple negative | 2 (18.2%) | 0 | 2 (2%) | |
| ki67 | <=14 | 3 (27.3%) | 9 (45%) | 32 (32%) | |
| | 14 - 50 | 5 (45.4%) | 7 (35%) | 52 (52%) | 0.545 |
| | 50 -70 | 0 | 2 (10%) | 11 (11%) | |
| | > 70 | 3 (27.3%) | 2 (10%) | 3 (3%) | |
| | Not done | 0 | 0 | 2 (2%) | |
| Affected nodes | 0 | 5 (45.5%) | 10 (50%) | 69 (69%) | 0.121 |
| | 1 - 3 | 3 (27.3%) | 9 (45%) | 20 (20%) | |
| | >3 | 2 (18.2%) | 1 (5%) | 11 (11%) | |

| | | | | | |
|---|---|---|---|---|---|
| | Not done | 2 (18.2%) | 0 | 0 | |
| Adyuvant chemotherapy | AC + taxol | 3 (27.3%) | 10 (50%) | 30 (30%) | 0.223 |
| | AC+ taxotere | 6 (54.5%) | 0 | 0 | |
| | Taxol + trastuzumab | 2 (18.2%) | 0 | 4 (4%) | |
| | Aramidex | 0 | 0 | 0 | |
| | Epirrubicin | 0 | 0 | 0 | |
| | Capecitabin | 0 | 0 | 0 | |
| | None | 0 | 10 (50%) | 66 (66%) | |
| Adyuvant hormonotherapy | Tamoxifen | 5 (45.5%) | 8 (40%) | 26 (26%) | |
| | Anastrozol | 2 (18.2%) | 2 (10%) | 26 (26%) | 0.51 |
| | Letrozol | 2 (18.2%) | 9 (45%) | 22 (22%) | |
| | Trastuzumab | 0 | 0 | 11 (11%) | |
| | Trastu+Pertu zumab | 0 | 0 | 2 (2%) | |
| | Exemestan | 0 | 0 | 3 (3%) | |
| | Giredestrant | 0 | 0 | 7 (7%) | |
| | None | 2 (18.2%) | 1 (5%) | 3 (3%) | |
| Radiotherapy | Yes | 9 (81.8%) | 2 (10%) | 92 (92%) | 0.583 |

| | | | | |
|---|---|---|---|---|
| | No | 2 (18.2%) | 18 (90%) | 8 (8%) |
| Exitus | Yes | 0 | 1 (5%) | 1 (1%) |
| | No | 11 (100%) | 19 (95%) | 99 (99%) |

Table 3. Clinic-pathological characteristics of patients with early breast cancer included in this study in the different phases (discovery, validation and blind validation phase).

| Characteristics | | Sequencing (N=6) | Marker Validation (N=11) | P value |
|---|---|---|---|---|
| Age | Mean ± SD | 46±6.38 | 58±10.14 | 0.01 |
| | <45 | 3 (50%) | 0 | |
| | 45-64 | 3 (50%) | 7 (63.6%) | |
| | >65 | 0 | 4 (36.4%) | |
| Menopause | Yes | 2 (33.3%) | 11 (100) | 0.246 |
| | No | 4 (66.7%) | 0 | |
| cTNM | T1N0M0 | 1 (16.7%) | 1 (9.1%) | 0.244 |
| | T1n0M1 | 0 | 3 (27.3%) | |
| | T2N0M0 | 1 (16.7%) | 3 (27.3%) | |
| | T2N1M0 | 0 | 1 (9.1%) | |
| | T2N2M0 | 1 (16.7%) | 0 | |
| | T2N3M1 | 0 | 2 (18.2%) | |
| | T3N1M0 | 3 (50.0%) | 0 | |
| | T2N1M1 | 0 | 1 (9.1%) | |
| Tumor size | Mean ± SD | 35±14.31 | 28±16.11 | 0.256 |
| | < 10 | 0 | 2 (18.2%) | |
| | 10 - 20 | 6 (100%) | 1 (9.1%) | |
| | 20 - 50 | 0 | 5 (45.5%) | |
| | > 50 | 0 | 1 (9.1%) | |
| | Unknown | 0 | 2 (18.2%) | |
| Tumor stage | I | 0 | 1 (9.1%) | 0.078 |
| | II | 0 | 1 (9.1%) | |
| | III | 4 (66.7%) | 1 (9.1%) | |
| | IV | 2 (33.3%) | 6 (54.5%) | |
| | Not done | 0 | 1 (9.1%) | |
| Perou Classification | Luminal A | 6 (100%) | 1 (9.1%) | 0.808 |
| | Luminal B | 0 | 4 (36.4%) | |

| | | | | |
|---|---|---|---|---|
| | HER2 | 0 | 1 (9.1%) | |
| | Triple negative | 0 | 1 (9.1%) | |
| | Not done | 0 | 3 (27.3%) | |
| ki67 | <=14 | 0 | 0 | 0.961 |
| | 14 - 50 | 4 (66.6%) | 7 (63.6%) | |
| | 50 - 70 | 1 (16.7%) | 4 (36.4%) | |
| | > 70 | 1 (16.7%) | 0 | |
| | Not done | 0 | 4 (36.4%) | |
| Adyuvant quimiotherapy | AC + taxol | 1 (16.7%) | 2 (18.2%) | 0.015 |
| | AC+ taxotere | 1 (16.7%) | 1 (9.1%) | |
| | Epirubicin | 0 | 2 (18.2%) | |
| | Arimidex | 0 | 2 (18.2%) | |
| | Docetaxel | 0 | 1 (9.1%) | |
| | Taxol + trastu | 0 | 1 (9.1%) | |
| | None | 4 (66.6%) | 2 (18.2%) | |
| Adyuvant hormonotherapy | Tamoxifen | 1 (16.7%) | 2 (18.2%) | 0.037 |
| | Anastrozol | 1 (16.7%) | 0 | |
| | Fulvestrant | 0 | 1 (9.1%) | |
| | Letrozol | 2 (33.3%) | 2 (18.2%) | |
| | Exemestan | 2 (33.3%) | 0 | |
| | None | 0 | 6 (54.5%) | |
| Radiotherapy | Yes | 4 (66.7%) | 7 (63.6%) | 0.35 |
| | No | 2 (33.3%) | 4 (36.4%) | |
| Organ metastasis | Liver | 5 (83.4%) | 2 (18.2%) | 0.462 |
| | Brain | 2 (33.3%) | 1 (9.1%) | |
| | Bone | 0 | 8 (72.7%) | |
| | Lung | 2 (33.3%) | 4 (36.4%) | |

| | | | | |
|---|---|---|---|---|
| | None | 0 | 0 | |
| Exitus | Yes | 6 (100%) | 1 (9.1%) | 0.001 |
| | No | 0 | 10 (90.9%) | |

Table 4. Clinic-pathological characteristics of patients with metastatic breast cancer included in this study in the different phases (discovery and validation phase).

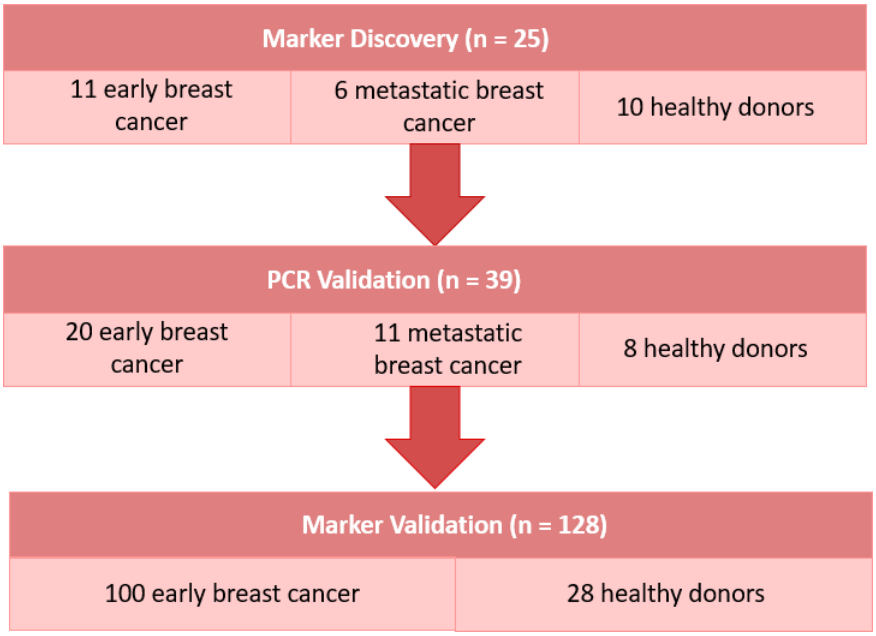



Figure 10. Schematic representation of the study's workflow.

### 2.4 CELL CULTURE

The MCF-7 human breast cancer cell line was sourced from the American Type Culture Collection (ATCC, U.S) and cultured in DMEM (Biowest, France) supplemented with 10% fetal bovine serum (Biowest, France), 100 U/ml penicillin, and 100 ng/ml streptomycin (Biowest, France). The cells were maintained in a humidified incubator at 37°C with 5% CO2. For detachment, TrypLE Express (ThermoFisher, U.S) was used, as it is gentler on cells than trypsin or other dissociation reagents, thereby better preserving the extracellular domains of transmembrane proteins.

### 2.5 EV ISOLATION AND CHARACTERIZATION

EVs from BC patients, healthy donors and cell culture were isolated using protocols previously developed by our group[205]. Initially, blood samples and culture samples were centrifuged at 2000 x g for 10 minutes to collect plasma and culture supernatant. The supernatants were then centrifuged at 10,000 x g for 30 minutes to remove cellular debris. The resulting supernatants were transferred into 6 ml polyallomer ultracentrifuge tubes, filled with 1X PBS, and ultracentrifuged using a TFT 80.4 Rotor (ThermoFisher, U.S) at 100,000 x g for 1 hour and 30 minutes at 4°C using Sorvall VX Ultracentrifuge (ThermoFisher, U.S).

After that, supernatants were removed and EV pellets were resuspended in different solutions depending on their posterior analyzes. For the microRNAs extraction, TEM and NTA, EVs were resuspended in 1X PBS, while for western blot analysis, EV pellets were directly lysed by adding ice-cold 1X Cell Lysis Buffer (Cell Signaling Technology, U.S).

### 2.5.1 TRANSMISSION ELECTRON MICROSCOPY

The vesicle-containing fraction was applied to carbon-activated coated grids. These grids were placed on 10 µL sample drops for 5 minutes and then washed five times with 100 µL PBS drops. The samples were fixed using a 1% glutaraldehyde-PBS solution. Afterward, the grids were rinsed with ddH2O and immediately transferred onto drops of uranyl-methylcellulose solution (pH 4) on a cooled metal plate for 5 minutes. Finally, the grids were air-dried at room temperature. The samples were then loaded onto a FEI Tecnai 12 Transmission Electron Microscope (JEOL, Germany) for imaging and analysis at 80 kV.

### 2.5.2 NANOPARTICLE TRACKING ANALYSIS

The concentration and size distribution of plasma-derived and cell culture EVs were assessed using a NanoSight NS300 system (Malvern Panalytical, UK), which measures individual particles in a physiological buffer by capturing videos. Each video recorded five different positions within the EV sample. Prior to analysis, the samples were diluted 1000-fold in 1X PBS, and size distribution was measured using Blue 405, the 405 nm laser at room temperature. Each video was 60 seconds long, with a frame rate of 1 frame per second, a sensitivity setting of 80, and a shutter speed of 100. An automatic measurement report was generated for each sample. The analysis was conducted using NTA 3.2 software (Malvern Panalytical, UK). The detection threshold was set to 6, with all other settings left at their default values.

### 2.5.3 WESTERN BLOT

EVs from patient plasma and the MCF-7 cell line were sonicated for 30 seconds and then centrifuged at 14,000 x g for 10 minutes at 4°C. The resulting supernatants were collected and stored for subsequent protein analysis. Protein concentration was measured using the Pierce™ BCA Protein Assay Kit (ThermoFisher, U.S) following the manufacturer's instructions, using the Infinite® 200 PRO NanoQuant plate reader (TECAN, Switzerland).

For each sample, 30 μg of protein were loaded onto Mini-PROTEAN TGX 4–20% precast gels (Bio-Rad, U.S) and electrophoresed at 120 mV for protein separation. Then, proteins were transferred to nitrocellulose membranes using the X-Cell II Blot module (Invitrogen, U.S). Membranes were blocked with 5% non-fat dry milk in Tris-buffered saline containing 1% Tween 20 for 1 hour, followed by overnight incubation at 4°C with primary antibodies diluted 1:1000 in the same solution employed for blocking the membranes.

For EV characterization, primary antibodies used include: mouse monoclonal $IgG_{2b}$ κ anti-CD81 (Santa Cruz Biotechnology, U.S - sc-166029), mouse monoclonal IgG2b anti-Hsp70 (BD Biosciences, U.S - 554243) and rabbit monoclonal IgG Anti-EpCAM antibody (Abcam, UK - ab223582). Next day, the membranes were incubated with a goat anti-mouse and goat anti-rabbit HRP-conjugated secondary antibody (Abcam, UK - ab97023) at a 1:5000 dilution in blocking buffer for 1 hour at room temperature. Finally, the membranes were revealed using Clarity Max™ Western ECL Substrate (Bio-Rad, U.S) using the ImageQuant LAS 4000 system (GE Healthcare Life Sciences, U.S).

## 2.6 TRANSCRIPTOMIC PROFILING OF MICRORNAs

### 2.6.1 EVs TOTAL RNA EXTRACTION

EVs samples were thaw on ice before total RNA extraction with miRNeasy Micro Kit (Qiagen, Germany). This kit is employed to extract the total RNA in the sample, but enriched in microRNAs. Briefly, 700 ul of QIAzol Lysis Reagent was added to the EV samples, followed by homogenization and incubation at room temperature. Chloroform is then added, and the mixture is centrifuged at 12.000 x g for 15 min at 4ºC to separate the aqueous phase, which contains the RNA.

This phase is transferred to a new tube, mixed with 100% ethanol, and loaded onto a spin column for RNA purification. The sample is subjected to several washes with specific buffers (RWT and RPE) and after that a wash with 80% ethanol to remove impurities. Later, the RNA is then eluted with 28 ul of RNase-free water.

The quality of the RNA extraction was validated using the Qubit RNA HS Assay (ThermoFisher, Wilmington, DE, U.S). Additionally, RNA integrity was assessed with the RNA Pico 6000 Assay Kit on the Agilent Bioanalyzer 2100 system (Agilent Technologies, U.S) prior to sequencing. The purified RNA is stored at -80°C until further applications.

### 2.6.2 MICRORNA SEQUENCING

After the quality control with the Qubit Assay and the Bioanalyzer, samples were sent to Macrogen (South Korea) for library construction and RNA sequencing. Small RNA libraries were prepared using the

TruSeq Small RNA Library Prep Kit (Illumina, Inc., U.S). The process began with the ligation of an adenylated single-strand DNA 3' adapter, followed by a 5' adapter to the small RNAs using a ligating enzyme. These adapters selectively captured small RNAs with 5' phosphate groups, a characteristic feature of microRNAs. The miRNA fragments with ligated adapters were then converted into cDNA fragments and amplified by PCR. Post-amplification, the target DNA fragments were separated on an agarose gel, and the band containing the miRNA fragments was excised. The cDNA libraries were subsequently recovered, purified, and sequenced on the Illumina HiSeq 2500 platform.

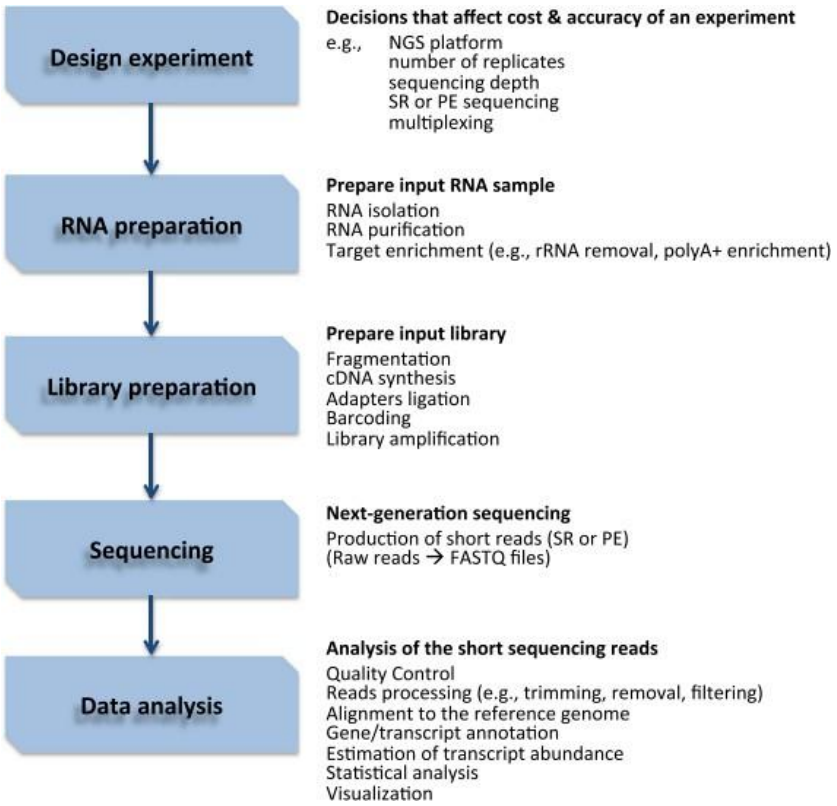



Figure 11. RNA Sequencing Experiment Workflow[206]

The quality of the raw sequences was assessed using FastQC v0.11.7. Adapter sequences, poly-A tails, and primers were removed with Cutadapt 2.8[207]. miRDeep2[208] was then used to identify microRNA genes. For alignment, the short-read aligner Bowtie 1.1.2[209] and HISAT2[210] v.2.1.0 were employed, using the reference human genome GRCh37. Finally, known and published miRNAs were identified and compared with the miRBase and TarBase databases[211,212].

### 2.6.3 BIOINFORMATICAL ANALYSIS

MicroRNA expression levels were quantified using the miRDeep2 package in R, based on logFC, logCPM, P-value, and False Discovery Rate (FDR) parameters. Pairwise comparisons were conducted between the different groups (Control vs. Early, Control vs. Metastatic, and Early vs. Metastatic). Significantly differentially expressed miRNAs were identified using a threshold of $P < 0.05$ and $logFC > 0$. After that, we employed $FDR < 0.01$ as a more restricted filter in order to find the most differentially expressed microRNAs. The resulting microRNAs were only found in the healthy-metastatic comparison, but we also used the values of these microRNAs in the healthy-early and early-metastatic comparisons, as they all showed a gradual and significant (p value < 0.05) increase in expression with disease progression.

#### FUNCTIONAL ANALYSIS

Functional analysis of differentially expressed microRNAs was conducted using DIANA tools, specifically miRPath v.3.0, to predict biological pathways based on the microT-CDS algorithm and the KEGG database. The analysis was performed with a significance threshold of P

< 0.05 and a microT score threshold of <0.8, with FDR correction applied. Dotplots were generated using the molecular pathways resulting from KEGG database.

A list of target genes of the differentially expressed microRNAs was extracted from DIANA's TarBase v8. Only genes with validated positive targets were retained. The frequency with which each gene was targeted was then calculated. Relationship networks between microRNAs and their target genes were constructed using Cytoscape. To be included in the network, a gene had to be regulated by the three microRNAs in the network.

Next, we used the Variant Interpreter tool from Hipathia v1.7.4 to analyzed the impact of the knockdown of each gene involve in the principal molecular pathways of breast cancer. This tool compares the sample expression values of the GTeX for breast cancer against the same samples but making a knock-out (reducing by 0.0001) in the expression of the chosen target gene. Only pathways with FDR and p-value < 0.05 were included.

## 2.7 DIAGNOSTIC AND PROGNOSTIC VALUE OF THE EV-DERIVED MICRORNAS

### 2.7.1 VALIDATION: MICRORNA PANEL SELECTION

To validate the markers identified from the sequencing results, microRNA levels derived from EVs were measured using qRT-PCR assays in an independent cohort. We focused on a signature of three microRNAs (miR-423-5p, miR-141-3p and miR-320b) out of the most differentially and gradually expressed. In the blind validation phase, we decided to increase the cohort and continue with only one

microRNA (miR-423-5p), as it was the one that gave the best results in the previous phase.

To ensure statistical relevance, we used the GPower 3.1 program to calculate the necessary sample size, setting an α error probability of 0.05 and power of 0.95. Furthermore, to normalize for RNA extraction variability, spike-in cel-mir-39 was added to all samples prior to total RNA extraction, in a final concentration of 10pM. miR- 16 was chosen as the endogenous control for the qRT-PCR assays based on an extensive literature revision.

### 2.7.2 cDNA SYNTHESIS AND QUANTITATIVE REAL TIME PCR

After total RNA extraction using miRNeasy Micro Kit (Qiagen, Germany), complementary DNA (cDNA) was synthesized with the TaqMan™ Advanced miRNA cDNA Synthesis Kit (Applied Biosystems, U.S), according to manufacturer protocol. This kit extends the mature miRNA by first adding a 3' poly-A tail and then ligating a 5' adaptor (Figure 12).

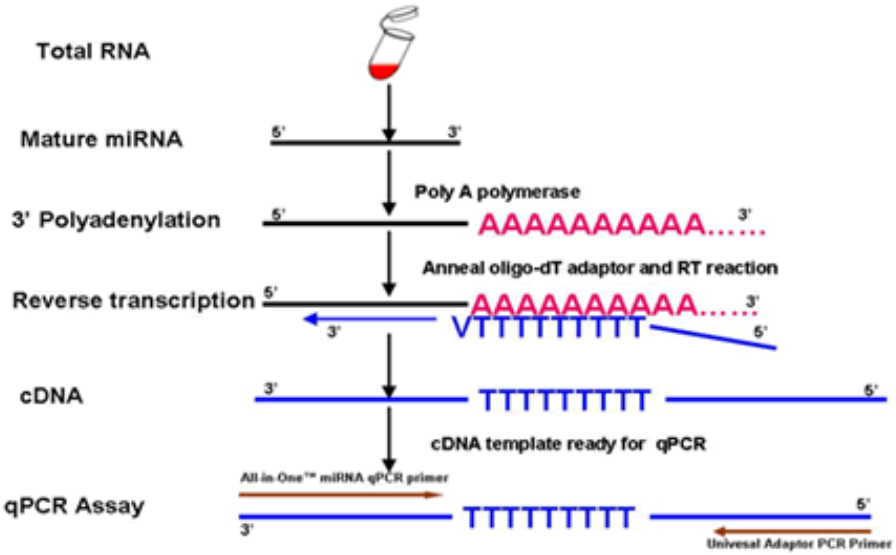



Figure 12. Overview of steps involved in the cDNA synthesis from miRNAs and qPCR detection. Genecopoeia.

microRNA expression levels were analyzed in triplicate and included non-template controls (NTC) using TaqMan™ MicroRNA assay probes and TaqMan™ Universal PCR Master Mix (Applied Biosystems), following the manufacturer's guidelines. Sequences can be found in Table 5. The assays were conducted on an Applied Biosystems 7900HT Fast Real-Time PCR System. The PCR cycling protocol included an initial denaturation and enzyme activation at 95°C for 2 minutes, followed by 40 cycles of 15 seconds at 95°C for denaturation and 30 seconds at 60°C for annealing and extension. Expression levels were calculated first subtracting the Ct mean of the spike-in of each sample to all the microRNAs in order to normalise the RNA extraction error, and then by the 2-ΔΔCt method, using miRNA-16 as housekeeping[213].

| microRNA | Mature Sequence |
|---|---|
| **hsa-miR-141-3p** | UAACACUGUCUGGUAAAGAUGG |
| **hsa-miR-423-5p** | UGAGGGGCAGAGAGCGAGACUUU |
| **hsa-miR-320b** | AAAAGCUGGGUUGAGAGGGCAA |

Table 5. microRNAs and primer sequences.

### 2.8 MICRORNA LOCATION IN THE VESICLE

As we said before, due to the recent discovery of the existence of co-isolated proteins or a corona surrounding the EV that can influence the biodistribution of the EVs in the body and the progression of cancer, it is important to know the location of the microRNA in the vesicle. Furthermore, if these microRNAs are mediators of the intercellular communication as a result of a pathological condition as cancer, it is suspected that the microRNA would be more likely to be inside the vesicle to provide information to the receptor cell.

To do this experiment, MCF-7 breast cancer cell line, healthy donor, early and metastatic BC EV samples were treated with four conditions:

1. Control Buffer (PBS 1x).
2. Control Buffer (PBS 1x), proteinase inhibitor (Sigma, Germany, P8340) and RNase A Purelink (Thermo Fisher, U.S)
3. Control Buffer (PBS 1x), proteinase K Ambion (Thermo Fisher, AM2546), proteinase inhibitor and RNase
4. Lysis buffer (Triton 1%), proteinase K, proteinase inhibitor and RNase.

Then, RT-qPCRs of microRNAs of interest were performed in each sample under the four conditions. Condition one is expected to present the maximum signal of amplification, second condition will remove free

RNA in plasma, and with third condition we will eliminate the free RNA and the RNA protected by proteins. Finally, in fourth condition we will degrade also the RNA inside the vesicle, so it is the condition with the less signal expected. The microRNA expression was normalized subtracting to the Ct mean of each microRNA the Ct mean of each experiment (ΔCt), and then 2^-ΔCt was calculated.

### 2.9 TUMORAL RELATION OF THE VESICLE

To determine whether our EV-derived microRNAs predominantly comes from tumoral environment or from other tissues, we tried to enrich our sample in EpCAM+ EVs. Despite of not all EpCAM+ EVs come necessary from the tumor, the probability is high.

To do that, we employed Dynabeads Protein G (Thermofisher, U.S, 10003D) conjugated with anti-EpCAM antibody (Abcam, UK, ab71916). Samples were incubated overnight with the EpCAM-conjugated Dynabeads, Then, EpCAM-EV complexes were isolated using a magnetic separator. The EVs were then detached from the Dynabeads by resuspension in either QIAzol or 0.1x RIPA buffer, depending on the subsequent analysis requirements. To verify the successful isolation, we performed Western blot analysis for CD9 and EpCAM on both the isolated EVs and the remaining supernatant. Finally, RT-qPCR was conducted on the EpCAM-enriched EVs to quantify the levels of the microRNAs of interest.

### 2.10 STATISTICAL METHODS

Statistical analyzes and graph generation were conducted using SPSS, Version 22.0 (IBM Corp.) and GraphPad Prism, Version 8.02 (GraphPad Software). microRNA expression levels were analyzed both as

continuous variables (absolute numbers) and as dichotomous variables (high/low expression). The optimal cut-off points for microRNA expression level were determined using ROC curve analysis and the Youden's J statistic.

Descriptive analyses were performed for continuous clinic-pathological variables and miRNA expression levels. Associations between microRNA expression and clinical characteristics were evaluated using the non-parametric test Mann-Withney U for two independent variables or Kruskal-Wallis test when there are more than two independent variables. Spearman’s rank was used to assess the correlations between variables. Statistical significance was set at $p < 0.05$. Additionally, ROC curves were generated to assess the diagnostic potential of miRNA expression levels.

# CHAPTER V. RESULTS

# 1. SYSTEMATIC REVIEW AND META-ANALYSIS

## 1.1 LITERATURE SEARCH OUTCOME

In order to analyze the differential expression of miRNAs in LNM in breast cancer, we performed an extensive review of the literature. As Figure 13 show, an initial total of 2,584 published records were retrieved from three databases: PubMed (n = 570), SCOPUS (n = 894), and Web of Science (n = 1,120). After removing 922 duplicates, the remaining 1,290 records were screened manually by reviewing their titles and abstracts to determine eligibility. At this stage, 1,095 records were excluded for being off-topic. Consequently, 195 articles underwent full-text review for inclusion in the systematic review, out of which 143 were excluded for the following reasons: 9 focused on localized breast cancer, 58 on distal metastasis, 5 on other organs, 5 were reviews or meta-analyses, 22 explored different genes, 13 were based on cell culture, animal studies, or databases, 10 reported responses to therapies, 12 were retracted, and 8 were not relevant. Ultimately, 52 articles were included in the systematic review.

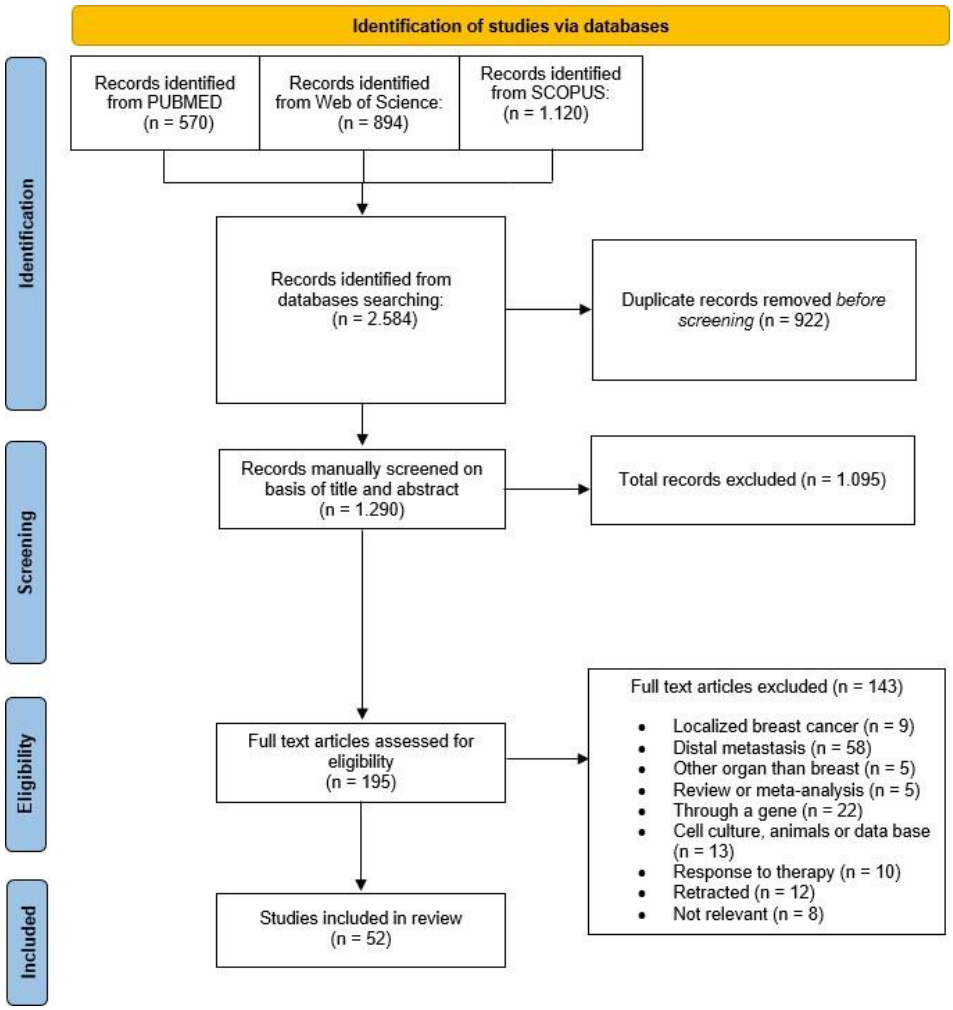



Figure 13. Flow diagram summary of item selection for this systematic review (PRISMA).

### 1.1.1 INCLUDED STUDIES

The main characteristics of the included studies are summarized in Table 6. Across the 52 articles included in the systematic review, a total of 84 miRNAs were identified as differentially expressed in breast cancer patients with LNM. All studies were retrospective and conducted

between 2008 and 2023 (Figure 14A). Regarding the molecular subtypes of breast cancer analyzed, most studies included samples from all subtypes. Specifically, 8 studies focused on the luminal subtype, 3 on TNBC, 1 on both luminal and HER2 subtypes, 1 on both luminal and TNBC subtypes, and 11 did not specify the molecular subtype (Figure 14B).

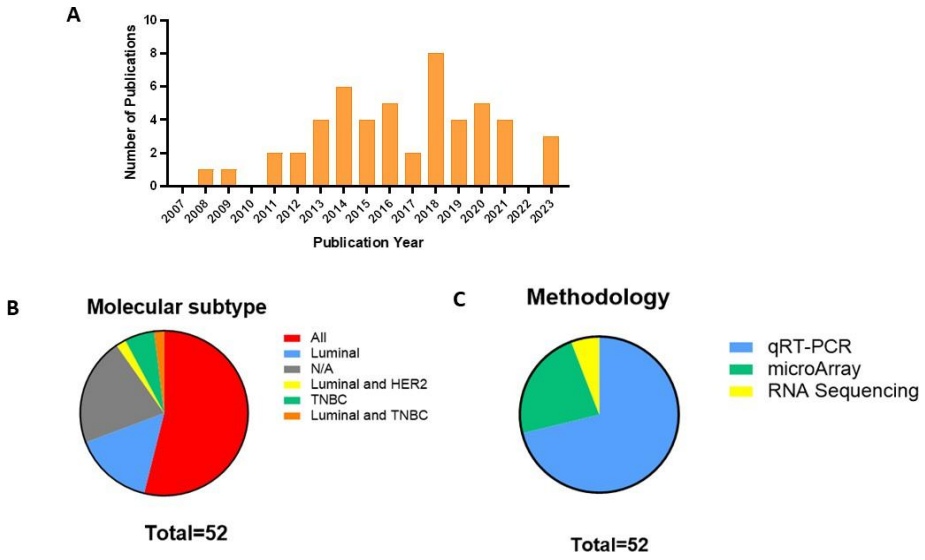



Figure 14. A) Distribution of the publication per year of the records included in the systematic review. B) Molecular subtype of breast cancer sample and C) methodology employ to assess the microRNAs involve in LNM.

The majority of studies employed RT-qPCR for miRNA analysis (37 records), followed by 12 studies using microarray and 3 using RNA sequencing (Figure 14C). Notably, all studies utilizing microarray or RNA sequencing validated their findings, with the exception of the study by Minemura et al., [214].

The direction of regulation (upregulation or downregulation) of these miRNAs is detailed in Figure 15, with a 2-item cut-off for liquid biopsy

samples and a 3-item cut-off for all biological samples. Here, 24 of these studies were performed on fresh tissue, 11 on Formalin-Fixed Paraffin-Embedded (FFPE), 12 on liquid biopsy samples (7 from serum, 4 from plasma, and 1 in exosomes from serum) and 5 articles used more than one type of sample for their study (3 combined tissue and serum and 2 combined tissue and plasma).

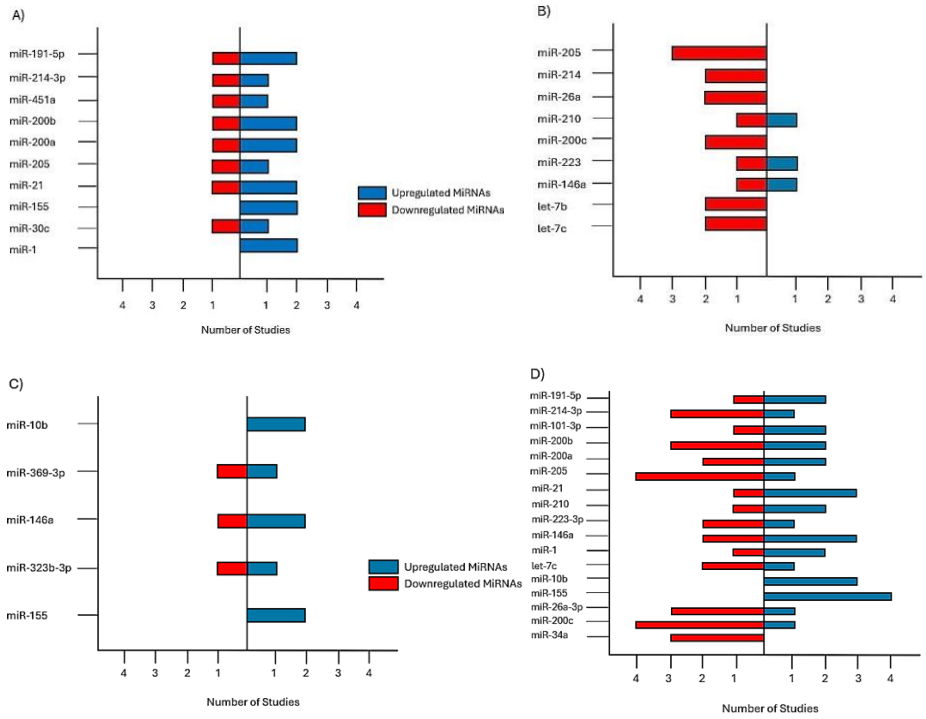



Figure 15. Pyramidal graph showcasing the direction of expression of miRNAs in A) Tissue B) Formalin-fixed, paraffin-embedded (FFPE Tissue C) Liquid Biopsy (Blood, Serum, Plasma) in at least 2 independent studies, and D) in all biological samples in at least 3 independent studies.

| Article | miRNA | Breast Molecular Subtype | Patient number (validation patients) | Controls number (validation patients) | Sample | Method | Validated | Treatment | Regulation |
|---|---|---|---|---|---|---|---|---|---|
| **Chen et al.,2018**[215] | miR-191-5p, miR-214-3p, miR-451a, miR-489 | All | 108 | 63 | Tissue | microArray | qRT-PCR | N/A | Up-regulated miR-191-5p; Down-regulated: miR-214-3p, miR-451a, miR-489 |
| **Zhou et al., 2014**[216] | miR-215 | All | 88 | 55 | Tissue | qRT-PCR | No | Surgery | Down-regulated |
| **Minemura et al., 2015**[214] | miR-1, miR-200a, miR-200b, miR-429, miR-206, miR-155, miR-152, let-7d, let-7c, miR-214, let-7g, | All | 141 | 22 | FFPE tissue | MicroRNA PCR array | No | Adjuvant endocrine therapy afterthe surgery, and tamoxifen and aromatase inhibitors | Up-regulated miR-1, miR-155, miR-152, let-7d, let-7c, miR-214, let-7g, miR-98, miR-204, miR-495, let-7f, miR-497, let-7a, miR- |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | miR-98, miR-204, miR-495, let-7f, miR-497, let-7a, miR-27b, miR-100, miR-130ª | | | | | | | | 27b, miR-100, miR-130a; Down-regulated: miR-200a, miR-200b, miR-429, miR-206, miR-1 (vs stage IV) |
| **Zhang et al., 2018**[217] | miR-1247-5p | Luminal and HER2 | 52 | 60 | FFPE tissue | qRT-PCR | No | Surgery | Down-regulated |
| **Ibrahim et al., 2020**[218] | miR-10b, miR-21 | All | 19 | 11 | Plasma | qRT-PCR | No | Neo-adjuvant chemotherapy and Taxol | Up-regulated |
| **Si et al., 2013**[219] | miR-92a, miR-21 | N/A | 20 | 28 | Tissue and serum | qRT-PCR | No | N/A | Down-regulated miR-92a; Up-regulated miR-21 |
| **Escuin et al., 2021**[220] | miR-339-5p, miR-133a-3p, miR-326, miR- | Luminal | 12 | 4 | Plasma | RNA sequencing | qRT-PCR | Surgery | Down-regulated miR-339-5p, miR-133a-3p, |

| | | |
|---|---|---|
| | 331-3p, miR369-3p, miR-328-3p, miR-26a-3p, miR-139-3p, miR-493-3p, miR664a-5p, miR-101-3p, miR-146a-5p, miR-144-3p, miR-323b-3pmiR-1307-3p and miR-423-3p, mR-376c-3p, miR-1, miR.1908-5p, miR-744-5p, miR-584-5p, miR-6721-5p, miR-432- | miR-326, miR-331-3p, miR369-3p, miR-328-3p, miR-26a-3p, miR-139-3p, miR-493-3p, miR664a-5p, miR-146a-5p, miR-323b-3pmiR-1307-3p, miR-423-3p, miR-376c-3p, miR-1, miR-1908, miR-744-5p, miR-584-5p, miR-6721-5p, miR-432-5p, miR-28-3p; Up-regulated miR- |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 5p, miR-28-3p, miR-29b-3p | | | | | | | | 101-3p, miR-144-3p, miR-29b-3p |
| **Antolín et al., 2015**[221] | miR-141 | All | 37 | 20 | Blood | qRT-PCR | | Surgery | Up-regulated |
| **Escuin et al., 2023**[222] | 643a-3p, miR-223 | Luminal | 12 | 18 | Plasma and tissue | RNAsequencing | No | Surgery | Up-regulated |
| **Huang et al., 2009**[223] | 101-3p, miR-144-3p | All | 19 | 21 | Tissue | qRT-PCR | No | Surgery | Up-regulated |
| **Liu et al., 2021**[224] | miR-367 | All | 31 | 32 | Serum | qRT-PCR | No | Surgery | Down-regulated |
| **Avery-Kiejda et al., 2014**[225] | let-7a, let-7b, let-7c, miR-100, miR-101, miR-10a, miR-125b, miR-126-5p, miR-126-3p, miR-130a, miR-135b, miR-136, miR-143, | TNBC | 16 | 15 | FFPE tissue | microArray | qRT-PCR | N/A | Down-regulated let-7a, let-7b, let-7c, miR-100, miR-101, miR-10a, miR-125b, miR-126-5p, miR-126-3p, miR-130a, miR-136, miR-143, miR-195, miR- |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | miR-195, miR-1977, miR-199a, miR-199a-3p, miR-205, miR-214, miR-26a-5p, miR-26b, miR-29c, miR-320c, miR-34a, miR-497, miR-210, miR-135b | | | | | | | | 1977, miR-199a, miR-199a-3p, miR-205, miR-214, miR-26a-5p, miR-26b, miR-29c, miR-320c, miR-34a, miR-497; Up-regulated miR-210, miR-135b |
| **Dong et al., 2015**[226] | miR-124 | All | 59 | 74 | Tissue | qRT-PCR | No | Surgery | Down-regulated |
| **Xu et al., 2016**[227] | miR-200b, miR-200a, miR-141, miR-429 | All | 40 | 49 | Tissue | qRT-PCR | No | Surgery | Down-regulated miR-200b, miR-200ª; Up-regulated miR-141, miR-429 |
| **Li et al., 2013**[228] | miR-206 | Luminal | 79 | 49 | Tissue | qRT-PCR | No | Surgery | Down-regulated |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Li et al.,2021**[229] | miR-3662, miR-146a, and miR-1290 | All | 40 | 20 | Serum exosomes | qRT-PCR | No | Surgery | Up-regulated |
| **Okuno et al., 2021**[230] | miR-98, miR-22, miR-223 | Luminal | 10 (31) | 10(69) | FFPE tissue | microArray | qRT-PCR | Surgery | Up-regulated |
| **Rask et al., 2014**[231] | miR-486-5p, miR-369-5p, miR-340, miR-139-5p, miR-504, miR-634, miR-509, miR-551a, miR-223, miR-16-1-3p, miR-181, miR-934, miR-455-3p | Luminal | 20(6) | 23(6) | FFPE tissue | microArray | qRT-PCR | Surgery | Down-regulated miR-139, miR-486, miR-369-5p, miR509, miR-223, miR-455-3p; Up-regulated miR-21, miR-340, miR-504, miR-634, miR-551a, miR-16-1-3p, miR-181d. miR-934 |
| **Fang et al., 2016**[232] | miR-199b-5p | Luminal | 62 | 69 | Tissue | qRT-PCR | No | Surgery | Down-regulated |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Jurkovicova et al., 2017**[233] | miR-17 and miR-20ª | Luminal | 42 | 85 | Plasma | qRT-PCR | No | Surgery | Down-regulated |
| **Savad et al., 2012**[234] | miR-205, and miR342 | All | 21 | 38 | Tissue | qRT-PCR | No | Surgery | Down-regulated |
| **Du et al., 2021**[235] | miR-92b-3p | N/A | 72 | 40 | Serum | qRT-PCR | No | Surgery | Up-regulated |
| **Chekhun et al.,2023**[236] | miR-182, -27a, -29b, and -34ª | Luminal and TNBC | 15 | 35 | Tissue | qRT-PCR | No | Surgery | Up-regulated miR-29b; Down-regulated miR-182, miR-27a, miR-34a |
| **Tan et al., 2016**[237] | miR-155, miR-34a | N/A | 21 | 19 | Serum and tissue | qRT-PCR | No | Surgery | Up-regulated miR-155; Down-regulated miR34a |
| **Wu et al., 2020**[238] | miR-21, miR-210 | N/A | 37 | 31 | Tissue | qRT-PCR | No | Surgery | Up-regulated |
| **Sun et al., 2019**[239] | miR-200a-3p, miR-96-5p, miR-1-3p, miR-486-3p , | N/A | 56(25) | 56(21) | Tissue | RNA sequencing | qRT-PCR | Surgery | Up-regulated miR-200a-3p and miR-96-5p, miR-200b-3p, |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | miR-200b, miR-196a-5p, miR223-3p, miR-145-5p, miR-205-5p | | | | | | | | miR-205-5p, miR-196a-5p; Down-regulated miR-1-3p and miR-486-3p, miR-223-3p, miR-145-5p |
| **Guo et al., 2018**[240] | miR-1915-3p, miR455-3p | N/A | 36(15) | 58 (15) | Serum | Affymetrix miRNA profiling array | In vitro | Surgery | Up-regulated miR-1915-3p; Down-regulated miR455-3p |
| **Yin et al., 2023**[241] | miR-338-3p | N/A | 16 | 28 | Serum | qRT-PCR | No | Surgery | Down-regulated |
| **Wang et al.,2018**[242] | miR-330-3p | Luminal | 79 | 154 | Tissue | qRT-PCR | No | Surgery | Up-regulated |
| **Cao et al., 2016**[243] | miR-409-3p | N/A | 103 | 87 | Tissue | qRT-PCR | No | N/A | Down-regulated |
| **Zhang et al., 2018**[244] | miR-597 | N/A | 55 | 135 | Tissue | qRT-PCR | No | Surgery | Down-regulated |
| **Elango et al., 2020**[245] | miR-155-5p, miR-150-5p, miR-146a-5p, | All | 44(32) | 44(32) | FFPE tissue | microArray | qRT-PCR | N/A | Up-regulated miR-155-5p, miR-150-5p, |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | miR-142-5p, miR-200a-3p, miR-200b-3p, miR-200c-3p, miR205-5p, miR-210-3p, miR-214-3p, miR-141-3p, miR-127-3p, miR-125a-5p, and et-7c-5p | | | | | | | | miR-146a-5p, miR-142-5p; Down-regulated miR-200a-3p, miR-200b-3p, miR-200c-3p, miR205-5p, miR-210-3p, miR-214-3p, miR-141-3p, miR-127-3p, miR-125a-5p, let-7c-5p |
| **Yan et al.,2008**[246] | miR-21 | N/A | 8(49) | 8(64) | Tissue | microArray | qRT-PCR | Surgery | Down-regulated |
| **Wu et al.,2021**[247] | miR-432 | All | 37 | 80 | Tissue | qRT-PCR | In vitro | Surgery | Down-regulated |
| **M'hamed et al., 2017**[248] | miR-146a, miR-26a, miR-10b | TNBC | 13 | 8 | FFPE tissue | qRT-PCR | No | N/A | Down-regulated miR-146a, miR-26a; Up- |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | regulated miR-10b |
| **Shahabi et al., 2019**[249] | miR-140 | All | 64 | 46 | Tissue | qRT-PCR | No | N/A | Down-regulated |
| **Berber et al., 2014**[250] | miR-205 and miR-200c: | TNBC | 16 | 16 | FFPE tissue | qRT-PCR | No | Surgery | Down-regulated |
| **Wang et al., 2014**[251] | miR-542-5p, miR-200a, miR-200b, miR-564, miR-451, miR-30c, miR-191-3p, miR-142-5p, miR-185-5p, miR-339-5p, miR-3923 | All | 6 | 6 | Tissue | microArray | qRT-PCR | N/A | Up-regulated miR-185-5p, miR-542-5p, miR-200a, miR-564, miR-451, miR-30c, miR-200b, miR-191-3p, miR-142-5p; Down-regulated miR-339-5p, miR-3923 |
| **Wang et al., 2013**[252] | miR-9 and miR-200c | N/A | 47 | 21 | Tissue | qRT-PCR | No | Surgery | Up-regulated |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Xie et al., 2018**[253] | miR-30b-5p, miR-148a-3p, miR-29c-3p, miR-29c-5p, miR-26a-5p, miR361-3p, miR-645, miR-652-5p and miR-934 | All | 12(122) | 12(196) | Tissue | microArray | qRT-PCR | Surgery | Up-regulated miR-645, miR-652-5p, miR-934; Down-regulated miR-30b-5p, miR-148a-3p, miR-29c-3p, miR-29c-5p, miR-26a-5p, miR361-3p |
| **Wang et al., 2014**[254] | miR-127 | All | 18 | 26 | Tissue | qRT-PCR | In vitro | Surgery | Down-regulated |
| **Zheng et al., 2015**[255] | miR-106b | All | 96 | 77 | Tissue and plasma | qRT-PCR | No | Surgery | Up-regulated |
| **Gao et al., 2016**[256] | miR-34c | All | 78 | 29 | Serum | qRT-PCR | No | N/A | Down-regulated |
| **Shiino et al., 2019**[257] | miR-629-3p, miR-4710 | All | 145(146) | 315(314) | Serum | microArray | qPCR | Neo-adjuvant chemotherapy | Up-regulated miR-629-3p, Down-regulated miR-4710 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Krell et al., 2012** [258] | miR-151-5p | All | 97 | 97 | FFPE tissue | qRT-PCR | No | Chemotherapy | Down-regulated |
| **Chen et al., 2013** [259] | miR-10b, miR-373 | All | 35 | 25 | plasma | qRT-PCR | No | Surgery | Up-regulated |
| **Chun et al., 2011** [260] | let-7b | All | 15 | 22 | FFPE tissue | qRT-PCR | No | N/A | Down-regulated |
| **Lukianova et al., 2019** [261] | miR-155, -320a | All | 30 | 89 | Serum and tissue | qRT-PCR | No | Surgery | Down-regulated miR-320a; Up-regulated miR-155 (both in tissue) |
| **Alunni-Fabbroni et al., 2018** [262] | miR-200b | All | 36 | 12 | Blood | qRT-PCR | No | Chemotherapy | Down-regulated |
| **Chernyy et al., 2018** [263] | miR-155 and miR-222 | All | 30 | 50 | Tissue | qRT-PCR | No | Neo-adjuvant chemotherapy | Up-regulated |
| **Smeets et al., 2011**[264] | miR-195, miR-191, miR-132, miR-203, miR-431, miR-16, | All | 48 | 48 | Tissue | microArray | No | N/A | Up-regulated miR-431; Down-regulated miR-195, miR-191, |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | miR-30c, miR-30a | | | | | | | | miR-132, miR-203, miR-16, miR-30c, miR-30a |
| **Niedźwiecki et al., 2019** [265] | miR-200c | All | 14 | 32 | Serum | qRT-PCR | No | Surgery | Down-regulated |

Table 6. Principal characteristics of the studies included in the systematic review on the microRNA dysregulation in LNM breast cancer.

1.2 FINDINGS FROM THE META-ANALYSIS

### 1.2.1 MICRORNAS INVOLVED IN LNM DIAGNOSIS AND PROGNOSIS

The miRNAs identified as potential biomarkers for lymph node metastasis in at least three independent studies, that was our cut-off for being included in the meta-analysis, comprise miR-191, miR-214, miR-101, miR-200a, miR-200b, miR-200c, miR-205, miR-21, miR-210, miR-223, miR-146a, miR-1, let-7c, miR-10b, miR-155, miR-36a, and miR-34a. We classified these resulting microRNAs into potential diagnostic or prognostic value, depending on whether the study from which they were derived analyzed their expression levels in lymph nodes against localized breast cancer or against distally metastatic or stage IV breast cancer (Table 7 and Table 8).

Among these, consistent expression patterns were observed for miR-155 (4 studies reporting upregulation and none reporting downregulation), miR-34a (3 down-regulated vs. 0 up-regulated), and miR-10b (3 up-regulated vs. 0 down-regulated). Additionally, a context-dependent expression pattern was noted for miR-205 (4 down-regulated vs. 1 up-regulated) and miR-200c (4 down-regulated vs. 1 up-regulated). However, miR-191, miR-214, miR-101, miR-200a, miR-200b, miR-210, miR-223, miR-146a, miR-1, let-7c, and miR-36a showed inconsistent expression directions across studies, leading to their exclusion as reliable biomarkers. Because of that, we finally chose miR-155, miR10b, and miR-34a to continue with further analysis for our meta-analysis.

### 1.2.2 POOL DIAGNOSITC VALUES

The meta-analysis was conducted on miRNAs that demonstrated consistent dysregulation in the direction of expression across at least three independent studies that met predefined inclusion criteria and provided adequate data on effect sizes and population samples for further analysis and interpretation (miR-155, miR-34a and miR-10b). However, not all studies included data on area under the curve (AUC), sensitivity, specificity, or fold change, particularly in those considered for diagnostic assessment and prognostic evaluations such as Overall Survival (OS) or Hazard Ratio (HR). The information about of AUC, OS and HR of the studies included in the meta-analysis was recruited in Supplementary Table II. Due to insufficient coincidence of valuable parameters, we were not able to perform the meta-analysis to analyze the prognosis value of the microRNAs.

To assess the diagnosis value of the microRNAs included, Fold Change was selected as the primary parameter (Tables 7 and 8), since the coincidence between articles for AUC, HR, and OS was too low. Also, as a consequence of insufficient data on Fold Change in the records included, miR-10b was also excluded of the meta-analysis. The forest plots for miR-155 and miR-34a (Figures 16A and 16B) revealed consistent dysregulation in expression direction. miR-155 was uniformly up-regulated across all studies, with an average Log2 Fold Change of 1.50, while miR-34a was consistently down-regulated, with an average Log2 Fold Change of -0.53. Both miRNAs demonstrated a coherent diagnostic value for lymph node metastasis in breast cancer compared to localized stages.

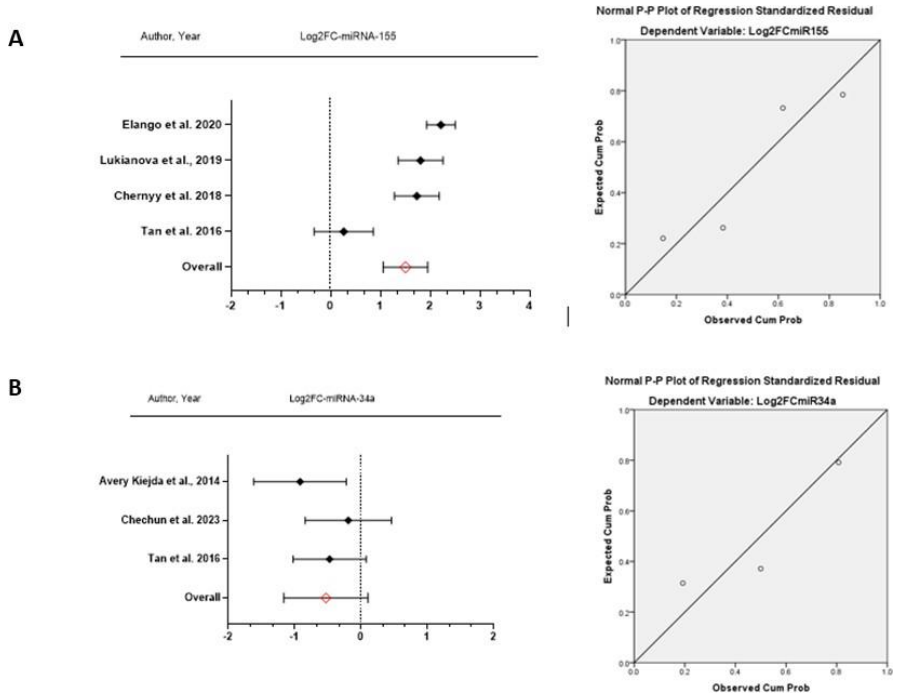



Figure 16: Forest plot of included studies assessing the fold change and plot of regression generated after Eager's test addressing A) publication bias in all biological samples in all included studies for miR-155 and B) miR-34a in LNM diagnosis (red diamond represents the pooled effect and error bars represent the 95% CI).

| Article | miRNA | Direction | Fold Change | Article | miRNA | Direction | Fold Change | Article | miRNA | Direction | Fold Change |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Cheng et al., 2018** | miR-191-5p | Up-regulated | > 1.7 | **Chen et al., 2013** | miR-10b | Up-regulated | 4.44 | **Chernyy et al., 2018** | miR-222 | Up-regulated | 1.46 |
| | miR-489 | Down-regulated | < 1.7 | | miR-373 | Up-regulated | 4.38 | | miR-155 | Up-regulated | 3.29 |
| | miR-214-3p | Down-regulated | < 1.7 | **Wang et al., 2013** | miR-200c | Up-regulated | 1.94 | **Niedźwiecki et al., 2019** | miR-200c | Down-regulated | 1.3 |
| | miR-451a | Down-regulated | < 1.7 | | miR-9 | Up-regulated | 3.16 | | miR-182 | Down-regulated | -3.6 |
| **Shiino et al., 2018** | miR-629-3p | Up-regulated | 2.62 | **Alunni-Faroni et al., 2018** | miR-200b | Down-regulated | 0.47 | | miR-27a | Down-regulated | -3 |
| | miR-4710 | Down-regulated | 0.33 | | let-7a | Down-regulated | -2.36 | **Checkun et al., 2023** | miR-34a | Down-regulated | -3.5 |
| | miR-200a-3p | Down-regulated | 0.95 | | let-7b | Down-regulated | -2.99 | | miR-29b | Up-regulated | 1.3 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Xu et al., 2016** | miR-429 | Down-regulated | 1.35 | **Avery-Kiejda et al., 2014** | let-7c-5p | Down-regulated | -3.84 | **Tan et al., 2016** | miR-34a | Down-regulated | 0.72 |
| | miR-141-3p | Down-regulated | 1.96 | | miR-100 | Down-regulated | -4.37 | | miR-155 | Up-regulated | 1.22 |
| | miR-200b | Down-regulated | 0.74 | | miR-101 | Down-regulated | -2.61 | **Yan et al., 2008** | miR-21 | Up-regulated | 2.84 |
| **Okuno et al. 2021** | miR-98 | Up-regulated | 2.52 | | miR-10a | Down-regulated | -2.37 | **Escuín et al., 2021** | miR-339-5p | Down-regulated | -1.8 |
| | miR-22 | Up-regulated | 3.37 | | miR-125b | Down-regulated | -5.18 | | miR-133a-3p | Down-regulated | -2 |
| | miR-223 | Up-regulated | 3.57 | | miR-126-5p | Down-regulated | -2.72 | | miR-326 | Down-regulated | -2.2 |
| | miR-369-5p | Down-regulated | -0.134 | | miR-126-3p | Down-regulated | -2.31 | | miR-331-3p | Down-regulated | -2.8 |
| | miR-340 | Up-regulated | 0.252 | | miR-130a | Down-regulated | -3.32 | | miR369-3p | Down-regulated | -1.7 |
| | miR-504 | Up-regulated | 0.116 | | miR-135b | Up-regulated | 14.14 | | miR-328-3p, | Down-regulated | -1.4 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | miR-634 | Up-regulated | 0.081 | miR-136 | Down-regulated | -6.18 | miR-26a-3p | Down-regulated | -2.5 |
| | miR-509 | Down-regulated | -0.105 | miR-143 | Down-regulated | -2.99 | miR-139-3p | Down-regulated | -1.4 |
| | miR-551a | Up-regulated | 0.081 | miR-195 | Down-regulated | -5.7 | miR-493-3p | Down-regulated | -2.1 |
| | miR-223 | Down-regulated | -0.597 | miR-1977 | Down-regulated | -2.28 | miR664a-5p | Down-regulated | -1.1 |
| | miR-16-1-3p | Up-regulated | 0.096 | miR-199a-3p | Down-regulated | -2.5 | miR-101-3p | Up-regulated | 0.9 |
| **Rask et al., 2014** | miR-181 | Up-regulated | 0.168 | miR-199a-5p | Down-regulated | -2.46 | miR-146a-5p | Down-regulated | -1.1 |
| | miR-934 | Up-regulated | 0.145 | miR-199-5p | Down-regulated | -5.66 | miR-144-3p | Up-regulated | 0.8 |
| | miR-455-3p | Down-regulated | -0.542 | miR-205 | Down-regulated | -4.65 | miR-323b-3p | Down-regulated | -1.4 |
| | miR-139 | Down-regulated | −0.184 | miR-210 | Up-regulated | 4.56 | miR-1307-3p | Down-regulated | -1.2 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | miR-486 | Down-regulated | −0.328 | miR-214 | Down-regulated | -2.64 | miR-423-3p | Down-regulated | -0.9 |
| | miR-21 | Up-regulated | 3.5 | miR-26a-5p | Down-regulated | -2.67 | mR-376c-3p | Down-regulated | -1.8 |
| **Jurkovicova et al., 2017** | miR-17 | Down-regulated | 0.3 | miR-26b | Down-regulated | -2.43 | miR-1 | Down-regulated | -1.3 |
| | miR-196a-5p | Up-regulated | 4.53 | miR-29c | Down-regulated | -2.33 | miR-1908-5p | Down-regulated | -1.3 |
| | miR-223-3p | Down-regulated | -2.66 | miR-320c | Down-regulated | -2.15 | miR-744-5p | Down-regulated | -1.1 |
| | miR-145-5p | Down-regulated | -2.57 | miR-34a | Down-regulated | -2.12 | miR-584-5p | Down-regulated | -1 |
| | miR-205 | Up-regulated | 5.95 | miR-497 | Down-regulated | -5.11 | miR-6721-5p | Down-regulated | -1.9 |
| | miR-200b | Up-regulated | 7.82 | miR-145 | Up-regulated | 3.6 | miR-432-5p | Down-regulated | -1.7 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Sun et al., 2019** | miR-200a-3p | Up-regulated | 6.46 | **Minemura et al., 2015** | miR-152 | Up-regulated | 3.3 | **Wang et al., 2014** | miR-28-3p | Down-regulated | -1 |
| | miR-96-5p | Up-regulated | 4.02 | | let-7d | Up-regulated | 3.1 | | miR-29b-3p | Up-regulated | 0.93 |
| | miR-1-3p | Down-regulated | -3.23 | | let-7c | Up-regulated | 3 | | miR-542-5p | Up-regulated | > 1.5 |
| | miR-486-3p | Down-regulated | -2.69 | | miR-214 | Up-regulated | 2.8 | | miR-200a | Up-regulated | > 1.5 |
| **Guo et al., 2018** | miR-1915-3p | Up-regulated | >5 | | let-7g | Up-regulated | 2.6 | | miR-200b | Up-regulated | > 1.5 |
| | miR455-3p | Down-regulated | <5 | | miR-98 | Up-regulated | 2.6 | | miR-564 | Up-regulated | > 1.5 |
| | miR-155-5p | Up-regulated | 4.67 | | miR-204 | Up-regulated | 2.5 | | miR-451 | Up-regulated | > 1.5 |
| | miR-150-5p | Up-regulated | 13.24 | | miR-495 | Up-regulated | 2.5 | | miR-30c | Up-regulated | > 1.5 |
| | miR-146a-5p | Up-regulated | 7.5 | | let-7f | Up-regulated | 2.4 | | miR-191-3p | Up-regulated | > 1.5 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | miR-142-5p | Up-regulated | 60.31 | | miR-497 | Up-regulated | 2.4 | | miR-142-5p | Up-regulated | > 1.5 | |
| | miR-200a-3p | Down-regulated | -53,79 | | let-7a | Up-regulated | 2.4 | | miR-185-5p | Up-regulated | > 1.5 | |
| | miR-200b-3p | Down-regulated | -50.23 | | miR-27b | Up-regulated | 2.3 | | miR-339-5p | Down-regulated | < 1.5 | |
| | miR-200c-3p | Down-regulated | -100.81 | | miR-100 | Up-regulated | 2.1 | | miR-3923 | Dow-regulated | < 1.5 | |
| | miR205-5p | Down-regulated | -117,24 | | miR-130a | Up-regulated | 2 | **Zheng et al., 2015** | miR-106b | Up-regulated | | |
| **Elango et al., 2020** | miR-210-3p | Down-regulated | -6.55 | **Berber et al., 2014** | miR-205 | Down-regulated | 0.591 | | miR-155 | Up-regulated | | 3.5 |
| | miR-214-3p | Down-regulated | -3.79 | | miR-200c | Down-regulated | 0.265 | **Liukanova et al., 2019** | miR-320a | Down-regulated | | 0.5 |
| | miR-141-3p | Down-regulated | -22.52 | | | | | | | | | |
| | miR-127-3p | Down-regulated | -56.28 | | | | | | | | | |

| | | | |
|---|---|---|---|
| | miR-125a-5p | Down-regulated | -2.51 |
| | let-7c-5p | Down-regulated | -2.23 |

Table 7. Direction of the regulation of microRNAs involved in LNM diagnosis and their Fold change value.

| Article | miRNA | Direction | Fold Change |
|---|---|---|---|
| **Minemura et al., 2015** | miR-200a | Down-regulated | |
| | miR-200b | Down-regulated | |
| | miR-429 | Down-regulated | |
| | miR-206 | Down-regulated | |
| | miR-1 | Down-regulated | |
| **Zhang et al., 2018** | miR-1247-5p | Up-regulated | |
| **Escuin et al., 2023** | 642a-3p | Up-regulated | 2.62 |
| | miR-223 | Up-regulated | 3.5 |
| **Dong et al., 2015** | miR-124 | Down-regulated | |
| **Li et al.,2021** | miR-3662 | Up-regulated | 1.52 |
| | miR-146a | Up-regulated | 1.62 |
| | miR-1290 | Up-regulated | 1.611 |
| **Wang et al.,2018** | miR-330-3p | Up-regulated | |
| **Cao et al., 2016** | miR-409-3p | Up-regulated | |
| **Zhang et al., 2018** | miR-597 | Down-regulated | |
| **Wu et al.,2021** | miR-432 | Down-regulated | |

| | | |
|---|---|---|
| **Shahabi et al., 2019** | miR-140 | Down-regulated |
| | miR-196a | Down-regulated |
| **Wang et al., 2014** | miR-127 | Down-regulated |

Table 8. Direction of the regulation of microRNAs involved in LNM prognosis and their Fold Change values.

### 1.3 QUALITY OF INCLUDED STUDIES

The methodological quality of the studies included in the meta-analysis was evaluated using the QUADAS-2 tool. The overall risk of bias and concerns regarding applicability were found to be low. Given the absence of significant bias, all studies were retained for inclusion in the meta-analysis (Figure 17).

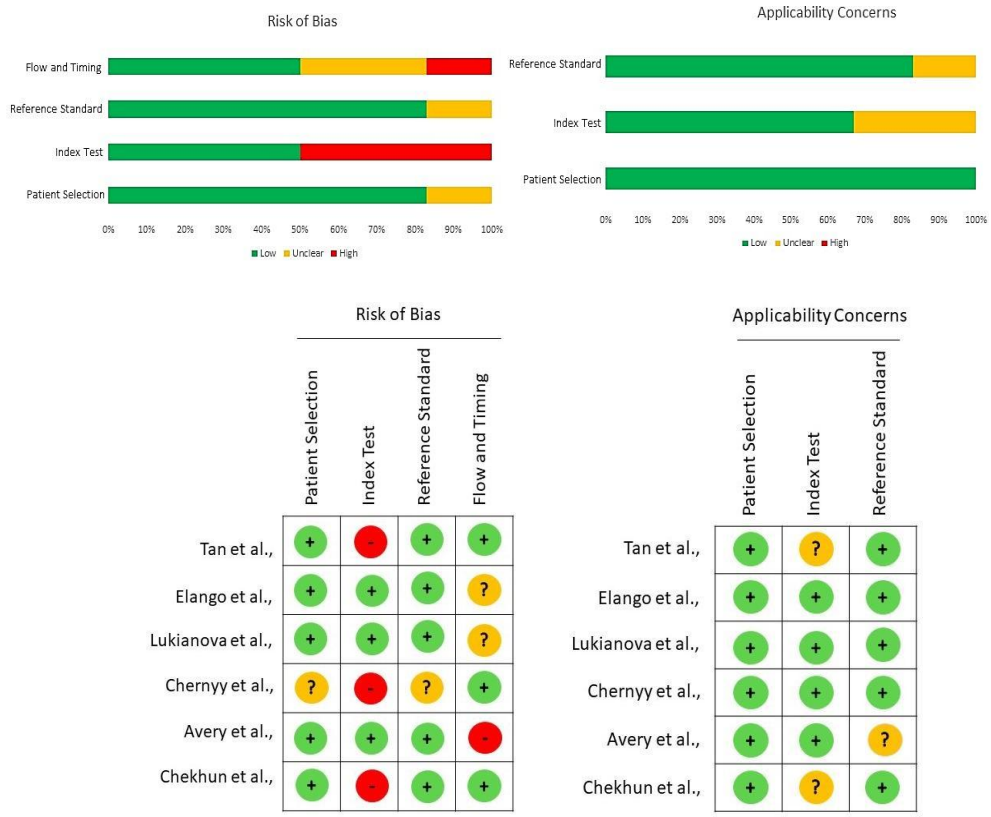



Figure 17. Quality assessment with the QUADAS-2 tool

## 1.4 PUBLICATION BIAS

Publication bias was assessed through Eager's and Begg's Tests. The output of Begg's tests includes the correlation co-efficient Kendall's Tau-b which is associated with p-value. Both tests provided a p-value >0.05, for both miR-155 and miR-34a analyzes, indicating the absence of publication bias (Supplementary Figures I and II).

## 2. IDENTIFICATION OF MICRONAs FOR BREAST CANCER DIAGNOSIS

In the experimental approach, we analyzed the differential expression patterns in the different stages of breast cancer. To this end, the EV-derived microRNAs of 11 early BC patients and 6 metastatic BC patients were sequenced, known as the "discovery cohort". Then, the most differentially expressed microRNAs were validated in 11 metastatic BC patients and 20 early BC patients. Lastly, we made a blind validation for a diagnosis microRNA candidate in a cohort of 100 early BC. The clinical-pathological characteristics of the patients involved in each step are summarized in Tables 3 (early BC patients) and Table 4 (metastatic BC patients) in Methodology Section.

### 2.1 EV ISOLATION AND CHARACTERIZATION

Following the MISEV guidelines[266], we determined the characteristics of our EVs and traced the source of the derived microRNAs.

#### 2.1.1 NANOPARTICLE TRACKING ANALYSIS

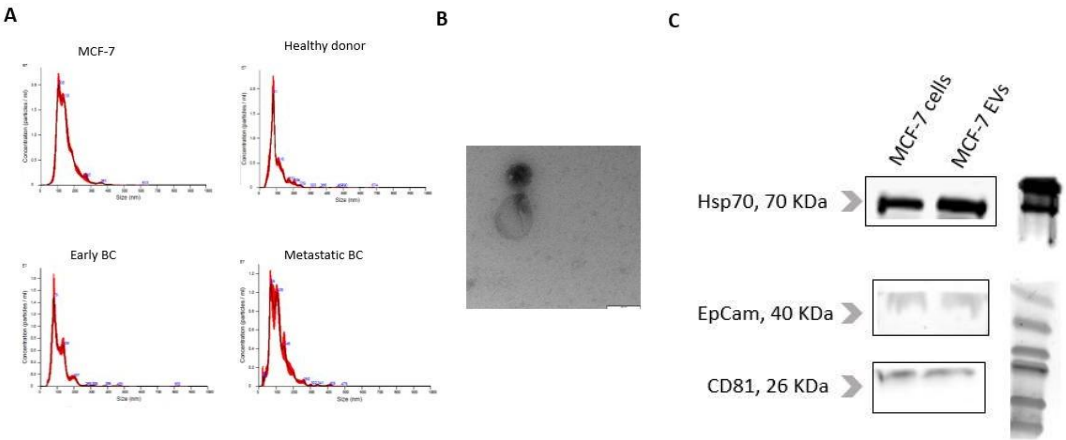
Nanoparticle tracking analysis revealed the concentration of particles/ml with a mode diameter for the different samples (MCF-7 cell line, healthy donor, early BC and metastatic BC) shown in Figure 18A. Vesicles from MCF-7 cell culture cells showed a larger diameter. The smallest but most abundant vesicles were from patients with metastatic breast cancer.

### 2.1.2 TRANSMISSION ELECTRON MICROSCOPY (TEM)

The transmission electron microscopy nanometre resolution allowed us to clearly identify individual extracellular vesicles. The images reveal double-membrane vesicles approximately 100 nm in diameter, containing electron-dense cargo and the presence of a corona surrounding the vesicle (Figure 18B).

### 2.1.3 WESTERN BLOT

The Western blot revealed high protein expression of: Cytosolic proteins recovered in EVs such as Hsp70, Transmembrane or glycosylphosphatidylinositol (GPI)-anchored proteins associated to plasmatic membrane and/or endosomes, such as the tetraspanin CD81, in derived EVs while lower in cell lysates, and EpCAM.




| Sample | Concentracion (particles / ml) | Diameter (nm) |
|---|---|---|
| MCF-7 | 1.76e+09 +/- 9.29e+07 | 105.2 |
| Healthy donor | 8.55e+08 +/- 1.03e+08 | 80.1 |
| Early Breast Cancer | 8.72e+08 +/- 1.09+08 | 78.8 |
| Metastatic Breast Cancer | 9.23e+08 +/- 6.06e+07 | 68.3 |

Figure 18. EV characterization following ISEV recommendations. Fig 18.A shows the concentration and the size distribution of the sample particles. Fig 18. B reveals a

vesicle with an electro-dense cargo under TEM and Fig 18. C reveals the EV characteristics markers by Western blot.

## 2.2 STATISTICS OF THE SEQUENCING DATA

A total of 390,429,166 raw reads were generated, with an average of 15,617,167 reads per sample. After trimming and filtering out low-quality reads, 228,995,664 clean reads were retained, averaging 9,159,826 reads per sample. The mean Q30 value was 95.97%, ranging from 94.2% to 96.58%. Detailed quality metrics for the sequencing data and reads per sample are provided in Supplementary Table III.

## 2.3 DIFFERENTIAL EXPRESSION OF EV-DERIVED MICRORNA BETWEEN NON-CANCER, EARLY AND METASTATIC BREAST CANCER

Analysis of the differential expression of EV-derived microRNAs identified a total of 2,656 known miRNAs by comparison with the miRBase and TarBase databases. Subsequently, a comprehensive pairwise comparison between groups was conducted, using a p-value threshold of <0.05 to identify differentially expressed miRNAs. In total, 137, 39, and 123 differentially expressed microRNAs were found in the comparisons between Early vs. Metastatic, Normal vs. Early, and Normal vs. Metastatic groups, respectively. All the microRNAs and their relative expression can be found in Supplementary Tables IV, V and VI. A $Log_2FC > 0$ indicates that the microRNA is overexpressed while a $Log_2FC < 0$ shows under-expression (Figure 19).

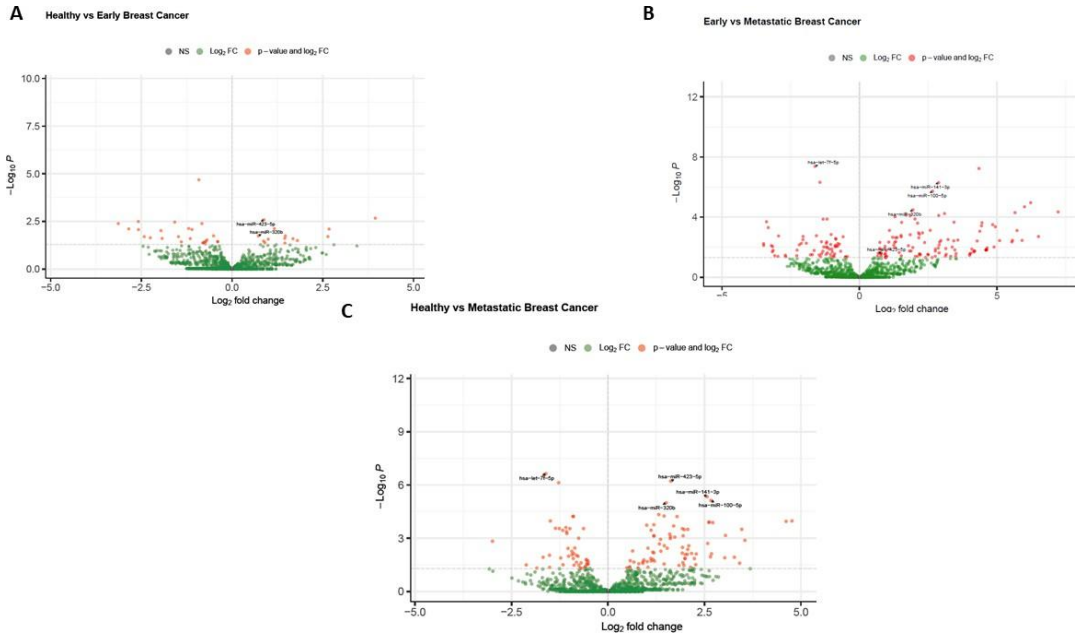



Figure 19. Volcano plots representing the differential expression of microRNAs by two groups: A) Healthy vs early BC; B) Early Vs Metastatic BC and C) Healthy vs Metastatic BC. Log fold change values above 0 indicate overexpression in cancer while values below 0 indicate down-regulation in cancer. Orange dots represent miRNAs with p value < 0.05 and log FC other than 0.

To further refine our results and identify microRNAs with higher statistical significance and a greater likelihood of true differentiation between groups, we applied an additional filter of FDR < 0.01, an even more stringent threshold. After that, we selected those microRNAs that showed a gradual over-expression from healthy to breast cancer, and from localized to metastatic breast cancer, with the aim of developing a diagnostic and prognostic microRNA signature.

This stringent criterion resulted in the identification of three microRNAs (Table 9). We also assessed differential expression between groups (Fig. 20A) and visualized the similarity of microRNA expression using an unsupervised heat map, which primarily clustered the metastatic and early samples in one side, and the control samples in the other side (Fig. 20B).

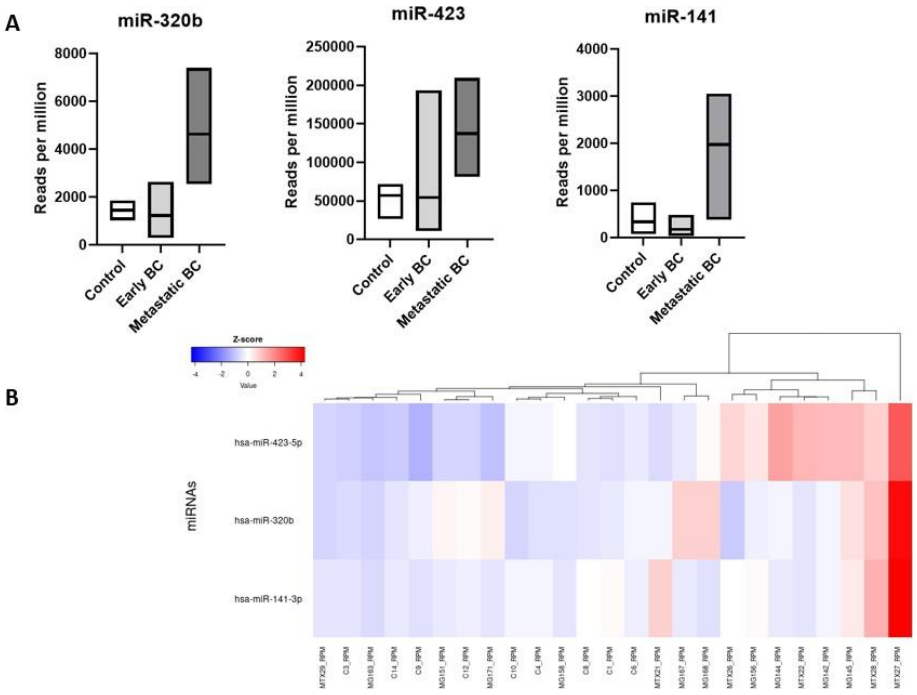



Figure 20. A) Expression of the 3 candidate microRNAs in reads per million in each of the groups and B) Unsupervised heatmap of all samples based on the 3 microRNAs.

| | Normal - Early | | | | Early - Metastatic | | | | Normal - Metastatic | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Log2FC | logCPM | PValue | FDR | Log2FC | logCPM | PValue | FDR | Log2FC | logCPM | PValue | FDR |
| **hsa-miR-423-5p** | 0.880123 | 14.16362 | 0.002597 | 1 | 0.746326 | 14.16362 | 0.018305 | 0.517201 | 1.626448 | 14.16362 | 6.05E-07 | 0.000495 |
| **hsa-miR-141-3p** | 0.16494 | 7.028973 | 0.742345 | 1 | 2.40611 | 7.028973 | 8.49E-06 | 0.003222 | 2.57119 | 7.028973 | 4.54E-06 | 0.00241 |
| **hsa-miR-320b** | 0.723552 | 8.858699 | 0.018082 | 1 | 0.786297 | 8.858699 | 0.017937 | 0.517201 | 1.509812 | 8.858699 | 1.01E-05 | 0.003833 |

Table 9. microRNAs with p-value < 0.05 and FDR < 0.01 with a gradual and significant expression levels increase between groups.

Then we analyze the KEGG molecular pathways of these three microRNAs (Fig. 21). Between the most regulated pathways are: prostate cancer, endometrial cancer, breast cancer and the estrogen signaling pathway, indicating the strong relationship of these microRNAs to hormone regulation.

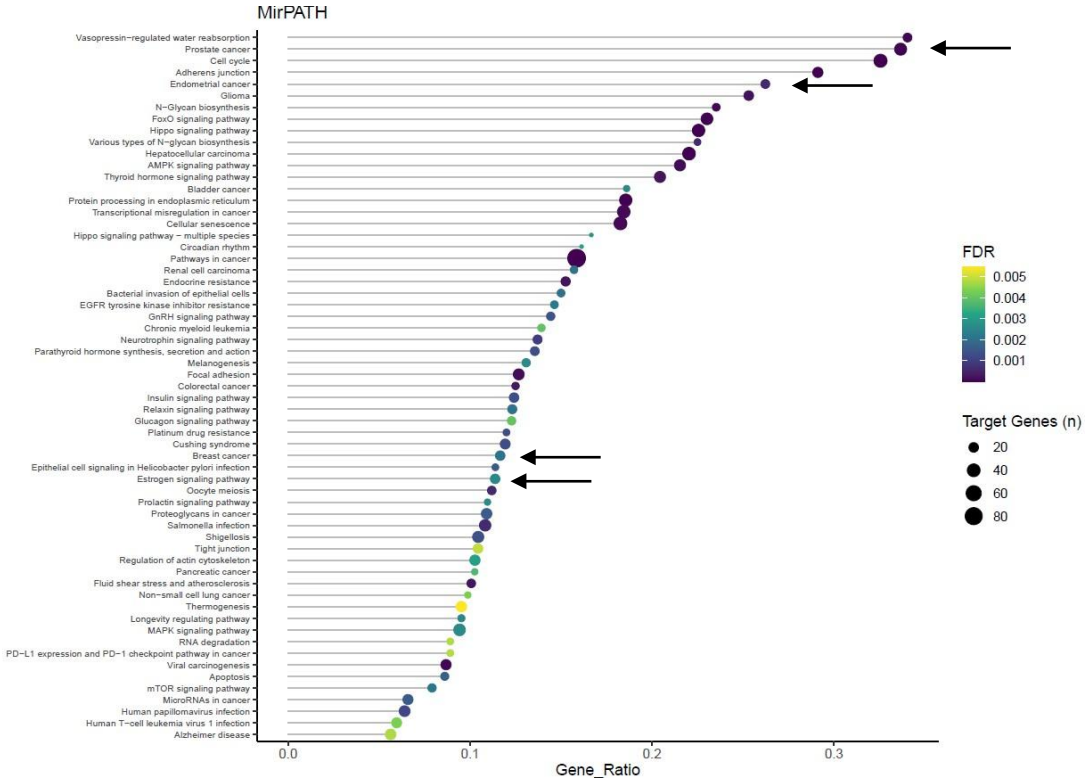



Figure 21. Most significant molecular pathways of the three microRNAs studied.

## 2.4 TARGET GENE PREDICTION OF THE EV-DERIVED MICRORNAs

The target genes of our microRNAs of interest were represented graphically using Cytoscape (Figure 22). The regulatory network consists of a total of 36 nodes and 99 edges. All the genes included in the network were regulated by the three microRNAs. The complete list of all regulated genes is in Supplementary Table VII.

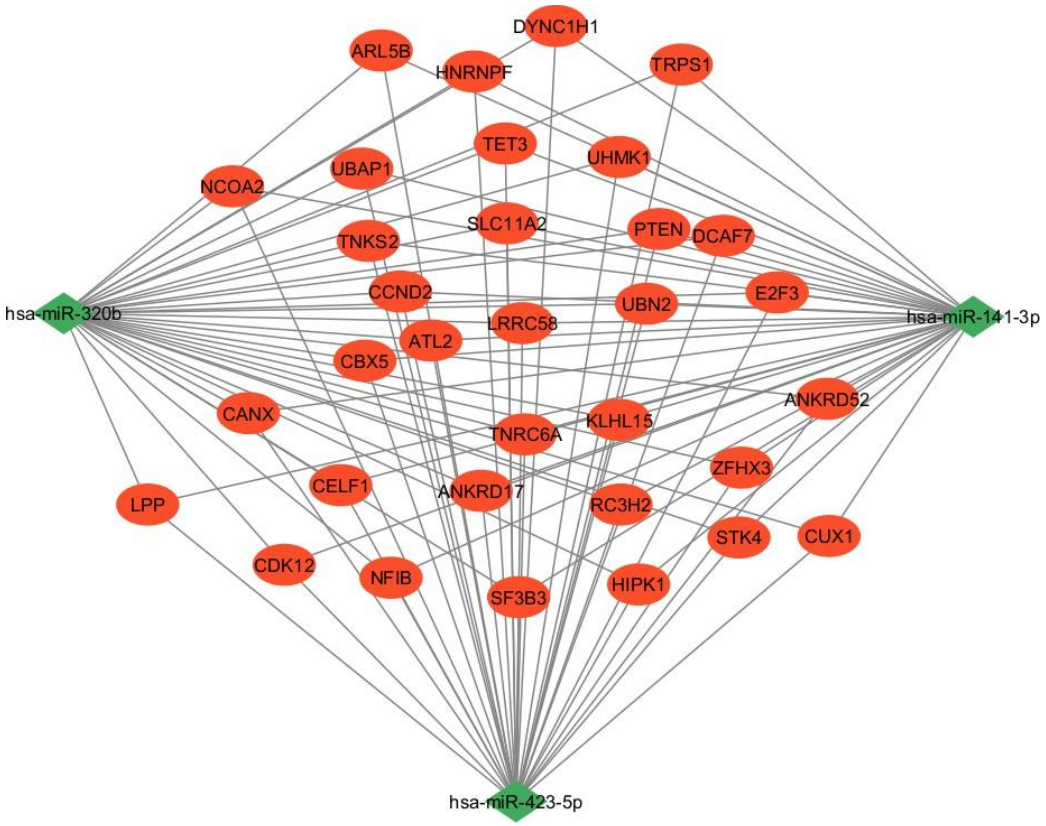



Figure 22. Network composed by the conjunction of the target genes of the three microRNAs.

Then, we analyzed which molecular pathways would be significantly altered if we silenced each of these genes individually in breast tissue using the Variant Interpreter tool, from the Hipathia software. This tool compares the expression values of the GTeX samples for the selected tissue (in this case, breast) against the same samples but making a knock-out (reducing by 0.0001) in the expression of the chosen target gene. The most affected molecular pathways and the silenced gene are shown in Table 10. The gene that was most affected by its silencing in the breast was IGF1R.

| Gene | pathName | Regulation | p.value | FDRp.value | FC | logFC |
|---|---|---|---|---|---|---|
| **IGF1R** | Ovarian steroidogenesis: CYP17A1 | UP | 1.94E-96 | 2.60E-94 | 0.014143466 | -6.1437216 |
| **IGF1R** | Ovarian steroidogenesis: HSD17B2 | UP | 1.94E-96 | 2.60E-94 | 0.0002603 | -11.909055 |
| **IGF1R** | Ovarian steroidogenesis: HSD17B1 | UP | 1.94E-96 | 2.60E-94 | 0.000248744 | -11.973096 |
| **IGF1R** | Proteoglycans in cancer: EIF4B RPS6 | UP | 1.94E-96 | 2.60E-94 | 0.0001078 | -13.179402 |
| **IGF1R** | Progesterone-mediated oocyte maturation: ARAF | UP | 1.94E-96 | 2.60E-94 | 0.0001 | -13.287712 |
| **IGF1R** | Proteoglycans in cancer: MAPK1***** | UP | 1.94E-96 | 2.60E-94 | 0.0001 | -13.287712 |
| **IGF1R** | AMPK signaling pathway: ULK1 | DOWN | 3.68E-96 | 4.07E-94 | 2.0148681 | 1.0106854 |
| **RAC1** | Proteoglycans in cancer: RAC1* | UP | 1.94E-96 | 5.20E-94 | 0.0001 | -13.287712 |
| **STK4** | Pathways in cancer: STK4 | UP | 1.94E-96 | 1.21E-93 | 0.0001 | -13.287712 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **STK4** | Non-small cell lung cancer: STK4 | UP | 1.94E-96 | 1.21E-93 | 0.0001 | -13.287712 |
| **PRLR** | Prolactin signaling pathway: PRLR** | UP | 3.57E-66 | 9.58E-64 | 1.09E-08 | -26.456501 |

Table 10. Most altered pathways resulting from silencing of target genes in breast tissue.

## 2.5 VALIDATION RESULT

For the marker validation phase, we focused on the three microRNAs (miR-423-5p, miR-141-3p and miR-320b) out of the most differentially and gradually expressed. For this step, we employed 20 early breast cancer patients, 11 metastatic breast cancer patients and 8 healthy donors. As 2^-(ΔΔCt) data did not follow a normal distribution (Kolmogorov-Smirnov and Shapiro Wilk tests), we tested the significance between groups two by two using the Mann-Whithey test, and the global significance with the Kruskal-Wallis test. The only microRNA with a global significant difference in expression level between groups was microRNA-423 (Figure 23A). Furthermore, the significance for the miR-423 between control and early stage was almost significant (p-value = 0.0716), so we decided to increase the control and early breast cancer cohorts to validate the diagnosis role of microRNA-423-5p.

In the blind validation phase, we continue only with miR-423-5p, as it was the one that gave the best results in the PCR validation. We test this microRNA in 100 early breast cancer patients and 28 healthy donors. After check the normality of the data, we employed the non-parametric

Mann-Whithey test to see the differences in miR-423-5p expression between healthy donors and early breast cancer patients (p value < 0.0001) (Figure 23B). Then, we performed the ROC curve, with an area under the curve of 0.767. For the contingency tables, we set the expression value cut-off of 2^-(ΔΔCt) at 0.858255, based on the highest sensitivity and specificity values (sensitivity = 0.890, specificity = 0.679).

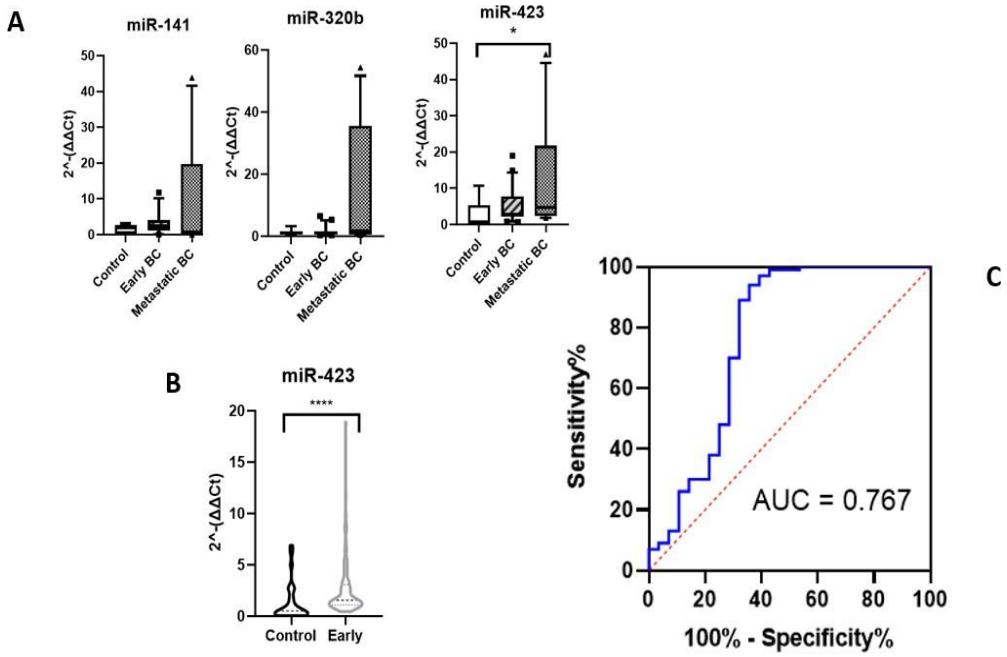



Figure 23. A) Differential expression of the 3 microRNAs in the different groups in the first validation phase; B) Differential expression of miR-423 between healthy and early breast cancer in the blinded validation and C) ROC curve for miR-423 elaborated with the data from the blinded validation.

## 2.6 MICRORNA-423-5p AND CLINIC-PATHOLOGICAL CHARECTERISTICS

After classifying early breast cancer patients into high or low miR-423 expression values, we correlated this with the different clinic-pathological variables (Table 11). High miR-423 values were associated in univariate analysis with the molecular type of breast cancer and with the presence or absence of cancer. Multivariate analysis showed a significant expression only with the presence of cancer variable.

| | | **Univariate** | | | **Multivariate** |
|---|---|---|---|---|---|
| **Variables** | | miR-423-5p | | | |
| | | Up | Down | p value | p value |
| **Age** | <45 | 15 | 1 | 0.238 | - |
| | 45-64 | 60 | 10 | | - |
| | >65 | 14 | 0 | | - |
| **Menopause** | Yes | 61 | 3 | 0.086 | - |
| | No | 28 | 8 | | - |
| **Tumor size** | < 10 | 19 | 1 | 0.471 | - |
| | 10 - 20 | 51 | 8 | | - |
| | 20 - 50 | 15 | 1 | | - |
| | > 50 | 4 | 1 | | - |
| **Tumor stage** | I | 44 | 6 | 0.587 | - |
| | II | 37 | 5 | | - |
| | III | 8 | 0 | | - |
| **Perou Classification** | Luminal A | 35 | 5 | 0.035 | 0.937 |
| | Luminal B | 52 | 5 | | |
| | HER2 | 0 | 1 | | |

| | | | | | |
|---|---|---|---|---|---|
| | Triple negative | 2 | 0 | | |
| **ki67** | <=14 | 30 | 3 | 0.484 | - |
| | 14 - 50 | 42 | 7 | | - |
| | 50 -70 | 11 | 0 | | - |
| | > 70 | 3 | 0 | | - |
| **Affected nodes** | 0 | 62 | 7 | 0.056 | - |
| | 1-3 | 25 | 0 | | - |
| | >3 | 1 | 1 | | - |
| **Adjuvant chemotherapy** | AC+ taxol | 27 | 3 | 0.659 | - |
| | Taxol + trastuzumab | 3 | 1 | | - |
| | None | 59 | 7 | | - |
| **Adjuvant hormonotherapy** | Tamoxifen | 23 | 3 | 0.617 | - |
| | Anastrozol | 25 | 1 | | - |
| | Letrozol | 20 | 2 | | - |
| | Trastuzumab | 9 | 2 | | - |
| | Trastu+Pertuzumab | 2 | 0 | | - |
| | Exemestan | 2 | 1 | | - |
| | Giredestrant | 6 | 1 | | - |
| | None | 2 | 1 | | - |
| **Radiotherapy** | No | 7 | 1 | 0.888 | - |
| | Yes | 82 | 10 | | - |
| **Breast Cancer** | Yes | 89 | 11 | 0.0001 | 0.0001 |
| | No | 9 | 19 | | |

Table 11. Correlation of miR-423 expression with clinic-pathological variables.

## 2.7 MICRORNA LOCATION IN THE VESICLE

We tested the location in the vesicle of our candidate microRNA: miR-423-5p, since the vesicle location has an important biological role. We performed the experiment in three metastatic breast cancer, three early breast cancer and three healthy donor samples. The experiment for the breast cancer cell line MCF-7 can be found in Supplementary Material Figure III. As we can see in Figure 24, the condition with higher levels of miR-423 expression was condition 1, except the Control sample, where the highest level was presented by condition 3. If we check the expression levels by sample (Figure 24B), there are no significant differences within each condition, except condition 3, which is significantly higher in Control sample. These data point that miR-423 is normally transported inside the vesicle and also outside the vesicle in cancer samples. Conversely, in healthy condition, miR-423 seems to be inside, with a cleaner expression than condition 2 thanks to the proteinase K.

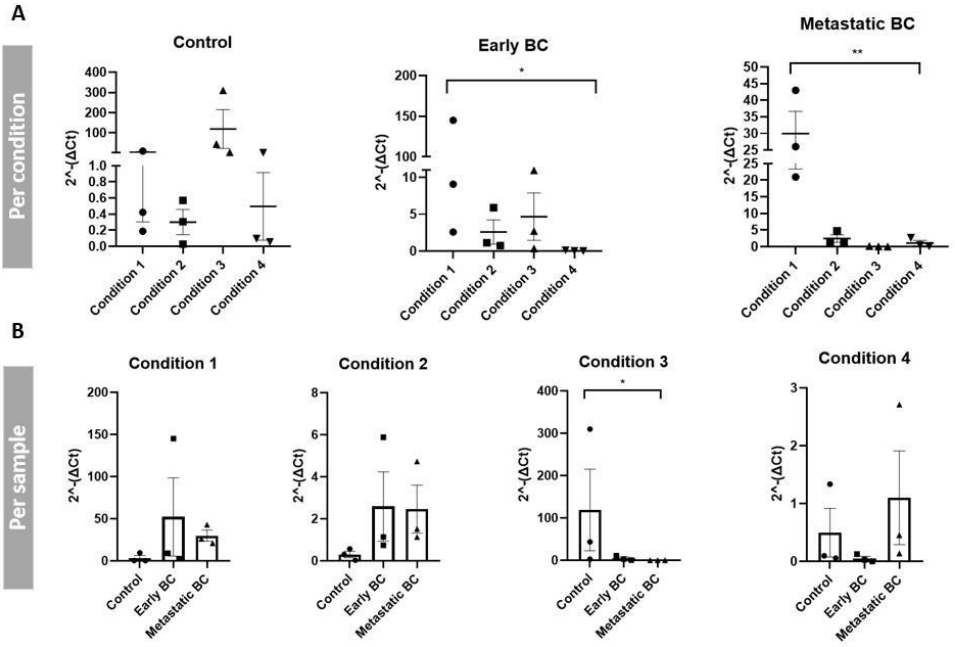



Figure 24. A) miR-423 expression values in the four conditions, separated by group type and B) miR-423 expression values in the three groups, separated by condition.

## 2.8 TUMORAL ORIGIN OF THE VESICLE

After isolation of EV-EpCAM+ using G protein-conjugated Dynabeads, miR-423 expression analysis was performed on healthy donor, early breast cancer and metastatic breast cancer samples. The correct isolation of EVs-EpCAM+ with the Dynabeads was verified by western blotting using the EpCAM antibody and the CD9 vesicle marker.

The samples with the highest levels of miR-423 expression were the metastatic BC patients, followed by the early BC patients and, lastly, the healthy donors, indicating the major EV-EpCAM+ presence in metastatic BC (Figure 25).

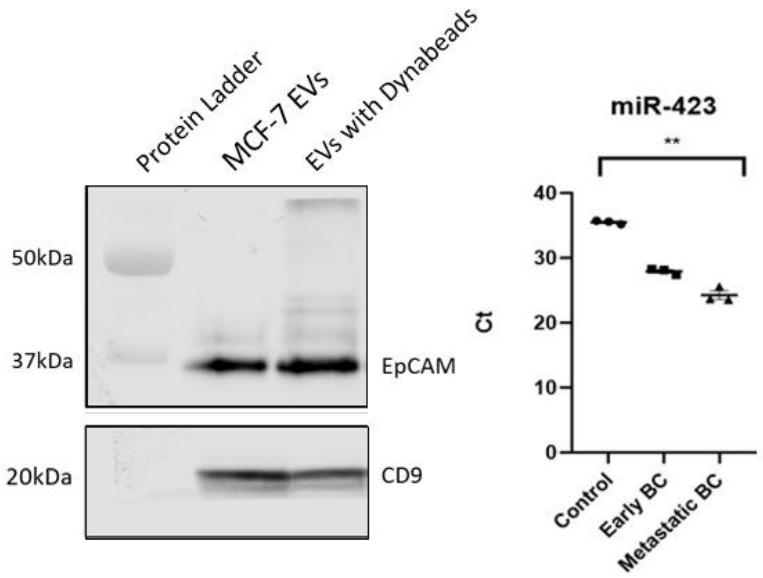



Figure 25. Isolation of EpCAM+ EVs by Dynabeads and analysis of miR-423 expression in this sample.

# CHAPTER VI: DISCUSSION

## 1. DISCUSSION IN THE SYSTEMATIC REVIEW AND META-ANALYSIS RESULTS

As we have seen, the primary challenge in breast cancer is achieving early diagnosis, which can boost the 5-year survival rate to 95% [267]. While mammograms and other imaging methods are valuable, they lack complete accuracy, often requiring a biopsy for confirmation[268]. However, this invasive procedure may not be ideal for all breast cancer patients. A key factor in determining breast cancer prognosis is the detection of lymph node metastases. Unfortunately, only 20-30% of LNM-positive patients avoid distant metastases in the long term, which significantly affects treatment, clinical decisions, and overall prognosis[132]. Currently, Sentinel Lymph Node Biopsy present some disadvantages, as is it is limited to patients undergoing initial surgery or neo-adjuvant chemotherapy, it is an invasive technique and has a false negative rate exceeding 10% in patients with LNM following preoperative systemic therapy[71].

Our work offers a complementary and less invasive biomarker to help mammography and SLNB to diagnose early breast cancer LNM. To do that, we reviewed the existing literature on dysregulated microRNAs in breast cancer patients with lymph node metastasis, with a specific focus on distinguishing these patients from those with localized breast cancer or distant metastasis.

*microRNAs found from the systematic review*

The role of microRNA dysregulation in breast cancer was first identified in 2005[269], and since then, numerous studies have investigated altered microRNA expression, underscoring their significance in cancer progression and metastasis[270,271].

After reviewing the studies included in this systematic review, we identified 84 microRNAs that were differentially expressed in breast cancer patients with LNM. Depending on the reference controls used—either comparing LNM to localized early-stage disease or patients with distant metastasis—these microRNAs were classified based on their 'diagnostic value' or 'prognostic value,' respectively. Finding so many microRNAs differentially dysregulated in this disease, we decided to focus only on those that showed consistency in dysregulation across all included studies.

Among these, only three microRNAs (miR-155, miR-34a, and miR-10) demonstrated consistent directional dysregulation across all studies (cut-off = 3 records). The lack of coherence in the directionality of regulation for most microRNAs could be due to inherent differences among the studies, such as variations in sample types, breast cancer subtypes, or analytical methodologies. For instance, in the systematic review, miR-205 was consistently down-regulated in FFPE tissue but up-regulated in fresh tissue [234]. Similarly, members of the miR-200 family (miR-200a, miR-200b, and miR-200c), which are widely implicated in breast cancer progression [272], exhibited inconsistent regulation across the studies included in this review [214,227,239,245,250,251,262,265].

Due to insufficient effect size data for miR-10b, we were only able to conduct a meta-analysis for miR-155, which was consistently up-regulated, and miR-34a, which was consistently down-regulated.

*Mir-155 and miR-34a*

MiR-155 functions as an oncogene and has been widely shown to promote breast cancer progression[273,274] and contribute to loco-regional metastasis[275]. It is also key regulator of EMT, facilitating the spread of

cancer cells through modulation of various EMT-related pathways and suppresses immune responses by altering immune cell function, allowing breast cancer cells to evade immune surveillance and spread to lymph nodes[276].

In contrast, miR-34a plays a key role in tumor suppression in breast cancer, acting as a crucial component of the p53 tumor suppressor network[277]. MiR-34a has been consistently found to be down-regulated in both breast cancer cell lines and tissues when compared to normal cell lines and adjacent non-tumor tissues[278]. Furthermore, higher miR-34a expression is associated with less aggressive breast cancers [279] and has been linked to drug resistance[280].

In our systematic review, miR-34a was found to be down-regulated in tissue [225,236,237] and serum [237], in the first two cases by qPCR and in the latter by microarray. This shows a consistency of low expression of this microRNA in different samples and by different techniques in case of LNM.

As for miR-155, it was found in tissue [237,245,261,263] and serum [237,261]. Three of the studies used qPCR[237,261,263] while Elango et al.,[245] used microArray. As in the case of mir-34a, consistent expression of miR-155 is evident in different samples and by different techniques.

*Results from the meta-analysis*

This work represents the first comprehensive meta-analysis to assess the diagnostic potential of microRNAs in breast cancer patients with lymph node metastasis. By analyzing Log2 miRNA Fold Change values, our meta-analysis generated a global pooled value of 1.50 for miR-155 and -0.53 for miR-34a, indicating significant dysregulation of these

microRNAs between localized breast cancer and those with loco-regional metastasis. Importantly, the absence of publication bias, as confirmed by Egger's and Begg's tests, strengthens the validity of our findings. Additionally, the application of the QUADAS-2 tool demonstrated a low risk of bias and high applicability, further supporting the strong correlation between the upregulation of miR-155 and the downregulation of miR-34a with LNM in breast cancer. These findings underscore the diagnostic value of miR-155 and miR-34a as biomarkers for detecting lymph node involvement in breast cancer.

*Limitations*

While we made efforts to minimize search bias by including grey literature from SCOPUS, we acknowledge that complete elimination of search bias is not possible. Furthermore, the inherent variability in the studies we included presents limitations that we could not fully control, such as differences in sample size, study timing, and analytical technology. These factors contribute to the heterogeneity observed across studies. A key challenge in this analysis was standardizing different cut-offs, effect sizes, and reference samples, given the diverse methodologies used. These variations in sample types, test methods, and breast cancer subtypes contributed to the heterogeneity of our results and should be considered when interpreting the overall findings of this meta-analysis.

*Future research*

Our study, which followed rigorous inclusion and exclusion criteria to ensure accuracy, identified two microRNAs with consistent dysregulation: miR-155 (up-regulated) and miR-34a (down-regulated). These findings are particularly noteworthy, as both microRNAs were

investigated simultaneously in multiple studies, highlighting the potential involvement of broader microRNA families in breast cancer metastasis. This pattern of dysregulation suggests the need for further exploration into related microRNAs, which could reveal additional diagnostic or prognostic biomarkers.

Moving forward, larger, well-designed studies with standardized methodologies are needed to validate these microRNAs across diverse breast cancer subtypes. Additionally, expanding research to explore the functional roles of these microRNAs in LNM could offer new insights into early detection and therapeutic strategies.

## 2. DISCUSSION IN IDENTIFICATION OF MICRONAs FOR BREAST CANCER DIAGNOSIS RESULTS

Despite being the most frequent and deadliest cancer in women, breast cancer is one of the cancers with the best 5-year survival rate, with 99% when localized. However, the 5-year survival rate drops to 31% when there is distal metastasis to other organs[7].

Currently, mammography is the gold standard for breast cancer screening. However, traditional mammography has notable limitations, including over-diagnosis, where benign tissue irregularities or nonaggressive tumors are misidentified, leading to unnecessary treatments and interventions.[112]. Additionally, the sensitivity of mammograms can vary based on factors like age and breast tissue density[113].

While mammography and other imaging techniques provide valuableinsights, they are not entirely definitive, often requiring a biopsy for confirmation. Unfortunately, this invasive procedure is not always

suitable for breast cancer patients. Detecting minimal residual disease (MRD) and managing intratumoral heterogeneity remain significant challenges, as these factors contribute to treatment resistance and impact prognosis, especially across different molecular subtypes of breast cancer[160,161].

To address these issues and enhance early detection, researchers are investigating emerging non-invasive techniques, such as liquid biopsies, which analyze circulating components like extracellular vesicles and microRNAs. MiRNAs, which broadly regulate gene expression, play a key role in cancer progression and phenotypic changes in tumor cells. Encapsulated within EVs and circulating in the bloodstream, these miRNAs are vital to intercellular communication and offer promising potential as biomarkers for early breast cancer diagnosis and to stratify the patients.

In this study, we sequenced microRNAs derived from EVs in plasma from two cohorts of breast cancer patients—those with localized and metastatic cancer—as well as a cohort of healthy donors, with the aim of identifying a diagnostic and prognostic biomarker to complement imaging techniques.

*microRNA sequencing results*

Our threshold of p value < 0.05 and FDR < 0.01, and a gradual over-expression from healthy to breast cancer, and from localized to metastatic breast cancer identified three EV-derived microRNAs: miR-141-3p, miR-320b and miR-423-5p.

The KEGG analysis for these microRNAs shows that between the most regulated pathways are: prostate cancer, endometrial cancer, breast cancer and the estrogen signaling pathway, indicating the strong

relationship of these microRNAs to hormone regulation. This is coherent with the results obtained from silencing specifically the target genes of these microRNAs, where Hipathia shown among the most important target genes *IGF1R*, upregulating the ovarian steroidogenesis or the progesterone-mediated oocyte maturation. These results together elucidate the importance of the three microRNAs in the hormonal pathways and across different types of hormonal cancers. *IGF1R* is a well-described target in breast cancer[281] and influences tumorigenic phenotypes and drug resistance across all breast cancer subtypes[282]. Another target gene of the microRNAs included in this work is *PRLR*, which is altered also upregulating the prolactin signaling pathway. The expression levels of *PRLR* in breast cancer cells and breast cancer tissues are elevated in most ER+ and ER− tumors. PRL activates downstream signaling pathways and affects endocrine therapy resistance by combining with *PRLR*[283]. All this shows the close link between our three microRNAs and the dysregulation of the hormonal network that occurs in breast cancer. However, direct studies with these microRNAs and genes in vitro are needed to consolidate these results.

*Potential clinical applications*

In the first validation phase, miR-423-5p was the only which shown significant differentially expression across the three cohorts, leading us to discard miR-141-3p and miR-320b as potential breast cancer biomarkers. Due to difficulties to increase the metastatic cohort, we focus on the diagnostic value of this microRNA, but we still working on the recruitment of metastatic breast cancer patients.

In the blind validation phase, miR-423-5p present markedly different expression levels between healthy donors and early breast cancer patients, showing upregulation in the latter group. Furthermore,

correlation analysis between miR-423-5p expression and clinic-pathological characteristics revealed that the only significant variable was the presence or absence of breast cancer. This suggests that miR-423-5p holds strong potential as a complementary diagnostic biomarker.

microRNA-423 is an oncogenic factor which is frequently up-regulated in breast cancer [284,285]. The majority of the studies point to its prognostic role, enhancing cell migration and invasion [284–286], while others study its role in drug resistance, like Zhong et al.[287], who evaluates the EV-derived miRNA levels in BC under chemotherapy and see the upregulation of miR-423-5p.

Also, there are other studies that have revealed their role in other hormone-related cancers. For example, in ovarian cancer, miR-423-5p was found to be down-regulated in both ovarian cancer tissues and plasma[288], while in prostate cancer, cancer-associated fibroblasts secreted exosomal miR-423-5p, promoting chemotherapy resistance[289].

This study represents the first evaluation of miR-423-5p diagnostic potential in breast cancer. However, we are aware that larger and independent cohorts of early breast cancer patients are needed to further validate and consolidate these findings.

*miR-423-5p location in the vesicle*

As we previously said, the corona of the EVs is suspected to significantly influence cancer biology and progression[174]. miRNAs associated with the vesicle membrane or its corona may exhibit different modes of uptake and functional roles in the recipient cells than miRNAs circulating free or in protein complexes. Thus, externally attached miRNAs may have more specific and individual effects, since EVs can contain multiple miRNAs and other components that act together[290]. Furthermore,

miRNAs attached to the outside of vesicles may act as ligands for surface receptors, triggering signaling pathways without needing to be internalised[291].

In our experiment, the data point that miR-423 in cancer samples is transported inside but also outside the vesicle. However, in healthy condition, our target microRNA seems to be fundamentally inside the vesicle. These result can be due to the increment in the expression of miR-423 in cancer condition, being reflect on the increment in the bloodstream as a consequence of the cell injury, inflammation and apoptotic processes that take place in cancer[292], or in protein–miRNA complexes secreted by tumoral cells or cells from the immune system[190], a subject that will address in the next section. These microRNAs outside the vesicle could have a more specific role and more specific recipient cells than miRNAs internalised[290]. On the other hand, in healthy condition, these inflammatory and apoptotic processes are absent, and the only signal inside the vesicle is because of the "normal location" of the miRNA.

*Cellular origin of the extracellular vesicle*

To abroad the miRNA location in the vesicle revealed the importance of knowing the cellular/tissue origin of the vesicle. EVs-derived tumor tissue are released into the blood, where they mix with EVs released by other healthy tissues and EVs released by immune system cells[178].

Consequently, when analyzing the EV cargo in plasma from cancer patients for potential biomarkers, it is crucial to determine whether these EVs originate from tumor cells or immune system cells[180]. This information could provide valuable insights into the underlying biological processes and improve clinical approaches, as therapies

selection.

Our results showed that the samples with the highest amount of EpCAM+ EVs were from metastatic patients, followed by those with early-stage breast cancer, and lastly, the healthy controls, where barely any miRNA amplification signal was detected. Although we cannot assume that these EpCAM+ EVs originate directly from the tumor, we can confirm that they do not come from any immune system cells, as all of them are EpCAM-. The possibility that these EVs come from CTCs or other healthy tissues cannot be ruled out, but their higher abundance in metastatic patients suggests a likely association with this process.

Taken together, all these data suggest that studying EV-derived microRNAs could be a valuable diagnostic and prognostic tool due to its non-invasive nature and the ease of monitoring patients over time. However, further studies are needed to validate the diagnostic strength of our candidate miRNA and to streamline EV and microRNA isolation techniques so they can be applied in clinical practice as a complementary tool to imaging techniques.

# CHAPTER VII. CONCLUSIONS

## CONCLUSIONS OF THE META-ANALYSIS

1. MicroRNAs exhibit differential expression patterns depending on the stage of breast cancer (early, loco-regional, or metastatic).
2. Within the same stage, miRNA expression can be influenced by variables such as tumor type or sample type, which were difficult to control in this study. However, several miRNAs demonstrated consistent regulation across studies, regardless of these variables.
3. The meta-analysis was conducted using the log Fold Change as the effect size, which was normalized, as it was the only common measure across all articles. No significant differences or publication biases were found, and the overall risk of bias, as assessed by QUADAS-2, was low.
4. Only three microRNAs (miR-155, miR-34a, and miR-10b) showed consistent dysregulation in lymph node metastasis (LNM) compared to early stages across all included studies. Due to insufficient data on Fold Change, miR-10b was excluded from the meta-analysis.
5. Forest plots for miR-155 and miR-34a (Figures 15A and 15B) showed consistent dysregulation in expression across studies. miR-155 was uniformly up-regulated, with an average Log2 Fold Change of 1.50, while miR-34a was consistently down-regulated, with an average Log2 Fold Change of -0.53.
6. Both miR-155 and miR-34a demonstrated potential diagnostic value for distinguishing lymph node metastasis from localized breast cancer stages. However, further studies are needed to confirm their diagnostic utility.

## CONCLUSIONS OF THE IDENTIFICATION OF MICRONAs FOR BREAST CANCER DIAGNOSIS

1. Sequencing results reveal distinct patterns of microRNA expression across healthy women, patients with early-stage breast cancer, and those with metastatic breast cancer.
2. Notably, three microRNAs—miR-320b, miR-423-5p, and miR-141-3p—met stringent selection criteria (p-value < 0.05, FDR < 0.01) and exhibited a stepwise increase in expression from healthy donors to patients with metastatic disease.
3. Bioinformatic analyses further highlighted that these microRNAs are implicated in pathways related to other hormone-driven cancers, such as prostate and endometrial cancer, as well as in estrogen and prolactin production pathways.
4. In the initial validation phase, miR-423 emerged as the only microRNA with a globally significant difference in expression between groups.
5. During the blinded validation phase, miR-423 showed markedly distinct expression levels between healthy donors and early-stage breast cancer patients, with an AUC of 0.767. Statistical analyses also revealed a significant correlation between elevated miR-423 levels and breast cancer presence.
6. Additionally, miR-423 was detected in early-stage and metastatic breast cancer samples, both inside and outside of vesicles. In contrast, it was primarily localized inside vesicles in samples from healthy donors.
7. miR-423 was found within EpCAM-positive extracellular vesicles, suggesting a possible tumor origin for these vesicles.

# CHAPTER VIII: REFERENCES

1. World Health Organization. Breast Cancer. 2024. https://www.who.int/news-room/fact-sheets/detail/breast-cancer.

2. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;(February):229-263. doi:10.3322/caac.21834

3. Brinton LA, Gaudet MM GG. Breast cancer. In: Thun MJ, Linet MS, Cerhan JR, Haiman C, Schottenfeld D, eds. Cancer Epidemiology and Prevention. In: *Cancer Epidemiology and Prevention*. 4th ed. Oxford University Press; 2017:861-888.

4. Heer E, Harper A, Escandor N, Sung H, McCormack V, Fidler-Benaoudia MM. Global burden and trends in premenopausal and postmenopausal breast cancer: a population-based study. *Lancet Glob Heal*. 2020;8(8):e1027-e1037. doi:10.1016/S2214-109X(20)30215-1

5. Soerjomataram I, Cabasag C, Bardot A et al. Cancer survival in Africa, Central and South America, and Asia (SURVCAN-3): a population-based benchmarking study in 32 countries. *Lancet Oncol*. 2023;24(1):22-32. doi:10.1016/s1470-2045(22) 00704-5

6. Duggan C, Trapani D, Ilbawi AM et al. National health system characteristics, breast cancer stage at diagnosis, and breast cancer mortality: a population-based analysis. *Lancet Oncol*. 2021;22(11):1632-1642. doi:10.1016/s1470-2045(21)00462-9

7. American Cancer Society. Understanding a Breast Cancer Diagnosis, Atlanta. *Https://WwwCancerOrg/Cancer/Breast-Cancer/About/Types-of-Breast-CancerHtml#References*. Published online 2016:1-36.

8. Abdelwahab-Yousef AJ. Male Breast Cancer: Epidemiology and Risk Factors. *Semin Oncol*. 2017;44(4):267-272. doi:10.1053/j.seminoncol.2017.11.002

9. Collaborative group E hormones and breast cancer., Fomby P, Cherlin AJ. Sex hormones ans breast cancer risk in premenopausal women. *Lancet Oncol*. 2013;14(10):1009-. doi:10.1016/S1470-2045(13)70301-2.Sex

10. Hamajima N, Hirose K, Tajima K, et al. Menarche, menopause, and breast cancer risk: Individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol*. 2012;13(11):1141-1151. doi:10.1016/S1470-2045(12)70425-4

11. Thomas HV., Reeves GK, Key TJA. Endogenous estrogen and postmenopausal breast cancer: A quantitative review. *Cancer Causes Control*. 1997;8(6):922-928. doi:10.1023/A:1018476631561

12. Merrill RM, Sloan A. Risk-adjusted female breast cancer incidence rates in the United States. *Cancer Epidemiol*. 2012;36(2):137-140. doi:10.1016/j.canep.2011.08.004

13. American Cancer Society. Breast Cancer Risk Factors You Cannot Change. Published 2021. https://www.cancer.org/cancer/types/breast-cancer/risk-and-prevention/breast-cancer-risk-factors-you-cannot-change.html

14. Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet.* 2009;358:1389-1399.

15. Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. *Jama*. 2017;317(23):2402-2416. doi:10.1001/jama.2017.7112

16. Chen S, Parmigiani G. Meta-Analysis of BRAC1 and BRAC2

Penetrance. *J Clin Oncol*. 2007;25(11):1329-1333.

17. Breast Cancer Association Consortium. Breast Cancer Risk Genes — Association Analysis in More than 113,000 Women. *N Engl J Med*. 2021;384(5):428-439. doi:10.1056/nejmoa1913948

18. Boraka Ö, Klintman M, Rosendahl AH. Physical Activity and Long-Term Risk of Breast Cancer, Associations with Time in Life and Body Composition in the Prospective Malmö Diet and Cancer Study. *Cancers (Basel)*. 2022;14(8). doi:10.3390/cancers14081960

19. De Cicco P, Catani MV, Gasperi V, Sibilano M, Quaglietta M, Savini I. Nutrition and breast cancer: A literature review on prevention, treatment and recurrence. *Nutrients*. 2019;11(7):1-28. doi:10.3390/nu11071514

20. Picon-Ruiz M, Morata-Tarifa C, Valle-Goffin JJ, Friedman ER, Slingerland JM. Obesity and adverse breast cancer risk and outcome: Mechanistic insights and strategies for intervention. *CA Cancer J Clin*. 2017;67(5):378-397. doi:10.3322/caac.21405

21. Liu Y, Nguyen N, Colditz GA. Links Between Alcohol Consumption and Breast Cancer. *Womens Heal*. 2015;11(1):65-77.

22. Admoun C, Mayrovitz HN. The Etiology of Breast Cancer. In: HN M, ed. *Breast Cancer*. Brisbane. 2022.

23. Jung S, Spiegelman D, Baglietto L, et al. Fruit and vegetable intake and risk of breast cancer by hormone receptor status. *J Natl Cancer Inst*. 2013;105(3):219-236. doi:10.1093/jnci/djs635

24. van den Brandt PA, Schulpen M. Mediterranean diet adherence and risk of postmenopausal breast cancer: results of a cohort study and meta-analysis. *Int J Cancer*. 2017;140(10):2220-2231. doi:10.1002/ijc.30654

25. Du M, Liu SH, Mitchell C, Fung TT. Associations between diet quality scores and risk of postmenopausal estrogen receptor-negative breast cancer: A systematic review. *J Nutr*. 2018;148(1):100-108. doi:10.1093/jn/nxx015

26. Nguyen HL, Geukens T, Maetens M, et al. Obesity-associated changes in molecular biology of primary breast cancer. *Nat Commun*. 2023;14(1). doi:10.1038/s41467-023-39996-z

27. Zhang SM, Lee IM, Manson JE, Cook NR, Willett WC, Buring JE. Alcohol consumption and breast cancer risk in the women's health study. *Am J Epidemiol*. 2007;165(6):667-676. doi:10.1093/aje/kwk054

28. Key TJ, Appleby PN, Reeves GK, et al. Circulating sex hormones and breast cancer risk factors in postmenopausal women: Reanalysis of 13 studies. *Br J Cancer*. 2011;105(5):709-722. doi:10.1038/bjc.2011.254

29. Gompel A. Cigarette smoking and the incidence of breast cancer. *Climacteric*. 2011;14(4):510-511. doi:10.3109/13697137.2011.586892

30. Fortner RT, Sisti J, Chai B, et al. Parity, breastfeeding, and breast cancer risk by hormone receptor status and molecular phenotype: Results from the Nurses' Health Studies. *Breast Cancer Res*. 2019;21(1):1-9. doi:10.1186/s13058-019-1119-y

31. Beral V, Bull D, Doll R, Peto R, Reeves G. Breast cancer and breastfeeding: Collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50 302 women with breast cancer and 96 973 women without the disease. *Lancet*. 2002;360(9328):187-195. doi:10.1016/S0140-6736(02)09454-0

32. Azubuike SO. Relationship between parity and breast cancer risk: A critical review of evidence (with focus on sub-Saharan Africa). *Int J Noncommunicable Dis*. 2023;8(2):66-74. doi:10.4103/jncd.jncd

33. Islami F, Liu Y, Jemal A, et al. Breastfeeding and breast cancer risk by receptor status-a systematic review and meta-analysis. *Ann Oncol*. 2015;26(12):2398-2407. doi:10.1093/annonc/mdv379

34. Manson JE, Chlebowski RT, Stefanick ML et al. Menopausal Hormone Therapy and Health Outcomes During the Intervention and Extended Poststopping Phases of the Women's Health Initiative Randomized Trials. *Jama*. 2013;10(13):1353–1368. doi:10.1001/jama.2013.278040

35. Weigelt B, Reis-Filho JS. Histological and molecular types of breast cancer: Is there a unifying taxonomy? *Nat Rev Clin Oncol*. 2009;6(12):718-730. doi:10.1038/nrclinonc.2009.166

36. Vuong D, Simpson PT, Green B, Cummings MC, Lakhani SR. Molecular classification of breast cancer. *Virchows Arch*. 2014;465(1):1-14. doi:10.1007/s00428-014-1593-7

37. Aster JC. *Robbins and Cotran Pathologic Basis of Disease.* 10th ed. Saunders/Elsevier; 2010.

38. Society AC. Special Section : Breast Carcinoma In Situ. *Cancer Facts Fig*. Published online 2015:26-36. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2015/special-section-breast-carcinoma-in-situ-cancer-facts-and-figures-2015.pdf

39. Coyne JD. DCIS and LCIS with multinucleated giant cells - A report of 4 cases [5]. *Histopathology*. 2007;50(5):669-671. doi:10.1111/j.1365-2559.2007.02639.x

40. Mastracci TL, Tjan S, Bane AL, O'Malley FP, Andrulis IL. E-cadherin alterations in atypical lobular hyperplasia and lobular carcinoma in situ of the breast. *Mod Pathol*. 2005;18(6):741-751. doi:10.1038/modpathol.3800362

41. Eheman CR, Shaw KM, Ryerson AB, Miller JW, Ajani UA, White MC. The changing incidence of in situ and invasive ductal and lobular breast carcinomas: United States, 1999-2004. *Cancer Epidemiol Biomarkers Prev*. 2009;18(6):1763-1769. doi:10.1158/1055-9965.EPI-08-1082

42. Makki J. Diversity of breast carcinoma: Histological subtypes and clinical relevance. *Clin Med Insights Pathol*. 2015;8(1):23-31. doi:10.4137/CPath.s31563

43. 5. Farid M. *Essentials of Diagnostic Breast Pathology, Practical Approach*. 1st Editio. Springer; 2007.

44. Cristofanilli M, Gonzalez-Angulo A, Sneige N, et al. Invasive lobular carcinoma classic type: Response to primary chemotherapy and survival outcomes. *J Clin Oncol*. 2005;23(1):41-48. doi:10.1200/JCO.2005.03.111

45. Fattaneh, AT; Peter D. WHO Classification Pathology and Genetics of Tumours of the Breast and Female Genital Organs. IARC Press; 2003.

46. Orrantia-Borunda E, Anchondo-Nuñez P, Acuña-Aguilar LE, Gómez-Valles FO, CAR-V. Subtypes of Breast Cancer. In: HN M, ed. *Breast Cancer*. Exon publications; 2022.

47. Sharma JD, Khanna S, Ramchandani S, Kakoti LM, Baruah A, Mamidala V. Prevalence of Molecular Subtypes of Breast Carcinoma and Its Comparison between Two Different Age Groups: A Retrospective Study from a Tertiary Care Center of Northeast India. *South Asian J Cancer*. 2021;10(4):220-224. doi:10.1055/s-0041-1731905

48. van Maaren MC, de Munck L, Strobbe LJA, Sonke GS, Westenend PJ, Smidt ML, Poortmans PMP. Ten-year recurrence rates for breast cancer subtypes in the Netherlands: A large population-based study.

*Int J Cancer*. 2019;144(2):263-272.

49. Higgins MJ, Stearns V. Understanding resistance to tamoxifen in hormone receptor-positive breast cancer. *Clin Chem*. 2009;55(8):1453-1455. doi:10.1373/clinchem.2009.125377

50. Lafci O, Celepli P, Oztekin PS KP. DCE-MRI Radiomics Analysis in Differentiating Luminal A and Luminal B Breast Cancer Molecular Subtypes. *Acad Radiol*. 2022;30(1):22-29. doi:https://doi.org/10.1016/j.acra.2022.04.004

51. Usman A. The Survival of Patients with Triple Negative Breast Cancer Undergoing Chemotherapy Along With Lifestyle Change Interventions. *Arch Breast Cancer*. 2022;10(1):66-73. doi:10.32768/abc.202310166-73

52. Krishnamurti U SJ. HER2 in breast cancer: a review and update. *Adv Anat Pathol*. 2014;21(2):100–107. doi:10.1097/PAP .0000000000000015

53. Figueroa-Magalhães MC, Jelovac D, Connolly RM, Wolff AC. Treatment of HER2-positive breast cancer. *Breast*. 2014;23(2):128-136. doi:10.1016/j.breast.2013.11.011

54. Loibl S, Gianni L. HER2-positive breast cancer. *Lancet*. 2017;389(10087):2415-2429. doi:10.1016/S0140-6736(16)32417-5

55. Fragomeni SM, Sciallis A, Jeruss JS. Molecular subtypes and local-regional control of breast cancer. *Surg Oncol Clin N Am*. 2018;27(1):95-120.

56. Oncohema Key. Breast Carcinogenesis. Etiology, Assessing Risk, Initial Evaluation, and Management: Introduction. Published 2019. https://oncohemakey.com/breast-carcinogenesis/#6405105

57. Brooks J, Cairns P, AZ-J. Promoter Methylation and the Detection of

Breast Cancer. *Cancer Causes Control*. 2009;20(9):1539-1550. doi:10.1007/s10552-009-9415-y

58. Sokolenko AP, Bogdanova N, Kluzniak W, et al. Double heterozygotes among breast cancer patients analyzed for BRCA1, CHEK2, ATM, NBN/NBS1, and BLM germ-line mutations. *Breast Cancer Res Treat*. 2014;145(2):553-562. doi:10.1007/s10549-014-2971-1

59. Zagouri F, Sergentanis TN, Zografos GC. Precursors and preinvasive lesions of the breast: The role of molecular prognostic markers in the diagnostic and therapeutic dilemma. *World J Surg Oncol*. 2007;5:1-11. doi:10.1186/1477-7819-5-57

60. Lebeau A, Unholzer A, Amann G, et al. EGFR, HER-2/neu, cyclin D1, p21 and p53 in correlation to cell proliferation and steroid hormone receptor status in ductal carcinoma in situ of the breast. *Breast Cancer Res Treat*. 2003;79(2):187-198. doi:10.1023/A:1023958324448

61. Milena G, Shukri S, Tim C. P53 Pathway in Breast Cancer. *Breast Cancer Res*. 2002;4:70-76.

62. Panigrahy D, Singer S, Shen LQ, et al. PPARγ ligands inhibit primary tumor growth and metastasis by inhibiting angiogenesis. *J Clin Invest*. 2002;110(7):923-932. doi:10.1172/JCI0215634

63. Cheung KJ, Padmanaban V, Silvestri V, et al. Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proc Natl Acad Sci U S A*. 2016;113(7):E854-E863. doi:10.1073/pnas.1508541113

64. Neophytou CM, Kyriakou TC, Papageorgis P. Mechanisms of metastatic tumor dormancy and implications for cancer therapy. *Int J Mol Sci*. 2019;20(24). doi:10.3390/ijms20246158

65. Pelon F, Bourachot B, Kieffer Y, et al. Cancer-associated fibroblast heterogeneity in axillary lymph nodes drives metastases in breast cancer through complementary mechanisms. *Nat Commun*. 2020;11(1). doi:10.1038/s41467-019-14134-w

66. Elizabeth D. Williams, Dingcheng Gao, Andrew Redfern, Erik W T. Controversies around epithelial–mesenchymal plasticity in cancer metastasis. *Nat Rev Cancer*. 2019;19(12):716-732. doi:10.1038/s41568-019-0213-x.

67. Siersbæk R, Scabia V, Nagarajan S, et al. IL6 / STAT3 signaling hijacks ER enhancers to drive breast cancer metastasis. *Cancer Cell*. 2020;38(3):412-423. doi:10.1016/j.ccell.2020.06.007.IL6/STAT3

68. Senkus E, Kyriakides S, Penault-Llorca F, et al. Primary breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2013;24(SUPPL.6):vi7-vi23. doi:10.1093/annonc/mdt284

69. Seely JM. Progress and Remaining Gaps in the Early Detection and Treatment of Breast Cancer. *Curr Oncol*. 2023;30(3):3201-3205. doi:10.3390/curroncol30030242

70. Weigelt B, Peterse JL, Van't Veer LJ. Breast cancer metastasis: Markers and models. *Nat Rev Cancer*. 2005;5(8):591-602. doi:10.1038/nrc1670

71. Humphrey KL, Saksena MA, Freer PE, Smith BL, Rafferty EA. To do or not to do: Axillary nodal evaluation after ACOSOG Z0011 trial. *Radiographics*. 2014;34(7):1807-1816. doi:10.1148/rg.347130141

72. Marino MA, Avendano D, Zapata P, Riedl CC, Pinker K. Lymph Node Imaging in Patients with Primary Breast Cancer: Concurrent Diagnostic Tools. *Oncologist*. 2020;25(2):e231-e242. doi:10.1634/theoncologist.2019-0427

73. Chang JM, Leung JWT, Moy L, Ha SM, Moon WK. Axillary nodal evaluation in breast cancer: State of the art. *Radiology*. 2020;295(3):500-515. doi:10.1148/radiol.2020192534

74. Trialists TS. The UK Standardisation of Breast Radiotherapy (START) Trial A of radiotherapy hypofractionation for treatment of early breast cancer: a randomised trial. *Lancet Oncol*. 2008;9(4):331-341. doi:10.1016/S1470-2045(08)70077-9

75. van Denderen BJW, Thompson EW. The to and fro of tumour spread. *Nature*. 2013;493:487-488.

76. Liu Q, Zhang H, Jiang X, Qian C, Liu Z, Luo D. Factors involved in cancer metastasis: A better understanding to "seed and soil" hypothesis. *Mol Cancer*. 2017;16(1):1-19. doi:10.1186/s12943-017-0742-4

77. Savci-Heijink CD, Halfwerk H, Koster J, van de Vijver MJ. A novel gene expression signature for bone metastasis in breast carcinomas. *Breast Cancer Res Treat*. 2016;156(2):249-259. doi:10.1007/s10549-016-3741-z

78. Wu Q, Li J, Zhu S, et al. Breast cancer subtypes predict the preferential site of distant metastases: A SEER based study. *Oncotarget*. 2017;8(17):27990-27996. doi:10.18632/oncotarget.15856

79. Kennecke H, Yerushalmi R, Woods R, et al. Metastatic behavior of breast cancer subtypes. *J Clin Oncol*. 2010;28(20):3271-3277. doi:10.1200/JCO.2009.25.9820

80. Xiao W, Zheng S, Yang A, et al. Breast cancer subtypes and the risk of distant metastasis at initial diagnosis: A population-based study. *Cancer Manag Res*. 2018;10:5329-5338. doi:10.2147/CMAR.S176763

81. Engla NEW, Journal ND. New engla nd journal. Published online

2015:1685-1687.

82. Karen, Daily; Emily, Douglas; Paul A, Romitti and Alexandra T. Epidemiology of De Novo Metastatic Breast Cancer. *Clin Breast Cancer*. 2021;21(4):302-308.

83. Seltzer S, Corrigan M, O'Reilly S. The clinicomolecular landscape of de novo versus relapsed stage IV metastatic breast cancer. *Exp Mol Pathol*. 2020;114(February 2020):104404. doi:10.1016/j.yexmp.2020.104404

84. Courtney D, Davey MG, Moloney BM, et al. Breast cancer recurrence: factors impacting occurrence and survival. *Ir J Med Sci*. 2022;191(6):2501-2510. doi:10.1007/s11845-022-02926-x

85. Riggio AI, Varley KE, Welm AL. The lingering mysteries of metastatic recurrence in breast cancer. *Br J Cancer*. 2021;124(1):13-26. doi:10.1038/s41416-020-01161-4

86. Ellsworth RE, Blackburn HL, Shriver CD, Soon-Shiong P, Ellsworth DL. Molecular heterogeneity in breast cancer: State of the science and implications for patient care. *Semin Cell Dev Biol*. 2017;64:65-72. doi:10.1016/j.semcdb.2016.08.025

87. Young RH, Louis DN. The Warrens and other pioneering clinician pathologists of the Massachusetts General Hospital during its early years: An appreciation on the 200th anniversary of the hospital founding. *Mod Pathol*. 2011;24(10):1285-1294. doi:10.1038/modpathol.2011.132

88. Zhang QX, Borg A, Wolf DM, Oesterreich S, Fuqua SA. An estrogen receptor mutant with strong hormone-independent activity from a metastatic breast cancer. *Cancer Res*. 1997;57(7):1244-1249.

89. Fumagalli C, Barberis M. Breast cancer heterogeneity. *Diagnostics*.

2021;11(9). doi:10.3390/diagnostics11091555

90. Harris LN, Ismaila N, McShane LM, et al. Use of biomarkers to guide decisions on adjuvant systemic therapy forwomenwith early-stage invasive breast cancer: American Society of clinical Oncology clinical practice guideline. *J Clin Oncol*. 2016;34(10):1134-1150. doi:10.1200/JCO.2015.65.2289

91. Zardavas D, Maetens M, Irrthum A, et al. The AURORA initiative for metastatic breast cancer. *Br J Cancer*. 2014;111(10):1881-1887. doi:10.1038/bjc.2014.341

92. Simpson PT, Reis-Filho JS, Gale T, Lakhani SR. Molecular evolution of breat cancer. *J Pathol*. 2005;205(2):248-254. doi:10.1002/path.1691

93. Lakhani SR, Ellis IO, Schnitt SJ, Tan PH, Van De Vijver MJ. *WHO Classification of Tumours of the Breast.* (Agency I, Cancer for R on, eds.).; 2013.

94. Vance GH, Barry TS, Bloom KJ, et al. Genetic heterogeneity in HER2 testing in breast cancer panel summary and guidelines. *Arch Pathol Lab Med*. 2009;133(4):611-612. doi:10.5858/133.4.611

95. Bardou VJ, Arpino G, Elledge RM, Osborne CK, Clark GM. Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases. *J Clin Oncol*. 2003;21(10):1973-1979. doi:10.1200/JCO.2003.09.099

96. Seol H, Lee HJ, Choi Y, et al. Intratumoral heterogeneity of HER2 gene amplification in breast cancer: Its clinicopathological significance. *Mod Pathol*. 2012;25(7):938-948. doi:10.1038/modpathol.2012.36

97. Abe O, Abe R, Enomoto K, et al. Relevance of breast cancer hormone

receptors and other factors to the efficacy of adjuvant tamoxifen: Patient-level meta-analysis of randomised trials. *Lancet*. 2011;378(9793):771-784. doi:10.1016/S0140-6736(11)60993-8

98. Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature*. 2012;486(7403):395-399. doi:10.1038/nature10933

99. Ellsworth RE, Toro AL, Blackburn HL, et al. Molecular Heterogeneity in Primary Breast Carcinomas and Axillary Lymph Node Metastases Assessed by Genomic Fingerprinting Analysis. *Cancer Growth Metastasis*. 2015;8:CGM.S29490. doi:10.4137/cgm.s29490

100. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012;486(7403):346-352. doi:10.1038/nature10983

101. Navin N, Krasnitz A, Rodgers L, et al. Inferring tumor progression from genomic heterogeneity. *Genome Res*. 2010;20(1):68-80. doi:10.1101/gr.099622.109

102. Sunami E, Shinozaki M, Sim MS, et al. Estrogen receptor and HER2/neu status affect epigenetic differences of tumor-related genes in primary breast tumors. *Breast Cancer Res*. 2008;10(3):1-11. doi:10.1186/bcr2098

103. Koren S, Bentires-Alj M. Breast Tumor Heterogeneity: Source of Fitness, Hurdle for Therapy. *Mol Cell*. 2015;60(4):537-546. doi:10.1016/j.molcel.2015.10.031

104. Gupta PB, Fillmore CM, Jiang G, et al. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell*. 2011;146(4):633-644. doi:10.1016/j.cell.2011.07.026

105. Quail DF, Joyce JA. Microenvironmental regulation of tumor

progression and metastasis. *Nat Med*. 2013;19(11):1423–1437. doi:10.1038/nm.3394.

106. Nowell PC. The clonal evolution of tumor cell populations. *Science (80- )*. 1976;194(4260):23-28. doi:10.1126/science.959840

107. Eguiara A, Elorriaga K, Rezola R, García Martín Á. Células madre tumorales: una diana terapéutica en el cáncer de mama. *Rev Senol y Patol Mamar*. 2012;25(3):107-115. doi:10.1016/s0214-1582(12)70024-5

108. Gilson P, Merlin JL, Harlé A. Deciphering tumour heterogeneity: from tissue to liquid biopsy. *Cancers (Basel)*. 2022;14(6).

109. Sant M, Allemani C, Berrino F, et al. Breast Carcinoma Survival in Europe and the United States: A Population-Based Study. *Cancer*. 2004;100(4):715-722. doi:10.1002/cncr.20038

110. Bhushan A, Gonsalves A, Menon JU. Current state of breast cancer diagnosis, treatment, and theranostics. *Pharmaceutics*. 2021;13(5):1-24. doi:10.3390/pharmaceutics13050723

111. Barba D, León-Sosa A, Lugo P, et al. Breast cancer, screening and diagnostic tools: All you need to know. *Crit Rev Oncol Hematol*. 2021;157. doi:10.1016/j.critrevonc.2020.103174

112. Løberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. *Breast Cancer Res*. 2015;17(1):1-12. doi:10.1186/s13058-015-0525-z

113. Gatta G, Somma F, Sardu C, et al. Automated 3D Ultrasound as an Adjunct to Screening Mammography Programs in Dense Breast: Literature Review and Metanalysis. *J Pers Med*. 2023;13(12). doi:10.3390/jpm13121683

114. Sehgal CM, Weinstein SP, Arger PH, Conant EF. A review of breast

ultrasound. *J Mammary Gland Biol Neoplasia*. 2006;11(2):113-123. doi:10.1007/s10911-006-9018-0

115. Geisel J, Raghu M, Hooley R. The Role of Ultrasound in Breast Cancer Screening: The Case for and Against Ultrasound. *Semin Ultrasound, CT MRI*. 2018;39(1):25-34. doi:10.1053/j.sult.2017.09.006

116. Cheng YC, Wu NY, Ko JS, et al. Breast cancers detected by breast MRI screening and ultrasound in asymptomatic Asian women: 8 years of experience in Taiwan. *Oncology*. 2012;82(2):98-107. doi:10.1159/000335958

117. Mann RM, Balleyguier C, Baltzer PA, et al. Breast MRI: EUSOBI recommendations for women's information. *Eur Radiol*. 2015;25(12):3669-3678. doi:10.1007/s00330-015-3807-z

118. Kuhl. CK. Abbreviated magnetic resonance imaging for breast cancer screening: Concept, early results, and considerations. *Annu Rev Med*. 2019;70:501-519. doi:10.3348/kjr.2018.0722

119. Gradishar WJ, Anderson BO, Abraham J, et al. Breast cancer, version 3.2020. *JNCCN J Natl Compr Cancer Netw*. 2020;18(4):452-478. doi:10.6004/jnccn.2020.0016

120. Elmore JG, Longton GM, Carney PA, et al. Diagnostic Concordance Among Pathologists Interpreting Breast Biopsy Specimens. *JAMA*. 2015;313(11):1122–1132. doi:doi:10.1001/jama.2015.1405.

121. Buono G, Gerratana L, Bulfoni M, et al. Circulating tumor DNA analysis in breast cancer: Is it ready for prime-time? *Cancer Treat Rev*. 2019;73(January 2019):73-83. doi:10.1016/j.ctrv.2019.01.004

122. American Joint Comitte of Cancer. Breast Cancer Staging AJCC Cancer Staging Manual 8. Published online 2018.

123. Koh J, Kim MJ. Erratum to: Introduction of a New Staging System of Breast Cancer for Radiologists: An Emphasis on the Prognostic Stage (Korean J Radiol 2019;20(1):69-82, 10.3348/kjr.2018.0231). *Korean J Radiol*. 2022;23(5):570. doi:10.3348/kjr.2022.0246

124. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging . *CA Cancer J Clin*. 2017;67(2):93-99. doi:10.3322/caac.21388

125. Geyer FC, Pareja F, Weigelt B, et al. The Spectrum of Triple-Negative Breast Disease: High- and Low-Grade Lesions. *Am J Pathol*. 2017;187(10):2139-2151. doi:10.1016/j.ajpath.2017.03.016

126. Beumer I, Witteveen A, Delahaye L, et al. Equivalence of MammaPrint array types in clinical trials and diagnostics. *Breast Cancer Res Treat*. 2016;156(2):279-287. doi:10.1007/s10549-016-3764-5

127. Sparano JA, Gray RJ, Makower DF, et al. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med*. 2018;379(2):111-121. doi:10.1056/nejmoa1804710

128. Wang J, Wu SG. Breast Cancer: An Overview of Current Therapeutic Strategies, Challenge, and Perspectives. *Breast Cancer Targets Ther*. 2023;15(October):721-730. doi:10.2147/BCTT.S432526

129. Senkus E, Łacko A. Over-treatment in metastatic breast cancer. *Breast*. 2017;31:309-317. doi:10.1016/j.breast.2016.06.024

130. Moo T, Sanford R, Dang C, et al. Overview of Breast Cancer Therapy. *PET Clin*. 2018;13(3):339-354. doi:10.1016/j.cpet.2018.02.006.Overview

131. Donker M , Slaets L, van Tienhoven G, Rutgers EJT. [Axillary lymph

node dissection versus axillary radiotherapy in patients with a positive sentinel node: the AMAROS trial]. *Ned Tijdschr Geneeskd*. 2015;159.

132. Giuliano AE, Morrow M. Effect of Axillary Dissection vs No Axillary Dissection on 10-Year Overall Survival Among Women With Invasive Breast Cancer and Sentinel Node Metastasis. *JAMA - J Am Med Assoc*. 2017;318(10):918-926. doi:10.1001/jama.2017.11470.Effect

133. Van der Hage JA, Van de Velde CJH, Julien JP, Tubiana-Hulin M, Vandervelden C, Duchateau L. Preoperative chemotherapy in primary operable breast cancer: Results from the European Organization for Research and Treatment of Cancer Trial 10902. *J Clin Oncol*. 2001;19(22):4224-4237. doi:10.1200/JCO.2001.19.22.4224

134. Broët P, Scholl SM, De la Rochefordière A, et al. Short and long-term effects on survival in breast cancer patients treated by primary chemotherapy: An updated analysis of a randomized trial. *Breast Cancer Res Treat*. 1999;58(2):151-156. doi:10.1023/A:1006339918798

135. Mano MS, Awada A. Primary chemotherapy for breast cancer: The evidence and the future. *Ann Oncol*. 2004;15(8):1161-1171. doi:10.1093/annonc/mdh302

136. Sparano JA, Gray RJ, Makower DF, et al. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med*. 2018;379(2):111-121.

137. Andre F, Ismaila N, Henry NL, et al. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: ASCO clinical practice guideline update-integration of results from TAILORx. *J Clin Oncol*. 2019;37(22):1956-1964. doi:10.1200/JCO.19.00945

138. Citron ML, Berry DA, Cirrincione C, et al. Randomized trial of dose-

dense versus conventionally scheduled and sequential versus concurrent combination chemotherapy as postoperative adjuvant treatment of node-positive primary breast cancer: First report of Intergroup Trial C9741/Cancer and Leukemia. *J Clin Oncol*. 2003;21(8):1431-1439. doi:10.1200/JCO.2003.09.081

139. Blum JL, Flynn PJ, Yothers G, et al. Anthracyclines in Early Breast Cancer: The ABC Trials. *J Clin Oncol*. 2017;35(23):2647-2655. https://doi.org/10.1200/JCO.2016.

140. Kunkler IH, Williams LJ, Jack WJL, Cameron DA, Dixon JM. Breast-conserving surgery with or without irradiation in women aged 65 years or older with early breast cancer (PRIME II): A randomised controlled trial. *Lancet Oncol*. 2015;16(3):266-273. doi:10.1016/S1470-2045(14)71221-5

141. Hughes KS, Schnaper LA, Bellon JR, et al. Lumpectomy plus tamoxifen with or without irradiation in women age 70 years or older with early breast cancer: Long-term follow-up of CALGB 9343. *J Clin Oncol*. 2013;31(19):2382-2387. doi:10.1200/JCO.2012.45.2615

142. Jacobson G, Kaidar-Person O, Haisraely O, et al. Palliative radiation therapy for symptomatic advance breast cancer. *Sci Rep*. 2021;11(1):5-10. doi:10.1038/s41598-021-84872-9

143. McBride A, Allen P, Woodward W, et al. Locoregional recurrence risk for patients with T1,2 breast cancer with 1-3 positive lymph nodes treated with mastectomy and systemic treatment. *Int J Radiat Oncol Biol Phys*. 2014;89(2):392-398. doi:10.1016/j.ijrobp.2014.02.013

144. Wu SG, Wang J, Lian CL, et al. Evaluation of the 8th edition of the American joint committee on cancer's pathological staging system in prognosis assessment and treatment decision making for stage T1-2N1 breast cancer after mastectomy. *Breast*. 2020;51:2-10.

doi:10.1016/j.breast.2020.02.012

145. Du L, Yau C, Brown-Swigart L, et al. Predicted sensitivity to endocrine therapy for stage II-III hormone receptor-positive and HER2-negative (HR+/HER2−) breast cancer before chemo-endocrine therapy. *Ann Oncol*. 2021;32(5):642-651. doi:10.1016/j.annonc.2021.02.011

146. Pagani O, Regan MM, Walley BA, et al. Adjuvant Exemestane with Ovarian Suppression in Premenopausal Breast Cancer. *N Engl J Med*. 2014;371(2):107-118. doi:10.1056/nejmoa1404037

147. Goss PE, Ingle JN, Pritchard KI, et al. Extending Aromatase-Inhibitor Adjuvant Therapy to 10 Years. *N Engl J Med*. 2016;375(3):209-219. doi:10.1056/nejmoa1604700

148. Burstein HJ, Temin S, Anderson H, et al. Adjuvant endocrine therapy for women with hormone receptor-positive breast cancer: American Society of Clinical Oncology clinical practice guideline focused update. *J Clin Oncol*. 2014;32(21):2255-2269. doi:10.1200/JCO.2013.54.2258

149. Tolaney SM, Barry WT, Dang CT, et al. Adjuvant Paclitaxel and Trastuzumab for Node-Negative, HER2-Positive Breast Cancer. *N Engl J Med*. 2015;372(2):134-141. doi:10.1056/nejmoa1406281

150. Gianni L, Pienkowski T, Im YH, et al. Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): A randomised multicentre, open-label, phase 2 trial. *Lancet Oncol*. 2012;13(1):25-32. doi:10.1016/S1470-2045(11)70336-9

151. Schneeweiss A, Chia S, Hickish T, et al. Pertuzumab plus trastuzumab in combination with standard neoadjuvant anthracycline-containing and anthracycline-free chemotherapy regimens in patients with HER2-

positive early breast cancer: A randomized phase II cardiac safety study (TRYPHAENA). *Ann Oncol*. 2013;24(9):2278-2284. doi:10.1093/annonc/mdt182

152. Warwas KM, Meyer M, Gonçalves M, et al. Co-Stimulatory Bispecific Antibodies Induce Enhanced T Cell Activation and Tumor Cell Killing in Breast Cancer Models. *Front Immunol*. 2021;12(August):1-22. doi:10.3389/fimmu.2021.719116

153. Mittendorf EA, Zhang H, Barrios CH, et al. Neoadjuvant atezolizumab in combination with sequential nab-paclitaxel and anthracycline-based chemotherapy versus placebo and chemotherapy in patients with early-stage triple-negative breast cancer (IMpassion031): a randomised, double-blind, phase 3 tria. *Lancet*. 2020;396(10257):1090-1100. doi:10.1016/S0140-6736(20)31953-X

154. Dastjerd NT, Valibeik A, Monfared SR, et al. Gene therapy: A promising approach for breast cancer treatment. *Cell Biochem Funct*. 2022;(1):28-48. doi:10.1002/cbf.3676

155. Pandha HS, Martin LA, Rigg A, et al. Genetic prodrug activation therapy for breast cancer: A phase I clinical trial of erbB-2-directed suicide gene expression. *J Clin Oncol*. 1999;17(7):2180-2189. doi:10.1200/jco.1999.17.7.2180

156. Hong DS, Kang YK, Borad M, et al. Phase 1 study of MRX34, a liposomal miR-34a mimic, in patients with advanced solid tumours. *Br J Cancer*. 2020;122(11):1630-1637. doi:10.1038/s41416-020-0802-1

157. Qiu J, Qian D, Jiang Y, Meng L, Huang L. Circulating tumor biomarkers in early-stage breast cancer: characteristics, detection, and clinical developments. *Front Oncol*. 2023;13(October):1-13. doi:10.3389/fonc.2023.1288077

158. Preat F, Simon P, Noel JC. Differences in breast carcinoma

immunohistochemical subtypes between immigrant arab and european women. *Diagn Pathol*. 2014;9(1):1-5. doi:10.1186/1746-1596-9-26

159. Safi F, Kohler I, Rottinger E, Suhr P, Beger HG. Comparison of CA 15-3 and CEA in diagnosis and monitoring of breast cancer. *Int J Biol Markers*. 1989;4(4):207-214. doi:10.1177/172460088900400405

160. Dai X, Li T, Bai Z, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. *Am J Cancer Res*. 2015;5(10):2929-2943.

161. Eliyatkin N, Yalcin E, Zengel B, Aktaş S, Vardar E. Molecular Classification of Breast Carcinoma: From Traditional, Old-Fashioned Way to A New Age, and A New Way. *J Breast Heal*. 2015;11(2):59-66. doi:10.5152/tjbh.2015.1669

162. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: Monitoring cancer-genetics in the blood. *Nat Rev Clin Oncol*. 2013;10(8):472-484. doi:10.1038/nrclinonc.2013.110

163. Wan JCM, Massie C, Garcia-Corbacho J, et al. Liquid biopsies come of age: Towards implementation of circulating tumour DNA. *Nat Rev Cancer*. 2017;17(4):223-238. doi:10.1038/nrc.2017.7

164. Heitzer E, Haque IS, Roberts CES, Speicher MR. Current and future perspectives of liquid biopsies in genomics-driven oncology. *Nat Rev Genet*. 2019;20(2):71-88. doi:10.1038/s41576-018-0071-5

165. Diaz LA, Bardelli A. Liquid biopsies: Genotyping circulating tumor DNA. *J Clin Oncol*. 2014;32(6):579-586. doi:10.1200/JCO.2012.45.2011

166. Boulanger CM, Loyer X, Rautou PE, Amabile N. Extracellular vesicles in coronary artery disease. *Nat Rev Cardiol*. 2017;14(5):259-272. doi:10.1038/nrcardio.2017.7

167. Willms E, Cabañas C, Mäger I, Wood MJA, Vader P. Extracellular vesicle heterogeneity: Subpopulations, isolation techniques, and diverse functions in cancer progression. *Front Immunol*. 2018;9(APR). doi:10.3389/fimmu.2018.00738

168. Marar C, Starich B, Wirtz D. Extracellular vesicles in immunomodulation and tumor progression. *Nat Immunol*. 2021;22(5):560-570. doi:10.1038/s41590-021-00899-0

169. Samuels M, Jones W, Towler B, Turner C, Robinson S, Giamas G. The role of non-coding RNAs in extracellular vesicles in breast cancer and their diagnostic implications. *Oncogene*. 2023;42(41):3017-3034. doi:10.1038/s41388-023-02827-y

170. Tian W, Lei N, Zhou J, et al. Extracellular vesicles in ovarian cancer chemoresistance, metastasis, and immune evasion. *Cell Death Dis*. 2022;13(1). doi:10.1038/s41419-022-04510-8

171. Loric S, Denis JA, Desbene C, Sabbah M, Conti M. Extracellular Vesicles in Breast Cancer: From Biology and Function to Clinical Diagnosis and Therapeutic Management. *Int J Mol Sci*. 2023;24(8). doi:10.3390/ijms24087208

172. Bebelman MP, Smit MJ, Pegtel DM, Baglio SR. Biogenesis and function of extracellular vesicles in cancer. *Pharmacol Ther*. 2018;188:1-11. doi:10.1016/j.pharmthera.2018.02.013

173. Prada I, Meldolesi J. Binding and fusion of extracellular vesicles to the plasma membrane of their cell targets. *Int J Mol Sci*. 2016;17(8). doi:10.3390/ijms17081296

174. Nguyen VH, Lee BJ. Protein corona: A new approach for nanomedicine design. *Int J Nanomedicine*. 2017;12:3137-3151. doi:10.2147/IJN.S129300

175. Buzás EI, Tóth E, Sódar BW, Szabó-Taylor K. Molecular interactions at the surface of extracellular vesicles. *Semin Immunopathol*. 2018;40(5):453-464. doi:10.1007/s00281-018-0682-0

176. Heidarzadeh M, Zarebkohan A, Rahbarghazi R, Sokullu E. Protein corona and exosomes: new challenges and prospects. *Cell Commun Signal*. 2023;21(1):1-15. doi:10.1186/s12964-023-01089-1

177. Abels ER, Breakefield XO. Introduction to Extracellular Vesicles: Biogenesis, RNA Cargo Selection, Content, Release, and Uptake. *Cell Mol Neurobiol*. 2016;36(3):301-312. doi:10.1007/s10571-016-0366-z

178. Li Y, He X, Li Q, et al. EV-origin: Enumerating the tissue-cellular origin of circulating extracellular vesicles using exLR profile. *Comput Struct Biotechnol J*. 2020;18:2851-2859. doi:10.1016/j.csbj.2020.10.002

179. Hoshino A, Costa-Silva B, Shen TL, et al. Tumour exosome integrins determine organotropic metastasis. *Nature*. 2015;527(7578):329-335. doi:10.1038/nature15756

180. Yu W, Hurley J, Roberts D, et al. Exosome-based liquid biopsies in cancer: opportunities and challenges. *Ann Oncol*. 2021;32(4):466-477. doi:10.1016/j.annonc.2021.01.074

181. Barlin M, Erdmann-Gilmore P, Mudd JL, et al. Proteins in Tumor-Derived Plasma Extracellular Vesicles Indicate Tumor Origin. *Mol Cell Proteomics*. 2023;22(1):0-9. doi:10.1016/j.mcpro.2022.100476

182. Tian Y, Gong M, Hu Y, et al. Quality and efficiency assessment of six extracellular vesicle isolation methods by nano-flow cytometry. *J Extracell Vesicles*. 2020;9(1). doi:10.1080/20013078.2019.1697028

183. Chiriacò MS, Bianco M, Nigro A, et al. Lab-on-chip for exosomes and microvesicles detection and characterization. *Sensors (Switzerland)*.

2018;18(10). doi:10.3390/s18103175

184. De Sousa KP, Rossi I, Abdullahi M, Ramirez MI, Stratton D, Inal JM. Isolation and characterization of extracellular vesicles and future directions in diagnosis and therapy. *Wiley Interdiscip Rev Nanomedicine Nanobiotechnology*. 2023;15(1):1-29. doi:10.1002/wnan.1835

185. Gebert LFR, MacRae IJ. Regulation of microRNA function in animals. *Nat Rev Mol Cell Biol*. 2019;20(1):21-37. doi:10.1038/s41580-018-0045-7

186. Peng Y, Croce CM. The role of microRNAs in human cancer. *Signal Transduct Target Ther*. 2016;1(November 2015). doi:10.1038/sigtrans.2015.4

187. Shang R, Lee S, Senavirathne G, Lai EC. microRNAs in action: biogenesis, function and regulation. *Nat Rev Genet*. 2023;24(12):816-833. doi:10.1038/s41576-023-00611-y

188. Di Leva G, Garofalo M, Croce CM. MicroRNAs in cancer. *Annu Rev Pathol Mech Dis*. 2014;9(September 2013):287-314. doi:10.1146/annurev-pathol-012513-104715

189. Bartel DP. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*. 2004;116(2):281-297. doi:10.1016/S0092-8674(04)00045-5

190. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Front Endocrinol (Lausanne)*. 2018;9(AUG):1-12. doi:10.3389/fendo.2018.00402

191. Vasudevan S, Steitz JA. AU-Rich-Element-Mediated Upregulation of Translation by FXR1 and Argonaute 2. *Cell*. 2007;128(6):1105-1118. doi:10.1016/j.cell.2007.01.038

192. Tian T, Zhu YL, Zhou YY, et al. Exosome uptake through clathrin-mediated endocytosis and macropinocytosis and mediating miR-21 delivery. *J Biol Chem*. 2014;289(32):22258-22267. doi:10.1074/jbc.M114.588046

193. Chen X, Liang H, Zhang J, Zen K, Zhang CY. Secreted microRNAs: A new form of intercellular communication. *Trends Cell Biol*. 2012;22(3):125-132. doi:10.1016/j.tcb.2011.12.001

194. Mitchell PS, Parkin RK, Kroh EM, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A*. 2008;105(30):10513-10518. doi:10.1073/pnas.0804549105

195. Kosaka N, Iguchi H, Yoshioka Y, Takeshita F, Matsuki Y, Ochiya T. Secretory mechanisms and intercellular transfer of microRNAs in living cells. *J Biol Chem*. 2010;285(23):17442-17452. doi:10.1074/jbc.M110.107821

196. De Mauro A, Greco M, Grimaldi M. A formal definition of Big Data based on its essential features. *Libr Rev*. 2016;65(3):122-135. doi:10.1108/LR-06-2015-0061

197. Stanford Encyclopedia of Philosophy. Scientific Research and Big Data. Published 2020. https://plato.stanford.edu/entries/science-big-data/

198. Jiang P, Sinha S, Aldape K, Hannenhalli S, Sahinalp C, Ruppin E. Big data in basic and translational cancer research. *Nat Rev Cancer*. 2022;22(11):625-639. doi:10.1038/s41568-022-00502-0

199. Luiz Pinto A, Shirshahi S, Varmazyar R, González-Valiente CL. Business science: Overproduction, self-citations and Article processing charge (APC). *Salud, Cienc y Tecnol - Ser Conf*. 2023;2. doi:10.56294/sctconf2023444

200. Ferreira González I, Urrútia G, Alonso-Coello P. Systematic Reviews and Meta-Analysis: Scientific Rationale and Interpretation. *Rev Española Cardiol (English Ed*. 2011;64(8):688-696. doi:10.1016/j.rec.2011.03.027

201. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Med*. 2009;6(7). doi:10.1371/journal.pmed.1000100

202. Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein HR. *Introduction to Meta-Analysis*. John Wiley & Sons Ltd.; 2011.

203. Garmaa G, Bunduc S, Kói T, et al. A Systematic Review and Meta-Analysis of microRNA Profiling Studies in Chronic Kidney Diseases. *Non-Coding RNA*. 2024;10(3):30. doi:10.3390/ncrna10030030

204. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-1558. doi:10.1002/sim.1186

205. Rodríguez-Martínez A, De Miguel-Pérez D, Ortega FG, et al. Exosomal miRNA profile as complementary tool in the diagnostic and prediction of treatment response in localized breast cancer under neoadjuvant chemotherapy. *Breast Cancer Res*. 2019;21(1):1-9. doi:10.1186/s13058-019-1109-0

206. Giannopoulou EG, Elemento O, Ivashkiv LB. Use of RNA sequencing to evaluate rheumatic disease patients. *Arthritis Res Ther*. 2015;17(1):1-10. doi:10.1186/s13075-015-0677-3

207. Martin M. Cutadapt removes adapter sequences from high-troughput sequencing reads. *EMBnet J*. Published online 2011.

208. Friedländer MR, MacKowiak SD, Li N, Chen W, Rajewsky N. MiRDeep2 accurately identifies known and hundreds of novel

microRNA genes in seven animal clades. *Nucleic Acids Res*. 2012;40(1):37-52. doi:10.1093/nar/gkr688

209. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3). doi:10.1186/gb-2009-10-3-r25

210. Kim D, Langmead B, Salzberg SL. HISAT: A fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357-360. doi:10.1038/nmeth.3317

211. Kozomara A, Griffiths-Jones S. MiRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res*. 2014;42(D1):68-73. doi:10.1093/nar/gkt1181

212. Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, et al. DIANA-TarBase v8: A decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Res*. 2018;46(D1):D239-D245. doi:10.1093/nar/gkx1141

213. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. *Nat Protoc*. 2008;3(6):1101-1108. doi:10.1038/nprot.2008.73

214. Minemura H, Takagi K, Miki Y, et al. Abnormal expression of miR-1 in breast carcinoma as a potent prognostic factor. *Cancer Sci*. 2015;106(11):1642-1650. doi:10.1111/cas.12808

215. Chen X, Wang YW, Zhu WJ, et al. A 4-microRNA signature predicts lymph node metastasis and prognosis in breast cancer. *Hum Pathol*. 2018;76:122-132. doi:10.1016/j.humpath.2018.03.010

216. Zhou S wei, Su B bei, Zhou Y, et al. Aberrant miR-215 expression is associated with clinical outcome in breast cancer patients. *Med Oncol*. 2014;31(11):1-6. doi:10.1007/s12032-014-0259-2

217. Zhang P, Fan C, Du J, Mo X, Zhao Q. Association of miR-1247-5p expression with clinicopathological parameters and prognosis in breast cancer. *Int J Exp Pathol*. 2018;99(4):199-205. doi:10.1111/iep.12287

218. Ibrahim AM, Said MM, Hilal AM, Medhat AM, Abd Elsalam IM. Candidate circulating microRNAs as potential diagnostic and predictive biomarkers for the monitoring of locally advanced breast cancer patients. *Tumor Biol*. 2020;42(10):1-13. doi:10.1177/1010428320963811

219. Si H, Sun X, Chen Y, et al. Circulating microRNA-92a and microRNA-21 as novel minimally invasive biomarkers for primary breast cancer. *J Cancer Res Clin Oncol*. 2013;139(2):223-229. doi:10.1007/s00432-012-1315-y

220. Escuin D, López-Vilaró L, Mora J, et al. Circulating microRNAs in Early Breast Cancer Patients and Its Association With Lymph Node Metastases. *Front Oncol*. 2021;11(August):1-14. doi:10.3389/fonc.2021.627811

221. Antolín S, Calvo L, Blanco-Calvo M, et al. Circulating miR-200c and miR-141 and outcomes in patients with breast cancer. *BMC Cancer*. 2015;15(1):1-15. doi:10.1186/s12885-015-1238-5

222. Escuin D, López-Vilaró L, Bell O, et al. Circulating miRNA Expression Is Inversely Correlated with Tumor Tissue or Sentinel Lymph Nodes in Estrogen Receptor-Positive Early Breast Cancer Patients. *Int J Mol Sci*. 2023;24(17). doi:10.3390/ijms241713293

223. Huang GL, Zhang XH, Guo GL, et al. Clinical significance of miR-21 expression in breast cancer: SYBR-Green I-based real-time RT-PCR study of invasive ductal carcinoma. *Oncol Rep*. Published online 2009.

224. Liu B, Pan J, Fu C. Correlation of microRNA-367 in the clinicopathologic features and prognosis of breast cancer patients. *Med*

*(United States)*. 2021;100(22):E26103. doi:10.1097/MD.0000000000026103

225. Avery-Kiejda KA, Braye SG, Mathe A, Forbes JF, Scott RJ. Decreased expression of key tumour suppressor microRNAs is associated with lymph node metastases in triple negative breast cancer. *BMC Cancer*. 2014;14(1):1-14. doi:10.1186/1471-2407-14-51

226. Dong LL, Chen LM, Wang WM, Zhang LM. Decreased expression of microRNA-124 is an independent unfavorable prognostic factor for patients with breast cancer. *Diagn Pathol*. 2015;10(1):1-5. doi:10.1186/s13000-015-0257-5

227. Xu F, He H, Huang W, et al. Decreased expression of MicroRNA-200 family in human breast cancer is associated with lymph node metastasis. *Clin Transl Oncol*. 2016;18(3):283-288. doi:10.1007/s12094-015-1364-1

228. Li Y, Hong F, Yu Z. Decreased expression of microRNA-206 in breast cancer and its association with disease characteristics and patient survival. *J Int Med Res*. 2013;41(3):596-602. doi:10.1177/0300060513485856

229. Li S, Zhang1 M, Xu2 F, Wang Y, Leng D. Detection significance of miR-3662, miR-146a, and miR-1290 in serum exosomes of breast cancer patients. *J Cancer Res Ther*. Published online 2021.

230. Okuno J, Miyake T, Sota Y, et al. Development of Prediction Model Including MicroRNA Expression for Sentinel Lymph Node Metastasis in ER-Positive and HER2-Negative Breast Cancer. *Ann Surg Oncol*. 2021;28(1):310-319. doi:10.1245/s10434-020-08735-9

231. Rask L, Balslev E, Søkilde R, et al. Differential expression of miR-139, miR-486 and miR-21 in breast cancer patients sub-classified according to lymph node status. *Cell Oncol*. 2014;37(3):215-227.

doi:10.1007/s13402-014-0176-6

232. Fang C, Wang FB, Li Y, Zeng XT. Down-regulation of miR-199b-5p is correlated with poor prognosis for breast cancer patients. *Biomed Pharmacother*. 2016;84:1189-1193. doi:10.1016/j.biopha.2016.10.006

233. Jurkovicova D, Smolkova B, Magyerkova M, et al. Down-regulation of traditional oncomiRs in plasma of breast cancer patients. *Oncotarget*. 2017;8(44):77369-77384. doi:10.18632/oncotarget.20484

234. Savad S, Mehdipour P, Miryounesi M, et al. Expression analysis of miR-21, miR-205, and miR-342 in breast cancer in Iran. *Asian Pacific J Cancer Prev*. 2012;13(3):873-877. doi:10.7314/APJCP.2012.13.3.873

235. Du Y, Miao Z, Wang K, Lv Y, Qiu L, Guo L. Expression levels and clinical values of miR-92b-3p in breast cancer. *World J Surg Oncol*. 2021;19(1):1-8. doi:10.1186/s12957-021-02347-7

236. Chekhun V, Borikun T, Mushii O, et al. EXPRESSION PROFILE OF miR-145 , -182 , -21 , -27a , -29b , and -34a IN BREAST CANCER. *Exp Oncol*. 2023;45(4):421-431. doi: 10.15407/exp-oncology

237. Tan X, Guo L, Luo S, Wang W, Zhang K, Chen F. Expression profile of miR-155 and miR-34a in peripheral blood and tissues of breast cancer patients. *Int J Clin Exp Pathol*. 2016;9(7):7566-7570.

238. Wu X. Expressions of MiR-21 and MiR-210 in breast cancer and their predictive values for prognosis. *Iran J Public Health*. 2020;49(1):21-29. doi:10.18502/ijph.v49i1.3048

239. Sun D, Zhong J, Wei W, Chen X, Liu J, Hu Z. Identification of microRNA expression in sentinel lymph nodes from patients with breast cancer via RNA sequencing for diagnostic accuracy. *J Gene Med*. 2019;21(4). doi:10.1002/jgm.3075

240. Guo J, Liu C, Wang W, et al. Identification of serum miR-1915-3p and miR-455-3p as biomarkers for breast cancer. *PLoS One*. 2018;13(7):1-15. doi:10.1371/journal.pone.0200716

241. Yin L, Ding Y, Wang Y, Wang C, Sun K, Wang L. Identification of Serum miR-501-3p and miR-338-3p as Novel Diagnostic Biomarkers for Breast Cancer and Their Target Genes Associated with Immune Infiltration. *Int J Gen Med*. 2023;16(April):1279-1294. doi:10.2147/IJGM.S406802

242. Wang H, Chen SH, Kong P, Zhang LY, Zhang NQ, Gu H. MiR-330-3p in breast cancer. *Eur Rev Med Pharmacol Sci*. 2018;22:1726-1730.

243. Cao GH, Sun XL, Wu F, Chen WF, Li JQ, Hu WC. Low expression of miR-409-3p is a prognostic marker for breast cancer. *Eur Rev Med Pharmacol Sci*. 2016;20(18):3825-3829.

244. Zhang XY, Liu DJ, Yuan RB, et al. Low expression of miR-597 is correlated with tumor stage and poor outcome in breast cancer. *Eur Rev Med Pharmacol Sci*. 2018;22(2):456-460.

245. Elango R, Alsaleh KA, Vishnubalaji R, et al. MicroRNA Expression Profiling on Paired Primary and Lymph Node Metastatic Breast Cancer Revealed Distinct microRNA Profile Associated With LNM. *Front Oncol*. 2020;10(May):1-13. doi:10.3389/fonc.2020.00756

246. Yan LX, Huang XF, Shao Q, et al. MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. *Rna*. 2008;14(11):2348-2360. doi:10.1261/rna.1034808

247. Wu J, Zhou Z. MicroRNA-432 Acts as a Prognostic Biomarker and an Inhibitor of Cell Proliferation, Migration, and Invasion in Breast Cancer. *Clin Breast Cancer*. 2021;21(4):e462-e470. doi:10.1016/j.clbc.2021.01.014

248. Fkih M'hamed I, Privat M, Trimeche M, Penault-Llorca F, Bignon YJ, Kenani A. miR-10b, miR-26a, miR-146a And miR-153 Expression in Triple Negative Vs Non Triple Negative Breast Cancer: Potential Biomarkers. *Pathol Oncol Res*. 2017;23(4):815-827. doi:10.1007/s12253-017-0188-4

249. Shahabi A, Naghili B, Ansarin K, Montazeri V, Zarghami N. miR-140 and miR-196a as potential biomarkers in breast cancer patients. *Asian Pacific J Cancer Prev*. 2020;21(7):1913-1918. doi:10.31557/APJCP.2020.21.7.1913

250. Berber U, Yilmaz I, Narli G, Haholu A, Kucukodaci Z, Demirel D. MiR-205 and miR-200c: Predictive micro RNAs for lymph node metastasis in triple negative breast cancer. *J Breast Cancer*. 2014;17(2):143-148. doi:10.4048/jbc.2014.17.2.143

251. Wang B, Li J, Sun M, Sun L, Zhang X. MiRNA expression in breast cancer varies with lymph node metastasis and other clinicopathologic features. *IUBMB Life*. 2014;66(5):371-377. doi:10.1002/iub.1273

252. Wang J, Zhao H, Tang D, Wu J, Yao G, Zhang Q. Overexpressions of microRNA-9 and microRNA-200c in human breast cancers are associated with lymph node metastasis. *Cancer Biother Radiopharm*. 2013;28(4):283-288. doi:10.1089/cbr.2012.1293

253. Xie X, Tan W, Chen B, et al. Preoperative prediction nomogram based on primary tumor miRNAs signature and clinical-related features for axillary lymph node metastasis in early-stage invasive breast cancer. *Int J Cancer*. 2018;142(9):1901-1910. doi:10.1002/ijc.31208

254. Wang S, Li H, Wang J, Wang D, Yao A, Li Q. Prognostic and biological significance of MicroRNA-127 expression in human breast cancer. *Dis Markers*. 2014;2014. doi:10.1155/2014/401986

255. Zheng R, Pan L, Gao J, et al. Prognostic value of miR-106b expression

in breast cancer patients. *J Surg Res*. 2015;195(1):158-165. doi:10.1016/j.jss.2014.12.035

256. Gao Z, Li M, Zhang L, Zhao Q. Serum microRNA-34c acts as a novel diagnostic biomarker in breast cancer patients. *Int J Clin Exp Pathol*. 2016;9(6):6489-6494.

257. Shiino S, Matsuzaki J, Shimomura A, et al. Serum miRNA–based Prediction of Axillary Lymph Node Metastasis in Breast Cancer. *Clin Cancer Res*. 2019;25(6):1817-1827. doi:10.1158/1078-0432.CCR-18-1414

258. Krell J, Frampton AE, Jacob J, et al. The clinico-pathologic role of MicroRNAs miR-9 and miR-151-5p in breast cancer metastasis. *Mol Diagnosis Ther*. 2012;16(3):167-172. doi:10.1007/bf03262205

259. Chen W, Cai F, Zhang B, Barekati Z, Zhong XY. The level of circulating miRNA-10b and miRNA-373 in detecting lymph node metastasis of breast cancer: Potential biomarkers. *Tumor Biol*. 2013;34(1):455-462. doi:10.1007/s13277-012-0570-5

260. Chun SM, Park HJ, Kim CH, Kim I. The significance of MicroRNA let-7b, miR-30c, and miR-200c expression in breast cancers. *Korean J Pathol*. 2011;45(4):354-360. doi:10.4132/KoreanJPathol.2011.45.4.354

261. Lukianova NY, Borikun T V., Chekhun VF. Tumor microenvironment-derived miRNAs as prognostic markers of breast cancer. *Exp Oncol*. 2019;41(3):242-247. doi:10.32471/exp-oncology.2312-8852.vol-41-no-3.13615

262. Alunni-Fabbroni M, Majunke L, Trapp EK, et al. Whole blood microRNAs as potential biomarkers in post-operative early breast cancer patients. *BMC Cancer*. 2018;18(1):1-12. doi:10.1186/s12885-018-4020-7

263. Chernyy V, Pustylnyak V, Kozlov V, Gulyaeva L. Increased expression of miR-155 and miR-222 is associated with lymph node positive status. *J Cancer*. 2018;9(1):135-140. doi:10.7150/jca.22181

264. Smeets A, Daemen A, Vanden Bempt I, et al. Prediction of lymph node involvement in breast cancer from primary tumor tissue using gene expression profiling and miRNAs. *Breast Cancer Res Treat*. 2011;129(3):767-776. doi:10.1007/s10549-010-1265-5

265. Niedźwiecki S, Piekarski J, Szymańska B, Pawłowska Z, Jeziorski A. Expression level of serum circulating miRNA-21, miRNA-10b and miRNA-200c in breast cancer patients with sentinel lymph node metastasis. *Postepy Hig Med Dosw*. 2019;73(502):325-331. doi:10.5604/01.3001.0013.2562

266. Théry C, Witwer KW, Aikawa E, et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J Extracell Vesicles*. 2018;7(1). doi:10.1080/20013078.2018.1535750

267. Waks AG, Winer EP. Breast Cancer Treatment: A Review. *JAMA - J Am Med Assoc*. 2019;321(3):288-300. doi:10.1001/jama.2018.19323

268. Heller DR, Chiu AS, Farrell K, Killelea BK, Lannin DR. Why has breast cancer screening failed to decrease the incidence of de novo stage iv disease? *Cancers (Basel)*. 2019;11(4):1-12. doi:10.3390/cancers11040500

269. Iorio MV, Ferracin M, Liu CG, et al. MicroRNA gene expression deregulation in human breast cancer. *Cancer Res*. 2005;65(16):7065-7070. doi:10.1158/0008-5472.CAN-05-1783

270. Loh HY, Norman BP, Lai KS, Rahman NMANA, Alitheen NBM, Osman MA. The regulatory role of microRNAs in breast cancer. *Int J*

*Mol Sci*. 2019;20(19):1-27. doi:10.3390/ijms20194940

271. Nair MG, Somashekaraiah VM, Ramamurthy V, Prabhu JS, Sridhar TS. miRNAs: Critical mediators of breast cancer metastatic programming. *Exp Cell Res*. 2021;401(1):112518. doi:10.1016/j.yexcr.2021.112518

272. Day EO, Lal A. MicroRNAs and their target gene networks in breast cancer. *Breast Cancer Res*.. 2010;12(2):201. doi: 10.1186/bcr2484

273. Sharma S, Opyrchal M, Lu X. Harnessing tumorous flaws for immune supremacy: is miRNA-155 the weak link in breast cancer progression? *J Clin Invest*. 2022;132(19). doi:10.1172/JCI163010

274. Grimaldi AM, Nuzzo S, Condorelli G, Salvatore M, Incoronato M. Prognostic and clinicopathological significance of mir-155 in breast cancer: A systematic review. *Int J Mol Sci*. 2020;21(16):1-13. doi:10.3390/ijms21165834

275. Petrović N, Kolaković A, Stanković A, et al. MIR-155 expression level changes might be associated with initial phases of breast cancer pathogenesis and lymph-node metastasis. *Cancer Biomarkers*. 2016;16(3):385-394. doi:10.3233/CBM-160577

276. O'Connell RM, Rao DS, Chaudhuri AA, Baltimore D. Physiological and pathological roles for microRNAs in the immune system. *Nat Rev Immunol*. 2010;10(2):111-122. doi:10.1038/nri2708

277. Imani S, Zhang X, Hosseinifard H, Fu S, Fu J. The diagnostic role of microRNA-34a in breast cancer: A systematic review and meta-analysis. *Oncotarget*. 2017;8(14):23177-23187. doi:10.18632/oncotarget.15520

278. Li L, Xie X, Luo J, et al. Targeted expression of miR-34a using the T-VISA system suppresses breast cancer cell growth and invasion. *Mol*

*Ther*. 2012;20(12):2326-2334. doi:10.1038/mt.2012.201

279. Tokumaru Y, Eriko K, Oshi M, et al. High expression of mir-34a associated with less aggressive cancer biology but not with survival in breast cancer. *Int J Mol Sci*. 2020;21(9). doi:10.3390/ijms21093045

280. Li ZH, Weng X, Xiong QY, et al. miR-34a expression in human breast cancer is associated with drug resistance. *Oncotarget*. 2017;8(63):106270-106282. doi:10.18632/oncotarget.22286

281. Ekyalongo RC, Yee D. Revisiting the IGF-1R as a breast cancer target. *npj Precis Oncol*. 2017;1(1):1-6. doi:10.1038/s41698-017-0017-y

282. Obr AE, Kumar S, Chang YJ, et al. Insulin-like growth factor receptor signaling in breast tumor epithelium protects cells from endoplasmic reticulum stress and regulates the tumor microenvironment. *Breast Cancer Res*. 2018;20(1):1-17. doi:10.1186/s13058-018-1063-2

283. Li Y, Kong X, Xuan L, Wang Z, Huang YH. Prolactin and endocrine therapy resistance in breast cancer: The next potential hope for breast cancer treatment. *J Cell Mol Med*. 2021;25(22):10327-10348. doi:10.1111/jcmm.16946

284. Mir R, Al Balawi IA, Abu Duhier FM. Involvement of microRNA-423 gene variability in breast cancer progression in Saudi Arabia. *Asian Pacific J Cancer Prev*. 2018;19(9):2581-2589. doi:10.22034/APJCP.2018.19.9.2581

285. Morales-Pison S, Carrasco V, Gutiérrez-Vera C, et al. Genetic variation in microrna-423 promotes proliferation, migration, invasion, and chemoresistance in breast cancer cells. *Int J Mol Sci*. 2022;23(1). doi:10.3390/ijms23010380

286. Dai T, Zhao X, Li Y, et al. Mir-423 promotes breast cancer invasion by activating nf-κb signaling. *Onco Targets Ther*. 2020;13:5467-5478.

doi:10.2147/OTT.S236514

287. Zhong S, Chen X, Wang D, et al. MicroRNA expression profiles of drug-resistance breast cancer cells and their exosomes. *Oncotarget*. 2016;7(15):19601-19609. doi:10.18632/oncotarget.7481

288. Tang X, Zeng X, Huang Y, et al. MiR-423-5p serves as a diagnostic indicator and inhibits the proliferation and invasion of ovarian cancer. *Exp Ther Med*. 2018;15(6):4723-4730. doi:10.3892/etm.2018.6015

289. Shan G, Gu J, Zhou D, et al. Cancer-associated fibroblast-secreted exosomal miR-423-5p promotes chemotherapy resistance in prostate cancer by targeting GREM2 through the TGF-β signaling pathway. *Exp Mol Med*. 2020;52(11):1809-1822. doi:10.1038/s12276-020-0431-z

290. Albanese M, Chen YFA, Hüls C, et al. *MicroRNAs Are Minor Constituents of Extracellular Vesicles That Are Rarely Delivered to Target Cells*. Vol 17.; 2021. doi:10.1371/journal.pgen.1009951

291. O'Brien K, Breyne K, Ughetto S, Laurent LC, Breakefield XO. RNA delivery by extracellular vesicles in mammalian cells and its applications. *Nat Rev Mol Cell Biol*. 2020;21(10):585-606. doi:10.1038/s41580-020-0251-y

292. Mussbacher M, Pirabe A, Brunnthaler L, Schrottmaier WC, Assinger A. Horizontal MicroRNA Transfer by Platelets – Evidence and Implications. *Front Physiol*. 2021;12(June):1-11. doi:10.3389/fphys.2021.678362

# CHAPTER IX: APPENDIXES

# SUPPLEMENTARY MATERIAL

## FIGURES

Supplementary Figure I: Model summary and coefficients for Eager's Test, and Correlations for Begg's Test. The output of Begg's tests includes the correlation co-efficient Kendall's Tau-b which is associated with p-value

**Model Summary[b]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
|---|---|---|---|---|---|
| 1 | .907[a] | .822 | .733 | .44148 | 1.992 |

a. Predictors: (Constant), SEmiR155
b. Dependent Variable: Log2FCmiR155

**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 4.185 | .910 | | 4.601 | .044 |
| | SEmiR155 | -11.902 | 3.918 | -.907 | -3.038 | .093 |

a. Dependent Variable: Log2FCmiR155

Correlations

| | | | Log2FCmiR155 | VAR00002 |
|---|---|---|---|---|
| Kendall's tau_b | Log2FCmiR155 | Correlation Coefficient | 1.000 | -.913 |
| | | Sig. (2-tailed) | . | .071 |
| | | N | 4 | 4 |
| | VAR00002 | Correlation Coefficient | -.913 | 1.000 |
| | | Sig. (2-tailed) | .071 | . |
| | | N | 4 | 4 |

Supplementary Figure II: Model summary and coefficients for Eager's Test, and Correlations for Begg's Test. The output of Begg's tests includes the correlation co-efficient Kendall's Tau-b which is associated with p-value.

**Model Summary[b]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Durbin-Watson |
|---|---|---|---|---|---|
| 1 | .517[a] | .267 | -.465 | .44380 | 2.976 |

a. Predictors: (Constant), SEmiR34a
b. Dependent Variable: Log2FCmiR34a

**Coefficients[a]**

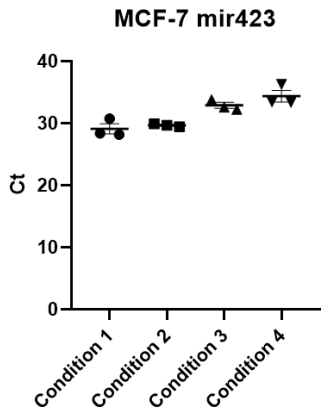
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | |
| 1 | (Constant) | 1.854 | 3.951 | | .469 | .721 |
| | SEmiR34a | -7.245 | 11.996 | -.517 | -.604 | .654 |

a. Dependent Variable: Log2FCmiR34a

**Correlations**

| | | | Log2FCmiR34a | SEmiR34a |
|---|---|---|---|---|
| Kendall's tau_b | Log2FCmiR34a | Correlation Coefficient | 1.000 | -.333 |
| | | Sig. (2-tailed) | . | .602 |
| | | N | 3 | 3 |
| | SEmiR34a | Correlation Coefficient | -.333 | 1.000 |
| | | Sig. (2-tailed) | .602 | . |
| | | N | 3 | 3 |

Supplementary Figure III: miR-423 expression in breast cancer cell line MCF-7 EVs under the four different treatments.



## TABLES

Supplementary Table I: Selective process of the records obtained from entering the keywords in the different databases for systematic review

https://docs.google.com/spreadsheets/d/1PFx0aJmnKEojQ2OZm78Lc-rvpr5-uPQO/edit?usp=sharing&ouid=111817647840889431552&rtpof=true&sd=true

Supplementaty Table II: AUC, Sensibility, Specificity and OS values for the articles included in the meta-analysis.

| Article | microRNA | AUC | AUC signature | Sens | Spec | OS Univar HR (95%CI) | Multivar HR (95% CI) |
|---|---|---|---|---|---|---|---|
| **Cheng et al., 2018** | miR-191-5p | 0.684 | 0.803 | 78.70% | 70.80% | 7.81 (3.79-16.1) | 3.10 0.91-10.43 |
| | miR-489 | 0.608 | | | | 0.33 0.12-0.92 | |
| | miR-214-3p | 0.626 | | | | 0.48 0.25-0.92 | |
| | miR-451a | 0.704 | | | | 1.01 0.16-6.28 | |
| **Ibrahim et al., 2020** | miR-10b | 0.73 | | 53.30% | 100% | x2 (0.01) LR (9.12) | |
| | miR-21 | 0.78 | | 63.30% | 100% | x2 (0.01) LR (9.96) | |
| **Xu et al., 2016** | miR-200a-3p | | | | | | |
| | miR-429 | | | | | | |
| | miR-141-3p | | | | | | |
| | miR-200b | 0.728 (0.629-0.827) | | | | | |
| **Okuno et al., 2021** | miR-98 | | 0.883 (0.807-0.958) | 90.30% | 53.60% | | |
| | miR-22 | | | | | | |
| | miR-223 | | | | | | |
| | miR-155-5p | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Elango et al., 2020** | miR-150-5p | | | | | |
| | miR-146a-5p | | | | | |
| | miR-142-5p | | | | | |
| | miR-200a-3p | | | | | |
| | miR-200b-3p | | | | | |
| | miR-200c-3p | | | | | |
| | miR205-5p | | | | 0.75 (0.61–0.91 | |
| | miR-210-3p | | | | | |
| | miR-214-3p | | | | 0.74 (0.59–0.93) | |
| | miR-141-3p | | | | | |
| | miR-127-3p | | | | | |
| | miR-125a-5p | | | | | |
| | let-7c-5p | | | | | |
| **M'hamed et al., 2017** | miR-153 | | | | | 0.651 (Odds ratio) |
| | miR-10b | | | | | -0.1498 (-0.333-0.333) |
| **Berber et al., 2014** | miR-205 | | 68.80% | 81.30% | | |
| | miR-200c | | 18.8 | 100 % | | |
| **Xie et al., 2017** | miR-645 | 0.721 (95% CI, 0.663–0.779) | 79.10% | 72.80% | 7.25 (4.29–12.24) | 2.49 (1.84–3.38) |
| | miR-652-5p | | | | | |
| | miR-934 | | | | | |
| | miR-30b-5p | | | | | |
| | miR-148a-3p | | | | | |
| | miR-29c-3p | | | | | |
| | miR-29c-5p | | | | | |
| | miR-26a-5p | | | | | |
| | miR361-3p | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Zheng et al., 2015** | miR-106b | Tissue (0.785 ( 0.674 - 0.896 )) Plasma (0.856 (0.775- 0.937 )) | | | Tissue (11.446), Plasma (13.77) | Tissue (4.882, 1.019-23.385) , Plasma (6.926, 1.447-33.143) |
| **Shiino et al., 2018** | miR-629-3p | | 0.75 (0.70–0.80) | 74% | 66% | |
| | miR-4710 | | | | | |
| **Chen et al., 2013** | miR-10b | | 0.8 | 71% | 72% | 2.19 (Odd ratio) |
| | miR-373 | | 0.84 | 68% | 89% | 2.62 (Odd ratio) |
| **Yan et al., 2008** | miR-21 | | | | (P = 0.006, HR = 2.752), | |

Supplementary Table III: Readings and quality values for each sequenced sample. C: controls; MG: early breast cancer; MTX: metastatic breast cancer.

| Sample ID | Total reads | Reads mapped | Reads unmapped | %mapped | GC (%) | Q20(%) | Q30(%) |
|---|---|---|---|---|---|---|---|
| **C1** | 9111113 | 6058529 | 3052584 | 66.496 | 53.52 | 98.38 | 96.33 |
| **C3** | 11494100 | 6938713 | 4555387 | 60.368 | 52.38 | 98.36 | 96.23 |
| **C4** | 8995578 | 5049394 | 3946184 | 56.132 | 53.61 | 98.26 | 96.25 |
| **C6** | 9395833 | 4885651 | 4510182 | 51.998 | 54.07 | 98.35 | 96.18 |
| **C8** | 8346173 | 4700439 | 3645734 | 56.318 | 54.65 | 98.35 | 96.18 |
| **C9** | 8463703 | 3986547 | 4477156 | 47.102 | 54.89 | 98.36 | 96.24 |
| **C10** | 8453667 | 5038914 | 3414753 | 59.606 | 53.72 | 98.33 | 96.24 |
| **C12** | 6169771 | 3435088 | 2734683 | 55.676 | 53.6 | 98.35 | 96.22 |
| **C14** | 7637371 | 4782414 | 2854957 | 62.619 | 54.46 | 98.29 | 96.16 |
| **MG 142** | 18302423 | 2447347 | 15855076 | 13.372 | 51.78 | 97.4 | 94.3 |
| **MG 144** | 8622771 | 6769604 | 1853167 | 78.508 | 57.71 | 98.35 | 96.28 |
| **MG 145** | 4451617 | 1729617 | 2722000 | 38.854 | 53.92 | 98.21 | 95.88 |
| **MG 151** | 5908769 | 3429236 | 2479533 | 58.036 | 54.11 | 98.34 | 96.19 |
| **MG 156** | 6079203 | 3449895 | 2629308 | 56.749 | 55.42 | 98.3 | 96.15 |
| **MG 158** | 9394243 | 5233124 | 4161119 | 55.706 | 54.9 | 98.27 | 96.07 |
| **MG 163** | 7828540 | 4185642 | 3642898 | 53.466 | 55.26 | 98.27 | 96.05 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **MG 167** | 6253060 | 2251464 | 4001596 | 36.006 | 53.69 | 98.27 | 96.05 |
| **MG 168** | 13666244 | 2466784 | 11199460 | 18.05 | 53.23 | 98.26 | 96 |
| **MG 171** | 5631894 | 1651595 | 3980299 | 29.326 | 51.93 | 98.29 | 96.07 |
| **MT X21** | 9417755 | 6569798 | 2847957 | 69.76 | 58.28 | 98.36 | 96.28 |
| **MT X22** | 8784347 | 3759836 | 5024511 | 42.802 | 54.5 | 97.4 | 94.2 |
| **MT X26** | 13577133 | 8542220 | 5034913 | 62.916 | 54.4 | 97.5 | 94.6 |
| **MT X27** | 6492638 | 3594379 | 2898259 | 55.361 | 54.04 | 98.29 | 96.09 |
| **MT X28** | 12397320 | 9683076 | 2714244 | 78.106 | 54.14 | 98.4 | 96.39 |
| **MT X29** | 14120398 | 11455995 | 2664403 | 81.131 | 57.69 | 98.49 | 96.58 |

Supplementary Table IV：Differential expressed miRNAs between Early vs Metastatic comparative

| | logFC | logCPM | PValue | FDR |
|---|---|---|---|---|
| **hsa-miR-183-5p\|MIMAT0000261\|Homo-sapiens\|miR-183-5p** | 1.87964767 | 7.18267275 | 6.18E-08 | 0.00016417 |
| **hsa-miR-200a-3p\|MIMAT0000682\|Homo-sapiens\|miR-200a-3p** | 3.36009438 | 3.55457593 | 9.75E-07 | 0.00129 46 |
| **hsa-miR-100-5p\|MIMAT0000098\|Homo-sapiens\|miR-100-5p** | 2.73977306 | 8.83422825 | 1.91E-06 | 0.00168 96 |
| **hsa-miR-500a-3p\|MIMAT0002871\|Homo-sapiens\|miR-500a-3p** | 1.22257475 | 7.36354916 | 4.15E-06 | 0.00246787 |
| **hsa-miR-122-5p\|MIMAT0000421\|Homo-sapiens\|miR-122-5p** | 5.53323777 | 8.45495 61 | 4.98E-06 | 0.00246787 |
| **hsa-miR-192-5p\|MIMAT0000222\|Homo-sapiens\|miR-192-5p** | 2.14206694 | 12.6900644 | 5.57E-06 | 0.00246787 |
| **hsa-miR-141-3p\|MIMAT0000432\|Homo-sapiens\|miR-141-3p** | 2.40611009 | 7.02897 28 | 8.49E-06 | 0.00322 24 |
| **hsa-let-7f-5p\|MIMAT0000067\|Homo-sapiens\|let-7f-5p** | -1.3150082 | 14.0507 2 | 1.37E-05 | 0.00454676 |
| **hsa-miR-193b-5p\|MIMAT0004767\|Homo-sapiens\|miR-193b-5p** | 3.90653202 | 2.86220143 | 3.48E-05 | 0.01026079 |
| **hsa-miR-92b-3p\|MIMAT0003218\|Homo-sapiens\|miR-92b-3p** | 1.28074009 | 10.1600 8 | 4.54E-05 | 0.01206293 |
| **hsa-miR-125b-5p\|MIMAT0000423\|Homo-sapiens\|miR-125b-5p** | 2.15818884 | 6.51511556 | 5.60E-05 | 0.01351435 |
| **hsa-miR-99a-5p\|MIMAT0000097\|Homo-sapiens\|miR-99a-5p** | 2.73805007 | 7.00052589 | 6.62E-05 | 0.01465577 |
| **hsa-miR-125b-2-3p\|MIMAT0004603\|Homo-sapiens\|miR-125b-2-3p** | 2.80122067 | 5.28208487 | 8.88E-05 | 0.01814 05 |
| **hsa-let-7a-5p\|MIMAT0000062\|Homo-sapiens\|let-7a-5p** | -1.2138772 | 12.9620 8 | 0.00011481 | 0.01994091 |
| **hsa-miR-29c-3p\|MIMAT0000681\|Homo-sapiens\|miR-29c-3p** | 1.19479425 | 6.51419957 | 0.00011738 | 0.01994091 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-222-3p\|MIMAT0000279\|Homo-sapiens\|miR-222-3p** | -1.1724081 | 9.18018019 | 0.00012013 | 0.01994091 |
| **hsa-miR-16-5p\|MIMAT0000069\|Homo-sapiens\|miR-16-5p** | 1.31784919 | 14.9045746 | 0.00014304 | 0.02234851 |
| **hsa-miR-193b-3p\|MIMAT0002819\|Homo-sapiens\|miR-193b-3p** | 3.85409001 | 2.44700732 | 0.00018699 | 0.02689351 |
| **hsa-miR-483-5p\|MIMAT0004761\|Homo-sapiens\|miR-483-5p** | 3.43687742 | 2.50327559 | 0.00019239 | 0.02689351 |
| **hsa-miR-188-5p\|MIMAT0000457\|Homo-sapiens\|miR-188-5p** | 3.10196791 | 2.45947978 | 0.00021323 | 0.02831697 |
| **hsa-miR-483-3p\|MIMAT0002173\|Homo-sapiens\|miR-483-3p** | 3.70919864 | 2.44608964 | 0.00026986 | 0.03413092 |
| **hsa-miR-8061\|MIMAT0030988\|Homo-sapiens\|miR-8061** | -3.1420761 | 3.39587112 | 0.00029688 | 0.03584099 |
| **hsa-miR-375-3p\|MIMAT0000728\|Homo-sapiens\|miR-375-3p** | 2.45543572 | 9.03593736 | 0.00041063 | 0.04741933 |
| **hsa-miR-1299\|MIMAT0005887\|Homo-sapiens\|miR-1299** | 3.50179045 | 2.64861233 | 0.00056661 | 0.05751296 |
| **hsa-miR-186-5p\|MIMAT0000456\|Homo-sapiens\|miR-186-5p** | 0.76639975 | 11.9812271 | 0.00057724 | 0.05751296 |
| **hsa-miR-451a\|MIMAT0001631\|Homo-sapiens\|miR-451a** | 1.48139856 | 14.6913498 | 0.00058093 | 0.05751296 |
| **hsa-miR-615-3p\|MIMAT0003283\|Homo-sapiens\|miR-615-3p** | 3.77201806 | 2.34075 32 | 0.00058466 | 0.05751296 |
| **hsa-miR-365b-3p\|MIMAT0022834\|Homo-sapiens\|miR-365b-3p** | 2.27471903 | 3.11482562 | 0.00065287 | 0.05803531 |
| **hsa-miR-365a-3p\|MIMAT0000710\|Homo-sapiens\|miR-365a-3p** | 2.27471253 | 3.11482562 | 0.00065311 | 0.05803531 |
| **hsa-miR-501-3p\|MIMAT0004774\|Homo-sapiens\|miR-501-3p** | 1.51512779 | 7.85257347 | 0.00065552 | 0.05803531 |
| **hsa-miR-98-5p\|MIMAT0000096\|Homo-sapiens\|miR-98-5p** | -1.1410463 | 8.48047385 | 0.00086168 | 0.07382 61 |
| **hsa-miR-11401\|MIMAT0044658\|Homo-sapiens\|miR-11401** | 2.83159165 | 2.49577704 | 0.00100286 | 0.08111069 |
| **hsa-miR-122-3p\|MIMAT0004590\|Homo-sapiens\|miR-122-3p** | 3.7089792 | 2.32175252 | 0.00100778 | 0.08111069 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-33b-5p\|MIMAT0003301\|Homo-sapiens\|miR-33b-5p** | 1.516429 | 4.10501733 | 0.0012046 | 0.09410029 |
| **hsa-miR-6815-5p\|MIMAT0027530\|Homo-sapiens\|miR-6815-5p** | 3.69102553 | 1.90016383 | 0.00127253 | 0.09656671 |
| **hsa-miR-96-5p\|MIMAT0000095\|Homo-sapiens\|miR-96-5p** | 1.62224208 | 5.72117116 | 0.00132417 | 0.09769405 |
| **hsa-miR-130b-3p\|MIMAT0000691\|Homo-sapiens\|miR-130b-3p** | 0.76721177 | 8.3926419 | 0.00139611 | 0.10021829 |
| **hsa-miR-502-3p\|MIMAT0004775\|Homo-sapiens\|miR-502-3p** | 0.96910314 | 5.14618893 | 0.00152685 | 0.10671847 |
| **hsa-let-7e-5p\|MIMAT0000066\|Homo-sapiens\|let-7e-5p** | -1.1152239 | 7.94613566 | 0.001633 | 0.10971308 |
| **hsa-miR-99b-5p\|MIMAT0000689\|Homo-sapiens\|miR-99b-5p** | 1.61281115 | 10.4298993 | 0.00165231 | 0.10971308 |
| **hsa-miR-340-5p\|MIMAT0004692\|Homo-sapiens\|miR-340-5p** | -0.9974466 | 8.90892494 | 0.00198945 | 0.12769607 |
| **hsa-miR-877-5p\|MIMAT0004949\|Homo-sapiens\|miR-877-5p** | 1.31185025 | 5.70308199 | 0.00201929 | 0.12769607 |
| **hsa-miR-885-5p\|MIMAT0004947\|Homo-sapiens\|miR-885-5p** | 3.04790994 | 3.3794144 | 0.00217862 | 0.13456748 |
| **hsa-let-7i-5p\|MIMAT0000415\|Homo-sapiens\|let-7i-5p** | -0.5229841 | 12.538629 | 0.00226853 | 0.13693684 |
| **hsa-miR-1285-3p\|MIMAT0005876\|Homo-sapiens\|miR-1285-3p** | 1.69615615 | 5.67398439 | 0.0024557 | 0.14233679 |
| **hsa-miR-6852-5p\|MIMAT0027604\|Homo-sapiens\|miR-6852-5p** | -1.082258 | 5.97417949 | 0.00246517 | 0.14233679 |
| **hsa-miR-1246\|MIMAT0005898\|Homo-sapiens\|miR-1246** | 2.67351449 | 2.44718051 | 0.00267199 | 0.1509961 |
| **hsa-miR-1304-3p\|MIMAT0022720\|Homo-sapiens\|miR-1304-3p** | -0.9046419 | 5.97213987 | 0.00294919 | 0.16032084 |
| **hsa-miR-331-5p\|MIMAT0004700\|Homo-sapiens\|miR-331-5p** | 1.50929591 | 3.05659972 | 0.00295773 | 0.16032084 |
| **hsa-miR-874-3p\|MIMAT0004911\|Homo-sapiens\|miR-874-3p** | 1.43729352 | 3.60156383 | 0.003044 | 0.16169722 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-1273h-3p\|MIMAT0030416\|Homo-sapiens\|miR-1273h-3p** | -0.9691526 | 5.35685214 | 0.00355907 | 0.18153369 |
| **hsa-miR-210-3p\|MIMAT0000267\|Homo-sapiens\|miR-210-3p** | 1.3172385 | 7.03925914 | 0.00359 35 | 0.18153369 |
| **hsa-miR-148b-3p\|MIMAT0000759\|Homo-sapiens\|miR-148b-3p** | -1.0217723 | 9.64724 08 | 0.00362247 | 0.18153369 |
| **hsa-miR-1277-5p\|MIMAT0022724\|Homo-sapiens\|miR-1277-5p** | -2.1695701 | 3.23364891 | 0.00409983 | 0.20165112 |
| **hsa-miR-342-3p\|MIMAT0000753\|Homo-sapiens\|miR-342-3p** | 1.08896839 | 6.95075553 | 0.00429 76 | 0.20753489 |
| **hsa-miR-486-5p\|MIMAT0002177\|Homo-sapiens\|miR-486-5p** | 0.99490699 | 19.2362797 | 0.00450163 | 0.21350571 |
| **hsa-miR-205-5p\|MIMAT0000266\|Homo-sapiens\|miR-205-5p** | 1.83191063 | 4.76221273 | 0.00463 42 | 0.21593754 |
| **hsa-miR-26b-3p\|MIMAT0004500\|Homo-sapiens\|miR-26b-3p** | 0.79051881 | 6.26419512 | 0.00481767 | 0.22061582 |
| **hsa-miR-130a-3p\|MIMAT0000425\|Homo-sapiens\|miR-130a-3p** | 0.72559107 | 9.73074663 | 0.00564101 | 0.25192707 |
| **hsa-miR-106b-5p\|MIMAT0000680\|Homo-sapiens\|miR-106b-5p** | 1.11647808 | 7.61505071 | 0.00569112 | 0.25192707 |
| **hsa-let-7d-5p\|MIMAT0000065\|Homo-sapiens\|let-7d-5p** | -0.7840655 | 9.40609233 | 0.00688854 | 0.29993384 |
| **hsa-miR-33a-3p\|MIMAT0004506\|Homo-sapiens\|miR-33a-3p** | 3.31082948 | 1.91016736 | 0.00734703 | 0.31473745 |
| **hsa-miR-374a-3p\|MIMAT0004688\|Homo-sapiens\|miR-374a-3p** | -1.0123748 | 5.16814953 | 0.00790623 | 0.33331669 |
| **hsa-miR-4646-5p\|MIMAT0019707\|Homo-sapiens\|miR-4646-5p** | 3.26081158 | 1.91986555 | 0.00844598 | 0.34145266 |
| **hsa-miR-149-5p\|MIMAT0000450\|Homo-sapiens\|miR-149-5p** | 3.21254095 | 1.81195586 | 0.00877699 | 0.34145266 |
| **hsa-miR-548az-5p\|MIMAT0025456\|Homo-sapiens\|miR-548az-5p** | 2.73469503 | 2.08390275 | 0.00895116 | 0.34145266 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-26b-5p\|MIMAT0000083\|Homo-sapiens\|miR-26b-5p** | -0.7773529 | 10.5233865 | 0.00896346 | 0.34145266 |
| **hsa-miR-181c-3p\|MIMAT0004559\|Homo-sapiens\|miR-181c-3p** | -0.6921532 | 6.18677259 | 0.00897933 | 0.34145266 |
| **hsa-miR-194-3p\|MIMAT0004671\|Homo-sapiens\|miR-194-3p** | 2.78935425 | 1.95037899 | 0.00900407 | 0.34145266 |
| **hsa-miR-28-5p\|MIMAT0000085\|Homo-sapiens\|miR-28-5p** | -0.85308 | 6.96288237 | 0.00911391 | 0.34145266 |
| **hsa-miR-1247-5p\|MIMAT0005899\|Homo-sapiens\|miR-1247-5p** | 2.74029009 | 2.29696731 | 0.00912769 | 0.34145266 |
| **hsa-miR-4732-5p\|MIMAT0019855\|Homo-sapiens\|miR-4732-5p** | 1.69710903 | 3.51796142 | 0.00931719 | 0.34289962 |
| **hsa-miR-328-3p\|MIMAT0000752\|Homo-sapiens\|miR-328-3p** | -0.7630192 | 6.34666503 | 0.00942458 | 0.34289962 |
| **hsa-miR-1468-5p\|MIMAT0006789\|Homo-sapiens\|miR-1468-5p** | 1.12406379 | 5.16202964 | 0.00996936 | 0.35781935 |
| **hsa-miR-30a-5p\|MIMAT0000087\|Homo-sapiens\|miR-30a-5p** | 1.08671795 | 10.9437677 | 0.01065353 | 0.37557954 |
| **hsa-miR-10226\|MIMAT0041128\|Homo-sapiens\|miR-10226** | -3.7238088 | 2.6113875 | 0.01074701 | 0.37557954 |
| **hsa-miR-196b-5p\|MIMAT0001080\|Homo-sapiens\|miR-196b-5p** | -1.2713989 | 4.12521213 | 0.01095367 | 0.37783049 |
| **hsa-miR-4446-3p\|MIMAT0018965\|Homo-sapiens\|miR-4446-3p** | -1.0311723 | 5.495024 | 0.01212036 | 0.41271367 |
| **hsa-miR-9-5p\|MIMAT0000441\|Homo-sapiens\|miR-9-5p** | -2.7471531 | 2.91454703 | 0.01231792 | 0.41413167 |
| **hsa-miR-19b-3p\|MIMAT0000074\|Homo-sapiens\|miR-19b-3p** | 0.80609654 | 8.42615674 | 0.01327243 | 0.44064045 |
| **hsa-miR-194-5p\|MIMAT0000460\|Homo-sapiens\|miR-194-5p** | 1.01541762 | 5.85117296 | 0.013611 | 0.44624479 |
| **hsa-miR-125a-5p\|MIMAT0000443\|Homo-sapiens\|miR-125a-5p** | 1.0125543 | 9.21321215 | 0.01377714 | 0.44624479 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-505-3p\|MIMAT0002876\|Homo-sapiens\|miR-505-3p** | 1.31431132 | 3.73387747 | 0.01456583 | 0.46610648 |
| **hsa-miR-592\|MIMAT0003260\|Homo-sapiens\|miR-592** | 3.10630824 | 1.87917328 | 0.01512809 | 0.47833576 |
| **hsa-miR-378a-3p\|MIMAT0000732\|Homo-sapiens\|miR-378a-3p** | 0.89756891 | 10.2437215 | 0.01534569 | 0.47950764 |
| **hsa-miR-22-3p\|MIMAT0000077\|Homo-sapiens\|miR-22-3p** | 0.63164656 | 15.0288099 | 0.01558669 | 0.48137507 |
| **hsa-miR-34a-5p\|MIMAT0000255\|Homo-sapiens\|miR-34a-5p** | 1.8834351 | 2.83001531 | 0.01616598 | 0.49352679 |
| **hsa-miR-30e-3p\|MIMAT0000693\|Homo-sapiens\|miR-30e-3p** | -0.5449845 | 7.73594479 | 0.01680767 | 0.50728614 |
| **hsa-miR-320c\|MIMAT0005793\|Homo-sapiens\|miR-320c** | 1.13321142 | 4.24369154 | 0.0174366 | 0.51621333 |
| **hsa-miR-369-3p\|MIMAT0000721\|Homo-sapiens\|miR-369-3p** | -2.4915631 | 2.77234965 | 0.01749217 | 0.51621333 |
| **hsa-miR-320b\|MIMAT0005792\|Homo-sapiens\|miR-320b** | 0.78629702 | 8.85869913 | 0.01793707 | 0.51720052 |
| **hsa-miR-27b-3p\|MIMAT0000419\|Homo-sapiens\|miR-27b-3p** | 0.9698597 | 13.3261292 | 0.01798785 | 0.51720052 |
| **hsa-miR-1290\|MIMAT0005880\|Homo-sapiens\|miR-1290** | 3.34102748 | 1.81055491 | 0.01818478 | 0.51720052 |
| **hsa-miR-423-5p\|MIMAT0004748\|Homo-sapiens\|miR-423-5p** | 0.746326 | 14.1636247 | 0.01830454 | 0.51720052 |
| **hsa-miR-5706\|MIMAT0022500\|Homo-sapiens\|miR-5706** | 3.05733334 | 1.91085292 | 0.01888766 | 0.52380888 |
| **hsa-miR-19a-3p\|MIMAT0000073\|Homo-sapiens\|miR-19a-3p** | 0.81034897 | 6.37691043 | 0.01893285 | 0.52380888 |
| **hsa-miR-199a-3p\|MIMAT0000232\|Homo-sapiens\|miR-199a-3p** | -0.5061146 | 10.0788492 | 0.02039892 | 0.55412566 |
| **hsa-miR-199b-3p\|MIMAT0004563\|Homo-sapiens\|miR-199b-3p** | -0.5061185 | 10.0788492 | 0.0204459 | 0.55412566 |
| **hsa-miR-331-3p\|MIMAT0000760\|Homo-sapiens\|miR-331-3p** | 1.87736645 | 2.36151704 | 0.02066475 | 0.55439984 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-671-3p\|MIMAT0004819\|Homo-sapiens\|miR-671-3p** | -0.6682366 | 6.61066632 | 0.02223682 | 0.59060984 |
| **hsa-miR-11400\|MIMAT0044657\|Homo-sapiens\|miR-11400** | -0.7144306 | 6.15939453 | 0.02312083 | 0.60800902 |
| **hsa-miR-6802-5p\|MIMAT0027504\|Homo-sapiens\|miR-6802-5p** | 3.66524874 | 1.88969066 | 0.02361863 | 0.61180452 |
| **hsa-miR-12136\|MIMAT0049032\|Homo-sapiens\|miR-12136** | -2.4598988 | 3.66302998 | 0.02372585 | 0.61180452 |
| **hsa-miR-202-3p\|MIMAT0002811\|Homo-sapiens\|miR-202-3p** | 3.14945648 | 1.88348436 | 0.02429923 | 0.62056506 |
| **hsa-miR-126-3p\|MIMAT0000445\|Homo-sapiens\|miR-126-3p** | -0.4859942 | 9.66834117 | 0.02523791 | 0.63596824 |
| **hsa-miR-3200-5p\|MIMAT0017392\|Homo-sapiens\|miR-3200-5p** | 1.33604917 | 2.92542527 | 0.02538126 | 0.63596824 |
| **hsa-miR-582-3p\|MIMAT0004797\|Homo-sapiens\|miR-582-3p** | -1.1115224 | 6.01931208 | 0.02672767 | 0.6621694 |
| **hsa-miR-363-3p\|MIMAT0000707\|Homo-sapiens\|miR-363-3p** | 0.89910484 | 10.380701 | 0.02730 97 | 0.6621694 |
| **hsa-miR-146b-3p\|MIMAT0004766\|Homo-sapiens\|miR-146b-3p** | 1.01497436 | 4.89048818 | 0.02739317 | 0.6621694 |
| **hsa-miR-7706\|MIMAT0030021\|Homo-sapiens\|miR-7706** | 0.85225749 | 4.97902636 | 0.02742418 | 0.6621694 |
| **hsa-let-7g-5p\|MIMAT0000414\|Homo-sapiens\|let-7g-5p** | -0.6347506 | 10.8157853 | 0.02802391 | 0.67055412 |
| **hsa-miR-1260a\|MIMAT0005911\|Homo-sapiens\|miR-1260a** | -1.1687913 | 4.17451137 | 0.02867599 | 0.6800305 |
| **hsa-miR-574-3p\|MIMAT0003239\|Homo-sapiens\|miR-574-3p** | 0.92045548 | 4.52208729 | 0.02899699 | 0.68155763 |
| **hsa-miR-191-3p\|MIMAT0001618\|Homo-sapiens\|miR-191-3p** | -0.8106586 | 4.55395863 | 0.02954465 | 0.6828037 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-130b-5p\|MIMAT0004680\|Homo-sapiens\|miR-130b-5p** | -0.8146477 | 5.89096342 | 0.02956417 | 0.6828037 |
| **hsa-miR-27a-5p\|MIMAT0004501\|Homo-sapiens\|miR-27a-5p** | -1.8162578 | 3.29188545 | 0.03023131 | 0.69219271 |
| **hsa-miR-550a-3p\|MIMAT0003257\|Homo-sapiens\|miR-550a-3p** | 0.82962532 | 5.20895335 | 0.0309234 | 0.69905692 |
| **hsa-miR-484\|MIMAT0002174\|Homo-sapiens\|miR-484** | 0.65784347 | 10.5331889 | 0.0310575 | 0.69905692 |
| **hsa-miR-421\|MIMAT0003339\|Homo-sapiens\|miR-421** | 0.46063603 | 7.52006564 | 0.03241932 | 0.72357733 |
| **hsa-miR-216a-5p\|MIMAT0000273\|Homo-sapiens\|miR-216a-5p** | 2.72742742 | 1.88215611 | 0.03337996 | 0.73880978 |
| **hsa-miR-4433b-5p\|MIMAT0030413\|Homo-sapiens\|miR-4433b-5p** | -0.8264835 | 6.90322098 | 0.03477111 | 0.76324025 |
| **hsa-miR-425-5p\|MIMAT0003393\|Homo-sapiens\|miR-425-5p** | 0.50369403 | 9.75805703 | 0.03642592 | 0.79116032 |
| **hsa-miR-345-5p\|MIMAT0000772\|Homo-sapiens\|miR-345-5p** | 0.68401362 | 8.53437473 | 0.03671101 | 0.79116032 |
| **hsa-miR-7705\|MIMAT0030020\|Homo-sapiens\|miR-7705** | 1.77277288 | 2.27012048 | 0.0369367 | 0.79116032 |
| **hsa-miR-3651\|MIMAT0018071\|Homo-sapiens\|miR-3651** | -3.2170809 | 2.27537885 | 0.03737188 | 0.79215081 |
| **hsa-miR-769-5p\|MIMAT0003886\|Homo-sapiens\|miR-769-5p** | -0.475075 | 7.36521816 | 0.03757944 | 0.79215081 |
| **hsa-miR-4751\|MIMAT0019888\|Homo-sapiens\|miR-4751** | 3.61591536 | 1.79122654 | 0.03879663 | 0.81136898 |
| **hsa-miR-32-5p\|MIMAT0000090\|Homo-sapiens\|miR-32-5p** | 0.90283039 | 5.94891895 | 0.03934878 | 0.81546587 |
| **hsa-miR-10a-5p\|MIMAT0000253\|Homo-sapiens\|miR-10a-5p** | 0.80606703 | 12.8684411 | 0.03960659 | 0.81546587 |
| **hsa-miR-221-3p\|MIMAT0000278\|Homo-sapiens\|miR-221-3p** | -0.5785662 | 9.51222115 | 0.04019635 | 0.8194987 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-93-3p\|MIMAT0004509\|Homo-sapiens\|miR-93-3p** | 0.93726345 | 3.99341133 | 0.04066822 | 0.8194987 |
| **hsa-miR-652-3p\|MIMAT0003322\|Homo-sapiens\|miR-652-3p** | 0.7043325 | 8.09834632 | 0.0407281 | 0.8194987 |
| **hsa-miR-624-5p\|MIMAT0003293\|Homo-sapiens\|miR-624-5p** | 2.00402049 | 2.22772105 | 0.04166893 | 0.83212543 |
| **hsa-miR-30c-1-3p\|MIMAT0004674\|Homo-sapiens\|miR-30c-1-3p** | -0.8276411 | 4.39842682 | 0.04299206 | 0.85214106 |
| **hsa-miR-182-5p\|MIMAT0000259\|Homo-sapiens\|miR-182-5p** | 0.64084 | 11.7899958 | 0.04367532 | 0.85927145 |
| **hsa-miR-29a-3p\|MIMAT0000086\|Homo-sapiens\|miR-29a-3p** | 0.75349968 | 7.82312575 | 0.04414567 | 0.862139 |
| **hsa-miR-185-5p\|MIMAT0000455\|Homo-sapiens\|miR-185-5p** | 1.08406608 | 5.86443828 | 0.04954305 | 0.96048426 |

Supplementary Table V: Different expressed miRNAs between Normal vs Early comparative

| | logFC | logCPM | PValue | FDR |
|---|---|---|---|---|
| **hsa-miR-144-5p\|MIMAT0004600\|Homo-sapiens\|miR-144-5p** | -0.9141744 | 9.70413685 | 2.08E-05 | 0.05528025 |
| **hsa-miR-3651\|MIMAT0018071\|Homo-sapiens\|miR-3651** | 3.94731282 | 2.27537885 | 0.00214514 | 1 |
| **hsa-miR-423-5p\|MIMAT0004748\|Homo-sapiens\|miR-423-5p** | 0.88012317 | 14.1636247 | 0.00259661 | 1 |
| **hsa-miR-3174\|MIMAT0015051\|Homo-sapiens\|miR-3174** | -2.5815534 | 2.14429158 | 0.00320488 | 1 |
| **hsa-miR-190a-5p\|MIMAT0000458\|Homo-sapiens\|miR-190a-5p** | -1.5800419 | 2.61100787 | 0.00345536 | 1 |
| **hsa-miR-26a-2-3p\|MIMAT0004681\|Homo-sapiens\|miR-26a-2-3p** | -3.1350912 | 1.94483398 | 0.00409908 | 1 |
| **hsa-miR-182-5p\|MIMAT0000259\|Homo-sapiens\|miR-182-5p** | -0.8306516 | 11.7899958 | 0.00409968 | 1 |
| **hsa-miR-378i\|MIMAT0019074\|Homo-sapiens\|miR-378i** | -1.1948965 | 4.40348348 | 0.00745416 | 1 |
| **hsa-miR-5010-5p\|MIMAT0021043\|Homo-sapiens\|miR-5010-5p** | 1.16586913 | 4.2347821 | 0.00759014 | 1 |
| **hsa-miR-6892-5p\|MIMAT0027684\|Homo-sapiens\|miR-6892-5p** | -2.8462506 | 1.90075519 | 0.0077908 | 1 |
| **hsa-miR-184\|MIMAT0000454\|Homo-sapiens\|miR-184** | 2.67403618 | 3.37798717 | 0.00795426 | 1 |
| **hsa-miR-33b-5p\|MIMAT0003301\|Homo-sapiens\|miR-33b-5p** | -1.1048971 | 4.10501733 | 0.00842445 | 1 |
| **hsa-miR-580-3p\|MIMAT0003245\|Homo-sapiens\|miR-580-3p** | -2.5778088 | 1.91060413 | 0.00855115 | 1 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-1306-5p\|MIMAT0022726\|Homo-sapiens\|miR-1306-5p** | -1.958036 | 2.44119029 | 0.00961633 | 1 |
| **hsa-miR-320b\|MIMAT0005792\|Homo-sapiens\|miR-320b** | 0.72355233 | 8.85869913 | 0.01808211 | 1 |
| **hsa-miR-93-5p\|MIMAT0000093\|Homo-sapiens\|miR-93-5p** | -0.5097253 | 10.7055883 | 0.01812107 | 1 |
| **hsa-miR-8061\|MIMAT0030988\|Homo-sapiens\|miR-8061** | 1.46779941 | 3.39587112 | 0.01822 69 | 1 |
| **hsa-miR-605-3p\|MIMAT0026621\|Homo-sapiens\|miR-605-3p** | 2.64160991 | 2.16367548 | 0.01980013 | 1 |
| **hsa-miR-1262\|MIMAT0005914\|Homo-sapiens\|miR-1262** | -2.4195171 | 2.01990699 | 0.01995276 | 1 |
| **hsa-miR-4772-5p\|MIMAT0019926\|Homo-sapiens\|miR-4772-5p** | -1.4771492 | 2.90932 84 | 0.02002751 | 1 |
| **hsa-miR-1299\|MIMAT0005887\|Homo-sapiens\|miR-1299** | -2.253584 | 2.64861233 | 0.02288683 | 1 |
| **hsa-miR-30c-2-3p\|MIMAT0004550\|Homo-sapiens\|miR-30c-2-3p** | 1.45357 43 | 3.15571053 | 0.02478644 | 1 |
| **hsa-miR-11401\|MIMAT0044658\|Homo-sapiens\|miR-11401** | -1.9271492 | 2.49577704 | 0.02488109 | 1 |
| **hsa-miR-548bc\|MIMAT0039765\|Homo-sapiens\|miR-548bc** | -1.1512153 | 2.75003526 | 0.02531659 | 1 |
| **hsa-miR-543\|MIMAT0004954\|Homo-sapiens\|miR-543** | 1.67856609 | 2.45795505 | 0.02562071 | 1 |
| **hsa-miR-4286\|MIMAT0016916\|Homo-sapiens\|miR-4286** | 0.9960536 | 4.21173909 | 0.02719041 | 1 |
| **hsa-miR-183-5p\|MIMAT0000261\|Homo-sapiens\|miR-183-5p** | -0.6902564 | 7.18267275 | 0.02975689 | 1 |
| **hsa-miR-12136\|MIMAT0049032\|Homo-sapiens\|miR-12136** | 1.80070886 | 3.66302998 | 0.03197859 | 1 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-146a-5p\|MIMAT0000449\|Homo-sapiens\|miR-146a-5p** | 0.86944967 | 11.8027856 | 0.03434631 | 1 |
| **hsa-miR-548d-5p\|MIMAT0004812\|Homo-sapiens\|miR-548d-5p** | -0.7252084 | 3.57461624 | 0.03518552 | 1 |
| **hsa-miR-199a-3p\|MIMAT0000232\|Homo-sapiens\|miR-199a-3p** | -0.3952492 | 10.0788492 | 0.03639565 | 1 |
| **hsa-miR-199b-3p\|MIMAT0004563\|Homo-sapiens\|miR-199b-3p** | -0.3952499 | 10.0788492 | 0.0364661 | 1 |
| **hsa-miR-548ay-5p\|MIMAT0025452\|Homo-sapiens\|miR-548ay-5p** | -0.7518364 | 3.37164745 | 0.03716031 | 1 |
| **hsa-miR-760\|MIMAT0004957\|Homo-sapiens\|miR-760** | -1.3975785 | 2.3547162 | 0.0373659 | 1 |
| **hsa-miR-146b-5p\|MIMAT0002809\|Homo-sapiens\|miR-146b-5p** | 0.90612305 | 10.1287112 | 0.04157754 | 1 |
| **hsa-miR-548ad-5p\|MIMAT0032114\|Homo-sapiens\|miR-548ad-5p** | -0.8006226 | 3.15550773 | 0.04242463 | 1 |
| **hsa-miR-548ae-5p\|MIMAT0032115\|Homo-sapiens\|miR-548ae-5p** | -0.8006293 | 3.15550773 | 0.04244369 | 1 |
| **hsa-miR-589-3p\|MIMAT0003256\|Homo-sapiens\|miR-589-3p** | 1.45675821 | 2.89339304 | 0.04353574 | 1 |
| **hsa-miR-454-3p\|MIMAT0003885\|Homo-sapiens\|miR-454-3p** | -0.7400652 | 4.97523339 | 0.04796849 | 1 |

Supplementary Table VI: Differential expressed miRNAs between Normal vs Metastatic comparative

| | logFC | logCPM | PValue | FDR |
|---|---|---|---|---|
| **hsa-let-7f-5p\|MIMAT0000067\|Homo-sapiens\|let-7f-5p** | -1.604887 | 14.05072 | 2.30E-07 | 0.00037705 |
| **hsa-let-7a-5p\|MIMAT0000062\|Homo-sapiens\|let-7a-5p** | -1.669016 | 12.96208 | 2.84E-07 | 0.00037705 |
| **hsa-miR-423-5p\|MIMAT0004748\|Homo-sapiens\|miR-423-5p** | 1.62644786 | 14.1636247 | 6.05E-07 | 0.00049493 |
| **hsa-miR-144-5p\|MIMAT0004600\|Homo-sapiens\|miR-144-5p** | -1.2769296 | 9.70413685 | 7.45E-07 | 0.00049493 |
| **hsa-miR-141-3p\|MIMAT0000432\|Homo-sapiens\|miR-141-3p** | 2.5711904 | 7.0289728 | 4.54E-06 | 0.00240979 |
| **hsa-miR-100-5p\|MIMAT0000098\|Homo-sapiens\|miR-100-5p** | 2.66884063 | 8.83422825 | 7.48E-06 | 0.00331016 |
| **hsa-miR-320b\|MIMAT0005792\|Homo-sapiens\|miR-320b** | 1.50981171 | 8.85869913 | 1.01E-05 | 0.00383268 |
| **hsa-miR-92b-3p\|MIMAT0003218\|Homo-sapiens\|miR-92b-3p** | 1.31630528 | 10.16008 | 4.62E-05 | 0.01323122 |
| **hsa-miR-486-5p\|MIMAT0002177\|Homo-sapiens\|miR-486-5p** | 1.45413061 | 19.2362797 | 5.61E-05 | 0.01323122 |
| **hsa-miR-199a-3p\|MIMAT0000232\|Homo-sapiens\|miR-199a-3p** | -0.9013621 | 10.07888492 | 5.88E-05 | 0.01323122 |
| **hsa-miR-199b-3p\|MIMAT0004563\|Homo-sapiens\|miR-199b-3p** | -0.9013666 | 10.07888492 | 5.92E-05 | 0.01323122 |
| **hsa-miR-451a\|MIMAT0001631\|Homo-sapiens\|miR-451a** | 1.78885692 | 14.6913498 | 5.98E-05 | 0.01323122 |
| **hsa-miR-122-5p\|MIMAT0000421\|Homo-sapiens\|miR-122-5p** | 4.7678994 | 8.4549561 | 0.00010603 | 0.0194186 |
| **hsa-miR-374a-3p\|MIMAT0004688\|Homo-sapiens\|miR-374a-3p** | -1.4930605 | 5.16814953 | 0.00010637 | 0.0194186 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-615-3p\|MIMAT0003283\|Homo-sapiens\|miR-615-3p** | 4.61659 34 | 2.34075 32 | 0.00011 181 | 0.01941 86 |
| **hsa-miR-200a-3p\|MIMAT0000682\|Homo-sapiens\|miR-200a-3p** | 2.62810 591 | 3.55457 593 | 0.00012 349 | 0.01941 86 |
| **hsa-miR-205-5p\|MIMAT0000266\|Homo-sapiens\|miR-205-5p** | 2.61173 011 | 4.76221 273 | 0.00012 714 | 0.01941 86 |
| **hsa-miR-99a-5p\|MIMAT0000097\|Homo-sapiens\|miR-99a-5p** | 2.71742 018 | 7.00052 589 | 0.00013 16 | 0.01941 86 |
| **hsa-miR-29c-3p\|MIMAT0000681\|Homo-sapiens\|miR-29c-3p** | 1.18969 464 | 6.51419 957 | 0.00017 334 | 0.02423 157 |
| **hsa-miR-500a-3p\|MIMAT0002871\|Homo-sapiens\|miR-500a-3p** | 1.00570 193 | 7.36354 916 | 0.00018 412 | 0.02445 079 |
| **hsa-miR-501-3p\|MIMAT0004774\|Homo-sapiens\|miR-501-3p** | 1.70245 486 | 7.85257 347 | 0.00020 344 | 0.02573 03 |
| **hsa-let-7g-5p\|MIMAT0000414\|Homo-sapiens\|let-7g-5p** | - 1.08859 06 | 10.8157 853 | 0.00024 959 | 0.02835 433 |
| **hsa-miR-130b-5p\|MIMAT0004680\|Homo-sapiens\|miR-130b-5p** | - 1.35970 34 | 5.89096 342 | 0.00027 675 | 0.02835 433 |
| **hsa-miR-98-5p\|MIMAT0000096\|Homo-sapiens\|miR-98-5p** | - 1.26261 2 | 8.48047 385 | 0.00028 402 | 0.02835 433 |
| **hsa-let-7i-5p\|MIMAT0000415\|Homo-sapiens\|let-7i-5p** | - 0.63381 08 | 12.5386 29 | 0.00028 717 | 0.02835 433 |
| **hsa-miR-99b-5p\|MIMAT0000689\|Homo-sapiens\|miR-99b-5p** | 1.92805 305 | 10.4298 993 | 0.00028 779 | 0.02835 433 |
| **hsa-miR-181c-3p\|MIMAT0004559\|Homo-sapiens\|miR-181c-3p** | - 0.98758 27 | 6.18677 259 | 0.00028 824 | 0.02835 433 |
| **hsa-miR-193b-3p\|MIMAT0002819\|Homo-sapiens\|miR-193b-3p** | 3.46752 612 | 2.44700 732 | 0.00031 271 | 0.02881 306 |
| **hsa-miR-185-5p\|MIMAT0000455\|Homo-sapiens\|miR-185-5p** | 2.06898 382 | 5.86443 828 | 0.00031 785 | 0.02881 306 |
| **hsa-miR-125b-5p\|MIMAT0000423\|Homo-sapiens\|miR-125b-5p** | 1.97740 042 | 6.51511 556 | 0.00032 545 | 0.02881 306 |
| **hsa-miR-484\|MIMAT0002174\|Homo-sapiens\|miR-484** | 1.11779 909 | 10.5331 889 | 0.00035 338 | 0.02982 822 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-340-5p\|MIMAT0004692\|Homo-sapiens\|miR-340-5p** | -1.1723659 | 8.90892494 | 0.00035938 | 0.02982822 |
| **hsa-miR-26b-5p\|MIMAT0000083\|Homo-sapiens\|miR-26b-5p** | -1.0738431 | 10.5233865 | 0.00042181 | 0.03394956 |
| **hsa-miR-222-3p\|MIMAT0000279\|Homo-sapiens\|miR-222-3p** | -1.0646154 | 9.18018019 | 0.00054252 | 0.04238074 |
| **hsa-miR-210-3p\|MIMAT0000267\|Homo-sapiens\|miR-210-3p** | 1.59960404 | 7.03925914 | 0.00058 15 | 0.04412731 |
| **hsa-miR-193b-5p\|MIMAT0004767\|Homo-sapiens\|miR-193b-5p** | 3.04669826 | 2.86220143 | 0.00070077 | 0.05073365 |
| **hsa-miR-1285-3p\|MIMAT0005876\|Homo-sapiens\|miR-1285-3p** | 1.94805 | 5.67398439 | 0.00072655 | 0.05073365 |
| **hsa-miR-652-3p\|MIMAT0003322\|Homo-sapiens\|miR-652-3p** | 1.1909739 | 8.09834632 | 0.00072956 | 0.05073365 |
| **hsa-miR-183-5p\|MIMAT0000261\|Homo-sapiens\|miR-183-5p** | 1.18972843 | 7.18267275 | 0.00074496 | 0.05073365 |
| **hsa-miR-192-5p\|MIMAT0000222\|Homo-sapiens\|miR-192-5p** | 1.62108219 | 12.6900643 | 0.00080 96 | 0.05375738 |
| **hsa-miR-96-5p\|MIMAT0000095\|Homo-sapiens\|miR-96-5p** | 1.73148192 | 5.72117116 | 0.00086 29 | 0.05589877 |
| **hsa-miR-30e-3p\|MIMAT0000693\|Homo-sapiens\|miR-30e-3p** | -0.758716 | 7.73594479 | 0.00100018 | 0.06324961 |
| **hsa-miR-5010-5p\|MIMAT0021043\|Homo-sapiens\|miR-5010-5p** | 1.61689445 | 4.23478 21 | 0.00103837 | 0.06413771 |
| **hsa-miR-363-3p\|MIMAT0000707\|Homo-sapiens\|miR-363-3p** | 1.36366904 | 10.380701 | 0.00115 84 | 0.06992 54 |
| **hsa-miR-184\|MIMAT0000454\|Homo-sapiens\|miR-184** | 3.54842139 | 3.37798717 | 0.00130957 | 0.07729356 |
| **hsa-miR-369-3p\|MIMAT0000721\|Homo-sapiens\|miR-369-3p** | -2.9928015 | 2.77234965 | 0.00147028 | 0.08489289 |
| **hsa-miR-197-5p\|MIMAT0022691\|Homo-sapiens\|miR-197-5p** | 2.58413805 | 2.35887302 | 0.00196662 | 0.11113489 |
| **hsa-miR-877-5p\|MIMAT0004949\|Homo-sapiens\|miR-877-5p** | 1.33230417 | 5.70308199 | 0.00208915 | 0.11559988 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-874-3p\|MIMAT0004911\|Homo-sapiens\|miR-874-3p** | 1.45969876 | 3.60156383 | 0.00262707 | 0.14065545 |
| **hsa-miR-188-5p\|MIMAT0000457\|Homo-sapiens\|miR-188-5p** | 2.0934215 | 2.45947977 | 0.00264788 | 0.14065545 |
| **hsa-miR-4732-5p\|MIMAT0019855\|Homo-sapiens\|miR-4732-5p** | 1.93096614 | 3.51796142 | 0.00328898 | 0.17128478 |
| **hsa-let-7d-5p\|MIMAT0000065\|Homo-sapiens\|let-7d-5p** | -0.8597629 | 9.40609233 | 0.0035494 | 0.1811598 |
| **hsa-miR-16-5p\|MIMAT0000069\|Homo-sapiens\|miR-16-5p** | 1.03668035 | 14.9045746 | 0.00363063 | 0.1811598 |
| **hsa-miR-3605-5p\|MIMAT0017981\|Homo-sapiens\|miR-3605-5p** | 1.53569505 | 4.19679618 | 0.00368386 | 0.1811598 |
| **hsa-let-7e-5p\|MIMAT0000066\|Homo-sapiens\|let-7e-5p** | -1.0352015 | 7.94613566 | 0.00375143 | 0.1811598 |
| **hsa-miR-125b-2-3p\|MIMAT0004603\|Homo-sapiens\|miR-125b-2-3p** | 2.06822474 | 5.28208487 | 0.00425152 | 0.20164359 |
| **hsa-miR-186-5p\|MIMAT0000456\|Homo-sapiens\|miR-186-5p** | 0.63828024 | 11.9812271 | 0.00520734 | 0.2426439 |
| **hsa-miR-146b-5p\|MIMAT0002809\|Homo-sapiens\|miR-146b-5p** | 1.37089106 | 10.1287112 | 0.00535727 | 0.24531186 |
| **hsa-miR-1273h-3p\|MIMAT0030416\|Homo-sapiens\|miR-1273h-3p** | -0.9357433 | 5.35685214 | 0.00544932 | 0.24531186 |
| **hsa-miR-125a-5p\|MIMAT0000443\|Homo-sapiens\|miR-125a-5p** | 1.16047599 | 9.21321215 | 0.00613776 | 0.27169814 |
| **hsa-miR-30a-5p\|MIMAT0000087\|Homo-sapiens\|miR-30a-5p** | 1.18969516 | 10.9437677 | 0.00664258 | 0.28922443 |
| **hsa-miR-375-3p\|MIMAT0000728\|Homo-sapiens\|miR-375-3p** | 1.92514223 | 9.03593736 | 0.00688856 | 0.29509711 |
| **hsa-miR-4508\|MIMAT0019045\|Homo-sapiens\|miR-4508** | 2.16670074 | 3.53986509 | 0.00718234 | 0.30242367 |
| **hsa-miR-28-5p\|MIMAT0000085\|Homo-sapiens\|miR-28-5p** | -0.8884773 | 6.96288237 | 0.00728732 | 0.30242367 |
| **hsa-miR-885-5p\|MIMAT0004947\|Homo-sapiens\|miR-885-5p** | 2.68236447 | 3.3794144 | 0.00758102 | 0.30938956 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-1246\|MIMAT0005898\|Homo-sapiens\|miR-1246** | 2.28800144 | 2.44718051 | 0.00768814 | 0.30938956 |
| **hsa-miR-34a-5p\|MIMAT0000255\|Homo-sapiens\|miR-34a-5p** | 2.05753178 | 2.83001531 | 0.00799624 | 0.31698508 |
| **hsa-miR-429\|MIMAT0001536\|Homo-sapiens\|miR-429** | 2.00344398 | 2.9346468 | 0.00835424 | 0.32630671 |
| **hsa-miR-148b-3p\|MIMAT0000759\|Homo-sapiens\|miR-148b-3p** | -0.9328425 | 9.6472408 | 0.00861407 | 0.33157939 |
| **hsa-miR-134-5p\|MIMAT0000447\|Homo-sapiens\|miR-134-5p** | -1.5143719 | 5.06393719 | 0.00909152 | 0.34430327 |
| **hsa-miR-11400\|MIMAT0044657\|Homo-sapiens\|miR-11400** | -0.8156523 | 6.15939453 | 0.00920389 | 0.34430327 |
| **hsa-miR-4433b-5p\|MIMAT0030413\|Homo-sapiens\|miR-4433b-5p** | -1.0218966 | 6.90322098 | 0.0103357 | 0.3777029 |
| **hsa-miR-301a-3p\|MIMAT0000688\|Homo-sapiens\|miR-301a-3p** | -0.774071 | 7.16120664 | 0.01038114 | 0.3777029 |
| **hsa-miR-6815-5p\|MIMAT0027530\|Homo-sapiens\|miR-6815-5p** | 2.65501673 | 1.90016383 | 0.01122325 | 0.4028237 |
| **hsa-miR-592\|MIMAT0003260\|Homo-sapiens\|miR-592** | 3.27586604 | 1.87917328 | 0.01179449 | 0.4146474 |
| **hsa-miR-6741-3p\|MIMAT0027384\|Homo-sapiens\|miR-6741-3p** | -1.6811885 | 3.26647439 | 0.01186491 | 0.4146474 |
| **hsa-miR-149-5p\|MIMAT0000450\|Homo-sapiens\|miR-149-5p** | 3.02757919 | 1.81195586 | 0.01268132 | 0.43742327 |
| **hsa-miR-196b-5p\|MIMAT0001080\|Homo-sapiens\|miR-196b-5p** | -1.2247946 | 4.12521213 | 0.01303713 | 0.44393113 |
| **hsa-miR-1468-5p\|MIMAT0006789\|Homo-sapiens\|miR-1468-5p** | 1.11185528 | 5.16202964 | 0.01373821 | 0.45015679 |
| **hsa-miR-223-5p\|MIMAT0004570\|Homo-sapiens\|miR-223-5p** | -0.8992462 | 6.32867255 | 0.01384523 | 0.45015679 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-589-3p\|MIMAT0003256\|Homo-sapiens\|miR-589-3p** | 1.99371854 | 2.89339304 | 0.01389422 | 0.45015679 |
| **hsa-miR-483-5p\|MIMAT0004761\|Homo-sapiens\|miR-483-5p** | 1.99778058 | 2.50327559 | 0.01392372 | 0.45015679 |
| **hsa-miR-483-3p\|MIMAT0002173\|Homo-sapiens\|miR-483-3p** | 2.23921097 | 2.44608964 | 0.01406 74 | 0.45015679 |
| **hsa-miR-181c-5p\|MIMAT0000258\|Homo-sapiens\|miR-181c-5p** | -0.5630765 | 8.07509427 | 0.01572607 | 0.49485498 |
| **hsa-miR-320c\|MIMAT0005793\|Homo-sapiens\|miR-320c** | 1.16697182 | 4.24369154 | 0.01583685 | 0.49485498 |
| **hsa-miR-342-3p\|MIMAT0000753\|Homo-sapiens\|miR-342-3p** | 0.93184713 | 6.95075553 | 0.01678 56 | 0.51840179 |
| **hsa-miR-181a-5p\|MIMAT0000256\|Homo-sapiens\|miR-181a-5p** | -0.5478649 | 13.6341652 | 0.01777307 | 0.54258922 |
| **hsa-miR-130b-3p\|MIMAT0000691\|Homo-sapiens\|miR-130b-3p** | 0.57710903 | 8.39264 19 | 0.01848968 | 0.55805224 |
| **hsa-miR-331-5p\|MIMAT0004700\|Homo-sapiens\|miR-331-5p** | 1.11740 64 | 3.05659971 | 0.01909475 | 0.56983881 |
| **hsa-miR-7-1-3p\|MIMAT0004553\|Homo-sapiens\|miR-7-1-3p** | 1.20241 58 | 3.14513 59 | 0.01954721 | 0.57685985 |
| **hsa-miR-126-3p\|MIMAT0000445\|Homo-sapiens\|miR-126-3p** | -0.5128558 | 9.66834117 | 0.01990822 | 0.58105755 |
| **hsa-miR-16-2-3p\|MIMAT0004518\|Homo-sapiens\|miR-16-2-3p** | 1.01698342 | 9.35770309 | 0.02084597 | 0.60181396 |
| **hsa-miR-7706\|MIMAT0030021\|Homo-sapiens\|miR-7706** | 0.9093351 | 4.97902636 | 0.0212807 | 0.60775843 |
| **hsa-miR-502-3p\|MIMAT0004775\|Homo-sapiens\|miR-502-3p** | 0.67217595 | 5.14618893 | 0.02207489 | 0.623733 |
| **hsa-miR-6852-5p\|MIMAT0027604\|Homo-sapiens\|miR-6852-5p** | -0.8156949 | 5.97417949 | 0.02280016 | 0.63744459 |
| **hsa-miR-6735-5p\|MIMAT0027371\|Homo-sapiens\|miR-6735-5p** | 1.94235341 | 2.02612379 | 0.02373729 | 0.65673173 |
| **hsa-miR-26a-5p\|MIMAT0000082\|Homo-sapiens\|miR-26a-5p** | -0.6991324 | 14.0471851 | 0.02425534 | 0.66414621 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-1290\|MIMAT0005880\|Homo-sapiens\|miR-1290** | 3.41137803 | 1.81055491 | 0.02537323 | 0.68766633 |
| **hsa-miR-15a-5p\|MIMAT0000068\|Homo-sapiens\|miR-15a-5p** | 0.90046448 | 11.6134744 | 0.02589132 | 0.69461952 |
| **hsa-miR-126-5p\|MIMAT0000444\|Homo-sapiens\|miR-126-5p** | -0.5507693 | 13.493865 | 0.02686679 | 0.71358191 |
| **hsa-miR-107\|MIMAT0000104\|Homo-sapiens\|miR-107** | 0.8005586 | 12.8085356 | 0.02736774 | 0.71969026 |
| **hsa-miR-769-5p\|MIMAT0003886\|Homo-sapiens\|miR-769-5p** | -0.5091718 | 7.36521816 | 0.0282608 | 0.7305509 |
| **hsa-miR-1294\|MIMAT0005884\|Homo-sapiens\|miR-1294** | 1.45471042 | 3.06943214 | 0.02833085 | 0.7305509 |
| **hsa-miR-142-3p\|MIMAT0000434\|Homo-sapiens\|miR-142-3p** | -0.5359993 | 8.09298232 | 0.03084653 | 0.78367114 |
| **hsa-miR-449c-5p\|MIMAT0010251\|Homo-sapiens\|miR-449c-5p** | 2.2752872 | 2.10756338 | 0.03098098 | 0.78367114 |
| **hsa-miR-1260a\|MIMAT0005911\|Homo-sapiens\|miR-1260a** | -1.1643864 | 4.17451137 | 0.03134152 | 0.78531198 |
| **hsa-miR-2355-3p\|MIMAT0017950\|Homo-sapiens\|miR-2355-3p** | -2.1181094 | 2.35386915 | 0.03212329 | 0.79737819 |
| **hsa-miR-4446-3p\|MIMAT0018965\|Homo-sapiens\|miR-4446-3p** | -0.8795219 | 5.495024 | 0.03310787 | 0.81220527 |
| **hsa-miR-181a-3p\|MIMAT0000270\|Homo-sapiens\|miR-181a-3p** | -0.7516211 | 5.73535605 | 0.03351633 | 0.81220527 |
| **hsa-miR-6740-5p\|MIMAT0027381\|Homo-sapiens\|miR-6740-5p** | 1.96946827 | 1.953817 | 0.03363802 | 0.81220527 |
| **hsa-miR-4687-5p\|MIMAT0019774\|Homo-sapiens\|miR-4687-5p** | 1.84083031 | 2.27575094 | 0.03413055 | 0.81667337 |
| **hsa-miR-221-3p\|MIMAT0000278\|Homo-sapiens\|miR-221-3p** | -0.5879869 | 9.51222115 | 0.03998644 | 0.94824982 |

| | | | | |
|---|---|---|---|---|
| **hsa-miR-26b-3p\|MIMAT0004500\|Homo-sapiens\|miR-26b-3p** | 0.56654159 | 6.26419512 | 0.04121553 | 0.96874739 |
| **hsa-miR-10b-5p\|MIMAT0000254\|Homo-sapiens\|miR-10b-5p** | 1.03748105 | 14.0358383 | 0.04348054 | 1 |
| **hsa-miR-5001-3p\|MIMAT0021022\|Homo-sapiens\|miR-5001-3p** | 2.13677035 | 2.29591704 | 0.04396265 | 1 |
| **hsa-miR-1277-5p\|MIMAT0022724\|Homo-sapiens\|miR-1277-5p** | -1.5128617 | 3.23364891 | 0.04409414 | 1 |
| **hsa-miR-328-3p\|MIMAT0000752\|Homo-sapiens\|miR-328-3p** | -0.5968961 | 6.34666503 | 0.04434523 | 1 |
| **hsa-miR-92a-3p\|MIMAT0000092\|Homo-sapiens\|miR-92a-3p** | 0.56275338 | 15.7899188 | 0.0461056 | 1 |
| **hsa-miR-425-5p\|MIMAT0003393\|Homo-sapiens\|miR-425-5p** | 0.49039083 | 9.75805703 | 0.04637307 | 1 |
| **hsa-miR-4449\|MIMAT0018968\|Homo-sapiens\|miR-4449** | -1.8428257 | 2.30420875 | 0.04804692 | 1 |
| **hsa-miR-550a-3p\|MIMAT0003257\|Homo-sapiens\|miR-550a-3p** | 0.77990837 | 5.20895335 | 0.04811027 | 1 |
| **hsa-miR-301b-3p\|MIMAT0004958\|Homo-sapiens\|miR-301b-3p** | -0.7193683 | 4.68939013 | 0.04865736 | 1 |
| **hsa-miR-181d-5p\|MIMAT0002821\|Homo-sapiens\|miR-181d-5p** | -0.6778672 | 5.37263153 | 0.04997824 | 1 |

Supplementary Table VII: All genes regulated by our microRNAs of interest and the number of microRNAs regulating them.

https://docs.google.com/spreadsheets/d/19oygiy9BKyqz3YdVvq5eZ-twGp9adIkJ/edit?usp=sharing&ouid=111817647840889431552&rtpof=true&sd=true