U.PORTO

MESTRADO EM ONCOLOGIA

ONCOLOGIA LABORATORIAL

# CHARACTERIZATION OF HOXB7 AND HOXB9 EXPRESSION IN BREAST CANCER PATIENTS

JOÃO ANDRADE MOUTA

M

2019

U.PORTO

João Andrade Mouta. CHARACTERIZATION OF HOXB7 AND HOXB9 EXPRESSION IN BREAST CANCER PATIENTS

M.ICBAS 2019

CHARACTERIZATION OF HOXB7 AND HOXB9 EXPRESSION IN BREAST CANCER PATIENTS

JOÃO ANDRADE MOUTA

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

João Andrade Mouta

# Characterization of HOXB7 and HOXB9 expression in breast cancer patients

Dissertação de Candidatura ao grau de Mestre em Oncologia submetida ao Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto

**Orientadora** Doutora Carla Renata Gonçalves de Freitas
**Categoria** Investigadora Auxiliar e Professora Auxiliar Convidada.
**Afiliação** Grupo “Cell Growth and Differentiation”
Instituto de Biologia Molecular e Celular
Instituto de Investigação e Inovação em Saúde- Universidade do Porto.
Instituto de Ciências Biomédicas Abel Salazar - Universidade do Porto

**Co-orientador** Doutora Carmen de Lurdes Fonseca Jerónimo
**Categoria** Investigadora Auxiliar e Professora Catedrática convidada
**Afiliação** Grupo de Epigenética e Biologia do Cancro
CI-IPO-Instituto Português de Oncologia
Instituto de Ciências Biomédicas Abel Salazar - Universidade do Porto

# Table of Contents

# Acknowledgements

The journey to accomplish this master's thesis was not easy, and every day I have learnt something new and grown at professional level as well as personal, however this was only possible with the contribution of several people.

First, I would like to thank to Professor Carmen Jerónimo of the Master Course in Oncology for accepting my application in the course and for collaborating in this project as co-supervisor, encouraging me to always give my best.

To Dr. Renata Freitas, for welcoming me into the project and for always having the patience to guide me. You have given me strength and motivation throughout this course, as well as numerous tools to achieve my objective and to be successful in my professional career. For that, I will always be grateful.

To the Cancer Biology and Epigenetics group at IPO-Porto, for receiving me as one of your own and for helping me whenever I needed, without exception. To Sofia Salta and Claúdia Lima for introducing me to immunohistochemistry, both of your help was fundamental for this work.

I would also like to show my gratitude to all Cell Growth and Differentiation group, for always helping me see the bright side. To Simone Bessa, for all your calmness and consideration and for helping me in so many different ways. To Mafalda Pereira, one of my dearest friends, for keeping me on tracks and for always make me laugh during the hardest times. I sincerely hope that you achieve everything that you want.

To my family, you were the ones who make this achievement possible, without your support any of this would be possible. Thank you for everything, and I am sorry for all the mood swings and no responses during this time. To Luís, Diogo and Marta, thank you all for your comprehension, making sure that I was doing fine and for never letting me walk alone, for that I am lucky to call you my friends.

Last but not least, I would like to thank my girlfriend for never letting me quit, for your motivation lectures and for every kind gesture that you made during this time, without your support this would not be the same. For your encouragement, toleration and inspiration I will always be grateful.

This master's thesis does not only belong to me, but for everyone who helped me in every single way.

# Resumo

Os genes *HOX* codificam fatores de transcrição cruciais para o desenvolvimento embrionário, que afetam a proliferação celular, diferenciação, migração e apoptose. No entanto alguns destes genes são também expressos no tecido adulto e frequentemente desregulados em processos cancerígenos, como é o caso do cancro de mama, a principal causa de morte nas mulheres, devido a doença oncológica. Resultados preliminares do nosso laboratório, demonstraram que o gene *HOXB7* é sobre-expresso em linhas celulares de cancro da mama representativas dos distintos subtipos molecular usados na clínica. No entanto estes resultados não foram replicáveis num grupo de pacientes com cancro de mama do IPO-Porto.

O principal objetivo desta dissertação foi aprofundar a caracterização da expressão dos genes *HOXB7* e *HOXB9* nestes pacientes. Com esta finalidade, realizaram-se análises de qPCR para avaliar os níveis de mRNA do *HOXB9* em biopsias de tecido normal e cancerígeno. De seguida, foi otimizado o protocolo para analisar a expressão proteica do HOXB9 por *western blot*. O nível proteico do HOXB7 e HOXB9 em pacientes, foi explorado também por imunohistoquímica, o que envolveu igualmente otimização de protocolos.

A analise quantitativa dos níveis de mRNA e de proteína revelaram níveis de expressão mais altos nos tecidos normais do que nos tecidos cancerígenos, o que contraria os resultados preliminares realizados em linhas celulares no nosso laboratório e também outros estudos publicados.

Este trabalho permitiu a otimização de protocolos visando a caracterização da expressão do HOXB7 e HOXB9 em cancros de mama, tendo em conta distintos subtipos moleculares, e enalteceu a necessidade de aumentar o número de casos estudados e de incorporar controlos visando uma menor heterogeneidade genética. Adicionalmente, os dados obtidos por imunohistoquímica permitiram uma caracterização das células e das estruturas mamárias que demonstram aberrante expressão destes genes durante os desenvolvimentos do cancro da mama, tendo em conta os diferentes subtipos moleculares, o que é relevante para a consolidação do estudo da sua função nestes contextos oncológicos.

# Abstract

HOX genes encode transcription factors crucial for embryonic development that affect cell proliferation, differentiation, migration, and apoptosis. However, some of these genes are also expressed in adult tissue and are often deregulated in cancerous processes, such as breast cancer, the leading cause of death in women due to cancer. Preliminary results from our laboratory demonstrated that the *HOXB7* gene is overexpressed in breast cancer cell lines representing the distinct molecular subtypes used in the clinic. However, these results were not replicable in a group of IPO-Porto breast cancer patients.

The main objective of this dissertation was to further characterize the expression of *HOXB7* and also *HOXB9* in these patients. To this end, qPCR analyzes were performed to evaluate *HOXB9* mRNA levels in normal and carcinogenic tissues. Then, the protocol to analyze HOXB9 protein levels by western blot was optimized. The protein level of HOXB7 and HOXB9 in patients was also explored by immunohistochemistry, which also involved protocol optimization.

Quantitative analysis of mRNA and protein levels revealed higher expression levels in normal tissues than in cancerous tissues, which contradicts preliminary results performed on cell lines in our laboratory and also other published studies.

This work allowed the optimization of protocols aiming to characterize HOXB7 and HOXB9 expression in breast cancers, taking into account different molecular subtypes, and emphasized the need to increase the number of cases studied and to incorporate controls aiming at less genetic heterogeneity. Additionally, data obtained by immunohistochemistry will allow characterization of breast cells and structures showing aberrant expression of these genes during the development of breast cancer, considering the distinct molecular subtypes, which is relevant to consolidate study addressing their function in these oncological contexts.

# List of Abbreviations

ATM – Ataxia telangiectasia mutated

AUC – Area under the curve

bFGF – Basic fibroblast growth factor

BrCa – Breast cancer

*CDH1* – E-cadherin gene

ChIP – Chromatin immunoprecipitation

COMMD7 – Copper Metabolism MURR1 Domain-containing 7

CRC – Colo rectal cancer

DCIS – Ductal carcinoma *in situ*

DNA – Deoxyribonucleic acid

DNTM3B – DNA methyltransferase 3 beta

EMT – Epithelial-mesenchymal transition

ER – Estrogen receptor

FW - Forward

GAPDH – Glyceraldehyde-3-phosphate dehydrogenate

GC – Gastric cancer

GF – Growth factors

GH – Growth hormone

HER2 – Human epidermal growth factor 2 receptor

HMW – High molecular weight

HOX – Homeobox

HOXB7 – Homeobox B7

HOXB8 – Homeobox B8

HOXB9 – Homeobox B9

HRT – Hormonal replacement therapy

IDC – Invasive ductal carcinoma

IGF1 – Insulin-like growth factor 1

IHC – Immunohistochemistry

ILC – Invasive lobular carcinoma

IMD – Intratumoral microvessel density

LCIS – Lobular carcinoma *in situ*

MRI – Magnetic resonance imaging

NBr – Normal breast

NRG2 – Neuregulin 2

PCR – Polymerase Chain Reaction

PgR – Progesterone receptor

qPCR – Quantitative real-time PCR

RNA – Ribonucleic acid

ROC – Receiver Operating Characteristic

RV – Reverse

SNPs – Single nucleotide polymorphisms

TDLU – Terminal duct lobular unit

TGF-β – Transforming growth factor beta

TNBC – Triple negative breast cancer

TP53 – Tumour protein 53

VEGF – Vascular endothelial growth factor

# Figures Index

# Tables Index

# 1. Introduction

## 1.1Breast Cancer

### 1.1.1 Epidemiology

Oncological diseases are extremely common in Europe, affecting 1 in 4 people throughout their lifetime. In addition, these diseases figure among the most common causes of death in Europe and, as for other main causes of death, affect most significantly the men population (Fig.1) [1]. The only exception, in which a main cause of death affects more women than men, is breast cancer (BrCa). After lung cancer, which has the highest number of new cases and deaths per year, BrCa is the second most common cancer type, affecting women in the world [2, 3]. In 2018, 2 million new cases of this oncological disease have been identified worldwide and it was also the most incident type of cancer among women, and also the deadliest [4, 5]. However, men also develop BrCa but in a much lower percentage, which represents about 1% of the diagnosed cases [6].

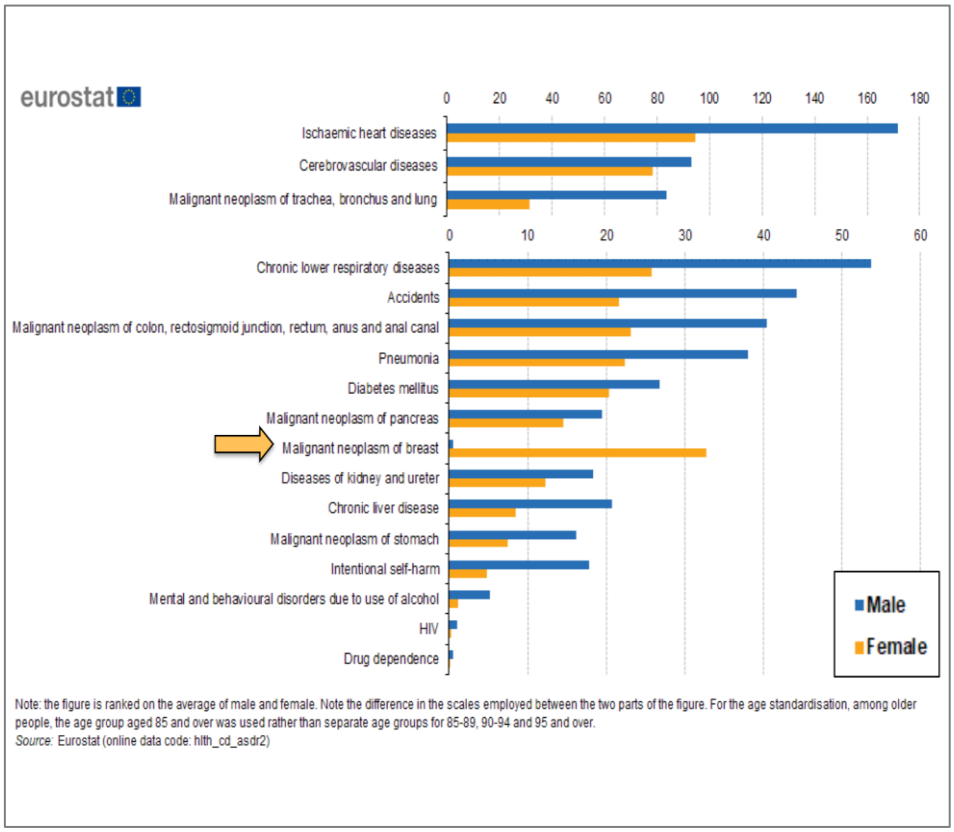


Figure 1 - Main causes of death in Europe (Last Eurostat data from 2015, EU), standardized rates per 100 000 inhabitants. Note higher mortality for the women population, in comparison with the men population, detected only for malignant neoplasm of breast (arrow).

Within the first 25 countries with the highest rate of BrCa, 19 are from Europe and 24 are from countries classified as having “upper middle” to “high” per capita gross national income

(Tab.1). This has been related with risk factors more prevalent in these countries such as aging of the population, low parity, increasing age of the first pregnancy, use of birth control and post-menopausal medication, obesity, physical inactivity, among others [7]. In countries with high income, BrCa incidence and mortality are expected to decrease, due to advances in early diagnosis and therapy, while in underdeveloped countries the opposite tendency is expected [8]. Recent studies suggest that the increasing incidence of BrCa in Northern and Eastern Europe is associated to mammography screening and reduction of post-menopausal hormonal replacement therapy [2]. However, these countries have available the best early detection programs, neo-adjuvant therapies and health care in general, which is contributing to decrease BrCa mortality. In contrast, in underdeveloped countries the number of BrCa cases is expected to increase, associated to risk factors such as obesity, physical inactivity, late first-pregnancy or smoking but also due to lack of routine screenings and poor health care [2]. In these countries, the number of new BrCa cases is expected to ascends to 2.4 million per year worldwide in 2030 [9].

Table 1 - Top 25 countries with highest rates of breast cancer worldwide. Three first columns show data from Continuous Update Project for 2018 [5] and last column indicates the economic status of each country/region, according to the report on World Economic Situation and prospect in 2014 (United Nations) [10].

| Rank | Country | Age-standardised rate/100,000 | Economic status |
|---|---|---|---|
| 1 | Belgium | 113.2 | High Income |
| 2 | Luxembourg | 109.3 | High Income |
| 3 | Netherlands | 105.9 | High Income |
| 4 | France (metropolitan) | 99.1 | High Income |
| 5 | New Caledonia (France) | 98.0 | High Income |
| 6 | Lebanon | 97.6 | Upper middle income |
| 7 | Australia | 94.5 | High Income |
| 8 | UK | 93.6 | High Income |
| 9 | Italy | 92.8 | High Income |
| 10 | New Zealand | 92.6 | High Income |
| 11 | Ireland | 90.3 | High Income |
| 12 | Sweden | 89.8 | High Income |
| 13 | Finland | 89.5 | High Income |
| 14 | Denmark | 88.8 | High Income |
| 15 | Switzerland | 88.1 | High Income |
| 16 | Montenegro | 87.8 | Upper middle income |
| 17 | Malta | 87.6 | High Income |
| 18 | Norway | 87.5 | High Income |
| 19 | Hungary | 85.5 | Upper middle income |
| 20 | Germany | 85.4 | High Income |
| 21 | Iceland | 85.2 | High Income |
| 22 | US | 84.9 | High Income |
| 23 | Canada | 83.8 | High Income |
| 24 | Cyprus | 81.7 | High Income |
| 25 | Samoa | 80.1 | Underdeveloped |

In Portugal, BrCa incidence rates have been rising for years (Fig.2) [11, 12]. Also in Portugal, this oncological disease is the most common cancer type and the principal cause of cancer mortality in women, accounting for 30% of all cancer cases and 16% of all cancer

deaths in 2012 [7]. Two recent studies addressed the incidence in distinct country regions and found a higher number of cases for most age groups in Lisbon area [11, 12]. In general, the South presents the highest age-standardized rate (155.8/100 000 inhabitants), while the North presents the quickest rate of increase (3.6%/year). Therefore, these studies suggest that the Northern region could be number one among all regions in a few years. For now, younger women from the North have lower risk to develop BrCa in comparison to women with the same age gap from South and Centre. However, this risk is reversed in older women. The authors of these studies suggest that differences found in distinct regions of the country might be due to a combination of different diagnostic practices and/or exposure to different risk factors. The heterogeneity of the disease in younger and older women may also explain partially the differences in age-specific rates [12]. The number of diagnosed cases in Azores and Madeira was much lower than in the continental region, representing only 4,2% of the cases occurring in the time period from 1998 through 2011 [11].

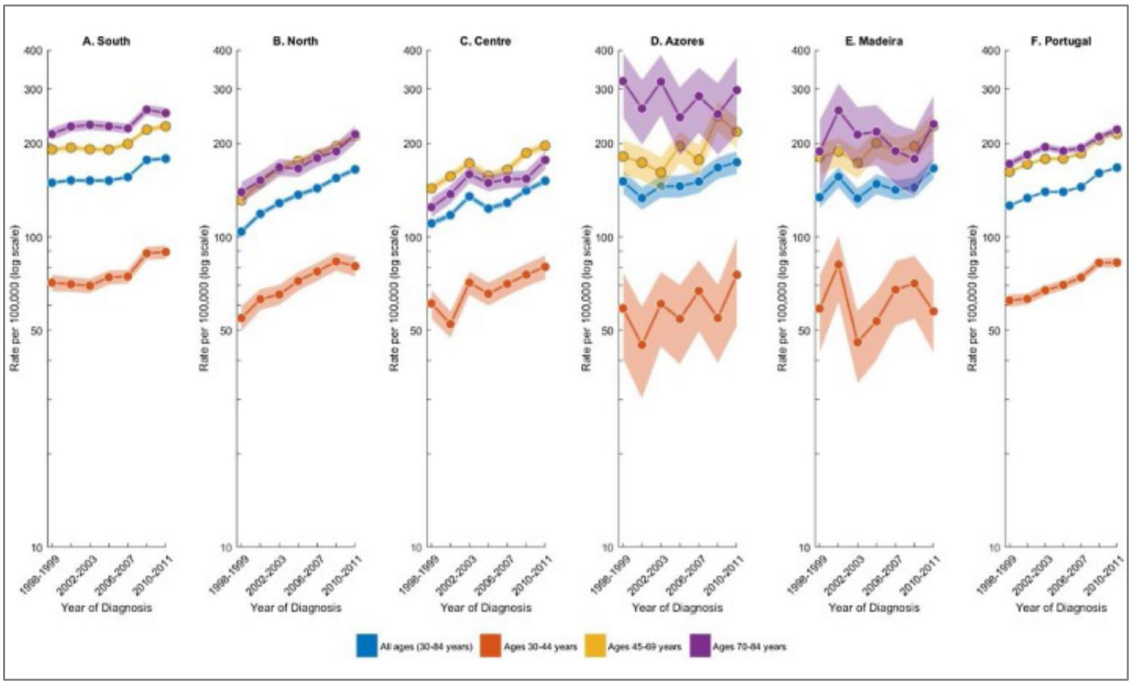


Figure 2 -BrCa incidence in Portugal 1998–2011 (overall and by region) [11]. Age-standardized rates from women aged 30-84 years and in specific age-groups (30–44, 45–69, 70–84 years).

### 1.1.2 Risk Factors

The identification of risk factors associated to the development of BrCa is a critical step for the prevention of this disease. There are 5 major categories of risk factors that have been associated to BrCa 1) aging, 2) family history, 3) genetics, 4) estrogen levels and 5) life habits.

1) *Aging*: It is the major risk factor for BrCa, which has an increasing incidence throughout a woman´s lifetime [13]. For example, 2 in 3 women with an Invasive BrCa are over 55 years old and post-menopausal. This is explained by accumulation of genetic damage (mutations) throughout life that eventually promotes oncological processes. However, it is also true that BrCa that occurs in younger ages, prior to 40 years old, are more difficult to treat, have higher risk of relapse and cause more death [14]. The first screening age, ideal according to several studies, is between 40-45 years, in order to detect BrCa with worse prognosis [14]. However, when women have family history of BrCa, they are advised to start the screening much earlier, at the age of 25. Thus, the genetic background alters the age of BrCa risk. An intriguing example was found in African-American women, which seems more prone to develop aggressive BrCa types before the age of 40, contrarily to Caucasian-American women [13].

2) *Family history:* It is strongly relevant for the risk to develop BrCa. As above mentioned, certain types of BrCa (5% to 10%) are related with hereditary genetic background and those are the ones affecting younger women and responsible for higher mortality rate. Thus, the risk for BrCa directly relates with the number of first-degree relatives with this disease. Moreover, even in older women the genetic background is relevant for the outcome of the disease. Thus, post-menopausal women with familiar history of BrCa have much higher risk to develop this disease. In certain cases, a strong family history of BrCa is linked to mutated genes, such as *BRCA1*, *BRCA2* gene or *CHEK2* gene, which play a central role in the development of different types of BrCa [15].

3) *Genetics*: DNA damage causes abnormal cell growth and function and is strongly associated to oncogenesis. It can be inherited or appear *de novo* over the course of life, resulting from natural aging or exposure to chemicals in the environment. Most inherited cases of BrCa are associated to mutations in *BRCA1* and *BRCA2* genes, involved in cell damage repair [16]. Women with *BRCA1* or *BRCA2* mutations (or both) have up to a 72% risk of being diagnosed with BrCa, which occurs particularly in younger women. They are tumour suppressors crucial for the repair of double-strand DNA breaks. *BRCA1* is also necessary for cell replication and DNA synthesis, but due to epigenetic inactivation or mutations, increases BrCa risk [17-19]. Single nucleotide polymorphisms (SNPs) seem to increase BrCa risk in women with or without *BRCA1* mutation. Several other gene mutations have been associated to high risk of BrCa: *ATM* (DNA repair); *BARD1*

(DNA repair); *BRIP1* (DNA repair); *CDH1* (cell adhesion); *CHEK2* (cell growth regulation); *MRE11A* (DNA repair); *MSH6* (DNA repair); *NBN* (DNA repair); *PALB2* (DNA repair); *PMS2* (DNA repair); *PTEN* (cell growth regulation); *RAD50* (DNA repair); *RAD51C* (DNA repair); *STK11* (cell growth regulation); *TP53* (cell growth regulation); *HER2* (epidermal growth factor receptor, which activates cell proliferation, survival and adhesion) [20-24].

4) *Estrogen levels:* Early menarche and late menopause increase BrCa risk, given the increased exposition to estrogens [25]. The breast cells from women that did not have a full-term pregnancy remain very active and responsive to estrogen. Thus, these women have a higher BrCa risk [25, 26]. Due to an increased exposure to estrogen, women who have their first child after age of 30 also have higher BrCa risk compared to women who gave birth before that. Pregnancy also reduces the total number of lifetime menstrual cycles, which may explain its protective effect against BrCa. Lactation also decreases BrCa risk, especially if it lasts longer than 1 year. Women who started menstruating younger than age 12 have higher BrCa risk later in life, as well as women who go through menopause older than 55. The earlier breasts form, the sooner they interact with internal or external hormones and with chemicals in hormone disruptor products, which increases BrCa risk. In addition, the hormonal replacement therapy (HRT) used by post-menopausal women have been strongly associated to BrCa diagnosis, increasing the risk up to 75% [27].

5) *Life habits*: Diets rich in carbohydrates and fat, combined with lack of exercise, frequently conduct to overweight and obesity, which increases the risk of BrCa, especially after menopause. In addition, overweight and obesity also increase the risk of recurrence after treatment. This relation between body weight and BrCa risk is due to the capacity of the adipocytes to produce estrogen, which may trigger BrCa develop and grow. The higher is the exposition to extra estrogen over time, the higher is the risk to develop BrCa. Several studies also show an association between exercising regularly (4 to 7 hours/week) to a lower BrCa risk. Regular exercise seems to lower the levels of insulin growth factor in the blood, hormone that can affect BrCa development. Moreover, people that exercise regularly tends to maintain a healthy weight, therefore, having less body fat than inactive people. Regarding diet, it is also known that women with an abundant processed meat or fast food intake have higher BrCa risk than women with a diet rich in fruits and vegetables [28]. Alcohol consumption is associated with BrCa risk in women, by increasing the levels of estrogen and other hormones associated with hormone-

receptor-positive BrCa or by damaging the DNA. The consumption of three alcoholic drinks per week seems to be enough to increase the risk of BrCa 15% and the risk goes up another 10% for each additional drink women regularly have each day. There are also some evidences that alcohol consumption, even a few beverages per week, increase the risk of relapse after BrCa treatment. Importantly, girls from 9 to 15 years old that consume 3 to 5 drinks per week have 3 times more probability to develop benign breast lumps, which later in life can become carcinogenic. In addition, smoking is linked to a higher risk of BrCa in younger, premenopausal women and very heavy second-hand smoke exposure increases BrCa risk in postmenopausal women. Moreover, Catsburg and colleagues, found a positive association between duration and intensity of cigarette smoking, as well as being a smoker prior to first pregnancy, with the risk to develop BrCa [29].

### 1.1.3 Man Breast Cancer

Male BrCa is rare and it can affect 1 in 100,000 men, accounting for 1% of all BrCa cases diagnosed [30]. It has some similarities with females BrCa, however, upon its discovery, it tends to present a more advanced stage, and consequently a worse prognosis [31]. In all male BrCa cases, the most common type of invasive carcinoma is the ductal, being the *in situ* and invasive lobular carcinomas the most rare cases in males [32]. Incidence of male BrCa cancers increases with age (>60) and is influenced by the genetic background, particularly associated to the *BRCA1* and *BRCA2* mutations [6]. About of 20% male BrCa cases have a first-degree family member with BrCa [32, 33]. The Klinefelter syndrome, gynecomastia, prostate, testicular or liver diseases, obesity, or alcohol consume, can have an impact on the incidence of man BrCa [34]. In addition, occupational exposures to radiation or polycyclic aromatic hydrocarbons could have an implication on the appearance of these cancers in males [35]. Decrease in mortality in male breast cancer is particularly associated to adjuvant systemic treatments, since there is no screening implemented for men [36]. Due to the lack of mammary tissue in males, skin and chest wall involvement is not rare and mastectomy is the standard surgical treatment [32].

### 1.1.4 Screening

BrCa screening is done by evaluating the breast before any signs or symptoms of the disease and is a practice implemented in several countries [37]. These trials have proven that screen-detected, non-palpable, preclinical tumours of ≤ 15 mm have a better prognosis.

Therefore, BrCa screening is helping to diminish the death associated to this disease. However, incidence of breast cancer is still increasing, in part due to a more efficient detection using periodical screening, along with an increase of the risk factors [38].

Currently, there are four reliable methods for BrCa detection: mammography, breast magnetic resonance (MRI), clinical breast exam and breast self-awareness [9]. The most valid and reliable method is mammography, since it can detect BrCa at an early stage, when treatment is more likely to have a positive outcome [39]. This is a multidisciplinary method and accuracy rate can be from 85% to 90%, which will lead to a reduction in mortality rate from 30% to 50% [9, 40]. However, risk of over-diagnosing or false-positive screening has to be taken into account of this method [37]. In families with *BRCA* mutation, screening is performed annually with magnetic resonance imaging and mammography combined, to detect the cancer at less invasive stage [37]. Clinical exam is also important for early stage detection and surveillance, spatially in women that have a greater risk for BrCa. In this group are included women with radiographically dense breasts, breast implants, women who have been submitted to radiotherapy or patients without access to an optimal health care [9]. This examination consists in a manual palpation of the breasts and local regional lymph nodes [41]. In countries with low income and underdeveloped health services, breast self-exam is an alternative approach that should not be, however, neglected even when there are other resources [9].

In order to positively confirm BrCa, a sample of the primary tumour and cytology/histology of axillary nodes, if involvement is suspected, needs to be analyzed by a pathologist [37]. The sample should be taken with a core needle biopsy, and must be retracted before any treatment for pathological diagnosis, who should be done according to the World Health Organization classification and the tumour-node-metastasis staging system [37]. The exam should analyze histologic type and grade, and determination of estrogen receptor (ER), progesterone receptor (PgR) and HER2 receptor status by IHC or FISH/CISH assay [41].

### 1.1.5 Breast anatomy

The breast, or mammary gland, is a structure able to produce milk in females. It is composed mainly by adipose tissue and extends from the collarbone down to the underarm and across to the middle of the ribcage (Fig.3). It can present different anatomies and physiologies, depending on age, hormonal status, genetic background and daily life habits [42] and, as glandular structures, they are very perceptive to hormonal changes in the body [43]. They have lobes, which are divided into smaller lobules connected via milk ducts [43]. The adipose tissue of the breast is supplied by a network of nerves, blood vessels, lymph

vessels, lymph nodes, and is also composed by fibrous connective tissue and ligaments [43].

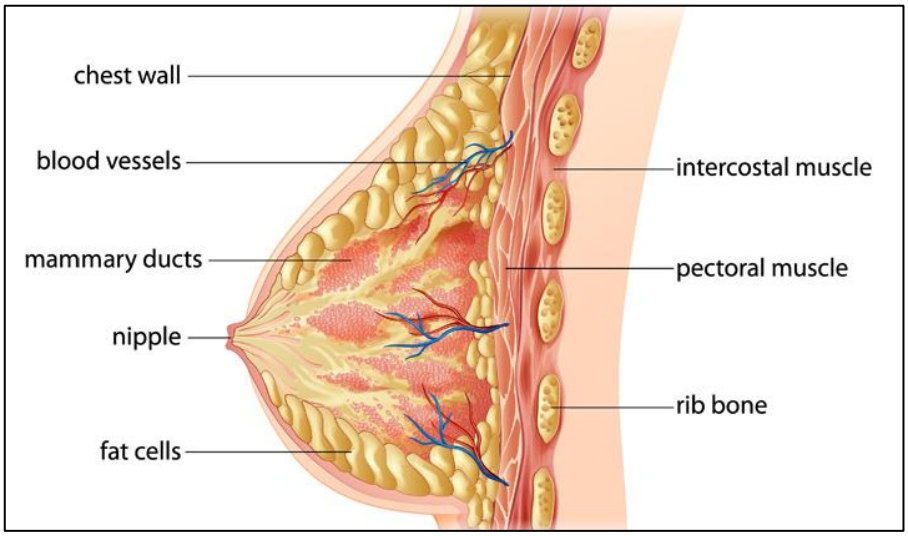


Figure 3 - Breast anatomy. All mammals have mammary glands that produce breast milk to feed their offspring [44].

### 1.1.6 Histopathological types

BrCa is a highly heterogeneous pathology that has different clinical outcomes and responsiveness to treatments [45, 46]. In fact, several lines of evidence suggest that BrCa is, in fact, a collection of different diseases with distinct risk factors, clinical presentation, pathological characteristics, response to therapy and outcomes, which originate in the same anatomical structure (the terminal duct-lobular unit) [47]. Most BrCa are adenocarcinomas (95%) and their phenotypic range can be correlated with diversity in gene expression patterns [46]. Bearing in mind the cell morphology, growth and architecture patterns, there are more than 21 histological subtypes described [48]. However, BrCa can be divided into *in situ* and invasive carcinomas, which are classified as ductal or lobular according to the origin site [49, 50].

1) Ductal carcinoma *in situ* (DCIS) is non-invasive potentially malignant intraductal proliferation of epithelial cells, characterized by cellular or nuclear atypia, that can evolve to invasive carcinoma [49]. Classification of DCIS depends on the degree of nuclear atypia, intraluminal necrosis, miotic activity and calcification, then it can be divided into three types, Low-Grade, Intermediate-Grade and High-grade, which are correlated to the risk of becoming invasive [49]. Historically, DCIS has been classified into five architectural subtypes: comedo, solid, cibriform, papillary and micropapillary. DCIS is consistently positive for E-cadherin and β-catenin.

2) Lobular carcinoma *in situ* (LCIS) is characterized by intra-lobular proliferation of small, and loosely cohesive cells, originated in the terminal duct lobular unit (TDLU). It constitutes a risk factor but a nonobligatory precursor for subsequent development of invasive carcinoma [49]. There are two immunohistochemical markers to distinguish LCIS, the lack of E-cadherin and β-catenin expressions and the positivity for high molecular weight (HMW) keratin [51].

3) Invasive ductal carcinoma (IDC) is a malignant ductal proliferation along with stromal invasion in the presence or absence of DCIS and can be classified according to cell type, number, type and location of secretion, architectural features and immunohistochemical profile [49]. It is the most frequent invasive breast cancer and accounts for 55% of breast cancer incidence upon diagnosis.

4) Invasive lobular carcinoma (ILC) is a malignant abnormal proliferation of neoplastic cells in the lobules, which invaded the stroma. ILC represents 5-15% of all breast cancer cases, and its incidence appears to be increasing, affecting postmenopausal women, and it could be related to hormone replacement treatment. ILC has 5 distinct histological groups, classic type, pleomorphic, histiocytoid, signet ring and tubolobular [49].

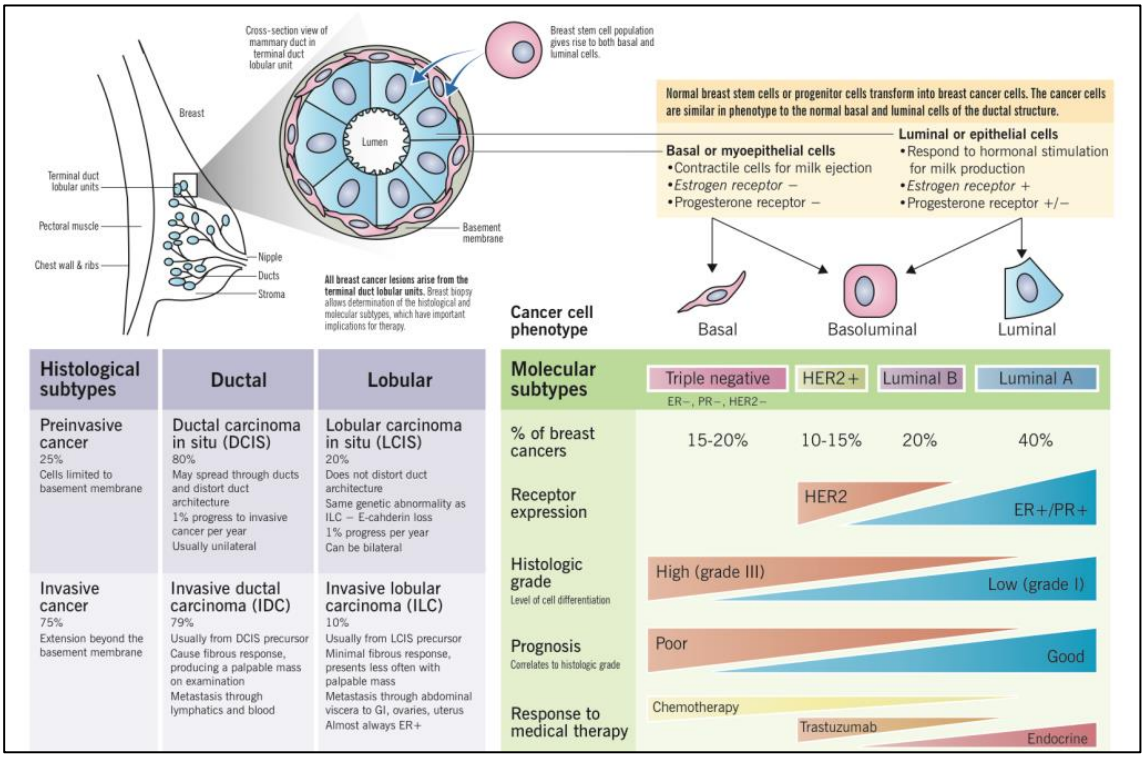


Figure 4 - BrCa pathogenesis and histologic versus molecular subtypes, by Eric Wong and Joana Rebelo.

### 1.1.7 Molecular subtypes

Human BrCa heterogeneity correlates with different gene expression patterns [52], which can be used as a prognostic marker [53]. Using hierarchical clustering, Perou and colleagues realized that part of the BrCa produce estrogen receptor (ER+) and the remaining one do not produce this hormone receptor (Fig.4,5) [52]. The ER+ tumours fall into luminal subtypes (A and B), given that the luminal cells have the most enriched levels of ER. The ER- tumours (low to absent) fall into Basal, HER2-enriched, Normal Breast-Like subtypes and Triple Negative [53, 54].

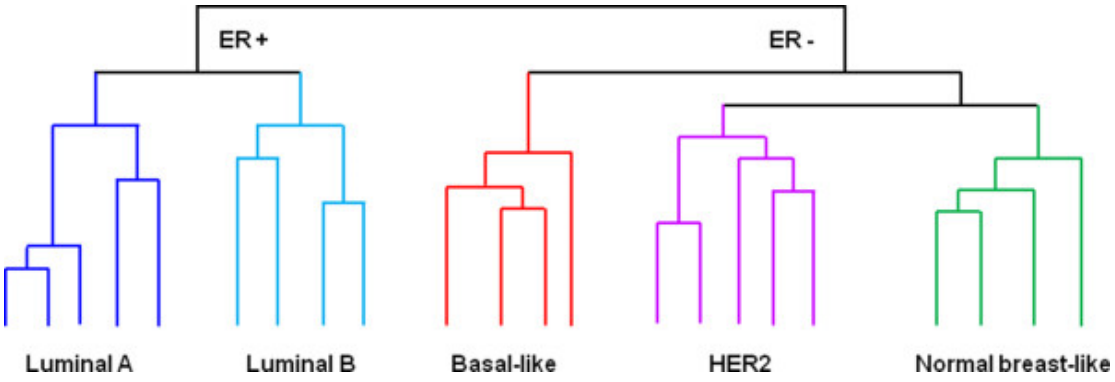


Figure 5 - Schematic illustrations of the five major clusters that represent the molecular subtypes of BrCa based on Perou and colleagues work carried out a cDNA microarray analysis of 38 invasive BrCa, 1 ductal carcinoma in situ, 1 fibroadenoma and 3 normal breast samples, and a number of biological replicates of tumours from the same patients [52]. Hierarchical clustering led to identify four subtypes: luminal, normal breast-like, human epidermal growth factor receptor 2 (HER2), and basal like. Later, it was shown that similar molecular subtypes can be identified in multiple cohorts and that luminal cancers can be subclassified into luminal A and B [53].

Thus, there are five molecular subtypes of BrCa according to their gene expression profiles belonging to five distinct groups:

1) Luminal tumours are the most common subtype of BrCa cases, and they can be divided into two subtypes, Luminal A and B, being the luminal A the most common. These subtypes express hormone receptors for estrogen (ER) and progesterone (PR). However, while Luminal A does not express the human epidermal growth factor receptor 2 (HER2) and have low levels of Ki-67 protein, which is involved in tumour growth regulation, the luminal B subtype may express or not HER2, always have high levels of the protein Ki-67 [50, 54]. Luminal B tumours have a higher expression of genes involved in cell proliferation when compared with luminal A subtypes [55]. In general, luminal tumours have a good prognosis, but luminal A subtype have a significant better prognosis than luminal B subtype, being a low-grade subtype, while luminal B usually is a higher-grade, having a lower survival rate [50, 54].

2) The HER2+ enriched subtype, are not only overexpressing this gene but are also amplifying other genes in the HER2 amplicon, 40 to 80% of tumours from this molecular subtype have a mutation related to p53 gene, subsequently being associated with a poor prognosis [50, 54, 56]. This group also has a low expression of basal-related genes such as keratin 5 [55].

3) Basal tumours have an increased expression of basal markers such as keratins 5, 6, 14 and 17, and EGFR. This molecular subtype is also known to have increased expression of genes related with proliferation and to higher probability to present

low *BRCA1* expression, and higher frequency of p53 mutation, which controls proliferation and maintains the integrity and stability of the genome, this gene is actually mutated in approximately in 80% of the cases [54, 57-59]. Most cases of basal tumours are triple negative breast cancer cases (TNBC), which account for 60-90%, of these tumours. TNBCs do not express ER and PR, and HER2 gene is not amplified or overexpressed [54, 60]. Overall, TNBC presents an aggressive phenotype, being histologically classified as high-grade, invasive, ductal carcinomas of no special type with basal-like features, thus it would not express biomarkers such as ER, PR and HER2, but would express cytokeratins, EGFR, vimentin and p-cadherin, which are characteristic of high-grade invasive breast carcinomas [60, 61].

4) Recent studies have defined a new molecular subtype of BrCa, claudin-low, which is typically triple negative. This subtype is characterized by downregulation of tight junction proteins, such as E-cadherin, and by overexpression of genes associated with epithelial-mesenchymal transition and BrCa cell phenotype [50, 59, 62].

5) Another subtype of BrCa is the normal-like subtype, which is a controversial group since some authors think that this subtype represents normal cell contamination on samples rather than a real subtype, also this group expresses genes associated with adipose tissue and other stromal cell types [50]. This subtype accounts for 7,8% of all breast cancer cases, and its gene expression is characterized by a high expression of genes of basal epithelial cells and adipose cells [52, 54].

### 1.1.8 Biological Markers

After BrCa diagnosis, there is the need to determine patient prognosis and identification of the most appropriate adjuvant systemic treatment. For a more accurate prognosis, some characteristics can be reviewed such as clinicopathological prognostic factors and predictive biological markers. Clinicopathological characteristics include lymph nodes with metastasis, tumour size and tumour grade. Although there are several new molecular tests, these factors are still very important for aiding therapy in newly diagnosed cases [63]. However, these markers are insufficient for proper patient treatment, especially in a time where personalized care is becoming a fundamental aim in biomedical research. Thus, with that in mind there was a need to discover additional molecular biomarkers that could correlate with clinical outcomes and treatment response. ER, PR biomarkers are currently used to predict response to endocrine therapy and HER2 to predict the benefit of trastuzumab treatment [64].

ER stimulates BrCa cell growth by binding to regulatory elements in the genome. Thus, when ER is activated in BrCa patients they can benefit from anti-estrogenic therapy [64] and currently, ER detection is used to identify patients for adjuvant treatments targeting ER and its downstream targets, preventing BrCa proliferation [64]. PR is used as a biomarker alongside with ER. Indeed, it was shown that ER could stimulate PR expression and that PR in the presence of progesterone could modify the location at which ER binds to chromatin. Thus, this alteration could lead to a switch in the regulation of genes that take part in cell proliferation, apoptosis and differentiation and so, when PR is present in ER+ BrCa patients the prognostic is better [64].

The third most used predictive biological marker is HER2. Its measurement is required on cases of invasive BrCa. HER2, when overexpressed, can activate MAPK and PI3K/AKT that leads to tumour growth, enhancing cell proliferation, invasion and metastization [20, 64]. HER2 is overexpressed in 15-20% of BrCa cases and, when ER is overexpressed, in nearly 80% of BrCa cases [64, 65]. This biomarker is used for predicting response to anti-HER2 therapy [64].

## 1.2 Hox genes and cancer

### 1.2.1 Homeobox Genes and homeodomain

HOX genes were primarily discovered in *Drosophila* and found to have a common sequence of 183bp, the homeobox, which encodes a 61-amino acid DNA-binding homeodomain (Fig.6) [66, 67]. This sequence was found to be expressed in all three kingdoms of multicellular organisms [66]. The conservation that is observed between all organisms, extends past the gene sequences into their relative positioning along the chromosome [68]. During vertebrate evolution, HOX genes underwent gene duplications followed by specific gene loss events, which resulted in 39 HOX genes divided in four clusters in mammals [69, 70]. Holland and colleagues divided the HOX genes into several groups: group 1 and 2: (HOX1 and HOX2, set to be the “anterior” genes), group 3 (HOX3), group 4 (HOX4-8 genes, set to be the “central” genes), group 5 (HOX9-13 genes, set to be the “posterior genes” [71]. In humans, HOX clusters are located in four distinct chromosomal loci: HOXA Chr 7p15.3; HOXB Chr 17q21.3; HOXC Chr 12q13.3 and HOXD Chr 2q31, and each can contain between 9 to 11 genes [72, 73].

These genes encode a family of transcription factors that regulate early morphogenesis, for example controlling the anterior-posterior and proximo-distal differentiation during development by interfering with the cell fate within these axes [68, 72, 74, 75]. In a certain way, HOX genes act as a selector gene, since their combinatorial expression can play a major role in identity of an individual body segment [76]. In addition, these genes were found to regulate cell proliferation, specification and cell death and are required for non-transcriptional functions such as DNA replication and repair, and mRNA translation [77]. Interestingly, their position within the cluster reflects the relative positions of the structures that they specify along this axis. Thus, if a gene is at the 3’ end of a cluster it is expressed in anterior regions of the embryo, while if it is located closer to the 5’ it is expressed in more posterior regions of the embryo. Akam and colleagues, suggested that during early *Drosophila* development HOX expression domains are determined by three regulatory inputs: transcriptional regulation from earlier segmentation genes, cellular memory system based on the action of Polycomb/tithorax group proteins and cross-regulatory interactions among the HOX genes themselves [78].

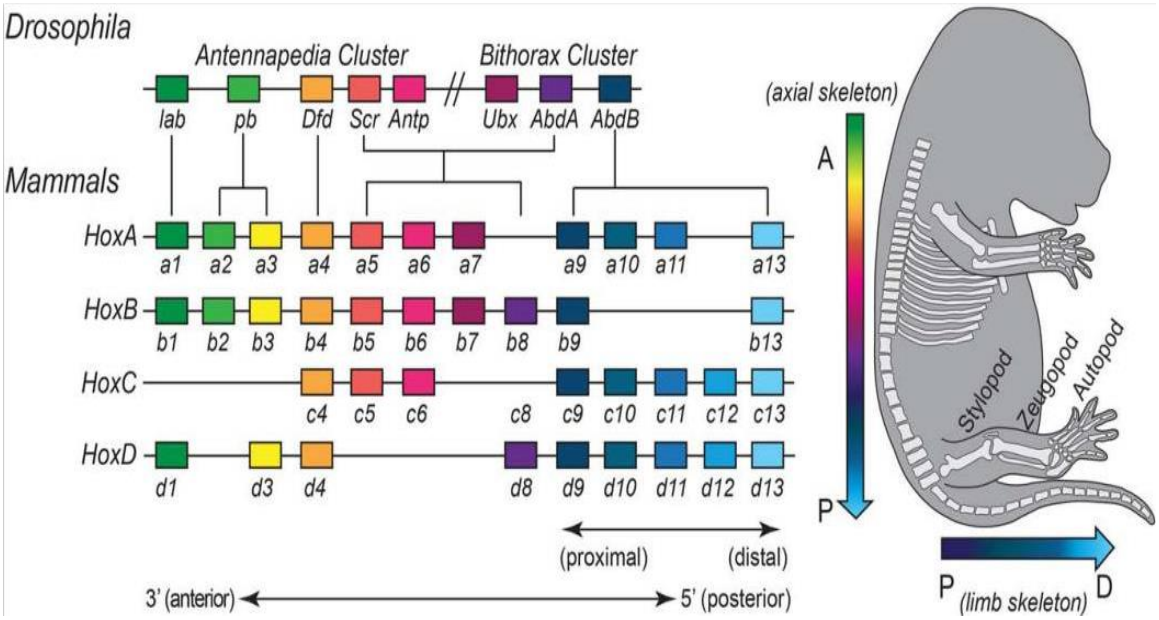


Figure 6 - Localization of the Hox gene within the clusters and their expression in mouse embryos. In the left, a coloured box represents each Hox gene and the line represents the non-coding regions linking them [70]. Color-coding highlights conserved relationships between *Drosophila* and mammalian *Hox* genes, and the paralogous relationships within the mammalian cluster. In the right, Hox expression along the antero-posterior and proximo-distal axes reflects the formation of distinct skeleton elements along these axes.

### 1.2.2 Altered HOX gene expression in cancer

As explained previously, these genes have a key role in development. However, they also are implicated in tumour initiation and progression during adulthood. In fact, atypical expression of homeobox genes have been described in tumours, and such can be due to re-expression of HOX genes that were only activated during embryonic development, and then gained expression in neoplastic cells [5, 79]. In some cases, there are homeobox genes that are expressed in neoplastic cells derived from tissues where that particular gene was not expressed in embryonic stages, or in other cases the homeobox genes can be downregulated in malignant cells that are derived from a tissue where a particular gene is expressed in the fully differentiated state [79]. The role of HOX genes in cancer seems to be highly tissue-specific, given that they can be up or downregulated, depending of the cancer tissue. For example, *HOXA9* is silenced in lung cancer cases through epigenetic mechanisms but is upregulated in acute lymphocystic leukemia [80]. In contrast, *HOXB13* is upregulated in BrCa, but it is silenced in prostate cancer [80]. These genes can also have two distinct roles in tumours, acting as oncogenes by promoting cell growth and invasion, or as tumour suppressors by modifying cell survival and morphogenesis [81, 82].

Actually, HOX genes are often deregulated in both solid tumours and liquid tumours. This family of genes are expressed during hematopoiesis and each cluster have a specific pattern of lineage-restricted expression, where HOXA genes are expressed in myeloid cells, HOXB genes in erythroid cells, HOXC genes in lymphoid cells and HOXD genes are not expressed in this process. *HOXA9* is one of the HOX genes that is widely studied, being a key regulator of hematopoiesis and as an oncogenic role in leukemia, enhancing the expression of proliferative genes and inhibiting apoptosis [83].

HOX genes expression can be regulated through epigenetic control, as described for example for *HOXA9* in lung cancer [80]. These genes have CpG islands in the promoters, and when silenced these islands can be found methylated. In addition, HOX gene expression can be affected by histone modifications [81]. In consequence of these mechanisms, among others, abnormal activity of HOX genes can affect pathways that promote tumourigenesis, or as a result of their altered expression they can maintain a more embryonic state through the activation of anti-apoptotic pathways or suppression of differentiation [81]. HOX genes can act as transcriptional activators in some types of cancer, but in other cases they have a role as transcriptional repressors, so upregulation and downregulation of these genes is crucial to promote tumourigenesis [84].

### 1.2.3 Altered HOX genes in breast cancer

The mammary glands start to develop during embryonic development with the formation of the mammary line, or milk line, and differentiation of placodes within this line (Fig.7A). At birth, these placodes already descend into the underlying mesenchyme forming the rudimentary ductal structure of the mammary glands. Then, subsequent stages of development are observed during pubertal, pregnancy, lactation, and involution states. During puberty, branching morphogenesis initiates under the influence of growth hormone (GH), estrogen and insulin-like growth factor 1 (IGF1). Upon pregnancy, progesterone and prolactin combined action generates alveoli able to secrete milk during lactation. The involution state initiates when lactation stops and involves remodelling of the gland back to the pre-pregnancy state. These processes require a variety of signalling pathways that regulate specialized subpopulation of mammary stem cells potentiating drastic changes in the gland occurring with each pregnancy. Particular HOX genes have been implicated in all these development states of the mammary glands, participating in their embryonic development, branching morphogenesis, alveoli formation and involution processes [82, 85] (Fig.7B).

Several studies report abnormal HOX gene expression in BrCa: *HOXA6, HOXA13, HOXB2, HOXB4, HOXB5, HOXB6, HOXB7, HOXB8, HOXB9, HOXC5, HOXC13, HOXD1*, and *HOXD8* [84]. Hur and colleagues quantified the expression levels of the HOX genes in BrCa using reverse transcriptase PCR showing that 14 out of 25 HOX genes, present statistically significant differences between malignant and non-malignant breast cancer tissues. Other studies regarding invasiveness and migration of BrCa cells, have demonstrated that 11 HOX genes were up-regulated in BrCa in comparison with normal breast tissues [67]. However, there are HOX genes that are downregulated in BrCa. *HOXA5*, for example, presents decreased expression in 60% of BrCa, which has been associated to apoptosis modulation in BrCa [84]. Similarly, *HOXD10* appears to have less expression in epithelial cells, as malignancy progresses. Indeed, when *HOXD10* expression is restored in cell lines, cell migration seems to diminish [81].

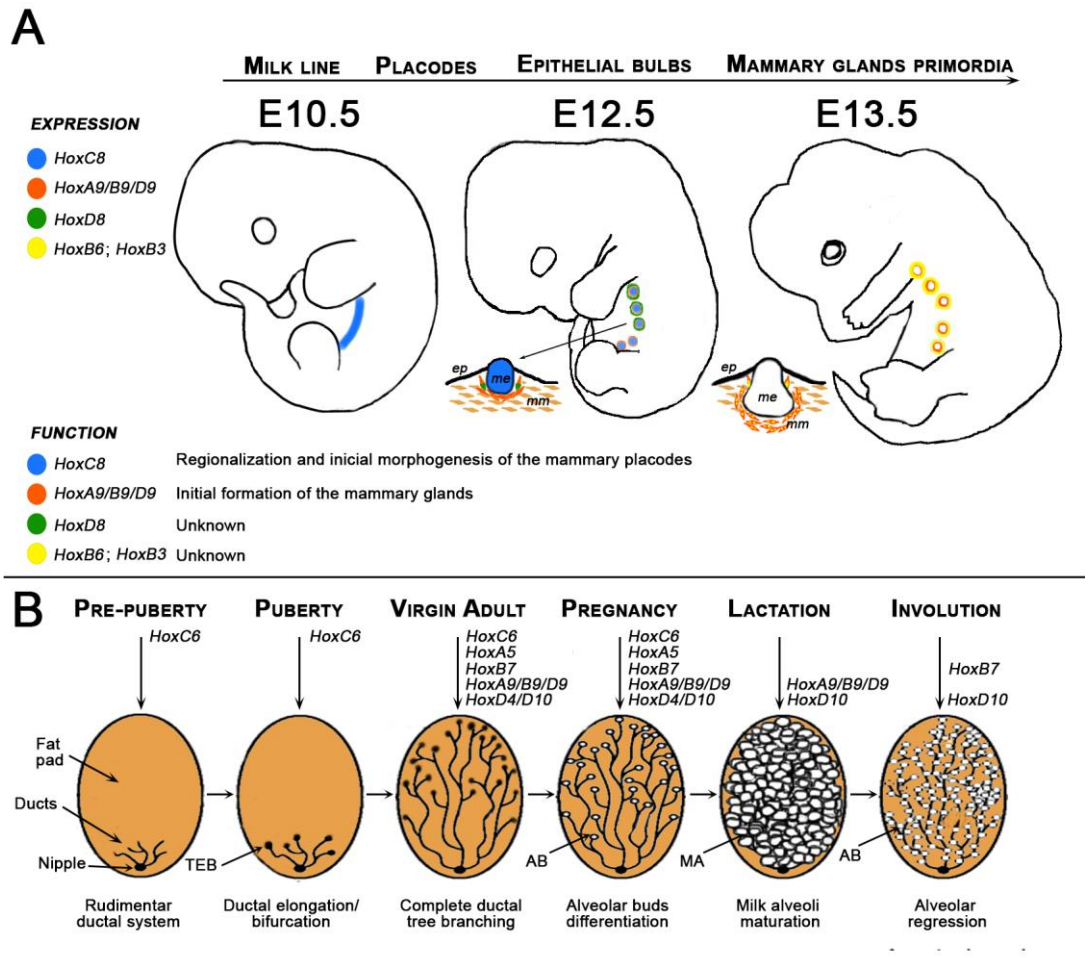


Figure 7 - *Hox* gene expression and function during mammary gland (MG) development in mice (*in* article submitted [86]). A. MG embryonic development. Gene expression is colour-coded in the image. *ep*: epithelium; *me*: mammary ectoderm; *mm*: mammary mesenchyme. Note ventrolateral surface ectoderm (milk line) marked by the expression of *HoxC8* at E.10.5 and then restriction of several Hox genes to mammary placodes or surrounding mesoderm in subsequent stages. B. MG development after birth and detected Hox gene expression. *TEB*: Terminal embryonic buds: *AB*: Alveolar buds; *MA*: mature milk alveoli.

### 1.2.4 HOXB7-9 genes in breast cancer

*HOXB7* has been associated with tissue development of the mammary gland and with BrCa development [85]. Microarray analyses suggest that *HOXB7* is significantly overexpressed in both the primary cancer and distant metastasis [87, 88]. Moreover, in cell lines, overexpression of this gene is directly or indirectly responsible for regulating the expression of vascular endothelial growth factor (*VEGF*) and basic fibroblast growth factor (*bFGF*), which promote BrCa progression [89]. Indeed, *HOXB7* overexpression in BrCa has been associated to BrCa progression and to the promotion of the epithelial-to-mesenchymal transition (EMT), through up-regulation of *bFGF* and activation of the MAPK pathway [81, 88, 90, 91]. Importantly, some studies suggest that *HOXB7* overexpression increases resistance to tamoxifen through activation of receptor tyrosine kinase pathways.

In other solid tumours, *HOXB7* function has been associated to cancer progression and poor clinical outcome and suggested as a potential prognostic factor and therapeutic target [90]. Moreover, there are studies that reveal that this gene is overexpressed in nearly 60% of BrCa cases. This was seen either in cell lines (MDA-MB-231, MDA-MB-468, MDA-MB-453, MCF-7, SkBr-3) or BrCa specimens [90, 91]. Interestingly, HOXB7 protein seems to have an important role in DNA repair. Rubin and colleagues have shown that BrCa cell lines expressing *HOXB7* present an enhanced survival after irradiation, so this resistance could allow these cells to accumulate mutations that might lead to tumourigenesis [90].

*HOXB8*, which is localized immediately after *HOXB7*, within the HOXB cluster, is known to be overexpressed in colorectal cancer (CRC), when compared to normal colon tissue [92]. A study performed by Wang and colleagues, revealed that when *HOXB8* is silenced can lead to a decreased proliferation, migration and invasion of CRC cells [93]. In BrCa, *HOXB8* does not seem to be deregulated in cell lines or breast cancer tissues, taking in consideration gene expression detection performed using reverse transcriptase PCR [67].

*HOXB9*, located after *HOXB7* and *HOXB8* in the cluster, and potentially sharing regulatory mechanisms with them, encode a transcription factor with major roles in mammary gland development [85]. Apparently, the deregulation of this gene is strongly associated to cancer progression, being involved in increment of cell migration, invasion, angiogenesis, and in the formation of lung metastasis derived from primary BrCa [94-96]. The relevance of its deregulation for cancer progression makes it an important prognostic factor in different cancer types [94, 95, 97].

*HOXB9* expression and function, has been studied mainly in the gastric cancer (GC) context, where it has been associated with a poor survival rate, when overexpressed,

inducing tumour progression, invasion and occurrence of metastasis. Kato and colleagues revealed that *HOXB9* expression increased in 51,7% of GC patients, and its expression was localized in the nucleus of the GC cells. Also, patients with "positive" *HOXB9* expression, seemed to exhibit a poorer prognosis when compared to those with a lower expression, being highly associated with lymphatic invasion and lymph node metastasis [98]. However, another study suggests that overexpression of *HOXB9* in GC cells leads to an increase in the apoptotic rate, thus, reducing the metastatic ability [99]. In endometrial cancer, *HOXB9* overexpression correlates with cell migration and to the upregulation of an oncogenic transcription factor (E2F3), which controls cell cycle, differentiation, apoptosis and response to stress [100]. The overexpression of *HOXB9* was also linked to colon cancer, where it seems to contribute for the formation of metastasis and a poor survival rate of patients [101]. This transcription factor was studied in BrCa, being upregulated when compared to normal breast tissues and it was associated with lower overall survival, relapse-free survival, distant metastasis-free survival and post-progression survival for patients diagnosed with BrCa [102].

Seki and colleagues studied *HOXB9* expression in BrCa through immunohistochemical staining [103]. Regarding those results, it was notable that 69 (48,9%) of a total of 141 tumour specimens stained positive and were associated with a higher nuclear grade [103]. Hayashida and colleagues, found *HOXB9* to be overexpressed in 42% of all BrCa cases analyzed, especially on those with high histological grade. *HOXB9* abnormal expression in BrCa was proposed to be a tumour progression factor rather than a tumour-initiating event, since its expression has oncogenic effects, such as inducing growth factors that can alter tumour-specific cell fates, angiogenic factors and also leading to the formation of metastasis [103, 104]. Indeed, *HOXB9* levels of expression relate to an increased expression of VEGFs that contributes to tumour progression and influence intratumoural microvessel density (IMD), a measure of neovascularization surrounding the tumour tissue [105]. Thus, higher *HOXB9* expression can increase IMD and, therefore, contribute to a worse prognostic in BrCa patients. Shrestha and colleagues also demonstrated that *HOXB9* is upregulated in BrCa tissues, being its expression also higher in various tumour cells including ER+ BrCa cells, MCF7 [106]. Thus, *HOXB9* overexpression in BrCa can be due to a higher estrogenic activity present in BrCa tissues [106].

Other studies attempted to apply ionizing radiation to *HOXB*-overexpressing breast cells showing that they become more resistant due to an enhanced DNA damage response, due to an increase of baseline ataxia telangiectasia mutated (ATM) phosphorylation [107]. So, in non-irradiated cells *HOXB9* can induce detectable DNA damage, such as double-strand DNA breaks, that will increase ATM phosphorylation, leading to DNA damage response.

Thus upon radiation, BrCa cells that are overexpressing *HOXB9*, have their ATM hyperactivated in the early stages of the double-stranded DNA break response, leading to a more rapid response. In fact, production of growth factors (GF) and cytokines can alter tumour microenvironment and modulates cell fate, resulting on an increase of tumour aggressiveness and chemotherapy resistance. Thus, *HOXB9* overexpression triggers the production of TGF-β, which is GF, and it can alter cell fate by inducing EMT [107].

In summary, *HOXB9* overexpression can lead to an increased production of 1) TGF-β controlling cell cycle, inducing cell differentiation and promoting apoptosis; 2) ErbB ligands inducing epithelial to mesenchymal transition and invasion; 3) bFGF supporting differentiation 4) VEGFs, promoting angiogenesis; 5) NRG2 influencing growth, differentiation and survival of different cell types [104, 106]. Also, *HOXB9* influences BrCa progression, triggering ErbB ligands and TGF-β production, which causes phosphorylation and posterior activation of HER2/3 and TGF-β pathways, respectively. Although, activation of HER2/3 promotes cell motility, the TGF-β pathway promotes EMT [107].

# 2. Preliminary Results

Our lab has been investigating the role of HOXB genes in BrCa, particularly *HOXB7* and *HOXB8*. Primarily, the HOX expression profile was characterized through qPCR and western blot in MDA-MB-231 and MDA-MB-468 BrCa cell lines, followed by HOX functional analysis through knockdown of *HOXB7* using specific siRNAs. These assays made us suggest that HOXB7 Tf may regulate CDH1. However, our lab then performed aggregation assays, after successfully knocking down *HOXB7* in MDA-MB-231 cell line, using as control cells that form compact aggregation as spheroids (MCF-7). The results demonstrated that silencing of *HOXB7* in MDA-MB-231 cells does not affect the rate of aggregate formation in the experimental condition, in comparison with the control. Thus, the silencing of *HOXB7* seems to be insufficient to modify the function of CDH1 proteins.

In order to further explore the role of *HOXB7* in BrCa, we went on to characterize its expression in distinct BrCa molecular subtypes, with the aim to update these analyses using the qPCR approach and also to facilitate the selection of the most suitable cellular models for future functional assays *in vitro* and *in vivo*. To this end, the lab expanded the analyses of gene/protein expression analyses to cell lines MCF-7, BT-474, SKBR3, MDA-MB-231 and MDA-MB-468 (Fig.8A). Human tissues were also analyzed, belonging to a cohort of BrCa patients from IPO-Porto. The evaluation of the cell lines leads us to conclude that *HOXB7* was overexpressed in all of them, in comparison with normal cells (MCF10A). However, the expression was particularly high in MCF-7 and BT-474, in accordance with previous studies by Hur and colleagues using RT-PCR instead of qPCR [67]. Complementary Western blot analyses further confirmed these results. Contradictory results emerged, however, when analyzing the expression levels of *HOXB7* in the cohort of BrCa patients. Where, *HOXB7* appeared to be more significantly expressed in normal breast (NBr) than in BrCa tissues. In addition, no significant differences were found in the expression levels of distinct BrCa molecular subtypes. Analyzing CDH1 expression in cell lines and tissues, we also noticed lack of differences between tissues belonging to distinct molecular subtypes, which were very clearly identified in the cell lines. Thus, these results suggested the need to increase the cohort of patients analyzed and to test methods of protein detection that might be more efficient to evaluate differential expression levels in NBr and BrCa tissues (immunohistochemistry).

*HOXB9* expression was also re-analyzed in our lab using qPCR and aiming to fulfill the same aims above described for *HOXB7*. The results showed that this gene is overexpressed in cell lines MCF-7, BT-474 and MDA-MB-231, when compared to MCF10A normal breast cells (Fig.8B). However, the expression levels of this gene were still uncharacterized in the cohort of BrCa patients from IPO-Porto previous to the initiation of this master thesis.

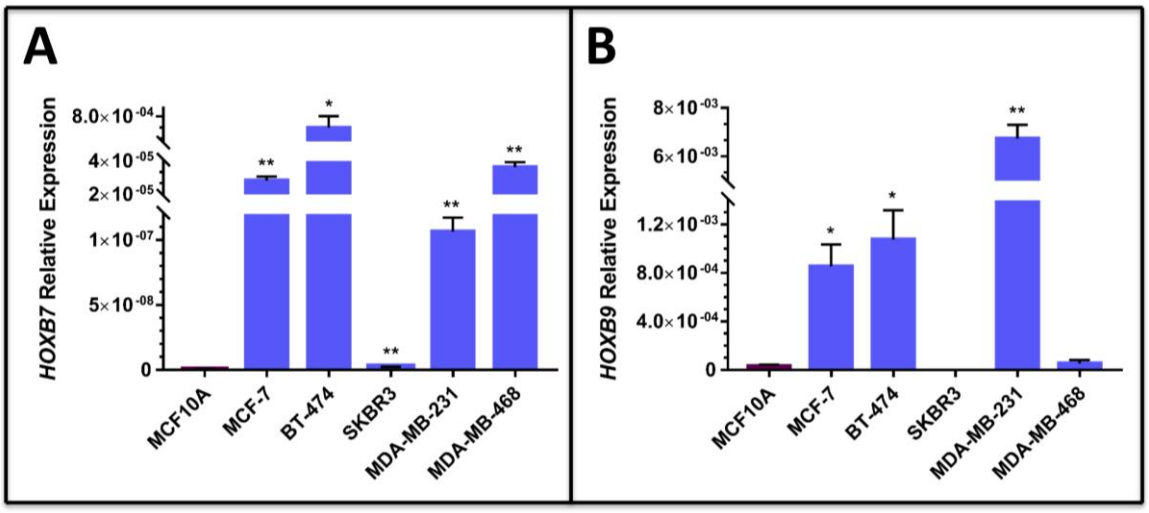


Figure 8 - *HOXB7* (A) and *HOXB9* (B) basal mRNA expression levels, analyzed by qPCR [108]. Expression levels of BrCa cells (MCF-7, BT-474, SKBR3, MDA-MB-231, MDA-MB-468) compared to normal breast cells (MCF10A). The * denotes *p*-value < 0,05, the ** denotes *p*-value < 0,01 by unpaired T test with Welch's correction.

Table 2 - BrCa cell lines that were used, represents different BrCa molecular subtypes.

| Breast cancer cell lines | Molecular subtypes |
|---|---|
| **MCF10A** | Normal breast |
| **MCF-7** | Luminal A |
| **BT-474** | Luminal B |
| **SKBR3** | HER2+ |
| **MDA-MB-231** | Triple negative |
| **MDA-MB-468** | Triple negative |

# 3. Aims

Overall, this project aims to overcome limitations from the literature regarding the characterization of *HOXB7* and *HOXB9* expression, at a transcriptional and protein level and considering BrCa cell lines and patients' biopsies from distinct molecular subtypes. This work will be important for the planning of future experiments aiming to further comprehend the function of these genes in BrCa.

Our specific aims were:

Aim 1 – Characterize basal expression profiles of *HOXB9* in BrCa tissues from a cohort of patients from IPO-Porto (n=128) representatives of the distinct molecular subtypes, using qPCR and characterize protein expression profile of HOXB9 using cell lines representing all recognized molecular subtypes of BrCa by western blot;

Aim 2 – Characterize HOXB7 and HOXB9 protein level in patients' biopsies, through immunohistochemistry.

# 4. Material and Methods

## 4.1 Patients sample

All human breast cancer and normal RNA samples were received from IPO-Porto. The BrCa samples were used to quantify HOXB9 protein and mRNA expression and HOXB7 protein expression. mRNA expression was analyzed through qPCR and protein expression through immunohistochemistry assay. This cohort group consisted of 127 patients, being 11 patient's cancer free and 117 patients with BrCa (table 8).

Table 3 - Clinicopathological characteristics of the cohort group from IPO-Porto.

| Characteristic | Values (%) |
|---|---|
| **Age (year), median (range)** | 60 (30-86) |
| **Primary tumour size (T)** | |
| T1 | 13 (10,2%) |
| T2 | 38 (29,7%) |
| T3 | 55 (43,0%) |
| Unknown | 11 (8,6%) |
| **Recidive** | 33 (28,2%) |
| Bone | 10 (32,3%) |
| Brain | 2 (1,6%) |
| Liver | 3 (2,3%) |
| Lung | 2 (1,6%) |
| Locoregional | 2 (1,6%) |
| Ganglionar | 1 (0,8%) |
| Ganglionar + Lung | 1 (0,8%) |
| Bone + Lung | 2 (1,6%) |
| Lung + Liver | 1 (0,8%) |
| Locoregional + Bone | 1 (0,8%) |
| Ganglionar + Bone | 1 (0,8%) |
| Breast + Bone | 1 (0,8%) |
| **Histological Type** | |
| Ductal invasive | 99 (77,3%) |
| Lobular invasive | 7 (5,5%) |
| Ductal invasive + DCIS | 8 (6,3%) |
| Unkwown | 2 (1,6%) |
| Metaplasic | 1 (0,8%) |
| **Molecular subtype** | |
| Luminal A | 33 (26,0%) |
| Luminal B | 42 (33,0%) |
| Triple negative | 32 (25,2%) |
| HER2+ | 9 (7,1%) |
| Normal | 11 (8,7%) |

## 4.2 cDNA synthesis

Total RNA was subjected to reverse transcription for cDNA synthesis, using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Vilnius, Lithuania) (Tab.2). To validate the cDNA synthesis, it was performed a PCR with GAPDH primers, adding reagents mix (New England Biolabs) to 0.5 µl of cDNA (Tab.3-4). A negative control was also used by adding 0.5 µl of RNase-free water instead of cDNA. The amplified products were then loaded into a 2% agarose gel, prepared by adding 1g of agarose in 50 mL of 1x TAE and dissolved in the microwave. The gel was cooled down and then 2 µl of Green Safe Premium were added (Nzytech, Lisboa, Portugal). The samples were prepared for the electrophoresis by adding 1 µl loading dye (Thermo Scientific, Lithuania, EU) to 5 µl of the amplified product. This mix was loaded into the agarose gel and run in 1x TAE buffer at 100 V alongside a 1 kilobase (Kb) ladder (Solis Biodyne). PCR was performed in a DNA thermal cycler (Tgradient, Biometra, Germany) and PCR electrophoresed products were visualized using an ultraviolet spectroscopy (Molecular Imager ® Gel Doc™ XR+ Imaging System,

Table 4 - cDNA synthesis.

| Reagents | Volume (ul) |
|---|---|
| RNA | 10 |
| 10x RT | 2 |
| 25 x dNTP | 0.8 |
| 10x | 2 |
| Multiscribe | 1 |
| Ribolock | 0.5 |
| $H_2O$ | 3.7 |
| Total | 20 |

Table 5 - Components of the PCR reaction to validate cDNA synthesis.

| Reagents | (µl) |
|---|---|
| cDNA | 0,5 |
| GAPDH F | 0,5 |
| GAPDH R | 0,5 |
| Master Mix | 10 |
| $H_2O$ | 8,5 |
| Total | 20 |

Table 6 - PCR amplification protocol for GAPDH. Primer Forward: ACTGGCGTCTTCACCACCAT; Primer Reverse: TCTTGAGGCTGTTGTCATACTTC; Product size: 142 bp.

| **Conditions** | *Step 1* | *Step 2* | | | *Step 3* | *Step 4* |
|---|---|---|---|---|---|---|
| Temperature | 95ºC | 94ºC | 60ºC | 72ºC | 72ºC | 4ºC |
| Time | 15 min | 3 sec | 90 sec | 90 sec | 10 min | Forever |
| Cycles | | | 30 | | | |

## 4.2 Quantitative real-time PCR analysis

Quantitative real-time PCR experiments (qPCR) were performed using glyceraldehyde-3-phosphate dehydrogenate (GAPDH), as a reference gene. qPCR reactions for *HOXB9* were done with 0.5µl of cDNA, 0.5µl of primer forward and reverse (Tab.5), 5µl of iTaq™ Universal SYBR ® Green SuperMix (Bio-rad, Hercules, CA, USA) and 3.5µl of DNase/RNase-free water. All experiments were carried in triplicates and were performed in real-time thermal cycler CFX96™ (Bio-rad, California, USA).

Table 7 - HOXB9 and GAPDH primers sequence, product length and annealing temperature.

| **Genes** | | *Primer Sequence (5'-3')* | *Product* | *Annealing* |
|---|---|---|---|---|
| *HOXB9* | FW | CTACGGTCCCTGGTGAGGTA | 198bp | 60º |
| | RV | TAATCAAAGACCCGGCTACG | | |
| *GAPDH* | FW | ACTGGCGTCTTCACCACCAT | 142bp | |
| | RV | TCTTGAGGCTGTTGTCATACTTC | | |

## 4.3 Immunohistochemistry

Immunohistochemistry assays were performed using NovoLink Max Polymer Detection System, Leica, using antibodies against HOXB7 (n=71) and HOXB9 (n=76) proteins, which recurred protocol optimization. Primarily, slides were de-paraffinized with xylene and re-hydrated in an ethanol series. Antigen retrieval was performed using sodium citrate and EDTA for HOXB7 and HOXB9 antibodies, respectively, in the microwave for 20 minutes in high power. Then, in order to neutralize endogenous peroxidase, peroxidase was used at 0,6%, thus, the non-specific background staining becomes significantly decreased. Afterwards, the slides were washed in TBST. To block albumin and immunoglobulins, horse serum was used in 1:50 for 20 minutes. Subsequently, the slides were incubated with the primary antibody (Tab.6) overnight, approximately 18 hours. Following this first incubation, slides were incubated with post primary and polymer, which is used as a substitute for the secondary antibody, for 30 minutes each. Next, they were washed in TBST and PBS, 3 minutes each. To develop peroxidase activity, DAB was used for 10 minutes. Hematoxylin was used, for 3 minutes, as a counterstain. Lastly, the slides were de-hydrated in an ethanol

series and posteriorly were mounted. To assess the quality of the staining, positive controls have to be used. As positive control for the staining we used bladder carcinoma for HOXB7 and colon adenocarcinoma for HOXB9. H-score was used for semiquantitive analysis of immunoreactivity of these antibodies. In summary, more than 2000 cells were counted in each sample, and the H-score was generated by adding the percentage of strongly stained nuclei, the percentage of moderately stained nuclei, and the percentage of weak stained nuclei, giving a possible range of 0-300 [109]. Receiver operating characteristic (ROC) curves analyzes were performed to select relevant thresholds for tumour positivity. The cut-off score with both maximum sensitivity and specificity was selected in order to have the greatest number of cases correctly classified and to decrease the number of false positives [110].. In addition, overall survival curves were created by Kaplan-Meier estimates and compared by log-rank tests, and it was analyzed with percentile 50 [98].

Table 8 - HOXB7 and HOXB9 Antibodies characteristics for immunohistochemistry assays.

| *Protein* | *Antibody* | *Species* | *Type* | *Incubation* | *Dilution* |
|---|---|---|---|---|---|
| HOXB7 | Sigma | Rabbit | Polyclonal | 18h | 1:100 |
| HOXB9 | Santa Cruz | Mouse | Monoclonal | 18h | 1:150 |

## 4.4 Western-Blot

The purified proteins (25 ug) were loaded into the wells of a 12% SDS-PAGE gel, along with molecular weight marker (Color Prestained Standard Broad Range, New England Biolabs). The proteins were then transferred from the gel to a nitrocellulose membrane (Invitrogen by Thermo Fisher Scientific) and afterwards blocked with 5% non-fat milk in TBS-T for 1 hour and 30 minutes at room temperature (RT) under agitation. Blots of the protein extract were probed with β-tubulin antibody (Santa Cruz Biotechnology), laminin antibody (Invitrogen) and HOXB9 antibody (Santa Cruz Biotechnology) primary antibodies and incubated overnight at 4ºC under agitation. The blots were washed with TBS-T, and TBS and incubated with secondary antibodies at RT for 1 hour and 30 minutes. The secondary antibody used was ECL anti-mouse IgG (GE Healthcare). The blots were washed in TBS-T and TBS. Revelation was performed using the Clarity™ Western ECL Blotting Substrate chemiluminescent reagent (Bio-Rad) and the Chemidoc Imaging System (Bio-Rad).

Table 9 - Antibodies characteristics for western blot.

| *Protein* | *Primary Ab* | *Dilution* | *Secondary Ab* | *Dilution* | *Protein kDa* |
|---|---|---|---|---|---|
| HOXB9 | Santa Cruz | 1:200 | Mouse | 1:1000 | 32 |
| B-tubulin | Santa Cruz | 1:200000 | Mouse | 1:1000 | 55 |
| Laminin | Invitrogen | 1:250 | Mouse | 1:1000 | 66 |

# 5. Results

## 5.1 *HOXB9* in BrCa

*HOXB9* has been studied in different types of cancer, being associated with poorer prognosis in colon and gastric cancer. In addition *HOXB9* deregulation in cancer appears associated to higher migration and invasiveness in oral squamous cell carcinoma [94, 98]. The limited number of studies regarding the function of *HOXB9* in BrCa suggests an involvement in tumour growth and angiogenesis, being frequently associated with a worse prognosis [104, 106]. Its expression levels have been characterized in the past using RT-PCR, which revealed that HOXB9 was overexpressed in BT474, MDA-MB-231 BrCa cell lines [67, 106]. In addition, the protein levels were analyzed, and 48,9% of the BrCa specimens stained positive for HOXB9 [103, 106].

After analyzing *HOXB9* expression by qPCR in BrCa cell lines belonging to all clinically relevant molecular subtypes (in Preliminary Results), we used the same method to *HOXB9* expression in human tissues from a cohort of patients (n=127), belonging to distinct molecular subtypes and compare it with NBr (Tab.8; Fig.9,10).

The median age at diagnosis of the cohort group was 60 years (range 30-86 years) and the majority of the tumours were pathologically diagnosed as ductal invasive (77,3%) (Tab.8). However some samples belong to other histological types such as lobular invasive (5,5%) and ductal invasive with ductal carcinoma *in situ* (DCIS) (6,3%), which was the second most common. This cohort group represents four distinct molecular subtypes: luminal A (n=33), luminal B (n=42), triple negative (n=32) and HER2+ (n=9). In addition, the metastasis occurrence was accountable for 33% of the patients, being the most common in bone (32,3%) followed by liver (2,3%).

*HOXB9* had slightly lower expression in BrCa tissue than NBr, however, statistical significance was not reached ($p > 0,076$) (Fig.9A). Furthermore, when we look at the molecular subtypes separately, it seems that triple negative and HER2+ have lower *HOXB9* mRNA expression, in comparison with other molecular subtypes, although statistical significance was not reached (Fig.9B). Previous studies addressing *HOXB* expression in BrCa suggested us that not all, but just a particular sub-set of patients may present aberrant *HOXB9* expression. In fact, analyses of *HOXB7* in BrCa tissues fail to detect over-expression in 100% of the cases, but instead showed higher expression in a percentage of 60% of the breast tumours analyzed [91]. Therefore, we depicted a sub-set of patients from our cohort (n=28), based on the median value of expression level of NBr and found significant differences (p value < 0,05) with the NBr (Fig.9C). Analyzes considering the molecular subtypes within this subset of patients, suggest significant overexpression (p

value < 0,001) of *HOXB9* in all types, with the exception of HER2+, which had only one case with increased *HOXB9* expression (Fig. 9D).

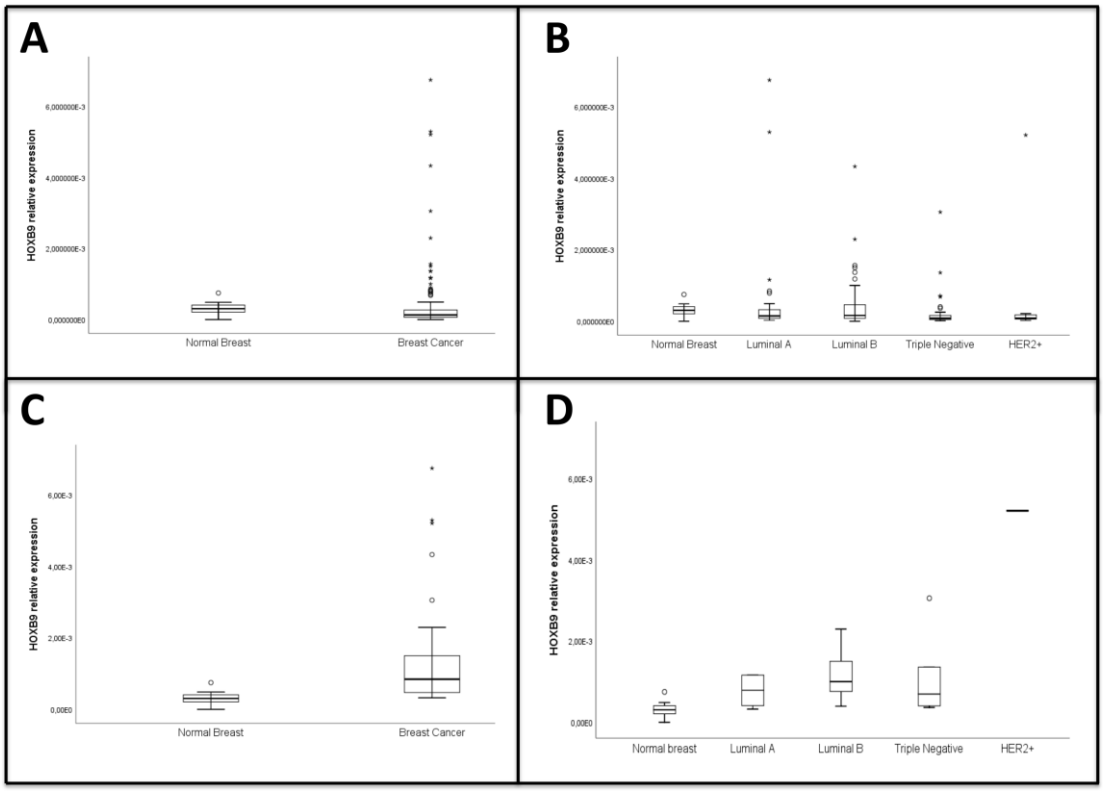


Figure 9 - Box-plots of mRNA expression levels of *HOXB9* in normal (NBr=11) and breast cancer tissues (BrCan=117) analyzed by qPCR. Y-axis as the $2^{-\Delta CT}$ values of *HOXB9* expression. Outliers were identified by the SPSS software and indicated with ˚ when considered as "Mild Values" and with * when considered as "Extreme Values". **A.** Comparison between NBr and BrCa tissues reveals no significant differences. **B.** Comparison between distinct molecular subtypes reveals that BrCa Triple negative and HER2+ seem to have lower expression than NBr tissue. Luminal A (n=33), Luminal B (n=43), Triple negative (n=32) and HER2+ (n=9). **C.** A group of 28 patients shown higher ($p < 0,001$) *HOXB9* expression in comparison with NBr (n=11). **D.** That group was analysed considering their molecular subtype. Luminal A (n=9), Luminal B (n=13), Triple negative (n=6) and HER2+ (n=1).

Thus, our results suggest that, or we have a considerably small percentage of cases in which *HOXB9* is aberrantly expressed or the normal controls do not allow us to detect the real differential expression between normal and BrCa tissues. Indeed, when we plot together *HOXB9* expression from normal breast cell lines (PCS-600-010, MCF10A) and our data from tissues, we detect much lower levels of expression in the cell lines than in NBr (Fig.10). Through, Mann-Whitney analysis with Bonferroni correction, BrCa tissue samples presents significantly differences in expression levels in comparison with both cell lines (*p* value < 0,05).

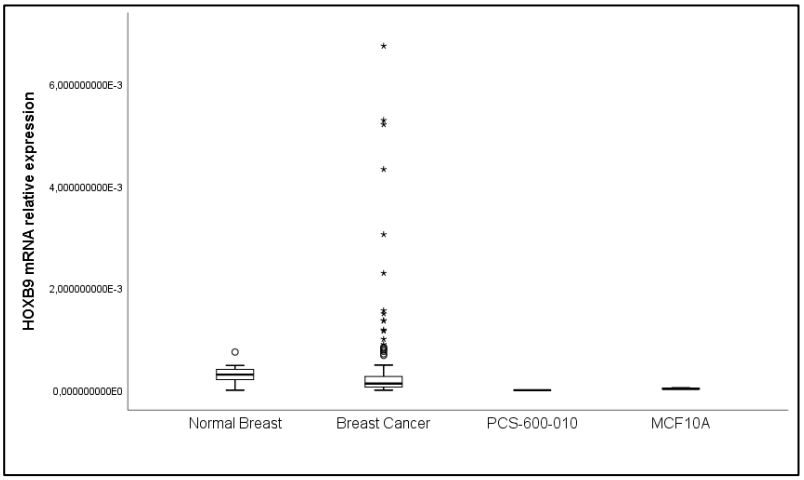


Figure 10 - Box-plots of mRNA expression levels of *HOXB9* in human NBr and BrCa tissues in comparison with normal breast cells (PCS-600-010, MCF10A). *HOXB9* is significant upregulated in NBr, when compared to the cell lines. Outliers were identified by the SPSS software and indicated with ˚ when considered as "Mild Values" and with * when considered as "Extreme Values". Y-axis as the $2^{-\Delta CT}$ values of *HOXB9* expression.

Furthermore, we have analyzed the differences of *HOXB9* expression regarding the grade of the tumour. Thus, as we can observe through Kruskal-Wallis analysis, there were not any significant differences (p values > 0,05) between the differents grades of the tumour (Fig. 11).

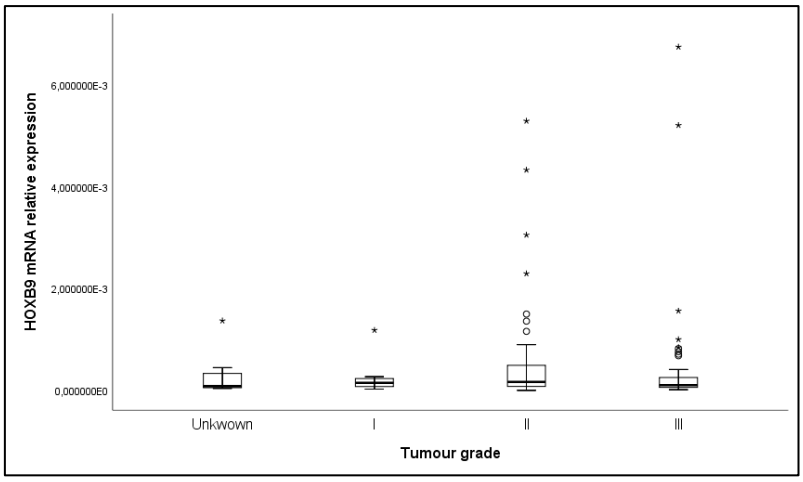


Figure 11 - Box-plots of mRNA expression levels of *HOXB9* in the different tumour grades. Outliers were identified by the SPSS software and indicated with ˚ when considered as "Mild Values" and with * when considered as "Extreme Values". Y-axis as the $2^{-\Delta CT}$ values of *HOXB9* expression.

## 5.2 Basal protein profile of *HOXB9* in BrCa cell lines

In order to contribute to ongoing work in the lab aiming to characterize HOXB9 expression in cellular models' representatives of the distinct BrCa molecular subtypes, I participated in the optimization of the western blot protocol for HOXB9 antibody. In terms of mRNA levels,

our team detected higher HOXB9 expression in MCF-7, BT474 and MDA-MB-231 cell lines when compared to MCF10A normal cells (*in* Preliminary Data). Here, we first conduct the protocol in one of these cell lines with higher expression levels of HOXB9 (MDA-MB-231), considering cytoplasm and nucleus extracts. Our results revealed higher levels of expression of HOXB9 in the nucleus than in the cytoplasm (Fig.12). Other replicates will now be conducted and densitometry analyses.

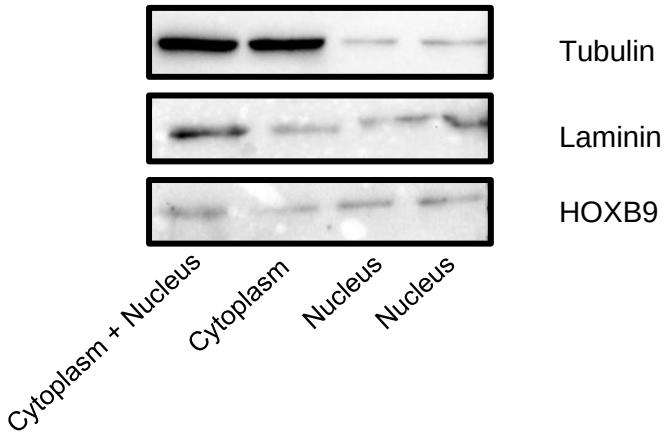


Figure 12 - HOXB9 basal protein expression level in MDA-MB-231 cell line by western blotting.

## 5.3 HOXB9 protein expression

In order to further explore the expression of *HOXB9* in our cohort of patients, we also characterized the protein levels through immunohistochemistry (n=76) (Fig.13). NBr tissue (Fig. 13B), presents strong staining for HOXB9, meaning that a high percentage of those cells shows positivity for HOXB9. In contrast, we can observe that there are some BrCa tissues that do not show any positivity for HOXB9 (Fig. 13C). Moreover, HOXB9 was strongly stained in some BrCa tissues (Fig. 13D), additionaly staining was observed in the nucleus of BrCa cells (Fig. 13E), as mentioned in previous reports [98, 103]. Analyzing the number of cells with nuclear positivity with different stainings, allowed to calculate an H-score value used to compare levels of expression in NBr and BrCa tissues These quantitative evaluations further suggest that NBr seems to have a higher expression of *HOXB9* in NBr (n=9) than in BrCa (Fig.14A). In addition, a ROC curve analysis was performed in order to understand the better threshold between NBr and BrCa (Fig.14B). Therefore, a threshold with high percentage of sensibility and specifity would be the cut-off score leading to greatest number of cases which were correctly classified as having or not the clinical outcome, in our case the threshold set was 142,3, with AUC of 0,902, meaning that cases with a H-score of 142,3 were considered as NBr. Thus, BrCa cases with a H-

score above that cut-off, were considered to be overexpressed. Overall survival (Fig. 15) was also analyzed with percentile 50, where patients above that percentile were considered HOXB9 positive, however there were not any significant statistical differences between *HOXB9* positive and negative patients, however it seems that HOXB9-positive patients have a slightly lower survival. Our results contradict previous studies reporting that *HOXB9* is overexpressed in 49% of breast tumours analyzed [103].

In addition, we also analyzed the protein expression of HOXB9 in different molecular subtypes of BrCa and we did not observe any significant differences between them (p value > 0,05) (Fig. 16), the same results were obtained when we analyzed tumour grade (Fig. 17).

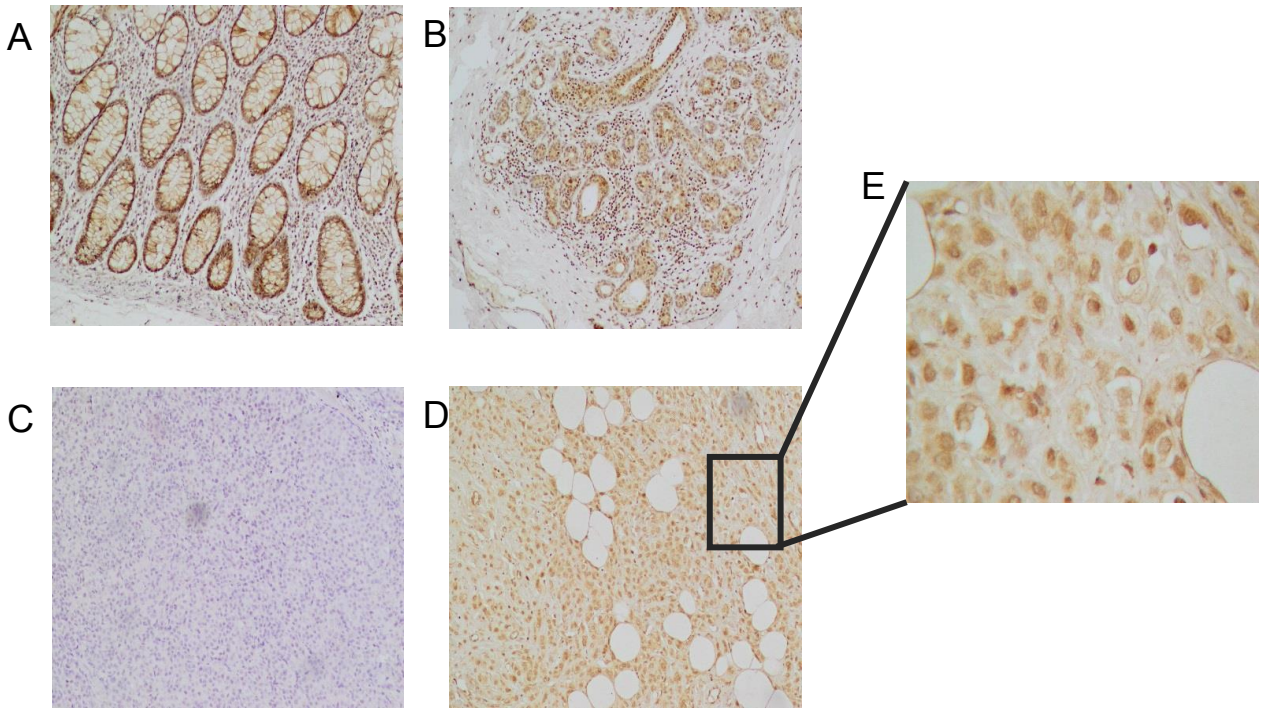


Figure 13 - HOXB9 expression in different tissues by immunohistochemistry. (A) HOXB9 expression in colon cancer (positive control), magnification of 10x; (B) HOXB9 stained positive in NBr, magnification of 10x; (C) HOXB9 stained negative in BrCa tissue, magnification of 10x; (D) HOXB9 stained positive in BrCa tissue, magnification of 10x; (E) Closer magnification (40x) of (D), HOXB9 stains positive in the nucleus of breast cancer cells.

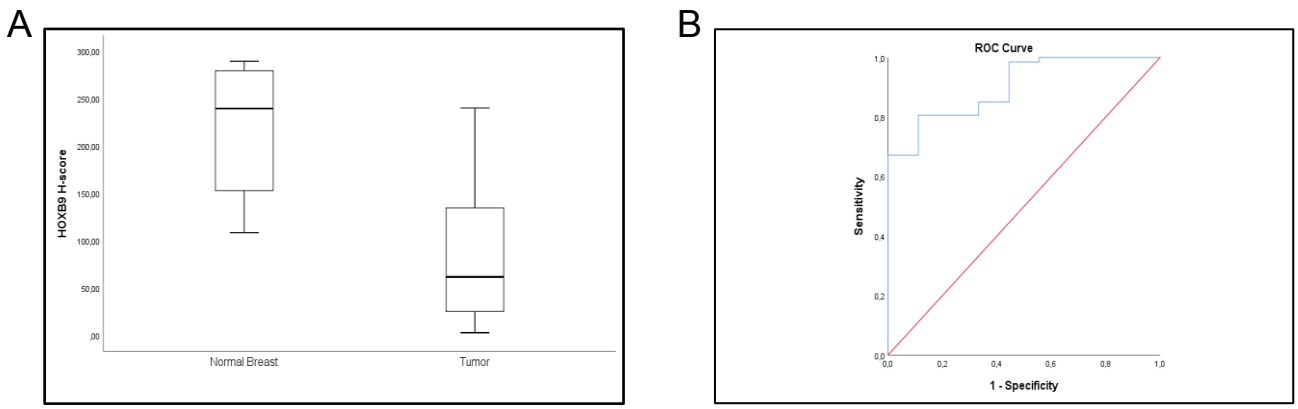


Figure 14 - HOXB9 protein expression in BrCa tissue. (A) Box-plots of HOXB9 H-score in human NBr and BrCa tissues analyzed through immunohistochemistry. (B) ROC curve for HOXB9 and differences between NBr and BrCa.

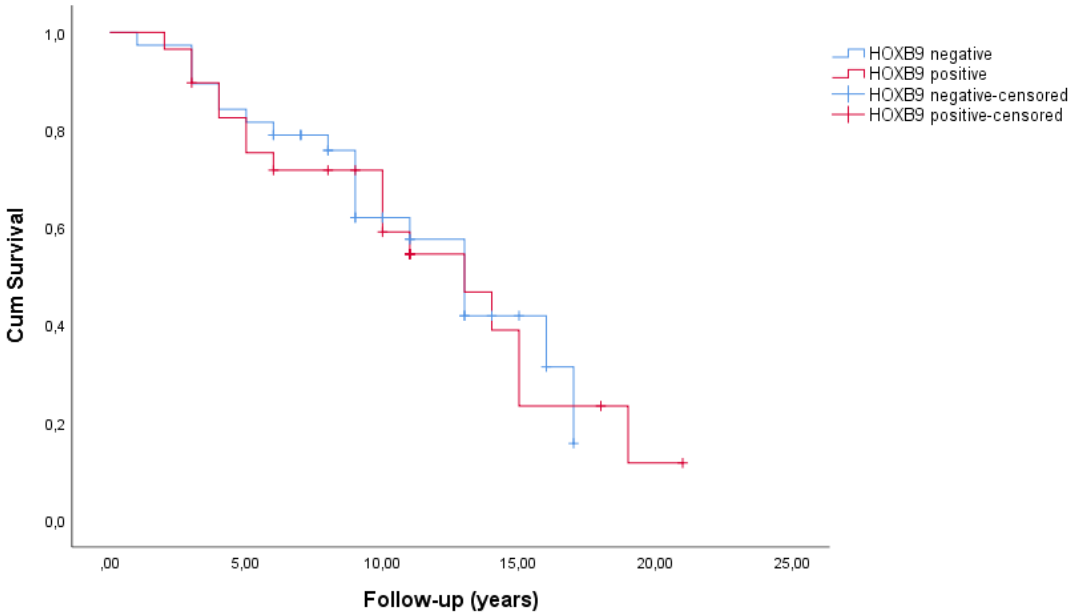


Figure 15 - Kaplan-Meier plots of overall survival in HOXB9-positive and HOXB9-negative patients with Breast Cancer.

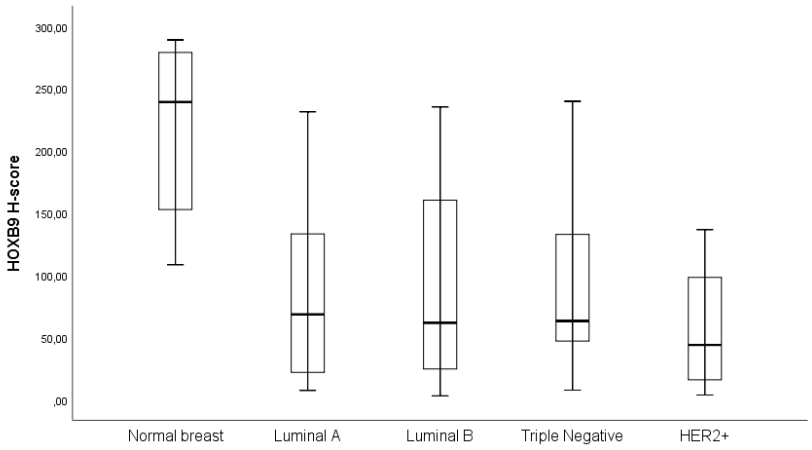


Figure 16 - Box-plots of hoxb9 H-score in different molecular subtypes of BrCa. Through Kruskal-Wallis analysis, we can observe that there are not any significant differences (p value > 0,05) between BrCa molecular subtypes.

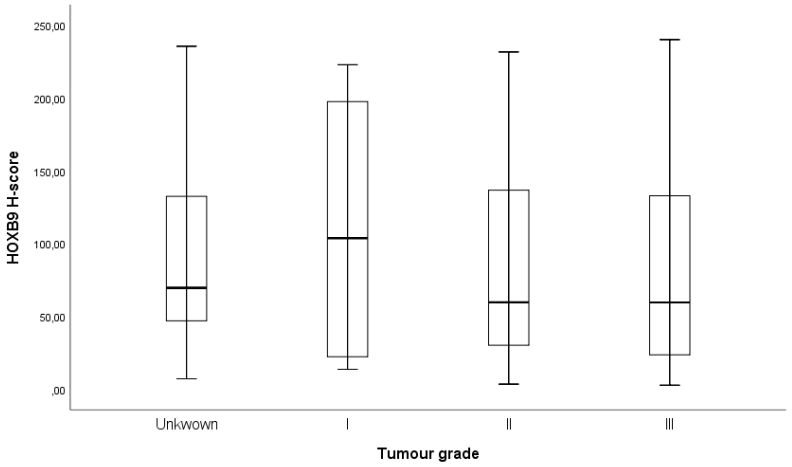


Figure 17 - Box-plots of HOXB9 H-score in different tumour grades. Through Kruskal-Wallis analysis, we can observe that there are not any significant differences (p value > 0,05) between different tumour grades.

## 5.4 HOXB7 protein expression

In order to complete the analyses previously done in the lab aiming to characterize the expression of *HOXB7* in a cohort of BrCa patients from IPO (*in* Preliminary Results), we made use of the expertise gained while analyzing HOXB9 expression in the course of this work and also performed immunohistochemistry analyses for HOXB7 proteins (NBr= 6; BrCa= 64) (Fig.18). Thus, we can observe that HOXB7 marked positively for our positive control (Fig. 18A), bladder cancer, as well for NBr tissue (Fig. 18B). However, there were cases where HOXB7 stained negatively (Fig.18C). However, we can observe that are cases of BrCa that stained positively for HOXB7 (Fig. 18D), and in a closer magnification we can

observe that HOXB7 stains positive in the nucleus of BrCa cells (Fig. 18E). As for HOXB7, quantitative analyses of protein expression suggest higher levels of expression in NBr than in BrCa (Fig.19). However, the limited number of NBr samples cannot be neglected. Regarding the threshold to distinguish NBr from BrCa tissues, we analyzed a ROC curve (Fig.19B), and the cut-off score was 284,4 with an AUC 0,989. In addition, through Kaplan-Meier survival analysis, with the use of percentile 50, we determine that patients that were above that percentile, were considered as HOXB7 positive. Thus, we could verify that patients that were classified as HOXB7 negative, presented slightly lower overall survival in comparison HOXB7 positive patients, however statistical significance was not reached (Fig.20). As for *HOXB7*, our results contradict previous studies reporting that *HOXB7* is overexpressed in 60% of BrCa cases analyzed [91]. In addition, we analyzed the protein expression of HOXB7 in different molecular subtypes through Kruskal-Wallis analysis, and as we can observe there are not any significant differences between them (p value > 0,05) (Fig. 21).

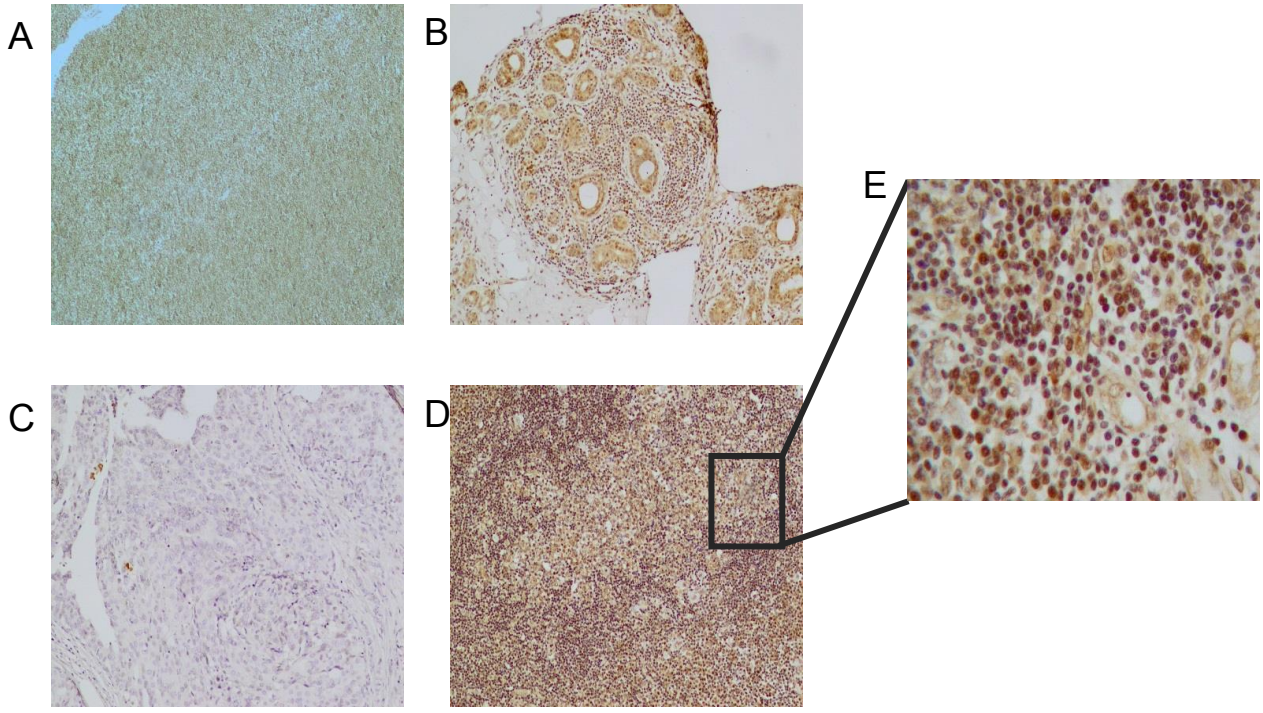


Figure 18 - HOXB7 expression in different tissues by immunohistochemistry. (A) HOXB7 expression in bladder cancer (positive control), magnification of 10x; (B) HOXB7 stained positive in NBr, magnification of 10x; (C) HOXB7 stained negative in BrCa, magnification of 10x; (D) HOXB7 stained positive in BrCa, magnification of 10x; (E) Closer magnification (40x) of (D), HOXB7 shows positive nuclear staining in breast cancer cells.

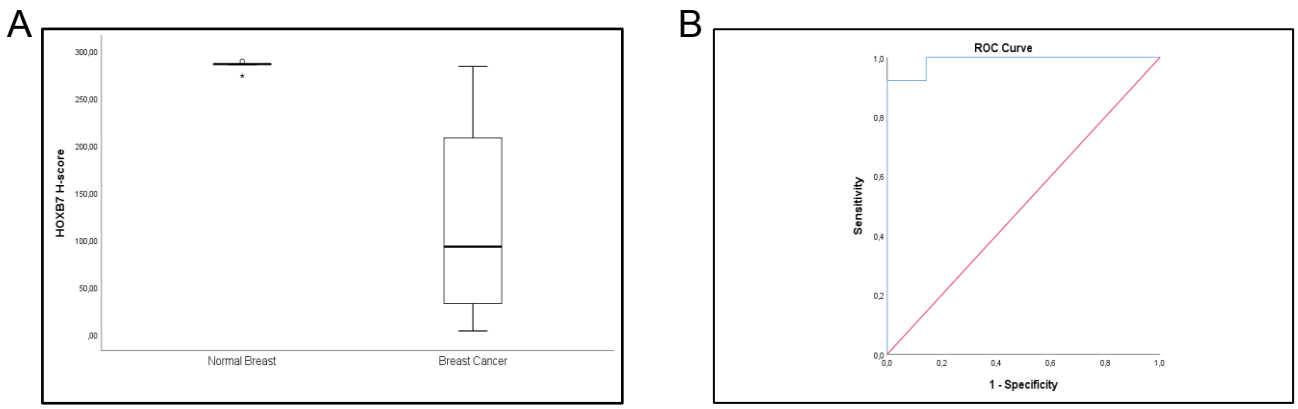


Figure 19 - HOXB7 protein expression in BrCa tissue. (A) Box-plots of HOXB9 H-score in human NBr and BrCa tissues analyzed through immunohistochemistry. (B) ROC curve for HOXB7 and differences between NBr and BrCa.

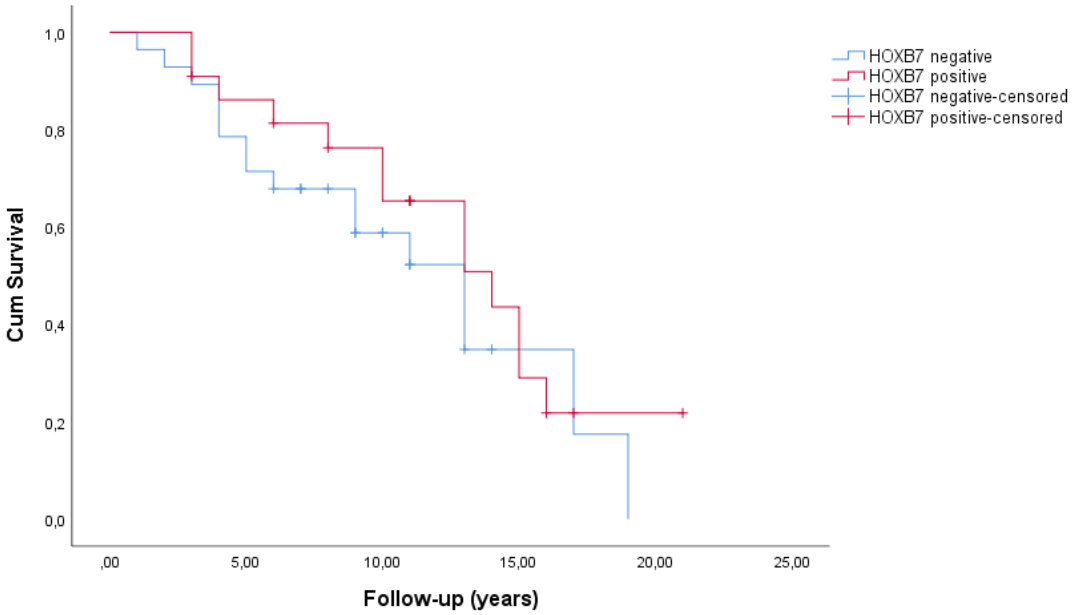


Figure 20 - Kaplan-Meier plots of overall survival in HOXB7-positive and HOXB7-negative patients with Breast Cancer.

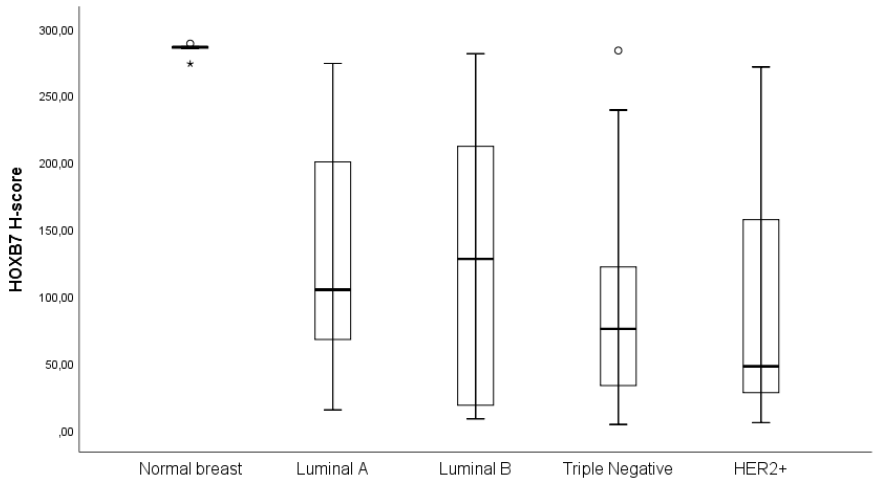


Figure 21 - Box-plots of HOXB9 protein expression in different molecular subtypes of BrCa. Through Kurskal-Wallis analysis, we can observe that there are not any significant differences (p value > 0,05) between BrCa molecular subtypes Outliers were identified by the SPSS software and indicated with ˚ when considered as “Mild Values” and with * when considered as “Extreme Values”.

# 6. Discussion

## 6.1 *HOXB* genes and BrCa

BrCa is the leading cause of death due to cancer among women, being the second most common cancer type that affects females [2, 103]. The role of HOX genes in embryogenesis has been focus of considerable attention, however HOX genes may also play a fundamental part in adult tissue, being essential in the cell memory program or for the formation of post-embryonic structures, such as the mammary glands [67, 106, 111]. These genes have been described to be frequently deregulated in several types of cancer [84], affecting proliferation, migration, invasion, differentiation and angiogenesis [81, 103]. Thus, the aberrant HOX gene expression can be critical for tumour progression [84].

The HOXB cluster have been studied in BrCa and was found to be deregulated in BrCa cell lines using RT-PCR analyses [67]. Moreover, deregulation of some of these transcription factors can lead to tumour progression through stimulation of epithelial-to-mesenchymal transition and angiogenesis [112]. Furthermore, it has been shown that HOXB cluster genes were upregulated in tamoxifen resistant MCF-7 cell line, conferring sensitivity to the cells when knockdown occurs [5]. In addition, many *HOXB* genes were found to be associated with BrCa progression through regulation of growth and angiogenic factors [91, 106]. Despite all the reports, the molecular mechanisms by which *HOXB* genes affect BrCa development remains unclear.

## 6.2 Role of *HOXB9* in BrCa tissue

As previously mentioned (*in* preliminary results), our lab started studying the expression of *HOXB9* in BrCa cell lines and the results achieved using qPCR were in accordance with the ones obtained in the past using RT-PCR [67]. However, our analyses allowed us to compare the expression levels in distinct molecular subtype cellular models. MDA-MB-231, representing the triple negative subtype, had the highest levels of *HOXB9* mRNA expression. However, qPCR analyses of BrCa tissues performed during this thesis revealed a different *HOXB9* expression profile. In contrast to the cell line results, triple negative subtype tissues did not appear to be overexpressing *HOXB9,* when compared with normal tissues. Furthermore, HER2+ had the lowest level of *HOXB9* expression, considering a limited number of for this subtype (n=9).

Thus, our results higlight the need to increase the cohort of patient tissues representing distinct molecular subtypes, in order to characterize the expression of *HOXB9* in distinct BrCa contexts. In additon, differences observed between cell lines and BrCa tissues can be

due to the use of random NBr tissue, which can increase variability. In fact, these tissues were retracted from mamoplasty procedures, from patients that did not have BrCa, but have an increased genetic risk due to hereditary factors. Moreover, optimal NBr controlos should be retracted from adjacent tumour tissue to be examined together with its corresponding BrCa tissue, due to BrCa genetic heterogenity [106]. Also, we did not have any control of the place within the tumour that was sampled to biopsy. Thus, the exact location of the BrCa biopsy might be an important parameter to compare molecular characteristics given the intra-tumoural heterogenity. These results will be, however, important to conduct future analyses integrating more NBr tissues, collected from locations adjacent to the tumours or from health human with no obvious genetic predesposition to develop cancer.

Furthermore, if we take only the *HOXB9*-overexpressing group into consideration, we have a significant difference between BrCa tissue and NBr tissue ($p< 0,05$). In addition, through statistical analysis we can observe that there are significant differences between all BrCa molecular subtypes ($p< 0,01$) in comparision with NBr. Through Kruskal-Wallis test with Bonferroni correction, we detected that Luminal B subtype was significant overexpressed in comparison with NBr tissue ($p < 0,05$). In addition, when we take a closer look in tumour tumour grading, we verify that  thaht variable did not present significant differences (p value > 0,05) (Fig. 11). Thus, these results suggest that *HOXB9* overexpresion is not a common feature of all BrCa patients, but most likely be specific for a sub-set of patients. However, our results makes us question if the use of random NBr tissue could , in fact, increase show a higher number of *HOXB9*-overexpressing patients. Thus, a randomized group and a greater number of control samples seems to be mandatory to better characterize *HOXB9* in BrCa tissue.

In addition, given that use of our cohort sample of NBr tissue could be increasing variability to our analysis, we have decided to compare the levels of expression of our controls with cell lines representative of normal breast (PCS-600-010, MCF10A). Through Kruskal-Wallis non-parametric analysis with Bonferroni correction, we found that *HOXB9* levels were significantly increased in NBr tissues in comparison to both cell lines ($p< 0,05$). This, further suggest us that NBr tissue variability could be interfering with our overall results.

## 6.3 HOXB9 Protein profile in BrCa biopsies

HOXB9 protein expression has been previously reported to be overexpressed in different cancer types, such in GC (51%) and BrCa (49%) [98, 103]. Our results suggest, however, that *HOXB9* more expressed in NBr than in Brca, taking in consideration evaluation of

mRNA and protein levels. However, our immunohistochemistry analyses clearly shows HOXB9 in the nucleus of BrCa cells, in accordance with Seki and colleagues [103]. However, the study that found 49% of *HOXB9*-overexpressing patients used distinct methods, given that they scored their results in accordance to the staining (none, weak or strong), not using our method of H-score [103]. Our results seem to be also contraditory to the ones presented by Shrestha and colleagues, where they compared cell densities in each BrCa case and compared them to its adjacent NBr tissue, indicating that *HOXB9* was overexpressed in BrCa [106]. In this case they had a smaller number of samples than us and their control group was from adjacent tissue from the tumour. This distinct results highlight the need to use appropriated controls to properly charcaterize *HOXB9* expression in distinct moelcular BrCa subtypes. Furthermore, we did not observe any significant difference (p value > 0,05) in HOXB9 protein expression between different molecular subtypes or tumour size (Fig. 16 and 17). Additionally, Kaplan-Meier curves for overall survival shows no signficant difference in our cohort group. However HOXB9 positive patients seemed to have a slightly lower survival, which is in accordance with another report [103]. Thus, this main differences between our results and others presented in previous reports could be due to variations in immunohistochemistry evaluation methods but most likely are due to the limited sample number. BrCa is a highly heterogeneous disease, so it is important to increase the number of tumour tissue samples and to acquire more NBr tissues, in order to reduce variability and charcaterize *HOXB9*-overexpression in light of BrCa molecular subtypes.

## 6.4 Basal protein profile of HOXB9 in BrCa cell lines

Taking in consideration our preliminary results of *HOXB9* expression levels on BrCa cell lines, we decided to start protocol optimization allowing western blot analyses of HOXB9 levels in MDA-MB-231 cells, the cell line with highest mRNA expression level. With that in mind, both the protein from the cytoplasm and the nucleus were used seperadetely, as well as protein from the cytoplasm and the nucleus together. Since, *HOXB9* is a transcription factor, we thought that protein would have an higher level of expression in the nucleus. As we can see from (Fig. 12), nucleus protein has slightly more HOXB9 expression than the rest of the samples. However, we still consider that this protocol could be further optimized and then used to characterize HOXB9 expression on the rest of BrCa cell lines trepresenting all recognized BrCa molecular subtypes. Thus, these results contributed to complete the carcaterization of *HOXB9* expressionconsidering the distinct molecular subtypes, that will be advantage to explore their function in distinct oncological contexts in the future.

## 6.5 HOXB7 protein expression

*HOXB7* was reported to be overexpressed in BrCa tissues and cell lines [90, 91], which was in accordence to our preliminary results *in vitro*, wherein we demonstrate that MCF-7, BT-474, MDA-MB-231 and MDA-MB-468 BrCa cell lines overexpresse *HOXB7.* Furthermore, it was also reported that *HOXB7* was overexpressed in BrCa tissue samples of IDC [91], However our preliminary data demonstrated an opposite tendency, whereby *HOXB7* was found to be significantly upregulated in NBr in comparion with BrCa [108]. The quantitative evaluation of the immunohistochemistry data, demonstrated that NBr has higher levels of HOXB7 protein than BrCa tissues. Thus, these results are in accordence with our preliminary results evaluating mRNA levels, which also show higher expression of *HOXB7* in NBr than BrCa tissues. Regarding, differences between all BrCa molecular subtypes, it was not visible any significant statistical variations (p value > 0,05) among them (Fig. 21).Thus, as mentioned while discussing reults for HOXB9, we believe that increment of NBr tissue samples could potentially diminuish variability in our study. In addition, given that BrCa is highly heterogenety disease, it would be important to increase sample number to better characterize HOXB7 expression patterns in BrCa tissues representing distinct molecular subtypes. In addition, Kaplan-Meier curves for overall survival were analyzed with percentile 50, and demonstrated that patients with lower HOXB7 expression had sligthly lower overall survival, however, statistical signifcance was not reached.

# 7. Conclusions

In this work, we attempted to provide the first quantitative analyses of HOXB9 expression in normal and breast cancer tissues representing all major molecular subtypes using qPCR.

In addition, we aimed to optimize immunohistochemistry protocols allowing HOXB7 and HOXB9 protein detection by immunohistochemistry in these samples, which we consider to be important to further characterize their degree of de-regulation in distinct BrCa molecular subtypes. We successful obtained positive staining of HOXB7 and HOXB9 in the nucleus of the cells, as reported in other studies using the same antibodies. Protein and mRNA quantitative analyses, however, provide results that contradict the literature and the data obtained by our lab using in cell lines: HOXB7 and HOXB9 seemed to be in higher levels in normal breast than in breast cancer tissues.

The data collected will be valuable with the addition of more samples, both “normal” and from breast cancers from distinct molecular subtypes. Due to the heterogeneity of breast cancer, it would be optimal to obtain control tissues adjacent to the tumour to reduce the variability of the data.

# 8. References

1. *Eurostat-EU* Eurostats-Statistics Explained 2015; Available from: https://ec.europa.eu/eurostat/statistics

2. DeSantis, C.E., et al., *International Variation in Female Breast Cancer Incidence and Mortality Rates.* Cancer Epidemiol Biomarkers Prev, 2015. **24**(10): p. 1495-506.

3. Ferlay, J., et al., *Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012.* International Journal of Cancer, 2015. **136**(5): p. E359-E386.

4. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA: A Cancer Journal for Clinicians, 2018. **68**(6): p. 394-424.

5. Yang, S., et al., *Up-regulation of HOXB cluster genes are epigenetically regulated in tamoxifen-resistant MCF7 breast cancer cells.* BMB reports, 2018. **51**(9): p. 450-455.

6. Ferzoco, R.M. and K.J. Ruddy, *The Epidemiology of Male Breast Cancer.* Curr Oncol Rep, 2016. **18**(1): p. 1.

7. Ferlay, J., et al., *Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.* Int J Cancer, 2015. **136**(5): p. E359-86.

8. Winters, S., et al., *Breast Cancer Epidemiology, Prevention, and Screening.* Prog Mol Biol Transl Sci, 2017. **151**: p. 1-32.

9. Coleman, C., *Early Detection and Screening for Breast Cancer.* Seminars in Oncology Nursing, 2017. **33**(2): p. 141-155.

10. *United Nations*, in *World Economic Situation and Prospectives in 2014.* 2014.

11. Forjaz de Lacerda, G., et al., *Breast cancer in Portugal: Temporal trends and age-specific incidence by geographic regions.* Cancer Epidemiol, 2018. **54**: p. 12-18.

12. Gomes, I., A. Miranda, and C. Nunes, *Spatiotemporal Analysis of Breast Cancer Incidence: A Study in Southern Portugal Between 2005 and 2012.* Anticancer Res, 2018. **38**(3): p. 1797-1805.

13. Vogel, V.G., *Epidemiology, genetics, and risk evaluation of postmenopausal women at risk of breast cancer.* Menopause, 2008. **15**(4 Suppl): p. 782-9.

14. Fu, J., et al., *How Young Is Too Young in Breast Cancer?-Young Breast Cancer Is Not a Unique Biological Subtype.* Clin Breast Cancer, 2018. **18**(1): p. e25-e39.

15. Apostolou, P. and F. Fostira, *Hereditary breast cancer: the era of new susceptibility genes.* BioMed research international, 2013. **2013**: p. 747318-747318.

16. Xie, Y., et al., *Mutation screening of 10 cancer susceptibility genes in unselected breast cancer patients.* Clin Genet, 2018. **93**(1): p. 41-51.

17. Phillips, K.A., *Current perspectives on BRCA1- and BRCA2-associated breast cancers.* Intern Med J, 2001. **31**(6): p. 349-56.

18. Narod, S.A. and L. Salmena, *BRCA1 and BRCA2 mutations and breast cancer.* Discov Med, 2011. **12**(66): p. 445-53.

19. Romagnolo, A.P., D.F. Romagnolo, and O.I. Selmin, *BRCA1 as target for breast cancer prevention and therapy.* Anticancer Agents Med Chem, 2015. **15**(1): p. 4-14.

20. Matsumoto, A., et al., *Biological markers of invasive breast cancer.* Jpn J Clin Oncol, 2016. **46**(2): p. 99-105.

21. Silwal-Pandit, L., A. Langerod, and A.L. Borresen-Dale, *TP53 Mutations in Breast and Ovarian Cancer.* Cold Spring Harb Perspect Med, 2017. **7**(1).

22. Darb-Esfahani, S., et al., *Role of TP53 mutations in triple negative and HER2-positive breast cancer treated with neoadjuvant anthracycline/taxane-based chemotherapy.* Oncotarget, 2016. **7**(42): p. 67686-67698.

23. Wang, L.L., et al., *PTEN/PI3K/AKT protein expression is related to clinicopathological features and prognosis in breast cancer with axillary lymph node metastases.* Hum Pathol, 2017. **61**: p. 49-57.

24. Zhou, X.J., et al., *PTEN expression is upregulated by a RNA-binding protein RBM38 via enhancing its mRNA stability in breast cancer.* J Exp Clin Cancer Res, 2017. **36**(1): p. 149.

25. Key, T.J., P.K. Verkasalo, and E. Banks, *Epidemiology of breast cancer.* The Lancet Oncology, 2001. **2**(3): p. 133-140.

26. Kapil, U., et al., *Reproductive factors and risk of breast cancer: A Review.* Indian J Cancer, 2014. **51**(4): p. 571-6.

27. *Type and timing of menopausal hormone therapy and breast cancer risk: individual participant meta-analysis of the worldwide epidemiological evidence.* The Lancet.

28. Kim, J.H., et al., *Dietary Factors and Female Breast Cancer Risk: A Prospective Cohort Study.* Nutrients, 2017. **9**(12).

29. Catsburg, C., A.B. Miller, and T.E. Rohan, *Active cigarette smoking and risk of breast cancer.* Int J Cancer, 2015. **136**(9): p. 2204-9.

30. Farr, D.E., et al., *Male Breast Cancer as a Second Primary Cancer: Increased Risk Following Lymphoma.* Oncologist, 2017. **22**(8): p. 895-900.

31. Doebar, S.C., et al., *Male breast cancer precursor lesions: analysis of the EORTC 10085/TBCRC/BIG/NABCG International Male Breast Cancer Program.* Mod Pathol, 2017. **30**(4): p. 509-518.

32. Shaaban, A.M., *Pathology of the male breast.* Diagnostic Histopathology, 2019.

33. Pritzlaff, M., et al., *Male breast cancer in a multi-gene panel testing cohort: insights and unexpected results.* Breast Cancer Res Treat, 2017. **161**(3): p. 575-586.

34. Serdy, K.M., et al., *Male Breast Cancer.* Am J Clin Pathol, 2017. **147**(1): p. 110-119.

35. Ruddy, K.J. and E.P. Winer, *Male breast cancer: risk factors, biology, diagnosis, treatment, and survivorship.* Ann Oncol, 2013. **24**(6): p. 1434-43.

36. Anderson, W.F., et al., *Male breast cancer: a population-based comparison with female breast cancer.* J Clin Oncol, 2010. **28**(2): p. 232-9.

37. Committee, o.b.o.t.E.G., et al., *Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†.* Annals of Oncology, 2015. **26**(suppl_5): p. v8-v30.

38. Kataja, V. and M. Castiglione, *Primary breast cancer: ESMO clinical recommendations for diagnosis, treatment and follow-up.* Ann Oncol, 2009. **20 Suppl 4**: p. 10-4.

39. Torre, L.A., et al., *Global cancer statistics, 2012.* CA Cancer J Clin, 2015. **65**(2): p. 87-108.

40. Sankaranarayanan, R., *Screening for cancer in low- and middle-income countries.* Ann Glob Health, 2014. **80**(5): p. 412-7.

41. Aebi, S., et al., *Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Ann Oncol, 2010. **21 Suppl 5**: p. v9-14.

42. Thomsen, S. and D. Tatman, *Physiological and pathological factors of human breast disease that can influence optical diagnosis.* Ann N Y Acad Sci, 1998. **838**: p. 171-93.

43. Akram, M., et al., *Awareness and current knowledge of breast cancer.* Biol Res, 2017. **50**(1): p. 33.

44. *The American Society of Breast Surgeons Foundation.* . Breast360.org 2019; Available from: https://breast360.org/topics/2017/01/01/breast-anatomy/.

45. Tang, Y., et al., *Classification, Treatment Strategy, and Associated Drug Resistance in Breast Cancer.* Clin Breast Cancer, 2016. **16**(5): p. 335-343.

46. Masood, S., *Breast cancer subtypes: morphologic and biologic characterization.* Women's health (London, England), 2016. **12**(1): p. 103-119.

47. Weigelt, B., F.C. Geyer, and J.S. Reis-Filho, *Histological types of breast cancer: how special are they?* Mol Oncol, 2010. **4**(3): p. 192-208.

48. Dieci, M.V., et al., *Rare breast cancer subtypes: histological, molecular, and clinical peculiarities.* Oncologist, 2014. **19**(8): p. 805-13.

49. Makki, J., *Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance.* Clin Med Insights Pathol, 2015. **8**: p. 23-31.

50. Vuong, D., et al., *Molecular classification of breast cancer.* Virchows Arch, 2014. **465**(1): p. 1-14.

51. Mastracci, T.L., et al., *E-cadherin alterations in atypical lobular hyperplasia and lobular carcinoma in situ of the breast.* Mod Pathol, 2005. **18**(6): p. 741-51.

52. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**: p. 747.

53. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.

54. Dai, X., et al., *Breast cancer intrinsic subtype classification, clinical use and future trends.* Am J Cancer Res, 2015. **5**(10): p. 2929-43.

55. Prat, A., et al., *Clinical implications of the intrinsic molecular subtypes of breast cancer.* Breast, 2015. **24 Suppl 2**: p. S26-35.

56. Bertheau, P., et al., *p53 in breast cancer subtypes and new insights into response to chemotherapy.* Breast, 2013. **22 Suppl 2**: p. S27-9.

57. Li, J.P., et al., *Association of p53 expression with poor prognosis in patients with triple-negative breast invasive ductal carcinoma.* Medicine (Baltimore), 2019. **98**(18): p. e15449.

58. Prat, A., et al., *Molecular features of the basal-like breast cancer subtype based on BRCA1 mutation status.* Breast Cancer Res Treat, 2014. **147**(1): p. 185-91.

59. Li, X., G.M. Oprea-Ilies, and U. Krishnamurti, *New Developments in Breast Cancer and Their Impact on Daily Practice in Pathology.* Archives of Pathology & Laboratory Medicine, 2017. **141**(4): p. 490-498.

60. Marotti, J.D., et al., *Triple-Negative Breast Cancer: Next-Generation Sequencing for Target Identification.* Am J Pathol, 2017. **187**(10): p. 2133-2138.

61. Barnard, M.E., C.E. Boeke, and R.M. Tamimi, *Established breast cancer risk factors and risk of intrinsic tumor subtypes.* Biochim Biophys Acta, 2015. **1856**(1): p. 73-85.

62. Prat, A. and C.M. Perou, *Deconstructing the molecular portraits of breast cancer.* Mol Oncol, 2011. **5**(1): p. 5-23.

63. Patani, N., L.A. Martin, and M. Dowsett, *Biomarkers for the clinical management of breast cancer: international perspective.* Int J Cancer, 2013. **133**(1): p. 1-13.

64. Nicolini, A., P. Ferrari, and M.J. Duffy, *Prognostic and predictive biomarkers in breast cancer: Past, present and future.* Semin Cancer Biol, 2018. **52**(Pt 1): p. 56-73.

65. Rakha, E.A. and A.R. Green, *Molecular classification of breast cancer: what the pathologist needs to know.* Pathology, 2017. **49**(2): p. 111-119.

66. Mark, M., F.M. Rijli, and P. Chambon, *Homeobox Genes in Embryogenesis and Pathogenesis.* Pediatric Research, 1997. **42**: p. 421.

67. Hur, H., et al., *Analysis of HOX Gene Expression Patterns in Human Breast Cancer.* Molecular Biotechnology, 2014. **56**(1): p. 64-71.

68. Gummalla, M., et al., *Hox gene regulation in the central nervous system of Drosophila.* Frontiers in Cellular Neuroscience, 2014. **8**(96).

69. Pineault, K.M. and D.M. Wellik, *Hox genes and limb musculoskeletal development.* Current osteoporosis reports, 2014. **12**(4): p. 420-427.

70. Rux, D.R. and D.M. Wellik, *Hox genes in the adult skeleton: Novel functions beyond embryonic development.* Dev Dyn, 2017. **246**(4): p. 310-317.

71. Holland, P.W.H., H.A.F. Booth, and E.A. Bruford, *Classification and nomenclature of all human homeobox genes.* BMC biology, 2007. **5**: p. 47-47.

72. Quinonez, S.C. and J.W. Innis, *Human HOX gene disorders.* Molecular Genetics and Metabolism, 2014. **111**(1): p. 4-15.

73. Procino, A., *HOX Genes and Oncogenesis*. Vol. 8. 2014. 6.

74. Bhatlekar, S., J.Z. Fields, and B.M. Boman, *Role of HOX Genes in Stem Cell Differentiation and Cancer.* Stem cells international, 2018. **2018**: p. 3569493-3569493.

75. Zandvakili, A. and B. Gebelein, *Mechanisms of Specificity for Hox Factor Activity.* Journal of developmental biology, 2016. **4**(2): p. 16.

76. Daftary, G.S. and H.S. Taylor, *Endocrine Regulation of HOX Genes.* Endocrine Reviews, 2006. **27**(4): p. 331-355.

77. Sauvegarde, C., et al., *Dynamic Pattern of HOXB9 Protein Localization during Oocyte Maturation and Early Embryonic Development in Mammals.* PLoS One, 2016. **11**(10): p. e0165898.

78. Akam, M., *The molecular basis for metameric pattern in the Drosophila embryo.* Development, 1987. **101**(1): p. 1.

79. Grier, D.G., et al., *The pathophysiology of HOX genes and their role in cancer.* J Pathol, 2005. **205**(2): p. 154-71.

80. Joo, M.K., J.J. Park, and H.J. Chun, *Impact of homeobox genes in gastrointestinal cancer.* World J Gastroenterol, 2016. **22**(37): p. 8247-8256.

81. Shah, N. and S. Sukumar, *The Hox genes and their roles in oncogenesis.* Nat Rev Cancer, 2010. **10**(5): p. 361-71.

82. Gilbert, P.M., et al., *HOXA9 regulates BRCA1 expression to modulate human breast tumor phenotype.* The Journal of clinical investigation, 2010. **120**(5): p. 1535-1550.

83. Alharbi, R.A., et al., *The role of HOX genes in normal hematopoiesis and acute leukemia.* Leukemia, 2013. **27**(5): p. 1000-8.

84. Bhatlekar, S., J.Z. Fields, and B.M. Boman, *HOX genes and their role in the development of human cancers.* J Mol Med (Berl), 2014. **92**(8): p. 811-23.

85. Lewis, M.T., *Homeobox genes in mammary gland development and neoplasia.* Breast cancer research : BCR, 2000. **2**(3): p. 158-169.

86. Bessa-Garcia S., M.A., Pereira T., Mouta J., Freitas R., *HOXB function in Breast Cancer*. 2019.

87. Jin, K., et al., *The HOXB7 protein renders breast cancer cells resistant to tamoxifen through activation of the EGFR pathway.* Proc Natl Acad Sci U S A, 2012. **109**(8): p. 2736-41.

88. Jin, K. and S. Sukumar, *A pivotal role for HOXB7 protein in endocrine resistant breast cancer.* Oncoscience, 2015. **2**(11): p. 917-9.

89. Carè, A., et al., *HOXB7: A Key Factor for Tumor-associated Angiogenic Switch.* Cancer Research, 2001. **61**(17): p. 6532.

90. Errico, M.C., et al., *The Widening Sphere of Influence of HOXB7 in Solid Tumors.* Cancer Res, 2016. **76**(10): p. 2857-62.

91. Wu, X., et al., *HOXB7, a homeodomain protein, is overexpressed in breast cancer and confers epithelial-mesenchymal transition.* Cancer Res, 2006. **66**(19): p. 9527-34.

92. Stavnes, H.T., et al., *HOXB8 expression in ovarian serous carcinoma effusions is associated with shorter survival.* Gynecol Oncol, 2013. **129**(2): p. 358-63.

93. Wang, T., et al., *HOXB8 enhances the proliferation and metastasis of colorectal cancer cells by promoting EMT via STAT3 activation.* Cancer Cell Int, 2019. **19**: p. 3.

94. Xue, M., et al., *HoxB9 promotes the migration and invasion via TGF-beta1/Smad2/Slug signaling pathway in oral squamous cell carcinoma.* Am J Transl Res, 2017. **9**(3): p. 1151-1161.

95. Zhan, J., et al., *Elevated HOXB9 expression promotes differentiation and predicts a favourable outcome in colon adenocarcinoma patients.* Br J Cancer, 2014. **111**(5): p. 883-93.

96. Zhan, J., et al., *High expression of transcriptional factor HoxB9 predicts poor prognosis in patients with lung adenocarcinoma.* Histopathology, 2015. **66**(7): p. 955-65.

97. Deb, P., et al., *Endocrine disrupting chemical, bisphenol-A, induces breast cancer associated gene HOXB9 expression in vitro and in vivo.* Gene, 2016. **590**(2): p. 234-43.

98. Kato, F., et al., *Experimental and clinicopathological analysis of HOXB9 in gastric cancer.* Oncol Lett, 2019. **17**(3): p. 3097-3102.

99. Zhang, L., et al., *HOXB9 inhibits proliferation in gastric carcinoma cells via suppression of phosphorylated-Akt and NF-kappaB-dependent Snail expression.* Dig Liver Dis, 2019. **51**(1): p. 157-165.

100. Wan, J., et al., *HOXB9 promotes endometrial cancer progression by targeting E2F3.* Cell Death Dis, 2018. **9**(5): p. 509.

101. Huang, K., et al., *Overexpression of HOXB9 promotes metastasis and indicates poor prognosis in colon cancer.* Chin J Cancer Res, 2014. **26**(1): p. 72-80.

102. Li, Y., et al., *Expression patterns of E2F transcription factors and their potential prognostic roles in breast cancer.* Oncol Lett, 2018. **15**(6): p. 9216-9230.

103. Seki, H., et al., *HOXB9 expression promoting tumor cell proliferation and angiogenesis is associated with clinical outcomes in breast cancer patients.* Ann Surg Oncol, 2012. **19**(6): p. 1831-40.

104. Hayashida, T., et al., *HOXB9, a gene overexpressed in breast cancer, promotes tumorigenicity and lung metastasis.* Proc Natl Acad Sci U S A, 2010. **107**(3): p. 1100-5.

105. Seki, H., et al., *HOXB9 Expression Promoting Tumor Cell Proliferation and Angiogenesis Is Associated with Clinical Outcomes in Breast Cancer Patients.* Annals of Surgical Oncology, 2012. **19**(6): p. 1831-1840.

106. Shrestha, B., et al., *Homeodomain-containing protein HOXB9 regulates expression of growth and angiogenic factors, facilitates tumor growth in vitro and is overexpressed in breast cancer tissue.* The FEBS journal, 2012. **279**(19): p. 3715-3726.

107. Chiba, N., et al., *Homeobox B9 induces epithelial-to-mesenchymal transition-associated radioresistance by accelerating DNA damage responses.* Proc Natl Acad Sci U S A, 2012. **109**(8): p. 2760-5.

108. Pereira, M.A., *HOXB genes function in breast cancer*, in *Molecular Biology*. 2018, Instituto de Ciências Biomédicas Abel Salazar - Universidade do Porto.

109. Ishibashi, H., et al., *Sex steroid hormone receptors in human thymoma.* J Clin Endocrinol Metab, 2003. **88**(5): p. 2309-17.

110. Zlobec, I., et al., *Selecting immunohistochemical cut-off scores for novel biomarkers of progression and survival in colorectal cancer.* Journal of clinical pathology, 2007. **60**(10): p. 1112-1116.

111. Soshnikova, N., *Hox genes regulation in vertebrates.* Developmental Dynamics, 2014. **243**(1): p. 49-58.

112. Zhussupova, A., et al., *An E2F1-HOXB9 transcriptional circuit is associated with breast cancer progression.* PLoS One, 2014. **9**(8): p. e105285.