ARNAUD DA CRUZ PAULA

CHARACTERIZATION OF DIFFERENT BREAST CANCER STEM CELL PHENOTYPES IN PROLIFERATIVE, PRE-MALIGNANT AND NEOPLASTIC LESIONS OF THE BREAST: ASSOCIATIONS WITH BREAST CANCER BEHAVIOR AND PROGRESSION

Tese de Candidatura ao grau de Doutor em Patologia e Genética Molecular submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador – Professor Doutor Carlos Alberto da Silva Lopes

Categoria – Professor Catedrático Jubilado

Afiliação – Departamento de Patologia e Imunologia Molecular, Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto

Co-Orientador – Professor Doutor Jorge Sérgio Reis-Filho

Categoria – Professor Convidado

Afiliação – Departamento de Patologia, Memorial Sloan-Kettering Cancer Center, Nova Iorque, E.U.A.

Financial support:

PhD fellowship: SFRH/BD/74307/2010

Fundação para a Ciência e Tecnologia (FCT)

DIRETIVAS LEGAIS

Ao abrigo do disposto do nº 2, alínea a) do artigo 31º do Decreto-Lei n.º 230/2009 declara-se que o autor desta dissertação contribuiu activamente na conceptualização, execução, interpretação e escrita dos seguintes manuscritos aceites, submetidos e em preparação:

A. Da Cruz Paula, O. Marques, A. Rosa, F. Faria, A. Rêma, C. Lopes, Co-expression of stem cell markers ALDH1 and CD44 in non-malignant and neoplastic lesions of the breast, Anticancer Res., 3 (2014) 1427-1434

A. Da Cruz Paula, O. Marques, R. Sampaio, A. Rosa, J. Garcia, A. Rêma, F. Faria, P. Silva, R. Vizcaíno, C. Lopes, Characterization of $CD44^+$ $ALDH1^+$ $Ki\text{-}67^-$ cells in non-malignant and neoplastic lesions of the breast, Anticancer Res., 36 (2016) 4629-4638

A. Da Cruz Paula, C. Leitão, O. Marques, A. Rosa, A. Santos, A. Rêma, F. Faria, A. Rocha, J. Luís Costa, M. Lima, C. Lopes, Characterization of $CD44^+/CD24^-/Ck^+/CD45^-$ cells in non-malignant and neoplastic lesions of the breast through flow-cytometry and massive parallel sequencing (Submitted for publication)

This thesis includes unpublished results.

AGRADECIMENTOS

Foi preciso ter percorrido um longo caminho para ter conseguido finalizar esta Tese. Durante esse percurso, tive o privilégio de ter a ajuda de inúmeras pessoas que me ensinaram muito a nível profissional e que me deram muita força a nível pessoal. A todas as pessoas que estiveram incansavelmente ao meu lado durante os últimos anos, o meu maior agradecimento.

Ao Professor Doutor Carlos Lopes, orientador deste trabalho, por me ter acompanhado e guiado durante os últimos oito anos. Muito obrigado pelas oportunidades dadas para o meu desenvolvimento profissional, pelos conselhos científicos partilhados e também pela motivação incutida e direcionada para a autoaprendizagem. Independentemente daquilo que o futuro me reserva, o Professor será sempre uma referência para mim, a todos os níveis.

Ao Professor Doutor Jorge Reis-Filho, co-orientador desta tese, pela transmissão dos seus conhecimentos científicos que em muito contribuíram para a minha aprendizagem enquanto aluno de doutoramento. Foi um enorme privilégio ter sido seu aluno no Memorial Sloan-Kettering Cancer Center e ter feito parte da sua equipa de investigação no seu laboratório. Será sempre um prazer ouvir as suas palestras e ideias que tanto alargaram a minha visão científica. Por tudo isto, o meu maior agradecimento.

À Doutora Margarida Lima pela disponibilidade e por me ter dado todas as condições para a realização da técnica de citometria de fluxo. Muito obrigado por ter sempre acreditado em mim e por me ter acompanhado e ajudado quando mais necessitei para conseguir levar esta tese até ao fim. Terei sempre uma enorme admiração pelo seu excelente trabalho, rigor e dedicação.

À Professora Doutora Fátima Gärtner pela constante disponibilidade demonstrada ao longo deste percurso e por ter sempre facilitado a realização das experiências científicas inerentes a este projecto.

To Doctor Britta Weigelt, I want to thank you for your kindness and support during the development of highly demanding scientific projects in the Laboratory of

Professor Jorge Reis-Filho. Thank you so much for the transmission of valuable knowledge and sense of responsibility.

À Doutora Catarina Leitão pela constante disponibilidade e dedicação para com este trabalho. Muito obrigado por teres sempre acreditado em mim e no projecto desenvolvido bem como por todos os teus conhecimentos científicos transmitidos. Foram muitas as adversidades que surgiram durante o desenvolvimento deste trabalho, mas graças ao esforço e à dedicação tudo se consegue. Por todo o teu esforço e dedicação, o meu maior agradecimento.

To Doctor Luciano Martelotto, thank you so much for having been my "Master" during the time we worked together, it was definitely a great pleasure for me. Thank you for everything you taught me, for the strength you gave me in the bad times and for the patience you had with me. More than a great professional, you are also a great friend.

To Doctor Salvatore Piscuoglio I wish to thank you for having helped integrate me into the research team and for everything you taught me. Thank you also for your companionship.

To all others elements of the research team from the Laboratory of Professor Jorge Reis-Filho I want to thank you all for the good work that we did together, for your support and transmission of knowledge.

À Professora Doutora Fátima Carneiro por amavelmente ter permitido a utilização do microscópio de microdissecção a laser e também pelas suas valiosas críticas referentes à análise da tripla-imunohistoquímica desenvolvida neste trabalho.

À Doutora Paula Silva pela ajuda e dedicação durante as horas intermináveis de trabalho. Muito obrigado por todos os conselhos dados durante o meu longo percurso que me ajudaram a tornar-me num aluno e numa pessoa mais responsável.

Ao Doutor José Luís Costa pelos conselhos científicos fornecidos que permitiram aumentar o meu campo de conhecimento na área da sequenciação e pela disponibilidade demonstrada relativamente ao desenvolvimento do projecto.

Ao Doutor José Ramon Vizcaíno, Director do Serviço de Anatomia Patológica do Hospital de Santo António, por todas as permissões e autorizações dadas para a obtenção de amostras valiosas e primordiais para todas as investigações que foram realizadas neste trabalho, assim como pela sua disponibilidade para as discussões científicas em torno da patologia.

Aos internos do Serviço de Anatomia Patológica do Hospital de Santo António, nomeadamente à Doutora Rita Sampaio, ao Doutor José Garcia, à Doutora Filipa Moreno e à Doutora Renata Dias pela amizade, acompanhamento e por todas as nossas enriquecedoras discussões científicas.

Aos patologistas e técnicos do Serviço de Anatomia Patológica do Hospital de Santo António pelas críticas construtivas e disponibilidade sempre demonstrada ao longo de todo este trabalho. Um profundo agradecimento também para as "Senhoras da Secretaria" do mesmo Serviço pela amabilidade e carinho evidenciado durante estes últimos anos. Foi um prazer ter feito parte desta casa.

A todos os elementos do Laboratório de Citometria do Hospital de Santo António pela amabilidade, apoio prestado e pela transmissão de conhecimentos inerentes à citometria de fluxo. Um agradecimento especial à Helena Santos que me acompanhou de um modo exemplar.

À Alexandra Rêma e Fátima Carvalho do Laboratório de Patologia Veterinária do Instituto de Ciências Biomédicas Abel Salazar por todo o esforço, dedicação e ajuda durante o desenvolvimento deste projecto e também por todos os conhecimentos imunológicos e histopatológicos transmitidos.

À Ana Rosa, pela incansável ajuda e pelo profissionalismo demonstrado na realização de importantes técnicas laboratoriais, contribuindo e muito para o desenvolvimento e finalização da Tese. Muito obrigado também pela amizade e apoio constante.

Sendo impossível agradecer individualmente a todos os meus amigos por serem numerosos, agradeço a todos aqueles que estiveram sempre ao meu lado durante esta caminhada. Um agradecimento especial à Oriana, não apenas por ser a grande amiga que é, mas também por me ter acompanhado ao longo do meu percurso académico e me ter apoiado constantemente. Também ao Rui

Rebelo, obrigado pela tua ajuda artística na apresentação de trabalhos científicos relativos a esta tese. Muito obrigado aos que me deram força para nunca desistir acreditando sempre em mim, tenho uma sorte tremenda em ter-vos a todos até porque muito daquilo que sou hoje é graças a vocês. À vila de Ponte de Lima por me ter proporcionado momentos de descontração e lazer, sobretudo em alturas de maior nervosismo e angústia. A todos vós, amigos, o meu maior agradecimento.

Por fim e o mais importante, aos meus pais e às minhas irmãs que me apoiaram incansavelmente durante toda a minha vida e que de tudo fizeram para me tornarem na pessoa que hoje sou. Depois deste árduo caminho percorrido espero que possam ficar orgulhosos de mim! Por todo o carinho demonstrado e pela força transmitida, o meu enorme agradecimento.

“If you understand failure you won't be afraid of it anymore.
You only fail when you decide to not try again, so it's entirely in your control.
Once you understand failure, it's impossible to fail”

Humans of New York, 2013

TABLE OF CONTENTS

FIGURES INDEX

Chapter 1 Figures

Chapter 3 Figures

Chapter 4 Figures

Chapter 5 Figures

Chapter 6 Figures

TABLES INDEX

Chapter 1 Tables

Chapter 3 Tables

Chapter 4 Tables

Chapter 5 Tables

ABBREVIATIONS

ABC – ATP-binding cassette
AD – Adenosis
ADH – Atypical Ductal Hyperplasia
AKT – v-Akt Murine Thymoma Viral Oncogene Homolog 1 and 2
ALDH1 – ALdehyde DeHydrogenase 1
ALH – Atypical Lobular Hyperplasia
APC – Adenomatous Poliposis Coli
APC – AlloPhyCocyanin
ASCO – American Society of Clinical Oncology
ASR – Age Standardized Rate
ATRA – All-Trans Retinoic Acid
BCSC – Breast Cancer Stem Cell
BRCA – BReast CAncer genes 1 and 2
BSA – Bovine Serum Albumin
BSC – Breast Stem Cell
CAF – Cancer Associated Fibroblast
CCND1 – CyCliN D1
CDKN2A – Cyclin-Dependent Kinase Inhibitor 2A
CHK1 – Checkpoint Homolog 1
CHPv2 – Ion Torrent Cancer Hotspot Panel v2
CGH – Comparative Genomic Hybridization
CK – Cytokeratin
CMF – Cyclophosphamide, Methotrexate, 5-Fluorouracil
CSF1 – Colony Stimulating Factor 1
CSFR1 – Colony Stimulating Factor Receptor 1
CSC – Cancer Stem Cell
CXCL7 – C-X-C chemokine Ligand 7
DAB - DiAminoBenzidine
DCIS – Ductal Carcinoma *In Situ*
DEAB – DiEthylAminoBenzaldehyde
DHP – Ductal HyperPlasia
DFS – Disease Free Survival
ECM – ExtraCellular Matrix
e.g – *exampli gratia*
EGFR – Epidermal Growth Factor Receptor
EMT – Epithelial-Mesenchymal Transition
ER – Estrogen Receptor
EU28 – European Union Member States
EZH2 – Enhancer of Zeste Homolog 2
FACS – Fluorescence-Activating Cell Sorting
FAD – Fibroadenoma

FCM – Flow CytoMetry
FEA – Flat Epithelial Atypia
FFPE – Formalin-Fixed Paraffin-Embedded
FGFR1 – Fibroblast Growth Factor Receptor 1
FITC – Fluorescein IsoThioCyanate
FMO – Fluorescence-Minus-One
FSC – Forward SCatter
GSTpi – Glutathione-S-Transferase pi
H&E – Hematoxylin and Eosin
HER2 – Epidermal Growth factor receptor 2
HIF-1 and 2 – Hipoxia Inducible Factor 1 and 2
HOXA1 – HOmeoboX A1
HR – Hormone Receptors
HRAS – Harvey Rat SArcoma viral oncogene homolog
IBC – Invasive Breast Cancer
IDC – Invasive Ductal Carcinoma
IDC-NOS – Invasive Ductal Carcinoma Not Otherwise Specified
ILC – Invasive Lobular Carcinoma
IHC - ImmunoHistoChemistry
INDEL – INsertion and DELetion
iPSC – induced Pluripotent Stem Cell
IL – Interleukin
LCIS – Lobular Carcinoma *In Situ*
LOH – Loss Of Heterozigosity
KRAS – Kirsten RAt Sarcoma viral oncogene homolog
mAb – monoclonal Antibody
MAP2K4 – Mitogen-Activated Protein Kinase kinase 4
MAP3K1 – Mitogen-Activated Protein Kinase kinase kinase 1
MET – Mesenchymal-Epithelial Transition
MGA – MicroGlandular Adenosis
MNP – Multi-Nucleotide Polymorphism
MPS – Massive Parallel Sequencing
mTOR – mammalian Target of Rapamycin
MUC4 – MUCin 4
NF1 – Neurofibromin 1
NGS – Next Generation Sequencing
NOD/SCID – Non-Obese Diabetic/Severe Combined ImmunoDeficiency
NOTCH1 – NOTCH homolog 1
NST – No Special Type
OCT – Optimum Cutting Temperature
OS – Overall Survival
PBS – Phosphate Buffered Saline
PE – PhycEerythrin
PerCP – Peridinin CloroPhyll

PI3K – PhosphoInositide 3-Kinase
PIK3CA – PhosphatIdylinositol-4,5-bisphosphate 3-Kinase Catalytic subunit Alpha
Pgp – P-glycoprotein
PR – Progesterone Receptor
PTEN – Phosphatase and TENsin homolog
RB1 – Retinoblastoma 1
RET – RET proto-oncogene
ROS – Reactive Oxygen Species
SC – Stem Cell
SD – Standard Deviation
SEM – Standard Error of the Mean
SERM – Selective Estrogen Receptors Modulator
SKT11 – Serine/Threonine Kinase 11
SMARCB1 - SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily B, Member 1
SNAIL1 – SNAIL Family Zinc Finger 1
SNP – Single-Nucleotide Polymorphism
SNV – Single Nucleotide Variant
SSC – Side SCatter
STAT1 – Signal Transducer and Activator of Transcription 1
STAT3 – Signal Transducer and Activator of Transcription 3
TDLU – Terminal Duct Lobular Unit
TGFβ2 – Transforming Growth Factor Beta 2
TMA – Tissue Microarray
TP53 – Tumor Supressor p53
UDH – Usual epithelial Ductal Hyperplasia
UN – Unknown
uPA – urokinase-type Plasminogen Activator
WHO – World Health Organization

ABSTRACT

The Cancer stem cell (CSC) model became an attractive concept to explain several poorly understood clinical phenomena due to its inherent theoretical properties. Such properties are based on the molecular features of normal stem cells (SCs). Thus, CSCs are believed to have the ability to renew themselves and to last a lifetime and to be resistant to electromagnetic and chemical insults. This resistance ability allows them to stagnate for long periods of time and consequently, to colonize other parts of the body. With this notion, a search for specific surface and intracellular biomarkers has been ongoing in recent years for the identification, isolation and characterization of CSCs in several cancers, like in breast cancer. In fact, breast cancer stem cells (BCSCs) were initially defined by the presence and absence of the cell-surface proteins CD44 and CD24, respectively. The $CD44^{+}/CD24^{-/low}$ phenotype has been demonstrated to have tumor initiating properties and has been associated with stem cell-like characteristics, enhanced invasive properties, radiation resistance and with distinct genetic profiles suggesting an association with adverse prognosis. However, and due to the high levels of heterogeneity associated with this disease, some breast tumors were shown not to harbor any $CD44^{+}/CD24^{-/low}$ breast cell. As a consequence, additional SC markers like ALDH1 have been reported.

Hence, our first goal in this study was to compare by immunohistochemistry (IHC), two of the most reliable SC markers ALDH1 and CD44 and to correlate their expression in different breast lesions. Moreover, we combined these markers with Ki-67 to evaluate quiescence and to identify, assess its distribution and estimate the mean percentages of $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ cells in non-malignant and malignant lesions.

CD44 and ALDH1 expression was commonly observed in distinct breast lesions and a higher combined expression of these markers was noticed in ductal carcinomas *in situ* (DCIS) when compared with invasive ductal carcinomas (IDCs). Such result was subsequently strengthened by the enrichment of $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ tumor cells observed in DCIS. Besides that, our results also demonstrated that this phenotype may favor distant metastasis being able of predicting overall survival (OS).

Despite these results, the $CD44^{+}/CD24^{-/low}$ phenotype continues to be by far the most studied phenotype in breast cancer. The latter was recently associated with epithelial-mesenchymal transition (EMT) markers and was demonstrated to have several signaling pathways dysregulated. Thus, massive parallel sequencing (MPS) can be regarded as an interesting approach to deepen the molecular characterization of $CD44^{+}/CD24^{-/low}$ cells since it allows the analysis of hundreds of genes in just one population of cells. In fact, standardized Next-Generation Sequencing (NGS) kits are currently available providing reliable sequencing results in routine cancer diagnostics, like the Ion Torrent Cancer Hotspot Panel v2 (CHPv2). Such assay includes 50 genes known to be involved in the pathogenesis of many human cancers.

In this way, our second goal was to characterize $CD44^{+}/CD24^{-}$ /Cytokeratin(Ck)$^{+}$/$CD45^{-}$ cells through flow cytometry (FCM) in a cohort containing non-malignant and malignant breast lesions. The CHPv2 assay was used for the identification of somatic mutations in the DNA extracted from isolated $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ cells. The expression of E-Cadherin and vimentin was also analyzed in the malignant lesions.

The percentage of $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ cells increased significantly from non-malignant to malignant lesions and was negatively correlated with tumor size. A significant association with vimentin was also observed. From the MPS analysis, the non-malignant lesion harbored only a single-nucleotide polymorphism (SNP). Mutations in the tumor suppressor p53 (*TP53*), NOTCH homolog 1 (*NOTCH1*), Phosphatase and tensin homolog (*PTEN*) and v-akt murine thymoma viral oncogene homolog 1 (*AKT1*) genes were found in isolated $CD44^{+}/CD24^{-}$ /Ck^{+}/$CD45^{-}$ cells from DCIS. Additional mutations in the colony-stimulating factor 1 receptor (*CSF1R*), ret proto-oncogene (*RET*) and *TP53* genes were also identified in IDCs.

In conclusion, $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ tumor cells may have a higher tumorigenic effect in breast cancer than $CD44^{+}/CD24^{-/low}$ tumor cells. Due to its role, ALDH1 can be determinant for the behavior of BCSCs, for their ability to resist to chemotherapeutic agents and their dissemination to other parts of the body, which can be aided by the role of CD44. Additionally, quiescence seems to be crucial for tumor progression, resistance to chemotherapeutic agents and metastatic spread of BCSCs. Further studies to infer about the tumorigenic and

metastatic ability of $CD44^{+}/ALDH1^{+/high}$ tumor cells combined with their quiescence status still have to be depicted.

The characterization of $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ cells supports the existence of a tumor initiation capability from these cells which can be strengthened by the acquisition of an EMT state. All of the mutated genes that were found in this study can play important roles for the development and transformation of $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ breast cells into a malignant phenotype, for stemness maintenance and activation of the EMT state. Functional analyses are now required to determine the tumorigenic effect of each mutation. Subsequent applications of NGS technologies are also demanded to better understand the malignant progression of breast stem cells (BSCs) and to design an effective mutational profile of these cells.

RESUMO

O modelo das células estaminais cancerosas tornou-se num conceito atractivo para explicar vários fenómenos clínicos pobremente compreendidos devido às suas propriedades teóricas inerentes. Tais propriedades baseiam-se nas características moleculares das células estaminais normais. Deste modo, acredita-se que as células estaminais cancerosas são capazes de se renovarem e de durarem uma vida inteira e de serem resistentes aos insultos químicos e electromagnéticos. Esta capacidade de resistência permite-lhes estagnar por longos períodos de tempo e consequentemente, de colonizarem outras partes do corpo. Com esta noção, tem sido feito nos últimos anos uma procura por biomarcadores de superfície e intracelulares específicos para a identificação, isolamento e caracterização das células estaminais cancerosas em vários tipos de cancros, como no cancro da mama. De facto, as células estaminais cancerosas da mama foram inicialmente definidas pela presença e ausência das proteínas de superfície celular CD44 e CD24, respectivamente. O fenótipo $CD44^{+}/CD24^{-/baixo}$ foi demonstrado ter propriedades de iniciação tumoral e tem sido associado com características típicas das células estaminais, propriedades invasoras aumentadas, resitência à radiação e com distintos perfis genéticos que sugerem uma associação com um prognóstico adverso. No entanto e devido aos elevados níveis de heterogeneidade associados com esta doença, alguns tumores da mama foram demonstrados não conter nenhuma célula $CD44^{+}/CD24^{-/baixo}$. Consequentemente, outros marcadores de células estaminais foram reportados como o ALDH1.

Deste modo, o nosso primeiro objectivo neste estudo foi comparar, por imunohistoquímica, dois dos marcadores de células estaminais mais fiáveis, o ALDH1 e o CD44 e correlacionar as suas expressões em diferentes lesões da mama. Mais ainda, combinámos estes marcadores com o Ki-67 para avaliar a quiescência e para identificar, determinar as suas distribuições e estimar as percentagens das células $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ em lesões benignas e malignas.

A expressão do CD44 e do ALDH1 foi comummente observada em diferentes lesões da mama e uma elevada expressão combinada destes marcadores foi observada em carcinomas *in situ* quando comparada com carcinomas invasores. Este resultado foi subsequentemente reforçado pelo

enriquecimento de células tumorais $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ verificado nos carcinomas *in situ*. Além disso, os nossos resultados também demonstraram que este fenótipo pode favorecer metástases à distância sendo capaz de prever a sobrevida global.

Apesar destes resultados, o fenótipo $CD44^{+}/CD24^{-/baixo}$ continua a ser de longe o fenótipo mais estudado no cancro da mama. Este último foi recentemente associado com marcadores de transição epitélio-mesenquimal e foi demonstrado ter várias vias de sinalização desreguladas. Assim sendo, a sequenciação massiva em paralelo pode ser vista como uma abordagem interessante para aprofundar a caracterização molecular deste fenótipo já que permite a análise de centenas de genes apenas numa população celular. De facto, encontram-se disponíveis kits de sequenciação de última geração fornecendo resultados de sequenciação fidedignos em rotinas de diagnóstico do cancro, como o Ion Torrent Cancer Hotspot Panel v2. Este teste inclui 50 genes conhecidos por estarem envolvidos na patogénese de vários cancros humanos.

Desta forma, o nosso segundo objectivo foi caracterizar células $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ por citometria de fluxo num grupo contendo lesões benignas e malignas da mama. O teste Ion Torrent Cancer Hotspot Panel v2 foi utilizado para a identificação de mutações somáticas no DNA extraído das células $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ isoladas. A expressão da E-Caderina e da vimentina foi também analisada nas lesões malignas.

A percentagem das células $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ aumentou significativamente a partir das lesões benignas para as malignas e foi negativamente correlacionada com o tamanho tumoral. Uma associação significativa com a vimentina foi também observada. A partir da análise da sequenciação massiva em paralelo, a lesão benigna conteve apenas um polimorfismo de nucleótido único. Foram encontradas mutações nos genes *TP53*, *NOTCH1*, *PTEN* e *AKT1* nas células $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ isoladas dos carcinomas *in situ*. Mutações adicionais nos genes *CSFR1*, *RET* e *TP53* foram também identificadas nos carcinomas invasores.

Em conclusão, as células tumorais $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ podem ter um maior efeito tumorigénico no cancro da mama do que as células tumorais $CD44^{+}/CD24^{-/baixo}$. Devido à sua função, o ALDH1 pode ser determinante para o comportamento das células estaminais cancerosas da mama, para a capacidade

destas células em resistirem aos agentes quimioterapêuticos e disseminação para outras partes do corpo, a qual pode ser auxiliada pela função do CD44. Mais ainda, o estado de quiescência parece ser crucial para a progressão tumoral, para a resistência aos agentes quimioterapêuticos e pela propagação metastática das células estaminais cancerosas da mama. Estudos posteriores têm ainda de serem descritos para inferir sobre a capacidade tumorigénica e metastástica das células tumorais $CD44^{+}/ALDH1^{+/alto}$ combinadas com o seu estado de quiescência.

A caracterização das células $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ suporta a existência de uma capacidade de iniciação tumoral por parte destas células a qual pode ser fortalecida pela aquisição de um estado de transição epitélio-mesenquimal. Todos os genes mutados que foram detectados neste estudo podem desempenhar funções cruciais para o desenvolvimento e transformação das células de mama $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ para um fenótipo maligno, para a manutenção da estaminalidade e activação do estado de transição epitélio-mesenquimal. Análises funcionais são agora necessárias para determinar os efeitos tumorais de cada mutação. Aplicações subsequentes das tecnologias de sequenciação de última geração são também necessárias para melhor compreender a progressão maligna das células estaminais da mama e para traçar um perfil mutacional efectivo destas células.

Chapter 1

General Introduction

1. Breast Anatomy

The mammary system is very different from other organ systems. From birth through puberty, pregnancy, and lactation, the breast is affected by several dramatic changes in size, shape and function [1].

The adult breast lies on the anterior chest wall between the second and sixth ribs, and from the sternal edge medially to the mid-axillary line laterally. Breast tissue also projects into the axilla as the axillary tail of Spence [2].

Anatomically, the breast lies in a space within the superficial fascia, although microscopic extensions of glandular parenchyma occasionally traverse these boundaries. Superiorly this layer is continuous with the cervical fascia and inferiorly with the superficial abdominal fascia of Cooper. Extensions of fibrous strands from the dermis into the breast form the suspensory ligaments of Cooper, which attach the skin and nipple to the breast [3].

An interesting work regarding the anatomy of the nipple suggested the existence of more than 20 lobes that are defined by the major lactiferous ducts that open on the nipple [4]. A lobe resembles to a tree, whose trunk, branches and leaves are hollow. These arborizing networks transport milk from the lobules to the nipple and are called the terminal portions of the duct system. One lobule is formed by multiple blunt-ending ducts in a cluster like the fingers of a glove. These fingers form the glandular acini of the lobule being surrounded by specialized connective tissue, histologically different from the stromal connective tissue existent in the rest of the breast. The lobule is then formed together by the glandular acini and specialized connective tissue. A terminal duct and its lobule are collectively called the terminal duct lobular unit (TDLU) which are found as immediate branches of the major ducts [1, 2] (Figure 1).

The lobular acini are invested by a loose, fibrovascular intralobular stroma with different numbers of lymphocytes, plasma cells, macrophages, and mast cells. This specialized intralobular stroma is sharply delimited from the surrounding denser, more highly colagenized, paucicellular interlobular stroma and stromal adipose tissue [5].

The number of acini per lobule and the size of mammary lobules have been found to be extremely variable. During the menstrual cycle, the morphological changes displayed by the lobules are seen in both epithelial and stromal

components [6-9]. Although the high variability of such changes that exist among lobules in the same breast or even among immediately adjacent lobules, a dominant morphologic pattern is typically current in each phase [5].

The nipple areola complex is a circular area of skin that exhibits increased pigmentation and contains numerous sensory nerve endings. The nipple is placed centrally and is prominent above the surrounding areola. Near the periphery of the areola are elevations (tubercles of Montgomery) formed by the openings of modified sebaceous gland, whose secretion protect the nipple during breastfeeding [5, 10]. A keratinizing, stratified squamous epithelium covers both the nipple and the areola and extends for a short distance into the terminal portions of the lactiferous ducts. During lactation, epithelial cells in both the terminal duct and lobule endure secretory changes. Thus, the terminal ducts are responsible for both secretion and transport of the secretions to the extra-lobular portion of the ductal system [11].

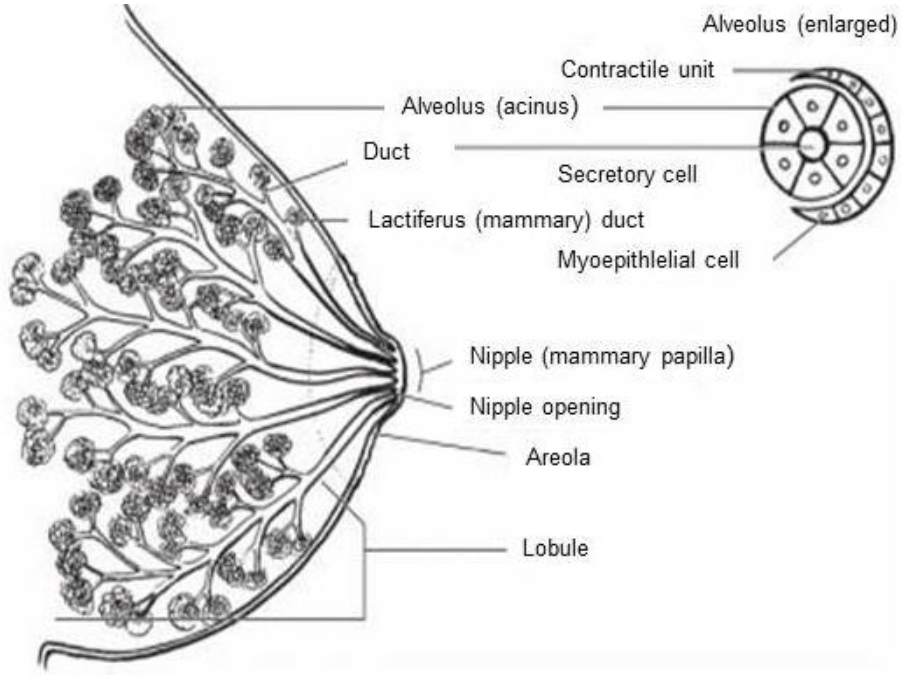

Figure 1. Schematic diagram of the breast. *Adapted from [1].*

Cells that form the duct epithelium are of two types: columnar cells, lining the lumen having cytoplasm endowed with abundant organelles involved in secretion and expressing a variety of low-molecular weight CKs including CKs 7, 8, 18 and 19 [12-16], and myoepithelial cells being distributed in a discontinuous manner in the epithelium [17]. Myoepithelial cells lay between the epithelial layer

and the basal lamina forming a network of slender processes investing the overlying epithelial cells. These cells range in appearance from barely visible, flattened cells with compressed nuclei to prominent epithelioid cells with abundant clear cytoplasm. In some cases, the myoepithelial cells have a myoid appearance featuring a spindle cell shape and dense, eosinophilic cytoplasm, reminiscent of smooth muscle cells. Immunohistochemical stains for several markers are used to discriminate these cells, including S-100 protein, actins, calponin, smooth muscle myosin heavy chain, p63, and CD10, among others [18-21]. A third cell type in normal breast tissue was also recently proposed. These cells were seen to be dispersed irregularly throughout the ductal lobular system expressing the basal cytokeratin CK5. Due to their ability to differentiate into both glandular epithelial and myoepithelial cells, such cells are believed to be progenitor cells [15].

The epithelial-stromal junction comprises an epithelial-mesenchymal layer within the duct, the basal lamina, and a surrounding zone of delimiting fibroblasts and capillaries. Elastic tissue fibers are variably present around normal ducts but in the premenopausal breast, these fibers tend to be less pronounced. Besides elastic fibers, the normal periductal stroma contains a sparse scattering of lymphocytes, plasma cells, mast cells, and histiocytes. Ochrocytes are periductal histiocytes with cytoplasmic accumulation of lipofuscin pigment. Such pigmented cells tend to be more frequent in the post-menopausal breast and are associated with inflammatory or proliferative conditions [3].

Considering that the mature breast is subjected to deep changes associated with the menstrual cycle, pregnancy, lactation and menopause, the normal microscopic anatomy of the lobules is not constant. Furthermore, and regardless of the physiologic conditions, there are variations in the functional state of individual lobules suggesting that these individual lobules or lobules in regions of the breast may have intrinsic differences in response to hormonal stimuli [22]. Due to this intrinsic dynamic ability of breast cells to be continuously influenced and remodeled, it is believed that they can be susceptible to carcinogenesis. In fact, loss of heterozygosity (LOH) is one of the genetic alterations that have been detected in histologically normal-appearing lobular epithelium [23], but other genetic changes have also been found in epithelium and myoepithelial cells. Although the frequency of these alterations has not been established yet,

increasing existing data suggest that they are detected more often in histologically normal lobules from patients with carcinoma than in breasts without carcinoma [3].

2. Breast Cancer

2.1. Epidemiology

Breast cancer is a global health problem and one of the major causes of female morbidity and mortality [24-26]. Its incidence, prevalence and the economic burden it imposes on national health services make it a major public health including both the developed and developing countries [27].

This disease affects women aged less than 45 years being more prevalent in the 45-65 years age group. Currently, breast cancer continues to be the leading cause of female death from cancer worldwide. In 2012, ~1.67 million new breast cancer diagnoses were made throughout the world and 522000 women died of breast cancer. In the same year, the incidence in Europe Union Member States was ~362000 new cases with 92000 deaths, accounting for a mean incidence rate of 66.5 and a mean mortality rate of 16.0 per 100000 women (world age-standardized rate, ASR-W) [28]. The differences in the incidence, mortality and survival rates are due to different risk factors, availability of organized screening programs and access to effective treatments [24]. Although the mortality rate is higher in less developed countries [29], the incidence of breast cancer in Western and Northern Europe is between the highest in the world [28]. For this reason, breast cancer prevention continues to be a major public health goal (Figure 2A).

Portugal applied its first region-based screening program in 1990 but nationwide screening was only attained in 2005. The Health Ministry aims at 60% coverage by the end of 2016 [30] implementing a program that offers digital mammography to women aged 45-69 years. Nonetheless, ~6000 new breast cancer cases are diagnosed annually and ~1500 women die due to this disease (Figure 2B). Indeed, the incidence rate of 67.6 cases per 100000 people is currently higher when compared with the mean established for European Union countries. Even if the mortality rates (13.1 cases per 100000 people) are still lower than the European ratio, early diagnosis of abnormalities is extremely important to better understand the risk of breast cancer progression [28].

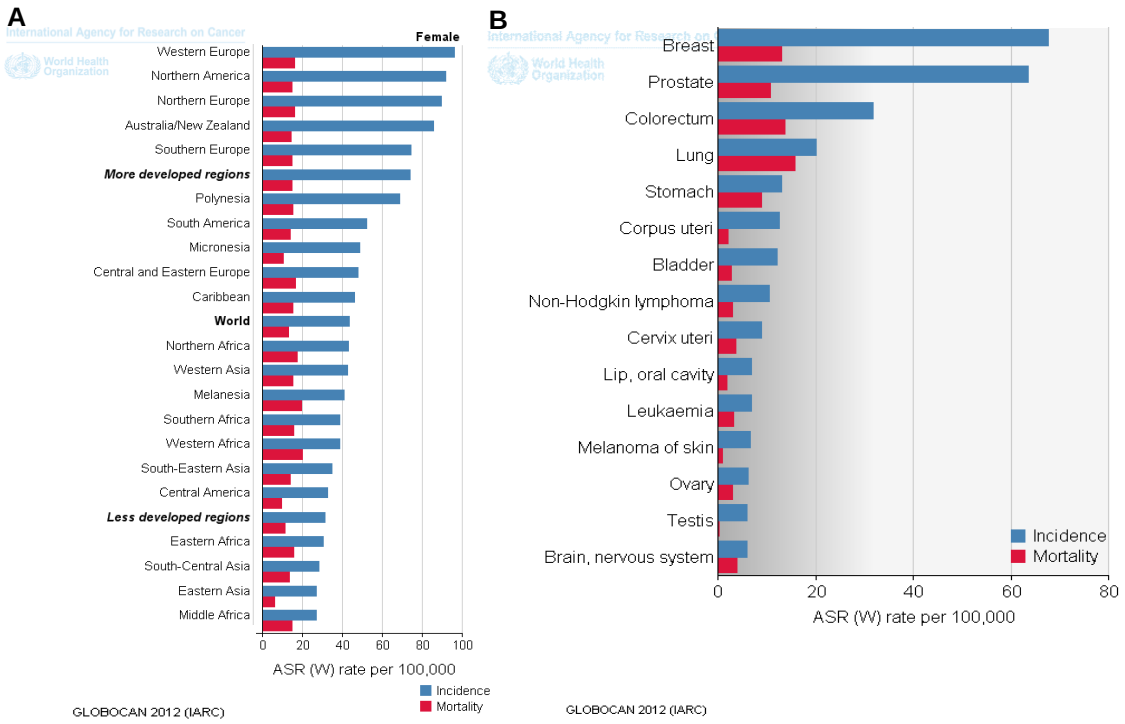

Figure 2. A. Breast cancer estimated age-standardized incidence and mortality rates in the World per 100000. **B.** Incidence and mortality rates of the most common types of cancer for both sexes in Portugal – estimates presented for the year 2012. *Adapted from [28].*

2.2. Risk Factors

In the past several years, significant improvements have occurred in our understanding of the causes and prevention of breast cancer. Factors like age, family history of breast cancer and experiences of reproductive life have long been known to be associated with breast cancer risk [31]. Age is by far the strongest risk factor for breast cancer in women. The incidence of this disease increases steeply with the greatest rate increment seen in postmenopausal women, where the risk doubles with each decade of life up to age 80. Race and ethnicity also constitute a marked risk factor, being highest in white women, following by black and Hispanic women [32]. Such dissimilarities may be explained by multifactorial inherited factors, genetic variations in the biology of the tumors or even cultural differences [33, 34]. Reproductive events, such as menarche, pregnancies and live births, lactation and menopause, all mark substantial changes that can influence breast cancer risk. Incidence may be affected by the effects of physical alterations due to reproduction and long-term modifications in hormonal exposures [32]. Increasing age at menarche is associated with decreasing breast cancer

incidence; the risk of breast cancer decreases 5% for each year increment in age at menarche [35]. Concomitantly, increasing the reproductive span with a late age at menopause increases the risk of developing breast cancer due to a greater lifetime exposure to circulating hormones [36]. Parity, especially at an early age is associated with a decreased risk of breast cancer. Women whose first birth occurred at age 35 years or older have a higher risk of developing the disease when compared to women under 18 at the time of first pregnancy [35]. In parous women, lactation further decreases the risk. Nevertheless, the overall reduction in risk varies considerably within the population studied [37].

Recently, new risk factors have emerged including low physical activity, obesity, alcohol intake, and exogenous hormone use. In fact, some of these new factors seem to be related to perturbations in circulating estrogens, which are believed to be the major cause of breast cancer [31]. Regarding the exogenous hormones, studies concerning the use of oral contraceptives and hormonal therapy for the menopause are still inconclusive. However, the last is associated with increased risk of breast cancer, especially when comparing its use during short and long periods of time [38, 39].

Women with a family history of breast cancer, particularly in a first-degree relative, have approximately double the risk of developing the disease compared to women without such a history [40]. Studies of high risk families provided evidence of an autosomal dominant inheritance of breast cancer [41, 42]. Gene linkage studies and cloning [43, 44] identified two genes, BRCA1 and BRCA2 that appear to be associated with the majority of inherited breast cancers accounting for 2-5% of all breast cancers [45]. Lifetime risk of diseases associated with BRCA1 and BRCA2 mutations ranges from 20 to 80% [46], since that these genes are tumor suppressor genes with several important cell functions like transcription, regulation of cell cycle checkpoints, genomic stability and DNA repair [47-49].

Benign breast lesions can also influence risk. Overall, women with benign breast lesions without hyperplasia have a 1.5-fold increased risk of breast cancer compared to women without any benign breast lesions. The risk between women with hyperplasia varies by whether or not atypia is present: women having atypical hyperplasia and hyperplasia without atypia have a 2.6-fold and a 1.8-fold increased risk, respectively. Women harboring fibroadenomas have an independent increased risk for breast cancer [50]. Benign breast disease may be

sensitive to the same risk factor as for invasive cancer. Hence, benign lesions should be considered as a first step in breast cancer progression [31].

2.3. Prognostic Markers and Therapeutic Strategies in Breast Cancer

Breast cancer progression evolves through a series of intermediate processes, beginning with ductal hyperproliferation, followed by subsequent evolution to carcinoma *in situ*, invasive carcinoma, and finally into metastatic disease [51]. Considering the marked heterogeneity of breast cancer, the identification of markers that can predict tumor behavior is particularly important. In fact, these markers can be seen as a useful tool for the clinical management of cancer patients, assisting in diagnostic procedures, staging, evaluation of therapeutic response, detection of recurrence, distant metastasis and prognosis [52].

Currently, the standard for assessing the prognosis of patients with newly diagnosed stage I-III breast cancer is to use an integrated prognostic model that includes information about tumor size, tumor grade, tumor proliferation index, lymph node involvement, distant metastasis, estrogen- and progesterone-receptors (ER and PR) and epidermal growth factor receptor type 2 (HER2) status. Such combination and integration converge into a single risk prediction score for prognostic stratification [53-55].

One of the most important prognostic factors is the gross measurement of tumor size. Several studies have shown that survival time decreases gradually with increasing tumor size. Patients with stage I tumors smaller than 1 cm have a 20-year disease-free survival (DFS) rate of 88%, whereas patients with tumors 1-2 cm have a 20-year DFS rate of only 68%. The incidence of lymph node metastases is 10% for patients with tumors having 1 cm or less, compared with 75-80% for patients with tumors larger than 10 cm [56].

Another important prognostic factor for breast cancer is the presence or absence of axillary lymph node metastases and the number of axillary lymph nodes with metastatic tumor. Overall survival time, recurrence and time to recurrence, metastases and treatment failure, all significantly correlate with the number of positive axillary lymph nodes [57-59]. The 10-year DFS rate is around

70-80% when axillary lymph nodes are negative for metastatic tumor. With each positive node, the OS rate gradually decreases [60].

In patients with invasive breast cancer (IBC), histologic grading has been constantly shown to predict DFS and OS and is recommended for all invasive carcinomas to provide an estimation of differentiation. A histologic grading system endorsed by World Health Organization (WHO) and based on criteria established by Bloom and Richardson and Eslton and Ellis evaluates three parameters: tubule formation, degree of nuclear pleomorphism and number of mitotic figures identified on histologic sections. Elston and Ellis found that patients with grade 1 tumors had a significantly better OS than those with grade 2 or grade 3 tumors [61].

In respect to breast cancer treatment, all cancer therapy can be divided into two basic types: local and systemic therapy. Local therapy refers to those treatments that control disease in the specific area of the tumor. Surgery to remove the breast mass is one of the first treatment methods but the extent of surgery depends on the features of a specific tumor. Even if a mastectomy may be mandatory for some patients, in several cases, only a small portion of the breast at the tumor site is removed. Surgery may be followed by radiotherapy in order to kill any cancer cells that may remain after surgery. In fact, clinical studies have shown that for the majority of breast cancers, removal of the breast lump followed by radiotherapy is as effective as surgery to remove the whole breast [62].

After the surgery and before radiation begins, it is important to discuss the role of systemic adjuvant therapy that includes endocrine therapy, chemotherapy, and/or targeted therapy. After examining the risks versus the benefits of the therapy, the use of such therapy is recommended or not. The benefit of adjuvant or endocrine therapy is in proportion to the risk of breast cancer recurrence; however, and despite the existence of prognostic factors, no parameter is completely predictive of recurrence [63, 64]. The majority of the clinicians agree that many women with node-negative breast cancer should receive adjuvant chemotherapy, especially those with larger tumors. Women with tumors smaller than 1 cm, a low-grade malignancy, positive ER/PRs, negative HER2 status and a low proliferative rate have the lowest risk of recurrence. Conversely, women with tumors larger than 2 cm, a high-grade malignancy, negative ER/PRs, positive HER2 status, triple negative receptor status (ER, PR and HER2 negative) and a high rate of proliferation are at highest risk for tumor recurrence [63]. Currently,

ER, PR and HER2 scores have been used along with the Ki-67 proliferation marker as predictive factors for identifying a high-risk phenotype and also for the selection of the most efficient therapies [65].

- **Ki-67**

Ki-67 is a nuclear protein expressed in all phases of the cell cycle other than the G0 phase and is used to assess proliferation in breast cancer. A high percentage of Ki-67-positive cells are associated with poor prognosis and the optimal cutoff between high and low risk of developing distant metastasis has been suggested to be 14% of positive cells [66].

The analysis of the value of Ki-67 as a predictive and prognostic tool is useful for neoadjuvant setting. Some studies examining complete pathological response have identified a high Ki-67 proliferation rate as a predictive factor for a higher rate of complete pathological response. It was found however, that patients in whom progression occurred had a higher proliferation rate than those who responded to neoadjuvant chemotherapy. These results suggest a nonlinear effect of Ki-67 on treatment response and probably on prognosis as well [67, 68]. Nonetheless, proliferation as assessed by the percentage of Ki-67 positive cells combined with ER, PR and HER2 scores has a strong prognostic power especially in patients with ER-positive breast cancers.

- **Hormonal Receptors**

The clinical significance of ER and PR status in IBC has been well recognized for decades and the measurement of ER and PR levels became a standard practice when evaluating patients with primary breast cancer [69-74].

ERs and PRs belong to a super family of nuclear hormone receptors that function as transcription factors when they are bound to their respective ligands [75, 76]. These receptor proteins are generally divided into six functional regions. The ligand-binding now called ERα contains an amino-terminal, hormone-independent activation function (AF-1) domain, a centrally located DNA-binding domain followed by a hinge region, the hormone-binding domain with its integral hormone-dependent region (AF-2) and a carboxyl-terminal F domain. Binding of hormone to ERα triggers the activation of the receptor leading to the

disassociation of receptor coregulatory repressor proteins and histone deacetylase complexes. Conversely, the recruitment of several coactivator protein complexes with histone acetylase activity acts in coordination to induce the transcription of estrogen-responsive genes. ERβ contains a similar structure, encoding however a smaller AF-1 domain as well as the PR-A and PR-B isoforms which also differ predominantly in their amino-terminal regions [77]. The detailed structure-function studies of AF-1 and AF-2 regions were fundamental insights into a molecular basis for hormone receptors due to the demonstration of different activities contained within these two regions [78]. It was shown that in a cell where AF-2 domain is dominant, antiestrogens such as tamoxifen act as pure antagonists. However, in a cell where AF-1 domain is dominant, tamoxifen behave as a partial agonist stimulating ERα transcriptional activity and proliferation [79, 80]. This modulation is the reason why different antiestrogens like tamoxifen or raloxifene are also termed selective ER modulators or SERMs: they can function as antagonists in the breast, but also as ERα agonists in other tissues, like the bone and cardiovascular systems [78].

Although the majority of studies suggest that like ERα, ERβ protein expression is also associated with a better outcome in untreated patients with breast cancer, the prognostic value of ERβ still need to be depicted. When tamoxifen binds to ERα, the estrogen-stimulated growth of tumor cells is inhibited leading to a significant reduction of cancer recurrences and to an increment in survival in patients with ERα positive IBCs of all stages [81]. More recently, tamoxifen has also been shown to reduce subsequent breast cancer in patients with ERα positive DCIS [82] and in patients who are cancer-free but at high risk for developing breast cancer [83]. ERα loss as a significant mechanism of acquired hormonal resistance is still an important question to be solved because even if ERα is reduced in tamoxifen-resistant tumors, the development of hormonal resistance is more frequently associated with the maintenance of ERα at the time of progression [84]. For this reason, the clinical use of other endocrine agents with distinct mechanisms of action, such as the steroidal antagonist faslodex has emerged. Currently, about two-thirds of tamoxifen-resistant breast cancer patients respond to second-line therapy with faslodex, exhibiting no ERα agonist activity in cells [85]. Additionally, a third generation of anastrozole and letrozole was shown to be effective in postmenopausal tamoxifen-resistant patients [86]. Thus,

tamoxifen cannot be assumed as a global hormonal resistance antagonist, suggesting a multiple mechanism of hormonal resistance and involving a combined pool of SERMs.

- **HER2**

The HER2 gene (also known as c-ErbB2) encodes a 185-kDa transmembrane tyrosine kinase receptor [87] belonging to the epidermal growth factor receptor (EGFR) family. Such family also includes HER1, HER3, and HER4. These receptors are sensitive to signals that stimulate cell growth being expressed in a variety of tissues of epithelial, mesenchymal, and neuronal origin [87]. The homo- and heterodimerization with one another allows the auto-phosphorylation of tyrosine molecules, and consequently, the propagation of intracellular signaling through the mitogen-activated protein kinases (HER1/HER2 heterodimers) and phosphoinositide 3-kinase (PI3K) pathways (HER2/HER3 heterodimers). HER2 is unique due to its ability to dimerize with any of the other three receptors and besides that, it does not require ligand binding for activation [88].

This oncogene is amplified in 20 to 30% of breast cancers and since that its overexpression is associated with an aggressive phenotype of tumor cells, resistance to antihormonal and cytotoxic therapies and poor OS, it is considered as a marker of poor prognosis [52]. Currently, the humanized monoclonal antibody trastuzumab targets the extracellular domains of HER2 being thus indicated for the treatment of HER2-positive breast cancers. Several studies demonstrated the high efficacy of trastuzumab through its significant inhibitory effect on tumor growth and chemotherapy sensitizer [89]. Such anti-tumor effect is believed to be conferred by the inhibition of receptor-receptor interaction, receptor decreasing by endocytosis, blockade of extracellular domain cleavage of the receptor and activation of antibody-dependent cellular cytotoxicity [90, 91]. Besides trastuzumab, other therapeutic strategies have been designed to target HER2 protein such as lapatinib, a tyrosine kinase inhibitor which showed enhanced efficacy after failure of trastuzumab therapy [92].

2.4. Molecular Subtyping of Breast Cancer

Although heterogeneity in breast cancer has long been recognized, recent microarray-based gene expression profiling analysis [93, 94] brought this concept to the vanguard of breast cancer research.

Studies published by Perou *et al* and Sorlie *et al* have demonstrated that breast carcinomas can be classified according to the similarity between their transcriptomic profiles. For instance, the authors developed an “intrinsic gene set” (i.e genetic differences between samples from different patients and samples from the same patient) and through hierarchical clustering analysis, tumors were classified into four main groups: (A) luminal, comprising ER-positive tumors and expressing genes belonging to the ER pathway with profiles corresponding to “normal luminal epithelial cells”; (B) basal-like cancers, which are hormone receptor negative/low tumors expressing genes that are usually expressed by basal/myoepithelial cells; (C) HER2 tumors with HER2 overexpression and genes belonging to the HER2 amplicon, and (D) normal breast-like group, which clusters together with normal breast samples and fibroadenomas [95, 96].

However, the molecular taxonomy of breast cancer is constantly being modified since that in each publication different intrinsic gene lists are demonstrated. As a consequence, slightly different subtypes have emerged like an interferon-rich group [97], the molecular apocrine subtype [98, 99] and the claudin-low subtype [100] (table 1).

Table 1. Characteristics of molecular subtypes of IBCs. *Adapted from [101].*

Molecular subtype	ER	PR	HER2	Basal markers	Proliferation	Histology
Luminal A	+++	+++	-	-	Low	Low grade ductal, cribiform, tubular, classic lobular
Luminal B	+	±	±	±	High	Ductal , micropapillary
HER2-enriched	+/-	+/-	+	±	High	High-grade ductal
Basal-like	-	-	-	+	High	High-grade ductal, metaplastic, medullary, adenoid-cystic
Molecular apocrine	-	-	±	±	High	Apocrine, ductal, lobular
Claudin-low	-	-	-	±	High	High-grade ductal, metaplastic

+++ High expression, + moderate expression, ± variable expression, - no expression

ER-positive expression markedly defines the luminal molecular subtype, which is also characterized by the relatively high expression of several genes expressed by breast luminal cells like CK8/18, GATA3 and estrogen-regulated genes. Luminal tumors are divided into luminal A and luminal B tumors due to their different ERα expression levels, proliferation rates assessed by Ki-67 and clinical outcomes [102]. However, patients with luminal tumors have better outcomes and a broader range of treatment options than patients with other types of breast cancer [103].

The overexpression of the HER2 protein on the cell membranes is due to the genomic amplification of the 17q22.24 region, which constitute the HER2-enriched molecular subtype. HER2-enriched tumors can be ER-positive and/or PR-positive and have a more aggressive phenotype than HER2 normal tumors. Still, overexpression of HER2 makes the majority of these tumors highly responsive to HER2 inhibitors such as trastuzumab [104]. Intrinsically, ER-positive, HER2-enriched tumors have a poorer response to systemic therapy than ER-negative, HER2-enriched tumors [103].

Basal-like breast cancers have the worst clinical outcome when compared with all others breast cancer groups. They are characterized by an aggressive phenotype with high histologic grade, high proliferation levels pushing borders of

invasion and large areas of tumor necrosis, representing 15-20% of breast cancers [105]. Such tumors express high levels of basal-markers like CKs 5, 14 and 17 but do not express ERα, PR or HER2. Consequently, anti-hormonal approaches for these tumors are unfeasible and also, the poor genetic knowledge regarding the transformation and progression of this tumor subtype makes it difficult to target with current therapies [106].

Accounting for 8-14% of breast cancers, the molecular apocrine subgroup was shown to be distinct from luminal and basal-like breast cancers. It is characterized by increase androgen signaling and is frequently associated with apocrine histological features. This molecular phenotype is thought to be associated with poor long-term survival and has specific surrogate immunohistochemical markers, including androgen receptor- and gamma-glutamyltransferase 1 [103, 107].

Claudin-low tumors account for approximately 5% of IBCs and are characterized by low or absent expression of EMT markers, presence of immune response genes and cancer stem cell-like features [108]. Clinically, the majority of claudin-low tumors carry poor prognosis and are ER-negative, PR-negative and HER2-negative (triple-negative) tumors that have a higher frequency displaying metaplastic and medullary differentiation [109].

It should be noted however, that this molecular taxonomy, as referred above, remains a working model in progress and not a definitive classification system, as additional molecular subtypes have been and may be identified in the future.

3. Breast Tumorigenesis

As previously discussed, the normal microscopic anatomy of the lobules in a mature breast is not constant and such fluctuation strongly influence how breast cells respond to a certain hormonal stimuli [22]. Such influence can render these cells susceptible to carcinogenesis. Considering that high responses to hormonal stimulus occur at the level of TDLUs, many of the known epithelial benign and malignant lesions arise from these terminal units.

3.1. Benign Lesions

Benign breast disease can be classified into non-proliferative, proliferative, atypical hyperplastic and miscellaneous lesions. Such categories are defined depending on the degree of atypia and cellular proliferation of a lesion as well as its susceptibility for developing into breast cancer [110]. Although the pathway from normal terminal ductal lobule to premalignant breast lesions, noninvasive and invasive cancer has been well defined morphologically, the major cause of benign breast diseases remains unknown [111]. Non-proliferative lesions are common and almost never increase the risk of developing breast cancer, while proliferative lesions without atypia are the most common type of benign breast lesions being associated with a small increase in breast cancer risk. Conversely, atypical hyperplasia is associated with a moderate to high risk to evolve into breast cancer [110].

The frequency of benign lesions is particularly variable with age. Fibroadenoma is frequent in younger patients while other localized benign and cysts occur most frequently in women between the ages of 30 and 50 [112]. The cumulative incidence of such diseases is approximately 10–20%, although autopsy studies have reported a much higher prevalence, at approximately 50% [113]. Whether some of the conditions of benign lesions (such as proliferative disease with atypia) are direct precursors to premalignant and invasive malignancy, as hypothesized by the modified Wellings Jensen model remains unclear [114]. In this way, investigation of management strategies is still needed to infer about which patients could potentially benefit from closer adherence to

existing screening recommendations, from additional screening modalities or even from chemoprevention [115].

3.2. *In situ* Breast Carcinomas

In situ breast carcinomas are a group of malignant lesions that can be confined inside the ducts (DCIS) or lobules (lobular carcinoma *in situ* or LCIS) of the breast. Such classification was previously based on the similarity of the involved spaces to normal ducts or lobules. Nonetheless, it is now well established that different patterns of *in situ* growths mainly reflect differences in tumor cell biology rather than being associated to the site or cell of origin. Presently, “lobular” refers to carcinomas of a specific type and “ductal” is more generally used for adenocarcinomas that have no other designation [116]. While LCIS is a rare lesion, DCIS currently account for 20-25% of all newly diagnosed breast cancer [117]. DCIS is defined as a premalignant proliferation of neoplastic epithelial cells contained within the lumen of mammary ducts. These carcinomas are lined by a layer of semi-continuous myoepithelial cells and surrounded by an intact basement membrane [118].

The risk of DCIS is rare in women younger than 30 years, is low in women younger than 40 years, but increases steadily from age 40 to 50 years. After that, the risk increases much more slowly and plateaus after 60 years of age. As a generalized risk factor, DCIS has an approximately 8- to 10-fold increased risk for subsequent invasive cancer [119].

There is no universally accepted histopathological classification for DCIS, but the majority of pathologists traditionally divide these carcinomas into five major architectural subtypes (papillary, micro-papillary, cribiform, solid and comedo), often discriminating the first four (non-comedo), from comedo [120-122]. The concept of nuclear grading has also assumed greater importance and significance in classification of DCIS lesions (low versus intermediate versus high). Actually, nuclear grade is considered to be a better biological predictor than architecture and has emerged as key histopathological factor for identifying aggressive behavior [123-125]. Comedo DCIS for example, is frequently associated with a higher nuclear grade aneuploidy, a higher proliferation rate, HER2 amplification or

protein overexpression, and clinically more aggressive behavior in opposition to non-comedo lesions [120-122, 126-128].

There are strong correlations between histological differentiation in DCIS and prognostic biomarkers in breast cancer. For example, almost all well differentiated or low-grade DCIS express high levels of ERα and PR in nearly all cells [129-131]. In the majority of poorly differentiated lesions, the proportion of cases expressing these receptors gradually decays to about 20% as well as the average proportion of positive cells [132]. Conversely, overexpression and amplification of HER2 and inactivating mutations of p53 are rare (5-10%) in well differentiated DCIS, but in the most poorly differentiated lesions, such alterations gradually increase to about 60% [130, 133-136].

Moreover and from lowest to highest grade, the average proliferation gradually increases from <5% to nearly 40% respectively and apoptosis varies in the same direction from <1% to over 5% [120, 137, 138].

Although there have been several efforts towards an improvement of clinical or molecular tests [139] to predict which patients are likely to develop invasive disease following a diagnosis of DCIS, currently, no test is clinically useful to identify this population. As a consequence, the vast majority of patients are still subjected to surgical treatment (breast reduction/total mastectomy) followed by radiation and/or prophylactic systemic therapies (e.g. tamoxifen) [140].

3.3. Invasive Breast Carcinoma

By definition, invasive breast tumors invade adjacent benign breast parenchyma and are capable of metastasis to distant sites. Classification of IBC falls into two broad subtypes: ductal and lobular carcinoma. Beyond these subtypes, several histologic patterns or phenotypes have a tendency to behave similarly and for that reason, such patterns are of great importance to clinicians [141].

Invasive ductal carcinoma not otherwise specified (IDC-NOS), comprises the largest subgroup of IBCs, accounting for 40-75% of all invasive breast tumors. The histological appearance of IDC is variable between individual tumors since that tumor cells may infiltrate as well-formed tubules, cords, clusters, or they may be arranged in solid sheets [141].

In IBCs, a strong correlation is markedly seen between tumor size and the incidence of axillary lymph node metastases and survival rates [61, 142, 143]. Characterization of such IBCs is best defined by grade and phenotype and by the most common immunohistochemical profile that is, a tumor positive for ER (60-70% of cases) and negative for HER2 (75-85% of cases) [141].

The majority of the special types of IBCs display relatively good prognosis and have an ER-positive and HER2-negative phenotype including ILC and tubular, cribiform, muscinous and papillary variants of IDC. The identification of such types of IBC is of great importance and in conjunction with other major prognostic factors, the prediction of likely behavior and response to treatment can be significantly improved [144]. Therefore and upon the retrieve of all relevant information together with patient age and overall health status, a treatment course can then be discussed. Adjuvant versus neoadjuvant therapy and breast conservation versus mastectomy remain the major decisions at the beginning of breast cancer treatment. After the appropriate treatment, almost all breast cancer patients are clinically free of disease. However, some of these initially cured patients will have a systemic treatment failure over the course of 5–10 years. The rates of such treatment failure vary between 5% up to 50%, depending on prognostic factors [145].

In patients treated with adjuvant chemotherapy, several prospective studies showed a measurable decrease in long-term systemic failure and mortality between 10% and 20%. Nonetheless and until now, it has been impossible to precisely determine which patient will benefit from this treatment culminating in an overtreatment of thousands of patients worldwide. Such limitation leads to an interesting pursuit for genetic tumor profiles benefiting from chemotherapy and profiles already cured by local therapy [145].

Several chemotherapeutic procedures have been developed over the years starting with the cyclophosphamide, methotrexate, and 5-fluorouracil (CMF) regimen [146]. The introduction of targeted therapy as another major oncologic breakthrough specific for the treatment of breast cancer allowed the use of antiestrogens such as tamoxifen and fulvestrant and, later, aromatase inhibitors. Indeed, adjuvant treatment with antihormonal medication has proven to be successful for more than 60% of patients (which are positive for ER and PR) [147]. An additional targeted therapy was introduced for those 15% of breast cancer

patients positive for HER2 which in fact turned to be extremely effective. With the implementation of trastuzumab, other substances such as lapatinib, pertuzumab, and trastuzumab-emtansine were shown to be also effective in treatment of HER2-positive breast cancers [145].

In advanced health systems, 80−85% of breast cancer patients will be cured. Due to long-term relapse, follow-up became a major issue since that patients with previous breast cancer are at high risk of recurrences. At the same time, breast cancer survivors are breast cancer advocates since they show how this increasingly common disease can be cured. The decline in breast cancer mortality has been attributed to the combination of early detection with screening programs as well as the existence of more effective adjuvant systemic therapy. Improvements in our understanding of the molecular biology of breast cancer progression allowed the discovery of novel pathway-specific targeted therapeutics. Indeed, the continuous development of such effective therapeutics is currently demanding for a molecular-based, ‘patient-tailored’ treatment planning through continuous studies of the molecular pathology of human breast cancer progression [148].

4. Breast Cancer Progression

4.1. Theories of Breast cancer Progression

The first classifications for pre-invasive lesions were based on the site or cell of origin giving rise to different names of *in situ* proliferation like DCIS or LCIS. However, the demonstrations done by Wellings *et al* regarding the origin of such pre-invasive lesions in the TDLUs constituted a paradigm shift of the histogenetic implications of the ductal and lobular terminology, marking the beginning of epidemiological and morphological observations [111, 149, 155]. Such observations led to the formulation of various linear models of breast cancer initiation, transformation and progression (Figure 3).

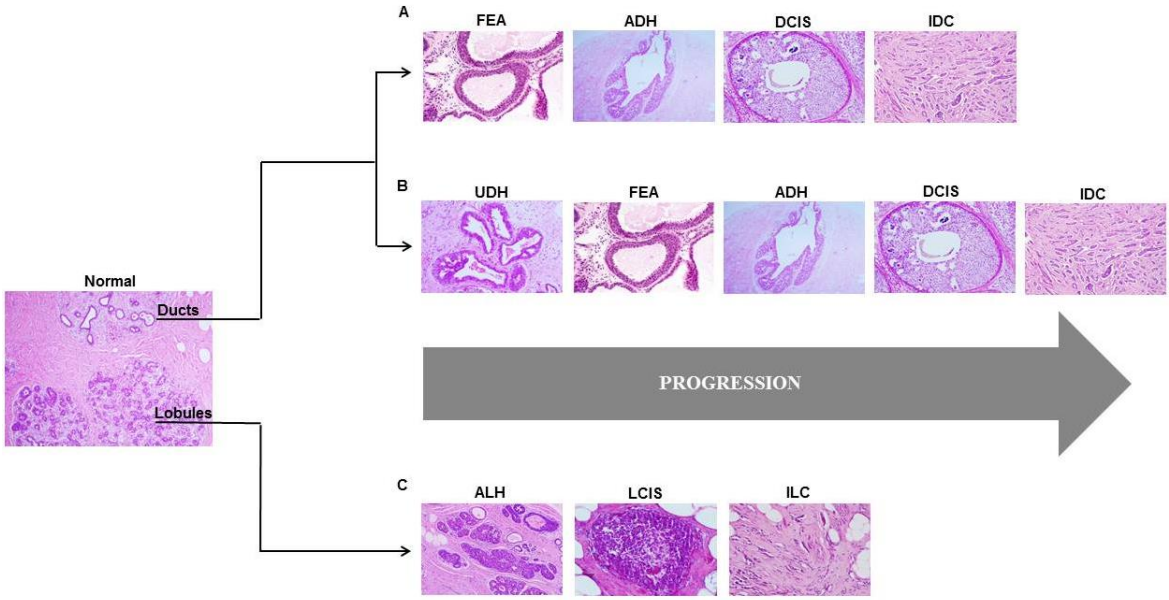

Figure 3. Historical model of breast cancer development and evolution based on morphological features and epidemiological studies. **A.** Classic model of breast cancer progression of the ductal type; **B.** Alternative model of breast ductal cancer progression; **C.** Model of lobular neoplasia. Abbreviations: ADH, atypical ductal hyperplasia; ALH, atypical lobular hyperplasia; DCIS, ductal carcinoma *in situ*; LCIS, lobular carcinoma *in situ*; FEA, flat epithelial atypia; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; UDH, usual epithelial ductal hyperplasia. *Adapted from [118, 148].*

Indeed, the continuous work of Wellings and colleagues culminated in a breast cancer progression model, where for the ductal subtype, flat epithelial atypia (FEA), atypical ductal hyperplasia (ADH) and DCIS were the non-obligate

precursors of invasive and metastatic ductal carcinoma (Figure 3A) [149-153]. A second “ductal subtype” model proposed by Page and Dupont through epidemiological studies postulated that usual epithelial ductal hyperplasia (UDH) was an intermediate stage of progression between FEA and DCIS (Figure 3B) [154, 155]. This second model was posteriorly disbelieved due to immunohistochemical and molecular biological evidences against UDH as an intermediate stage of progression [156-158]. Regarding the lobular subtype, the progression scheme recognizes atypical lobular hyperplasia (ALH) and LCIS as the non-obligate precursor lesions of ILC [159, 160] (Figure 3C).

With the continuous improvements of genetic-based technologies, chromosomal (CGH) and microarray-based comparative genomic hybridization (aCGH) and microarray-based expression profiling studies [161] have deeply changed the perspective of breast cancer progression: low-grade IDCs were shown to display fewer overall chromosomal aberrations as compared with high-grade IDCs [162-164]. In fact, recurrent chromosomal loss of 16q and gains of 1q,16p and 8q observed in low-grade tumors IDCs and the recurrent losses of 8p, 11q, 13q, 1p and 18q and gains of 8q, 17q, 20q and 16p observed in high-grade tumors suggested a different breast cancer progression model than previously proposed [163, 164]. The absence of 16q deletions in the majority of high-grade breast cancers and expression profiling analysis of matched *in situ* and invasive breast lesions led to the assumption that lesions would cluster according to their histological grade rather than stage of progression [165, 166]. In this way, a low- and high-grade multistep model of breast cancer progression based on morphological, immunophenotipical and molecular features was proposed (Figure 4). The low-grade group comprises lesions that present molecular characteristics such as expression of hormone receptors, lack of HER2 overexpression and expression of basal markers, containing simple, diploid/near diploid karyotypes with recurrent changes like the deletion of 16q (>80%) and gains of 1q (>75%) and 16p (>50%) [118].

The high-grade group comprises microglandular adenosis (MGA) lesions, high-grade DCIS and high-grade invasive carcinomas. This group, contrarily to the low-grade group is more heterogeneous comprehending lesions classified by microarray expression profiling as luminal B, HER2-enriched, basal-like, claudin-low and molecular apocrine [93, 167]. At the molecular level, high-grade

carcinomas display losses of 1p (60%), 8p (60%) and 17p (60%) and gains of 1q (60%) and 8q (75%). Amplifications are commonly observed in the high-grade lesions, occurring most frequently at 1q, 8q, 17q and 20q chromosome arms [168]. As described above, the most notorious difference between the two groups rely on the deletion of 16q chromosome arm, which in the high-grade group is seen in less than 30% of cases [151, 163, 164, 169, 170].

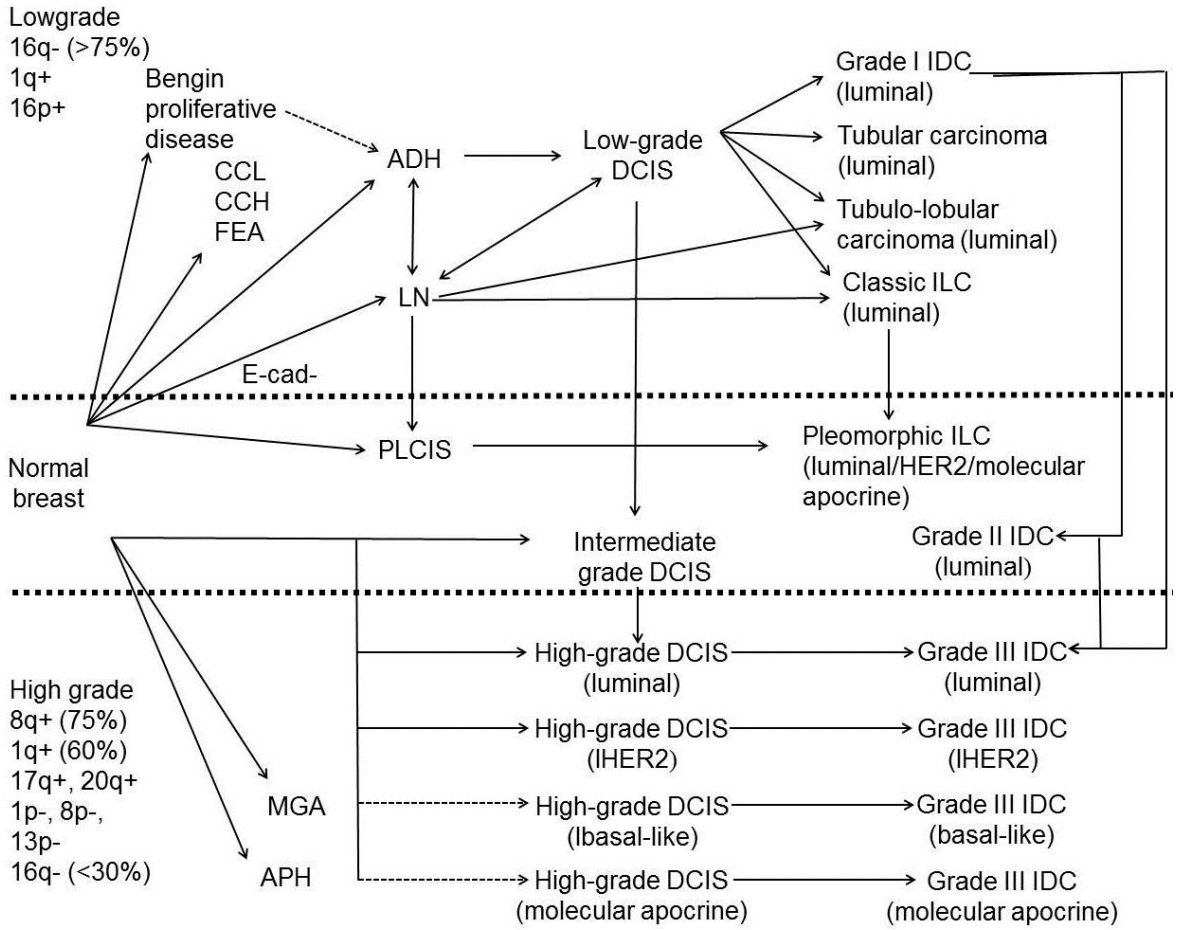

Figure 4. Low- and high-grade multistep model of breast cancer progression based on morphological, immunophenotypical and molecular features. Solid arrows represent links between morphological entities demonstrated by morphological and/or molecular data. Dotted arrows represent hypothetical links still needed to be demonstrated. Abbreviations: ADH, atypical ductal hyperplasia; APH, atypical apocrine hyperplasia; CCH, columnar cell hyperplasia; CCL, columnar cell lesion; DCIS, ductal carcinoma *in situ*; E-cad, E-cadherin; FEA, flat epithelial atypia, IDC: invasive ductal carcinoma; ILC, invasive lobular carcinoma; LN, lobular neoplasia; MGA: microglandular adenosis; PLCIS, pleomorphic lobular carcinoma *in situ*. *Adapted from [118].*

Even with this well characterized breast cancer progression model supported and confirmed by genome-wide profiling studies, recent outcomes

suggested that breast cancers can be also clustered in two main groups, depending on the expression of ER and ER-regulated genes. Such clustering highlights the importance of the ER pathway as having a determinant role in breast cancer progression, supported by the associations seen between the patterns and regions of genomic amplification with the ER status of a given lesion [171, 172]. Nonetheless, even with the notion that ER-positive and ER-negative breast cancers are fundamentally different diseases, marked evidences strongly support that genetic instability displayed by a given lesion is related to its histological grade [173, 174].

4.2. Transition from DCIS to IBC

For several decades it has been accepted that DCIS constitutes a non-obligate precursor of IBC. Since the demonstration of the breast cancer progression model by Wellings and Jensen that a plethora of succeeding work aimed to characterize DCIS and IBCs at the molecular level, which have shown their genetic similarity and likely common origin [164, 170, 175, 176]. Further studies supported the hypothesis that DCIS is a precursor of IBC, mainly due to the clinical observation that these lesions affect the same anatomical site, representing thus progressive stages of an evolutionary scale. Indeed, and at the time of diagnosis, DCIS is frequently found adjacent to the majority of IBCs so such coexistence would support the notion of DCIS as the precursor lesion [177, 178]. Nonetheless, this coexistence varies according to the subtype of breast cancer [179].

Regarding the different molecular subtypes, DCIS can also be classified like IBCs are, through the expression patterns of ER, PR, HER2, EGFR and CKs 5/6 [180-185]. Interestingly, associations of *in situ* and invasive components were shown to not always display a similar immunophenotype: in some HER2-positive tumors, HER2 amplification was found to be present in the DCIS but not in the associated IBC and even within the DCIS component, an evident heterogeneity of HER2 overexpression was seen [175, 184-187]. Explanations for such data rely on two different interpretations: loss of HER2 amplification during progression to IBC or; a clone from the DCIS component not harboring HER2 amplification gave rise to the IBC component [140]. The recent evidences for a convergent phenotypic

evolution in tumor progression and metastasis can be applied to explain the progression from DCIS to IBC, which actually can be regarded as a possible explanation for the negative results in the genomic and transcriptomic comparisons between these two lesions [188] (Figure 5A).

The theories of progression from DCIS to IDC can be divided in two major hypotheses: the first one is based on the occurrence of genetic aberrations in neoplastic cells as the cause of invasiveness, and the second one suggests that tumor progression is not due to additional genetic changes within the lesion but rather due to the microenvironment or tumor stroma. Gene expression profiling evidences supported this last idea through the demonstration of substantial changes in several cell types of the tumor microenvironment (e.g, fibroblasts, myoepithelial cells and leukocytes) during progression from DCIS to IBC [140, 189-191]. However, with the lack of a clear cause for this difference in gene expression and with the expected absence of clonal genetic aberrations in these cell types surrounding DCIS and IBC, epigenetic alterations in the stroma emerged as a complementary explanation for tumor progression [192]. One of the major histological characteristic of DCIS lesion is the presence of an outer layer of myoepithelial cells and an intact basement membrane. Looking at the theories of tumor progression, it was obviously acceptable to assume that this layer of myoepithelial cells were responsible to exert tumor suppressive effects on the *in situ* lesion, but the loss of such ability would in turn trigger invasion through the release of tumor cells [193-196]. Indeed, myoepithelial cells were shown to secrete several extra-cellular matrix (ECM) components and protease inhibitors such as Maspin, further proposed to inhibit the invasive capacity of DCIS in a paracrine manner [193, 197, 198]. Epigenetic changes have also emerged as an explanation for the lack of significant differences detected between DCIS and IBC at DNA sequence level. However, even if DNA methylation was seen to be higher in DCIS lesions when compared with the normal breast epithelium, similar levels of promoter hypermethylation in DCIS and IBC were detected in the majority of studies done so far [199-201]. Consequently, such findings explain early events in breast carcinogenesis rather than later events for the transition to an invasive disease. Alternative epigenetic changes other than methylation were proved to be more important in the progression from *in situ* to IBCs. EMT for example, regulated by global changes in histone modifications that mark heterochromatin and

euchromatin was reported to be associated with progression from DCIS to IBC [202, 203]. As a consequence, additional epigenetic studies still need to be done concerning the progression of breast tumors.

Another concern in breast cancer progression is to determine both intra-tumor genetic heterogeneity in both components of DCIS and IBC and also the genetic differences acquired during progression to IBC. Recent studies have shown a wide genetic heterogeneity regarding the number of DCIS tumor cells harboring amplification of specific loci. Interestingly, these tumor cells also presented some evidence suggesting a clonal selection during the progression to an invasive state. In other words, the transition from DCIS to IBC can be regarded as an evolutionary bottleneck, following a Darwinian evolution [204-207] (Figure 5B). Consequently, the employment of MPS and single cell analyses is an important next step to discern the contribution of genomic alterations and Darwinian evolution to the transition from DCIS to IBC.

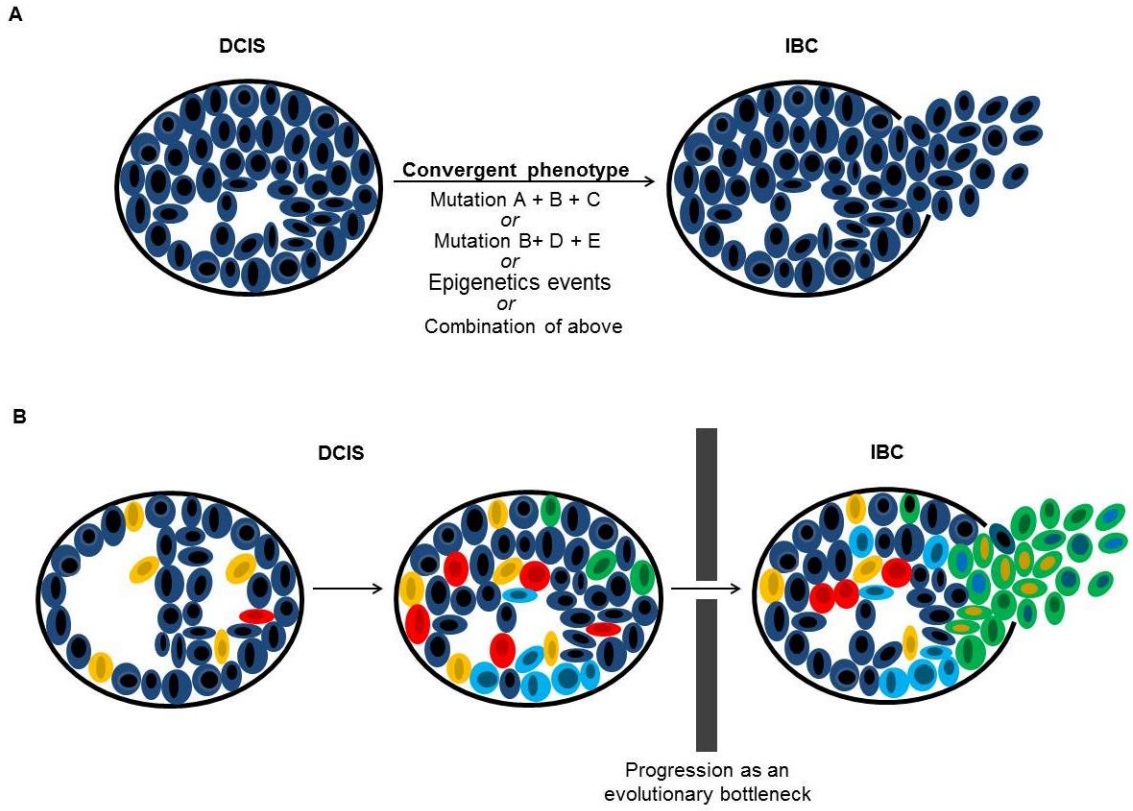

Figure 5. Hypothetical models of progression from *in situ* to invasive breast cancer. **A.** Progression from DCIS to IDC as a convergent phenotype. **B.** Progression from DCIS to IDC as an evolutionary bottleneck. Abbreviations: DCIS, ductal carcinoma *in situ*; IBC, invasive breast carcinoma. *Adapted from [140].*

4.3. Breast Cancer Progression and Next-Generation Sequencing

In recent years, whole-genome sequencing has become a powerful tool to dissect the evolutionary history of a cancer. Analogous to species comparisons, an evolutionary lineage three based on somatic genomes from a single individual can be delineated, giving essential insights on somatic evolution studies, even with the low rate of mutations in normal tissues [208]. With such improvements, it is now possible to detect the genomic changes during tumor evolution by comparing the genome of a cancer to clinically precursor lesions. Considering the frequent associations of early neoplastic breast lesions with advanced breast cancers, this type of cancer became the most suitable for whole-genomic sequencing studies providing windows into the earliest stages of tumor evolution [118, 148, 160, 179].

With the introduction of NGS by sequencing both the tumor and the germline DNA, somatic genetic changes can be classified in two major ways: driver or passenger mutations. By definition, driver mutations are those that give survival advantages to tumor cells due to the tumor microenvironment changes contributing thus for tumor development, while passenger mutations are the product of the genomic instability of the tumor [208, 209]. However, the distinction between driver and passenger mutations can change during the carcinogenesis: looking at the Darwinian evolution, a passenger mutation can become a driver mutation after therapy or changes in nutritional or immune status of a tumor leading to resistant clones that will prevail and progress along the tumorigenesis.

Another classification of mutations is also related to the type of genetic alteration: nucleotide substitution mutations, small insertions and deletions (also called “indels”), copy number gains and losses, chromosomal rearrangements as well as nucleic acids from foreign origin [210]. Analysis of the molecular features of early stage breast cancer using NGS has led to a genomic landscape portrait of this disease: *TP53* and *PIK3CA* mutations were demonstrated to be the most frequent genomic alterations in all breast cancer subtypes (28% for both genes), while amplifications in ERBB2, FGFR1 and CCND1 were observed in 10-20%. Clinical relevance of *PTEN* mutations and deletions as well as *AKT1*, *RB1*, *BRCA1* or *BRCA2* mutations was reported and subsequent sequencing analysis of mutations in other genes like *KRAS*, *APC*, *NF1*, *SKT11*, *MAP2K4*, *MAP3K1* and *AKT2* were also shown to have some clinical relevance [211]. Moreover, and

within the breast cancer subtypes, a marked heterogeneity regarding the molecular alterations was also reported in triple-negative breast cancers [212].

At the RNA level, a study involving the comparison of RNA sequencing with the genomes/exomes has revealed that only 36% of validated somatic single nucleotide variants (SNVs) were observed in a transcriptome sequence, which has questioned the use of NGS alone to identify potential drivers of breast cancer [212]. Regarding the proteomic analyses in early stage breast cancer, three pathways were reported to be predominantly activated in all subtypes: PI3K/AKT/mTOR, p53 and CCND1/CDK4/Rb [211].

In the near future, NGS will allow the development of catalogues of cancer-related genes as well as the identification of driver mutations that are responsible for tumor development. The characterization of the genomic landscape of tumors and of the activated protein networks will guide combination therapies to optimize therapeutic effects. In fact, such technologies are being directed for a personalized medicine improving considerably the management of patients [213].

5. Theories of cancer evolution

Breast cancer heterogeneity is now largely established and categorized by tumor segregation into distinct molecular subtypes defined by gene expression profiles that correlate with clinical behavior [214, 215]. Despite such improvements, the existence of tumor heterogeneity still has to be fully elucidated.

The two major theories that explain the existence and maintenance of tumor heterogeneity are the CSC hypothesis and the clonal evolution model. These two models, initially thought to be mutually exclusive, are now currently believed to be the complement of each other [216, 217]. Actually, both hypotheses support the notion that tumors arise from single cells affected by several molecular alterations and developing unlimited proliferative potential. Moreover, both models also rely on the microenvironmental features that shape the composition of a cancer; nonetheless, these two concepts have important differences (Table 2). The clonal evolution model is not based on a hierarchical model but explains intra-tumor diversity by natural selection instead, where tumor progression and resistance to therapy follow the Darwinian guidelines. Consequently, the emergence of resistant clones that progress along the tumorigenesis depend on the genetic instability and microenvironment and/or different selective pressures. In turn, the CSC hypothesis assumes the existence of a hierarchical organization of cancer cells where only a small fraction of cells are able to drive tumor progression and are inherently therapy-resistant [206].

Table 2. Tenets of the clonal evolution and cancer stem cell models. *Adapted from [218].*

	Clonal evolution model	Cancer stem cell model
Tumorigenic cells	Any cell	CSCs
Tumor cells organization	Stochastic	Hierarchical
Capacity of self-renewal with asymmetric divisions	Not applicable	CSCs have the ability to self-renew indefinitely while terminally differentiated cells have a limited proliferative propensity
Progression	Driven by the fittest clone under an assemblage of selective pressures	Driven by CSCs which correspond to a small subpopulation of the tumor bulk
Source of heterogeneity	Genetic and phenotypic heterogeneity	Initially believed as largely phenotypic, although recent studies have demonstrated that CSCs may be genetically heterogeneous within a tumor
Source of resistance to therapy	Selection of resistant subclones containing specific genetic or epigenetic aberrations	CSCs

Abbreviations: CSCs, cancer stem cells.

5.1. Clonal Evolution Model

The clonal evolution model assumes that during tumorigenesis and tumor evolution, cancer cells display different levels of genetic instability, harboring additional genetic aberrations. Such alterations lead to the development of a cancer cell population branched in different subpopulations, where each subpopulation acquires specific genetic aberrations. Thus, the most frequent clone in the tumor cell population is assessed by the tumor features and by the assemblage of selective pressures it is subjected to. Looking at the type and effect of each mutation that can occur along time, some of these mutations can confer a biological advantage for a certain cancer cell population leading consequently to clonal expansion [216, 218]. As denoted previously, clonal expansion may be also triggered by different epigenetic mechanisms that can increase the malignant potential of a given clone [219].

As also depicted above, this model is mostly related with the effect of a certain mutation that can drive progression, like driver mutations, which can arise from passenger mutations, depending on the different selective pressures continuously existent in the tumor microenvironment [208, 209].

The mutational rate during tumor progression can vary, increasing thus the heterogeneity within the tumor. As a consequence, tumors having a higher genetic complexity have a greater diversity of genetic aberrations that can be subjected to selective pressures. Thus, the chances of having a clone that can prevail under a new set of selective pressures are higher than in tumors where intra-tumor heterogeneity is lower [216].

5.2. Cancer Stem Cell Model

According to the CSC model, heterogeneity may occur as a result of cancer being hierarchically organized with a particular subset of cancer cells called CSCs at the apex, having the ability of stemness [220, 221]. Such cells can theoretically explain certain phenomena of cancer such as resistance to chemo-radiation therapy, tumor relapse and metastasis [222]. Due to their different capabilities (self-renewal, multi-lineage differentiation through symmetric and asymmetric divisions), CSCs can be fundamentally important in predicting the biological aggressiveness of a cancer since it was believed that progenitor and differentiated cancer cells lose the stemness ability not being able to contribute to tumor aggressiveness [223]. The identification of the population able to form tumors in a relatively permissive environment is mostly based on cell-surface markers, obtained from the analysis of normal SCs in the tissue of origin. In breast cancer, as few as 100 $CD44^{+}/CD24^{-/low}/Lineage^{-}$ breast cancer cells were demonstrated to efficiently form tumors when injected into mice [224], whereas the efficiency of cells with other phenotypes was inexistent. Such demonstration has highlighted the tenets of the CSC model by the existence of stem-like cells. It was clear, however, that the $CD44^{+}/CD24^{-/low}$ surface markers are enriched for tumorigenic cells in some, but not all breast cancers [226]. As a consequence, the concepts of the CSC hypothesis are being doubtful due to the recent demonstration of a dynamic equilibrium between differentiated cells and CSCs: as CSCs can differentiate into terminally differentiated cells, such terminally differentiate cells

have also the ability to de-differentiate into a CSC state [225]. Moreover, the overlap between some phenotypic characteristics of CSCs and the phenomenon of EMT has embraced a definition of the CSC phenotype as a dynamic cell state rather than a distinct cell type [226]. In addition, there is now direct evidence to demonstrate that at least in some types of cancer (that is, acute lymphoblastic leukemia) intra-tumor genetic heterogeneity is present not only in terminally differentiated cells, but also in the CSC population as defined by xenograft experiments [227].

Nonetheless and at least for certain malignant diseases, the CSC model is undoubtedly proved to be responsible for tumor progression and aggressiveness. Considering also the evidence for clonal evolution in the pathogenesis of cancer, it appears inescapable that both models should be integrated in order to explain intra-tumor heterogeneity [228].

5.3. Co-existence of CSC and Clonal Evolution Models

It is currently well known that most solid tumors show a wide genomic instability [229]. In fact, a large diversity of molecules responsible for the maintenance of the integrity of the genome is affected by genetic alterations which have become well-known drivers of oncogenesis [230]. For example, in a disease like chronic myeloid leukemia, already proved to be driven by SCs, clonal evolution can co-exist when imatinib is administered: the malignancy becomes tumor-resistant due to the emergence of clones carrying mutations in the target of imatinib [231]. Besides that, the progression of chronic myeloid leukemia into acute lymphoblastic leukemia blast crises is caused by the development of subclones that contain inactivating lesions in the cyclin-dependent kinase inhibitor 2A (*CDKN2A*, also known as *ARF*) gene [232]. Such evidences suggest that both the CSC and the clonal evolution models can be seen as a major complex model rather than being mutually exclusive. Even with the notion that cancer cells may be genetically similar, they may occupy different positions in a differentiation hierarchy, reflecting thus the physiological hierarchy of the tissue of origin. Also, a single tumor may comprehend several CSC clones that are genetically different, but these cells can still have a common ancestor that sustained the first oncogenic

mutation. In this way, the assays and understandings recently brought into cancer research by the CSC field may be covered over the genetic data [228] (Figure 6).

However, and as we develop a more complete understanding of genetic heterogeneity within tumors, there may be some cancers in which genetic heterogeneity is the main driver of phenotypic and functional heterogeneity. As so, new models of cancer heterogeneity and plasticity may emerge that can be more important than the clonal evolution or the CSC model [233].

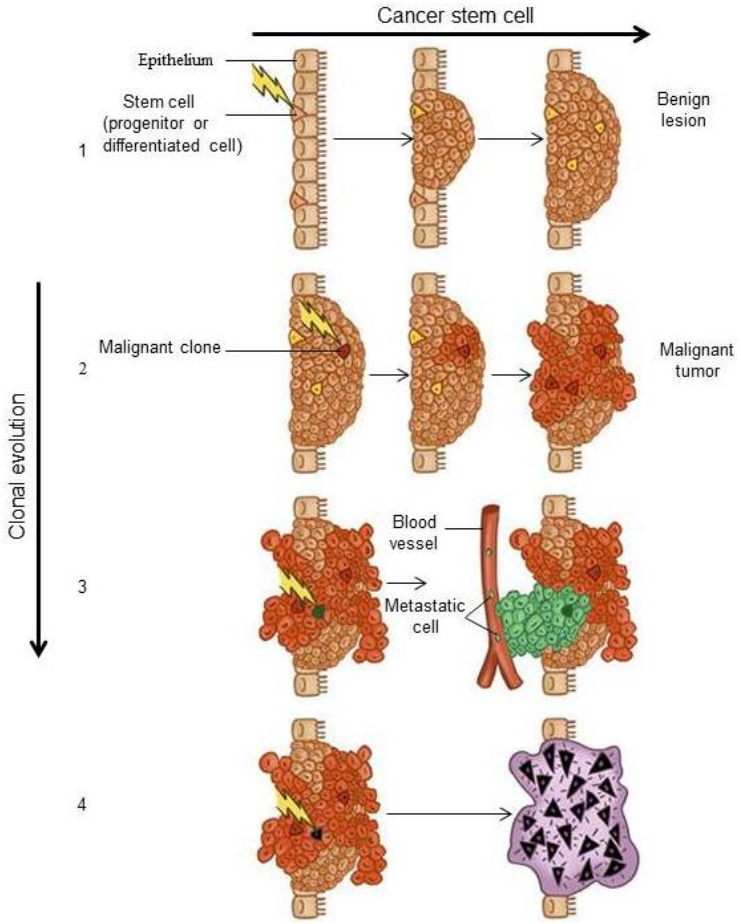

Figure 6. Theoretical synthesis of the clonal evolution and CSC concepts. Top to bottom: clonal evolution model. **1.** A first oncogenic mutation (lightning arrow) occurs in a SC of a normal epithelium, leading to the formation of a genetically homogeneous benign lesion. **2.** A second mutation occurs in one of the cells in the benign lesion leading to the growth of a malignant clone with invasive potential within the primary tumor. **3.** A third mutation occurs in a cell within the malignant subclone resulting in the entry into the blood vessel for distant metastasis. **4.** A final mutation give rise to a tumor completely taken over by cells that behave as CSCs. Left to right: at each stage of this clonal evolution process, tumors and subclones within tumors contain cells that behave as CSCs. Since the final hit (4) causes all cells to behave as CSCs, the CSC concept become meaningless at this stage. *Adapted from [228].*

6. Cancer Stem Cells

6.1. Origin of the CSCs

Although the concept that cancer may arise from a small cell population with SC characteristics has been proposed since more than 150 years, new evidence have given an impetus to it through new advancements on SC research. At the present, the origin of CSCs can be explained by three major hypotheses including the (1) malignant transformation of normal SCs, (2) dedifferentiation of mature cancer cells through EMT and (3) induced pluripotent cancer cells. The first hypothesis lies on the belief that CSCs arise from somatic stem or progenitor cells with genetic or epigenetic alterations [234-236]. The long-lived nature of somatic SCs means that they are potential targets for the multiple accumulated mutations that lead to malignancy [51]. Although a SC origin for CSCs has to be determined empirically, a number of adult SCs have been shown to undergo spontaneous transformation *in vitro* to generate tumorigenic SCs [237-241]. Theoretical evidences that supports this assumption is the notion that only adult SCs with self-renewal and differentiation abilities could accumulate several mutations along an acceptable time to acquire a malignant potential, considering the low mutational rate in a normal somatic cell. Also, and due to the functional similarities between CSCs and SCs, it is suitable to consider SCs as the origin of CSCs [242]. However, the demonstration of the ability that a differentiated tumor cell obtains to self-renew through the EMT gave rise to a new assumption that CSCs can be enriched within an existing malignancy [243]. The suggested mesenchymal phenotype in normal BSCs and BCSCs and the reported gene expression characteristic of EMT led to this hypothesis, which was further corroborated in other cancer types with the co-expression reported of EMT and SC markers [243-245].

The third theory of the origin of CSCs postulates that endogenous reprogramming or residues of the embryo are the source of CSCs. The recent improvements of induced pluripotent stem cells (iPSCs) artificially derived from an adult somatic cell have been used to explain the transformation of normal cells into tumor cells [246]. Nonetheless and has related above, induced pluripotency and oncogenic transformation can be seen as a connected process rather than being

models mutually exclusive. In fact, the specific tumor microenvironment existing in each cancer types also influence the behavior of several subsets of cancer cells that can act in different ways [247, 248].

6.2. CSC Niche

With the discovery of CSCs, attentions naturally turned to the importance of the interactions of cells with their microenvironment [249]. The so called CSC niche comprehend different types of cells (niche cells, stromal cells, immune cells and vasculature) that surrounds the CSCs and also, secreted factors derived from these cells offering a “fertile territory” for CSCs to propagate [250]. Hence, for CSCs that exhibit stem cell-like features and have the capability to regenerate the bulk of tumor cells without losing their self-renewal propensity, the CSC niche act as a microenvironment regulatory system for these cells. [251]. Given the complexity of SCs, internal and external signals must exist in order to assure a correct balance between SCs and their progeny. While the internal signals (molecular pathways) are responsible for differentiation monitoring of SCs, external signals (cells from the microenvironment) must assure the correct anchorage of SCs to their niche and also to segregate factors responsible for quiescence and an undifferentiated state maintenance of SCs [252]. Considering that normal SC niches are also known to serve as CSC niches, the detachment of CSCs from their niche leads to their asymmetric or symmetric division (Figure 7). Numerous studies suggested that the CSC niche is essential for the production and development of CSCs [253-255], like their existence near the endothelial cells (supposed to stimulate stemness through certain factors like Notch), in gliomas [254, 255]. Likewise, activin and nodal secretory proteins were shown to be produced, apart from CSCs, by stellate cells, stimulating CSCs in a paracrine fashion [256]. Concretely, in breast cancer, mesenchymal cells were proved to support CSCs through a signaling loop modulated by Interleukin-6 (IL-6) and chemokine C-X-C motif ligand 7 (CXCL7) [257].

Other important studies observed that CSCs respond to anti-tumor agents differently, both *in vitro* and *in vivo*, strengthening the role of CSC niches in response of CSCs to a given therapy [258]. In fact, the demonstration of the inability of less malignant tumors to form new tumors following implants into mice

supports the notion that tumor microenvironment can be determinant for the fate of certain tumor cells [259]. Considering that the major analysis studies of CSCs are dependent on CSC marker selection, the success rate of xenotransplantation assays may be compromised due to the loss of niche-derived factors upon the isolation of tumor cells for experimental designs [260]. In this way, it is plausible to assume that tumor microenvironment is crucial for the malignant progression of CSCs. It remains to be more fully elucidated what role stromal cells play in shaping this unique microenvironment, even if these cells were reported to regulate the invasive phenotype of BCSCs and to control tumor invasion. The mechanisms by which CSC regulation occurs have yet to be fully determined [242].

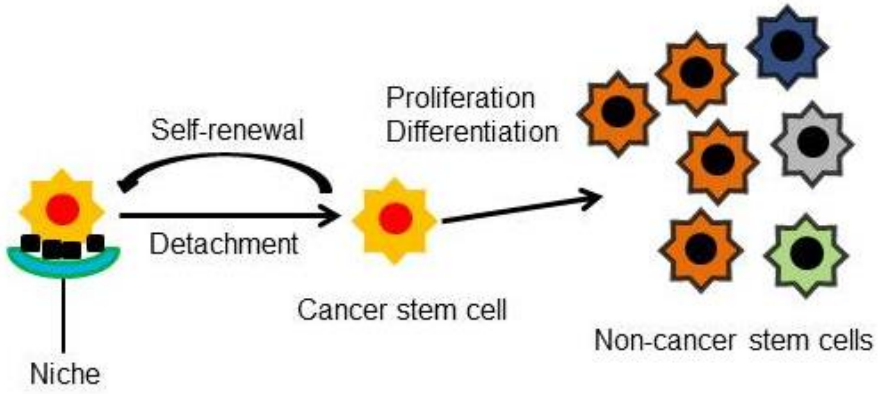

Figure 7. CSCs have two abilities, self-renewal and differentiation potency. Thus, the detachment of CSCs from their niche leads to their asymmetric or symmetric division. *Adapted from [261]*.

6.3. The Role of Quiescence in CSCs

Stem cell quiescence is extremely important for cancer therapy since that quiescent CSCs are frequently resistant to both conventional chemotherapy and targeted therapies contributing for tumor relapse and discontinuation of treatment [262]. Sparse knowledge about quiescence mechanisms has however limited significant advances in targeting quiescent CSC populations that are specifically resistant to drugs. Nonetheless, recent discoveries suggest that the quiescent state is not just a passive state but, instead, actively regulated by several intrinsic mechanisms. Quiescent SCs are believed to be sensitive to environmental modifications responding through the re-entrance in the cell cycle for proliferation. In fact, it is proposed that CSCs may adopt the quiescent state to resist metabolic stress and to preserve their genomic integrity. Quiescent CSCs may be prompted

for activation by specific energetically favorable mechanisms that are compatible with the low metabolic state of quiescence [263]. Such modulation of these cells can consequently generate rapid and global responses needed for activation. This assumption is based through recent studies that have provided a better understanding of the regulatory mechanisms that control SC quiescence. A family of miRNAs was shown to have a preponderant role in regulating G0 to G1 transition and cell cycle progression of SCs [264]. Related to breast cancer, different signaling pathways such as Hedgehog, Notch and Wnt have been implicated in multiple aspects of CSC quiescence [265-268].

High-throughput analyses of various SC subpopulations have recently provided valuable information on the molecular signatures of quiescent SCs in different tissue compartments. However, each pathway needs to be tested in studies of SC quiescence *in vivo* to determine the functional relevance. Such task is still extremely challenging due to the high levels of heterogeneity existing in cancer, especially in breast cancer, demanding thus more efforts in this field [269].

6.4. CSCs and Intra-tumor Heterogeneity

Intra-tumor heterogeneity denotes the coexistence of subpopulations of cancer cells that differ in their genetic, phenotypic or behavioral characteristics within a given primary tumor and between a given primary tumor and its metastasis. This diversity can be attributed to genetic and epigenetic factors and to non-hereditary mechanisms such as adaptive responses or fluctuation in signaling pathways [206, 270]. CSCs have different combinations of genetically derived cells with particular predispositions for growth, survival and dominance in the tumor micro-environment [270]. Looking at the CSC model, which postulates that disparities in the differentiation status between individual tumor cells lead to different functional properties within these cells, non-genetic factors may also contribute to tumor heterogeneity [242].

Different clinical outcomes are currently being explained by intra-tumoral phenotypic heterogeneity [257]. However, the frequency of CSCs in different types of cancers are also associated with aggressive forms of cancer, what in turn led to the assumption that CSCs are more prone to be resistant to treatment with a higher metastatic propensity than non-CSCs. In this way, it is plausible to assume

that subpopulations of cancer cells within a tumor may carry their own group of CSCs, but these cells and non-CSCs will share heritable genetic and epigenetic alterations. Inter-tumor heterogeneity of CSCs has also been identified due to the variations seen between features and prevalence of CSCs among different stages of cancer [271].

As referred above, the attempts to identify, isolate and characterize CSC populations are mostly dependent on cell-surface markers, obtained from the analysis of normal SCs in the tissue of origin [247]. Thus, a correct validation and selection of stem-cell marker is demanded in order to achieve solid and consistent results regarding this field. However, and due to the existence of a plethora of stem-cell markers for each type of cancer, different stem-like phenotypes were already been demonstrated to have different clinical outcomes for the same cancer. Specifically, in breast cancer, several proteins were reported to have stem properties that are currently being used to predict breast cancer aggressiveness and behavior [272].

7. Breast Cancer Stem Cells

In the last two decades, breast cancer research has majorly focused in the identification, isolation and characterization of BCSCs. In order to do so, some genes having stem properties were studied and their corresponding proteins were subsequently validated as BCSC markers [224, 273]. As a consequence, a plethora of studies were published describing the impact of BCSCs identified by these established BCSC markers, like CD44, CD24 and ALDH1 as tumor initiating cells in breast cancer progression with high propensity to metastasize and to be resistant to therapeutic treatments [273-275]. With the increasing evidences for such ability, researchers attempted to demonstrate which genes or gene signaling pathways could potentially contributed for the tumorigenic potential of BCSCs. In fact, Notch, Wnt/β-catenin, or Hedgehog signaling pathways were shown to be deregulated in subpopulations of these cells [276, 277]. With the upcoming evidences regarding the effects of the stroma and the microenvironment in breast tumor progression, several genes were also reported to be associated with BCSCs [278]. The phenomenon of EMT and mesenchymal-epithelial transition (MET) in breast cancer cells during tumor progression was also an important discovery associated to BCSCs as an explanation for their ability to invade and to colonize other parts of the body, which has indeed questioned some of the tenets of the CSC model [225]. With all this knowledge, targeting BCSCs for breast cancer treatment was demanded and currently, some important inhibitors targeting subpopulations of BCSCs or gene signaling pathways that regulate these subpopulations are reported to be strongly effective [279, 280].

One concern has however changed the definition of BCSCs: the breast cancer heterogeneity [218]. Due to the observations that not all BCSC markers were expressed in all breast cancer types, a research for different BCSC markers and different combinations of these markers that could be restricted to a specific breast cancer subtype or associated to aggressive forms of this disease has been ongoing [281, 282]. As a consequence, different BCSC phenotypes were described and characterized and in the future, other molecuels will be reported to have stem properties. Beyond the tenets of the CSC model, it is important to define which BCSC phenotypes have high tumorigenic potential and also high

ability to resist to therapeutic agents [283]. Moreover, it is also crucial to determine which oncogenes or tumor suppressor genes, other than those already described, are consistently mutated within these phenotypes and are able to drive tumorigenesis.

In ductal breast cancers, a large study comprehending 12 different markers were immunohistochemically characterized showing that the prevalence of stem cell-like and more differentiated markers varies according to tumor subtype and histologic stage [284]. For that reason, a concise review is here presented regarding the implications of the most studied BCSC markers and BCSC phenotypes in breast cancer progression and treatment as well as the description of promising inhibitors able to target these cells.

7.1. $CD44^{+}/CD24^{-/low}$ Phenotype

The combination of the BCSC markers CD44 and CD24 is by far the most extensively studied and undeniably the most contentious. The pioneering study by Al-Hajj *et al* showed that as few as 100 $CD44^{+}/CD24^{-/low}$ $Lineage^{-}$ cells in patients with breast cancer could form tumors in mice, whereas tens of thousands of cells with alternative phenotypes failed to do so [224].

Immunohistochemically, breast cancer tissues were investigated for the prevalence of $CD44^{+}CD24^{-/low}$ tumor cells and their prognostic value. In a study including 136 patients with and without recurrences, the prevalence of $CD44^{+}/CD24^{-/low}$ cells was ≤10% in 78% of cases and >10% in the other 22%. However, no significant correlation between the prevalence of this phenotype and tumor progression was noticed nor significant differences was seen between recurrence, DFS or OS [285]. In another study comprehending 95 patients with IDCs subjected to mastectomy, radiotherapy, chemotherapy, and axillary lymph node dissection, $CD44^{+}/CD24^{-/low}$ cancer cells were shown to be abundant in the basal subgroups and absent in HER2-positive tumors [286]. This phenotype was also associated with BRCA1 mutational status which was correlated with basal-like tumor status and despite its association with increased poor prognostic features, it was not able to predict OS [287]. Regarding such important studies, the $CD44^{+}/CD24^{-/low}$ phenotype has not a distinct prognostic value but is enriched in basal-like breast cancer subtype.

Gene expression profiling of $CD44^{+}/CD24^{-/low}$ breast cancer cells revealed a gene signature of 186 genes associated with invasion and poor prognosis [288, 289]. This signature was enriched in genes related to cell cycle, calcium-ion binding, chemotaxis, differentiation, protein transport, signal transduction and ubiquitination. Among these genes, this phenotype was observed to express high levels of IL-1 alpha, IL-6, IL-beta and urokinase plasminogen activator which predisposed to distant metastases (table 3).

The enrichment of $CD44^{+}/CD24^{-/low}$ cells demonstrated in primary breast tumors following radiation and chemotherapy has suggested an innate resistance to standard treatments [275]. The presence of ATP-binding cassette (ABC) transporters (which confers resistance to chemotherapeutic agents) highly expressed in a subpopulation of CSCs containing these markers has led to such assumption [290, 291]. In fact, the ability of these cells to reproduce, in an *in vitro* model following at least four generations of xenograft transplanted mice, has also suggested a significant role in tumor relapse and metastasis [292]. Potential mechanisms of chemotherapy and radiation resistance associated to this phenotype were shown to include the presence of lower concentration of reactive oxygen species, cell dormancy, efficient DNA repair mechanisms, overexpression of EMT markers, Wnt/β-catenin, Hedgehog and Notch signaling pathways and STAT1 and STAT3 signaling activation [276, 243, 293-298]. As a consequence, the $CD44^{+}/CD24^{-/low}$ phenotype in breast cancer is currently being assessed as a therapeutic target.

One of the most promising therapeutic agents belongs to the class of gamma-secretase inhibitors (MK0752) that was recently administered in patient-derived tumor xenograft in combination with docetaxel. This inhibitor was reported to improve the docetaxel activity leading to a decrease of $CD44^{+}/CD24^{-/low}$ tumor cells, to a reduced mammosphere forming activity and consequently, to the inhibition of tumor formation after serial transplantations. With these results, a phase I clinical trial in advanced breast cancers that did not respond to standard treatment was developed culminating in a decrease of $CD44^{+}/CD24^{-/low}$ tumor cells and in the reduction of the tumor bulk [279]. This inhibitor is also being tested in combination with endocrine therapy (tamoxifen or letrazole in early stage breast cancer patients) and chemotherapy (docetaxel in locally advanced or metastatic breast cancer patients) [299].

Other inhibitor from the same class such as PF-03084014 was also administered in a Phase I trial for the treatment of advanced breast cancers being able to reduce Notch activity and to considerably decrease tumor cell migration and mammosphere forming efficiency [299]. Its ability in decreasing self-renewal capability and expression levels of Notch target genes was also demonstrated in *in vivo* studies [300].

Another promising compound for breast cancer treatment is metformin, a drug generally used for anti-diabetic therapy. Metformin have been shown to target preferentially $CD44^+/CD24^{-/low}$ cell subpopulations in different molecular subtypes of breast cancer cell lines and to have a synergistic effect in eradicating CSCs upon its administration with doxorubicin [301, 302]. Indeed, this compound is believed to interfere directly with the tumorigenesis of $CD44^+/CD24^{-/low}$ tumor cells and to prevent neoplastic mammary lesions [303]. Other promising compounds are being tested in order to target $CD44^+/CD24^{-/low}$ tumor cells (table 3), although their potential still need to be proved [304-306].

With all the experimental studies done regarding this phenotype it became clear that the $CD44^+/CD24^{-/low}$ surface markers are enriched for tumorigenic cells in some, but not all breast cancers. Hence, the validity of the combination of these markers as a definition of BCSC has been called into question and additional markers have been reported like ALDH1 [307].

7.2. ALDH1

ALDH1 is broadly used as a functional marker in various types of cancer. Ginestier *et al* were the first to demonstrate the ALDH1 activity as a marker of stemness in normal and malignant breast cells. They were able to generate a stable tumor via orthotopic injection of 500 ALDH1-positive cells (evaluable by the ALDEFLUOR assay) into the mammary fat pads of Non-obese diabetic/severe combined immunodeficiency (NOD/SCID) mice. Such tumorigenic ability of ALDEFLUOR positive cells was also seen to be increased when shared with the $CD44^+/CD24^{-/low}$ phenotype, since that only 20 of such cells were sufficient to generate tumors in animals [273]. Nonetheless, functional studies revealed that $ALDH1^+$ cells were more prone to form colonies and tumors than $CD44^+/CD24^{-/low}$ cells and also to be more chemoresistant [273, 308].

Several immunologic studies have attempted to assess ALDH1 as a prognostic marker in breast cancer. In the same study of Ginestier *et al*, a cohort comprehending 481 breast tumors from 2 independent breast tumors showed a prevalence of 30% for ALDH1 positivity and a correlation with high histologic grade, HER2 overexpression and absence of ER and PR expression [2]. ALDH1 was also demonstrated to be an independent prognostic factor and the same was obtained in a group of 80 patients who received breast-conserving therapy. However, no association between ALDH1 tumor staining and micrometastatic disease was noticed [309]. In a larger cohort of more than 200 primary breast cancer patients, Morimoto *et al* reported a tendency for a worse prognosis with high ALDH1 expression [310]. Nonetheless and regarding the type of cells expressing ALDH1, a well-designed study from Resetkova and colleagues found that ALDH1 expression in the stromal compartment of 2 cohorts of triple-negative breast tumors had prognostic significance although being associated with good, rather than poor DFS [311].

ALDH1 expression and clinical outcomes were also explored in inflammatory breast cancers (a particularly lethal form of breast cancer characterized by exaggerated lymphovascular invasion) revealing that ALDH1 expression was a predictive factor for early metastasis and decreased survival [312]. Contrarily, in another study, no significant correlation between ALDH1 expression and clinicopathological variables was obtained, despite a trend toward OS association [313].

With all these contradictory results, the reliability of ALDH1 expression as a clinical predictor of response to treatment is doubtful enhancing thus the need for a standard protocol and evaluation process as well as the necessity to consider the differences between whole tissue staining vs tissue microarray staining [314]. Consistent findings between the reported studies are shown in table 3.

Functional observations associated with ALDH1 are the increased levels of Notch and β-catenin which regulate the deacetylation of ALDH1 increasing its tumorigenicity *in vivo* and contributing to a poor clinical outcome. Increased expression of Hipoxia Inducible Factors 1 and 2α was also shown to be associated with ALDH1 activity which is believed to raise the metastatic propensity of $ALDH1^{high}$ cells [277, 315, 316]. Besides that, increased expression of HOXA1 and MUC4 were associated with high ALDH1 activity contributing also for tumor

relapse and metastasis. TGFβ2 signaling pathway activation was also shown to be involved in the pathological regulation of ALDH1 in breast cancer [317-319].

Using the $ALDH^{high}/CD44^{+}/CD24^{-/low}$ and $ALDH^{high}$ phenotypes, Charafe-Jauffret *et al* and Croker *et al* provided the first direct experimental evidence implicating $ALDH^{high}$ cells in breast cancer metastases *in vivo*, respectively. Cells with a CSC phenotype characterized by ALDH activity were shown to have an improved ability for metastatic behavior *in vitro* (adhesion, colony formation, migration, and invasion) and/or metastases *in vivo*, supporting the hypothesis that CSCs might act as metastasis-initiating cells [283, 317, 320-322].

ALDH1 is also involved in metabolizing chemotherapeutic drugs so its inhibition was believed to contribute for the reduction or elimination of BCSCs. Indeed, significant resistance to sequential paclitaxel- and epirubicin-based chemotherapy was found in tumor cells expressing ALDH1 [304]. The inhibition of ALDH1 activity was also demonstrated to reduce stem-like properties and resistance to drugs and radiotherapy. As a consequence, these findings have emphasized the need to target $ALDH1^{+}$ tumor cells in breast cancer treatment [322].

Even though the sparse existence of studies evaluating the impact of pharmacological or immune targeting of ALDH on metastases *in vivo*, most of them showed a decrease of the metastatic burden. In this way, rationalized small molecule discovery has been proposed as a viable methodology in order to overcome these difficulties and such improvement led to the current development and generation of isoform-specific ALDH inhibitors (table 2). Salinomycin (an inhibitor of Wnt pathway) is currently the most promising chemotherapeutic drug demonstrated to inhibit the distinctive phenotypic properties of CSCs rather than by inducing apoptosis of these cells [280]. These findings underline the potential therapeutic value of targeting these properties to reduce the likelihood of tumor recurrence following chemotherapy.

Despite the enrichment of CSCs in $ALDH1^{+}$ populations reported in several tissues, enzymatic activity measured by ALDEFLUOR alone is much more transient than the expression of traditional surface markers. The usefulness of ALDH1 activity as a sole CSC marker may than be limited but can be increased if staining cells simultaneously for ALDH1 activity and for more stable markers like CD44 or CD133 [323].

7.3. $CD44^{+}/ALDH1^{+/high}$ Phenotype

The importance of ALDH1 activity in breast cancer has been explored alone or in combination with the $CD44^{+}/CD24^{-/low}$ phenotype. However, analysis of tumor samples revealed that only 1% of the ALDH-positive cell population also had the $CD44^{+}/CD24^{-/low}/Lin^{-}$ phenotype reported for BCSCs [273]. In fact, an increase in the population of $ALDH1^{+}$ cells but not $CD44^{+}/CD24^{-/low}$ cells has been observed in breast cancer tumor biopsies after neoadjuvant treatment [223]. Considering the limited usefulness of ALDH1 activity as a sole BCSC marker, other combinations have been studied like the $CD44^{+}/ALDH1^{+/high}$ phenotype. This phenotype was recently demonstrated to have high tumorigenic ability in breast cancer cell lines and also high metastatic propensity, being resistant to standard cancer therapies [322].

An interesting *in situ* method to define CSCs in FFPE breast cancer tissues through a quantitative immunofluorescence method was designed by Neumeister and colleagues in order to measure the co-expression of CD44, ALDH1 and cytokeratin. Using a retrospective collection of 321 node-negative and 318 node-positive patients, localization of $CD44^{+}/ALDH1^{+}$ cells was shown within the epithelial (cytokeratin) compartment of breast tumor tissue. Even if this co-expression was seen in variable sized clusters and only in 6% of cases, such combination showed significantly worse outcome, being able to identify high risk patients in breast cancer [281]. Our research group has also studied the co-expression of CD44 and ALDH1, and remarkably, such combined expression was seen to be higher in DCIS lesions when compared with IDCs of the breast, enhancing the tumorigenic potential of these putative BCSCs [324]. More interestingly, we also demonstrated, in a cohort comprehending 250 patients having different benign and malignant breast lesions, that the $CD44^{+}/ALDH1^{+}$ phenotype was significantly increased in high-grade DCIS when compared with IDCs. Moreover, this phenotype was found to be predominantly in a quiescent state (negativity for Ki-67 proliferation marker) rising some questions about the real role of dormancy in BCSCs [325].

A gene expression analysis study revealed an increased expression of Notch and Wnt/β-catenin signaling pathways in $CD44^{+}/ALDH1^{+}$ breast cancer cells isolated from an IBC cell line [326]. However, and concerning this phenotype,

further genetic and even epigenetic analysis are demanded to better understand its behavior in breast cancer progression. Like $CD44^+/CD24^{-/low}$ and/or $ALDH1^+$ tumor cells, it would be also interesting to depict the mechanisms that drive the progression of $CD44^+/ALDH1^+$ tumor cells.

Indeed, Croker *et al* have directly inhibited ALDH activity with the specific ALDH inhibitor diethylaminobenzaldehyde (DEAB) and indirectly through all-trans retinoic acid (ATRA). They isolated $ALDH^{high}/CD44^+$ and $ALDH^{low}/CD44^-$ populations and demonstrated that $ALDH^{high}/CD44^+$ human breast cancer cells were resistant to certain chemotherapy drugs [322]. For that, they subjected both populations to treatment with doxorubicin, paclitaxel or radiation in the presence or absence of DEAB or ATRA and concluded that the reduction in cell viability was significantly greater in the $ALDH^{high}/CD44^+$ population. Furthermore and in contrast to $ALDH^{low}/CD44^-$ cells, $ALDH^{high}/CD44^+$ cells showed increased basal activity in a series of DNA response proteins including p-glycoprotein, glutathione-S-transferase pi and/or CHK1 checkpoitn homolog 1 (table 3).

Although such results enhance the need of targeting $CD44^+/ALDH1^+$ tumor cells in breast cancer, no other agents or drugs have been developed to directly target this phenotype. Only inhibitors that target $CD44^+/CD24^{-/low}$ tumor cells or $ALDH1^+$ tumor cells were shown to be promising. Despite the current improvements regarding the adverse effects of ALDH1 and CD44 for breast cancer treatment, additional studies in order to infer about the tumorigenic and metastatic ability of $CD44^+/ALDH1^{+/high}$ tumor cells combined with their quiescence status still have to be depicted. Moreover, the development of additional promising inhibitors to target this phenotype is also demanded.

7.4. Prominin-1 (CD133)

CD133 has been recently included in CSC research. It is also named Prominin-1 for its prominent location on the protrusion of cell membranes being the first identified gene in a class of novel pentaspan transmembrane glycoproteins. Although it was initially considered to be a marker of hematopoietic SCs, CD133 mRNA transcript is also found in normal non-lymphoid hematopoietic tissue [327] and has been shown to play a role in SC migration and asymmetric division [328]. CD133 was reported to be overexpressed in several solid tumors,

[329, 330] including colon cancer and glioblastoma [331, 332]. In IBCs, CD133 expression was demonstrated by Liu *et al* [333] where they assumed that its expression could be of help in a more accurately prediction of breast cancer aggressiveness and determination of the most suitable treatment. Actually, in BRCA1-associated breast cancer cell lines, $CD133^+$ sorted cells were shown to contain CSC properties including a greater colony-forming efficiency, higher proliferative output and greater capability to form tumors in NOD/SCID mice [334]. Moreover, CD133 was also proved to be suitable in the identification of CSCs in triple-negative breast cancers through several *in vitro* [335, 336] and *in vivo* studies [337]. In addition, the recent use of CD133 to detect circulating tumor cells in triple-negative breast cancer patients [338, 339] has increased the attention of this marker, emphasizing its role in the determination of the prognostic and predictive value in this breast cancer subtype. Expression of CD133 was also recently reported in 22 out of 25 cases of inflammatory breast cancer [282]. Taken together, these interesting results increase the need of more advanced research to understand the role of CD133 in BCSCs.

Expression of SC associated genes, such as Notch1, ALDH1, Fgfr1 and Sox1, was shown to be increased not only in $CD44^+/CD24^{-/low}$ but also in $CD133^+$ breast cancer cells [334]. Xenograft-initiating breast cancer cells enriched in $CD44^+/CD49f^{high}/CD133/2^{high}$ cells were also shown to have elevated levels of Nanog, Sox2, and/or Bmi-1 [340]. Further extensive CD133 profiling in breast cancer have hence to be performed to define $CD133^+$ breast cancer cells as tumor initiating cells in breast cancer tumors.

Due to the increasing importance of CD133 expression in breast cancer progression, attempts have been made to correlate its expression with tumor relapse and resistance to chemotherapeutic agents. In fact, CD133 expression was reported to be correlated with tumor recurrence in breast cancer patients [341]. In drug-sensitive MCF-7 cells, only a small fraction of cells was found to be $CD133^+$ [342]. In an interesting study, polymeric nanoparticles loaded with paclitaxel and surface functionalized with anti-CD133 antibody demonstrated efficient elimination of tumor initiating cells *in vitro* and significant inhibition of tumor-regrowth *in vivo* [343]. With such results, CD133 is regarded as a potential target for anticancer therapeutics, being possible to reduce tumor recurrence in breast cancer through the elimination of $CD133^+$ cells. Thus, additional studies

investigating specific drugs that could efficiently target this protein are required in order to be applied in the clinic.

7.5. Integrins

The use of the integrins CD29 (*β*1) and CD49f (*α*6) in combination with CD24 was recently demonstrated to be able to identify mouse mammary SCs [344, 345]. Since that all previously described SC markers were shown not only to identify normal mammary SCs but also to isolate mammary CSCs, Vassilopolus *et al* used CD24/CD29 and CD24/CD49f to successfully identify a subpopulation of mammary tumor cells [346]. Such demonstration highlighted hence the importance of CD29 and CD49f in BCSCs. CD49f heterodimerizes with either the CD29 or CD104 (*β*4 integrin) subunits to generate the CD49fCD29 and CD49fCD104 integrins, which function primarily as laminin receptors [347]. Besides that, CD49f cooperates with receptor tyrosine kinases to communicate, bidirectionally, between the cell and the ECM. Interestingly, however, the CD104 subunit appears to be expressed at very low levels, if at all, in CSCs when compared to non-CSCs indicating that CD49fCD29 is the dominant integrin expressed by CSCs [348, 349].

In mice, CD29 represents the predominant expressed integrin in mammary epithelial cells and is aberrantly expressed in human breast carcinomas contributing to diverse malignant phenotypes, including EMT, metastases and angiogenesis [349-352]. Moreover, in patients with IBCs, high CD29 expression was found to be associated with significantly shorter OS and DFS [353, 354]. In human breast cancers, CD49f integrin is overexpressed and was shown to be an independent prognostic factor [355]. $CD49f^{+}$ cancer cells were also associated with a higher probability of distant metastasis after initial surgery and poor clinical outcomes with respect to both DFS and OS [356]. Additionally, normal human SCs and myoepithelial progenitor cells characterized by $CD49^{high}/EpCAM^{-}$ cells were shown to express vimentin, a common EMT marker suggesting that some cells may be undergoing EMT [357]. Interestingly, an aberrant luminal progenitor cell population ($EpCAM^{+}/CD49f^{+}$) was also proposed to be the cell origin of BRCA1 associated basal breast cancers [358].

Functional analysis revealed that while knockdown of CD29 or CD49f alone slightly decreased cell migration ability, knockdown of both genes caused a

profound effect to block their migration, suggesting an overlapping, yet critical function of both genes in the migration of BCSCs [346]. Such interesting finding supports the notion that both integrins can pair with each other in order to form heterodimers for ECM components such as fibronectin and laminin [347]. Consistent with the assumption that a malignant social network mediates cell–cell adhesion and communication between CSCs and their microenvironment [359], both integrins may be implicated in mediating such network. Specifically, CD29/CD49f integrins may mediate CSCs–stroma interaction, relaying ECM signaling to cellular machinery leading to the increased activity of CSCs in terms of viability, differentiation and metastasis [346].

Although the CD29/CD49f integrins have been implicated in the function of breast and other CSCs [348, 349, 360], much needs to be learned about the contribution of these integrins to the genesis of BCSCs. It has been shown that either CD49f or CD29 contributes to therapy resistance, tumor relapse and metastases in breast cancer. As a consequence, the development of inhibitors that could potentially target these two integrins in breast cancer is required (table 3). Currently, interesting studies have been published with promising results in targeting these integrins like the use of short hairpin RNAs or micro-RNAs [361]. Targeting gene signaling pathways associated to these integrins or even specific kinases like FER (feline sarcoma-related) kinase that controls migration and metastasis of IBC cell lines by regulating CD49f- and CD29-integrin-dependent adhesion is also an interesting approach [362].

7.6. BCSCs and Next-Generation Sequencing

The continuous improvements of NGS technologies currently allow the analysis of hundreds of genes in just one population of cells, or even in one single cell [204, 363]. Such application opened a new window in the genomic field where a mutational, time-based lineage tree can now be delineated for a specific subtype of cancer considered to be highly aggressive. In this way and with NGS, it is possible to determine other genes from those already associated to be oncogenic, or, more importantly, to assess which mutated genes can be responsible to drive tumorigenesis considering the high levels of heterogeneity in cancers, especially in breast cancer [364]. In fact, NGS has recently been used for the analysis of the

molecular features of early stage breast cancer, leading to a genomic portrait of this disease. Within such portrait, *TP53* and *PIK3CA* mutations were the most frequent genomic alterations found in all breast cancer subtypes. Clinical relevance of *PTEN* mutations and deletions as well as *AKT1*, *BRCA1* or *BRCA2* mutations was also highlighted [211].

With this in mind and related to BCSCs, such technology would allow the definition of a mutational repertoire of each subpopulation of BCSCs here presented. Even using a small cohort, our research group was able to detect, through NGS, somatic mutations in $CD44^{+}/CD24^{-/low}/Ck^{+}/CD45^{-}$ breast cells isolated from non-malignant and malignant breast lesions. Mutations affecting the *TP53*, *NOTCH1*, *HRAS*, *AKT1*, *PTEN*, *CSF1R* and *RET* genes were detected in the malignant lesions suggesting a heterogeneous molecular profile of these BCSCs [365]. Thus, a practical example would be the application of NGS in isolated BCSCs (defined by different combinations of BCSC markers) from primary tumors and their corresponding metastases in order to determine which gene is more frequently mutated (hotspot mutations) in each subpopulation of BCSCs. Such approach would be of great importance for the development of additional therapeutic drugs that could be promising for the most known BCSC markers (CD44, CD24 and ALDH1), but also for the discovery of new targets directly associated with other BCSC markers like CD133 or integrins.

In the future, this growing technology will definitely revolutionize the CSC research with the upcoming of new gene signaling pathways directly involved in the progression of tumor initiating cells already proved to have stem properties, particularly in breast cancer [290].

Table 3. Characteristics of the different assessed BCSC phenotypes and BCSC markers.

BCSC phenotypes/ BCSC Markers	Tumorigenic potential	Clinicopathological features	Functional/mechanistic observations	Clinical observations	Inhibitors
$CD44^{+}/CD24^{-/low}$	Able to drive tumor formation when inoculated into NOD/SCID mice [224]	Enriched in basal-like and claudin-low breast cancer subtypes [286] Associated with BRCA1 mutational status [287] Poor prognosis [285]	Increased levels of IL-1α, IL-.6, IL-β and uPA [288, 289] Increased expression of TWIST and SNAIL1 [243, 293] Increased levels of Wnt/β-catenin, Hedgehog and Notch signaling pathways [294, 299] Low levels of ROS [276] Cell dormancy and efficient DNA repair mechanisms [296] Increased levels of ABC transporters, STAT1 and STAT3 [291, 297, 298]	Tumor recurrence [293] Resistance to radiation and standard treatments [275] High metastatic propensity [292]	Short hairpin RNAs [366] Gamma-secretase inhibitors [279, 299, 300] Metformin.[301, 302] ATRA or vorinostat [367] Niclosamide [304] Disulfiram/copper [305] Cyclophosphamide [306]
ALDH1	Able to generate a stable tumor via orthotopic injection of $ALDH1^{+}$ cells into NOD/SCID mice [273]	Associated with poorer clinical outcomes including ER negativity, basal subtype and HER2-amplification [273, 310]	Increased expression of Ki-67 and EZH2 [310, 368] Increased expression of HIF-1/2α [315] Increased expression of HOXA1 and MUC4 [317] Increased levels of TGFβ2, Notch and Wnt/β-catenin signaling pathways [316, 319, 369]	Tumor recurrence [284] Enhanced capacity for metastatic behavior [284, 320] Resistance to sequential paclitaxel-and epirubicin-based chemotherapy [304]	DEAB [322] ATRA [322] Salinomycin [280] Disulfiram/copper [370]
$CD44^{+}/ALDH1^{+/high}$	Able to drive tumor formation when inoculated into NOD/SCID mice [283]	Able to identify high risk patients in breast cancer [281] Enriched in high-grade DCIS [325]	Increased levels of Notch and Wnt/β-catenin signaling pathways [326] High levels of p-glycoprotein, GSTpi, and/or CHK1 [322] Cell dormancy [325]	Predicts distant metastases and OS [325] High metastatic propensity [322] Resistant to standard cancer therapies [322]	DEAB [322] ATRA [322]
CD133	Ability to form tumors in NOD/SCID mice from BRCA1-associated breast cancer cell lines [334]	Enriched in IBCs and particularly in TNBCs [335-337] CTCs detection in triple-negative breast cancer patients [338, 339]	Increased expression of Notch1, ALDH1, Fgfr1 and Sox1 [334] High levels of Nanog, Sox2, and/or Bmi-1 in $CD44^{+}/CD49f^{high}/CD133/2^{high}$ breast cancer cells [340]	Tumor recurrence [341] Resistant to standard cancer therapies [342]	Paclitaxel and surface anti-CD133 antibody [343]
CD29/CD49f	Enrichment of CD29 and CD49 in combination with CD24 for cancer-initiating cells in primary breast tumors [346]	CD29 associated with shorter OS and DFS [353, 354] CD49f associated with poor clinical outcomes and regarded as a prognostic factor [355, 356]	Increased expression of EMT markers [82, 357] $EpCAM^{+}/CD49f^{+}$ breast cancer cells proposed to be the cell origin of BRCA1 associated basal breast cancers [358]	Tumor relapse [372, 373] High metastatic propensity [359] Resistance to therapy [374, 375]	Combination of miR-9-3p with AZD6244 for CD29 [360]

References

1. K. Wambach, J. Riordan, Breastfeeding and Human Lactation, 5th ed., Jones & Bartlett Learning, LLC, 2014.

2. D.B. Kopans, Breast Imaging, 3th ed., Lippincott Williams & Wilkins, Philadelphia, 2007.

3. P.P. Rosen, Rosen's Breast Pathology, 2nd ed., Lippincott Williams & Wilkins, Philadelphia, 2001.

4. J.E. Rusby, E.F. Brachtel, J.S. Michaelson, F.C. Koerner, B.L. Smith, Breast duct anatomy in the human nipple: three-dimensional patterns and clinical implications. Breast Cancer Res. Treat., 106 (2007) 171-179.

5. S.R. Lakhani, E.I. Schnitt, P. Hoo Tan, M. van der Vijver, WHO Classification of Tumours of the Breast, 4th ed., I.A.R.C., Lyon, 2012.

6. P.M. Vogel, N.G. Georgiade, B.F. Fetter, F.S. Vogel, K.S. McCarty, The correlation of histologic changes in the human breast with the menstrual cycle, Am. J.Pathol., 104 (1981) 23.

7. T.A. Longacre, S.A. Bartow, A correlative morphologic study of human breast and endometrium in the menstrual cycle, Am. J. Sur. P., 10 (1986) 382-393.

8. R. Ramakrishnan, S.A. Khan, S. Badve, Morphological changes in breast tissue with menstrual cycle, Mod. Pathol., 15 (2002) 1348-1356.

9. T.J Anderson, Normal breast: myths, realities, and prospects, Mod. Pathol., 11 (1998) 115-119.

10. P. Butler, A. Mitchell, H. Ellis, Applied Radiological Anatomy for Medical Students, Cambridge University Press, Cambridge, 2007.

11. D.L. Page, T.J. Anderson, Diagnostic Histopathology of the Breast, Churchill Livingstone, New York, 1987.

12. E.D. Jarasch, R.B. Nagle, M. Kaufmman, C. Maurer, W.J. Bocker, Differential diagnosis of benign epithelialproliferations and carcinomas of the breast using antibodies to cytokeratins, Human Pathol., 19 (1988) 276-289.

13. W. Böcker, B. Bier, G. Freytag, B. Brommelkamp, E.D. Jarasch, G. Edel, B. Dockorn-Dworniczak, K.W. Schmid, An immunohistochemical study of the breast using antibodies to basal and luminal keratins, alpha-smooth muscle actin, vimentin, collagen IV and laminin, Virchows Archiv. A., 421 (1992) 315-322.

14. M. Heatley, P. Maxwell, C. Whiteside, P. Toner, Cytokeratin intermediate filament expression in benign and malignant breast disease, J. Clin. Pathol., 48 (1995) 26-32.

15. W. Böcker, R. Moll, C. Poremba, R. Holland, P.J. van Diest, H. Burger, D. Wai, R.I. Diallo, B. Brandt, H. Herbst, A. Schmidt, M.M. Lerch, I.B. Buchwallow, Common adult stem cells in the human breast give rise to glandular and myoepithelial cell lineages: a new cell biological concept, Laboratory Investigation, 82 (2002) 737-746.

16. D.M. Abd El-Rehim, S.E. Pinder, C.E. Paish, J. Bell, R.W. Blamey, J. Robertson, R.I. Nicholson, I.O. Ellis, Expression of luminal and basal cytokeratins in human breast carcinoma, J. Pathol., 203 (2004) 661-671.

17. C. Pechoux, T. Gudjonsson, L. Ronnov-Jessen, M.J. Bissell, Human mammary luminal epithelial cells contain progenitors to myoepithelial cells, Dev. Biol., 206 (1999) 88-99.

18. H. Yaziji, A.M. Gown, N. Sneige, Detection of stromal invasion in breast cancer: the myoepithelial markers, Adv. Anat. Pathol., 7 (2000) 100-109.

19. M. Barbareschi, L. Pecciarini, M.G. Cangi, E. Macri, A. Rizzo, G. Viale, C. Doglioni, p63, a p53 homologue, is a selective nuclear marker of myoepithelial cells of the human breast, Am. J. Surg. Pathol., 25 (2001) 1054-1060.

20. S. Moritani, R. Kushima, H. Sugihara, M. Bamba, T.K. Kobayashi, T. Hattori, Availability of CD10 immunohistochemistry as a marker of breast myoepithelial cells on paraffin sections. Mod. Pathol., 15 (2002) 397-405.

21. N.K Popnikolov, S.M. Cavone, P.M. Schultz, F.U. Garcia, Diagnostic utility of p75 neurotrophin receptor (p75NTR) as a marker of breast myoepithelial cells. Mod. Pathol., 18 (2005) 1535-1541.

22. R. Ali-Fehmi, K. Carolin, T. Wallis, D.W. Visscher, Clinicopathologic analysis of breast lesions associated with multiple papillomas, Hum. Pathol., 34 (2003) 234-239.

23. A. Estabrook, Are patients with solitary or multiple intraductal papillomas at higher risk of developing breast cancer, Surg. Oncol. Clin. North. Am., 2 (1993) 45-56.

24. M.D. Althuis, J.M. Dozier, W.F. Anderson, S.S. Devesa, L.A. Brinton, Global trends in breast cancer incidence and mortality 1973–1997. Int. J. Epidemiol., 34 (2005) 405-412.

25. A. Jemal, F. Braym, M.M. Center, J. Ferlay, E. Ward, D. Forman, Global cancer statistics, CA Cancer J. Clin., 61 (2011) 69-90.

26. G.N Hortobagyi, J. de la Garza Salazar, K. Pritchard, D. Amadori, R. Haidinger, C.A. Hudis, M.C. Liu, M. Martin, M. Namer, J.A. O'Shaugnessy, Z.Z. Shen, K.S. Albain, The global breast cancer burden: variations in epidemiology and survival, Clin. Breast Cancer, 6 (2005) 391-401.

27. J.F. Forbes, The incidence of breast cancer: the global burden, public health considerations, Seminars in oncology, 24 (1997) S1-20.

28. J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D.M. Parkin, D. Forman, F. Bray, Cancer incidence and mortality worldwide: sources, methods

and major patterns in GLOBOCAN 2012. Int. J. Cancer, 136 (2015) E359-E386, http://globocan.iarc.fr.

29. O.M. Khatib, A. Modjtabai, Guidelines for the early detection and screening of breast cancer, World Health Organization, Regional Office for the Eastern Mediterranean, 2006.

30. E. Altobelli, A. Lattanzi, Breast cancer in European Union: An update of screening programmes as of March 2014 (Review), Int. J. Oncol., 45(2014) 1785-1792.

31. E.R. Sauter, M.B. Daly, Breast Cancer Risk Reduction and Early Detection, Springer, New York, 2010.

32. R. Lag, D. Harkins, M. Krapcho, A. Mariotto, B.A. Miller, E.J. Feuer, M. Hayat, Seer cancer statistics review, Bethesda, National Cancer Institute, 1975 1975-2003.

33. K.E. Heck, E.R. Pamuk, Explaining the relation between education and postmenopausal breast cancer, Am. J. Epidemiol., 145 (1997), 366-372;

34. N. Krieger, S.K. Van Den Eeden, D. Zava, A. Okamoto, Race/ethnicity, social class, and prevalence of breast cancer prognostic biomarkers: a study of white, black, and Asian women in the San Francisco bay area, Ethnicity & Disease, 7 (1996) 137-149.

35. J.L. Kelsey, M.D. Gammon, E.M. John, Reproductive factors and breast cancer, Epidemiol. Rev., 15 (1993) 36.

36. H.B. Nichols, A. Tretham-Dietz, J.M. Hampton, L. Titus-Ernstoff, K.M. Egan, W.C. Willet, P.A. Newcomb, From menarche to menopause: trends among US women born from 1912 to 1969. Am. J. Epidemiol., 164 (2006) 1003-1011.

37. P.A. Newcomb, B.E. Storer, M.P. Longnecker, R. Mittendorf, E.R. Greenberg, R.W. Clapp, K.P. Burke, W.C. Willett, B. MacMahon, Lactation and a reduced risk of premenopausal breast cancer, New Eng. J. Med., 330 (1994) 81-87.

38. T.J. Key, P.K. Verkasalo, Endogenous hormones and the aetiology of breast cancer. Breast Cancer Res., 1 (1999) 18-21.

39. Million Woman Study Collaborators, Breast cancer and hormone-replacement therapy in the Million Women Study, The Lancet, 362 (2003), 419-427.

40. Collaborative Group of Hormonal Factors in Breast Cancer, Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58 209 women with breast cancer and 101 986 women without the disease, The Lancet, 358 (2001) 1389-1399.

41. B. Newman, M.A. Austin, M. Lee, M.C. King, Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families, Proc. Natl. Acad. Sci. U.S.A., 85 (1988) 3044-3048.

42. R.C. Go, M.C. King, J. Bailey-Wilson, R.C. Elston, H.T. Lynch, Genetic epidemiology of breast cancer and associated cancers in high-risk families. I. Segregation analysis. J. Natl. Cancer Inst., 71 (1983) 455-461.

43. G. Lenoir, H. Lynch, P. Watson, T. Conway, J. Lynch, S. Narod, J. Feunteun, Familial breast-ovarian cancer locus on chromosome 17q12-q23, The Lancet, 338 (1991) 82-83.

44. J.M. Hall, M.K. Lee, B. Newman, J.E. Morrow, L.A. Anderson, B. Huey, M.C. King, Linkage of early-onset familial breast cancer to chromosome 17q21, Science, 250 (1990) 1684-1689.

45. D. Easton, D. Ford, J. Peto, Inherited susceptibility to breast cancer, Cancer Surveys, 18 (1992) 95-113.

46. A. Antoniou, P.D. Pharoah, S. Narod, H.A. Rish, J.E. Eyfjor, J.L. Hopper, N. Loman, H. Olsson, O. Johansson, A. Borg, B. Pasini, P. Radice, S. Manoukian, D.M. Eccles, N. Tang, E. Olah, H. Anton-Culver, E. Warner, J. Lubinski, J. Gronwald, B. Gorski, H. Tulinius, S. Thorlacius, H. Eerola, H. Nevanlinna, K. Syrjakoski, O.P. Kallioniemi, D. Thompson, C. Evans, J. Peto, F. Lalloo, D.G. Evans, D.F. Easton , Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies, Am. J. Hum. Gen., 72 (2003) 1117-1130.

47. D.W. Abbott, M.E. Thompson, C. Robinson-Benion, G. Tomlinson, R.A. Jensen, J.T. Holt, BRCA1 expression restores radiation resistance in BRCA1-defective cancer cells through enhancement of transcription-coupled DNA repair, J. Biol. Chem., 274 (1999) 18808-18812.

48. L.C. Gowen, A.V. Avrutskaya, A.M. Latour, B.H. Koller, S.A. Leadom, BRCA1 required for transcription-coupled repair of oxidative DNA damage, Science, 281 (1998) 1009-1012.

49. R. Scully, J. Chen, A. Plug, Y. Xiao, D. Weaver, J. Feuteun, T. Ashley, D.M. Livingston, Association of BRCA1 with Rad51 in mitotic and meiotic cells, Cell, 88 (1997) 265-275.

50. R.W. McDivitt, J.A. Stevensm, N.C. Lee, P.A. Wingo, G.L. Rubin, D. Gersell, Histologic types of benign breast disease and the risk for breast cancer, Cancer, 69 (1992) 1408-1414.

51. K. Polyak, Breast cancer: origins and evolution, J. Clin. Invest., 117 (2007) 3155-3163.

52. B.K. Banin Hirata, J.M. Oda, R.L. Guembarovski, C.B. Ariza, C.E. Coral de Oliveira, M.A. Watanabe, Molecular markers for breast cancer: prediction on tumor behavior. Disease Markers, 2014 (2014).

53. G. D'Eredita, C. Giardina, M. Martelotta, T. Natale, F. Ferrarese, Prognostic factors in breast cancer: the predictive value of the Nottingham Prognostic Index in patients with a long-term follow-up that were treated in a single institution, Euro. J. Cancer, 37 (2001) 591-596.

54. Ravdin, P.M. L.A. Siminoff, G.J. Davis, M.B. Mercer, J. Hewlett, N. Gerson, H.L. Parker, Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J. Clin. Oncol., 19 (2001) 980-991.

55. J.R. Benson, The TNM staging system and breast cancer, Lancet Oncol., 4 (2003) 56-60.

56. S.E. Singletary, Breast Cancer, Springer, New York, 2012.

57. A.E. Giuliano, D.M. Kirgan, J.M. Guenther, D.L. Morton, Lymphatic mapping and sentinel lymphadenectomy for breast cancer, Annals Surg., 220 (1994) 391.

58. I. De Mascarel, F. Bonichon, J.M. Coindre, M. Trojani, Prognostic significance of breast cancer axillary lymph node micrometastases assessed by two special techniques: reevaluation with longer follow-up, Br. J. Cancer, 66 (1992) 523.

59. K. Dowlatshahi, M. Fan, H.C. Snider, F.A. Habib, Lymph node micrometastases from breast carcinoma, Cancer, 80 (1997) 1188-1197.

60. I.A. Nasser, A.K.C. Lee, S. Bosari, R. Saganich, G. Heatley, M.L. Silverman, Occult axillary lymph node metastases in “node-negative” breast carcinoma, Hum. Pathol., 24 (1993) 950-957.

61. C.W. Elston, I.O. Ellis, Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up, Histopathology, 19 (1991) 403-410.

62. T.L. Smith, Breast Cancer: Current and Emerging Trends in Detection and Treatment, 1st ed., Rosen Publishing Group, New York, 2006.

63. C.H. Yarbro, D. Wujcik, B.H. Gobel, Cancer Nursing: Principles and Practice, 7th ed., Jones & Bartlett Learning, Massachusetts, 2010.

64. A. Goldhirsch, J.H. Glick, R.D. Gelber, A.S. Coates, B. Thurlimann, H.J. Senn, P. Members, Meeting highlights: international expert consensus on the primary therapy of early breast cancer 2005, Annals Oncol., 16 (2005) 1569-1583.

65. B Weigelt, J.S. Reis-Filho, Molecular profiling currently offers no more than tumour morphology and basic immunohistochemistry, Breast Cancer Res., 12 (2010) S5.

66. A. Goldhirsch, W.C. Wood, A.S. Coates, R.D. Gelber, B. Thurliman, P. Members, Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011, Annals Oncol, (2011) mdr304.

67. A.S. Caudle, A.M. Gonzalez-Angulo, K.K. Hunt, L. Pusztai, W.F. Symmans, H.M. Kuerer, E.A. Mittendorf, G.N. Hortobagyi, F. Meric-Bernstam, Predictors of tumor progression during neoadjuvant chemotherapy in breast cancer, J. Clin. Oncol., 28 (2010) 1821-1828.

68. P.A. Fasching, K. Heusinger, L. Haeberle, M. Niklos, A. Hein, C.M. Bayer, C. Rauh, R. Schulz-Wendtland, M.R. Bani, M. Schrauder, L. Kahmann, M.P. Lux, J.D. Strehl, A. Hartmann, A. Dimmler, M.W. Beckmann, D.L. Wachter, Ki67, chemotherapy response, and prognosis in breast cancer patients receiving neoadjuvant treatment, BMC cancer, 11 (2011) 486.

69. G.M. Clark, C.K. Osborne, W.L. McGuire, Correlations between estrogen receptor, progesterone receptor, and patient characteristics in human breast cancer, J. Clin. Oncol., 2 (1984) 1102-1109.

70. N. Sneige, Utility of cytologic specimens in the evaluation of prognostic and predictive factors of breast cancer: current issues and future directions, Diagn. Cytopathol., 30 (2004) 158-165.

71. M. Cianfrocca, L.J. Goldstein, Prognostic and predictive factors in early-stage breast cancer. The oncologist, 9 (2004) 606-616.

72. G. Arpino, Estrogen receptor–positive, progesterone receptor–negative breast cancer: association with growth factor receptor expression and tamoxifen resistance, J. Natl. Cancer Inst., 97 (2005) 1254-1261.

73. A.A. Onitilo, J.M. Engel, R.T. Greenlee, B.N. Mukesh, Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival, J. Clin. Med., 7 (2009) 4-13.

74. A.C. Wolff, M.E.H. Hammond, J.N. Schwartz, K.L. Hagerty, D.C. Allred, R.J. Cote, M. Dowsett, P.L. Fitzgibbons, W.M. Hanna, A. Langer, L. Mcshane, S. Paik, M.D. Pegram, E.A. Perez, M.F. Press, A. Rhodes, C. Sturgeon, S.E. Taube, R. Tubbs, G.H. Vance, M. van de Vijver, T.M. Wheeler, D.F. Hayes, American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer, J. Clin. Oncol., 25 (2006) 118-145.

75. S. Thénot, C. Henriquet, H. Rochefort, V. Cavaillès, Differential interaction of nuclear receptors with the putative human transcriptional coactivator hTIF1, J. Biol. Chem., 272 (1997) 12062-12068.

76. C.K. Osborne, H.H. Zhao, S.A. Fuqua, Selective estrogen receptor modulators: structure, function, and clinical use, J. Clin. Oncol., 18 (2000). 3172-3186.

77. M. Piccart, W.C. Wood, C.M. Hung, L.J. Solin, F. Cardoso, Breast Cancer Management and Molecular Medicine Springer Berlin Heidelberg, 2007.

78. M. Bocquel, V. Kumar, C. Stricker, P. Chambon, H. Gronemeyer, The contribution of the N-and C-terminal regions of steroid receptors to activation of transcription is both receptor and cell-specific, Nucleic Acids Res., 17 (1989) 2581-2595.

79. T. Fujita, Y, Kobayashi, O. Wada, Y. Tateishi, L. Kitada, Y. Yamamoto, H. Takashima, A. Murayama, T. Yano, T. Baba, S. Kato, Y.I. Kawabe, J. Yanagisawa, Full activation of estrogen receptor α activation function-1 induces proliferation of breast cancer cells, J. Biol. Chem., 278 (2003) 26704-26714.

80. T.J. Peterson, S. Karmakar, M.C. Pace, T. Gao, C.L. Smith, The silencing mediator of retinoic acid and thyroid hormone receptor (SMRT) corepressor is required for full estrogen receptor α transcriptional activity. Mol. Cell. Biol., 27 (2007) **27**(17) 5933-5948.

81. S.K. Gruvberger-Saal, P.O. Bendahl, L.H. Saal, M. Laakso, C. Hegardt, P. Edén, C. Peterson, P. Malmstrom, J. Isola, A. Borg, M. Ferno, Estrogen receptor β

expression is associated with tamoxifen response in ERα-negative breast carcinoma, Clin. Cancer Res., 13 (2007) 1987-1994.

82. D.C. Allred, S.J. Anderson, S. Paik, D.L. Wickerham, I.D. Nagtegaal, S.M. Swain, E.P. Mamounas, T.B. Julian, C.E. Geyer, J.P. Constantino, S.R. Land, N. Wolmark, Adjuvant tamoxifen reduces subsequent breast cancer in women with estrogen receptor–positive ductal carcinoma in situ: A study based on NSABP Protocol B-24, J. Clin. Oncol, 30 (2012) 1268-1273.

83. B. Fisher, J.P. Constantino, D.L. Wickerham, R.S. Cecchini, W.M. Cronin, A. Robidoux, T.B. Bevers, M.T. Kavanah, J.N. Atkins, R.C. Margolese, C.D. Runowicz, J.M. James, L.C. Ford, N. Wolmark, Tamoxifen for the prevention of breast cancer: current status of the National Surgical Adjuvant Breast and Bowel Project P-1 study, J. Natl. Cancer Inst., 97 (2005) 1652-1662.

84. M.C. Abba, Y. Hu, H. Sun, J.A. Drake, S. Gaddis, K. Baggerly, A. Sahin, C.M. Aldaz, Gene expression signature of estrogen receptor α status in breast cancer, BMC genomics, 6 (2005) 37.

85. A. Howell, D.J. Defriend, J.F. Roberson, R.W. Blamey, L. Anderson, E. Anderson, F.A. Sutcliffe, Pharmacokinetics, pharmacological and anti-tumour effects of the specific anti-oestrogen ICI 182780 in women with advanced breast cancer, Br. J. of Cancer, 74 (1996) 300.

86. A. Howell, J.F. Robertson, J.Q. Albano, A. Aschermannova, L. Mauriac, U.R. Kleeberg, I. Vergote, B. Erikstein, A. Webster, C. Morris, Fulvestrant, formerly ICI 182,780, is as effective as anastrozole in postmenopausal women with advanced breast cancer progressing after prior endocrine treatment. J. Clin. Oncol, 20 (2002) 3396-3403.

87. N. Elster, D.M. Collins, S. Toomey, J. Crown, A.J. Eustace, B.T. Hennessy, HER2-family signalling mechanisms, clinical implications and targeting in breast cancer, Breast Cancer Res. Treat., 149 (2015) 5-15.

88. B. Browne, N. O'Brien, M.J. Duffy, J. Crown, N, O'Donovan, HER-2 signaling and inhibition in breast cancer, Current Cancer Drug Targets, 9 (2009) 419-438.

89. D.J. Slamon, B. Leyland-Jones, S. Shak, H. Fuchs, V. Paton, A. Bajamonde, T. Fleming, W. Elermann, J. Wolter, M. Pegram, J. Baselga, L. Norton, Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2, New Eng. J. Med., 344 (2001) 783-792.

90. M.A. Molina, J. Codony-Servat, J. Albanell, F. Rojo, J. Arribas, J. Baselga, Trastuzumab (herceptin), a humanized anti-Her2 receptor monoclonal antibody, inhibits basal and activated Her2 ectodomain cleavage in breast cancer cells, Cancer Res, 61 (2001) 4744-4749.

91. C.A. Hudis, Trastuzumab—mechanism of action and use in clinical practice, New Eng. J. Med., 357 (2007) 39-51.

92. S. Johnston, A. Leary, Lapatinib: a novel EGFR/HER2 tyrosine kinase inhibitor for cancer, Drugs Today (Barc), 42 (2006) 441-453.

93. B. Weigelt, F.L. Baehner, J.S. Reis-Filho, The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. J. Pathol., 220 (2010) 263-280.

94. F.C. Geyer, J.S. Reis-Filho, Microarray-based gene expression profiling as a clinical tool for breast cancer management: are we there yet?, Int. J. Sur. Pathol, 17 (2009) 285-302.

95. C.M. Perou, T. Sorlie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey, C.A. Rees, J.R. Pollack, D.T. Ross, H. Johnsen, L.A. Akslen, O. Fluge, A. Pergamenschikov, C. Williams, S.X. Zhu, P.E. Lonning, A.L. Borresen-Dale, P.O. Brown, D. Botstein, Molecular portraits of human breast tumours, Nature, 466 (2000) 747-752.

96. T. Sørlie, C.M. Perou, R. Tibshirani, T. Aas, Geislerg, H. Johnsen, T. Hastie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey, T. Thorsen, H. Quist, J.C. Matese, P.O. Brown, D. Botstein, P.E. Lonning, A.L. Borresen-Dale, Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications, Proc. Natl. Aca. Sci., 98 (2001) 10869-10874.

97. Z. Hu, C. Fan, D.S. Oh, J.S. Marron, X. He, B. Qaqish, C. Livasy, L.A. Carey, E. Reynolds, L. Dressier, A. Nobel, J. Parker, M. Ewend, L. Sawyer, J. Wu, Y. Liu, R. Nanda, M. Tretiakova, A. Orrico, D. Dreher, J. Palazzo, L. Perreard, E. Nelson, M. Mone, H. Hansen, M. Mullins, J. Quackenbush, M. Ellis, O. Olopade, P. Bernard, C.M. Perou, The molecular portraits of breast tumors are conserved across microarray platforms, BMC genomics, 7 (2006) 96.

98. P. Farmer, H. Bonnefoi, V. Becette, M. Tubiana-Hollin, P. Fumoleau, D. Larsimont, G. MacGrogan, J. Bergh, D. Cameron, D. Goldstein, S. Duss, A.L. Nicoulaz, C. Brisken, M. Fiche, M. Delorenzi, R. Iggo, Identification of molecular apocrine breast tumours by microarray analysis, Oncogene, 24 (2005) 4660-4671.

99. A. Doane, M. Danso, P. Lai, M. Donaton, L. Zhang, C. Hudis, W.L. Gerald, An estrogen receptor-negative breast cancer subset characterized by a hormonally regulated transcriptional program and response to androgen, Oncogene, 25 (2006) 3994-4008.

100. B.T. Hennessy, A.M. Gonzalez-Angulo, K. Stemke-Hale, M.Z. Gilcrease, S. Krishnamurty, J.S. Lee, J. Fridlyand, A. Sahin, R. Agarwal, C. Joy, W. Liu, D. Stivers, K. Baggerly, M. Carey, A. Llush, C. Monteagudo, X. He, V. Weigman, C. Fan, J. Palazzo, G.N. Hortobagyi, L.K. Nolden, N.J. Wang, V. Valero, J.W. Gray, C.M. Perou, G.B. Mills, Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics, Cancer Res., 69 (2009) 4116-4124.

101. Z. Kahán, T. Tot, Breast Cancer, a Heterogeneous Disease Entity: The Very Early Stages, Springer, Netherlands, 2011.

102. M.C. Cheang, S.K. Chia, D. Voduc, D. Gao, S. Leung, J. Snider, M. Watson, S. Davies, P.S. Bernard, J.S. Parker, C.M. Perou, M. J. Ellis, T.O: Nielsen, Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer, J. Natl. Cancer Inst., 101 (2009) 736-750.

103. L.M. McManus, R.N. Mitchell, Pathobiology of Human Disease: A Dynamic Encyclopedia of Disease Mechanisms, 1st ed., Science, Elsevier, 2014.

104. L. Gianni, W. Eiermann, V. Semiglazov, A. Manikhas, A. Lluch, S. Tjulandin, M. Zambetti, F. Vazquez, M. Byakhow, M. Lichinitser, M.A. Climent, E. Ciruelos, B. Ojeda, M. Mansutti, A. Bozhok, R. Baronio, A. Feyereislova, C. Barton, P. Valagussa, J. Baselga, Neoadjuvant chemotherapy with trastuzumab followed by adjuvant trastuzumab versus neoadjuvant chemotherapy alone, in patients with HER2-positive locally advanced breast cancer (the NOAH trial): a randomised controlled superiority trial with a parallel HER2-negative cohort, The Lancet, 375 (2010) 377-384.

105. F. Yehiely, J.V. Moyano, J.R. Evans, T.O. Nielsen, V.L. Cryns, Deconstructing the molecular portrait of basal-like breast cancer, Trends Mol. Med., 12 (2006) 537-544.

106. A. Prat, B. Adamo, M.C.U Cheang, C.K. Anders, L.A. Carey, C.M. Perou, Molecular characterization of basal-like and non-basal-like triple-negative breast cancer. The oncologist, 18 (2013) 123-133.

107. Y.J. Cha, W.H. Jung, J.S. Koo, The clinicopathologic features of molecular apocrine breast cancer. Korean J. Pathol., 46 (2012) 169.

108. A. Prat, J.S. Parker, O. Karginova, C. Fan, C. Livasy, J.I. Herschkowitz, Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer, Breast Cancer Res., 12 (2010) R68.

109. L.A, Carey, E.C. Dees, L. Sawyer, L. Gatti, D.T. Moore, F. Collichio, D.W. Ollila, C.I. Sartor, M.L. Graham, C.M. Perou, The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes, Clin. Cancer Res., 13 (2007) 2329-2334.

110. D. Vyas, Comprehensive Textbook of Surgery, 1st ed., Jaypee Brothers,Medical Publishers Pvt. Limited, 2012.

111. S. Wellings, H. Jensen, R. Marcum, An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. J. Natl. Cancer Inst., 55 (1975) 231-273.

112. F.A. Tavassoli, P. Devilee, Pathology and Genetics of Tumours of the Breast and Female Genital Organs, IARC Press, Lyon, 2003.

113. C. Friedenreich, H.E. Bryant, F. Alexander, J. Hugh, J. Danyluk, D.L. Page, Risk factors for benign proliferative breast disease, Int. J. Epidemiol., 29 (2000) 637-644.

114. J.G. Elmore, G. Gigerenzer, Benign breast disease-the risks of communicating risk. New Eng. J. Med., 353 (2005) 297-299.

115. S.W. Dyrstad, Y. Yan, A.M. Fowler, G.A. Colditz, Breast cancer risk associated with benign breast disease: systematic review and meta-analysis, Breast cancer Res. Treat., 149 (2015) 569-575.

116. V. Kumar, A.K. Abbas, N. Fausto, J.C. Aster, Robbins and Cotran Pathologic Basis of Disease, 9th ed., Professional Edition, Elsevier Health Sciences, 2009.

117. K. Kerlikowske, Epidemiology of ductal carcinoma in situ. J. Natl. Cancer Inst. Monographs, 2010 (2010) 139.

118. M.A. Lopez-Garcia, F.C. Geyer, M. Lacroix-Triki, C. Marchió, J.S. Reis-Filho, Breast cancer precursors revisited: molecular features and progression pathways, Histopathology, 57 (2010) 171-192.

119. D.J. Dabbs, Breast Pathology, Elsevier/Saunders, Philadelphia, 2012.

120. L.G. Bobrow, L.C. Happerfield, W.M. Gregory, R.D. Springall, R.R. Millis, The classification of ductal carcinoma in situ and its association with biological markers, Sem. Diag. Pathol., 11 (1994) 199-207.

121. R. Holland, J.L. Peterse, R.R. Millis, V. Eusebi, D. Faverly, M.J. van de Vijver, B. Zafrani, Ductal carcinoma in situ: a proposal for a new classification, Sem. Diag. Pathol., 11 (1994) 167-180.

122. D. Poller, A. Barth, D.J. Slamon, M.J. Silverstein, E.D. Gierson, W.J. Coburn, J.R. Waisman, P. Gamagami, B.S. Lewinsky, Prognostic classification of breast ductal carcinoma-in-situ, The lancet, 345 (1995) 1154-1157.

123. M. Silverstein, A. Barth, D. Poller, E.D. Gierson, W.J. Colburn, J.R. Waisman, P. Gamagami, Ten-year results comparing mastectomy to excision and radiation therapy for ductal carcinoma in situ of the breast, Eur. J. Cancer, 31 (1995) 1425-1427.

124. C.K. Kuhl, S. Schrading, H.B. Bieling, E. Warderlmann, C.C. Leuther, R. Koening, W. Kuhn, H.H. Schild, MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. The Lancet, 370 (2007) 485-492.

125. N. Sneige, M.D. Lagios, R. Schwarting, W. Colburn, E. Atkinson, D. Weber, A. Sahin, B. Kemp, A. Hoque, S. Risin, A. Sabichi, C. Boone, Interobserver reproducibility of the Lagios nuclear grading system for ductal carcinoma in situ, Hum. Pathol., 30 (1999) 257-262.

126. S.H Cho, A. Sahin, G.N. Hortobagyi, W.N. Hittelman, K. Dhingra, Sialyl-Tn antigen expression occurs early during human mammary carcinogenesis and is associated with high nuclear grade and aneuploidy, Cancer Res., 54 (1994) 6302-6305.

127. A. Douglas-Jones, S.K. Gupta, R.L. Attanoos, J.M. Morgan, R.E. Mansel, A critical appraisal of six modern classifications of ductal carcinoma in situ of the breast (DCIS): correlation with grade of associated invasive carcinoma, Histopathology, 29 (1996) 397-409.

128. M.J. van de Vijver, J.L. Peterse, W.J. Mooi, P. Wisman, J. Lomans, O. Dalesio, R. Nusse, Neu-protein overexpression in breast cancer, New Eng. J. Med., 319 (1988) 1239-1245.

129. D. Giri, S.A. Dundas, J.F. Nottingham, J.C. Underwood, Oestrogen receptors in benign epithelial lesions and intraduct carcinomas of the breast: an immunohistological study, Histopathology, 15 (1989) 575-584.

130. H.J. Helin, M.J. Helle, O.P. Kallioniemi, J.J. Isola, Immunohistochemical determination of estrogen and progesterone receptors in human breast carcinoma. Correlation with histopathology and DNA flow cytometry, Cancer, 63 (1989) 1761-1767.

131. A.J. Karayiannakis, E.A. Bastounis, E.B. Chatzigianni, G.G. Makri, D. Alexiou, P. Karamanakos, Immunohistochemical detection of oestrogen receptors in ductal carcinoma in situ of the breast, Eur. J. Surg. Oncol. (EJSO), 22 (1996) 578-582.

132. B. Zafrani, M.H. Aubriot, E. Mouret, P. De Cremoux, Y. De Rycke, A. Nicolas, E. Boudou, A. Vincent-Salomon, H. Magdelénat, X. Sastre-Garau, High sensitivity and specificity of immunohistochemistry for the detection of hormone receptors in breast carcinoma: comparison with biochemical determination in a prospective study of 793 cases, Histopathology, 37 (2000) 536-545.

133. C.B. Leal, F.C. Schmitt, M.J. Bento, N.C. Maia, C.S. Lopes, Ductal Carcinoma in-situ of the breast-histologic categorization and its relationship to ploidy and immunohistochemical expression of hormone receptors, p53, and c-erbb-2 protein, Cancer, (1995) 2123-2131.

134. L. Pallis, N. Wilking, B. Cedemark, L.E. Rutqvist, L. Skoog, Receptors for estrogen and progesterone in breast carcinoma in situ, Anticancer Res., 12 (1991) 2113-2115.

135. D. Poller, E.C. Roberts, J.A. Bell, C.W. Elston, R.W. Blamey, p53 protein expression in mammary ductal carcinoma in situ: relationship to immunohistochemical expression of estrogen receptor and c-erbB-2 protein, Hum. Pathol., 24 (1993) 463-468.

136. B. Zafrani, A. Leroyer, A. Fourquet, M. Laurent, D. Trophilme, P. Validire, X. Sastre-Garau, Mammographically-detected ductal in situ carcinoma of the breast analyzed with a new classification. A study of 127 cases: correlation with estrogen and progesterone receptors, p53 and c-erbB-2 proteins, and proliferative activity, Sem. Diag. Pathol.. 11 (1994) 208-214.

137. D. Poller, M.G. Silverstein, M. Galea, A.P. Locker, C.W. Elston, R.W. Blamey, I.O. Ellis, Ideas in pathology. Ductal carcinoma in situ of the breast: a proposal for a new simplified histological classification association between cellular proliferation and c-erbB-2 protein expression, Mod. Pathol., 7 (1994) 257-262.

138. A.P Locker, C. Horrocks, A.S. Gilmour, I.O. Ellis, C.S. Dowle, C.W. Elston, R.W. Blarney, Flow cytometric and histological analysis of ductal carcinoma in situ of the breast, Br. J. Surg., 77 (1990) 564-567.

139. L.J. Solin, R. Gray, F.L. Baehner, S.M. Butler, L.L. Hughes, C. Yoshizawa, D.B. Cherbavaz, S. Shak, D.L. Page, G.W. Sledge, N.E. Davidson, J.N. Ingle, E.A. Perez, W.C. Wood, J.A. Sparano, S. Badve, A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ of the breast, J. Natl Cancer Inst., (2013) djt067.

140. C.F. Cowell, B. Weigelt, R.A. Sakr, C.K.Y. Niang, J. Hicks, T.A. King, J.S. Reis-Filho, Progression from ductal carcinoma in situ to invasive breast cancer: revisited, Mol. Oncol., 7 (2013) 859-869.

141. F. Dirbas, C. Scott-Conner, Breast Surgical Techniques and Interdisciplinary Management, Springer, New York, 2011.

142. A.E. Giuliano, R.C. Jones, M. Brennan, R. Statman, Sentinel lymphadenectomy in breast cancer, J. Clin. Oncol., 15 (1997) 2345-2350.

143. H.M. Kuerer, L.A. Newman, T.L. Smith, F.C. Ames, K.K. Hunt, K. Dhingra, R.L. Theriault, G. Singh, S.M. Binkley, N. Sneige, T.A. Buchholz, M.I. Ross, M.D. Mcneese, A.U. Buzdar, G.N. Hortobagyi, S.E. Singletary, Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy, J. Clin. Oncol., 17 (1999) 460-460.

144. A.D. Corben, Pathology of invasive breast disease, Surg. Clin. North Am., 93 (2013) 363-392.

145. S. Becker, A historic and scientific review of breast cancer: The next global healthcare challenge, Int. J. Gynecol. Obst., 131 (2015) S36-S39.

146. J. Greene, B. Hennessy, The role of anthracyclines in the treatment of early breast cancer, J. Oncol. Pharm. Practice, (2014) 1078155214531513.

147. Early Breast Cancer Trialist's Collaborative Group, Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials, The lancet, 378 (2011) 771-784.

148. A. Bombonati, D.C. Sgroi, The molecular pathology of breast cancer progression, J. Pathol, 223 (2011) 308-318.

149. S. Wellings, H. Jensen, On the origin and progression of ductal carcinoma in the human breast, J. Natl. Cancer Inst., 50 (1973) 1111-1118.

150. S. Wellings, V. Jentoft, Organ cultures of normal, dysplastic, hyperplastic, and neoplastic human mammary tissues, J. Natl. Cancer Inst., 49 (1972) 329-338.

151. T. Oyama, H. Maluf, F. Koerner, Atypical cystic lobules: an early stage in the formation of low-grade ductal carcinoma in situ, Virchows Archiv, 435 (1999) 413-421.

152. T. Oyama, H. Takei, J. Horigushi, Y. Lino, T. Nakajima, F. Koerner, Atypical cystic lobule of the breast: An early stage of low-grade ductal carcinomain-situ, Breast Cancer, 7 (2000) 326-331.

153. M.F. Lerwill, Flat epithelial atypia of the breast, Arch. Pathol. Lab. Med., 132 (2008) 615-621.

154. D.L. Page, Atypical hyperplastic lesions of the female breast. A long-term follow-up study, Cancer, (1959).

155. W.D. Dupont, D.L. Page, Risk factors for breast cancer in women with proliferative breast disease, New Eng. J. Med., 312 (1985) 146-151.

156. R. Kusama, M. Fujimori, I. Matsuyama, L. Fu, K. Ishii, Y. Hama, K. Asanuma, K. Shingu, S. Kobayashi, S.I. Tsuchiya, Clinicopathological characteristics of atypical cystic duct (ACD) of the breast: assessment of ACD as a precancerous lesion, Pathol. Int., 50 (2000) 793-800.

157. F. Otterbach, A. Bankfalvi, S. Bergner, T. Decker, R. Krech, W. Boeker, Cytokeratin 5/6 immunohistochemistry assists the differential diagnosis of atypical proliferations of the breast, Histopathology, 37 (2000) 232-240.

158. W. Boecker, R. Moll, P. Dervan, H. Buerger, C. Poremba, R.I. Diallo, H. Herbst, A. Schmidt, M.M. Lerch, I.B. Buchwalow, Usual ductal hyperplasia of the breast is a committed stem (progenitor) cell lesion distinct from atypical ductal hyperplasia and ductal carcinoma in situ, J. Pathol., 198 (2002) 458-467.

159. J.S. Reis-Filho, P.T. Simpson, T. Gale, S.R, Lakhani, The molecular genetics of breast cancer: the contribution of comparative genomic hybridization, Pathol. Res. Pract., 201 (2005) 713-725.

160. P.T. Simpson, J.S. Reis-Filho, T. Gale, S.R. Lakhani, Molecular evolution of breast cancer. J. Pathol., 205 (2005) 248-254.

161. R.F. Bonner, M. Emmert-Buck, K. Cole, T. Pohida, R. Chuaqui, S. Goldstein, L.A. Liotta, Laser capture microdissection: molecular analysis of tissue. Science, 278 (1997) 1481.

162. M. Tirkkonen, M. Tanner, R. Karhu, A. Kalloniemi, J. Isola, O.P. Kalloniemi, Molecular cytogenetics of primary breast cancer by CGH. Genes Chr. Cancer, 21 (1998) 177-184.

163. R. Roylance, P. Gorman, W. Harris, R. Liebmann, D. Barnes, A. Hanby, D. Sheer, Comparative genomic hybridization of breast tumors stratified by histological grade reveals new insights into the biological progression of breast cancer, Cancer Res., 59 (1999) 1433-1436.

164. H. Buerger, F. Otterbach, R. Simon, K.L. Shafer, C. Poremba, R. Diallo, C. Brinkschmidt, B. Dockhorn-Dworniczak, W. Boecker, Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes, J. Pathol, 189 (1999) 521-526.

165. X.J. Ma, R. Salunga, J.T. Tuggle, J. Gaudet, E. Enright, P. McQuary, Y.X. Zhou, Gene expression profiles of human breast cancer progression, Proc. Natl. Aca. Sci., 100 (2003) 5974-5979.

166. R.L. Balleine, L.R. Webster, S. Davis, E.L. Salisbury, J.P. Palazzo, G.F. Schwartz, D.B. Cornfield, R.L. Walker, K. Byth, C.L. Clarke, P.S. Meltzer, Molecular grading of ductal carcinoma in situ of the breast, Clin. Cancer Res., 14 (2008) 8244-8252.

167. B. Weigelt, J.S. Reis-Filho, Histological and molecular types of breast cancer: is there a unifying taxonomy?, Nature Rev. Clin. Oncol., 6 (2009) 718-730.

168. R. Natrajan, M.B. Lambros, S.M. Rodriguez-Pinilla, G. Moreno-Bueno, D. Tan, C. Marchió, R. Vatcheva, S. Rayter, B. Mahler-Araujo, L.G. Fulford, D. Hungermann, A. Mackay, A. Grigoriadis, K. Fenwick, N. Tamber, D. Hardisson, A. Tutt, J. Palacios,

C.J. Lord, H. Buerger, A. Ashworth, J.S. Reis-Filho, Tiling path genomic profiling of grade 3 invasive ductal breast cancers, Clin. Cancer Res., 15 (2009) 2711-2722.

169. A.M. Cleton-Jansen, H. Buerger, N. Haar, K. Philippo, M.J. van de Vijver, W. Boecker, V. Smit, C.J. Cornelisse, Different mechanisms of chromosome 16 loss of heterozygosity in well-versus poorly differentiated ductal breast cancer, Genes Chr. Cancer, 41 (2004) 109-116.

170. H. Buerger, F. Otterbach, R. Simon, C. Poremba, R. Diallo, T. Decker, L. Riethdorf, C. Brinkschmidt, B. Dockhorn-Dworniczak, W. Boecker, Comparative genomic hybridization of ductal carcinoma in situ of the breast—evidence of multiple genetic pathways, J. Pathol., 187 (1999) 396-402.

171. R. Natrajan, B. Weigelt, A. Mackay, F.C. Geyer, A. Grigoriadis, D. Tan, C. Jones, C.J. Lord, R. Vatcheva, S.M. Rodriguez-Pinilla, J. Palacios, A. Ashworth, J.S. Resi-Filho, An integrative genomic and transcriptomic analysis reveals molecular pathways and networks regulated by copy number aberrations in basal-like, HER2 and luminal cancers, Breast Cancer Res. Treat, 121 (2010) 575-589.

172. Y. Nikolsky, E. Sviridov, J. Yao, D. Dosymbekov, V. Ustyansky, V. Kaznacheev, Z. Dezso, L. Mulvey, L.E. Macconaill, W. Winckler, T. Serebryiskaya, T. Nikolskaya, K. Polyak, Genome-wide functional synergy between amplified and mutated genes in human breast cancer, Cancer Res., 68 (2008) 9532-9540.

173. J. Hicks, A. Krasnitz, N.E. Navin, M. Riggs, E. Leibu, D. Esposito, J. Alexander, J. Troge, V. Crubor, S. Yoon, M. Wigler, K. Ye, A.L. Borresen-Dale, B. Naume, E. Schlicting, L. Norton, T. Hagerstrom, L. Skoog, G. Auer, S. Manér, P. Lundin, A. Zetterberg, Novel patterns of genome rearrangement and their association with survival in breast cancer, Genome Res., 16 (2006) 1465-1479.

174. K. Chin, K, S. Devries, J. Fridly, P.T. Spellman, R. Roydasgupta, W.L. Kuo, A. Lapuk, R.M. Neve, Z. Qian, T. Ryder, F. Chen, H. Feiler, T. Tokuyaso, C. Kingsley, S. Dairkee, Z. Meng, K. Chew, D. Pinkel, A. Jain, B.M. Ljung, L. Esseman, D.G. Albertson, F.M. Waldman, J. W. Gray, Genomic and transcriptional aberrations linked to breast cancer pathophysiologies, Cancer Cell, 10 (2006) 529-541.

175. L. Burkhardt, T.J. Grob, I. Hermann, E. Burandt, M. Choschzick, F. Janicke, V. Muller, C. Bockemeyer, R. Simon, G. Sauter, W. Wilczak, A. Lebeau, Gene amplification in ductal carcinoma in situ of the breast, Breast Cancer Res. Treat., 123 (2010) 757-765.

176. P. O'Connell, V. Pekkel, D.C. Allred, Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci, J. Natl. Cancer Inst., 90 (1998) 697-703.

177. A. Evans, S.E. Pinder, D.R. Snead, A.R. Wilson, I.O. Ellis, C.W. Elston, The detection of ductal carcinoma in situ at mammographic screening enables the diagnosis of small, grade 3 invasive tumours, Br. J. Cancer, 75 (1997) 542.

178. E.R. Fisher, R.M. Gregorio, B. Fisher, F. Vellios, S.C. Sommers, The pathology of invasive breast cancer A Syllabus Derived from Findings of the National Surgical Adjuvant Breast Project (Protocol No. 4), Cancer, 36 (1975) 1-85.

179. T.M. Abdel-Fatah, D.G. Powe, Z. Hodi, A. Lee, J.S. Reis-Filho, I.O. Ellis, High frequency of coexistence of columnar cell lesions, lobular neoplasia, and low grade ductal carcinoma in situ with invasive tubular carcinoma and invasive lobular carcinoma, Am. J. Surg. Pathol., 31 (2007) 417-426.

180. B.B. Bryan, S.J. Schnitt, L.C. Collins, Ductal carcinoma in situ with basal-like phenotype: a possible precursor to invasive basal-like breast cancer, Mod. Pathol., 19 (2006) 617-621.
181. S. Clark, J. Warwick, R. Carpenter, R.L. Bowen, S.W. Duffy, J.L. Jones, Molecular subtyping of DCIS: heterogeneity of breast cancer reflected in pre-invasive disease, Br. J. of Cancer, 104 (2011) 120-127.

182. C.A. Livasy, C.M. Perou, G. Karaca, D.W. Cowan, D. Maia, S. Jackson, C.K. Tse, S. Nyante, R.C. Millikan, Identification of a basal-like subtype of breast ductal carcinoma in situ, Hum. Pathol., 38 (2007) 197-204.

183. M.E. Sanders, P.A. Schuyler, W.D. Dupont, D.L. Page, The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up, Cancer, 103 (2005) 2481-2484.

184. S. Steinman, J. Wang, P. Bourne, Q. Yang, P. Tang, Expression of cytokeratin markers, ER-alpha, PR, HER-2/neu, and EGFR in pure ductal carcinoma in situ (DCIS) and DCIS with co-existing invasive ductal carcinoma (IDC) of the breast, Annals Clin. Lab. Sci., 37 (2007) 127-134.

185. R.M. Tamimi, H.J. Baer, J. Marotti, M. Galan, L. Galaburda, Y. Fu, A.C. Deitz, J.L. Connolly, S.J. Schnitt, G.A. Golditz, L.C. Collins, Comparison of molecular phenotypes of ductal carcinoma in situ and invasive breast cancer, Breast Cancer Res., 10 (2008) R67.

186. E. Latta, S. Tjan, R.K. Parkes, F.P. O'Malley, The role of HER2/neu overexpression/amplification in the progression of ductal carcinoma in situ to invasive carcinoma of the breast, Mod. Pathol., 15 (2002) 1318-1325.

187. Park, K, S. Han, H.J. Kim, E. Shin, HER2 status in pure ductal carcinoma in situ and in the intraductal and invasive components of invasive ductal carcinoma determined by fluorescence in situ hybridization and immunohistochemistry, Histopathology, 48 (2006) 702-707.

188. M. Gerlinger, A.J. Rowan, S. Horswell, J. Larkin, D. Endesfelder, E. Gronroos, P. Martinez, N. Matthews, A. Stuart, P. Tarpey, I. Varela, B. Phillimore, S. Begum, N.Q. McDonald, A. Butler, D. Jones,K. Raine, C. Latimer, C.R. Santos, N. Nohadani, A.C. Eklund, B. Spencer-Dene, G. Clark, L. Pickering, G. Stamp, M. Gore, Z. Szallasi, J. Downward, A. Futreal, C. Swanton, Intratumor heterogeneity and branched evolution revealed by multiregion sequencing, New Eng. J. Med., 366 (2012) 883-892.

189. M. Allinen, R. Beroukhim, L. Cai, C. Brenann, J. Lahti-Domenici, H. Huang, D. Porter, M. Hu, A. Richardson, S. Schnitt, W.R. Sellers, K. Polyak, Molecular characterization of the tumor microenvironment in breast cancer, Cancer Cell, 6 (2004) 17-32.

190. X.J. Ma, S. Dahiya, E. Richardson, M. Erlander, D.C. Sgroi, Gene expression profiling of the tumor microenvironment during breast cancer progression, Breast Cancer Res., 11 (2009) 1-18.

191. A.C. Vargas, A.E.M.C. Reed, N. Waddell, A. Lane, L.E. Reid, C.E. Smart, S. Cocciardi, L. da Silva, S. Song, G. Chenevix-Trench, P.T. Simpson, S.R. Lakhani, Gene expression profiling of tumour epithelial and stromal compartments during breast cancer progression, Breast Cancer Res. Treat., 135 (2012) 153-165.

192. M. Hu, J. Yao, D.K. Carroll, S. Weremowicz, H. Chen, D. Carrasco, A. Richardson, S. Violette, T. Nikolskaya, Y. Nikolsky, E.L. Bauerlein, W.C. Hahn, R.S. Gelman, C. Alfred, M.J. Bissell, S. Schnitt, K. Polyak, Regulation of in situ to invasive breast carcinoma transition, Cancer Cell, 13 (2008) 394-406.

193. S.H. Barsky, N.J. Karlin, Myoepithelial cells: autocrine and paracrine suppressors of breast cancer progression. J. Mammary Gland Biol. Neoplasia, 10 (2005) 249-260.

194. T. Gudjonsson, M.C. Adriance, M.D. Sternlicht, O.W. Petersen, M.J. Bissell, Myoepithelial cells: their origin and function in breast morphogenesis and neoplasia, J.Mammary Gland Biol. Neoplasia, 10 (2005) 261-272.

195. A.E. Place, S. Jin Huh, K. Polyak, The microenvironment in breast cancer progression: biology and implications for treatment. Breast Cancer Res., 13 (2011) 227.

196. K. Polyak, M. Hu, Do myoepithelial cells hold the key for breast tumor progression? J. Mammary Gland Biol. Neoplasia, 10 (2005) 231-247.

197. S.H. Barsky, N.J. Karlin, Mechanisms of disease: breast tumor pathogenesis and the role of the myoepithelial cell, Nature Clin. Pract. Oncol., 3 (2006) 138-151.

198. M.D. Sternlicht, P. Kedeshian, Z.M. Shao, S. Safarians, S.H. Barsky, The human myoepithelial cell is a natural tumor suppressor, Clin. Cancer Res., 3 (1997) 1949-1958.

199. C.B. Moelans, A.H. Verschuur-Maes, P.J. van Diest, Frequent promoter hypermethylation of BRCA2, CDH13, MSH6, PAX5, PAX6 and WT1 in ductal carcinoma in situ and invasive breast cancer, J. Pathol., 225 (2011) 222-231.

200. S.Y. Park, H.J. Kwon, H.E. Lee, H.S. Ryu, S.W. Kim, J.H. Kim, I.A. Kim, N. Jung, N.Y. Cho, G.H. Kang, Promoter CpG island hypermethylation during breast cancer progression, Virchows Archiv, 458 (2011) 73-84.

201. A.H. Verschuur-Maes, P.C. de Bruin, P.J. van Diest, Epigenetic progression of columnar cell lesions of the breast to invasive breast cancer, Breast Cancer Res. Treat., 136 (2012) 705-715.

202. O.G. McDonald, H. Wu, W. Timp, A. Doi, A.P. Feinberg, Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition, Nat. Struct. Mol. Biol., 18 (2011) 867-874.

203. E.S. Knudsen, A. Ertel, E. Davicioni, J. Kline, G.F. Schwartz, A.K. Witkiewicz, Progression of ductal carcinoma in situ to invasive breast cancer is associated with

gene expression programs of EMT and myoepithelia, Breast Cancer Res. Treat., 133 (2012) 1009-1024.

204. L. Hernandez, P.M. Wilkerson, M.B. Lambros, A. Campion-Flora, D.N. Rodrigues, A. Gauthier, C. Cabral, V. Pawar, A. Mackay, R. A'Hern, C. Marchió, J. Palacios, R. Natrajan, B. Weigelt, J.S. Reis-Filho, Genomic and mutational profiling of ductal carcinomas in situ and matched adjacent invasive breast cancers reveals intra-tumour genetic heterogeneity and clonal selection, J. Pathol., 227 (2012) 42-52.

205. K. Heselmeyer-Haddad, L.Y.B. Garcia, A. Bradley, C. Ortiz-Melendez, W.J. Lee, R. Christensen, S.A. Prindville, K.A. Calzone, P.W. Soballe, Y. Hu, S.A. Chowdhury, R. Schwartz, A.A. Schaffer, T. Ried, Single-cell genetic analysis of ductal carcinoma in situ and invasive breast cancer reveals enormous tumor heterogeneity yet conserved genomic imbalances and gain of MYC during progression. Am. J. Pathol., 181 (2012) 1807-1822.

206. N.C. Turner, Reis-Filho, *Genetic heterogeneity and cancer drug resistance.* Lancet Oncol., 13 (2012) e178-e185.

207. T.A. Yap, M. Gerlinger, P.A. Futreal, L. Pusztai, C. Swanton, Intratumor heterogeneity: seeing the wood for the trees, Sci. Transl. Med., 4 (2012) 127ps10-127ps10.

208. D. Hanahan, R.A. Weinberg, The hallmarks of cancer. cell, 100 (2000) 57-70.

209. D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell, 144 (2011) 646-674.

210. M.R. Stratton, P.J. Campbell, P.A. Futreal, The cancer genome, Nature, 458 (2009) 719-724.

211. P.J. Stephens, P.S. Tarpey, H. Davies, P. Van Loo, C. Greenman, D.C. Wedge, S. Nik-Zainal, S. Martin, I. Varela, G.R. Bignell, L.R. Yates, E. Papaemmanuil, D. Beare, A. Butler, A. Cheverton, J. Gamble, J. Hinton, M. Jia, A. Jayakumar, D. Jones, C. Latimer, K. Wai Lau, S. McLaren, D.J. McBride, A. Menzies, M.R. Strattton, The landscape of cancer genes and mutational processes in breast cancer, Nature, 486 (2012) 400-404.

212. S.P. Shah, A. Roth, R. Goya, A. Oloumi, G. Ha, Y. Zhao, G. Turashvili, J. Ding, K. Tse, G. Haffari, A. Bashashati, L.M. Prentice, J. Khattra, A. Burleigh, D. Yap, V. Bernard, A. McPherson, K. Shumansky, A. Crisan, R. Giuliany, A. Heravi-Moussavi, J. Rosner, D. Lai, I. Birol, R. Varhol, S. Aparicio,The clonal and mutational evolution spectrum of primary triple-negative breast cancers, Nature, 486 (2012) 395-399.

213. M. Arnedos, C. Vicier, S. Loi, C. Lefebvre, S. Michiels, H. Bonnefoi, F. Andre, Precision medicine for metastatic breast cancer [mdash] limitations and solutions, Nat. Rev. Clin. Oncol., 12 (2015) 693-704.

214. C.L. Arteaga, Progress in breast cancer: overview, Clin. Cancer Res., 19 (2013) 6353-6359.

215. C.A. Pinto, E. Widodo, M. Waltham, E.W. Thompson, Breast cancer stem cells and epithelial mesenchymal plasticity–implications for chemoresistance, Cancer lett., 341 (2013) 56-62.

216. A. Marusyk, K. Polyak, Tumor heterogeneity: causes and consequences. B.B.A.-Rev. Cancer, 1805 (2010) 105-117.

217. L.L. Campbell, K. Polyak, Breast tumor heterogeneity: cancer stem cells or clonal evolution? Cell Cycle, 6 (2007) 2332-2338.

218. L.G. Martelotto, C.K. Ng, S. Piscuoglio, B. Weigelt, J.S Reis-Filho, Breast cancer intra-tumor heterogeneity, Breast Cancer Res., 16 (2014) 210.

219. M. Greaves, C.C. Maley, Clonal evolution in cancer. Nature, 481 (2012) 306-313.

220. M.F. Clarke, J.E. Dick, P.B. Dirks, C.J. Eaves, C.H.M. Jamieson, D.L. Jones, J. Visvader, I.L. Weissman, G.M. Wahl, Cancer stem cells—perspectives on current status and future directions: AACR Workshop on cancer stem cells, Cancer Res., 66 (2006) 9339-9344.

221. L. Vermeulen, M.R. Spricks, K. Kemper, G. Stassi, J.P. Medema, Cancer stem cells–old concepts, new insights, Cell Death Diff., 15 (2008) 947-958.

222. L. Vermeulen, F.S. Melo, D.J Richel, J.P. Medema, The developing cancer stem-cell model: clinical challenges and opportunities, Lancet Oncol., 13 (2012) e83-e89.

223. H.E. Lee, J.H. Kim, Y.J. Kim, S.Y. Choi, S.W. Kim, E. Kang, I.Y. Chung, I.A. Kim, Y. Choi, H.S. Ryu, S.Y. Park, An increase in cancer stem cell population after primary systemic therapy is a poor prognostic factor in breast cancer, Br. J. Cancer, 104 (2011) 1730-1738.

224. M. Al-Hajj, M.S. Whicha, A. Bentio-Hernandez, S.J. Morrison, M.F. Clarke, Prospective identification of tumorigenic breast cancer cells, Proc. Natl. Aca. Sci., 100 (2003) 3983-3988.

225. P.B. Gupta, C.M. Filmore, G. Jiang, S.D. Shapira, K. Tao, C. Kuperwasser, E.S. Lander, Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells, Cell, 146 (2011) 633-644.

226. A.D. Donnenberg, J.B. Hicks, M. Wigler, V.S. Donnenberg, The cancer stem cell: cell type or cell state? Cytometry A, 83 (2013) 5-7.

227. K. Anderson, C. Lutz, F.W. Van Delft, C.M. Bateman, Y. Guo, S.M. Colman, H. Kempski, A.V. Moorman, I. Titley, J. Swansbury, L. Kearney, T. Enver, M. Greaves, Genetic variegation of clonal architecture and propagating cells in leukaemia, Nature, 469 (2011) 356-361.

228. H. Clevers, The cancer stem cell: premises, promises and challenges, Nat. Med., (2011) 313-319.

229. G.S. Charames, B. Bapat, Genomic instability and cancer, Curr. Mol. Med., 3 (2003) 589-596.

230. B. Vogelstein, K.W. Kinzler, Cancer genes and the pathways they control, Nat. Med., 10 (2004) 789-799.

231. N.P. Shah, B.J. Skaggs, S. Branford, T.P. Hughes, J.M. Nicholl, R.L. Paquette, C.L. Sawyers, Sequential ABL kinase inhibitor therapy selects for compound drug-resistant BCR-ABL mutations with altered oncogenic potency, J. Clin. Inv., 117 (2007) 2562-2569.

232. B. Calabretta, D. Perrotti, The biology of CML blast crisis. Blood, 103 (2004) 4010-4022.

233. C.E. Meacham, S.J. Morrison, Tumour heterogeneity and cancer cell plasticity. Nature, 501 (2013) 328-337.

234. C.M. Fillmore, C. Kuperwasser, Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy, Breast Cancer Res., 10 (2008) R25.

235. R. Swords, J. Quinn, M. Fay, R. O'Donnel, J. Goldman, P.T. Murphy, CML clonal evolution with resistance to single agent imatinib therapy, Clin. Lab. Haematol., 27 (2005) 347-349.

236. M. Todaro, M.P. Alea, A.B. Di Stefano, P. Cammareri, L. Vermeulen, F. Lovino, C. Tripodo, A. Russo, G. Gulotta, J.P. Medema, G. Stassi, Colon cancer stem cells dictate tumor growth and resist cell death by production of interleukin-4, Cell stem Cell, 1 (2007) 389-402.

237. J.S. Burns, B.M. Abdallah, P. Guldberg, J. Rygaard, H.D. Schroder, M. Kassem, Tumorigenic heterogeneity in cancer stem cells evolved from long-term cultures of telomerase-immortalized human mesenchymal stem cells, Cancer Res., 65 (2005) 3126-3135.

238. G.V. Røsland, A. Svendesen, A. Torsvik, E. Sobala, E. McCormack, H. Immervoll, J. Mysliwietz, J.C. Tonn, R. Goldbrunner, P.E. Lonning, R. Bjerkvig, C. Schichor, Long-term cultures of bone marrow–derived human mesenchymal stem cells frequently undergo spontaneous malignant transformation, Cancer Res., 69 (2009) 5331-5339.

239. C. Shi, Y. Mai, Y. Zhu, T. Cheng, Y. Su, Spontaneous transformation of a clonal population of dermis-derived multipotent cells in culture, In Vitro Cell. Dev. Biol., 43 (2007) 290-296.

240. F.A. Siebzehnrubl, I. Jeske, D. Muller, R. Buslei, R. Coras, E. Hahnen, H.B. Huttner, D. Corbeil, J. Kaesbauer, T. Appl, S. Von Horsten, I. Blumcke, Spontaneous in vitro transformation of adult neural precursors into stem-like cancer cells, Brain Pathol., 19 (2009) 399-408.

241. D. Rubio, J. Garcia-Castro, M.C. Martin, R. de la Fuente, J.C. Cigudosa, A.C. Lloyd, A. Bernard, Spontaneous human adult stem cell transformation, Cancer Res., 65 (2005) 3035-3039.

242. F. Islam, B. Qiao, R.A. Smith, V. Gopalan, A. Lam, Cancer stem cell: Fundamental experimental pathological concepts and updates, Exp. Mol. Pathol., 98 (2015) 184-191.

243. S.A. Mani, W. Guo, M.J. Liao, E.N. Eaton, A. Ayyanan, A.Y. Zhou, M. Brooks, F. Reinhard, C.C. Zhang, M. Shipitsin, L.L. Campbell, K. Polyak, C. Brisken, J. Yang, R.A. Weinberg, The epithelial-mesenchymal transition generates cells with properties of stem cells, Cell, 133 (2008) 704-715.

244. K. Polyak, R.A. Weinberg, Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits, Nat. Rev. Cancer, 9 (2009) 265-273.

245. C.H. Heldin, M. Landström, A. Moustakas, Mechanism of TGF-β signaling to growth arrest, apoptosis, and epithelial–mesenchymal transition, Curr. Op. Cell Biol., 21 (2009) 166-176.

246. K. Takahashi, S. Yamanaka, Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors, Cell, 126 (2006) 663-676.

247. I. Malanchi, A. Santamaria-Matinez, E. Susanto, H. Peng, H.A. Lehr, J.F. Delaloye, J. Huelsken, Interactions between cancer stem cells and their niche govern metastatic colonization, Nature, 481 (2012) 85-89.

248. C. Odoux, H. Fohrer, T. Hoppo, L. Guzik, D.B. Stolz, D.W. Lewis, S.M. Gollin, T.C. Gamblin, D.A. Geller, E. Lagasse, A stochastic model for cancer stem cell origin in metastatic colon cancer, Cancer Res., 68 (2008) 6932-6941.

249. G.H. Heppner, Tumor heterogeneity, Cancer Res., 44 (1984) 2259-2265.

250. S. Chen, E.H. Huang, The colon cancer stem cell microenvironment holds keys to future cancer therapy, J. Gastro. Surg., 18 (2014) 1040-1048.

251. F. Pajonk, E. Vlashi, Characterization of the stem cell niche and its importance in radiobiological response, Sem. Rad. Oncol., 23 (2013) 237-241.

252. I. Malanchi, J. Huelsken, Cancer stem cells: never Wnt away from the niche, Curr. Op. Oncol., 21 (2009) 41-46.

253. J.P. Medema, L. Vermeulen, Microenvironmental regulation of stem cells in intestinal homeostasis and cancer, Nature, 474 (2011) 318-326.

254. T. Borovski, E.M. Felipe De Sousa, L. Vermeulen, J.P. Medema, Cancer stem cell niche: the place to be, Cancer Res., 71 (2011) 634-639.

255. L. Ricci-Vitiani, R. Pallini, M. Biffoni, M. Todaro, G. Invernici, T. Cenci, G. Maira, E.A. Parati, G. Stassi, L.M. Larocca, R. De Maria, Tumour vascularization via endothelial differentiation of glioblastoma stem-like cells, Nature, 468 (2010) 824-828.

256. E. Lonardo, J. Frias-Aldeguer, P.C. Hermann, C. Heeschen, Pancreatic stellate cells form a niche for cancer stem cells and promote their self-renewal and invasiveness, Cell Cycle, 11 (2012) 1282-1290.

257. S. Liu, C. Ginestier, S.J. Ou, S.G. Clouthier, S.H. Patel, F. Monville, H. Korkaya, A. Heath, J. Dutcher, C.G. Kleer, Y. Jung, G. Dontu, R. Taichman, M.S. Whicha, Breast cancer stem cells are regulated by mesenchymal stem cells through cytokine networks, Cancer Res., 71 (2011) 614-624.

258. M.R. Junttila, F.J. de Sauvage, Influence of tumour micro-environment heterogeneity on therapeutic response, Nature, 501 (2013) 346-354.

259. P. Valent, D. Bonnet, R. De Maria, T. Lapidot, M. Copland, J.V. Melo, C. Chomienne, F. Ishikawa, J.J. Schuringa, G. Stassi, B. Huntly, H. Herrmann, J. Soulier, G.J. Schuurhuis, S. Wohrer, M. Arock, J. Zuber, S. Cerny-Reiterer, H.E. Johnsen, M. Andreef, C. Eaves, Cancer stem cell definitions and terminology: the devil is in the details, Nat. Rev. Cancer, 12 (2012) 767-775.

260. J.P. Medema, Cancer stem cells: the challenges ahead, Nat. Cell Biol., 15 (2013) 338-344.

261. K. Kai, Y. Arima, T. Kamiya, Breast cancer stem cells, Breast Cancer, 17 (2010) 80-85.

262. J.E. Dick, Stem cell concepts renew cancer research, Blood, 112 (2008) 4793-4807.

263. T.H. Cheung, T.A. Rando, Molecular regulation of stem cell quiescence, Nat. Rev. Mol. Cell Biol., 14 (2013) 329-340.

264. P.S. Linsley, J. Schelter, J. Burchard, M. Kibukawa, M.M. Martin, S.R. Bartz, J.M. Johnson, J.M. Cummins, C.K. Raymond, H. Dai, N. Chau, M. Cleary, A.L. Jackson, M. Carleton, L. Lim, Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression, Mol. Cell. Biol., 27 (2007) 2240-2252.

265. N. Li, S. Singh, P. Cherukuri, H. Li, Z. Yuan, L. Ellisen, B. Wang, D. Robbins, J. DiRenzo, Reciprocal intraepithelial interactions between TP63 and hedgehog signaling regulate quiescence and activation of progenitor elaboration by mammary stem cells, Stem Cells, 26 (2008) 1253-1264.

266. L. Sang, H.A. Coller, J.M. Roberts, Control of the reversibility of cellular quiescence by the transcriptional repressor HES1, Science, 321 (2008) 1095-1100.

267. H. Liu, A.S. Adler, E. Segal, H.Y. Chang, A transcriptional program mediating entry into cellular quiescence, PLoS Genet., 3 (2007) e91.

268. T. Reya, H. Clevers, Wnt signalling in stem cells and cancer, Nature, 434 (2005) 843-850.

269. L. Li, R. Bhatia, Stem cell quiescence. Clin. Cancer Res., 17 (2011) 4936-4941.

270. A. Marusyk, V. Almendro, K. Polyak, Intra-tumour heterogeneity: a looking glass for cancer? Nat. Rev. Cancer, 12 (2012) 323-334.

271. D.R. Pattabiraman, R.A. Weinberg, Tackling the cancer stem cells—what challenges do they pose? Nat. Rev. Drug Discov., 13 (2014) 497-512.

272. V. Angeloni, P. Tiberio, V. Appierto, M.G. Daidone, Implications of stemness-related signaling pathways in breast cancer response to therapy, Sem. Cancer Biol, 31 (2015) 43-51.

273. C. Ginestier, M.H. Hur, E. Charafe-Jauffret, F. Monville, J. Ducther, M. Brown, J. Jacquemier, P. Viens, C.G. Kleer, S. Liu, A. Schott, D. Hayes, D. Birnbaum, M.S.

Wicha, G. Dontu, ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome, Cell Stem Cell, 1 (2007) 555-567.

274. C. Sheridan, H. Kishimoto, R.K. Fuchs, S. Mehrotra, P. Bhat-Nakshatri, C.H. Turner, R. Goulet, S. Badve, H. Nakshatri, CD44+/CD24-breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis, Breast Cancer Res., 8 (2006) R59.

275. X. Li, M.T. Lewis, J. Huang, C. Gutierrez, C.K. Osborne, M.F. Hu, S.G. Hilsenbeck, A. Pavlick, X. Zhang, G.C. Chamness, H. Wong, J. Rosen, J.C. Chang, Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. J. Natl. Cancer Inst., 100 (2008) 672-679.

276. M. Diehn, R.W. Cho, N.A. Lobo, T. Kalisky, M.J. Dorie, A.N. Kulp, D. Qian, J.S. Lam, L.E. Ailles, M. Wong, B. Joshua, M.J. Kaplan, I. Wapnir, F.M. Dirbas, G. Somlo, C. Garberoglio, B. Paz, J. Shen, S.K. Lau, S.R. Quake, J.M. Brown, I.L. Weissman, M.F. Clarke, Association of reactive oxygen species levels and radioresistance in cancer stem cells, Nature, 458 (2009) 780-783.

277. G.B Jang, I.S. Hong, R.J. Kim, S.Y. Lee, S.J. Park, E.S. Lee, J.H. Park, C.H. Yun, J.U. Chung, K.J. Lee, H.Y. Lee, J.S. Nam, Wnt/β-catenin small-molecule inhibitor CWP232228 preferentially inhibits the growth of breast cancer stem-like cells, Cancer Res., 75 (2015) 1691-1702.

278. M. Rodriguez-Torres, A.L. Allan, Aldehyde dehydrogenase as a marker and functional mediator of metastasis in solid tumors, Clin. Exp. Met., 33 (2015) 1-17.

279. A.F. Schott, M.D. Landis, G. Dontu, K.A. Griffith, R.M. Layman, I. Krop, L.A. Paskett, H. Wong, L.E. Dobrolecki, M.T. Lewis, A.M. Froehlich, J. Paranilam, D.F. Hayes, M.S. Wicha, J.C. Chang, Preclinical and clinical studies of gamma secretase inhibitors with docetaxel on human breast tumors, Clin. Cancer Res., 19 (2013) 1512-1524.

280. H. An, J.Y. Kim, N. Lee, Y. Cho, E. Oh, J.H. Seo, Salinomycin possesses anti-tumor activity and inhibits breast cancer stem-like cells via an apoptosis-independent pathway. Biochem. Biophys. Res., 466 (2015) 696-703.

281. V. Neumeister, In situ identification of putative cancer stem cells by multiplexing ALDH1, CD44, and cytokeratin identifies breast cancer patients with poor prognosis, Am. J. Pathol., 176 (2010) 2131-2138.

282. Y. Xiao, Y. Ye, K. Yearsley, S. Jones, S.H. Barsky, The lymphovascular embolus of inflammatory breast cancer expresses a stem cell-like phenotype, Am. J. Pathol, 173 (2008) 561-574.

283. A.K. Croker, D. Goodale, J. Chu, C. Postenka, B.D. Hedley, D.A. Hess, A.L. Allan, High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability, J. Cell. Mol. Med., 13 (2009) 2236-2252.

284. S.Y, Park, H.E. Lee, H. Li, M. Shipitsin, R. Gelman, K. Polyak, Heterogeneity for stem cell–related markers according to tumor subtype and histologic stage in breast cancer. Clin. Cancer Res., 16 (2010) 876-887.

285. B.K. Abraham, P. Fritz, M. McClellan, P. Hauptvogel, M. Athelogou, H. Brauch, Prevalence of CD44+/CD24−/low cells in breast cancer may not be associated with clinical outcome but may favor distant metastasis, Clin. Cancer Res., 11 (2005) 1154-1159.

286. M. Bernardi, A.F. Logullo, F.S. Pasini, S. Nonogaki, C. Blumke, F.A. Soares, M.M. Brentani, Prognostic significance of CD24 and claudin-7 immunoexpression in ductal invasive breast cancer, Oncol. Rep., 27 (2012) 28-38.

287. A. Bane, A. Viloria-Petit,. D. Pinnaduwage, A.M. Mulligan, F.P. O'Malley, I.L. Andrulis, Clinical–pathologic significance of cancer stem cell marker expression in familial breast cancers, Breast Cancer Res. Treat., 140 (2013) 195-205.

288. R. Liu, X. Wang, G.Y. Chen, P. Dalerba, A. Gurney, T. Hoey, G. Sherlock, J. Lewicki, K. Shedden, M.F. Clarke, The prognostic role of a gene signature from tumorigenic breast-cancer cells, N. Eng. J. Med., 356 (2007) 217-226.

289. F. Bertucci, N. Cervera, D. Birnbaum, A gene signature in breast cancer, N. Eng. J. Med., 356 (2007) 1887-8.

290. X.W. Ding, J.H Wu, C.P. Jiang, ABCG2: a potential marker of stem cells and novel target in stem cell and cancer therapy, Life Sciences, 86 (2010) 631-637.

291. L.A. Doyle, W. Yang, L.V. Abruzzo, T. Krogmann, Y. Gao, A.K. Rishi, D.D. Ross, A multidrug resistance transporter from human MCF-7 breast cancer cells, Proc. Natl. Acad. Sci., 95 (1998) 15665-15670.

292. C. Lagadec, E. Vlashi, L.D. Donna, Y.H. Meng, C. Dekmezian, K. Kim, F. Pajonk, Survival and self-renewing capacity of breast cancer initiating cells during fractionated radiation treatment, Breast Cancer Res., 12 (2010) R13.

293. A. Trumpp, O.D. Wiestler, Mechanisms of disease: cancer stem cells—targeting the evil twin, Nat. Clin. Pract. Oncol., 5 (2008) 337-347.

294. W.A. Woodward, M.S. Chen, F. Behbod, M.P. Alfaro, T.A. Buchholz, J.M. Rosen, WNT/β-catenin mediates radiation resistance of mouse mammary progenitor cells. Proc. Natl Aca. Sci., 104 (2007) 618-623.

295. M.T. Lewis, J.M. Veltmaat, Next stop, the twilight zone: hedgehog network regulation of mammary gland development, J. Mammary Gland Biol. Neoplasia, 9 (2004) 165-181.

296. T.M. Phillips, W.H. McBride, F. Pajonk, The response of CD24−/low/CD44+ breast cancer–initiating cells to radiation. J. Natl. Cancer Inst., 98 (2006) 1777-1785.

297. J.F. Zhan, L. Chen, Y. Yuan, G. Xie, A. Sun, Y. Liu, Z.X. Chen, STAT1 promotes radioresistance of CD44+/CD24−/low cells in breast cancer, Exp. Biol. Med., 236 (2011) 418-422.

298. W. Wei, D.J. Tweardy, M. Zhang, J. Landua, I. Petrovic, W. Bu, K. Roarty, S.G. Hilsenbeck, J.M. Rosen, M.T. Lewis, STAT3 Signaling Is Activated Preferentially in Tumor-Initiating Cells in Claudin-Low Models of Human Breast Cancer, Stem Cells, 32 (2014) 2571-2582.
299. R. Olsauskas-Kuprys, A. Zlobin, C. Osipo, Gamma secretase inhibitors of Notch signaling. Onco. Targets Ther., 6 (2013) 943-955.

300. C.C. Zhang, A. Pavlicek, Q. Zhang, M.E. Lira, C.L. Painter, Z. Yan, X. Zheng, N.V. Lee, M. Ozek, M. Qiu, Q. Zong, P.B. Lappin, A. Wong, P.A. Rejto, T. Smeal, J.G. Christensen, Biomarker and pharmacologic evaluation of the γ-secretase inhibitor PF-03084014 in breast cancer models, Clin. Cancer Res., 18 (2012) 5008-5019.

301. H.A. Hirsch, D. Lliopoulos, P.N. Tsichlis, K. Struhl, Metformin selectively targets cancer stem cells, and acts together with chemotherapy to block tumor growth and prolong remission, Cancer Res., 69 (2009) 7507-7511.

302. S. Cufí, B. Coromina-Faja, A. Vazquez-Martin, O.F. Cristina, D.R. Joan, B.B. Joaquim, M.C. Begona, M.M. Abel, Metformin-induced preferential killing of breast cancer initiating CD44+ CD24-/low cells is sufficient to overcome primary resistance to trastuzumab in HER2+ human breast cancer xenografts, Oncotarget, 3 (2012) 395-398.

303. P. Zhu, M. Davis, A.J. Blackwelder, N. Bachman, B. Liu, S. Edgerton, L.L. Williams, A.D. Thor, X. Yang, Metformin selectively targets tumor-initiating cells in ErbB2-overexpressing breast cancer models, Cancer Prev. Res., 7 (2014) 199-210.

304. T.C. Wang, T.K. Chao, C.C. Chang, Y.T. Yo, M.H. Yu, H.C. Lai, Drug screening identifies niclosamide as an inhibitor of breast cancer stem-like cells, PloS One, 8 (2013) e74538.

305. N. Yip, I.S. Fonbom, P. Liu, S. Brown, V. Kannappan, A.L. Armesilla, B. Xu, J. Cassidy, J.L. Darling, W. Wang, Disulfiram modulated ROS–MAPK and NFκB pathways and targeted breast cancer cells with cancer stem cell-like properties, Br. J. Cancer, 104 (2011) 1564-1574.

306. S.P. Zielske, A.C. Spalding, T.S. Lawrence, Loss of tumor-initiating cell activity in cyclophosphamide-treated breast xenografts, Trans. Oncol, 3 (2010) 149-152.

307. M. Shipitsin, L.L. Campbell, P. Argani, S. Weremowicz, N. Bloushtain-Qimron, J. Yao, T. Nikolskaya, T. Serebryiskaya, R. Beroukhim, M. Hou, M.K Haloushka, S. Sukumar, L.M. Parker, K.S. Anderson, L.N. Harris, J.E. Garber, A.L. Richardson, S.J. Schnitt, Y. Nikolsky, R.S. Gelman, K. Polyak, Molecular definition of breast tumor heterogeneity, Cancer Cell, 11 (2007) 259-273.

308. T. Tanei, K. Morimoto, K. Shimazu, S.J. Kim, Y. Tanji, T. Tagushi, Y. Tamaki, S. Nogushi, Association of breast cancer stem cells identified by aldehyde dehydrogenase 1 expression with resistance to sequential Paclitaxel and epirubicin-based chemotherapy for breast cancers, Clin. Cancer Res., 15 (2009) 4234-4241.

309. W.A. Woodward, S. Krishnamurthy, A. Lodhi, L. Xiao, Y. Cong, M. Cristofanilli, T.A. Buchholz, A. Lucci, Aldehyde dehydrogenase1 immunohistochemical staining in primary breast cancer cells independently predicted overall survival but did not

correlate with the presence of circulating or disseminated tumors cells, J. Cancer, 5 (2014) 360-367.

310. K. Morimoto, S.J. Kim, T.Tanei, K. Shimazu, Y. Tanji, T. Tagushi, Y. Tamaki, N. Terada, S. Nogushi, Stem cell marker aldehyde dehydrogenase 1-positive breast cancers are characterized by negative estrogen receptor, positive human epidermal growth factor receptor type 2, and high Ki67 expression, Cancer Sci., 100 (2009) 1062-1068.

311. E. Resetkova, J.S. Reis-Filho, R.K. Jain, R. Mehta, M.A. Thorat, H. Nakshatri, S. Badve, Prognostic impact of ALDH1 in breast cancer: a story of stem cells and tumor microenvironment, Breast Cancer Res. Treat., 123 (2010) 97-108.

312. E. Charafe-Jauffret, C. Ginestier, F. Lovino, C. Tarpin, M. Diebel, B. Esterni, G. Houvenaeghel, J.M. Extra, F. Bertucci, J. Jacquemier, L. Xerri, G. Dontu, G. Stassi, Y. Xiao, S.H. Barsky, D. Birnbaum, P. Viens, M.S. Whicha, Aldehyde dehydrogenase 1–Positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer, Clin. Cancer Res., 16 (2010) 45-55.

313. Y. Gong, J. Wang, L. Huo, W. Wei, N.T. Ueno, Aldehyde dehydrogenase 1 expression in inflammatory breast cancer as measured by immunohistochemical staining. Clin. Breast Cancer, 14 (2014) e81-e88.

314. A.R. Wolfe, W.A. Woodward, Breast cancer stem cell correlates as predicative factors for radiation therapy. Sem. Rad. Oncol., 25 (2015) 251-259.

315. D.G. Tiezzi, W.S. Clagnan, L.M.R. Mandarano, C.B. de Sousa, H.R.C. Marana, M.G. Tiezzi, J.M. Andrade, Expression of aldehyde dehydrogenase after neoadjuvant chemotherapy is associated with expression of hypoxia-inducible factors 1 and 2 alpha and predicts prognosis in locally advanced breast cancer, Clinics, 68 (2013) 592-598.

316. D. Zhao, Y. Mo, M.T. Li, S.W. Zhou, Z.L. Cheng, Y.P. Sun, Y. Xiong, K.L. Guan, Q.Y. Lei, NOTCH-induced aldehyde dehydrogenase 1A1 deacetylation promotes breast cancer stem cells, J. Clin. Inv., 124 (2014) 5453.

317. P. Marcato, C.A. Dean, R.Z. Liu, K.M. Coyle, M. Bydoun, M. Wallace, D. Clements, C. Turner, E.G. Mathenge, S.A. Gujar, C.A. Giacomoantonio, J.R. Mackey, R. Godbout, P.W.K. Lee, Aldehyde dehydrogenase 1A3 influences breast cancer progression via differential retinoic acid signaling, Mol. Oncol., 9 (2015) 17-31.

318. N.E. Bhola, J.M. Balko, T.C. Dugger, M.G. Kuba, V. Sánchez, M. Sanders, J. Stanford, R.S. Cook, C.L. Arteaga, TGF-β inhibition enhances chemotherapy action against triple-negative breast cancer, J. Clin.l Inv.,, 123 (2013) 1348**.**

319. R. Zheng, J. Wang, Q. Wu, Y. Ou, Z. Wang, L. Ma, M. Wang, J. Wang, Y. Yan, Expression of ALDH1 and TGFβ2 in benign and malignant breast tumors and their prognostic implications, Int. J. Clin. Exp. Pathol., 7 (2014) 4173.

320. E. Charafe-Jauffret, C. Ginestier, F. Lovino, J. Wicinski, N. Cervera, P. Finetti, M.H. Hur, M.E. Diebel, F. Monville, J. Dutcher, M. Brown, P. Viens, L. Xerri, F. Bertucci, G. Stassi, G. Dontu, D. Birnbaum, M.S. Whicha, Breast cancer cell lines contain

functional cancer stem cells with metastatic capacity and a distinct molecular signature, Cancer Res., 69 (2009) 1302-1313.

321. P. Marcato, C.A. Dean, D. Pan, R. Araslanova, M. Gillis, M. Joshi, L. Helyer, L. Pan, A. Leidal, S. Gujar, C.A. Giacomoantonio, P.W.K. Lee, Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis, Stem Cells, 29 (2011) 32-45.

322. A.K. Croker, A.L. Allan, Inhibition of aldehyde dehydrogenase (ALDH) activity reduces chemotherapy and radiation resistance of stem-like ALDHhiCD44+ human breast cancer cells, Breast Cancer Res. Treat., 133 (2012) 75-87.

323. S.B. Keysar, A. Jimeno, More than markers: biological significance of cancer stem cell-defining molecules. Mol. Cancer Ther., 9 (2010) 2450-2457.

324. A.D.C. Paula, O. Marques, A.M. Rosa, M.D.F. Faria, A. Rêma, C. Lopes, Co-expression of stem cell markers ALDH1 and CD44 in non-malignant and neoplastic lesions of the breast, Anticancer Res, 34 (2014) 1427-1434.

325. A.D.C. Paula, O. Marques, R. Sampaio, A. Rosa, J. Garcia, A. Rêma, M.D.F. Faria, P. Silva, R. Vizcaíno, C. Lopes, Characterization of $CD44^{+}$/$ALDH1^{+}$/$Ki\text{-}67^{-}$ cells in non-malignant and neoplastic lesions of the breast (Submitted for publication)

326. J. Cui, P. ll, X. Liu, H. Hu, W. Wei, Abnormal expression of the Notch and Wnt/β-catenin signaling pathways in stem-like ALDHhiCD44+ cells correlates highly with Ki-67 expression in breast cancer, Onco. Lett., 9 (2015) 1600-1606.

327. S. Miraglia, W. Godfrey, A.H. Yin, K. Atkins, R. Warnke, J.T. Holden, R.A. Bray, E.K. Waller, D.W. Buck, A novel five-transmembrane hematopoietic stem cell antigen: isolation, characterization, and molecular cloning, Blood, 90 (1997) 5013-5021.

328. Y. Kosodo, K. Roper, W. Haubensak, A.M. Marzesco, D. Corbeil, W.B. Huttner, Asymmetric distribution of the apical plasma membrane during neurogenic divisions of mammalian neuroepithelial cells, EMBO J., 23 (2004) 2314-2324.

329. W. Hilbe, S. Dimhofer, F. Oberwasserlechner, T. Schmid, E. Gunsilius, G. Hilbe, E. Woll, C.M. Kahler, CD133 positive endothelial progenitor cells contribute to the tumour vasculature in non-small cell lung cancer, J. Clin. Pathol., 57 (2004) 965-969.

330. S.K. Singh, C. Hawkins, I.D. Clarke, J.A. Squire, J. Bayani, T. Hide, R.M. Henkelman, M.D. Cusimano, P.B. Dirks, Identification of human brain tumour initiating cells, Nature, 432 (2004) 396-401.

331. L. Ricci-Vitiani, D.G. Lombardi, E. Pilozzi, M. Biffoni, M. Todaro, C. Peschle, R. De Maria, Identification and expansion of human colon-cancer-initiating cells, Nature, 445 (2007) 111-115.

332. R.J. Schrot, J.H. Ma, C.M. Greco, A.D. Arias, J.M. Angelastro, Organotypic distribution of stem cell markers in formalin-fixed brain harboring glioblastoma multiforme, J. Neuro-Oncol., 85 (2007) 149-157.

333. Q. Liu, J.G. Li, X.Y. Zheng, F. Jin, H. Dong, Expression of CD133, PAX2, ESA, and GPR30 in invasive ductal breast carcinomas. Chin. Med. J., 122 (2009) 2763-2769.

334. M.H. Wright, A.M. Calcagno, C.D. Salcido, M.D. Carlson, S.V. Ambudkar, L. Varticovski, Brca1 breast tumors contain distinct CD44+/CD24-and CD133+ cells with cancer stem cell characteristics, Breast Cancer Res., 10 (2008) R10.

335. L. Liu, B.C. Sun, X.L. Zhao, T. Sun, Q. Gu, Z. Yao, X.Y. Dong, N. Zhao, N. Liu, CD133+ cells with cancer stem cell characteristics associates with vasculogenic mimicry in triple-negative breast cancer, Oncogene, 22 (2013) 544-553.

336. F. Brugnoli, S. Grassilli, M. Piazzi, M. Palomba, E. Nika, A. Bavelloni, S. Capitani, V. Bertagnolo, In triple negative breast tumor cells, PLC-β2 promotes the conversion of CD133 high to CD133 low phenotype and reduces the CD133-related invasiveness. Mol. Cancer, 12 (2013) 1.

337. P. Zhao, Y. Lu, X. Jiang, X. Li, Clinicopathological significance and prognostic value of CD133 expression in triple-negative breast carcinoma, Cancer Sci., 102 (2011) 1107-1111.

338. R. Nadal, F.G. Ortega, M. Salido, J.A. Lorente, M. Rodriguez-Rivera, M. Delgado-Rodriguez, M. Macià, A. Fernández, J.M. Corominas, J.L. Garcia-Puche, P. Sánchez-Rovira, F. Solé, M.J. Serrano, CD133 expression in circulating tumor cells from breast cancer patients: potential role in resistance to chemotherapy, Int. J. Cancer, 133 (2013) 2398-2407.

339. C. Bock, B. Rack, J. Huober, U. Andergassen, U. Jeschke, S. Doisneau-Sixou, Distinct expression of cytokeratin, N-cadherin and CD133 in circulating tumor cells of metastatic breast cancer patients, Fut. Oncol., 10 (2014) 1751-1765.

340. M.J. Meyer, J.M. Fleming, A.F. Lin, S.A. Hussnain, E. Ginsburg, B.K. Vonderhaar, CD44posCD49fhiCD133/2hi defines xenograft-initiating cells in estrogen receptor–negative breast cancer, Cancer Res., 70 (2010) 4624-4633.

341. N. Aomatsu, M. Yashiro, S. Kashiwagi, T. Takashima, T. Ishikawa, M. Ohsawa, K. Wakasa, K. Hirakawa, CD133 is a useful surrogate marker for predicting chemosensitivity to neoadjuvant chemotherapy in breast cancer, PloS One, 7 (2012) e45865.

342. X.Y. Wang, L.O.F. Penalva, H. Yuan, R.I. Linnoila, J. Lu, H. Okano, R.I. Glazer, Musashi1 regulates breast tumor cell proliferation and is a prognostic indicator of poor survival, Mol. Cancer, 9 (2010) 221.

343. S.K. Swaminathan, E. Roger, U. Toti, L. Niu, J.R. Ohlfest, J. Panyam, CD133-targeted paclitaxel delivery inhibits local tumor recurrence in a mouse model of breast cancer, J. Control. Rel., 171 (2013) 280-287.

344. M. Shackleton, F. Vaillant, K.J. Simpson, J. Stingl, G.K. Smyth, M.L. Asselin-Labat, Li. Wu, G.J. Lindeman, J.E. Visvader, Generation of a functional mammary gland from a single stem cell, Nature, 439 (2006) 84-88.

345. J. Stingl, P. Eirew, I. Ricketson, M. Shackleton, F. Vaillant, D. Choi, H.I. Li, C.J. Eaves, Purification and unique properties of mammary epithelial stem cells, Nature, 439 (2006) 993-997.

346. A. Vassilopoulos, R.H. Wang, C. Petrovas, D. Ambrozak, R. Koup, C.X. Deng, Identification and characterization of cancer initiating cells from BRCA1 related mammary tumors using markers for normal mammary stem cells. Int. J. Biol. Sci., 4 (2008) 133-142.

347. A.M. Mercurio, Laminin: multiple forms, multiple receptors, Curr. Op. Cell Biol., 2 (1990) 845-849.

348. H.L. Goel, B. Pursell, C. Chang, L.M. Shaw, J. Mao, K. Simin, P. Kumar, C.W.V. Kooi, L.D. Schultz, D.L. Greiner, J.H. Norum, R. Toftgard, C. Kuperwasser, A.M. Mercurio, GLI1 regulates a novel neuropilin-2/α6β1 integrin based autocrine pathway that contributes to breast cancer initiation, EMBO Mol. Med., 5 (2013) 488-508.

349. J.D. Lathia, J. Gallagher, J.M. Heddleston, J. Wang, C.E. Eyler, J. MacSwords, Q. Wu, A. Vasanji, R.E. McLendon, A.B. Hjelmeland, J.N. Rich, Integrin alpha 6 regulates glioblastoma stem cells, Cell Stem cell, 6 (2010) 421-432.

350. D.D. Schlaepfer, T. Hunter, Focal adhesion kinase overexpression enhances ras-dependent integrin signaling to ERK2/mitogen-activated protein kinase through interactions with and activation of c-Src, J. Biol. Chem., 272 (1997) 13189-13195.

351. Y. Imanishi, B. Hu, M.J. Jarzynka, P. Guo, E. Elishaev, I. Bar-Joseph, S.Y. Chen, Angiopoietin-2 stimulates breast cancer metastasis through the α5β1 integrin-mediated pathway, Cancer Res., 67 (2007) 4254-4263.

352. A. Jahangiri, M.K. Aghi, W.S. Carbonell, β1 integrin: Critical path to antiangiogenic therapy resistance and beyond, Cancer Res., 74 (2014) 3-7.

353. E.S. Yao, H. Zhang, Y.Y. Chen, B. Lee, K. Chew, D. Moore, C. Park, Increased β1 integrin is associated with decreased survival in invasive breast cancer, Cancer Res., 67 (2007) 659-664.

354. P.B. dos Santos, J.S. Zanetti, A. Ribeiro-Silva, E.I.C. Beltrão, Beta 1 integrin predicts survival in breast cancer: a clinicopathological and immunohistochemical study, Diag. Pathol., 7 (2012) 104.

355. H.R. Ali, S.J. Dawson, F.M. Blows, E. Provenzano, P.D. Pharoah, C. Caldas, Cancer stem cell markers in breast cancer: pathological, clinical and prognostic significance, Breast Cancer Res., 13 (2011) R118.

356. F. Ye, Y. Qiu, L. Li, L. Yang, F. Cheng, H. Zhang, B. Wei, Z. Zhang, L. Sun, H. Bu, The Presence of EpCAM-/CD49f+ Cells in Breast Cancer Is Associated with a Poor Clinical Outcome, J. Breast Cancer, 18 (2015) 242-248.

357. P. Eirew, J. Stingl, A. Raouf, G. Turasshvili, S. Aparicio, J.T. Emerman, C.J. Eaves, A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability, Nat. Med., 14 (2008) 1384-1389.

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

359. J.S. Hale, M. Li, J.D. Lathia, The malignant social network: cell-cell adhesion and communication in cancer stem cells, Cell Adhe. Mig., 6 (2012) 346-355.

360. M. Cariati, A. Naderi, J.P. Brown, M.J. Smalley, S.E. Pinder, C. Caldas, A.D. Purushottam, Alpha-6 integrin is necessary for the tumourigenicity of a stem cell-like subpopulation within the MCF7 breast cancer cell line, Int. J. Cancer, 122 (2008) 298-304.

361. J.S. Zawistowski, K. Nakamura, J.S. Parker, D.A. Granjer, B.T. Golitz, G.L. Johnson, MicroRNA 9-3p targets β1 integrin to sensitize claudin-low breast cancer cells to MEK inhibition, Mol. Cell. Biol., 33 (2013) 2260-2274.

362. I. Ivanova, J.F. Vermeulen, C. Ercan, J.M. Houthuijzen, F.A. Saig, E.J. Vlug, E. van der Wall, P.J. van Diest, M. Vooijs, P.W.B. Derksen, FER kinase promotes breast cancer metastasis by regulating α6-and β1-integrin-dependent cell adhesion and anoikis resistance, Oncogene, 32 (2013) 5582-5592.

363. N. Navin, J. Kendall, J. Troge, P. Andrews, L. Rodgers, J. Mclindoo, K. Cook, A. Stepansky, D. Levy, D. Esposito, L. Muthuswamy, A. Krasnitz, W.R. McCombie, J. Hicks, M. Wigler, Tumour evolution inferred by single-cell sequencing, Nature, 472 (2011) 472: 90-94.

364. J.S. Reis-Filho, Next-generation sequencing. Breast Cancer Res, 11 (2009) S12.

365. A.D.C. Paula, C. Leitão, O. Marques, A.M. Rosa, A.H. Santos, A. Rêma, M.D.F. Faria, A. Rocha, J.L. Costa, M. Lima, C. Lopes, Characterization of CD44+/CD24-/Ck+/CD45- cells in non-malignant and neoplastic lesions of the breast through flow cytometry and massive parallel sequencing, submitted for publication.

366. H. Harrison, G. Farnie, S.J. Howell, R.E. Rock, S. Stylianou, K.R. Brennan, N.J. Brundred, R.B. Clarke, Regulation of breast cancer stem cell activity by signaling through the Notch4 receptor, Cancer Res., 70 (2010) 709-718.

367. P.V. Pham, N.L.C. Phan, N.T. Nguyen, N.H. Truong, T.T. Duong, D.V. Le, K.D. Truong, N.K. Phan, Differentiation of breast cancer stem cells by knockdown of CD44: promising differentiation therapy, J. Trans. Med., 9(2011) 209.

368. E.S. Knudsen, A. Witkiewicz, EZH2 and ALDH1 expression in ductal carcinoma in situ: complex association with recurrence and progression to invasive breast cancer, Cell Cycle 12 (2013) 2042-2050.

369. H. Korkaya, A. Paulson, E. Charafe-Jauffret, C. Ginestier, M. Brown, J. Dutcher, S.G. Clouthier, M.S. Whicha, Regulation of mammary stem/progenitor cells by PTEN/Akt/β-catenin signaling, PLoS Biol., 7 (2009) e1000121.

370. D. Raha, T.R. Wilson, J. Peng, D. Peterson, P. Yue, M. Evangelista, C. Wilson, M. Merchant, J. Settleman, The cancer stem cell marker aldehyde dehydrogenase is

required to maintain a drug tolerant tumor cell subpopulation, Cancer Res., 74 (2014) 3579–3590.

371. C. Visus, D. Ito, A. Amoscato, M. Maciejewska-Franczak, A. Abdelsalem, R. Dhir, A.B. DeLeo, Identification of human aldehyde dehydrogenase 1 family member A1 as a novel CD8+ T-cell-defined tumor antigen in squamous cell carcinoma of the head and neck, Cancer Res., 67 (2007) 10538–45.

372. K.M. Ahmed, H. Zhang, C.C. Park, NF- κB regulates radioresistance mediated by β1-integrin in three-dimensional culture of breast cancer cells, Cancer Res., 73 (2013) 3737 –3748.

373. T. Narita, N. Kimura, M. Sato, N. Matsuura, R. Kannagi, Altered expression of integrins in adriamycin-resistant human breast cancer cells, Anticancer Res., 18 (1997) 257-262.

374. A. Fawzi, K. Vuori, Integrin signaling inhibits paclitaxel-induced apoptosis in breast cancer cells, Oncogene, 20 (2001) 4995-5004.

375. P. Nisticò, F. Di Modugno, S. Spada, M.J. Bissel, β1 and β4 integrins: from breast development to clinical practice, Breast Cancer Res., 16 (2014)10-1186.

Chapter 2

Rationale and Aims

Breast cancer continues to be one of the major causes of female morbidity and mortality worldwide. Despite the main diagnostic and therapeutic innovations, the effect on mortality has been modest. Over the last decade, the notion that tumors are maintained by their own SCs, the so-called CSCs, has created great excitement in the research community. With the growing idea that these self-renewing tumorigenic cells are mainly responsible for resistance to chemo-radiation therapy and cancer relapses, several studies were done to identify putative CSCs in several solid cancers, like in breast cancer.

However, and due to the high levels of heterogeneity associated with this disease, several BCSC markers have been identified and characterized, with some being associated to aggressive forms of breast cancer. Nonetheless, and among all BCSC markers identified until now, it is important to define which BCSC phenotypes have high tumorigenic potential and ability to resist to therapeutic agents.

With this in mind, the general aim of this doctoral thesis was to characterize different BCSC populations in different stages of breast cancer progression. Thus, and using breast reduction specimens as well as non-malignant and malignant tissues, the following studies were performed in order to assess specific objectives:

I. **To compare two of the most reliable BCSCs markers, ALDH1 and CD44 and to correlate their expression within different breast lesions.**

   We aimed to explore a possible association between a high expression of SC markers and a specific type of lesion during breast cancer progression. Thus, ALDH1 and CD44 expression was immunohistochemically evaluated in non-malignant, DCIS and IDCs sample cores.

II. **To characterize the $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ phenotype in non-malignant and malignant breast lesions.**

   A higher combined expression of CD44 and ALDH1 in DCIS was observed when compared with IDCs in objective I. In this way, we combined these BCSC markers with Ki-67 to evaluate quiescence in

order to identify, evaluate its distribution and estimate the mean percentages of $CD44^+/ALDH1^+/Ki\text{-}67^-$ cells in a series of normal, non-malignant and malignant breast tissues. Clinical relevance of this phenotype was inferred by associations with markers of breast cancer behavior, progression and survival.

III. To characterize $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells through flow cytometry and massive parallel sequencing.

Considering the well-defined $CD44^+/CD24^{-/low}$ phenotype in breast cancer and the current improvements done in NGS technologies, we aimed to characterize, through flow cytometry, $CD44^+/CD24^-/Ck^+/CD45^-$ cells in frozen biopsy samples harboring different breast lesions. More importantly, we also aimed to determine which somatic mutations were associated to this phenotype isolated from frozen mastectomy samples, using the Ion Torrent Ampliseq Cancer Hotspot panel v2.

Chapter 3

Co-expression of Stem Cell Markers ALDH1 and CD44 in Non-Malignant and Neoplastic Lesions of the Breast

ANTICANCER RESEARCH *34*: 1427-1434 (2014)

Co-expression of Stem Cell Markers ALDH1 and CD44 in Non-malignant and Neoplastic Lesions of the Breast

ARNAUD DA CRUZ PAULA[1], ORIANA MARQUES[2], ANA MARGARIDA ROSA[1], MARIA DE FÁTIMA FARIA[1], ALEXANDRA RÊMA[1] and CARLOS LOPES[3]

[1]*Department of Pathology and Molecular Immunogenetics, Institute of Biomedical Sciences Abel Salazar, Porto, Portugal;*
[2]*Unit for Multidisciplinary Biomedical Research (UMIB), Institute of Biomedical Sciences Abel Salazar, Porto, Portugal;*
[3]*Group of Pathology and Experimental Therapeutics, Portuguese Institute of Oncology – IPO Research Center, Porto, Portugal*

Abstract. *Background/Aim: The Cancer Stem Cell (CSC) model proposes that cancer is driven by a cellular component which possesses stem cell (SC) properties, cancer stem cells (CSCs), a distinct cell-type which is tumorigenic and capable of invasion and metastasis. Enzymatic activity of aldehyde dehydrogenase-1 (ALDH1), a de-toxifying enzyme that oxidizes intracellular aldehydes, has been used as a marker of normal and malignant breast stem cells (BSCs). CD44-transmembrane protein has already been shown to possess the ability to identify breast epithelial cells with stem properties. Materials and Methods: In order to compare two of the currently most reliable BSCs markers, ALDH1 and CD44, and to correlate their expression within different breast lesions, 190 samples from breast cancer specimens were analyzed by tissue microarrays. Results: ALDH1 expression was observed in 85.43% and CD44 in 90.3% of all samples. No overexpression was observed for ALDH1 between invasive tumors, ductal carcinomas in situ and non-malignant lesions of breast, although ALDH1 had a significant negative correlation with estrogen-receptor (ER) and progesterone-receptor (PR) status ($p=0.002$ and $p=0.001$, respectively) and a positive correlation with CD44 ($p<0.001$). Moreover, combined overexpression of ALDH1 and CD44 was observed in ductal in situ tumors ($p<0.001$). Conclusion: The combined overexpression of these markers in ductal carcinomas in situ is in agreement with the CSC model in breast cancer.*

Correspondence to: Arnaud Da Cruz Paula, MSc. and Carlos Lopes, MD, Ph.D., Department of Pathology and Molecular Immunogenetics, Institute of Biomedical Sciences Abel Salazar. Jorge Viterbo Ferreira Street, n° 288, 4050-313, Porto, Portugal. Tel: +35 1967693532, +35 1917270209, Fax: +35 1225511184, e-mail: arnaudcpaula@hotmail.com, calopes@icbas.up.pt

Key Words: Breast cancer, breast stem cells, ALDH1, CD44, cancer stem model.

Although the concept that cancer may arise from a small cell population with stem cell (SC) characteristics has been proposed since more than 150 years, new evidence has given an impetus to it through new advancements on SC research (1). According to the cancer stem cell (CSC) model, CSCs have the ability to maintain the growth and expansion of the tumor mass, and originate a differentiated cell population, with none or limited proliferation capacity (2), through de-regulation of the self-renewal process. There are two ways by which CSCs can be formed: mutations allowing for abnormal expansion of normal SCs or mutations in progenitor cells that re-acquire the ability to self-renew (3-6). Either way, the CSC model is capable of explaining the phenotypic heterogeneity observed in many tumors, since self-renewal conducts the tumorigenesis and differentiation (albeit aberrant) of the tumor (1, 3).

Developmental biology has advanced greatly through the discovery of markers that distinguish, phenotypically, SCs from their differentiated progeny. These discoveries have allowed SCs to be isolated and studied separately, using techniques such as flow cytometry or immunohistochemistry, while providing critical insight into the regulatory mechanisms of SC function (1). However, there is relatively little overlap between the different CSC markers reported in different tumor types and so, the choice of marker can vary greatly, depending on tissues or species (2-5). ALDH1 has been shown to be a reliable SC marker based on its ability to isolate a sub-population of cells displaying SC properties from normal human breast tissue and breast carcinomas (6). Furthermore, CD44 is also overexpressed in basal-like breast cancers having the ability to identify breast epithelial cells with stem properties (7).

ALDH1 is a de-toxifying enzyme that oxidizes intracellular aldehydes and confers resistance to alkylating agents (8-11). In fact, the de-toxification capacity of ALDH1, protecting SCs against oxidative insult, might underlie the well-recognized longevity of these cells. ALDH1 also converts retinol to retinoic acid, a modulator of cell proliferation, which may also control stem cell proliferation (12).

CD44 is a transmembrane receptor protein that participates in many cell–cell and cell–matrix interactions (13, 14), known to be expressed in most cell types (6, 7) and reported to enhance tumor invasion and metastasis (15).

In the present study, we aimed to investigate whether a higher expression of the SC markers ALDH1 and CD44 is associated with particular types of breast lesions, more specifically: non-malignant, ductal carcinomas *in situ* (DCIS) and invasive ductal carcinomas (IDC).

Materials and Methods

Tissue Microarrays. From the 190 patients selected for this study, 139 had invasive ductal carcinomas and 51 had *in situ* ductal carcinomas. The correspondent formalin-fixed paraffin-embedded (FFPE) tissue blocks were collected from 2004 to 2011, along with their corresponding hematoxylin and eosin (H&E)-stained slides, from the archives of Department of Pathology, Santo António Hospital (Porto Hospital Centre, Porto, Portugal). Hormone receptor status was obtained from routine performance on diagnostic specimens from the same Department of Pathology. Areas of DCIS and IDC of the breast were identified and marked from each block. Non-malignant lesions were selected from specimens that also contained invasive ductal carcinomas. DCIS sections were selected from patients with ductal carcinoma *in situ*. *In situ* and invasive lobular carcinomas were excluded from the study. A total of 414 FFPE breast tissue specimens were used for tissue microarray construction, out of which 350 cores were evaluable for ALDH1 and 365 cores for CD44. Tissue cores from human liver donor samples were included in each of these recipient blocks, for a correct slide orientation. 2-μm thick sections were routinely cut in a microtome. Pathological and clinical characteristics of the study population are provided in Tables I and II.

Immunohistochemistry. ALDH1 and CD44 immunohistochemistry was performed on tissue microarray paraffin-embedded sections with the monoclonal antibodies ALDH1 (Rabbit monoclonal, IgG isotype, 0.13 mg/ml concentrate, Abcam, Cambridge, United Kingdom) and CD44 (Mouse monoclonal, IgG isotype, 0.01 mg/ml concentrate, MRQ-13, Cell Marque, California, United States). The sections were de-paraffinized, rehydrated in a series of graded ethanol and washed in water. Target retrieval was achieved with citrate buffer (pH 6.0), in a microwave at 850 W until boiling. After cooling, slides were washed three times in PBST (phosphate buffered saline, pH 7.4 – 0.05% Tween 20) each for 5 min. To avoid unspecific tissue peroxidase activity, slides were incubated, for 10 min, with 3% peroxide hydrogen in methanol and after incubation with a blocking solution (Ultra Vision LP Detection System, Thermo Fisher Scientific, Cheshire, UK), for 5 min. The ALDH1 (dilution: 1/100) and CD44 (dilution: 1/100) antibodies were incubated in a humid chamber, for 1 h at room temperature. The slides were then washed once in PBST for 5 min and incubated in a humid chamber with Primary Antibody Enhancer (Ultra Vision LP Detection System, Thermo Fisher Scientific, Cheshire, UK), for 10 min, and further washed three times, for 5 min each in PBST. Horseradish Peroxidase Polymer (Ultra Vision LP Detection System, Ultra Vision LP Detection System, Thermo Fisher Scientific, Cheshire, UK) was added, for 20 minutes in a humid chamber and washed three times, 5 minutes each in PBST. Enzyme reactivity was visualized using 3,3'-Diaminobenzidine tetrahydrochloride (DAB, Sigma, Saint Louis, United States). Slides were counter-stained with Mayers hemalumen solution (Merck), de-hydrated and mounted with Entellan (Sigma).

Staining. The reaction obtained in all samples was observed in a Nikon Eclipse E400 bright-field microscope. Each core was analyzed and classified according to its histological type. For the evaluation of staining patterns only breast epithelial cells were considered, excluding also artifacts staining tumor cell debris. A semi-quantitative evaluation method was applied: the percentage of positive cells (0 points: 0%; 1 point: 1-10%, 2 points: 11-20%, 3 points: 21-50% and 4 points: >50%) and the staining intensity (0 points: no staining, 1 point: weak staining, 2 points: moderate staining and 3 points: strong staining) were considered and multiplied. The presence or absence of immunoexpression for ALDH1 and CD44 was also considered. A combined score was also created by addition of ALDH1 and CD44 dichotomized scores in order to produce 3 categories (0 to 2). According to the American Society of Clinical Oncology (ASCO) for the hormone receptor status criteria, specimens having at least 1% positive tumor cells for ER or PR were considered positive.

Table I. *Clinical characteristics of samples*.

Characteristics	Cohort population n (%)	CD44 expression samples n (%)	ALDH1 expression samples n (%)
Type of Lesion			
Non-Malignant	71 (17.2%)	67 (18.4%)	48 (13.7%)
DCIS	121 (29.2%)	115 (31.5%)	113 (32.3%)
IDC	222 (53.6%)	183 (50.1%)	189 (54.0%)
Total	414 (100%)	365 (100%)	350 (100%)

DCIS, Ductal carcinoma *in situ*; IDC, invasive ductal carcinoma.

Table II. *Invasive ductal carcinoma patients' hormone receptor status*.

Characteristics	n (%)
ER status	
Positive	118 (84.9%)
Negative	21 (15.1%)
PR status	
Positive	106 (76.3%)
Negative	33 (23.7%)

ER, Estrogen receptor; PR, progesterone receptor.

Table III. *ALDH1 and CD44 positivity in breast tissue samples.*

	ALDH1 expression score			CD44 expression score		
Characteristics	Positive cases (%)	Negative cases (%)	*p*-Value	Positive cases (%)	Negative cases (%)	*p*-Value
Type of Lesion						
Non-Malignant	87.5	12.5	0.064[a]	81.1	11.9	0.003[a]
DCIS	91.2	8.8		94.8	5.4	
IDC	81.5	18.5		80.9	19.1	

DCIS, Ductal carcinoma *in situ*; IDC, invasive ductal carcinoma. [a]Pearson Chi-Square.

Statistical analysis. Results of the immunohistochemical analyses were statistically examined using IBM SPSS Statistics Version 20.0 (SPSS Inc., IBM, Chicago, IL, USA). Statistical significance was calculated using the Mann-Whitney or Kruskal-Wallis test, according to the number of groups considered, and Fisher's Exact Test or Pearson's Chi-Square to assess relationships between data sets. Significance was set at $p<0.05$.

Results

Details on the baseline characteristics of the samples selected for analysis are provided on Table I. Samples from 190 patients' FFPE blocks were included in the study. The cohort comprised of 71 non-malignant (17.2%), including normal, hyperplastic and other non-malignant samples, 121 DCIS (29.2%) and 222 IDC (53.6%) lesions. The ALDH1 immunoexpression evaluation cohort comprised of 48 non-malignant (13.7%), 113 DCIS (32.3%) and 189 IDC (54.0%) lesions. For CD44, 67 non-malignant (18.4%), 115 DCIS (31.5%) and 183 IDC (50.1%) lesions were analyzed (Table I).

ALDH1 and CD44 immunostaining was observed in epithelial cells of the breast and also in stromal inflammatory cells, both in non-malignant and neoplastic samples. Figure 1a shows ALDH1 immunoexpression in a papilloma where expression is observed in the cytoplasm of most cells. In a high-grade DCIS and in an invasive ductal carcinoma, ALDH1 is seen in cytoplasm of rare cells (Figure 1b) and in most of cells presented (Figure 1c). Figure 1d demonstrates CD44 immunoexpression also in a papilloma with a strong expression in the cells' membrane. A strong CD44 expression is also seen in the membrane of most cells in a high-grade DCIS (Figure 1e) as well as in an invasive ductal carcinoma (Figure 1f).

ALDH1 and CD44 expression and correlation with clinicopathological variables. Using positivity criteria, ALDH1 and CD44 expression was evaluated for each lesion type. CD44 expression was observed in 299 (85.4%) of the 350 samples and CD44 in 316 (90.3%) out of 365. A higher percentage of samples expressing ALDH1 (91.2%) and CD44 (94.8%) was seen in the DCIS group, followed by the non-malignant sample group (ALDH1 87.5%; CD44 81.1%) and then IDC (ALDH1 81.5%; CD44 80.9%). Differences between the groups were significant for CD44 expression only ($p=0.003$) (Table III).

Table IV. *ALDH1 and CD44 expression in breast tissue samples.*

	ALDH1 expression score		CD44 expression score	
Characteristics	Mean	*p*-Value	Mean	*p*-Value
Type of Lesion				
Non-Malignant	155.21		163.07	
DCIS	169.43	0.147[a]	190.43	0.209[a]
IDC	184.28		185.63	

DCIS, Ductal carcinoma *in situ*; IDC, invasive ductal carcinoma. [a]Kruskal-Wallis test.

ALDH1 and CD44 mean rank expression and correlation with clinicopathological variables. ALDH1 and CD44 mean expression was assessed for each type of lesion and for clinicopathological variables. For ALDH1, mean rank expression was higher in the IDC group (185.63), succeed by DCIS (169.43) and the non-malignant group (155.21). On the other hand, the DCIS group presented the higher CD44 mean rank expression (190.43), with the non-malignant group presenting the lowest mean rank expression (163.07). None of the SC markers mean rank expression was significantly different between the groups (Table IV). Considering the ER- and PR-status analysis, ALDH1 mean rank expression was significantly higher for ER-negative (86.65) and PR-negative (83.33) patients ($p=0.002$ and $p=0.001$, respectively). A significant negative correlation was obtained between ALDH1 expression and ER-negative ($p=0.002$) and PR-negative status ($p=0.001$). Even though a higher mean rank for CD44 expression was observed in ER- and PR-negative patients, the differences were not significant

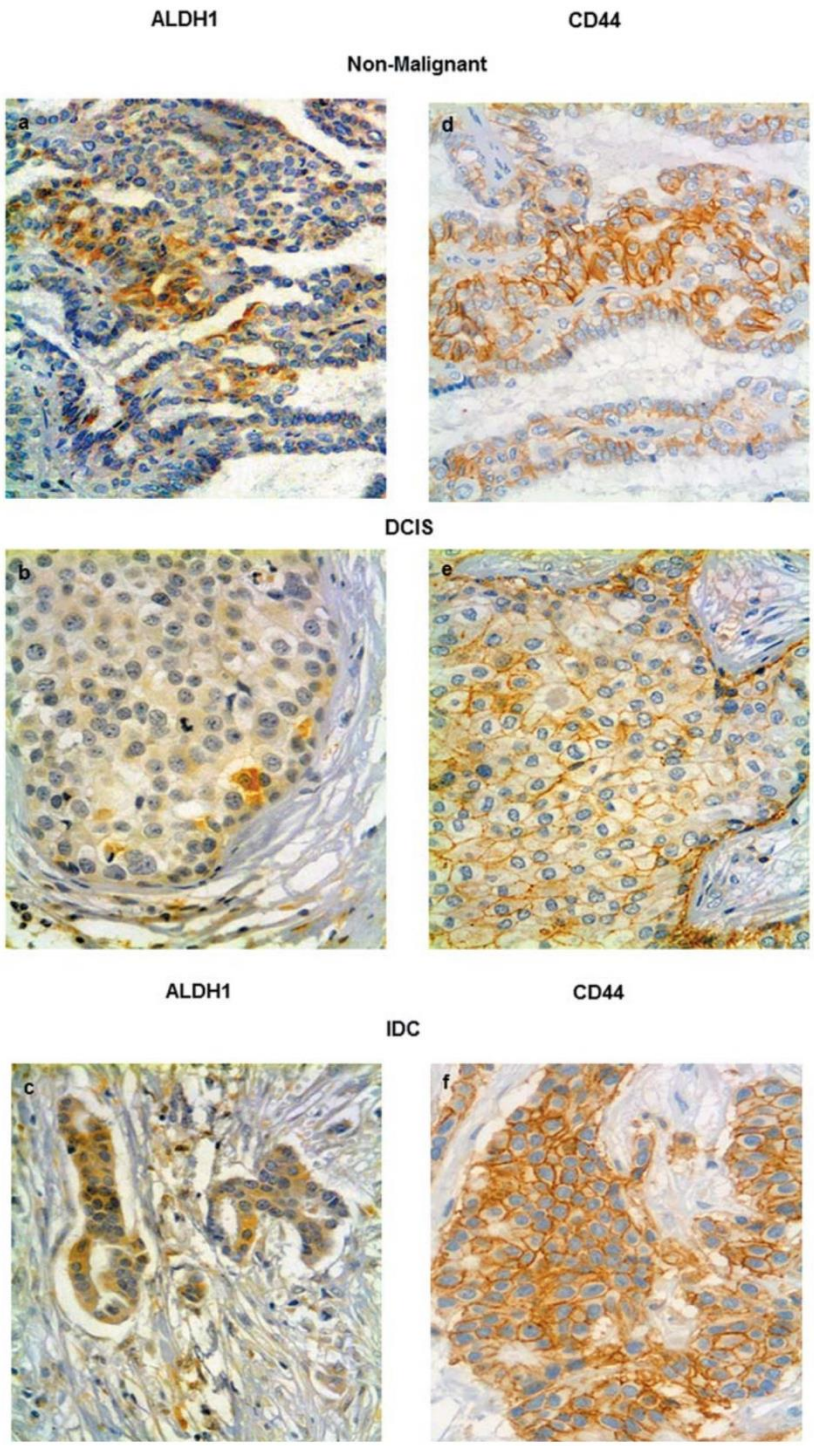

Figure 1. *Representatives images of ALDH1 staining in a) non-malignant (×20), b) DCIS (×40) and c) IDC samples (×20) and CD44 staining in d) non-malignant (×20), e) DCIS (×40) and f) IDC samples (×40).*

Table V. *ALDH1 and CD44 expression according to patients' hormone receptor status.*

	ALDH1 expression score				CD44 expression score			
Characteristics	Mean	*p*-Value	Correlation coefficient	*p*-Value	Mean	*p*-Value	Correlation coefficient	*p*-Value
ER status								
Positive	59.77	0.002[a]	–0.270	0.002[b]	61.11	0.511[a]	–0.059	0.514[b]
Negative	86.65				66.89			
PR status								
Positive	58.13	0.001[a]	–0.297	0.001[b]	62.19	0.915[a]	0.01	0.915[b]
Negative	83.33				61.40			

ER, Estrogen receptor; PR, progesterone receptor. [a]Mann-Whitney test; [b]Non-parametric correlation.

(p>0.05) (Table V). No significant correlation was observed between CD44 expression and ER- and PR-status.

Correlation and simultaneous expression of the stem-cell markers ALDH1 and CD44. The presence of ALDH1 and CD44 immunoreactivity was positively correlated (Spearman's R=0.153; p=0.008), as well as CD44 mean rank expression was correlated with ALDH1 mean rank expression (Spearman's R=0.261; p<0.001) (data not shown). Based on these results, we created a score representing a composite measure for the expression of both markers. A higher mean rank expression for the combined SC score was observed in DCIS samples, followed by non-malignant lesions and then IDC. The mean rank expression difference between the groups is highly significant (p<0.001) (Table VI).

Table VI. *ALDH1 and CD44 combined score mean rank expression.*

Characteristics Type of Lesion	Mean	*p*-Value
Non-Malignant	161.82	<0.001[a]
DCIS	169.57	
IDC	134.58	

DCIS, Ductal carcinoma *in situ*; IDC, invasive ductal carcinoma. [a]Kruskal-Wallis test

Discussion

The aim of the present study was to explore a possible association between a high expression of SC markers and a determined type of lesion during breast cancer progression. For this purpose, ALDH1 and CD44 expression was immunohistochemically-evaluated in non-malignant breast lesions, ductal carcinoma *in situ* and invasive ductal carcinoma sample cores. However, our study has some limitations. Since putative CSCs were originally identified using flow cytometry we have assumed that this assay can be translated into an immunohistochemistry-based equivalent. Although we would expect these modalities to identify a population with a high degree of overlapping, it is probable that there will be some discordance. Besides that we have used TMAs to detect for a sub-population of cells of reputed scarcity and as a result it is probable there is some sampling error. In order to minimize these errors, we selected more than one area of breast cancer specimens to be representative of the entire tissue block, thus increasing our cohort population. We decided to exclude *in situ* and invasive lobular carcinomas due to the reduced number of blocks containing these types of lesions, not being representative for this expression study. In fact, invasive lobular carcinomas counts for 5 to 15% of different races worldwide, contrarily to 70 to 90% of women having invasive ductal carcinomas (16). Moreover, these two types of breast cancer are histologically- and genetically-different, which explains the restriction to the immunoexpression analysis performed in *in situ* and invasive ductal carcinomas.

Several studies have assessed ALDH1 expression by immunohistochemistry in breast tissue, either in non-malignant or malignant breast tumors in order to define ALDH1 role and impact in predicting cancer development. Kunju *et al*. reported that ALDH1 is expressed in epithelial and stromal cells in benign breast tissues, and discovered that ALDH1 positivity in breast epithelial cells is associated with increased risk of breast cancer (17). Madjd and his colleagues verified a high percentage of breast tumors with ALDH1 positive cells (86%) (18). In our study, ALDH1 expression was observed in 85.43% of all samples, with a homogenous distribution in the different groups of assessed lesions. Considering the fact that ALDH1 was already demonstrated to be a potential marker of normal and malignant human BSCs (6) a high percentage of non-malignant breast tissues samples with ALDH1-positive cells would be expected. However, samples with less than 5% of

ALDH1-positive cells were also counted as ALDH1-positive which increases the possibility of having a high number of positive samples. In fact, the vast majority of samples had a relatively small amount of ALDH1-positive cells, which can be consistent with the theory that SCs comprise of a minority of tumor tissues. Furthermore, all normal and hyperplasic breast tissue samples were selected and collected from cancer specimens. These may not be representative of truly benign tissue, due to genetic alterations not affecting cell architecture (19). Moreover, no significant differences were observed between the lesion groups regarding ALDH1 expression, but we obtained a negative significant correlation between ALDH1 expression and ER-negative and PR-negative status. These results are in accordance with the majority of others studies done so far regarding ALDH1 expression in breast carcinogenesis (23, 25, 26) and studies that have correlated ALDH1$^+$ tumors with negativity for estrogen and progesterone receptors (12, 25, 27-29).

Interest in CD44 has stemmed from reports showing that high levels of CD44 expression in combination with low-level expression of CD24 can be used to prospectively identify a population of breast cancer cells enriched in SC-like properties and tumor-initiating capacity (3). CD44 comprises a family of cell surface receptors that recognize hyaluronan, a component of the extracellular matrix, as their principal ligand (20). Multiple splice variants of CD44 exist, which have a more restricted expression (21). Moreover, CD44 has been associated with SC in normal breast tissue (22). In breast carcinomas, Auvinen and colleagues demonstrated that CD44 expression was common in carcinoma cells and only in 8% of cases, were all carcinoma cells CD44-negative (23). The same authors detected the expression of an isoform of CD44, CD44v6 in 20-30% of ductal epithelial cells in benign lesions of breast (24). Moreover, Bankfalvi *et al*. investigated whether the CD44 immunophenotype of breast lesions correlated with the clinical evolution and prognosis of breast cancer. They found that in normal breast tissue luminal epithelial cells lacked detectable CD44 in contrast to basal cells, which constitutionally expressed CD44s, v.3, v.5 v.6 and v.9 isoforms (25). Our results showed positivity for CD44 in 90.3% of all samples with homogenous distribution between the lesion groups. No significant differences were observed between the lesion groups relatively to CD44 expression and no significant correlation was seen between CD44 expression and ER and PR status.

Cell surface markers and enzymatic activity detected by fluorescence-activated cell sorting (FACS) have been widely used for the prospective isolation of putative CSCs. The idea of combining markers to increase the purity of sub-populations for CSCs was utilized by Ginestier *et al*. who showed that the combination of CD44$^+$/CD24$^{-/low}$ and ALDEFLUOR activity enabled for isolation of cells able to form tumors in NOD/SCID mice from as few as 20 cells, compared to 500 cells when sorted by ALDEFLUOR activity alone (12). Neumeister and colleagues developed an *in situ* method to define CSCs in FFPE breast cancer tissues, with the goal of assessing the prognostic value of the presence of breast cancer stem cells (BCSCs) using a multiplexed assay for CD44 and ALDH1. They found a strong co-expression between these two markers in breast tumors associated with poor prognosis (26). In our study, and considering the fact that both markers are positively-correlated we decided to combine ALDH1 and CD44 expression in order to determine if both markers had a significant co-expression in the different lesion groups. This co-expression was significantly higher in *in situ* carcinomas when compared against the other two groups.

According to the CSC model, cancers originate from the malignant transformation of an adult SC or a progenitor through the de-regulation of the normally tightly-regulated self-renewal program. This leads to clonal expansion of stem/progenitor cells that undergo further genetic or epigenetic alterations to become fully-transformed. As a consequence, tumors contain a cellular component of CSCs which retains the key stem cell properties that may initiate and drive carcinogenesis (1). Regarding the roles of ALDH1 and CD44, it is plausible to associate their expression in transformation of BSCs in to BCSCs. The progression to a malignant phenotype involves local metastasis and invasion, two processes in which the cell-cell and cell-extracellular matrix adhesion are altered. Along the tumoral progression, CD44 can be involved in the cellular matrix adhesion, cellular matrix degradation, cancer cells migration and angiogenesis (27). On the other hand, ALDH1 plays a role in early differentiation of stem cells by promoting the formation of retinoic acid where the retinoic signaling is been directly implicated in modulating BCSCs differentiation (28). If this cellular component is present in breast tumors with stem cell properties that initiate and drive carcinogenesis, after their differentiation through symmetrical division, the formed cancer cells will have a lower expression of ALDH1 than BCSCs. Besides that, cancer cells other than BCSCs can metastasize *in vivo* or in experimental models, indicating that cells with an invasive phenotype can be found outside the BCSC pool (29). Thus, this could explain the combined overexpression between ALDH1 and CD44 in *in situ* carcinomas obtained in this study, a type of lesion supposed to be a transition of non-invasive to an invasive phenotype in breast cancer (30).

In conclusion, no ALDH1 and CD44 overexpression was obtained between the different types of assessed breast lesions. Moreover, the co-expression between these markers was shown to be significantly higher in *in situ* carcinomas when compared to the other two lesions groups. Regarding the CSC model theory, it is plausible to associate these findings with the malignant transformation of BSCs in the

carcinogenesis process of the breast. Nevertheless more studies have to be performed, other than immunoexpression analysis, to consolidate this association.

Acknowledgements

The Authors would like to acknowledge the Foundation for Sciences and Technology (Portugal) for funding this study. Scholarship reference: SFRH/BD/74307/2010.

Conflicts of Interest

The Authors declare that they have no conflicts of interest.

References

1 Shackleton M: Normal stem cells and cancer stem cells: similar and different. Semin Cancer Biol *20(2)*: 85-92, 2010.

2 Charafe-Jauffret E, Ginestier C, Bertucci F, Cabaud O, Wicinski J, Finetti P, Josselin E, Adelaide J, Nguyen TT, Monville F, Jacquemier J, Thomassin-Piana J, Pinna G, Jalaguier A, Lambaudie E, Houvenaeghel G, Xerri L, Harel-Bellan A, Chaffanet M, Viens P and Birnbaum D: ALDH1-Positive cancer stem cells predict engraftment of primary breast tumors and and are governed by a common stem cell program. Cancer Res *73(24)*: 7290-7300, 2013.

3 Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ and Clarke MF: Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci USA *100(7)*: 3983-3988, 2003.

4 Sleeman KE, Kendrick H, Robertson D, Isacke CM, Ashworth A and Smalley MJ: Dissociation of estrogen receptor expression and *in vivo* stem cell activity in the mammary gland. J Cell Biol *176(1)*: 19-26, 2007.

5 Croker AK, Goodale D, Chu J, Postenka C, Hedley BD, Hess DA and Allan DL: High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability. J Cell Mol Med *13(8B)*: 2236-2252, 2009.

6 Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, Jacquemier J, Viens P, Kleer CG, Liu S, Schott A, Hayes D, Birnbaum D, Wicha MS and Dontu G: ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell *1(5)*: 555-67, 2007.

7 Klingbeil P, Natrajan R, Everitt G, Vatcheva R, Marchio C, Palacios J, Buerger H, Reis-Filho JS and Isacke CM: CD44 is overexpressed in basal-like breast cancers but is not a driver of 11p13 amplification. Breast Cancer Res Tr *120(1)*: 95-109, 2010.

8 Magni M, Shammah S, Schiro R, Mellado W, Dalla-Favera R and Gianni AM: Induction of cyclophosphamide-resistance by aldehyde-dehydrogenase gene transfer. Blood *87(3)*: 1097-1103, 1996.

9 Sophos NA and Vasiliou V: Aldehyde dehydrogenase gene superfamily: the 2002 update. Chem Biol Interact *143-144*: 5-22, 2003.

10 Yoshida A, Rzhetsky A, Hsu LC and Chang C: Human aldehyde dehydrogenase gene family. Eur J Biochem *251(3)*: 549-557, 1998.

11 Dylla SJ, Beviglia L, Park IK, Chartier C, Raval J, Ngan L, Pickell K, Aguilar J, Lazetic S, Smith-Berdan S, Clarke MF, Hoey T, Lewicki J and Gurney AL: Colorectal cancer stem cells are enriched in xenogeneic tumors following chemotherapy. PLoS One *3(6)*: e2428, 2008.

12 Huang EH, Hynes MJ, Zhang T, Ginestier C, Dontu G, Appelman H, Fields JZ, Wicha MS, and Boman BM: Aldehyde Dehydrogenase 1 Is a Marker for Normal and Malignant Human Colonic Stem Cells (SC) and Tracks SC Overpopulation during Colon Tumorigenesis. Cancer Research *69(8)*: 3382-3389, 2009.

13 Jalkanen S, Bargatze RF, de los Toyos J and Butcher EC: Lymphocyte recognition of high endothelium: antibodies to distinct epitopes of an 85-95-kD glycoprotein antigen differentially inhibit lymphocyte binding to lymph node, mucosal, or synovial endothelial cells. J Cell Biol *105(2)*: 983-990, 1987.

14 Ponta H, Sleeman J, Dall P, Moll J, Sherman L and Herrlich P: CD44 isoforms in metastatic cancer. Invasion Metastasis *14(1-6)*: 82-86, 1994.

15 Farnie G, Clarke RB, Spence K, Pinnock N, Brennan K, Anderson NG and Bundred NJ: Novel cell culture technique for primary ductal carcinoma *in situ*: role of Notch and epidermal growth factor receptor signaling pathways. J Natl Cancer Ins *99(8)*: 616-627, 2007.

16 Zhao H, Langerod A, Ji Y, Nowels KW, Nesland JM, Tibshirani R, Bukholm IK, Karesen R, Botstein D, Borresen-Dale AL and Jeffrey SS: Different gene expression pattterns in invasive lobular and ductal carinomas of the breast. Mol Biol Cell *15(6)*: 2523-2536, 2004.

17 Kunju LP, Cookingham C, Toy KA, Chen W, Sabel MS and Kleer CG: EZH2 and ALDH-1 mark breast epithelium at risk for breast cancer development. Modern Pathol *24(6)*: 786-793, 2011.

18 Madjd Z, Ramezani B, Molanae S and Asadi-Lari M: High Expression of Stem Cell Marker ALDH1 is Associated with Reduced BRCA1 in Invasive Breast Carcinomas. Asian Pac J Cancer Prev *13(6)*: 2973-2978, 2012.

19 Tsai YC, Lu Y, Nichols PW, Zlotnikov G, Jones PA and Smith HS: Contiguous patches of normal human mammary epithelium derived from a single stem cell: implications for breast carcinogenesis. Cancer Res *56(2)*: 402-404, 1996.

20 Diaz LK, Zhou X, Wright ET, Cristofanilli M, Smith T, Yang Y, Sneige N, Sahin A and Gilcrease MZ: CD44 expression is associated with increased survival in node-negative invasive breast carcinoma. Clin Cancer Res *11(9)*: 3309-3314, 2005.

21 Mackay CR, Terpe HJ, Stauder R, Marston WL, Stark H and Gunthert U: Expression and modulation of CD44 variant isoforms in humans. J Cell Biol *124(1-2)*: 71-82, 1994.

22 Hebbard L, Steffen A, Zawadzki V, Fieber C, Howells N, Moll J, Ponta H, Hofmman M and Sleeman J: CD44 expression and regulation during mammary gland development and function. J Cell Sci *113(Pt 14)*: 2619-2630, 2000.

23 Auvinen P, Tammi R, Kosma VM, Sironen R, Soini Y, Mannermaa A Tumelius R, Uljas E and Tammi M: Increased hyaluronan content and stromal cell CD44 associate with HER2 positivity and poor prognosis in human breast cancer. Int J Cancer *132(3)*: 531-39, 2012.

24 Auvinen P, Tammi R, Tammi M, Johansson R and Kosma VM: Expression of CD 44 s, CD 44 v 3 and CD 44 v 6 in benign and malignant breast lesions: correlation and colocalization with hyaluronan. Histopathology *47(4)*: 420-428, 2005.

25 Bankfalvi A, Terpe HJ, Breukelmann D, Bier B, Rempe D, Pschadka G, Krech R and Bocker W: Gains and losses of CD44 expression during breast carcinogenesis and tumour progression. Histopathology *33(2)*: 107-116, 1998.

26 Neumeister V, Agarwal S, Bordeaux J, Camp RL and Rimm DL: *In situ* identification of putative cancer stem cells by multiplexing ALDH1, CD44, and cytokeratin identifies breast cancer patients with poor prognosis. Am J Pathol *176(5)*: 2131-2138, 2010.

27 Herrera-Gayol A and Jothy S: Adhesion proteins in the biology of breast cancer: contribution of CD44. Exp Mol Pathol *66(2)*: 149-156, 1999.

28 Ginestier C, Wicinski J, Cervera N, Monville F, Finetti P, Bertucci F, Wicha MS, Birnbaum D and Charafe-Jauffret M: Retinoid signaling regulates breast cancer stem cell differentiation. Cell Cycle *8(20)*: 3297-3302, 2009.

29 Velasco-Velazquez MA, Homsi N, De La Fuente M and Pestell RG: Breast cancer stem cells. Int J Biochem Cell B *44(4)*: 573-577, 2012.

30 Volinia S, Galasso M, Sana ME, Wise TF, Palatini J, Huebner K and Croce CM: Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA. P Natl Acad Sci USA *109(8)*: 3024-3029, 2012.

Received December 17, 2013
Revised January 13, 2014
Accepted January 15, 2014

Chapter 4

Characterization of $CD44^+$ $ALDH1^+$ $Ki\text{-}67^-$ Cells in Non-malignant and Neoplastic Lesions of the Breast

ANTICANCER RESEARCH *36*: 4629-4638 (2016)
doi:10.21873/anticanres.11013

Characterization of CD44+ALDH1+Ki-67− Cells in Non-malignant and Neoplastic Lesions of the Breast

ARNAUD DA CRUZ PAULA[1,2], ORIANA MARQUES[1,3,4], RITA SAMPAIO[5], ANA ROSA[1,4], JOSÉ GARCIA[5], ALEXANDRA RÊMA[1], MARIA DE FÁTIMA FARIA[1], PAULA SILVA[6,7], RAMÓN VIZCAÍNO[5] and CARLOS LOPES[1,2]

[1]Pathology and Molecular Immunology Department, [3]Unit for Multidisciplinary Biomedical Research, Abel Salazar Institute of Biomedical Sciences, [7]Science Diffusion, Institute of Molecular Pathology and Immunology, and [6]Faculty of Medicine, University of Porto, Porto, Portugal;
[2]Department of Pathology, Portuguese Oncology Institute, Porto, Portugal;
[4]Basic and Clinical Research on Iron Biology, Institute of Molecular and Cell Biology/i3S, Porto, Portugal;
[5]Department of Pathology, Porto Hospital Centre, Porto, Portugal

Abstract. *Background: Cancer stem cells are tumor cells that present self-renewal, clonal tumor initiation capacity and clonal long-term repopulation potential. We have previously demonstrated that the co-expression of the breast cancer stem cell (BCSC) markers hyaluronan receptor (CD44) and aldehyde dehydrogenase-1 (ALDH1) in ductal carcinomas in situ could be determinant for disease progression. Combining these established BCSC markers with Ki-67 to evaluate quiescence we sought to identify, evaluate the distribution and estimate the mean percentages of CD44+ALDH1+Ki-67− breast cells. Materials and Methods: Triple-immunohistochemistry for CD44, ALDH1 and Ki-67 was applied in a series of 16 normal, 54 non-malignant and 155 malignant breast tissues. Clinical relevance was inferred by associations with markers of breast cancer behavior, progression and survival. Results: The mean percentages of cells with this phenotype increased significantly from non-malignant lesions to high-grade ductal carcinomas in situ, decreasing in invasive ductal carcinomas, as also evidenced by an inverse correlation with histological grade and tumor size. The mean percentage of CD44+ALDH1+Ki-67− cells was also significantly higher in women who developed distant metastasis and died due to breast cancer, and a significant association with human epidermal growth factor type 2 (HER2) negativity was observed. Conclusion: Our novel findings indicate that CD44+ALDH1+Ki-67− tumor cells may favor distant metastasis and can predict overall survival in patients with ductal carcinomas of the breast. More importantly, quiescence may have a crucial role for tumor progression, treatment resistance and metastatic ability of BCSCs.*

Correspondence to: Arnaud Da Cruz Paula and Carlos Lopes, Pathology and Molecular Immunology Department, Abel Salazar Institute of Biomedical Sciences, Jorge Viterbo Ferreira Street, n° 288, 4050-313, Porto, Portugal. Tel: +351 967693532, +351 917270209, Fax: +351 225511184, e-mail: arnaudcpaula@hotmail.com/ calopes@icbas.up.pt

Key Words: Breast cancer, breast cancer stem cell markers, ALDH1, CD44, quiescence, cancer stem cell model.

The cancer stem cell (CSC) model is an attractive concept to explain several poorly understood clinical phenomena due to its inherent theoretical properties. Such properties are based on the molecular features of normal stem cells. Thus, CSCs have the ability to renew themselves and last a lifetime but also to be resistant to electromagnetic and chemical insults. Such resistance ability allows them to survive for long periods of time and, consequently, colonize other parts of the body (1). The existence of tumor cells presenting stem cell features has been well established in the literature for specific cancer types; however, and due to functional heterogeneity existing in all tumors, no single CSC phenotype can be generalized. As a consequence, distinct cell populations within a unique tumor may exhibit CSC features, which has led to identification of other putative CSC subsets in a diversity of solid tumor types (2). Despite this, no single protocol or even combined protocols are guaranteed to obtain pure CSC subsets and most recent iterations to define CSCs have embraced a definition of the CSC phenotype as a dynamic cell state rather than a distinct cell type (3).

Apart from the clonogenic features of CSCs (self-renewal and differentiation tumorigenicity), several associations have been made between putative CSCs and their normal counterparts. In order to identify CSCs within the heterogeneous tumor bulk, numerous factors such as surface marker expression, cell-cycle state, migratory properties, immune

escape, or metabolic and transporter activities are being studied (4). In breast cancer, a population of hyaluronan receptor (CD44)$^+$CD24^{-low} tumor cells has been demonstrated to have tumor-initiating properties. This tumorigenic phenotype has been associated with stem cell-like characteristics, enhanced invasive properties, radiation resistance and with distinct genetic profiles suggesting association with adverse prognosis (5-7). However, the notion that the CD44$^+$CD24^{-low} surface markers are enriched for tumorigenic cells cannot be applied to all cases of breast cancer. Thus, the validity of the combination of these markers as a definition of breast cancer stem cells (BCSCs) has been called into question and additional markers such as aldehyde dehydrogenase-1 (ALDH1) have been reported (8, 9).

More recently, subpopulations of ALDH1^{high}CD44$^+$ cells were identified in several human breast cancer cell lines, which contributed to both chemotherapy and radiation resistance, suggesting a much broader role for ALDH1 in treatment response than previously reported (10). Indeed, an increase in the population of ALDH1$^+$ cells but not CD44$^+$CD24^{-low} cells has been observed in breast cancer tumor biopsies after neoadjuvant treatment (11).

Nevertheless, a potentially more challenging problem is the recent observation that CSCs that display quiescent properties may exist. The isolation of adult stem cells revealed new insights on the epigenetic, transcriptional and post-transcriptional control of quiescence, proposing an actively preserved state of quiescence, which is regulated by signaling pathways that sustain a controlled state, allowing rapid activation (12). Although these quiescent cells are slowly dividing, they possess increased sphere-forming aptitude *in vitro*, suggesting that these cells are enriched in CSCs, being able to repair DNA damage induced by chemotherapeutic agents and radiotherapy (13). Such dormant cells were also identified in pancreatic adenocarcinoma and shown to be enriched for CSC markers such as CD133, CD44, CD24 and ALDH1 (12).

We recently reported a higher combined expression of CD44 and ALDH1 in ductal carcinomas *in situ* (DCIS) when compared with invasive ductal carcinomas (IDCs) (14). Regarding these results and combining these established BCSC markers with Ki-67 as a marker of quiescence, we aimed to identify, evaluate its distribution and estimate the mean percentage of CD44$^+$ALDH1$^+$Ki-67$^-$ breast cells in a series of normal, non-malignant and malignant breast tissues. This phenotype was further correlated with clinicopathological markers of breast cancer prognosis and overall survival.

Materials and Methods

Patient samples. Formalin-fixed, paraffin-embedded (FFPE) breast specimens of 229 patients who underwent breast surgery prior to systemic treatment from 2004 to 2012 were obtained from the archives of the Pathology Department of Santo António Hospital (Porto Hospital Centre, Porto, Portugal). At the time of diagnosis, 59 patients were considered to have benign breast lesions (mean age of 43.63 years), including adenosis (ADs), fibroadenomas (FADs) and ductal hyperplasias (DHPs) without signs of malignancy. Regarding the malignant cases, 170 were retrieved including 28 pure DCIS (without signs of invasion), 34 DCIS within IDC and 108 IDCs. As controls, normal breast tissue was obtained from 20 women who underwent reduction mammoplasty with no previous history of breast cancer (mean age of 32 years). Clinicopathological information was obtained by reviewing pathology reports, hematoxylin and eosin (H&E)-stained sections and also H&E-stained tissue microarray sections for a more accurate information of each tissue core. The following histopathological variables for the pure DCIS and DCIS within IDC samples were available in the interim records: nuclear grade, estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor type 2 (HER2) status. For IDC cases, we retrieved the Elston–Ellis histological grade, tumor size, lymph-node status, local recurrences, distant metastasis, mortality, ER, PR and HER2 status. All patients were female, with a mean age of 55.33 years. Regarding the presence of distant metastasis, none of the patients had distant metastasis at the time of diagnosis; hence, M1 was analyzed in patients who developed distant metastasis during follow-up, of which the mean time was 68.18 (SD=23) months. All cases were reviewed by two pathologists (JG and CL). FFPE tissue samples were arrayed using a tissue-arraying instrument (Thermo Scientific, USA) with a punch extractor of 2 mm. Each sample was arrayed in duplicate or in triplicate to minimize tissue loss and compensate for tumor heterogeneity. Representative areas with malignant lesions from DCIS and IDC were classified as pure DCIS or IDC alone, respectively. DCIS cores retrieved from IDC samples, without the invasive compartment, were classified as DCIS within IDC. The current study was approved by the following ethical boards: Porto Hospital Centre Research Ethics Health Committee (reference 203-CES) and by Porto Hospital Centre Department of Education, Development and Research (reference 135-DEFI).

Triple immunohistochemistry. Considering our previous results obtained by single-staining immunohistochemistry of CD44 and ALDH1, our cohort for this study was composed of some autologous samples of DCIS and IDCs from those previously used (14), which served as a control for triple immunohistochemistry validation.

For CD44, ALDH1 and Ki-67 triple immunohistochemistry, 3 μm-thick tissue microarray sections were cut, deparaffinized, rehydrated in a series of graded ethanol and washed in water. Target retrieval was achieved with citrate buffer (pH 6.0) in a microwave until boiling. Slides were then incubated for 10 min with 3% peroxide hydrogen in methanol, and further with blocking solution (Ultra Vision LP Detection System; Thermo Fisher Scientific, Cheshire, UK), for 5 min. CD44 (mouse monoclonal anti-human, 0.01 mg/ml concentration, MRQ-13, dilution: 1100, Cell Marque, Rocklin, CA, USA) and Ki-67 (rabbit monoclonal anti-human, 0.01 mgml concentrate, SP6, dilution: 1100, Cell Marque) were incubated together in a humid chamber for 2 hours at room temperature. Following this, immunohistochemistry was performed according to HiDef Detection™ Horseradish Peroxidase Polymer System procedures (Cell Marque). Enzyme reactivity was visualized using 3,3'-diaminobenzidine tetrahydrochloride (DAB; Sigma, Saint

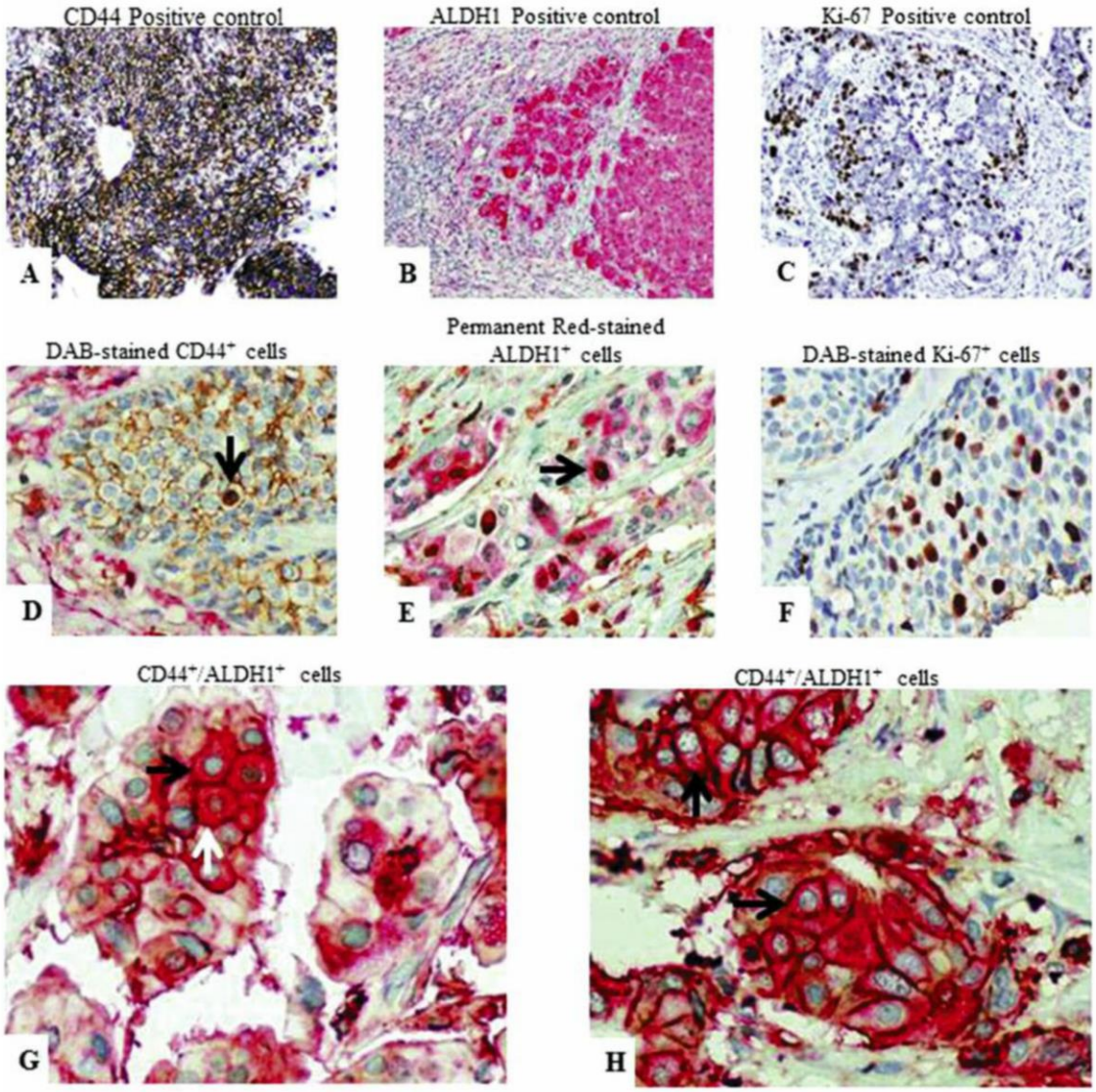

Figure 1. *Identification of hyaluronan receptor (CD44)+ aldehyde dehydrogenase-1 (ALDH1)+ Ki-67− breast cells. Positive controls for CD44 (A), ALDH1 (B) and Ki-67 (C) (×20). D: Illustrates a breast tissue sample without ALDH1 expression; black arrow points to a CD44+ALDH1−Ki-67+ breast cell (×40). E: A breast tissue sample without CD44 expression; black arrow points a CD44−ALDH1+Ki-67+ breast cell (×40). F: A breast tissue sample without CD44 and ALDH1 expression (×40). G, H: Breast tissue samples containing CD44+ALDH1+Ki-67− breast cells (black arrows); white arrow in (G) points to a CD44+ALDH1+Ki-67+ breast cell (×40).*

Louis, MO, USA). Slides were washed in running water and incubated with glycine buffer (pH 2.2) for 45 min at 37˚C. ALDH1 (rabbit monoclonal anti-human, 0.13 mgml concentrate, dilution: 1100; Abcam, Cambridge, UK) was then added and slides incubated in a humid chamber for 2 h at room temperature. HiDef DetectionTm Alkaline Phosphatase Polymer System procedures (Cell Marque) were followed and enzyme reactivity was visualized using Permanent Red chromogen (Permanent Red Chromogen Kit; Cell Marque). Finally, slides were counterstained with Mayer's hemalum solution (Merck Millipore, Darmstadt, Germany) and mounted with Aquatex Mounting Medium (Sigma-Aldrich, Darmstadt, Germany). For CD44, ALDH1 and Ki-67, positive controls were also used including chorion, liver and breast tissues, respectively (Figure 1A-C).

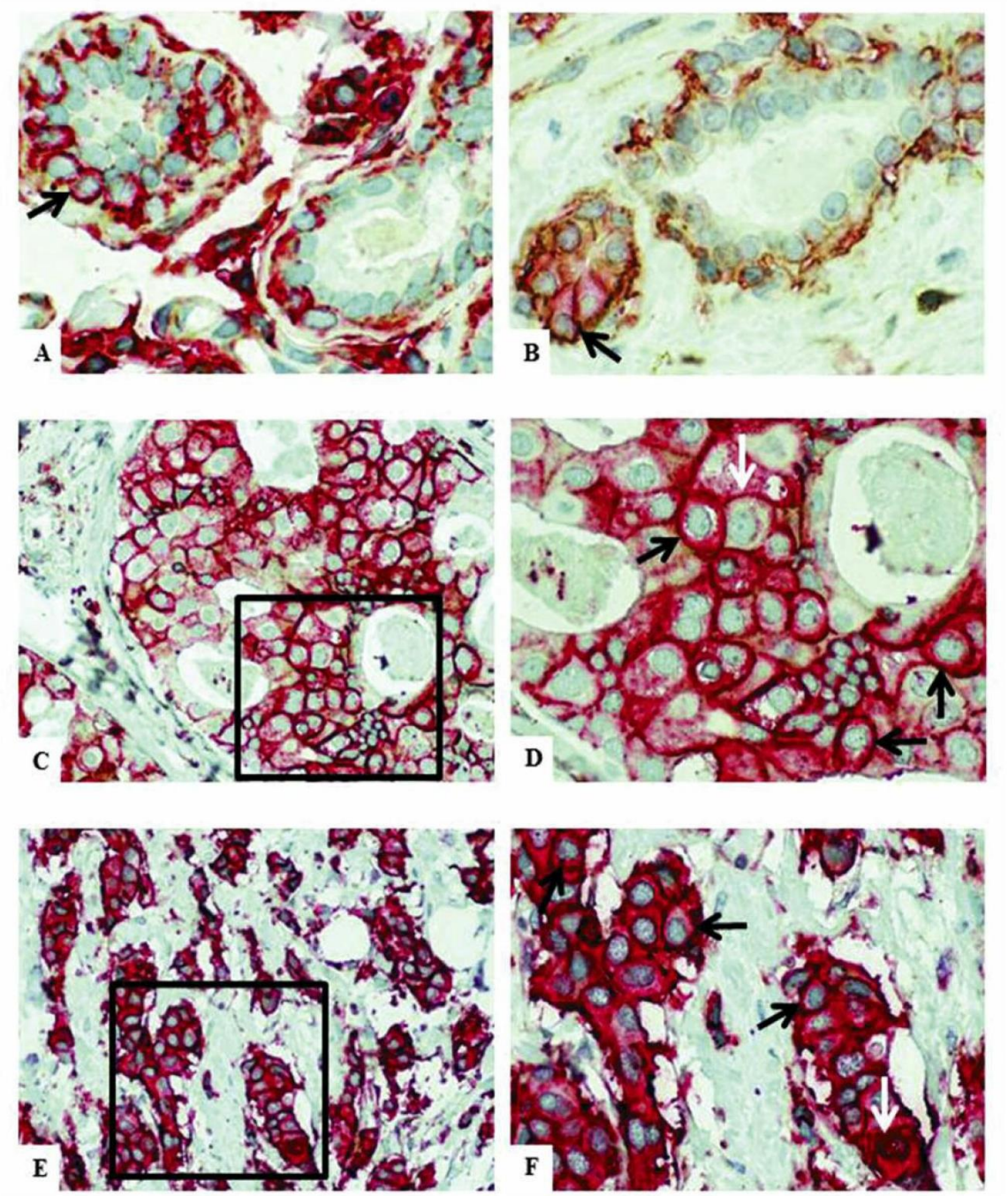

Figure 2. *Representative images of the triple immunohistochemistry for hyaluronan receptor (CD44), aldehyde dehydrogenase-1 (ALDH1) and Ki-67 in breast tissue sections. Normal (A) and non-malignant (Fibroadenoma) (B) breast tissue section; black arrows points to CD44⁺ALDH1⁺Ki-67⁻ breast cells (×40). Overview of ductal carcinoma in situ (C) and invasive ductal carcinoma (E) breast tissue (×20). Insets (D and F) show the same tissues at higher magnification (×40). D: Examples of CD44⁺ALDH1⁺Ki-67⁻ tumor cells (black arrows); the white arrow demonstrates an example of a cell without ALDH1 expression. F: Examples of CD44⁺ALDH1⁺Ki-67⁻ invasive tumor cells (black arrows); the white arrow demonstrates an example of a CD44⁺ALDH1⁺Ki-67⁺ cell.*

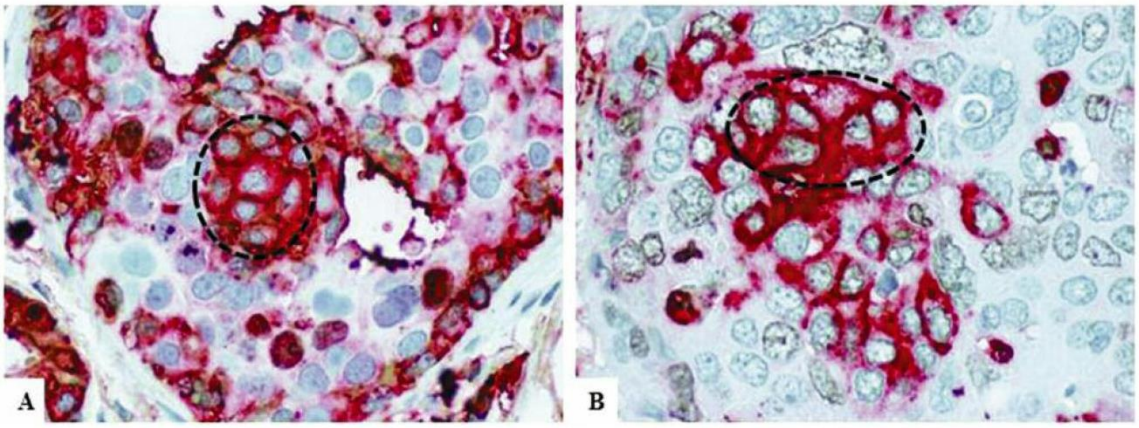

Figure 3. *Pools of hyaluronan receptor (CD44)$^+$ aldehyde dehydrogenase-1 (ALDH1)$^+$ Ki-67$^-$ breast cells (dotted circles) in ductal carcinoma in situ (A) and invasive ductal carcinoma (B) breast tissues respectively (×40).*

Definition of hormone receptor status. Primary breast tumors that expressed nuclear staining in 1% of tumor cells were regarded as positive for ER and PR. Immunostaining results for HER2 were scored as 0 when no staining was observed or when <10% of tumor cells had membranous staining; as 1 when a faint or barely membranous staining was present in >10% of tumor cells; as 2 when a weak to moderate membrane staining was observed in >10% of tumor cells and as 3 when a strong complete membrane staining was present in >10% of tumor cells. All cases classified previously as 2 were subjected to fluorescence *in situ* hybridization for determination of *HER2* gene amplification using the *HER2* DNA probe kit (Abbott Laboratories, Abbott Park, IL, USA). Scores of 0 and 1 were considered to be negative for HER2, while scores of 2 (if *HER2* gene amplification was observed) and 3 were considered to be positive.

Pathological evaluation. After omitting 20 cases with uninterpretable immunohistochemistry results, a total of 209 cases were informative (10 ADs, 21 FADs, 23 DHPs, 25 pure DCIS, 30 DCIS within IDC and 100 IDCs). Using light microscopy (Olympus U-SPO3;Olympus Corporation, Tokyo, Japan), stained tissue sections were inspected by two pathologists (RS and CL) and a trained scientist (ACP) without knowledge of the diagnosis. To unambiguously identify CD44$^+$ALDH1$^+$Ki-67$^-$ cells, we needed to distinguish double-stained cells labeled with both DAB and Permanent Red, considered to be CD44$^+$ALDH1$^+$ cells from DAB-stained CD44$^+$ cells (Figure 1D) and Permanent Red-stained ALDH1$^+$ cells (Figure 1E). As for Ki-67, a clear distinction was seen between nuclear DAB-stained Ki-67$^+$ from Ki-67$^-$ cells (Figure 1F). Consequently, cells with brown membranes, red cytoplasm and hematoxylin stained-nucleus were considered to be CD44$^+$ALDH1$^+$Ki-67$^-$ (Figure 1G and 1H). Regarding the previous results obtained from single-immunohistochemistry of CD44 and ALDH1 and considering that we used the same antibodies, we discriminated our cells of interest only in breast epithelial cells, excluding artifacts staining tumor cell debris. All slides were scanned (Leica SCN400, Meyer Instruments, Houston, TX, USA) and digital images of each core were visualized through the SlidePath Gateway 2.0 System (Leica Biosystems, Wetzlar, Germany) program. These digitalized images allowed us to count more accurately the number of cells of interest. Percentages of such cells were estimated from the entire lesion areas. Cores from the same donor tissue diagnosed with the same histological type were grouped and their mean score calculated.

Statistical analysis. Statistical analysis was performed using IBM SPSS Statistics version 20.0 software (IBM Corp., Armonk, NY, USA). Sample distributions were compared using Kruskal–Wallis or Mann–Whitney tests. Pearson's chi-square test was used to evaluate the differences between categorical variables. Spearman's rank correlation coefficient was used to evaluate the relationship between variables, and the Kaplan–Meier method was used for univariate survival analysis. Statistical significance was accepted at $p<0.05$.

Results

CD44$^+$ALDH1$^+$Ki-67$^-$ cell populations in breast samples. ALDH1, CD44 and Ki-67 immunostaining was observed in epithelial cells and also in stromal inflammatory cells, both in normal, non-malignant and malignant samples. While CD44 exhibited mainly membranous staining, ALDH1 was almost exclusively detected in the cytoplasm and Ki-67 in the nuclei. We were able to identify CD44$^+$ALDH1$^+$Ki-67$^-$ cells in normal (Figure 2A), non-malignant (Figure 2B) and malignant breast tissues (Figure 2C-F). Interestingly, we were also able to identify small pools of these cells in some malignant breast tissues and although in rare cases, these pools of cells were observed completely isolated (Figure 3A and 3B).

Considering the low percentages of CD44$^+$ALDH1$^+$Ki-67$^-$ cells identified in normal and non-malignant breast tissues, raw percentages of these cells for mean calculation

Table I. *Prevalence of hyaluronan receptor (CD44)+ aldehyde dehydrogenase-1 (ALDH1)+ Ki-67− cells in breast tissue samples.*

Characteristic	n (%)	CD44+ALDH1+Ki-67− breast cells <1%, n (%)	<10%, n (%)	≥10%, n (%)
Type of lesion				
Normal	16 (7.11)	14 (87.50)	2 (12.50)	0 (0)
Non-malignant	54 (24)	42 (77.78)	11 (20.37)	1 (1.85)
Malignant	155 (68.89)	52 (33.55)	67 (43.23)	36 (23.22)
Pure DCIS	25 (16.13)	7 (28)	14 (56)	4 (16)
DCIS within IDC	30 (19.35)	9 (30)	10 (33.33)	11 (36.67)
IDC alone	100 (64.52)	36 (36)	43 (43)	21 (21)

DCIS: Ductal carcinoma *in situ*; IDC: invasive ductal carcinoma.

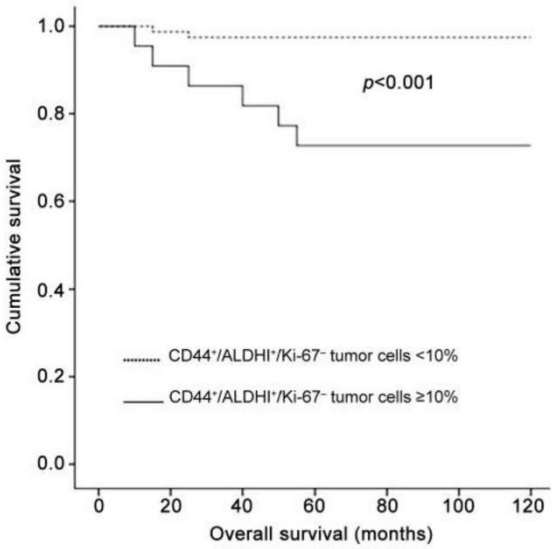

Figure 4. *Univariate analysis of overall survival according to hyaluronan receptor (CD44)+ aldehyde dehydrogenase-1 (ALDH1)+ Ki-67− tumor cell prevalence.*

were preserved instead of using cut-off values, except for prevalence description and overall survival representation. In control samples, obtained from reduction mammoplasties, percentages of CD44+ALDH1+Ki-67− cells ranged from 0% to 2%, with only two cases having more than 1%. Regarding the non-malignant tissues, these cells ranged from 0% to 11% with only one case having ≥10%. As for the malignant cases, CD44+ALDH1+Ki-67− tumor cells in pure DCIS (range: 0 to 50.7%), DCIS within IDC (range: 0 to 55.5%) and in IDC alone (range: 0 to 51%) were negative or presented scattered cells in 28%, 30% and in 36% of cases, respectively. In pure DCIS, 16% of cases had ≥10%; 36.67% of cases displayed ≥10% in DCIS within IDC and in IDC alone only 21% displayed ≥10% (Table I).

As described in Table II, although non-malignant tissues presented a higher mean percentage of CD44+ALDH1+Ki-67− cells when compared to normal breast tissues, this difference was not statistically significant. Conversely, the mean percentage of CD44+ALDH1+Ki-67− tumor cells in malignant specimens was significantly higher than that of normal (p=0.025) and non-malignant (p<0.001) breast tissues. Regarding only the malignant cases, DCIS within IDC cases presented a significantly higher percentage of CD44+ALDH1+Ki-67− tumor cells than IDC alone (p=0.016).

Clinical implications of CD44+ALDH1+Ki-67− tumor cells. Considering the percentage of CD44+ALDH1+Ki-67− tumor cells in DCIS lesions, a significant positive correlation was found between the mean percentage of these cells and nuclear grade (p=0.005, Table III).

From the 100 IDC cases analyzed, a significant negative correlation between the mean percentage of CD44+ALDH1+Ki-67− tumor cells and Elston-Ellis grading (p=0.027) and tumor size (p=0.007) were obtained. The follow-up analysis demonstrated that two patients developed local recurrences and 11 developed distant metastases in bone (n=3), liver (n=4) and lung (n=4); as for survival, eight patients died due to their disease. The mean percentage of CD44+ALDH1+Ki-67− tumor cells was strikingly higher in patients who developed distant metastases (p=0.001) and in those who died due to breast cancer (p=0.001, Table IV). Besides this, all patients presented CD44+ALDH1+Ki-67− tumor cells and those presenting ≥10% of these cells had a poorer overall survival (Figure 4).

We also explored the associations between the percentages of CD44+ALDH1+Ki-67− tumor cells and the hormone receptor status defined for each patient. The mean percentage of these cells was observed to be significantly higher in those with HER2− (p=0.002) breast tumors (Table V).

Discussion

One of the principles of the CSC model is that tumor growth, disease progression and the generation of heterogeneity in cancer is driven by a small population of tumorigenic cells within a tumor (8). Heterogeneity and plasticity in the phenotype of CSCs have been described in relation to their tissue of origin; as a consequence, few definitive markers have been isolated for CSCs from human solid tumors (15). In breast cancer, a plethora of studies involving the BCSC markers CD44, CD24 and ALDH1

Table II. *Mean percentage of hyaluronan receptor (CD44)$^{+}$ aldehyde dehydrogenase-1 (ALDH1)$^{+}$ Ki-67^{-} breast cells in breast tissue samples.*

	CD44^{+}ALDH1^{+}Ki-67^{-} breast cells		
Characteristic	n (%)	Mean±SEM (%)	*p*-Value
Type of lesion			
Normal	16 (7.11)	0.43±0.12	Normal *vs*. non-malignant: 0.975
Non-Malignant	54 (24)	1.06±0.25	Normal vs. malignant: 0.025
Malignant	155 (68.89)	7.58±1.00	Malignant *vs*. non-malignant: <0.001
Pure DCIS	25 (16.13)	7.08±2.51	Pure DCIS *vs*. DCIS within IDC: 0.162
DCIS within IDC	30 (19.35)	13.13±2.88	Pure DCIS *vs*. IDC alone: 0.924
IDC alone	100 (64.52)	6.05±1.08	IDC alone *vs*. DCIS within IDC: 0.016

DCIS: Ductal carcinoma *in situ*; IDC: invasive ductal carcinoma; SEM standard error of the mean.

Table III. *Mean percentage of hyaluronan receptor (CD44)$^{+}$ aldehyde dehydrogenase-1 (ALDH1)$^{+}$ Ki-67^{-} breast cells according to ductal carcinoma in situ nuclear grade.*

	CD44^{+}ALDH1^{+}Ki-67^{-} breast cells			
Characteristic	n (%)	Mean±SEM (%)	Correlation coefficient	*p*-Value
Nuclear grade				
Low	10 (18.18)	0.36±0.30		
Intermediate	19 (34.55)	9.10±3.17	0.374	0.005
High	26 (47.27)	15.16±3.15		

SEM: Standard error of the mean.

have demonstrated, essentially through immunological and genetic approaches, the implications of this type of cell in tumorigenesis and their effect after therapy failure. In fact, attempts to identify and characterize CD44^{+}CD24^{-low} or ALDH1^{+high} cells revealed enhanced ability to form tumors *in vitro* (16) and enhanced metastatic capacity (5). From an immunological point of view, however, ALDH1 seems to be a more effective predictive marker than the CD44^{+}CD24^{-low} phenotype. In fact, few studies have characterized CD44^{+}ALDH1^{+high} breast cancer cells and even a lack of correlation among these three major BCSC markers was recently reported (14, 17).

In this study, beyond the characterization of CD44 and ALDH1 in a cohort comprising of a series of normal, non-malignant and malignant breast tissues, we also aimed to identify CD44^{+}ALDH1^{+} breast cells in their quiescent state (Ki-67^{-}) to analyze not only their correlation with the clinicopathological information but also to localize them among the different breast lesions assessed. After our previous results, we were able to identify CD44^{+}ALDH1^{+}Ki-67^{-} breast cells through a triple-staining method along with the corresponding digital images which allowed us to better control color-staining differences and the precise proportion of these cells when comparing with the entire tumor tissue.

Normal breast tissues with no previous history of cancer presented very low percentages of CD44^{+}ALDH1^{+}Ki-67^{-} breast cells. In tissues with benign breast diseases, this percentage, although not significantly higher than in normal tissues, was greater than 1% and even in a few cases, these cells were present in more than 5% of the entire lesion. CD44 and ALDH1 have already been described in normal and non-malignant breast tissues and even a higher expression of ALDH1 in women who developed breast cancer from proliferative benign lesions was demonstrated (18). Considering the sequential progression of breast tumors, we sought to determine if there was any difference in the percentage of CD44^{+}ALDH1^{+}Ki-67^{-} cells between pure DCIS and DCIS within IDC. As expected, the mean percentage of these cells in DCIS within IDC was higher although not statistically significantly so. Nonetheless, when grouping all the DCIS cases, a significant positive correlation between the mean percentage of CD44^{+}ALDH1^{+}Ki-67^{-} cells and nuclear grade was observed. This enrichment of CD44^{+}ALDH1^{+}Ki-67^{-} tumor cells supports the CSC model regarding their tumor-initiating capacity but more interestingly, highlights that these cells exist in a quiescent state, raising some important questions about the real role of dormancy. Recent discoveries

Table IV. *Mean percentage of hyaluronan receptor (CD44)+ aldehyde dehydrogenase-1 (ALDH1)+ Ki-67− breast cells according to invasive ductal carcinoma clinicopathological markers of clinical progression and outcome.*

	CD44+ALDH1+Ki-67− breast cells			
Characteristic	n (%)	Mean±SEM (%)	Correlation coefficient	*p*-Value
Grade				
G1	15 (15)	9.40±3.50		
G2	53 (53)	7.12±1.64	−0.221	0.027[a]
G3	32 (32)	2.70±0.97		
Tumor size				
T1	51 (71)	7.47±1.57		
T2	41 (51)	5.04±1.72	−0.269	0.007[a]
T3-4	8 (8)	2.19±1.23		
Nodal status				
N0	55 (55)	5.36±1.16		
N1	30 (30)	7.13±2.43	−0.067	0.508[a]
N2-3	15 (15)	6.39±3.31		
Metastasis				
M0	89 (89)	4,64±0.91	-	0.001[b]
M1	11 (11)	17.48±5.34		
Status				
Alive	92 (92)	4,88±0.94	-	0.001[b]
Died	8 (8)	19.55±6.68		

SEM: Standard error of the mean. [a]Spearman's non-parametric correlation; [b]Mann–Whitney *U*-test.

Table V. *Mean percentage of hyaluronan receptor (CD44)+ aldehyde dehydrogenase-1 (ALDH1)+ Ki-67− breast cells according to the patient's hormone receptor and human epidermal growth factor receptor (HER)-2 status.*

	CD44+ALDH1+Ki-67− breast cells				
	Pure DCIS n (%)	DCIS within IDC n (%)	IDC alone n (%)	Mean±SEM (%)	*p*-Value
ER status					
ER+	20 (80)	27 (90)	80 (80)	7.85±1.17	0.114
ER−	5 (20)	3 (10)	20 (20)	6.39±1.56	
PR status					
PR+	14 (56)	22 (73.33)	70 (70)	6.81±1.51	0.319
PR−	11 (44)	8 (26.67)	30 (30)	7.94±1.28	
HER2 status					
HER2+	8 (32)	8 (26.67)	24 (24)	4.03±1.15	0.002
HER2−	17 (68)	22 (73.33)	76 (76)	8.82±1.27	

ER: Estrogen receptor; PR: progesterone receptor; SEM: standard error of the mean.

suggested that the quiescent state is not just a passive state but, instead, is actively regulated by different intrinsic mechanisms. As CSCs may adopt the quiescent state to resist metabolic stress and to preserve genomic integrity, quiescent CSCs may be prompted for activation by specific energetically favorable mechanisms that are compatible with the low metabolic state of quiescence. Such modulation of CSCs can consequently generate rapid and global responses needed for activation (19).

Another interesting point is the significant decrease of CD44+ALDH1+Ki-67− tumor cells from DCIS within IDC to IDC alone and the negative correlation obtained with histological grade and tumor size. Such difference can be partly explained by other theoretical properties of CSCs: if these cells have the ability to self-renew and differentiate generating non-tumorigenic cancer cells that form a tumor mass, only a few cells would be required for invasion. Regarding the important roles of CD44 (self-renewal, niche preparation and resistance to apoptosis) and ALDH1 (self-renewal, stem cell proliferation control, protection against oxidative insults) in the stemness maintenance of CSCs, a decrease in the expression of these markers would then be expected. Again, only a few studies have analyzed ALDH1 and CD44 expression in DCIS. A higher expression of CD44 in DCIS when compared to IDCs was reported (20) and ALDH1 expression in DCIS associated with tumor-initiating properties was also demonstrated (21). Moreover, the increment of non-tumorigenic cancer cells obtained by symmetric divisions of CSCs can also explain the lower number of CD44+ALDH1+Ki-67− tumor cells seen in larger tumors.

Apart from the number of these cells, isolated pools of CD44+ALDH1+Ki-67− tumor cells were also detected in DCIS and IDCs. Even with the limitations of immunological techniques and considering the features described in literature for the so-called CSC niches, such isolated cells in malignant cases can be highly tumorigenic. If these cells happen to resist and prevail due to specific microenvironmental conditions, this advantage can allow them to gain malignant potential and thus to progress along the tumorigenic process.

Furthermore, the higher mean percentage of CD44+ALDH1+Ki-67− tumor cells among women who developed distant metastases and in those who died due to their disease is also noteworthy. A large body of evidence points to the fact that CSCs are particularly resistant to radiotherapy and chemotherapy, which can be partly explained by the clonal evolution model and tumor heterogeneity. Tumor progression has been related to 'Darwinian' evolution. The expansion of an established tumor can be explained by the generation of new clones as a result of mutations, genetic instability or epigenetic alterations. The prevalence of new CSCs and their clones will be determined by different selective pressures (nutritional or immune status, oxygenation and therapy) that modify the tumor microenvironment. If selected, these cells

can be responsible for tumor relapse or metastasis (22). Another factor that can compromise a good therapeutic response is the assumption of a fluid existence of 'stemness', with cells having the ability to both acquire and lose stemness. The acquisition of a stem cell phenotype has been demonstrated through the epithelial–mesenchymal transition induced either by paracrine signaling from cancer-associated fibroblasts or neighboring tumor cells (8).

In agreement with our results, the $CD44^{+}ALDH1^{+}Ki\text{-}67^{-}$ phenotype can potentially contribute to both chemotherapy and radiation resistance. These BCSCs markers were already described to be determinant for treatment resistance, recurrences and metastasis development through expression of high levels of therapy-resistance proteins (10, 23). ALDH1 activity has been shown to render cancer cells exquisitely resistant to some chemotherapy agents mainly due to its well-characterized role in differentiation through the retinoic acid pathway (24, 25). Moreover, $ALDH1^{+}$ tumor cells were shown to be more likely negative for Ki-67 (26), which is in part in accordance with our present study. In fact, quiescence may also have a determinant role in tumor progression and relapse. Several chemotherapeutic agents as well as radiotherapy work by inducing DNA damage. Thus, cells that have the ability to repair DNA damage are more prone to survive chemotherapy. Regarding the properties of quiescence, quiescent cells may have the potential and time to repair the damage inflected on them. Although quiescence is not an essential characteristic that defines stem cells, in BCSCs, there is increased expression of DNA-repair genes, indicating that high DNA-repair activity may aid in making CSCs resistant to tumor therapy (12).

Regarding the distribution of $CD44^{+}ALDH1^{+}Ki\text{-}67^{-}$ tumor cells according to hormone receptor status, a significant association with HER2 negativity was observed. Some controversial studies regarding the associations between ALDH1 and CD44 expression and hormone receptor status raise the question about the prognostic ability of these markers (27). Moreover, the studies carried out so far regarding the characterization of $ALDH1^{high}CD44^{+}$ tumor cells were *in vitro* studies; hence the correlation of these cells with poor prognosis remains to be clarified.

Despite our solid results concerning the *in situ* identification of $ALDH1^{+}CD44^{+}Ki\text{-}67^{-}$ tumor cells, some limitations of this work need to be addressed: the lack of validation precludes us from concluding that we have truly identified stem cells and BCSCs, in any case we referred to our cells of interest as BCSCs. In addition, we were not able to infer anything about breast tumor progression and $CD44^{+}ALDH1^{+}Ki\text{-}67^{-}$ cells in the same patient as this was a retrospective study. Finally, the limited number of patients enrolled in our study may have influenced the results, especially for patients who developed distant metastases (n=11) and for those who died (n=8).

In conclusion, $CD44^{+}ALDH1^{+}Ki\text{-}67^{-}$ tumor cells may have a greater tumorigenic effect in breast cancer than $CD44^{+}CD24^{-low}$ tumor cells. Due to its role, ALDH1 can determine the behavior of CSCs, their ability to resist chemotherapeutic agents and their dissemination to other parts of the body, which can be aided by the role of CD44. Additionally, quiescence seems to have a more preponderant role than previously expected, which may be crucial for tumor progression, resistance to chemotherapeutic agents and the metastatic spread of BCSCs. Despite improvements in knowledge of the adverse effects of ALDH1 and CD44 in breast cancer treatment, additional studies on the tumorigenic and metastatic ability of $CD44^{+}ALDH1^{+high}$ tumor cells combined with their quiescence status are still needed.

Acknowledgements

This study was funding by the Foundation for Science and Technology (Portugal). Scholarships references: SFRHBD743 072010 (ACP) and SFRHBD201178184 (OM). The Authors would like to thank Dr. Jorge Reis-Filho (Memorial Sloan Kettering Cancer Center, Department of Pathology, New York, NY, USA) for scientific assistance concerning the triple-immunohistochemistry optimization and validation and to Filipa Moreno (Pathology Department of Santo António Hospital, Porto, Portugal) for clinical assistance during breast specimen collection.

Conflicts of Interest

The Authors state that no conflicts of interest exist in regard to this study.

References

1 Clevers H: The cancer stem cell: premises, promises and challenges. Nat Med *17*: 313-319, 2011.
2 Visvader JE and Lindeman GJ: Cancer stem cells: current status and evolving complexities. Cell stem Cell *10*: 717-728, 2012.
3 Donnenberg AD, Hicks JB, Wigler M and Donnenberg VS: The cancer stem cell: cell type or cell state? Cytometry Part A *1*: 5-7, 2013.
4 Park TS, Donnenberg VS, Donnenberg AD, Zambidis ET and Zimmerlin L: Dynamic interactions between cancer stem cells and their stromal partners. Current Pathol Rep *1*: 41-52, 2014.
5 Sheridan C, Kishimoto H, Fuchs RK, Mehrotra S, Nakshatri PB, Turner CH, Goulet R Jr., Badve S and Nakshatri H: CD44+CD24− breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. Breast Cancer Res 5: R59, 2006.
6 Phillips TM, William HM and Frank P: The response of CD24−lowCD44+ breast cancer–initiating cells to radiation. J Natl Cancer Inst *24*: 1777-1785, 2006.
7 Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J, Nikolskaya T, Serebryiskaya T, Beroukhim R, Hu M, Halushka MK, Sukumar S, Parker LM, Anderson KS, Harris LN, Garber JE, Richardson AL, Schnitt Nikoslky Y, Gelman RS and Polyak K: Molecular definition of breast tumor heterogeneity. Cancer Cell *3*: 259-273, 2010.

8 Martelotto LG, Ng CK, Piscuoglio S, Weigelt B and Reis-Filho JS: Breast cancer intra-tumor heterogeneity. Breast Cancer Res *3*: 210 (2014).

9 Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, Jacquemier J, Viens P, Kleer CG, Liu S, Schott A, Hayes D, Birnbaum D, Wicha MS and Dontu G: ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell stem Cell *1*: 555-567, 2007.

10 Croker AK and Alison LA: Inhibition of aldehyde dehydrogenase (ALDH) activity reduces chemotherapy and radiation resistance of stem-like ALDH^{hi}CD44+ human breast cancer cells. Breast Canc Res Treat *1*: 75-87, 2012.

11 Lee HE, Kim JH, Kim YJ, Choi SY, Kim SW, Kang E, Chung EY, Kim IA, Kim EJ, Choi Y, Ryu HS and Park SY: An increase in cancer stem cell population after primary systemic therapy is a poor prognostic factor in breast cancer. Br J Canc *11*: 1730-1738, 2011.

12 Colak S and Medema JP: Cancer stem cells–important players in tumor therapy resistance. FEBS Journal *21*: 4779-4791, 2014.

13 Roesch A, Fukanaga-Kalabis M, Schmidt EC, Zabierowski SE, Brafford PE, Vultur A, Basu D, Gimotty P, Vogt T and Herlyn M: A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. Cell *4*: 583-594, 2010.

14 Da Cruz Paula A, Marques A, Rosa AM, Faria MDF, Rema A and Lopes C: Co-expression of stem cell markers ALDH1 and CD44 in non-malignant and neoplastic lesions of the breast. Anticancer Res *3*: 1427-1434, 2014.

15 Maccalli C and De Maria R: Cancer stem cells: perspectives for therapeutic targeting. Cancer Immunol. Immunother *1*: 91-97, 2015.

16 Ponti D, Costa A, Zaffaroni N, Pratesi G, Pretrangolini G, Coradini D, Pilotti S, Pierotti MA and Daidone MG: Isolation and *in vitro* propagation of tumorigenic breast cancer cells with stem progenitor cell properties. Cancer Res *13*: 5506-5511, 2005.

17 Liu Y, Nenutil R, Appleyard MV, Murray K, Boylan M, Thompson AM and Coates PJ: Lack of correlation of stem cell markers in breast cancer stem cells. Br J Canc *8*: 2063-2071, 2014.

18 Kunju LP, Cookingham C, Toy KA, Chen W, Sabel MS and Kleer CG: EZH2 and ALDH-1 mark breast epithelium at risk for breast cancer development. Mod. Pathol *6*: 786-793, 2011.

19 Cheung TH and Thomas AR: Molecular regulation of stem cell quiescence. Nat Rev Mol Cell Biol *6*: 329-340, 2013.

20 Park SY, Lee HE, Li H, Shipitsin M, Gelman R and Polyak K: Heterogeneity for stem cell–related markers according to tumor subtype and histologic stage in breast cancer. Clin. Cancer Res *3*: 876-887, 2010.

21 Knudsen ES and Agnes W: EZH2 and ALDH1 expression in ductal carcinoma *in situ*: complex association with recurrence and progression to invasive breast cancer. Cell Cycle *13*: 2042-2050, 2013.

22 Alison MR, Lin WR, Lim SM and Nicholson LJ: Cancer stem cells: in the line of fire. Canc Treat Rev *6*: 589-598, 2012.

23 Croker AK, Goodale D, Chu J, Postenka C, Hedley BD, Hess, and Allan AL: High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability. J Cell Mol Med *8B*: 2236-2252, 2009.

24 Ginestier C, Wicinski J, Cervera N, Monville F, Finetti P, Bertucci F, Wicha MS, Birnbaum D and Charafe-Jauffret E: Retinoid signaling regulates breast cancer stem cell differentiation. Cell Cycle *8*: 3297-3302, 2009.

25 Tanei T, Morimoto K, Shimazu K, Kim SJ, Tanji Y, Taguchi T, Tamaki Y and Noguchi S: Association of breast cancer stem cells identified by aldehyde dehydrogenase 1 expression with resistance to sequential paclitaxel and epirubicin-based chemotherapy for breast cancers. Clin Cancer Res *12*: 4234-4241, 2009.

26 Morimoto K, Kim SJ, Tanei T, Shimazu K, Tanji Y, Taguchi T, Terada N, and Noguchi S: Stem cell marker aldehyde dehydrogenase 1-positive breast cancers are characterized by negative estrogen receptor, positive human epidermal growth factor receptor type 2, and high Ki67 expression. Cancer Sci *6*: 1062-1068, 2009.

27 Zhou L, Jiang Y, Yan T, Di G, Shen Z, Shao Z and Lu J: The prognostic role of cancer stem cells in breast cancer: a meta-analysis of published literatures. Breast Cancer Res Treat *3*: 795-801, 2010.

Received July 15, 2016
Revised August 6, 2016
Accepted August 17, 2016

Chapter 5

Characterization of $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ Cells in Non-malignant and Neoplastic Lesions of the Breast through Flow Cytometry and Massive Parallel Sequencing

Characterization of CD44+/CD24-/Ck+/CD45- cells in non-malignant and neoplastic lesions of the breast through flow cytometry and massive parallel sequencing

Arnaud Da Cruz Paula MSc[1,2], Catarina Leitão PhD[3,4], Oriana Marques MSc[1,5,6,7], Ana Margarida Rosa MSc[6,7], Ana Helena Santos MSc[4,8], Alexandra Rêma BSc[1], Maria de Fátima Faria BSc[1], Ana Rocha MSc[9,10], José Luís Costa PhD[7,9,10], Margarida Lima MD, PhD[4,8*], Carlos Lopes MD, PhD[1,2*]

[1] Pathology and Molecular Immunology Department, Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto, Porto, Portugal;
[2] Department of Pathology, Portuguese Oncology Institute (IPO), Porto, Portugal;
[3] Institute of Research and Health Innovation, University of Porto
[4] IBMC – Institute of Molecular and Cell Biology, Porto, Portugal
[5] Unit for Multidisciplinary Biomedical Research (UMIB), Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto, Porto, Portugal;
[6] Basic and Clinical Research on Iron Biology, Institute of Molecular and Cell Biology (IBMC)/i3s, Porto, Portugal
[7] Institute of Research and Health Innovation, University of Porto, Porto, Portugal
[8] Laboratory of Cytometry, Department of Hematology, Santo António Hospital (HSA), Porto Hospital Centre (CHP), Porto, Portugal [*] Contributed equally to the supervision of this work
[9] IPATIMUP - Institute of Molecular Pathology and Immunology of the University of Porto, Portugal
[10] Medical Faculty of the University of Porto, Portugal

Conflict of interest The authors declare that they have no conflict of interest.

Funding This study was funding by the Foundation for Science and Technology (Portugal). Scholarships references: SFRH/BD/74307/2010 (ACP) and SFRH/BD/2011/78184 (OM).

Corresponding authors: Arnaud Da Cruz Paula and Carlos Lopes, Pathology and Molecular Immunology Department, Institute of Biomedical Sciences Abel Salazar. Jorge Viterbo Ferreira Street, n° 288, 4050-313, Porto, Portugal. Tel: +351 967693532, +351 917270209, Fax: +351 225511184, e-mail: arnaudcpaula@hotmail.com, calopes@icbas.up.pt

Keywords Breast cancer, breast cancer stem cell markers, CD44, CD24, Next-Generation Sequencing, Ion Torrent AmpliSeq Cancer Hotspot panel v2.

Summary

Breast cancer epithelial cells with the CD44$^+$/CD24$^{-/low}$ phenotype have been shown to possess tumor-initiating cells and epithelial-mesenchymal transition (EMT) capacity. Considering such important features, massive parallel sequencing can be an interesting approach to deepen the molecular characterization of these cells. We characterized CD44$^+$/CD24$^-$/Cytokeratin(Ck)$^+$/CD45$^-$ breast cells through flow cytometry in 43 biopsy and 6 mastectomy samples harboring different breast lesions. The Ion Torrent Ampliseq Cancer Hotspot panel v2 (CHPv2) was used for the identification of somatic mutations in the DNA extracted from isolated CD44$^+$/CD24$^-$/Ck$^+$/CD45$^-$ cells. E-Cadherin and vimentin immunohistochemistry was performed in the correspondent formalin-fixed, paraffin-embedded (FFPE) blocks. The percentage of CD44$^+$/CD24$^-$/Ck$^+$/CD45$^-$ cells increased significantly from non-malignant to malignant lesions and was negatively correlated with tumor size. A significant association with estrogen receptor (ER) positivity, human epidermal growth factor type 2 (HER2) negativity and vimentin positivity was observed. The non-malignant lesion harbored only a single-nucleotide polymorphism (SNP). Mutations in the tumor suppressor p53 (*TP53*), NOTCH homolog 1 (*NOTCH1*), Phosphatase and tensin homolog (*PTEN*) and v-akt murine thymoma viral oncogene homolog 1 (*AKT1*) genes were found in isolated CD44$^+$/CD24$^-$/Ck$^+$/CD45$^-$ cells from ductal carcinomas *in situ* (DCIS). Additional mutations in the colony-stimulating factor 1 receptor (*CSF1R*), ret proto-oncogene (*RET*) and *TP53* genes were also identified in invasive ductal carcinomas (IDCs). The use of massive parallel sequencing technology for this type of application revealed to be extremely effective even when using small amounts of DNA extracted from a low number of cells. Additional studies are now required using larger cohorts to design an appropriate mutational profile for this phenotype.

Introduction

Although several studies have tried to identify breast cancer stem cells (BCSC) through cell surface marker profiles, agreement on their phenotypic characterization is still lacking. With the current demonstrations of several putative BCSC markers [1,2], it became unfeasible to obtain a universal combination of markers that could specifically identify BCSCs in all breast cancers. Breast cancer

heterogeneity as reflected by numerous histological subtypes, with variable clinical presentations and different molecular signatures also contributes to major drawbacks [3].

Indeed, intra-tumor heterogeneity leads to a single tumor to contain, at any given time, tumor cell populations displaying different molecular profiles and biological properties [4]. As a consequence, different BCSC phenotypes were described, with some being associated with aggressive forms of breast cancer [5, 6]. Although the validation of the CSC model remains an ongoing task, it is important to define which BCSC phenotypes have high tumorigenic potential and ability to resist to therapeutic agents [7].

We have previously demonstrated that the co-expression of the BCSC markers CD44 and ALDH1 in DCIS could be determinant for disease progression [8]. Nonetheless, the combination of the BCSC markers CD44 and CD24 continues to be the most extensively studied. The pioneering study by Al-Hajj et al showed that as few as 100 CD44$^+$/CD24$^{-/low}$ lineage$^-$ breast cancer cells from patients with breast cancer could form tumors in mice, whereas tens of thousands of cells with alternative phenotypes failed to do so [9]. Immunohistochemically, several studies have identified the CD44$^+$/CD24$^{-/low}$ phenotype as being associated with poor prognostic features [1, 10].

Gene expression profiling of CD44$^+$/CD24$^{-/low}$ breast cancer cells revealed a gene signature of 186 genes associated with invasion and poor prognosis [11]. This signature was enriched in genes related to cell cycle, calcium-ion binding, chemotaxis, differentiation, protein transport, signal transduction and protein ubiquitination.

The increment of CD44$^+$/CD24$^{-/low}$ cells demonstrated in primary breast tumors following radiation and chemotherapy has suggested an innate resistance to standard treatments [14]. Potential mechanisms of chemotherapy and radiation resistance associated with this phenotype were shown to include the presence of lower concentration of reactive oxygen species, cell dormancy, efficient DNA repair mechanisms, overexpression of EMT markers, Wnt/β-catenin, Hedgehog and Notch signaling pathways and STAT1 and STAT3 signaling activation [12-18].

The continuous improvements of Next-Generation Sequencing (NGS) technologies allow the analysis of hundreds of genes in just one population of cells, or even in one single cell [19, 20]. With NGS, it is possible to determine

which genes associated somatic mutations can be responsible to drive tumorigenesis, while considering the high levels of heterogeneity in cancers, especially in breast cancer [21]. Currently, standardized NGS kits are available and multiple studies have shown that these kits provide reliable sequencing results in routine cancer diagnostics, like the CHPv2 [22, 23]. Such assay includes 50 genes known to be involved in the pathogenesis of many human cancers.

In this study we aimed to identify $CD44^+/CD24^-/Ck^+/CD45^-$ cells and determine their mean percentages in a cohort of frozen breast specimens consisting of non-malignant and malignant lesions and to correlate their frequency with clinicopathological markers of breast cancer progression. Moreover and through the CHPv2, we aimed to determine which mutated genes were associated with this phenotype.

Materials and Methods

Patient samples

Biopsy and mastectomy samples from 49 patients were obtained from the Radiology, Surgery and Pathology Departments of Santo António Hospital (Porto Hospital Centre, Porto, Portugal) between 2013 and 2015 and the correspondent FFPE blocks of each patient were retrieved from the archives of the same Pathology Department. This study was approved by the local Ethics Committee and all patients gave the informed consent to participate. The fragments taken at the time of surgery were routinely processed by freezing in optimum cutting temperature (OCT) media. H&E-stained sections for each sample were microscopically analyzed by one experienced pathologist (CL). From this analysis and regarding the biopsy samples, 20 were considered to have benign breast lesions, including adenosis, fibroadenomas and ductal hyperplasias, 2 contained high-grade DCIS lesions and 21 harbored IDC lesions. Regarding the 6 mastectomy samples that were also used for molecular analysis, one contained a fibroadenoma, 3 harbored DCIS lesions (one pure DCIS, and two DCIS within IDC) and 2 were confirmed to have IDC lesions (one triple-negative tumor and one luminal tumor). None of these patients had a family history of breast cancer. Pathology reports of each patient were retrieved to (1) confirm the existence of

each lesion diagnosed in each sample, (2) to confirm the absence of malignancy in samples that contained only benign lesions and (3) to confirm the presence or absence of invasion in the observed DCIS samples. Moreover, the following histopathologic variables for the malignant cases were available in the interim records: Elston-Ellis histologic grade, tumor size, lymph-node status, local recurrences, distant metastasis, ER, PR and HER2 status and Ki-67 proliferation index. The definition of hormone receptor status was assessed by immunohistochemistry, as routinely done in the Pathology Service. HER2 ambiguous results were confirmed by Fluorescence *in situ* Hybridization. Low and high Ki-67 expression was defined according to the cut-off of ≤14% and >14%, respectively [24].

Flow cytometry and cell sorting

Frozen breast samples were thawed on ice and fat tissue was removed as far as possible. The remaining tissue was crudely minced with scalpels and disaggregated with 1 ml of Phosphate-buffered saline solution (PBS, pH 7.4) with 0.2% of Bovine serum albumin (BSA) and 0.1% of sodium azide (PBS-0.2%BSA-0.1%Azide). Cell suspensions were transferred to 5 ml, 75 x 12 mm, polypropylene tubes, centrifuged for 5 minutes at 400 x g and once again washed with PBS-0.2%BSA-0.1%Azide. The remaining pellets were suspended in the same buffer, and cells were counted. Cell fixation and permeabilization was made using the Fix & Perm® reagent kit (Invitrogen Corporation, Camarillo, CA) according to the manufacturer's instructions. Briefly, 1 x 10^6 cells were incubated for 15 minutes at room temperature in the dark, in the presence of saturating amounts of mouse anti-human monoclonal antibodies (mAb) specific for CD44 (clone G44-26, BD Biosciences, Franklin Lakes, NJ, USA), CD24 (clone ML5, BD Biosciences, Franklin Lakes, NJ, USA), and CD45 (clone 2D1, BD Biosciences, California, USA), conjugated with allophycocyanin (APC), phycoerythrin (PE) and Peridinin chlorophyll (PerCP), respectively. Fluorescence-minus-one (FMO) controls lacking either the anti-CD44-APC or anti-CD24-PE antibodies were used to distinguish positive and negative cell populations. After staining, all samples were washed twice with PBS-0.2%BSA-0.1%Azide, centrifuged (5 minutes, 400 x g), and incubated with 100 µl of Fix & Perm® reagent A (fixative medium) for 15

minutes at room temperature. After another two washes in the same conditions, the remaining pellets were incubated for 15 minutes at room temperature in the dark with 100 µl of Fix & Perm® reagent B (permeabilization medium), in the presence of saturating amounts of mouse anti-human cytokeratin mAb (clone J1B3, Beckman Coulter, Marseille, France) conjugated with fluorescein isothiocyanate (FITC). An appropriate IgG1-FITC isotype control was run in parallel (clone 679.1Mc7, Beckman Coulter, Marseille, France). The labeled cells were washed twice in the same conditions, and suspended in 0.5 ml of PBS before analysis. All measurements were performed on a BD FACSCanto™ II Flow Cytometer. For each sample run, a minimum of 10 000 events were recorded. Flow cytometric data was analyzed using FlowJo X 10.0.7r2 (TreeStar, Ashland, OR, USA). Briefly, cells were gated on a forward scatter (FSC) versus side scatter (SSC) dot plot to exclude dead cells and debris. After exclusion of doublets, epithelial mammary cells were identified and hematopoietic cells were eliminated from the analysis by gating cells on cytokeratin-FITC versus CD45-PercP dot plots (Fig. 1). The percentages of $CD44^{+/}CD24^-/Ck^+/CD45^-$ breast cells and the median fluorescence intensities of CD44, CD24 and Ck expression were retrieved for statistical analysis.

For sorting and considering the purpose of this study, $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells from six mastectomy samples were purified. Cell sorting was performed on a BD Biosciences FACSAria operating at Pressure of 70 psi using a 70 µm nozzle, using the gating strategy described above. Considering the low numbers of $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells sorted, cell purity was checked in cell populations defined as non-BCSCs (not $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells). A purity of more than 95% was obtained. Cell populations of interest were sorted into individual Eppendorf tubes and collected for DNA extraction (Table 5).

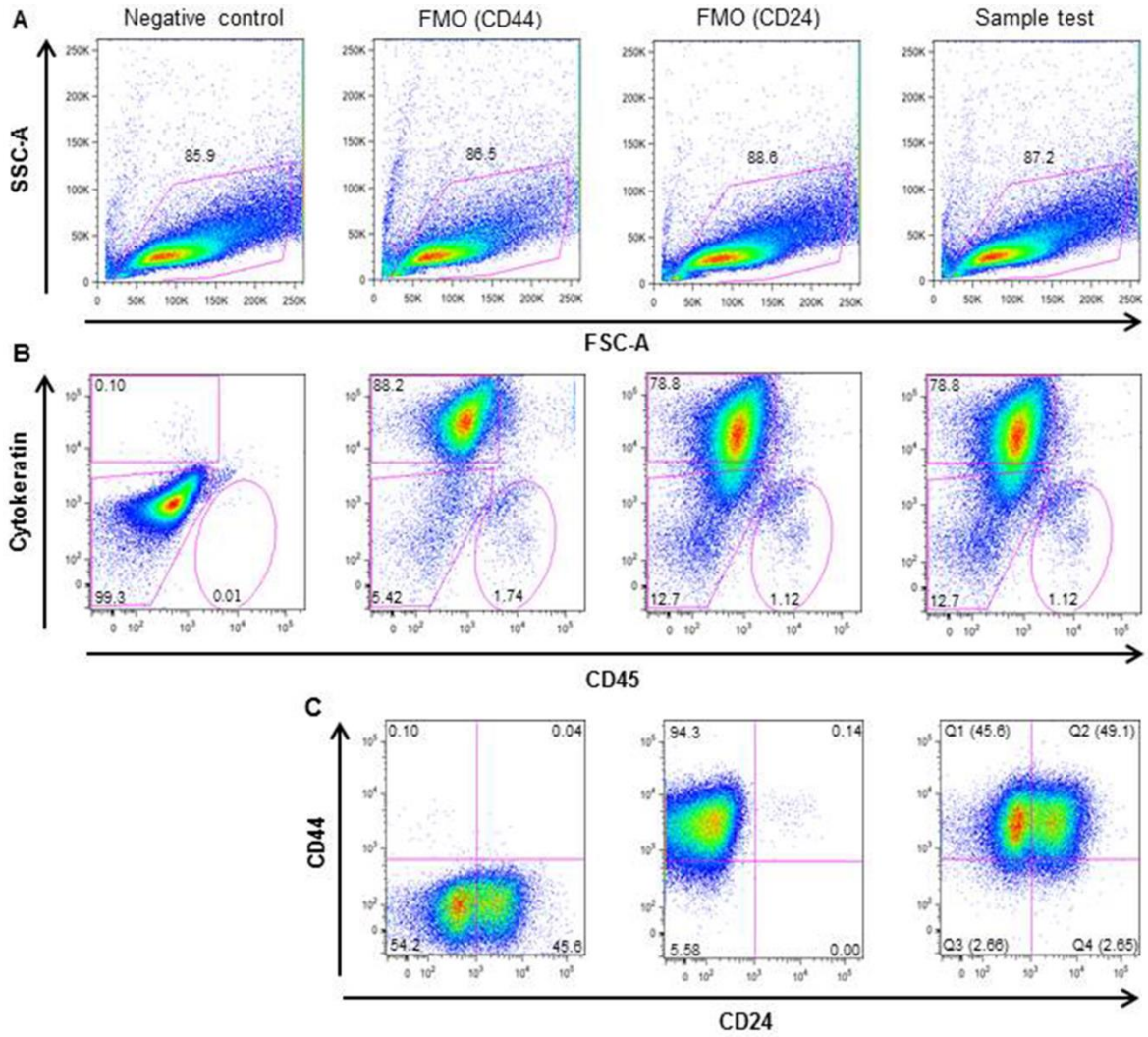

Fig. 1. Gating strategy for the identification of CD44+/CD24-/Ck+/CD45- breast cells: (A), all cells were gated on a forward scatter (FSC) versus side scatter (SSC) dot plot to exclude dead cells and debris (B), epithelial mammary cells were identified and hematopoietic cells were eliminated from the analysis by gating cells on cytokeratin-FITC versus CD45-PercP dot plots (C) and CD44+/CD24-/Ck+/CD45- breast cells were identified by gating cells on CD44-APC versus CD24-PE. Negative control: cells labeled only with the IgG1-FITC isotype control. FMO CD44: Fluorescence-minus-one control lacking the anti-CD44-APC antibody. FMO CD24: Fluorescence-minus-one control lacking the anti-CD24-PE antibody. Sample test: cells labeled with anti-C45-PerCp, anti-Cytokeratin-FITC, anti-CD44-APC and anti-CD24-PE antibodies. Q1: CD44+/CD24-/Ck+/CD45- cells; Q2: CD44+/CD24+/Ck+/CD45- cells; Q3: CD44-/CD24-/Ck+/CD45- cells; Q4: CD44-/CD24+/Ck+/CD45- cells.

DNA extraction

Genomic DNA was extracted using the Arcturus PicoPure DNA extraction Kit (Life Technologies, California, USA), according to the manufacturer's recommendations. Due to the reduced number of cells, only 50 µl of the Extraction

Solution was used for each sample. Extracted DNA samples were quantified using Nanodrop spectrophotometer (ThermoScientific, Wilmington, DE, USA) and stored at -20ºC until library preparation for sequencing.

AmpliSeq™ Cancer Hotspot panel v2

Libraries were generated using the CHPv2 (Thermo Fisher Scientific, California, USA). This panel consists of 207 amplicons covering over 20 000 bases of 50 genes with known cancer associations. DNA concentration of each sample was re-measured using Qubit 3.0 fluorometer (Thermo Fisher Scientific). All samples were processed in a vacuum concentrator (Thermo Fisher Scientific) to increase DNA concentration at room temperature and default vacuum pressure, with constant monitoring. Starting DNA (1–10 ng) from each sample was used to prepare barcoded libraries using IonXpress barcoded adapters (Thermo Fisher Scientific). Recommended additional cycling conditions were used for libraries with low molarities (<50 pM). The Ion Ampliseq library kit (Thermo Fisher Scientific) was used as per the manufacturer's recommendations. Libraries were further processed on Ion Chef System and the resulting Ion 540 chip was sequenced on Ion S5 XL System (Thermo Fisher Scientific).

Data analysis

Data from the S5 XL run was processed using the Ion Torrent platform specific pipeline software Torrent Suite v5.0.2. Reads generated were aligned using the Torrent Mapping Alignment Program (TMAP) - to the human reference genome build 19 (hg19). After alignment, coverage statistics were generated using Coverage Analysis plugin v5.0 (Life Technologies). Ion Reporter v5.0 was used to call somatic SNPs, multi-nucleotide polymorphisms (MNPs), single-nucleotide variants (SNVs) insertions, deletions, (INDELs) and block substitutions. FASTQ and/or BAM files were generated using the Torrent Suit plugin FileExporter v5.0 and use for Integrative Genomic Viewer (IGV).

Immunohistochemistry

E-Cadherin and vimentin IHC was performed on the FFPE blocks correspondent to malignant lesions (n=28). The staining method was carried out as previously described [8] with the monoclonal antibodies E-Cadherin (clone 4A2C7, Dilution: 1/50, Invitrogen, UK) and vimentin (clone V9, Dilution: 1/500, Dako, Denmark). The reaction obtained in all samples was microscopically analyzed (Olympus U-SPO3, Olympus Corporation, Japan) by one pathologist (CL). For E-Cadherin, the semi-quantitative evaluation method was applied as previously described [8] with only 2 categories considered (negative/low expression and high expression). For vimentin only presence or absence of immunoexpression was considered.

Statistical analysis

Statistical analyses of the flow cytometry and immunohistochemistry results were performed using IBM SPSS Statistics version 20.0 software. Sample distributions were compared using Kruskal-Wallis or Mann-Whitney U tests. Pearson's Chi-Square was used to evaluate the differences between categorical variables. The Spearman's rank correlation coefficient was used to evaluate the relationship between variables. The Kaplan-Meier method was used for univariate survival analysis. Statistical significance was considered for $P < 0.05$.

Results

Median Fluorescence Intensity of CD44, CD24 and Ck in breast samples

CD44, CD24 and Ck expression was observed in both non-malignant and malignant samples. For each sample, the median fluorescence intensity for CD44, CD24 and Ck within mammary epithelial cells was retrieved. Regarding the cell-surface marker CD44, the median fluorescence intensity was higher in malignant samples when compared with non-malignant ones, but was not statistically significant. For the fluorescence intensity of CD24 and Ck, a significant lower median fluorescence intensity was observed in malignant samples when

compared with non-malignant ones (*P* = 0.032 and *P* = 0.030, respectively, Table 1).

Table 1. Mean fluorescence intensities of CD44, CD24 and Cytokeratin expression in breast tissue samples

		CD44 intensity		CD24 intensity		Ck intensity	
Characteristics	n (%)	Mean ± SEM (%)	*P*	Mean ± SEM (%)	*P*	Mean ± SEM (%)	*P*
Type of lesion							
Non-malignant	21 (43)	2302 ± 362	0.337[a]	11369 ± 2890	0.032[a]	65149 ± 11300	0.030[a]
Malignant	28 (57)	3061 ± 619		5480 ± 620		39077 ± 5558	
DCIS	5 (18)	2337 ± 887	0.595[b]	5903 ± 2503	0.809[b]	42913 ± 19191	0.822[b]
IDC	23 (82)	3219 ± 732		5388 ± 832		26920 ± 5613	

Abbreviations: DCIS, ductal carcinoma *in situ;* IDC, invasive ductal carcinoma; SEM standard error of the mean
Values were approximated to the nearest full unit
[a] Non-malignant vs Malignant
[b] DCIS vs IDC

CD44+/CD24-/Ck+/CD45- cell population in breast samples and association with clinicopathological variables

We were able to identify CD44+/CD24-/Ck+/CD45- breast cells in non-malignant (Fig. 2A), DCIS (Fig. 2B) and IDC (Fig. 2C) lesions. Due to the low percentages of CD44+/CD24-/Ck+/CD45- cells identified in several samples, raw percentages of these cells for statistical analysis were preserved instead of using cut-off values, except for prevalence description. Actually, for non-malignant samples, percentages of CD44+/CD24-/Ck+/CD45- cells ranged from 0% to 7% with 12 cases having more than 1%. Regarding DCIS, CD44+/CD24-/Ck+/CD45- cells ranged from 2.6% to 14.1% with 3 cases having ≥10%. As for IDC lesions, these cells ranged from 0% to 53.8%, where 30.4% of the samples were negative or presented scattered cells and the same percentage of cases displayed ≥10% of these cells (Table 2).

As also presented in table 2, malignant samples had a higher mean percentage of CD44+/CD24-/Ck+/CD45- cells when compared to non-malignant samples (*P* = 0.007). Within the malignant cases, no significant differences were observed between DCIS and IDC lesions.

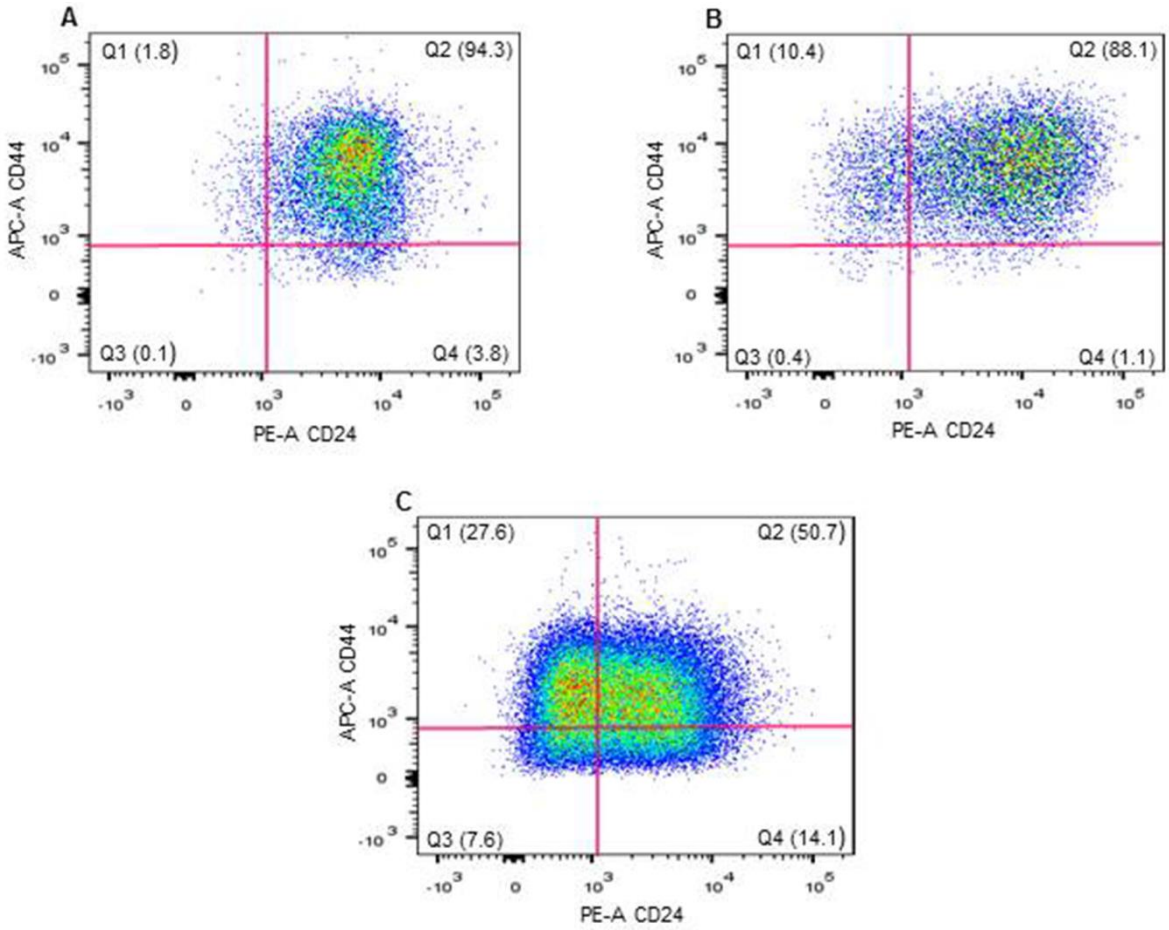

Fig. 2. Identification of $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells. (A) fibroadenoma; (B) ductal carcinoma *in situ* (DCIS) and (C) Invasive ductal carcinoma (IDC). Q1: $CD44^+/CD24^-/Ck^+/CD45^-$ cells; Q2: $CD44^+/CD24^+/Ck^+/CD45^-$ cells; Q3: $CD44^-/CD24^-/Ck^+/CD45^-$ cells; Q4: $CD44^-/CD24^+/Ck^+/CD45^-$ cells.

Table 2. Prevalence and mean percentages of $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells in breast tissue samples

	$CD44^+/CD24^-/Ck^+/CD45^-$ breast cells					
Characteristics	n (%)	<1%, n (%)	<10%, n (%)	≥10%, n (%)	Mean ± SEM (%)	*P*
Type of lesion						
Non-malignant	21 (42.9)	9 (42.9)	12 (57.1)	0 (0)	2.0 ± 0.4	0.007[a]
Malignant	28 (57.1)	7 (25)	11 (39.3)	10 (35.7)	11.7 ± 2.9	
DCIS	5 (17.9)	0 (0)	2 (40)	3 (60)	8.9 ± 2.0	0.832[b]
IDC	23 (82.1)	7 (30.4)	9 (39.1)	7 (30.4)	12.2 ± 3.5	

Abbreviations: DCIS, ductal carcinoma *in situ;* IDC, invasive ductal carcinoma; SEM standard error of the mean
Values were rounded to one decimal place
[a] Non-malignant vs Malignant
[b] DCIS vs IDC

Regarding the IDC cases, although grade 1 tumors presented a tendency to a higher percentage of $CD44^+/CD24^-/Ck^+/CD45^-$ cells, no significant correlation

was observed with tumor grade. However, a significant negative correlation was obtained between the mean percentage of these cells and tumor size (*P* = 0.034, Table 3).

Table 3. Mean percentages of $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells according to IDC clinicopathological markers of clinical progression.

	$CD44^+/CD24^-/Ck^+/CD45^-$ breast cells			
Characteristics	n (%)	Mean ± SEM (%)	Correlation coefficient	*P*
Grade				
G1	5 (21.7)	16.6 ± 7.6		
G2	11 (47.8)	14.5 ± 6.3	-0.233	0.284
G3	7 (30.4)	5.6 ± 2.4		
Tumor size				
T1	11 (47.8)	18.3 ± 5.2		
T2	7 (30.4)	8.8 ± 7.6	-0.443	0.034
T3	5 (21.7)	4.0 ± 1.5		
Nodal status				
N0	10 (43.5)	19.6 ± 5.6		
N1	6 (26.1)	3.2 ± 1.0	-0.401	0.058
N2-3	7 (30.4)	9.6 ± 7.4		

Abbreviations: IDC, invasive ductal carcinoma; SEM, standard error of the mean
Values were rounded to one decimal place

Concerning hormone receptor status, an association between the percentage of $CD44^+/CD24^-/Ck^+/CD45^-$ cells and ER^+ was found (*P* = 0.014) as observed in table 4. Furthermore, $HER2^-$ tumors also presented a significantly higher percentage of $CD44^+/CD24^-/Ck^+/CD45^-$ cells than $HER2^+$ breast tumors (*P* = 0.009). For E-Cadherin and vimentin immunostainings (Fig. 3), no significant differences were noticed between E-Cadherin expression and the percentage of such cells; however and for vimentin expression, the mean percentage of $CD44^+/CD24^-/Ck^+/CD45^-$ cells was higher in samples $Vimentin^+$ (*P* = 0.004, Table 4).

Table 4. Mean percentages of $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells according to patient's hormone receptor status and to E-Cadherin and vimentin expression.

	$CD44^+/CD24^-/Ck^+/CD45^-$ breast cells			
	DCIS, n (%)	IDC, n (%)	Mean ± SEM (%)	*P*
ER status				
ER+	3 (60.0)	18 (78.3)	14.2 ± 3.7	0.014
ER-	2 (40.0)	5 (21.7)	4.1 ± 1.6	
PR status				
PR+	3 (60.0)	16 (69.6)	14.1 ± 3.9	0.085
PR-	2 (40.0)	7 (30.4)	6.6 ± 3.2	
HER2 status				
HER2+	4 (80.0)	5 (21.7)	6.5 ± 2.1	0.009
HER2-	1 (20.0)	18 (78.3)	14.1 ± 4.1	
Ki-67 status				
Low	2 (40.0)	15 (65.2)	12.8 ± 4.1	0.214
High	3 (60.0)	8 (34.8)	9.9 ± 3.9	
Vimentin status				
Negative	0(0)	14 (60.9)	5.6 ± 2.9	0.004
Positive	5 (100)	9 (39.1)	17.6 ± 4.6	
E-Cadherin				
Negative/low expression	4 (80.0)	16 (69.6)	11.6 ± 3.6	0.995
High expression	1 (20.0)	7 (30.4)	11.9 ± 5.3	

Abbreviations: ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor; SEM standard error of the mean
Values were rounded to one decimal place

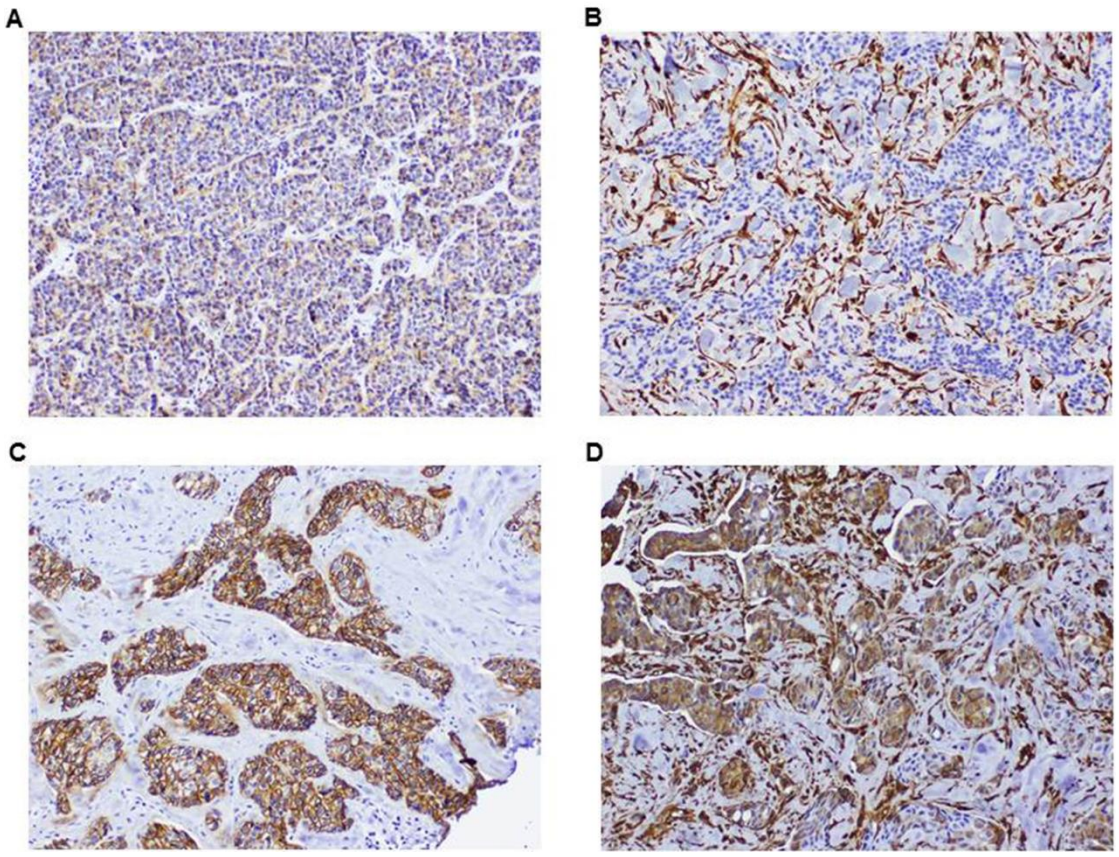

Fig. 3: Representative images of the immunohistochemistry for E-cadherin and Vimentin in breast tissue sections. (A) and (C) Negative and positive expression of E-Cadherin, respectively. (B) and (D) Negative and positive expression of vimentin, respectively.

Mutations identified in CD44+/CD24-/Ck+/CD45- breast cells

To determine the mutations associated to CD44+/CD24-/Ck+/CD45- breast cells from non-proliferative to invasive lesions, all cell populations were sequenced (n=6) using the CHPv2. Although the low numbers of CD44+/CD24-/Ck+/CD45- breast cells collected from the six mastectomy samples, we were able to obtain adequate libraries from all samples and subsequently sequenced without whole-genome amplification.

All target regions could be covered adequately (Table 5) with a mean read length of 106 bp except for one sample which showed a mean coverage below 100x and was consequently excluded from the analysis. All other samples were covered ≥500x on average. Non-synonymous somatic mutations identified in each sample are described in Table 6. The sample corresponding to a fibroadenoma only showed a SNP. In the pure DCIS samples, one frameshift mutation in the *TP53* gene was found. In DCIS within IDC, four different mutations were detected in the *NOTCH1, PTEN,* Harvey Rat Sarcoma Viral Oncogene Homolog (*HRAS*) and *AKT1* genes. Regarding the invasive cases, the luminal tumor harbored only one mutation in the *CSF1R* gene; for the triple-negative breast tumor, three mutations were identified affecting the *RET*, *TP53* and SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily B, Member 1 (*SMARCB1*) genes.

Table 5. Characteristics of the Targeted Sequencing applied in the DNA samples of CD44+/CD24-/Ck+/CD45- breast cells sorted from 6 mastectomies samples.

CD44+/CD24-/Ck+/CD45- breast cells				
Sample	Number of sorted cells	Mapped reads	On Target (%)	Uniformity (%)
FAD	126	2265645	96.3	99.6
Pure DCIS	553	1737683	95.8	97.9
DCIS within IDC	717	1291288	0.3	39.1
DCIS within IDC	1511	1408814	99.6	89.2
IDC (Luminal)	147	2775810	96.6	99.6
IDC (Triple-negative)	1900	2417404	98.7	92.8

Abbreviations: FAD, fibroadenoma; DCIS, ductal carcinoma *in situ;* IDC, invasive ductal carcinoma
Percentages were rounded to one decimal place

Table 6 Description of mutations identified in CD44+/CD24-/Ck+/CD45- breast cells of each mastectomy sample.

Sample	Locus	Genotype	Type	Genes	Variant ID	Transcript	Coding	Amino acid change	Variant effect
Fad	chr4:55593464	A/C	SNP	KIT	COSM28026	NM_000222.2	c.1621A>C	p.Met541Leu	Missense
Pure DCIS	chr17:7577099	CTCCCAGGACA GGCACAAACAC GCACCTCAAAG CTGTTCCG/C	INDEL	TP53	UN	NM_000546.5	c.799_838del40	p.Arg267fs	Frameshift Deletion
DCIS within IDC	chr9:139390819	G/A	SNV	NOTCH1	UN	NM_017617.3	c.7372C>T	p.Pro2458Ser	Missense
	chr10:89720802	TTAC/T	INDEL	PTEN	COSM4982	NM_000314.4	c.954_956delTAC	p.Thr319del	Non frameshift Deletion
	chr11:533908	T/C	SNV	HRAS	UN	NM_001130442.1	c.148A>G	p.Thr50Ala	Missense
	chr14:105246539	T/C	SNV	AKT1	UN	NM_001014431.1	c.61A>G	p.Thr21Ala	Missense
IDC (Luminal)	chr5:149433688	C/T	SNV	CSF1R	UN	NM_014983.2	c.2863G>A	p.Glu955Lys	Missense
IDC (Triple-negative)	chr10:43615605	T/C	SNV	RET	UN	NM_020975.4	c.2684T>C	p.Leu895Ser	Missense
	chr17:7577556	CAGG/AAGG	SNV	TP53	COSM10810	NM_000546.5	c.725G>T	p.Cys242Phe	Missense
	chr22:24145534	G/A	SNV	SMARCB1	UN	NM_003073.3	c.553G>A	p.Val185Met	Missense

Abbreviations: FAD, fibroadenoma; DCIS, ductal carcinoma *in situ;* IDC, invasive ductal carcinoma; SNP, single-nucleotide polymorphism; SNV, single.nucleotide variant; INDEL, insertion/deletion; UN Unknown.

Discussion

Considering that $CD44^+/CD24^{-/low}$ cells were already proved to be tumor initiating cells in breast cancer [10], our first purpose was to characterize $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells in non-malignant and malignant lesions. Regarding the results obtained, the median fluorescence intensity was significantly lower in malignant samples when compared with non-malignant ones for both CD24 and Ck. Considering that CD24 has been identified as a marker of differentiated normal mammary epithelial cells [10] and that along breast tumorigenesis, malignant cells become poorly differentiated (as seen with the lower fluorescence intensity of Ck), such result would then be expected.

The gradual enrichment of $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells from non-malignant lesions to DCIS and IDC lesions supports the notion of a tumor-initiation capability, which can be explained by the tenets of the CSC model. In fact, studies by others have already demonstrated the existence of $CD44^+/CD24^{-/low}$ breast cells in DCIS lesions but more importantly, that these cells exhibited enhanced invasive properties [25, 26].

Among the invasive cases, a negative correlation with the mean percentages of $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells and tumor size was obtained. The concept that BCSCs give rise to non-tumorigenic cancer cells that form a tumor mass through a symmetric division can explain the lower number of $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells seen in larger tumors.

Furthermore, and regarding the distribution of such cells according to hormone receptor status, a significant association with ER^+ and $HER2^-$ was obtained. This association can be misreported due to the small number of ER^- breast cancers used in this study. Nonetheless, and despite the reported association of $CD44^+/CD24^{-/low}$ cells with basal-like breast tumors, other studies did not find any association between this phenotype and hormone receptor status [1, 27]. Moreover, low expression of CD24 was also demonstrated to be associated with ER positivity and HER2 negativity, what can in part explain our results obtained [28, 29].

Additionally, $vimentin^+$ cases presented a higher mean percentage of $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells when compared with cases negative for this myoepithelial marker. Vimentin has recently been shown to be an important EMT

marker and a critical regulator of mesenchymal cell migration [30]. EMT is a developmental process during which epithelial cells lose epithelial characteristics and acquire mesenchymal properties by the disassembly of cell-cell junctions, loss of cell polarity and reorganization of the cytoskeleton, thereby acquiring increased motility. Because EMT endows cancer cells with migratory and invasive properties, it is implicated in tumor invasion and metastatic dissemination and has been linked with BCSCs [15]. Thus, our result can suggest a close association with $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells and the acquisition of an EMT state during breast cancer progression. What is noteworthy, however, is the fact that all DCIS samples were positive for vimentin and all of them contained $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells. Because EMT promotes tumor invasion by facilitating tumor cells to escape from the rigid constraints of the surrounding tissue architecture, such as basement membrane, the EMT state acquisition from these cells could be equally important for the progression to an invasive state [31].

Consistent with the notion that genetic alterations may lead to genomic instability and cancer development and considering the properties of BCSCs, which were already demonstrated to have several signaling pathways dysregulated [14-18], our second purpose was to identify, trough NGS, which somatic mutations were associated to $CD44^+/CD24^-/Ck^+/CD45^-$ cells in different breast lesions. We were able to identify somatic mutations in DNA samples from as little as 126 $CD44^+/CD24^-/Ck^+/CD45^-$ cells without the need to recur to whole-genome amplification technologies, thus reducing sequencing errors. Besides that, the CHPv2 assay was recently proved to be highly sensitive for manual inspection of such type of mutations and to be a reliable test in cancer diagnostic using both FFPE and frozen tissues even with low input DNA [32].

Regarding our samples, only a SNP in the tyrosine-protein kinase *KIT* gene was identified in our cell population taken from a fibroadenoma. Even with the notion that this tyrosine-protein kinase plays an essential role in stem cell maintenance [33], the clinical significance of this SNP in breast cancer remains unknown.

In pure DCIS, a frameshift deletion of *TP53* gene was found to be associated to $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells. In fact, loss of p53 in mammary epithelial cells was shown to activate the EMT program and to increase mammary stem cell population leading to the development of a high grade tumor

[34]. Considering that this gene is frequently mutated in breast cancer, it is plausible to assume that this frameshift deletion can aid $CD44^+/CD24^-/Ck^+/CD45^-$ cells in the development and progression to a malignant phenotype.

In DCIS within IDC, four different mutations were found to be associated to this phenotype affecting *NOTCH1*, *HRAS*, *PTEN*, and *AKT1* genes. Regarding the first one, aberrant activation of Notch signaling pathway was found to be an early event in breast cancer and high expression of NOTCH1 was seen to increase self-renewal capacity of BCSCs in DCIS lesions, contributing for breast cancer progression [35]. Hyperactivation of HRAS, in turn, was shown to induce cancer development in MCF10A human breast epithelial cells and to prompt EMT in CSCs via upregulation of vimentin [29]. In this way, the two missense mutations identified in these genes can eventually contribute for breast cancer progression of $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells by increasing their self-renewal ability and to promote the acquisition of an EMT state considering that vimentin was seen to be expressed in this particular case. As for *PTEN*, where a known INDEL was identified, an interesting study demonstrated that upon the inhibition of *PTEN*, the cancer stem-like fraction (characterized by $CD44^+/CD24^{-/low}$ cells) within the human breast cancer cell line MCF7 was significantly increased inducing activation of AKT in these cells, another gene found to be mutated in this case. In fact, increased AKT phosphorylation was shown to regulate BCSC expansion [36]. Besides that, vimentin phosphorylation regulated by AKT1 leads to its enhanced ability to induce motility and invasion [37]. Taken together, the mutations identified in these genes can strongly contribute for the expansion of $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells as well as for the progression of these cells towards an invasive state.

Concerning the invasive samples, only one mutation affecting the *CSF1R* gene was identified in $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells from the luminal tumor, probably due to the low number of these sorted cells (n=126). However, the Colony-stimulating factor-1 (CSF1) and its receptor CSF1R have been implicated in the pathogenesis and progression of various types of cancer, including breast cancer [38]. Indeed, high expression of CSF1 and CSFR1 in normal mammary epithelial cells was demonstrated to result in a dramatic stimulation of the invasive phenotype and anchorage-independent growth of these cells [39]. As referred in this study, breast tumors are comprised of phenotypically diverse populations of

breast cancer cells. Among them, BCSCs are important for tumor growth and metastasis. Thus, it is plausible to assume that mutations in the *CSF1R* gene can potentially stimulate $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells to do so.

Finally, cells sorted from the triple-negative tumor harbored three different mutations. *TP53* gene alterations occur in the majority of triple-negative tumors. [40]. Considering the reasons presented above between *TP53* mutations and the $CD44^+/CD24^{-/low}$ phenotype, this known missense mutation found in this cell population from this sample can have a crucial role for the progression of $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells along breast tumorigenesis. Concerning the mutation affecting the *RET* gene, its overexpression has recently been identified in triple-negative tumors [41]. RET is known to influence cellular proliferation, differentiation and migration [42]. Whether this identified mutation affects the behavior of BCSCs is currently unknown. Further studies are required to assess a possible association between mutated *RET* and the $CD44^+/CD24^-/Ck^+/CD45^-$ phenotype. Conversely, *SMARCB1* is considered to be highly conserved in solid tumors [43]. In fact, little is known about the effects of its deregulation in breast cancer Again, further demonstrations of existent mutations affecting this gene in breast cancer, specifically in BCSCs are required.

Apart from the mutations found in this study is the evidence that in any case where $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells were identified and sorted, the same mutation was noticed. Such marked heterogeneity can be the reason for the relative low impact of new pharmacological drugs targeting directly or indirectly BCSCs in breast cancer patients. Considering that the DCIS within IDC and IDC samples were taken from advanced breast cancers (grade 2 and 3, respectively), clonal evolution in BCSCs can be fundamentally important for such variability of mutations that were found. Important factors like time, tumor microenvironment and hormone exposure can trigger clonal selection among different populations of BCSCs harboring different mutations. Indeed, all genes that were found to be mutated in this study can play crucial roles for the development and transformation of $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells into a malignant phenotype, for stemness maintenance and acquisition of an EMT state.

Nonetheless, the lack of validation precludes us to conclude that we have truly identified stem cells and BCSCs. Moreover, our small number of cases used for this study limits some important conclusions like the statistical relevance

obtained for hormone receptor status or regarding the inter- and intra-tumor heterogeneity.

The use of massive parallel sequencing technology for this type of application revealed to be extremely effective even when using small amounts of DNA extracted from a low number of cells. Thus, such approach can be of particular interest to identify which mutations affects BCSCs for each patient in order to achieve more effective therapies for breast cancer treatment. Additional studies are now required involving functional analysis to infer about the tumorigenic effect of the most relevant mutations found. Besides that, the same approach is also needed using a larger cohort comprehending primary tumors and their correspondent metastasis to deepen the knowledge regarding the breast cancer progression of BCSCs and also to design an effective mutational profile of these cells.

Acknowledgments

The authors would like to thank the invaluable help of Manuela Certo and Maria José Oliveira during breast tissue collection.

Ethical Standard

This project was approved by the following ethical boards: Porto Hospital Centre Research Ethics Health Committee (reference 203-CES) and by Porto Hospital Centre Department of Education, Development and Research (reference 135-DEFI).

References

1. Abraham BK, Fritz P, McClellan M, Hauptvogel P, Athelogou M, Brauch H. Prevalence of $CD44^+/CD24^{-/low}$ cells in breast cancer may not be associated with clinical outcome but may favor distant metastasis. Clin Cancer Res 2005;11: 1154-1159.

2. Ginestier C, Min MH, Charafe-Jauffret E, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell stem cell 2007;1: 555-567.

3. Symmans WF, Liu J, Knowles DM, Inghirami G. Breast cancer heterogeneity: evaluation of clonality in primary and metastatic lesions. Hum Pathol 1995;26: 210-216.

4. Martelotto LG, Ng CK, Piscuoglio S, Weigelt B, Reis-Filho JS. Breast cancer intra-tumor heterogeneity. Breast Cancer Res Treat 2014;16: 210.

5. Neumeister v, Agarwal S, Bordeaux J, Camp RL, Rimm DL. In situ identification of putative cancer stem cells by multiplexing ALDH1, CD44, and cytokeratin identifies breast cancer patients with poor prognosis. Am J Pathol 2010;176: 2131-2138.

6. Xiao Y, Ye Y, Yearsley K, Jones S, Barsky SH. The lymphovascular embolus of inflammatory breast cancer expresses a stem cell-like phenotype. Am J Pathol 2008;173: 561-574.

7. Croker AK, Goodale D, Chu J, Postenka C, Hedley BD, Hess DA, Allan AL. High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability. J Cell Mol Med 2009;13: 2236–2252.

8. Da Cruz Paula A, Marques O, Rosa AM, Faria MDF, Rema A, Lopes C. Co-expression of Stem Cell Markers ALDH1 and CD44 in Non-malignant and Neoplastic Lesions of the Breast. Anticancer Res 2014;34:1427-1434.

9. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci USA 2003;100: 3983-3988.

10. Bane A, Viloria-Petit A, Pinnaduwage D, Mulligan AM, O'Malley FP, Andrulis IL. Clinical–pathologic significance of cancer stem cell marker expression in familial breast cancers. Breast Cancer Res Treat 2013;140: 195-205.

11. Liu R, Wang X, Chen GY, et al. The prognostic role of a gene signature from tumorigenic breast-cancer cells. N Engl J Med 2007;356: 217-226.

12. Li X, Lewis MT, Huang J, et al. Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. J Natl Cancer Inst 2008;100: 672-679.

13. Mani SA, Guo W, Liao MJ, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 2008;133: 704-715.

14. Diehn M, Cho RW, Lobo NA, et al. Association of reactive oxygen species levels and radioresistance in cancer stem cells. Nature 2009;458: 780-783.

15. Woodward WA, Chen MS, Behbod F, Alfaro MP, Buchholz T, Rosen JM. WNT/β-catenin mediates radiation resistance of mouse mammary progenitor cells. Proc Natl Acad Sci USA 2007;104: 618-623.

16. Phillips TM, McBride WH, Pajonk F. The response of CD24−/low/CD44+ breast cancer–initiating cells to radiation. J Natl Cancer Inst 2006;98: 1777-1785.

17. Zhan J, Chen L, Yuan Y. STAT1 promotes radioresistance of CD44+/CD24−/low cells in breast cancer. Exp Biol Med 2011;236: 418-422.

18. Wei W, Tweardy DJ, Zhang M et al. STAT3 Signaling Is Activated Preferentially in Tumor-Initiating Cells in Claudin-Low Models of Human Breast Cancer. Stem cells 2014;32: 2571-2582.

19. Hernandez L, Wilkerson PM, Lambros MB et al. Genomic and mutational profiling of ductal carcinomas in situ and matched adjacent invasive breast cancers reveals intra-tumour genetic heterogeneity and clonal selection. J Pathol 2012;227: 42-52.

20. Navin N, Kendall J, Troge J et al. Tumour evolution inferred by single-cell sequencing. Nature 2011;472: 90-94.

21. Jorge RF. Next-generation sequencing. Breast Cancer Res Treat 2009;11: S12.

22. Liu X, Mody K, de Abreu FB, et al. Molecular profiling of appendiceal epithelial tumors using massively parallel sequencing to identify somatic mutations. Clin Chem 2014;60: 1004-1011.

23. Amato E, dal Molin M, Mafficini A, et al. Targeted next-generation sequencing of cancer genes dissects the molecular profiles of intraductal papillary neoplasms of the pancreas. J Pathol 2014;233: 217-227.

24. Cheang MCU, Chia SK, Voduc D, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst 2009;101: 736–750.

25. Makki J, Myint O, Wynn AA, Samsudin AT, John DV. Expression distribution of cancer stem cells, epithelial to mesenchymal transition, and telomerase activity in breast cancer and their association with clinicopathologic characteristics. Clin Med Insights Pathol 2014;8: 1-16.

26. Sheridan C, Hiromitsu K, Robyn KF, et al. CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. Breast Cancer Res 2006;8:R59.

27. Mylona E, Giannopoulou I, Fasomytakis E, et al. The clinicopathologic and prognostic significance of CD44+/CD24−/low and CD44−/CD24+ tumor cells in invasive breast carcinomas. Hum Pathol 2008;39: 1096-1102.

28. Kaipparettu BA, Malik S, Konduri SD, et al. Estrogen-mediated downregulation of CD24 in breast cancer cells. Int J Cancer 2008;123: 66-72.

29. Hosonaga M, Arima Y, Sugihara E, Kohno N, Saya H. Expression of CD24 is associated with HER2 expression and supports HER2-Akt signaling in HER2-positive breast cancer cells. Cancer Sci 2014;105: 779-787.

30. Vuoriluoto K, Haugen H, Kiviluoto S, et al. Vimentin regulates EMT induction by Slug and oncogenic H-Ras and migration by governing Axl expression in breast cancer. Oncogene 2011;30: 1436-1448.

31. Choi Y, Lee HJ, Jang MH, et al. Epithelial-mesenchymal transition increases during the progression of in situ to invasive basal-like breast cancer. Hum Pathol 2013;44: 2581-2589.

32. de Leng W, Gadellaa-van Hooijdonk CG, Barendregt-Smouter FA, Koudijs MJ, et al. Targeted Next Generation Sequencing as a Reliable Diagnostic Assay for the Detection of Somatic Mutations in Tumours Using Minimal DNA Amounts from Formalin Fixed Paraffin Embedded Material. PloS one 2016;11: e0149405.

33. Zsebo KM, Williams DA, Geissler EN, et al. Stem cell factor is encoded at the SI locus of the mouse and is the ligand for the c-kit tyrosine kinase receptor. Cell 1990;63: 213-224.

34. Chang CJ, Chao CH, Xia W et al. p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. Nat Cell Biol 2011;13: 317-323.

35. Farnie G, Clarke RB, Spence K, et al. Novel cell culture technique for primary ductal carcinoma in situ: role of Notch and epidermal growth factor receptor signaling pathways. J Natl Cancer Inst 2007;99: 616-627.

36. Zhou J, Wulfkuhle J, Zhang H, et al. Activation of the PTEN/mTOR/STAT3 pathway in breast cancer stem-like cells is required for viability and maintenance. Proc Natl Acad Sci 2007;104: 16158-16163.

37. Zhu QS, Rosenblatt K, Huang KL, et al. Vimentin is a novel AKT1 target mediating motility and invasion. Oncogene 2011;30: 457-470.

38. Kacinski BM, Scata KA, Carters D, et al. FMS (CSF-l receptor) and CSF-l transcripts and protein are expiessed. Oncogene 1991;6: 941-952.

39. Sapi E, Flick MB, Rodov S, et al. Independent regulation of invasion and anchorage-independent growth by different autophosphorylation sites of the macrophage colony-stimulating factor 1 receptor. Cancer Res 1996;56: 5704-5712.

40. Gao J, Aksoy BA, Dresdner G, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013;269: pl1.

41. Gattelli A, Nalvarte I, Boulay A, et al. Ret inhibition decreases growth and metastatic potential of estrogen receptor positive breast cancer cells. EMBO Mol Med 2013;5: 1335-1350.

42. Kan Z, Jaiswal BS, Stinson J, et al. Diverse somatic mutation patterns and pathway alterations in human cancers. Nature 2010;466: 869-873.

43. Mimori K, Inoue H, Shiraishi T, et al. A single-nucleotide polymorphism of SMARCB1 in human breast cancers. Genomics 2002;80: 254-258.

Chapter 6

Discussion and Conclusions

1. General Discussion

A lot has been done in breast cancer research in order to reduce the high rates of incidence and mortality associated to this disease. From more advanced screening programs to more effective treatments, breast cancer continues to be the major concern in women health. Nonetheless, and despite all efforts, several limitations still exist preventing an elucidative comprehension regarding the progression of breast cancer cells as well as their ability to resist to certain therapeutic agents and to colonize other parts of the body.

Over the last decade, the notion that tumors are maintained by their own SCs, the so-called CSCs, has created great excitement in the research community [1]. In fact, the CSC hypothesis has become increasingly popular after the identification of defined tumor subsets endowed with tumorigenic activity and exhibiting phenotypic features of normal SCs. After its discovery in leukaemia, CSCs were isolated in many solid malignancies [2]. In breast cancer, the first report was by Al-Hajj *et al* where they isolated BCSCs designated as $CD44^{+}/CD24^{-/low}$ $lineage^{-}$. As few as 100 of these cells were sufficient to generate tumors when xenotransplanted into NOD/SCID mice, whereas tens of thousands of cells with alternative phenotypes failed to do so [3]. With such demonstration, a plethora of studies were published describing the impact of BCSCs identified by these established BCSC markers, as tumor initiating cells in breast cancer with high propensity to metastasize and to be resistant to therapeutic treatments [4-6]. However, and due to the high levels of heterogeneity associated with this disease, some breast cancers were shown not to contain any $CD44^{+}/CD24^{-/low}$ $lineage^{-}$ breast cell. As a consequence, other markers like ALDH1 or CD133 were reported [7, 8]. Indeed, a multiplexed method of *in situ* identification of putative BCSCs characterized by CD44 and ALDH1 was able to identify high risk patients in breast cancer [9]. Moreover, subpopulations of $ALDH1^{high}/CD44^{+}$ cells were recently identified in several human breast cancer cell lines which contributed to both chemotherapy and radiation resistance, suggesting a much broader role for ALDH1 in treatment response than previously reported [10, 11].

With this in mind and regarding the sequential progression of breast cancers seen from non-malignant to malignant lesions [12], we analyzed, by IHC, the expression of ALDH1 and CD44 in different breast lesions. What was

noteworthy in this study was the higher combined expression of these markers observed in DCIS when compared with IDCs. Looking at the concepts of the CSC model, cancers originate from the malignant transformation of an adult stem or progenitor cell through the deregulation of the normally tightly-regulated self-renewal program [13]. Regarding the important roles of CD44 (self-renewal, niche preparation and resistance to apoptosis) and ALDH1 (self-renewal, SC proliferation control, protection against oxidative insults), both markers can be responsible for the malignant transformation of BSCs into BCSCs. Moreover, if CSCs have the ability to self-renew and differentiate generating non-tumorigenic cancer cells that form a tumor mass [14], only a few cells would be enough for invasion, explaining thus the lower combined expression of these markers in IDCs. Associated to this explanation is the hypothetical model of an evolutionary bottleneck upon the transition from DCIS to IDC [15]. BCSCs can be subjected to a clonal selection during the progression to an invasive state where only a subpopulation of these cells will have the ability to invade. In other words, the low number of BCSCs that promote invasiveness can be the consequence of a clonal selection upon the transition of DCIS to IDC.

Among the important features of CSCs, we turned our attention to a particular one: the recent observation that CSCs mainly exist in a quiescent state. The isolation of adult SCs has revealed new insights about the epigenetic, transcriptional and post-transcriptional control of quiescence. It was proposed that an actively preserved state of quiescence would exist being regulated by signaling pathways that sustain a controlled state, allowing thus a rapid activation [16]. Besides that, SC quiescence is highly relevant for cancer therapy since that quiescent CSCs are often resistant to both chemotherapy and targeted therapies contributing to relapse following discontinuation of therapy [17]. In fact and more recently, such dormant cells were identified in pancreatic adenocarcinoma and shown to be enriched for CSC markers such as CD133, CD44, CD24 and ALDH1 [16]. Regarding the results obtained in the first study, we developed a triple-immunostaining method to identify $CD44^{+}/ALDH1^{+}$ breast cells in their quiescent state ($Ki\text{-}67^{-}$) in a cohort comprehending different breast lesions.

As expected, the mean percentage of $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ breast cells was higher in DCIS when compared with IDCs. Interestingly and apart from the number of these cells, some isolated pools of $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ tumor cells

were detected in DCIS and IDCs, highlighting another concept of the CSC model: the existence of a CSC niche. This specific niche comprehend different types of cells (niche cells, stromal cells, immune cells and vasculature) that surrounds the CSCs and also, secreted factors derived from these cells, which offer a “fertile territory” for CSCs to propagate [18]. Hence, for CSCs that exhibit stem cell-like features and have the capability to regenerate the bulk of tumor cells without losing their self-renewal propensity, the CSC niche act as a microenvironment regulatory system for these cells [19]. Both CD44 and ALDH1 display important roles in the regulation of BCSCs and the maintenance of their stemness properties [15]. Thus, $CD44^{+/}ALDH1^{+}/Ki\text{-}67^{-}$ breast cells identified in the malignant tissues can take advantage of the niche that supports normal SCs to progress along the tumorigenic process. Another pertinent point is their existence in a quiescent state. Quiescent SCs are able to respond to stimuli that originate from their niche environment by activating and entering the cell cycle [20]. In a similar way, quiescent CSCs may be prompted for activation by specific energetically favorable mechanisms from their niche that are compatible with the low metabolic state of quiescence [20]. Also like normal SCs, the detachment of CSCs from their niche leads to an asymmetric or symmetric division, another ability sustained by the roles of CD44 and ALDH1 [15]. In fact, this ability can also explain the negative correlation obtained from the mean percentages of $CD44^{+/}ALDH1^{+}/Ki\text{-}67^{-}$ tumor cells with tumor size. If these cells are capable to differentiate generating non-tumorigenic cancer cells, large tumors will proportionally have less tumor initiating cells than smaller ones.

Furthermore, the observation that $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ tumor cells can favor distant metastasis and are able to predict OS in breast cancer is also of great importance. These BCSCs markers were already described to be determinant for treatment resistance, recurrences and metastasis development due to the expression of high levels of therapy-resistance proteins [11, 21]. ALDH1 activity has been shown to render cancer cells exquisitely resistant to some chemotherapy agents mainly due to its well-characterized role in differentiation through the retinoic acid pathway [22, 23]. CD44 plays a potent role in every aspects of breast cancer involving cancer cell proliferation, progression and metastasis. High levels of CD44 expression enhance the invasion of BCSCs; CD44s dynamic association with cytoskeletal proteins enables motility, while CD44

clustering also facilitates localization to secondary metastatic sites through CD44/Hyaluronan acid binding [24]. With such features, it is plausible to assume that both CD44 and ALDH1 can play important roles for the resistance and dissemination of BCSCs.

Moreover**,** quiescence may also have a determinant role in tumor progression and relapse. Several chemotherapeutic agents as well as radiotherapy work by inducing DNA damage. Thus, cells that have the ability to repair DNA damages are more prone to survive chemotherapy. Regarding the properties of quiescence, quiescent cells may have the potential and/or time to repair the damage inflected to them. Although quiescence is not an essential characteristic that defines SCs, in BCSCs, there is increased expression of DNA repair genes, indicating that high DNA repair pathway activity may aid in making CSCs resistant to tumor therapy [16].

Apart from the roles of CD44 and ALDH1 and the effects of cell dormancy in breast cancer progression, is the recent suggestion of a co-existence between the CSC model and the clonal evolution model. From a different perspective of what was proposed by Clevers in 2011, tumor progression has been related to ‘Darwinian’ evolution. The expansion of an established tumor can be explained by the generation of new CSC clones as a result from mutations, genetic instability or epigenetic alterations. The prevalence of new CSCs and their clones will be determined by different selective pressures (nutritional or immune status, oxygenation and therapy) that modify the tumor microenvironment. If selected, these cells can be responsible for tumor relapse or metastasis (Figure 1) [25].

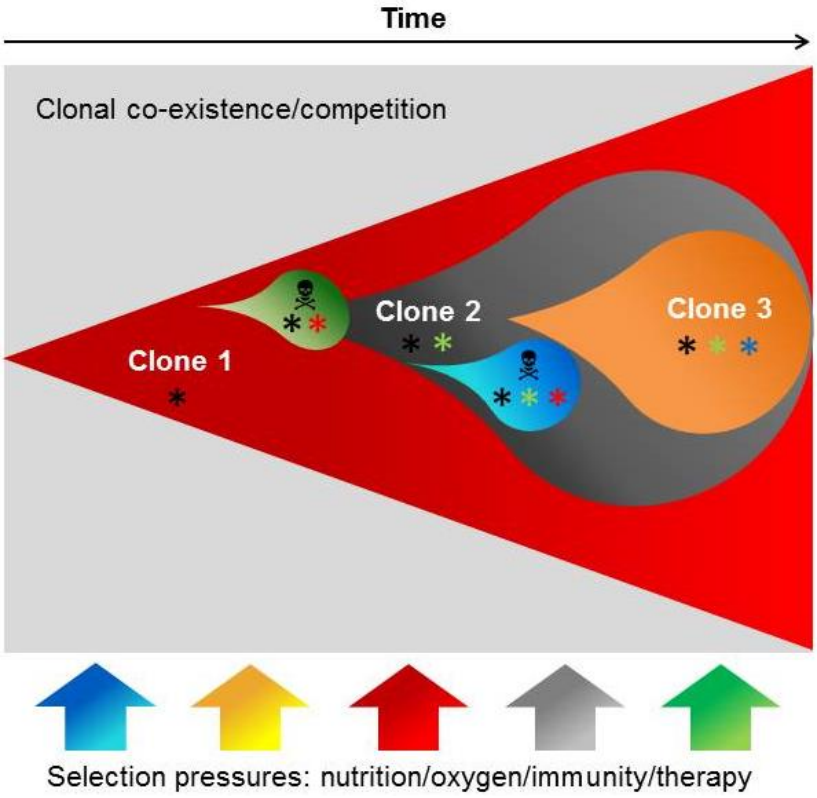

Figure 1. Clonal co-existence. Most tumors have their origin in a single mutated cell so the cells in the resultant clone will have the same founder mutation (black asterisk). During progression, different selection pressures (nutritional or immune status, oxygenation, therapy) could result in the emergence of new clones through genetic and epigenetic alteration. *Adapted from [25]*.

Despite these novel findings regarding the identification of $CD44^{+}/ALDH1^{+}$ breast cells in their quiescent state ($Ki\text{-}67^{-}$), the $CD44^{+}/CD24^{-/low}$ phenotype continues to be by far the most studied phenotype in breast cancer having several signaling pathways demonstrated to be dysregulated [6, 26-29]. Moreover, this phenotype was also demonstrated to harbor potential mechanisms of chemotherapy and radiation resistance including the presence of lower concentration of reactive oxygen species, cell dormancy, efficient DNA repair mechanisms, overexpression of EMT markers, and STAT1 and STAT3 signaling activation [30, 31, 28-29]. With all this knowledge, the use of more advanced technologies like NGS that could deepen the molecular characterization of this phenotype can be seen as an interesting approach.

NGS has recently been used for the analysis of the molecular features of early stage breast cancer, leading to a genomic portrait of this disease. Within such portrait, *TP53* and *PIK3CA* mutations are the most frequent genomic

alterations. Besides these mutated genes, clinical relevance of *PTEN* mutations and deletions as well as *AKT1*, *BRCA1* or *BRCA2* mutations is also highlighted [32]. With the current availability of standardized NGS kits providing reliable sequencing results in routine cancer diagnostics [33, 34], we used the CHPv2 to analyze the molecular features of $CD44^+/CD24^-/Ck^+/CD45^-$ cells in different breast lesions.

Considering that we used FCM and FACS technologies to identify and isolate $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells, we have also characterized these cells by FCM from a cohort comprehending non-malignant and malignant lesions. From this analysis, the gradual enrichment of $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells observed from benign to IDC lesions supports the idea of a tumor initiation capability from these cells. Interestingly and similar to the $CD44^+/ALDH1^+/Ki\text{-}67^-$ phenotype, a negative correlation was obtained from the mean percentages of $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells with tumor size, strengthening the idea that BCSCs are able to differentiate in non-tumorigenic cells that form a tumor mass [14].

Additionally and with the recent observations that the $CD44^+/CD24^{-/low}$ phenotype possess an epithelial-mesenchymal transition (EMT) capacity, vimentin and E-Cadherin expression was analyzed by IHC. Indeed, the association seen between vimentin expression and the mean percentages of $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells also strength the notion of an EMT state acquisition from these cells during breast cancer progression [31].

The novelty in this study was the application of MPS technologies in the CSC field. Despite the low number of cases used for this purpose, we were able to detect somatic mutations in as little as 126 $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells without the need to recur to whole-genome amplification technologies.

In DCIS lesions, mutations affecting the *TP53*, *NOTCH1*, *HRAS*, *PTEN* and *AKT1* genes were found to be associated to the $CD44^+/CD24^-/Ck^+/CD45^-$ phenotype. Considering the effects of such altered genes in breast tumorigenesis [35-38], these mutations can strongly contribute for the expansion of $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells as well as for the progression of these cells towards an invasive state.

Concerning the invasive samples, only one mutation affecting the *CSF1R* gene was identified in $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells isolated from the

luminal tumor. Such mutation can potentially stimulate these cells to proliferate and to colonize other parts of the body [39].

Finally, cells sorted from the triple-negative tumor harbored three different mutations affecting the *TP53*, *RET* and *SMARCB1* genes. While *TP53* alterations are well characterized in breast cancer [40], the functions of altered *RET* and *SMARCB1* genes in breast cancer are still undefined.

Looking at the number of mutations that was found in isolated $CD44^+/CD24^-/Ck^+/CD45^-$ cells from each sample, the DCIS within IDC lesion presented more mutations than the IDC lesions. Such result can be explained, once again, by the hypothetical model of an evolutionary bottleneck upon the transition from DCIS to IDC. Assuming the existence of different subpopulations of BCSCs, clonal selection may favor those who harbor advantageous mutations (driver mutations) for the progression to an invasive state.

Moreover, clonal evolution may also explain the variability of the mutations that were found in isolated $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells from the malignant samples. Important factors like time, tumor microenvironment and hormone exposure affect breast tumors differently. Hence, it is plausible to assume that each population of $CD44^+/CD24^-/Ck^+/CD45^-$ tumor cells that was isolated from each sample was subjected to different clonal selections harboring thus different mutations. With such variability, the identification of driver mutations can be extremely challenging. In fact, all genes that were found to be mutated in this study, especially in DCIS lesions, can play crucial roles for the development and transformation of $CD44^+/CD24^-/Ck^+/CD45^-$ breast cells into a malignant phenotype, for stemness maintenance and acquisition of an EMT state.

References

1. H. Clevers, The cancer stem cell: premises, promises and challenges, Nat. Med., (2011) 313-319.

2. J.E. Visvader JE, G.J Lindeman 2008) Cancer stem cells in solidtumours: accumulating evidence and unresolved questions, Nat. Rev. Cancer, 8 (2008) 755–768.

3. M. Al-Hajj, M.S. Whicha, A. Bentio-Hernandez, S.J. Morrison, M.F. Clarke, Prospective identification of tumorigenic breast cancer cells, Proc. Natl. Aca. Sci., 100 (2003) 3983-3988.

4. C Sheridan, H. Kishimoto, R.K. Fuchs, S. Mehrotra, P. Bhat-Nakshatri, C.H. Turner, R. Goulet, S. Badve, H. Nakshatri, CD44+/CD24-breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis, Breast Cancer Res., 8 (2006) R59..

5. T.M. Phillips, H.M. William, P. Frank, The response of $CD24^{-/low}/CD44^{+}$ breast cancer–initiating cells to radiation, J. Natl. Cancer Inst. 24 (2006) 1777-1785.

6. M. Shipitsin, L.L. Campbell, P. Argani, S. Weremowicz, N. Bloushtain-Qimron, J. Yao, T. Nikolskaya, T. Serebryiskaya, R. Beroukhim, M. Hou, M.K Haloushka, S. Sukumar, L.M. Parker, K.S. Anderson, L.N. Harris, J.E. Garber, A.L. Richardson, S.J. Schnitt, Y. Nikolsky, R.S. Gelman, K. Polyak, Molecular definition of breast tumor heterogeneity, Cancer Cell, 11 (2007) 259-273.

7. C. Ginestier, M.H. Hur, E. Charafe-Jauffret, F. Monville, J. Ducther, M. Brown, J. Jacquemier, P. Viens, C.G. Kleer, S. Liu, A. Schott, D. Hayes, D. Birnbaum, M.S. Wicha, G. Dontu, ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome, Cell Stem Cell, 1 (2007) 555-567.

8. M.H. Wright, A.M. Calcagno, C.D. Salcido, M.D. Carlson, S.V. Ambudkar, L. Varticovski, Brca1 breast tumors contain distinct CD44+/CD24-and CD133+ cells with cancer stem cell characteristics, Breast Cancer Res., 10 (2008) R10.

9. V. Neumeister, In situ identification of putative cancer stem cells by multiplexing ALDH1, CD44, and cytokeratin identifies breast cancer patients with poor prognosis, Am. J. Pathol., 176 (2010) 2131-2138.

10. A.K. Croker, A.L. Allan, Inhibition of aldehyde dehydrogenase (ALDH) activity reduces chemotherapy and radiation resistance of stem-like ALDHhiCD44+ human breast cancer cells, Breast Cancer Res. Treat., 133 (2012) 75-87.

11. H.E. Lee, J.H. Kim, Y.J. Kim, S.Y. Choi, S.W. Kim, E. Kang, I.Y. Chung, I.A. Kim, E.J. Kim, Y. Choi, H.S. Ryu, S.Y. Park, An increase in cancer stem cell population after primary systemic therapy is a poor prognostic factor in breast cancer, Br. J. Cancer., 104 (2011) 1730-1738.

12. A. Bombonati, D.C. Sgroi, The molecular pathology of breast cancer progression, J. Pathol, 223 (2011) 308-318.

13. M. Shackleton, Normal stem cells and cancer stem cells: similar and different. Semin. Cancer Biol., 20 (2010) 85-92.

14. K. Kai, Y. Arima, T. Kamiya, Breast cancer stem cells, Breast Cancer, 17 (2010) 80-85.

15. C.F. Cowell, B. Weigelt, R.A. Sakr, C.K.Y. Niang, J. Hicks, T.A. King, J.S. Reis-Filho, Progression from ductal carcinoma in situ to invasive breast cancer: revisited, Mol. Oncol., 7 (2013) 859-869.

16. S. Colak, J.P. Medema, Cancer stem cells–important players in tumor therapy resistance, FEBS J., 21 (2014) 4779-4791.

17. A. Roesch, M. Fukanaga-Kalabis, E.C. Schmidt, S.E. Zabierowski, P.E. Brafford, A. Vultur, D. Basu, P. Gimotty, T. Vogt, M. Herlyn, A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth, Cell, 4 (2010) 583-594.

18. S. Chen, E.H. Huang, The colon cancer stem cell microenvironment holds keys to future cancer therapy, J. Gastro. Surg., 18 (2014) 1040-1048.

19. F. Pajonk, E. Vlashi, Characterization of the stem cell niche and its importance in radiobiological response, Sem. Rad. Oncol., 23 (2013) 237-241.

20. T.H. Cheung, T.A. Rando, Molecular regulation of stem cell quiescence, Nat. Rev. Mol. Cell Biol., 14 (2013) 329-340.

21. A.K. Croker, D. Goodale, J. Chu, C. Postenka, B.D. Hedley, D.A. Hess, A.L. Allan, High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability, J. Cell. Mol. Med., 13 (2009) 2236-2252.

22. C. Ginestier, J. Wicinski, N. Cervera, F. Monville, P. Finetti, F. Bertucci, M.S. Wicha, D. Birnbaum, E. Charafe-Jauffret, Retinoid signaling regulates breast cancer stem cell differentiation, Cell Cycle, 8 (2009) 3297–3302.

23. T. Tanei, K. Morimoto, K. Shimazu, S.J. Kim, Y. Tanji, T. Tagushi, Y. Tamaki, S. Nogushi, Association of breast cancer stem cells identified by aldehyde dehydrogenase 1 expression with resistance to sequential Paclitaxel and epirubicin-based chemotherapy for breast cancers, Clin. Cancer Res., 15 (2009) 4234-4241.

24. G. Mattheolabakis, L. Milane, A. Singh, M.M. Amiji, Hyaluronic acid targeting of CD44 for cancer therapy: from receptor biology to nanomedicine, J. Drug Target., 23 (2015) 605-618.

25. M.R. Alison, W.R. Lin, S.M. Lim, L.J. Nicholson, Cancer stem cells: in the line of fire, Cancer Treat. Rev., 6 (2012) 589-598.

26. M. Diehn, R.W. Cho, N.A. Lobo, T. Kalisky, M.J. Dorie, A.N. Kulp, D. Qian, J.S. Lam, L.E. Ailles, M. Wong, B. Joshua, M.J. Kaplan, I. Wapnir, F.M. Dirbas, G. Somlo, C. Garberoglio, B. Paz, J. Shen, S.K. Lau, S.R. Quake, J.M. Brown, I.L. Weissman, M.F. Clarke, Association of reactive oxygen species levels and radioresistance in cancer stem cells, Nature, 458 (2009) 780-783.

27. W.A. Woodward, M.S. Chen, F. Behbod, M.P. Alfaro, T.A. Buchholz, J.M. Rosen, WNT/β-catenin mediates radiation resistance of mouse mammary progenitor cells. Proc. Natl Aca. Sci., 104 (2007) 618-623.

28. J.F. Zhan, L. Chen, Y. Yuan, G. Xie, A. Sun, Y. Liu, Z.X. Chen, STAT1 promotes radioresistance of CD44+/CD24−/low cells in breast cancer, Exp. Biol. Med., 236 (2011) 418-422.

29. W. Wei, D.J. Tweardy, M. Zhang, J. Landua, I. Petrovic, W. Bu, K. Roarty, S.G. Hilsenbeck, J.M. Rosen, M.T. Lewis, STAT3 Signaling Is Activated Preferentially in Tumor-Initiating Cells in Claudin-Low Models of Human Breast Cancer, Stem Cells, 32 (2014) 2571-2582.

30. X. Li, M.T. Lewis, J. Huang, C. Gutierrez, C.K. Osborne, M.F. Hu, S.G. Hilsenbeck, A. Pavlick, X. Zhang, G.C. Chamness, H. Wong, J. Rosen, J.C. Chang, Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. J. Natl. Cancer Inst., 100 (2008) 672-679.

31. S.A. Mani, W. Guo, M.J. Liao, E.N. Eaton, A. Ayyanan, A.Y. Zhou, M. Brooks, F. Reinhard, C.C. Zhang, M. Shipitsin, L.L. Campbell, K. Polyak, C. Brisken, J. Yang, R.A. Weinberg, The epithelial-mesenchymal transition generates cells with properties of stem cells, Cell, 133 (2008) 704-715.

32. M.R. Stratton, P.J. Campbell, P.A. Futreal, The cancer genome, Nature, 458 (2009) 719-724.

33. X. Liu, K. Mody, F.B. de Abreu, J.M. Pipas, J.D. Peterson, T.L. Gallagher, A.A. Suriawinata, G.H. Ripple, K.C. Hourdequin, K.D. Smith, R.J. Barth, T.A. Colacchio, M.J. Tsapakos, B.I. Zaki, T.B. Gardner, S.R. Gordon, C.I. Amos, W.A. Wells, G.J. Tsongalis, Molecular profiling of appendiceal epithelial tumors using massively parallel sequencing to identify somatic mutations, Clin. Chem., 60 (2014) 1004-1011.

34. E. Amato, M. dal Molin, A. Mafficini, J. Yu, G. Malleo, B. Rusev, M. Fassan, D. Antonello, Y. Sadakari, P. Castelli, G. Zamboni, A. Maitra, R. Salvia, R.H. Hruban, C. Bassi, P. Capelli, R.T. Lawlor, M. Coggins, A. Scarpa, Targeted next-generation sequencing of cancer genes dissects the molecular profiles of intraductal papillary neoplasms of the pancreas. J. Pathol., 233 (2014) 217-227.

35. K. Vuoriluoto, H. Haugen, S. Kiviluoto, J.P. Mpindi, J. Nevo, C. Gjerdrum, C. Tiron, J.B. Lorens, J. Ivaska, Vimentin regulates EMT induction by Slug and oncogenic H-Ras and migration by governing Axl expression in breast cancer, Oncogene, 30 (2011) 1436-1448.

36. G. Farnie, R.B. Clarke, K. Spence, N. Pinnok, K. Brennan, N.G. Anderson, N.J. Bundred, Novel cell culture technique for primary ductal carcinoma in situ: role of Notch and epidermal growth factor receptor signaling pathways. J. Natl. Cancer Inst., 99 (2007) 616-627.

37. J. Zhou, J. Wulfkuhle, H. Zhang, P. Gu, Y. Yang, J. Deng, Y. Zhang, Activation of the PTEN/mTOR/STAT3 pathway in breast cancer stem-like cells is required for viability and maintenance, Proc. Natl. Acad. Sci. 104 (2007) 16158-16163.

38. Q.S. Zhu, K. Rosenblatt, K.L. Huang, G. Lahat, R. Brobey, S. Bolshakov, T. Nguyen, Z. Ding, R. Belousov, K. Bill, X. Luo, A. Lazar, A. Dicker, G.B. Mills, M.C. Hung, D. Lev, Vimentin is a novel AKT1 target mediating motility and invasion, Oncogene, 30 (2011) 457-470.

39. E. Sapi, M.B. Flic, S. Rodov, M. Gilmore-Hebert, M. Kelley, S. Rockwell, B.M. Kacinski, Independent regulation of invasion and anchorage-independent growth by different autophosphorylation sites of the macrophage colony-stimulating factor 1 receptor, Cancer Res., 56 (1996) 5704-5712.

40. J. Gao, B.A. Aksoy, G. U. Dogrusoz, Dresdner, B. Gross, S.O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, N. Schultz, Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal, Sci. Signal 269 (2013) pl1.

2. Conclusions and Future Perspectives

In this thesis work, we aimed to characterize different BCSC phenotypes in non-malignant and malignant lesions. With the results obtained from the analysis of CD44 and ALDH1 expression in different breast specimens, where a combined overexpression of both markers was shown to be significantly higher in DCIS when compared with IDCs, a triple-immunostaining was developed. From this triple-immunostaining, we characterized the $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ phenotype and conclude that $CD44^{+}/ALDH1^{+}/Ki\text{-}67^{-}$ tumor cells may have a higher tumorigenic effect in breast cancer than $CD44^{+}/CD24^{-/low}$ tumor cells. In fact, the roles of ALDH1 and CD44 can be determinant for the behavior of CSCs and the ability to resist to chemotherapeutic agents and their dissemination to other parts of the body. Besides that, we also conclude that quiescence can have a more preponderant role than previously expected which can be crucial for tumor progression, resistance to chemotherapeutic agents and metastatic spread of BCSCs. From this first characterization, additional studies are now required to infer about the tumorigenic and metastatic ability of $CD44^{+}/ALDH1^{+/high}$ tumor cells combined with their quiescence status.

Considering the well-known features of the $CD44^{+}/CD24^{-/low}$ phenotype in breast cancer, we used FCM, FACS and NGS technologies to characterize $CD44^{+}/CD24^{-}/Ck^{+}/CD45^{-}$ cells in non-malignant and malignant breast tissues. The use of MPS for this type of application revealed to be extremely effective even when using small amounts of DNA extracted from a low number of cells. Considering the results obtained from this study, functional analyses are now needed to determine the tumorigenic effect of the most relevant mutations found in this work. Besides that, we conclude that the application of NGS technologies can be of particular interest to determine which mutations affect the behavior of BCSCs in order to achieve more effective therapies for breast cancer treatment. Due to the high levels of heterogeneity existing in breast tumors, a practical example would be the application of NGS in isolated BCSCs from primary tumors and their corresponding metastases in order to determine which gene is more frequently mutated (hotspot mutations) in each BCSC population.

In the future and with the continuous improvements of NGS, we will be able to deepen our knowledge regarding the progression of BCSCs and also to design an effective mutational profile of these cells.