U.PORTO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

STEM CELLS IN NORMAL AND MALIGNANT BREAST TISSUE: IS P-CADHERIN A STEM CELL MARKER AND A POSSIBLE TARGET FOR CANCER STEM CELL THERAPY?

André Filipe de Barros Vieira

Tese de Doutoramento em
Ciências Biomédicas

2012

Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto

Instituto de Patologia e Imunologia Molecular da Universidade do Porto

Paterson Institute for Cancer Research, University of Manchester, Manchester, UK

ANDRÉ FILIPE DE BARROS VIEIRA

STEM CELLS IN NORMAL AND MALIGNANT BREAST TISSUE: IS P-CADHERIN A STEM CELL MARKER AND A POSSIBLE TARGET FOR CANCER STEM CELL THERAPY?

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

Orientador:
Doutora Joana Cancela de Amorim Falcão Paredes

Categoria:
Investigadora Principal do Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP) e Docente Afiliada da Faculdade de Medicina da Universidade do Porto

Afiliação:
Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP)

Dedicated to all of those who lost the fight and to those who wont quit fighting

Dedicado a todos aqueles que perderam esta batalha e àqueles que não desistem de lutar

Agradecimentos

Sem prejuízo para as demais pessoas envolvidas na concretização deste grande objetivo, o primeiro agradecimento estará sempre reservado à minha orientadora de Doutoramento, Joana Paredes. Obrigado por aceitares este desafio! Juntos no mesmo barco, iniciamos um projeto arriscado e chegamos a bom porto! Foi uma experiência gratificante crescer cientificamente no teu grupo, pela tua inspiração e por quereres ver coisas que os outros não vêem. Agradeço-te a liberdade e a confiança depositada no desenvolvimento do meu trabalho científico. Simplesmente espero, com estas páginas, ter retribuído o investimento que fizeste em mim.

Um reconhecimento muito especial à Raquel Seruca. Não conheço ninguém que viva a ciência tão fervorosamente. Obrigado pelas discussões calorosas! Poucas pessoas conseguem tão brilhantemente reunir o saber científico, a capacidade de liderança, o altruísmo, a compreensão e a preocupação a nível pessoal e social. Juntamente com a Céu Figueiredo, o José Carlos, a Carla Oliveira e a Joana Paredes, a Raquel formou o grupo mais forte em investigação de cancro em Portugal.

Ao Professor Fernando Schmitt, um exemplo de liderança e de didática em Patologia. Com muito trabalho e imenso talento, construiu um grupo de investigação de referência em Portugal e no estrangeiro. É com muito orgulho que hoje partilho a sua experiência e saber.

To Robert Clarke, a major thank you for receiving me in your lab! Rob is an extraordinary researcher whose thorough criticisms made me a much better scientist.
I also appreciate enormously the cooperation form the Breast Biology group and the way I was welcomed in the Paterson Institute for Cancer Research. I especially wish to thank Gillian Farnie, who gave me many suggestions that were outstandingly important in constructing this study.
This thesis success is also the result of many people that helped me in the lab. I have to emphasize that no one could be more supportive than Mathew Ablett and a great deal of what I have learnt was thanks to him.

Ao Professor Manuel Sobrinho Simões, o Diretor, sempre com sorriso e uma palavra amiga, obrigado por me incutir este carinho tão especial pelo nosso Instituto!

Ao Professor Pedro Moradas-Ferreira, um amigo excecional, que não me deixou desistir deste grande projeto.

Ao grupo da Patologia Mamária, parte integrante da Divisão de Genética do Cancro do IPATIMUP. Em lugar nenhum no mundo, eu me senti tão acolhido! Posso dizer que aqui, formei grandes laços de amizade com todos sem exceção e juntos constituímos uma grande família… Por isso, aqui vai um grande obrigado a esta gente fantástica: Maria Rita (a minha grande amiga), Ana Sofia (sempre tão rápida e descontraída), Sara Ricardo (a minha camarada neste projeto), Bábi (a grande tratadora dos animais e... do lab), Ritinha (a princesa), Diana (a Sr.ª Professora), Madalena (a Leninha), Daniela (a pupila), Nair (a lutadora), Nuno Bernardes (o mouro), Carlinhos (um elemento chave da mama), André Albergaria (o meu homónimo e excelente relações públicas), Sílvia (a boa disposição em pessoa), Zé Luís (o nosso expert em DNA), e a lista continua... Renata, Cecília, Ângela, Ana Costa, Daniel, Joaninha, Rui, Soraia, Xu, Marta Correia, Joana Carvalho, Sofia, Patrícia Oliveira, Mafalda, Margarida, Hugo, Miguel, Inês. E uma palavra de agradecimento às meninas do INEB: Marta Pinto, Catarina Almeida, Maria Molinos, Patrícia, Ana Pinto e, claro... à Maria Oliveira, que sempre arranja um tempinho para mim.

A ajuda de algumas pessoas foi absolutamente critica para a concretização desta tese. Um grande agradecimento ao Renê Gerhard, Nuno Mendes, Jorge Camesele-Teijeiro, à Dina Leitão e à Regina Pinto.

My friends in the UK are truely amazing. I am missing those nights out... and nights in... that allowed me to make some friends for life!
To Guilherme and Sílvia, thank you for making me feel like home. Of all of the Portuguese people dispersed troughout the Globe, I was very fortunate to find you!
Sharmin 'Princesa', your fabulous Indian food makes me want to go back to Manchester.
Angélica, I am still waiting for your visit to Porto! Sorry I did not teach you how to speak Portuguese properly. That said, my Spanish was also a bit rusty.
Erik, I am missing the squash lessons in the freezing cold!
The luch group: Cristina, Flor, Monica, Andrej, Millena, Christophe, Olga, Eli, Sara.

The lab would never function properly without the precious contribution of Ron Blance, Kath Spence, Pamela Willan and Julia Resh. I miss you all so much. An enormous thank you to the girls from the lab: Ciara O'Brian, Jagdeep Singh, Kate, Hannah Harrison and Becky Lamb.

Deixo gravado um agradecimento à Fundação para a Ciência e Tecnologia, que financiou o meu custo de vida com uma bolsa individual durante os últimos quatro anos. Os apoios institucionais passam ainda pela Fundação Calouste Gulbenkian, que financiou o projeto científico e à Associação Luso-Britânica.

Finalmente, à minha família em especial à minha mãe, a quem devo um sincero pedido de desculpas, porque teve de me suportar nos dias em que o trabalho científico terminou em insucessos (que foram muitos...) e a quem exigi muitos sacrifícios. Um enorme beijinho para a Joana e um grande abraço para o Nuno!

Agradeço ainda a todas as pessoas que, por lapso ou omissão, não estão aqui citadas, mas com quem partilhei os bons e os maus momentos da minha vida nos últimos 4 anos.

Não posso deixar passar uma homenagem especial a todas as mulheres que enfrentaram o peso de sofrer desta terrível doença e que contribuíram para o estudo que aqui apresento. É graças a elas que o estudo do cancro da mama deu mais um passo em frente.

Este foi um grande desafio pessoal, que atingi com força e determinação!

Um Muito Obrigado...
☺

André F. Vieira

IPATIMUP
Instituto de Patologia e Imunologia Molecular da Universidade do Porto

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

FUNDAÇÃO CALOUSTE GULBENKIAN

ASSOCIAÇÃO LUSO-BRITÂNICA DO PORTO

Nos termos do artº. 34º. do Diário da República nº. 84-A/2007, declara-se que o trabalho descrito nesta tese foi suportado por uma bolsa individual da Fundação para a Ciência e Tecnologia (SFRH/BD/41363/2007), cofinanciada pelo Fundo Social Europeu

Diretivas legais

No cumprimento do disposto no Decreto-Lei 216/92 de 14 de Outubro, declara-se que o autor desta dissertação participou ativamente na conceção e na execução do trabalho experimental que esteve na origem dos resultados apresentados, bem como na sua interpretação e na redação dos respetivos manuscritos.
Fazem parte integrante desta tese os seguintes artigos científicos originais publicados ou submetidos para publicação:

I. Sara Ricardo, André Filipe Vieira, Renê Gerhard, Dina Leitão, Regina Pinto, Jorge F. Cameselle-Teijeiro, Fernanda Milanezi, Fernando Schmitt, Joana Paredes. **Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtypes**. *Journal of Clinical Pathology* 64:937-46, 2011.

II. André Filipe Vieira, Sara Ricardo, Matthew Paul Ablett, Maria Rita Dionísio, Nuno Mendes, André Albergaria, Gillian Farnie, Renê Gerhard, Jorge F. Cameselle-Teijeiro, Raquel Seruca, Fernando Schmitt, Robert B. Clarke, Joana Paredes. **P-cadherin is co-expressed with CD44 and CD49f and mediates stem cell properties in basal-like breast cancer.** *Stem Cells* 30:854-864, 2012.

III. André Filipe Vieira, Ana Sofia Ribeiro, Maria Rita Dionísio, Bárbara Sousa, Ana Rita Nobre, André Albergaria, Angélica Santiago-Gómez, Fernando Schmitt, Robert B. Clarke, Joana Paredes. **P-cadherin intracellular signalling is dependent on α6β4 integrin activation to induce breast cancer stem cell and invasive properties.** 2012 (*submitted to publication*).

Adicionalmente, o autor desta dissertação contribuiu para as seguintes publicações que são artigos de revisão e que incluem partes da introdução desta tese:

IV. André Albergaria, Ana Sofia Ribeiro, André Filipe Vieira, Bárbara Sousa, Ana Rita Nobre, Raquel Seruca, Fernando Schmitt, Joana Paredes. **P-cadherin role in normal breast development and cancer** *International Journal of Development Biology* 55:811-22, 2011.

V. Fernando Schmitt, Sara Ricardo, André Filipe Vieira, Maria Rita Dionísio, Joana Paredes. **Cancer stem cell markers in breast neoplasias: their relevance and distribution in distinct molecular subtypes** *Virchows Archives,* 460: 545-53, 2012.

Table of Contents

Abbreviations

ABCG2	ATP-binding cassette, sub-family G, member 2
ABCB1	ATP-binding cassette, sub-family B, member 1
ABCC1	ATP-binding cassette, sub-family C, member 1
AC	adenocarcinoma
ALDH	aldheyde dehydrogenase
APC	allophycocyanin
AKT	protein kinase B (PKB)
BAAA	BODIPY® - aminoacetaldehyde
BCRP1	breast cancer resistance protein 1
BRCA1	breast cancer 1 gene
BRCA2	breast cancer 2 gene
β-ctn	beta catenin
CD	cluster of differentiation
CDH	cadherin gene
Chk1/2	checkpoint kinase 1/2
CI	confidence interval
CMV	cytomegalovirus
cDNA	complementary deoxyribonucleic acid
CSC	cancer stem cell
DCIS	ductal carcinoma *in situ*
DEAB	diethylaminobenzaldehyde
DNA	deoxyribonucleic acid
EEM	ectodermal dysplasia, ectrodactyly, and macular (syndrome) dystrophy
E-cad	epithelial cadherin
ECM	extracellular matrix
EDTA	etilenediaminetetracetic acid
EGFR	epidermal growth factor receptor
EMT	epithelial to mesenchymal transition
EpCAM	epithelial cell adhesion molecule
ER	oestrogen receptor
ERK	extracellular signal-regulated kinase
ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2 gene
ESA	epithelial specific antigen
FACS	fluorescence activated cell sorting

FGF	fibroblast growth factor
FITC	fluorescein isothiocyanate
FAK	focal adhesion kinase
GPI	glycosylphosphatidylinositol
H&E	haematoxilin and eosin
HER-2	human epidermal growth factor receptor-2
HER2-OE	human epidermal growth factor receptor-2 – overexpressing
HJMD	hypotrichosis with juvenile macular dystrophy
HR	hazard ratio
HRP	horseradish peroxidase
IDAC	infiltrating ductal adenocarcinoma
IDC-NOS	invasive ductal carcinoma, not otherwise specified
ID4	inhibitor of DNA binding 4
Ig	immunoglobulin
IHC	immunohistochemistry
ITGA6	alpha-6 integrin gene
ITGB4	beta-4 integrin gene
JMD	juxtamembrane domain
K	cytokeratin
KRAB	kruppel-associated box domain
Lin	lineage marker
LN	lymph node
MFE	Mammosphere forming efficiency
MMP	matrix metalloprotease
MRP1	multidrug resistance-associated protein 1
MUC	mucin
MAPK	mitogen-activated protein kinase
N-cad	neural cadherin
NOD/SCID	non-obese diabetic/severe combined immunodeficiency
OS	overall survival
PAGE	polyacrilamyde gel electrophoresis
P-cad	placental cadherin
p120ctn	p120-catenin
PBS	phosphate buffered saline
PE	phycoerythrin
PgR	progesterone receptor

PI3K	phosphatidylinositol-3-kinase
PROCR	protein C receptor, endothelial
RNA	ribonucleic acid
RNAi	RNA interference
SHFM	split hand/foot malformation
shRNA	short hairpin RNA
siRNA	small interference RNA
SR	steroid receptor
TDLU	terminal ductal-lobular unit
TEB	terminal end bud
TIC	tumour initiating cell
TMA	tissue microarray
TN	triple negative
TRE	tetracycline responsive element
tTS	tetracycline-controlled transcriptional silencer
VEGF	vascular endothelial growth factor
Vim	vimentin

Abstract

Breast cancer is the most prevalent malignancy in the Portuguese population and it is expected that about 1 in 9 women will suffer from this disease at some time in their lifetime. Basal-like breast cancers are particularly aggressive, representing 15-20% of all breast cancer cases, especially affecting young women, frequently in a reproductive age. Except for a few agents still undergoing clinical trials, novel and improved anticancer agents are urgently needed to target this specific molecular subtype. Notably, basal-like breast carcinomas show increased expression of the basal marker P-cadherin. This is a type-I classical cadherin molecule that mediates important tumour promoting effects in breast carcinomas, being an indicator of poor differentiated carcinomas, high histological grade and decreased patient survival.

The *Cancer Stem Cell (CSC) Hypothesis,* states that the recurrence of breast cancer and the lack of therapy effectiveness is due to a small population of cells with stem-like characteristics present within the tumour mass. This population of cells is responsible for intratumoural heterogeneity, owing to the establishment of a cellular hierarchy. Although the cell surface phenotype $CD44^{+}CD24^{-/low}$ is the most successfully employed in the literature to describe CSCs, this is not a universal phenotype and it is urgent to better define these very important population of cells, as well as to characterize the distribution of different CSC markers within the distinct breast cancer molecular subtypes.

Herewith, we report the analysis of the expression of the CSC markers CD44, CD24, CD49f and Aldehyde Dehydrogenase (ALDH)-1 in breast cancer cells, as well as in a large series of invasive human primary tumours comprising a total of 466 breast cancer cases. Our data revealed that several markers are needed, as different phenotypes are able to detect cells with distinct levels of differentiation. The CSC phenotype $CD44^{+}CD24^{-/low}$ and ALDH1 were significantly increased in the basal-like molecular subtype of human breast carcinomas. Using flow cytometry analysis in a panel of breast cancer cell lines, we showed that basal cells with a mesenchymal phenotype are enriched in the CSC phenotype $CD44^{+}CD24^{-/low}$ whereas basal cells with an epithelial phenotype are mainly $CD44^{+}CD24^{+}$; and luminal and HER-2 overexpressing (HER-2 OE) cell lines are enriched for the $CD44^{-/low}CD24^{+}$ phenotype. Furthermore, we found that the phenotype $CD49f^{+}CD24^{+}$ was associated with the basal epithelial cell lines, whereas the phenotype $CD49f^{-}CD24^{+}$ was associated with luminal and HER-2 OE cells. The activity of ALDH1 was mainly found in HER-2 OE and basal/epithelial breast cancer cell lines.

We report that P-cadherin is directly associated with the expression of the breast stem cell markers CD44, CD49f and the activity of ALDH1 in the series of breast cancer cell lines. The expression of these stem cell markers was also found to be associated with P-cadherin and with the basal-like subtype. Moreover, cell populations enriched for P-cadherin expression by Fluorescence Activated Cell Sorting (FACS) comprised increased *in vitro* self-renewal capacity (measured by the mammosphere assay), increased ability to grow colonies in tridimensional matrigel cultures, as well as greater tumourigenicity in immune compromised mice (athymic nude). Using genetically manipulated breast cells, we showed that the over-expression of P-cadherin increased the mammosphere forming efficiency. Conversely, when we inhibited P-cadherin using siRNA transient knock-down a decrease in the capacity to form mammospheres was observed (*Vs.* control cells). Additionally, P-cadherin expression was shown to confer resistance to X-ray induced cell death, sustaining a role for this molecule in another stem cell property. An association was found between P-cadherin and the stem/progenitor-like phenotype: $CD44^+CD24^+$, as well as with the luminal progenitor phenotype of the normal breast differentiation hierarchy: $CD49f^+CD24^+$.

In this work, we have also focused our attention in the extracellular matrix (ECM) components and their receptors, integrins, that constitute major players in the maintenance of the stem cell niche. Given the highly aggressive behaviour of P-cadherin in breast cancer, it is important to clarify the downstream pathways that are triggered by this adhesion molecule and which mechanisms can be targeted for cancer therapy. Thus, using breast cancer cell lines we showed that P-cadherin is essential for the adhesion of cells to extracellular matrix substrates, such as laminin, vitronectin and fibronectin. The $\alpha6\beta4$ integrin heterodimer was implicated in the downstream signalling of P-cadherin in response to laminin, as well as in the stem cell activity of cancer cells (measured by the mammosphere assay). Importantly, the activation of FAK and Src signalling was dependent on P-cadherin expression. We showed for the first time that P-cadherin and integrin receptors do not signal alone and they have a central role in the oncogenic response.

In summary, the results generated in this research work indicated that in human breast cancer tissue, the basal-like subtype is enriched in the CSC phenotype $CD44^+CD24^{-/low}$ and ALDH1 expression, whereas in human cell lines the basal-like subtype is associated with ALDH1 activity and the luminal progenitor phenotype $CD44^+CD49f^+$. We demonstrated for the first time that the basal marker P-cadherin mediates stem cell properties in basal-like breast cancer, uncovering a novel oncogenic signalling pathway in which P-cadherin cross-talks with $\alpha6\beta4$ integrin. Therefore, P-

cadherin could be explored as a marker to better define the CSC phenotype of basal-like breast carcinomas, as well as the cell-of-origin for this malignancy.

Resumo

O cancro de mama representa a doença maligna mais prevalente na população portuguesa, prevendo-se que 1 em cada 9 mulheres irá sofrer desta doença no decurso da sua vida. O carcinoma de mama do tipo basal é particularmente agressivo, representando 15-20% dos cancros de mama, afectando em especial mulheres jovens, frequentemente em idade reprodutiva. Com exceção de alguns agentes ainda em ensaios clínicos, a pesquisa de novos e melhores agentes anti-tumorais dirigidos a este subtipo particular de cancro de mama torna-se essencial. Os carcinomas do subtipo basal apresentam um aumento de expressão do marcador basal P-caderina. Esta é uma molécula de adesão clássica do tipo-I, implicada em características agressivas no cancro de mama, sendo um importante indicador de indiferenciação, elevado grau histológico, e mau prognóstico.

Segundo o *Modelo das Células Estaminais Tumorais*, a recidiva de cancro de mama e a resistência aos agentes anti-tumorais devem-se à presença de uma pequena população de células com propriedades semelhantes às células estaminais normais, que se encontram na massa tumoral. Esta população de células é responsável pela heterogeneidade intratumoral devido ao estabelecimento de uma hierarquia de diferenciação, semelhante à hierarquia que se encontra no tecido normal. Apesar do fenótipo caracterizado pelos marcadores de superfície $CD44^{+}CD24^{-/low}$ ser o mais usado para descrever a população de células tumorais estaminais, este não se trata de um fenótipo universal e torna-se urgente definir especificamente esta população de células tumorais, assim como caracterizar a distribuição dos diferentes marcadores de células estaminais tumorais nos diferentes subtipos moleculares de cancro de mama.

Neste trabalho, reportamos a análise da expressão dos marcadores de células estaminais CD44, CD24, CD49f e Aldeído Desidrogenase (ALDH)-1 em linhas celulares de cancro de mama e numa série de carcinomas primários invasivos, contendo um total de 466 casos. Os nosso dados indicam que vários marcadores estaminais são necessários, pois os diferentes fenótipos são capazes de detetar células com diferentes níveis de diferenciação. O fenótipo das células estaminais tumorais $CD44^{+}CD24^{-/low}$ e o marcador ALDH1 encontram-se significativamente aumentados no subtipo molecular basal. Aplicando a técnica de citometria de fluxo numa série de linhas celulares de mama, mostramos que as células basais com um fenótipo mesenquimal se encontram enriquecidas no fenótipo $CD44^{+}CD24^{-/low}$, enquanto que as células basais com um fenótipo epithelial são $CD44^{+}CD24^{+}$, e as céluas luminais e com sobre-expressão de HER-2 estão enriquecidas no fenótipo $CD44^{-/low}CD24^{+}$. Adicionalmente, descobrimos que

o fenótipo $CD49f^{+}CD24^{+}$ está associado às linhas basais epiteliais, enquanto que o fenótipo $CD49f^{-}CD24^{+}$ se encontra relacionado com as células luminais e com sobre-expressão de HER-2.

Mostramos que a P-caderina está diretamente associada com a expressão dos marcadores estaminais CD44, CD49f e com a atividade da enzima ALDH em linhas celulares de cancro de mama. A expressão destes marcadores estaminais também surgiu associada à P-caderina e ao subtipo molecular basal na série de carcinomas de mama invasivos. Além disso, populações celulares enriquecidas para a expressão de P-caderina, por citometria de fluxo, apresentam maior capacidade de auto-renovação *in vitro* (medida pelo ensaio das mamosferas), maior capacidade clonogénica em culturas tridimensionais contendo matrigel, assim como maior tumorigenicidade em ratinhos imunocomprometidos (atímicos, modelo *nude*). Usando células geneticamente manipuladas, mostramos que a sobre-expressão de P-caderina aumenta a eficiência de formação de mamosferas. Por outro lado, a inibição transiente de P-caderina (usando *small interfering* RNA) leva a uma diminuição da atividade estaminal das células. Adicionalmente, a expressão da P-caderina confere resistência à morte celular induzida por raio-X, apoiando assim mais uma característica estaminal. Encontramos ainda uma associação marcante entre a P-caderina e o fenótipo estaminal/progenitor $CD44^{+}CD24^{+}$, assim como com o fenótipo progenitor luminal $CD49f^{+}CD24^{+}$.

Neste trabalho, focamos a nossa atenção no estudo dos componentes da matriz extracelular (ECM) e nos seus recetores, as integrinas, que constituem peças fundamentais na manutenção do nicho das células estaminais. Dada a importância da P-caderina no comportamento agressivo do cancro de mama, torna-se importante esclarecer as vias de sinalização desencadeadas por esta molécula de adesão e quais os mecanismos que poderão funcionar com alvo terapêutico. Usando linhas celulares de cancro de mama do tipo basal, mostramos que a P-caderina é essencial para a adesão das células aos substratos da ECM laminina, vitronectina e fibronectina. A integrina $\alpha 6\beta 4$ está implicada na via de sinalização desencadeada pela P-caderina em reposta à laminina, assim como na atividade estaminal das células tumorais. A ativação das tirosinas cinases FAK e Src em resposta à laminina é dependente da expressão de P-caderina.

Em suma, os resultados gerados neste estudo mostram que em carcinomas de mama primários, o subtipo basal está enriquecido no fenótipo da célula estaminal tumoral $CD44^{+}CD24^{-/low}$ e na expressão do marcador ALDH1. Nas linhas celulares, o fenótipo basal surge associado à atividade da enzima ALDH1 e ao fenótipo da célula progenitora luminal $CD44^{+}CD49f^{+}$. Demonstramos ainda, pela primeira vez, que a P-caderina medeia propriedades estaminais nos carcinomas de mama do subtipo basal, descobrindo uma

nova via de sinalização oncogénica, em que a P-caderina e a integrina $\alpha 6\beta 4$ interatuam. A P-caderina poderá ser explorada como um marcador que permitirá definir/otimizar o fenótipo da célula estaminal tumoral de carcinomas de mama basais, assim como a célula-de-origem para este subtipo de cancro de mama.

Thesis Outline

In Chapter I, a general introduction presents the state of the art in the field of breast development and cancer, with a special focus in the aggressive basal-like subtype of breast carcinoma. The hierarchical model of tumour progression, based in the Cancer Stem Cell hypothesis is presented. The relevance of cadherin molecules, particularly P-cadherin, in the cancer context, as well as its significance in stem cell biology is also described. A brief depiction of the importance of the tumour microenvironment is portrayed, with a particular focus in integrin signalling. This introduction section focuses the main topics related with the research data presented in the following chapters.

In Chapter II, the aims of the thesis are defined, in accordance with the state of the art.

In Chapter III, a description of the materials and the detailed methodology applied to perform the work leading to this thesis dissertation is specified.

Chapters IV, V and VI describe all the experimental work that gave rise to the original data presented in this thesis, already accepted or submitted to publication in international peer reviewed journals.

A general analysis and an integrated view of the results is presented in the General Discussion section of this thesis, with the conclusion remarks (Chapter VII) followed by Future Perspectives (Chapter VIII).

The Appendix 1 section describes the establishment of a human basal-like cell model stably expressing shRNA for the *CDH3*/P-cadherin gene, inducible with doxycycline (DOX). The publications that resulted from this thesis are included in the Appendix 2 section.

GENERAL INTRODUCTION

CHAPTER I

1. Breast Cancer

1.1 Epidemiology, risk factors and molecular classification

In western countries, breast cancer is the most frequent malignancy in women. Not surprisingly, this type of cancer is the most prevalent in the adult Portuguese female population. The latest Globocan estimates (in reference to the year 2008) indicate that in Portugal 5333 new cases emerge every year, with an incidence rate of 49.9 cases per 100.000 people. Mortality rates are in the order of 14.4 cases per 100.000 people, corresponding to 4.2 deaths daily in the whole Portuguese population (1537 deaths/ year) (**Figure 1**) (Ferlay *et al.*, 2010 and Liga Portuguesa contra o Cancro). This reality has turned breast cancer into one of the main research interests in the national cancer research community.

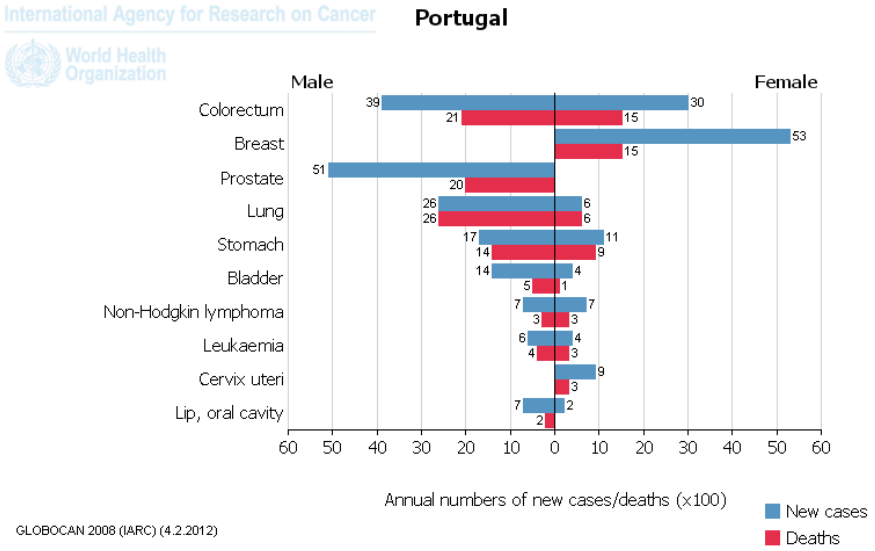


Figure 1 - GLOBOCAN estimates presented for 2008 regarding the annual numbers of incidence and mortality of the most common types of cancer for both sexes in Portugal (GLOBOCAN 2008 v1.2, Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 10, (Ferlay *et al.*, 2010)).

This malignancy has a multifactorial etiology, since a single causative agent is not sufficient to induce breast cancer development. The interaction between environmental factors and the individual genetic profile dictate the susceptibility for the development of this disease. Notably, breast cancer is more common among Caucasians living in the colder climates and the more highly industrialized countries of the western hemisphere (Macdonald *et al.*, 2004). One of the most relevant risk factors for breast cancer is age: older women are more likely to develop this type of malignancy. Certain breast alterations, such as atypical mammary hyperplasia or preneoplastic lesions can represent important risk factors. The previous history of developing a neoplastic breast lesion, as well as the

family history, such as inherited germline mutations in the breast cancer susceptibility genes *BRCA1* and *BRCA2* are also major predetermining factors. The earlier age at menarche (before 12 years of age), the late menopause (after the age of 55 years) and postmenopausal obesity are significant risk factors, since they increase the breast exposure to elevated levels of oestrogen. Other factors such as a late pregnancy, sedentary life style, alcohol consumption, the use of oral contraceptives and hormonal replacement therapy also contribute to increased breast cancer risk (Kufe *et al.*, 2010).

Breast cancers vary greatly in clinical behavior, outcome, morphology, and biologic characteristics. Histologically, invasive breast cancers are categorized into at least 18 different subtypes, based on growth patterns and cytological characteristics (Tavassoli & Devilee, 2003). This heterogeneity led to the development of a new classification system that more accurately relates the molecular characteristics, prognosis, and treatment options for this malignancy. Based in gene expression profiles, four main molecular subtypes of breast cancers were described: Luminal A, Luminal B, human epidermal growth factor receptor (HER)-2 overexpressing cancers and triple negative carcinomas (that comprise mainly basal-like, normal breast-like and claudin-low carcinomas) (**Figure 2**) (Perou *et al.*, 1999, Perou *et al.*, 2000, Sorlie *et al.*, 2001, Sorlie *et al.*, 2003, Herschkowitz *et al.*, 2007, Hennessy *et al.*, 2009).

Luminal breast cancers represent the bulk of breast cancers (60-75%); these express estrogen receptor (ER)-α and are generally of good prognosis, because directed therapy may be applied to inhibit the mitogenic paracrine effects of this receptor. Current therapies are based on the administration of tamoxifen or fulvestrant (ER antagonists) or aromatase inhibitors (oestrogen synthesis inhibitors), such as anastrazole or letrozole. Within this molecular group, Luminal A carcinomas are distinguished from Luminal B, since the latter cancer subtype usually expresses lower levels of ER-related genes, high levels of Ki-67 (proliferation), may present HER-2 overexpression and show a worse prognosis.

HER-2 overexpressing breast carcinomas represent 18-20% of breast cancers (Rakha & Reis-Filho, 2009) and present high levels of a member of the epidermal growth factor receptor family, the HER-2 receptor, and the neighboring genes found in the chromosome region 17q, frequently amplified in these cancers. Though of poor prognosis, therapeutic strategies are used in an attempt to target HER-2 receptor, either with the monoclonal antibody trastuzumab (Herceptin), or tyrosine kinase receptor inhibitors (e.g., lapatinib).

Triple negative carcinomas denote 10-17% of all breast carcinomas (Reis-Filho & Tutt, 2008). They represent a heterogeneous group of malignancies that are ER-negative, progesterone receptor (PgR)-negative and HER-2-negative. For this reason, this type of

breast cancer is usually insensitive to endocrine and anti-HER-2 therapies. The triple negative normal breast-like subgroup is typified by the high expression of genes characteristic of basal epithelial cells and adipose cells, and the low expression of genes characteristic of luminal epithelial cells. These breast cancers have usually a better prognosis and cluster together with benign lesions of the breast (Sorlie *et al.*, 2003). On the other hand, basal-like breast cancer is a triple negative carcinoma that includes tumors characterized by an expression signature similar to that of the basal cells of the breast, such as cytokeratins (CK) 5, 14 and 17, laminin-5, $\alpha 6\beta 4$-integrin, caveolin-1 and 2, P-cadherin, and calponin (Lakhani *et al.*, 1998, Lakhani *et al.*, 2005, Lakhani *et al.*, 2002). Finally, Claudin-low tumors are a recently described molecular subgroup within the triple negative tumors, defined by loss of a cluster of genes involved in cell-cell adhesion, namely claudins 3, 4, and 7, occludin and E-cadherin. The gene set used also describes low expression of luminal genes, inconsistent expression of basal-like genes, and high expression of lymphocyte and endothelial cell markers (Hennessy *et al.*, 2009, Herschkowitz *et al.*, 2007).

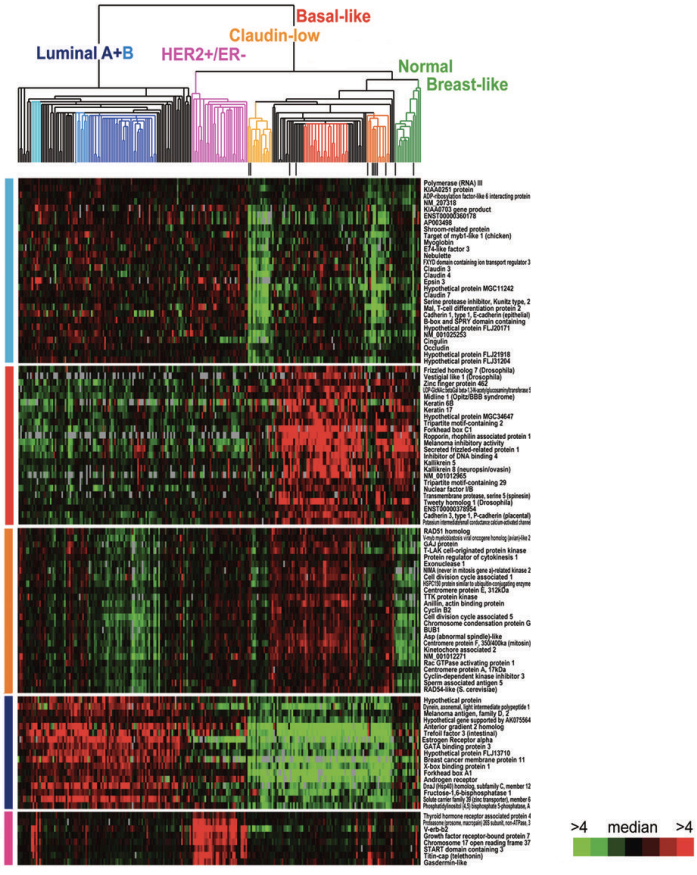


Figure 2 – Hierarchical clustering of 205 breast tissue samples into five subgroups according to the intrinsic gene set defined by Sorlie *et al.,* 2003 and Herschkowitz *et al.,* 2007. Dark blue – luminal A subtype, light blue – Luminal B subtype, pink – HER-2 overexpressing subtype, red – basal subtype, orange – claudin low subtype, green – normal breast like subtype. Branches corresponding to tumours with low correlation to any subtype are shown in black. Adapted from Hennessy BT, *Cancer Research,* 2009 (Hennessy *et al.*, 2009).

1.2. The breast cancer basal-like molecular subtype

Basal-like breast tumors account for 15% of all breast cancer cases, occur at a slightly younger age and have an extremely bad prognosis, presenting the shortest survival of all breast-cancer subtypes. This observation is in part due to early relapses within the first five years after diagnosis (Da Silva *et al.*, 2007, Korsching *et al.*, 2008, Sorlie *et al.*, 2001, Yehiely *et al.*, 2006). It mainly affects african-american, hispanic and obese women (Rakha & Reis-Filho, 2009). In addition, the basal tumour phenotype generates therapeutic dilemmas, since anti-oestrogens and trastuzumab are not adequate options for such patients.

Basal-like breast cancers encompass grade III invasive ductal carcinomas of no special type, (atypical) medullary carcinomas, metaplastic breast cancers and other rare histological subtypes, such as adenoid cystic, acinic cell carcinoma, pleomorphic lobular and secretory breast cancers (Rakha & Reis-Filho, 2009, Weigelt *et al.*, 2010). Pushing and non-infiltrative borders of invasion are observed, with large zones of geographic or comedo-type necrosis, stromal lymphocytic infiltrates, scant stromal content, lack of tubule formations, marked cellular pleomorphism, high nuclear–cytoplasmic ratios, vesicular chromatin, prominent nucleoli, high mitotic indices, and frequent apoptotic cells (Rakha & Reis-Filho, 2009). Basal-like cancer is associated with an aggressive clinical history, development of recurrence, shorter survival, and a specific pattern of hematogenous distant metastasis (Patanaphan *et al.*, 1988).

Using immunohystochemistry, basal-like cancers can be identified by a panel that comprises 4 markers (lack of ER and HER-2, and expression of EGFR and/or CK5/6) with 100% specificity and 76% sensitivity (Nielsen *et al.*, 2004). Basal-like cancers are also usually PgR-negative and may present a CK14-positive staining. P-cadherin is also used as an adjunct marker to identify this type of carcinomas (Matos *et al.*, 2005, Paredes *et al.*, 2007b, Sousa *et al.*, 2010), since its expression is highly associated with this molecular subgroup. Cyclin E, Ki-67, fascin, Sox-2, moesin, vimentin, nestin and laminin are also frequently expressed (Rakha & Reis-Filho, 2009). Basal-like tumors are predominantly negative for BRCA1, cyclin D1, p27 and MUC1 (Foulkes, 2004, Nielsen *et al.*, 2004, Potemski *et al.*, 2005), and may express luminal keratins (CK8/18), albeit at lower levels than those found in luminal cancers (Honrado *et al.*, 2007).

Familial breast tumors account for 5-10% of all breast tumors and about half of these cases are related to *BRCA1* gene mutations (Hwang-Verslues *et al.*, 2008). Breast tumors arising in *BRCA1* germline mutation carriers often exhibit a basal-like phenotype and frequently these tumors also harbor *p53* mutations (Foulkes *et al.*, 2003, Lakhani *et*

al., 2005, Turner & Reis-Filho, 2006). Although somatic *BRCA1* gene mutations are rare in sporadic basal-like cancers, there is evidence to suggest that BRCA1 mRNA (Turner *et al.*, 2007) and protein levels (Lambie *et al.*, 2003, Rakha *et al.*, 2008) are downregulated, either by epigenetic silencing of the *BRCA1* gene promoter (i.e., *BRCA1* gene promoter methylation) (Turner *et al.*, 2007) or by up-regulation of ID4, a *BRCA1*-negative regulator (Beger *et al.*, 2001). BRCA1 protein is involved in the DNA repair mechanism by homologous recombination and several lines of evidence suggest that cancer cells harboring this defective DNA repair mechanism show an exquisite sensitivity to cross-linking agents, such as carboplatin, and to inhibitors of the PARP enzyme (Edwards *et al.*, 2008, Farmer *et al.*, 2005). However, agents that effectively target basal-like breast cancers are needed.

Currently, chemotherapy and radiotherapy for systemic and local control remains the mainstay to treat basal-like cancer, although a high proportion of patients die in a shorter time frame due to metastatic disease to the lungs and brain (Da Silva *et al.*, 2007, Fulford *et al.*, 2007, Rakha *et al.*, 2006, Reis-Filho & Tutt, 2008, Yehiely *et al.*, 2006). The mechanisms involved in this aggressive behavior are under study. Based on the strong association with poor relapse-free and overall survival, numerous studies, including ours, attempt to further define the pathology, biology, clinical features and, essentially, therapeutic options for basal-like tumors. This is particularly challenging, since not much is known about the genes or the cells that are responsible for the aggressive phenotype of these tumors, which is a major barrier to develop target therapies. Additionally, data in the literature supports the idea that the poor outcome phenotype of basal-like breast cancers may reflect its cellular origin and a cancer stem cell component. One of the findings supporting this hypothesis is that BRCA1 regulates breast stem cell function and fate (Foulkes, 2004), as well as transcriptionally regulates genes associated with basal-like phenotype in breast cancer, namely P-cadherin, CK5 and CK17 (Gorski, 2009).

2. Normal breast development and mammary stem cells

The mammary gland in humans and other mammals is a dynamic organ that undergoes significant developmental changes during embryonic development, puberty, pregnancy, lactation and involution. The mature gland consists of a branching ductal-lobular system. The lobules of the human breast are organized into 15-20 lobes, which are drained by collecting ducts that converge at the nipple in a radial arrangement (Osborne, 2000) (**Figure 3A**). Each lobule in turn is made up of acini (also called alveoli) that form the functional secretory units of the mammary gland, the terminal duct lobular

units (TDLUs). The acini and ducts have a central lumen and are lined by two cell layers, an inner layer of epithelial cells and an outer layer of myoepithelial cells. Surrounding this structure is a basement membrane separating the epithelium from the stroma (Osborne, 2000) (**Figure 3B and 3C**). The mammary gland is embedded in stroma, which is composed of mesenchymal cells, such as fibroblasts, adipocytes, endothelial cells, immune cells, and extracellular matrix components.

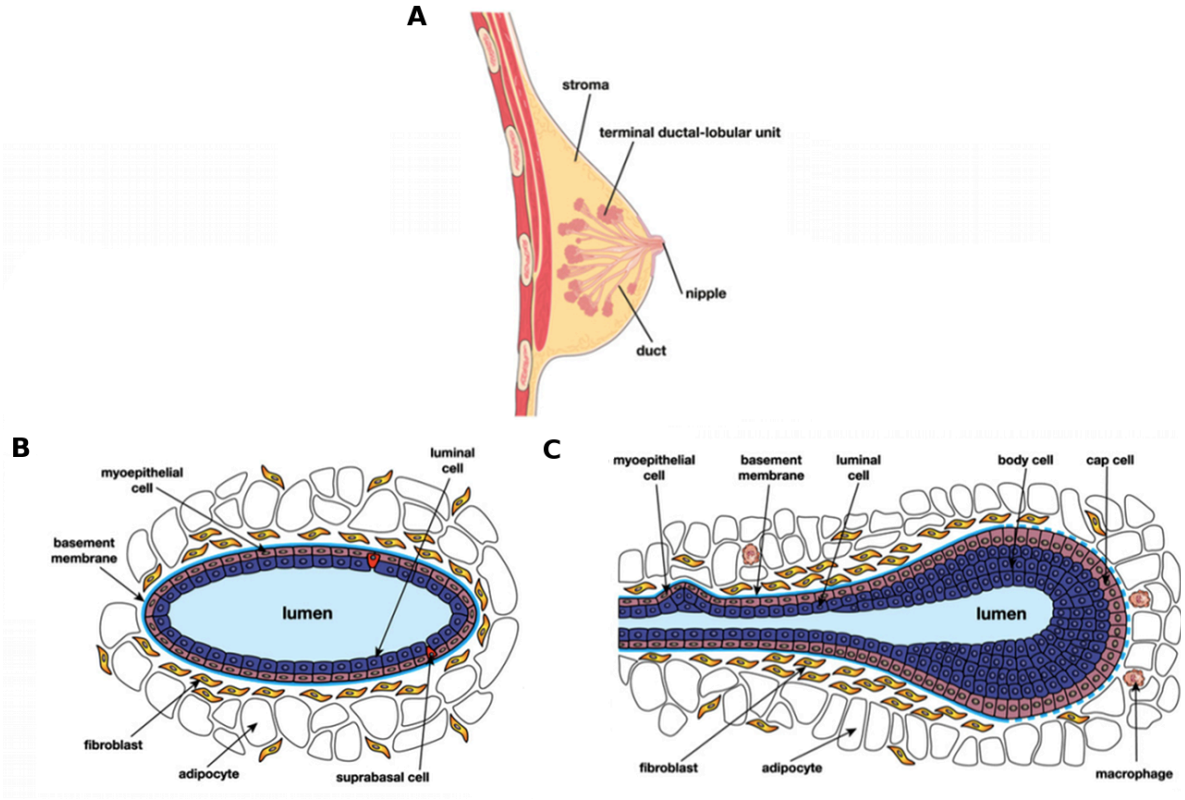


Figure 3 – Schematic representation of the human mammary gland (comprising 15-20 lobes) (A); Representation of the mammary duct (B) and a terminal end bud (C). A suprabasal region contains cells (represented in red) that sit on the myoepithelial layer, but do not reach the lumen. This is most likely the region where the mammary stem cells are located. Adapted from Visvader JE, *Genes and Development*, 2009 (Visvader, 2009).

The mammary epithelial tissue develops from the embryonic ectoderm, forming a primitive branched ductal system (Hennighausen & Robinson, 2001, Robinson *et al.*, 1999). The rudiments of the mammary gland remain quiescent until puberty. At this time, the increase in steroid hormones, oestrogen (Robinson *et al.*, 1999) and progesterone, stimulate significant ductal growth, which is driven by specialized structures at the tips of the elongating ducts, the terminal end buds (TEB) (**Figure 3C**). This region is enriched in stem cell activity and consists of two morphologically distinct cell types, an inner layer of body cells and an outer layer of cap cells, which give rise to luminal and basal cell layers

respectively (Ball, 1998, Sapino *et al.*, 1993). Once the mammary fat pad has been filled with glandular tissue, the TEBs regress and are converted to terminal ducts. During pregnancy, a surge of hormones results in major structural changes in the mammary gland. Proliferation and maturation of the side branches occurs to form alveoli, which contain the alveolar cells of TDLUs, responsible for milk production (Hennighausen & Robinson, 2001). In humans, at least some TDLUs are already present prior to pregnancy, although they are probably not as complex as during pregnancy (Molyneux *et al.*, 2007). The termination of breast-feeding initiates significant structural remodelling known as involution, resulting in apoptotic cell death and collapse of the alveoli. The mammary gland returns to a predominantly ductal structure resembling the mature virgin gland. The cycles of proliferation, differentiation and involution that occur with every pregnancy suggest that these development changes are supported by adult multipotent mammary stem cells.

In the last decade, it has been firmly established that, like most tissues, if not all, the mammary gland has also a hierarchical organization, similar to the hematopoietic system (Visvader & Lindeman, 2006) (**Figure 4**). At the top of this cell-hierarchy there is a small population of cells with self-renewal capability, named stem cells, responsible for generating and maintaining the tissue architecture and permitting tissue remodeling and repairing.

Stem cells in the breast are scarce and although they have been purified and isolated using different techniques, they have not yet been fully characterized. A number of different cell surface markers have been described to define stem and progenitor mammary cells from both human and mice origin. It is important to say that some biomarkers that are useful for human stem cell characterization are not effective in the mouse model. In 2001, using Fluorescence Activated Cell Sorting (FACS), Stingl *et al* fractionated the human mammary epithelial gland using the markers EpCAM and CD49f and the isolated subpopulations were characterized for lineage restricted *in vitro* colony forming ability (Stingl *et al.*, 2001). Since then several groups have used several markers to isolate and characterize *in vitro* colony forming ability as well as *in vivo* mammary repopulating capacity. Using FACS sorting of cell populations, Stingl *et al.*, Eirew *et al.* and Lim *et al.* showed that mammary stem cells of the human breast were enriched in the subpopulation $EpCAM^{-/low}CD49f^{+}$, whereas the phenotype $EpCAM^{+}CD49f^{+}$ was considered to have luminal progenitor features (Eirew *et al.*, 2008, Lim *et al.*, 2009, Stingl *et al.*, 2001). Villadsen *et al.* and Raouf *et al*. propose that $EpCAM^{+}CD49f^{+}$ ($CD10^{+}Thy1^{+}$) cells preferentially have a bipotent differentiation capacity (Raouf *et al.*, 2008, Villadsen *et al.*, 2007). Recently, Keller *et al.* fractionated the human breast tissue using the same

markers and demonstrated that, within the CD49f^{+} fraction, both EpCAM^{high} (CD10^{-}) and EpCAM^{low} (CD10^{+}) appear to have bipotent differentiation capacity, although the later phenotype was more associated with basal-like characteristics (Keller *et al.*, 2011). Other subsets of cells were also found to be enriched in mammary stem cells, expressing the following sets of markers: SSEA-4^{+}/K5^{+}/K6a^{+}/K15^{+}/Bcl-2^{+} (Villadsen *et al.*, 2007), K19^{+}/K14^{+} (Villadsen *et al.*, 2007), ALDH1^{+} (Ginestier *et al.*, 2007), CD44^{+}/CD24^{-} /PROCR^{+}/ER^{-}/MUC1^{-}/HER-2^{-} (Mani *et al.*, 2008, Shipitsin *et al.*, 2007).

It has also been shown that putative mammary stem cells in the mouse typically show the expression of EGFR and high molecular weight cytokeratins, weak or almost absent expression of CK18 and lack the expression of ER and HER-2 (Korsching *et al.*, 2008, Asselin-Labat *et al.*, 2006). Human breast stem cells are also likely negative for ER expression (O'Brien *et al.*, 2011, Raouf *et al.*, 2008, Shipitsin *et al.*, 2007).

All these studies are consistent with the idea that mammary stem cells reside within the basal compartment of the mammary gland. However, no definitive phenotype is described for the stem cell at the top of the hierarchy or the multipotent early and late progenitor cells. It is important to refer that increasing data show that the breast hierarchy is not static. This means that progenitor cells or even mature cells can dedifferentiate and spontaneously move up the hierarchy (Chaffer *et al.*, 2011, Gupta *et al.*, 2011). In this scenario, stem cells can originate non-stem cells and vice-versa.

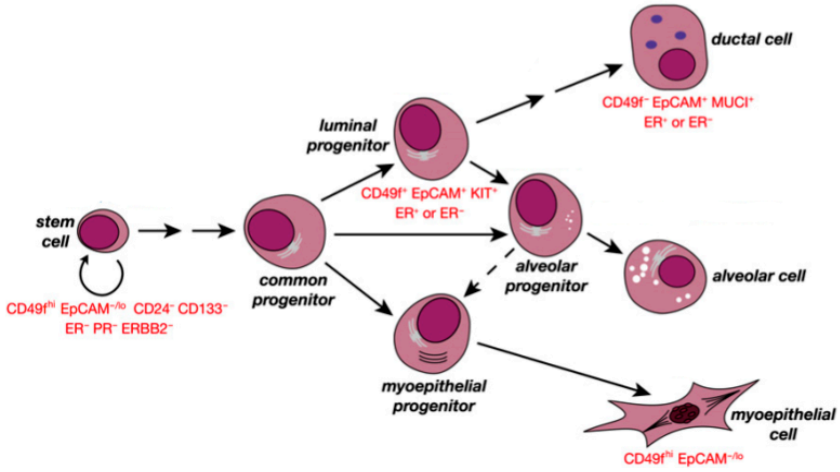


Figure 4 – Model of the differentiation hierarchy within human mammary epithelium. Primary cell surface markers used in the isolation of human epithelial cell subsets are shown. The common progenitor is also referred to as bipotent progenitor cell. Adapted from Visvader JE, *Genes and Development*, 2009 (Visvader, 2009).

Nothing is known about the location of mammary stem cells and their niche in the mature gland. A significant proportion of the total cells present in the mammary epithelium

of the rat, approximately 3%, were proposed to be stem/early progenitor cells. These small cells reside in a suprabasal position, between luminal and myoepithelial cell layers, lacking organelles and with a pale cytoplasm (Chepko & Dickson, 2003). Terminal end buds within the pubertal gland and the ductal branch points in the adult gland may be sites of residence of stem cells (Visvader & Lindeman, 2006). It is possible that adhesion molecules, which allow the interaction between these stem cells and neighboring cells, as well as with the extracellular matrix, may play an important role in maintaining the stem cell niche.

3. The Cancer Stem Cell Hypothesis

The cellular origin of most solid tumors is largely unknown, but it has been speculated that different subtypes of cancer reflect distinct cells of origin at the time of tumor initiation. Alternatively, the tumor phenotype may reflect the acquisition of specific genetic mutations and epigenetic alterations, that influence the intercation between tumor cells and their microenvironment (stroma, inflammatory cells and recruited vasculature) and have a profound influence on the tumorigenic process (Visvader, 2011). Although the definition of a cell of origin or a driver mutation for the different subtypes of breast cancer has been proven very difficult, there is evidence that tumors contain a minority of cells with normal stem cell-like characteristics that may be responsible for their propagation. Normal adult stem cells have self-renewal and differentiation capacity into several lineages; tumor cells have high proliferative capacity, phenotypic plasticity and aberrant differentiation (Reya *et al.*, 2001). These similarities between stem cells and tumor cells have given rise to the hypothesis that breast tumors should arise from undifferentiated stem or progenitor cells. On the other hand, alternatively, it is believed that cancer cells may result from progressive de-differentiation during tumor development (Beachy *et al.*, 2004, Reya *et al.*, 2001, Stingl & Caldas, 2007).

In this way, whatever the cell of origin for breast tumors may be, two hypothesis attempt to explain the observed heterogeneity in solid tumors, namely in breast cancer.
In the traditional hypothesis, or the *clonal evolution model* (**Figure 5**) (Campbell & Polyak, 2007, Nowell, 1976), any cell in the mammary gland can receive an oncogenic hit and mutant tumor cells with growth advantage are selected and expanded. The cells in the dominant population have a similar potential for regenerating tumor growth. In this model, most cancer cells can proliferate extensively and have tumuorigenic potential. Intratumoral heterogeneity is due to the accumulation of additional mutations during tumor growth. The acquisition of genetic events underpins this model, but epigenetic differences and microenvironmental changes are also likely to have an important role.

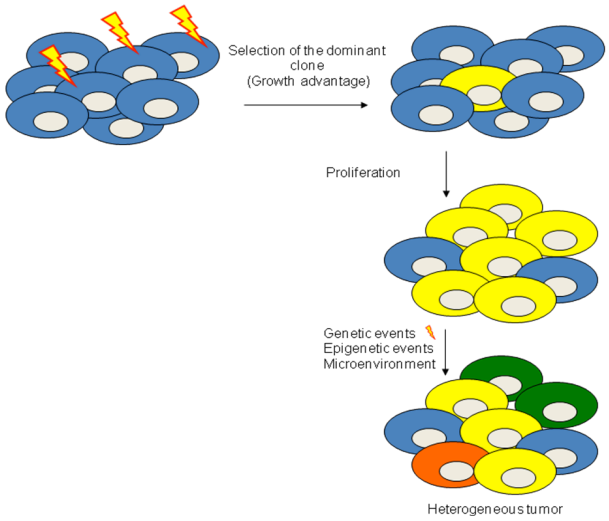


Figure 5 - The clonal evolution model of tumor progression. All undifferentiated cells have similar tumorigenic capacity (most cells can proliferate extensively and form new tumors). Cellular phenotypes are not stable and can change as the tumor evolves. Existing therapeutic approaches have been based largely in this model.

The most recent model, known as *cancer stem cell hypothesis* or *hierarchical model* (**Figure 6**) (Bonnet & Dick, 1997, Reya *et al.*, 2001), proposes that only a small subpopulation of cancer cells is tumourigenic and has the ability to self renew and generate the diverse cells that comprise the tumor. These cells, named cancer stem cells (CSC), share important properties with normal tissue stem cells, including self-renewal (by symmetric and asymmetric division) and differentiation capacity, albeit aberrant. Markers have been identified that distinguish the tumourigenic from the non-toumourigenic cells and the difference between these cells is epigenetic. Importantly, CSCs do not necessarily originate from the transformation of normal stem cells. The intratumural heterogeneity of human breast cancers is due to the establishment of a hierarchy with the CSC lying at the apex.

Importantly, the CSC model of tumor progression is consistent with the behaviour of some leukemias (Bonnet & Dick, 1997) and certain solid tumors, such as breast cancer (Al-Hajj *et al.*, 2003), brain cancers (Singh *et al.*, 2004), colon cancers (O'Brien *et al.*, 2007) and germ cell cancers (Illmensee & Mintz, 1976) in which cancer cells seem to be clearly organized in a hierarchical manner. However, some other tumors such as melanoma (Quintana *et al.*, 2008) and B-cell lymphoblastic leukaemia (Williams *et al.*, 2007) have a much more high frequency of cancer cells with tumourigenic ability and the clonal evolution model is a better description for these malignancies.

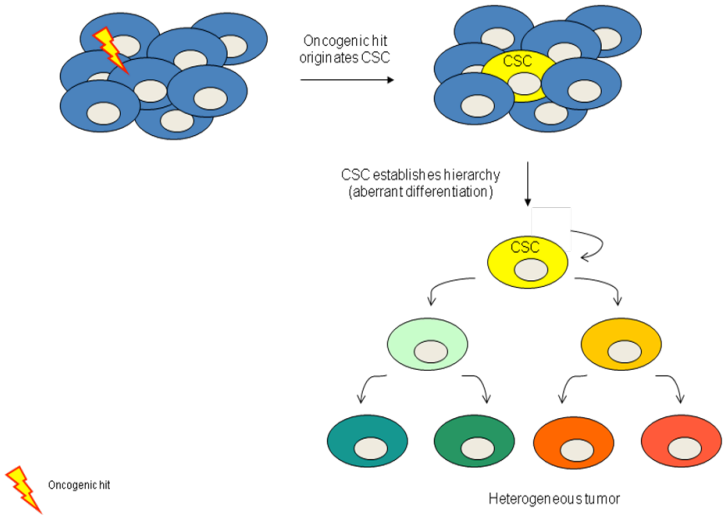


Figure 6 – The cancer stem cell hypothesis or hierarchical model of tumor progression. Only the cancer stem cell (CSC) can propagate a tumor, based on its extensive self-renewal and differentiation capacity. Most cancer cells have limited proliferative potential.

Although the CSC model offers a tempting explanation for the aggressive behaviour of certain tumours, several aspects are not clearly addressed. The unknown origin of the CSC, and the lack of a universal CSC phenotype for solid cancers complicate the establishment of this model. Furthermore, not a single CSC was able to reconstitute the whole solid tumour, only subpopulations of cells. Also, technical issues, like the unfeasibility of testing this model in humans, and the bias of *in vitro* stem cell assays, are major caveats of this model. In fact, the hierarchical model of tumour progression seems to be imperfect and probably cancer stem cells also evolve according to a clonal evolution model (Shackleton *et al.*, 2009). This would mean that the previous two models proposed do not have to be mutually exclusive, and their combination is also plausible.

3.1. The clinical relevance of Cancer Stem Cells

Classical antineoplastic treatments, such as chemotherapy or radiation, can efficiently eradicate the majority of proliferating and genetically unstable malignant cells within neoplastic lesions. However, these regimens frequently fail to eliminate the minor subpopulation of resistant CSCs. Therefore, the concept of the CSC leads to an important consequence in the clinical setting: more effective treatments are needed for cancers and clinicians probably need to focus therapy on the most malignant cells. The cancer stem cell model offers a good explanation for recurrence and metastasis, as well as the resistance of certain tumors to current therapy, indicating that targeting CSCs, in

combination with current therapies is goal to achieve in cancer treatment (Ablett *et al.*, 2012).

The recurrence of breast cancer, often in a secondary site, may reflect the inherent ability of CSCs to survive as circulating tumor cells and to form micrometastasis, remaining quiescent in distant sites for a long period. In this setting the pre-metastatic niche has a major role, since it gathers all the necessary conditions to the survival of the metastatic cells. Importantly, the metastatic behaviour of cancer cells may be associated to the acquisition of a more mesenchymal phenotype, which enables certain stem-like characteristics (Mani *et al.*, 2008, Morel *et al.*, 2008).

Concerning patient's resistance to current therapy, it is important to mention that CSCs possess innate resistance mechanisms against radiation- and chemotherapy-induced cell death, allowing them to survive and cause tumor recurrence.

The mechanisms of intrinsic therapy resistance in normal and malignant mammary stem cells involve enhanced DNA repair mechanisms, which confer resistance of stem cells to DNA-damaging agents (Cairns, 1975, Cairns, 2002, Park & Gerson, 2005, Potten *et al.*, 2002) and increased stem-cell maintenance signalling pathways, such as Wnt/β-ctn and Notch signalling (Wang *et al.*, 2010, Chen *et al.*, 2007). Novel targeted therapies against the DNA damage checkpoint or stem-cell maintenance pathways may sensitize cancer stem cells to radiation or other therapies. Additionally, it was found that many cancer stem cells are not cycling and are in G0 and, thus, are resistant to cell cycle–specific chemotherapy agents (Venezia *et al.*, 2004). Furthermore, it was postulated that during DNA replication, the parental ''immortal'' DNA strand always segregates with the stem cell and not with the differentiating progeny. This prevents the stem cell compartment from accumulating mutations associated with replication (Rambhatla *et al.*, 2005). Stem cells also express higher levels of anti-apoptotic proteins, such as members of the Bcl-2 family and inhibitors of apoptosis, than do the differentiated cells (Wang *et al.*, 2003). Stem cells still express high levels of transporter proteins, such as ABCG2 (BCRP1), ABCB1 (P-glycoprotein 1) and ABCC1 (MRP1), which represent the three principal multidrug resistance genes (Wicha *et al.*, 2006). Although certain regulators of stem cell function have been implicated in cancer pathogenesis (Beachy *et al.*, 2004), a broad characterization of stem cell-associated regulatory networks in tumors, as well as in normal and cancerous stem cells is still lacking.

3.2. Cancer Stem Cells in basal-like breast cancer

In 2003, Michael Clarke´s group described for the first time the phenotype of the breast CSCs. The cancer initiating cells were identified and isolated using cell surface markers as Lineage$^-$ ESA$^+$CD44$^+$CD24$^{-/low}$ in eight out of nine patients. As few as 200 cells with this phenotype were able to form tumours in NOD/SCID immunocompromised mice, whereas 20000 cells with alternate phenotypes failed to form a tumour. The tumourigenic cells gave rise to non-tumourigenic cells and could be serially passaged, proving the differentiation ability and self-renewal potential of these cells (Al-Hajj *et al.*, 2003).

Basal breast tumors are mainly composed of cells expressing the cancer stem cell marker CD44 (Fulford *et al.*, 2007, Polyak, 2007, Shipitsin *et al.*, 2007) and an association between the CD44$^+$/CD24$^-$ phenotype and basal-like breast cancer subtype has been reported (Ali *et al.*, 2011, Honeth *et al.*, 2008, Park *et al.*, 2010), indicating possible stem cell features for this cancer. Moreover, other cell markers associated with stem cell properties were identified in basal-like breast cancers, such as the enzyme aldehyde dehydrogenase-1 (ALDH1), which was correlated with poor prognosis cancers (Ginestier *et al.*, 2007, Park *et al.*, 2010). Notably, the CSC phenotype CD44$^+$CD24$^{-/low}$ is not universal and other phenotypes have been proposed to better describe CSCs present in basal-like mammary tumors. Thus, for mice breast tumors harbouring *BRCA1* deletion, additionally to the CD44+CD24$^{-/low}$ subpopulation, the population defined as CD133$^+$ (Prominin$^+$) correlated with stem cell activity, producing increased spheroids, being significantly more resistant to DNA damaging drugs, and having increased tumourigenic ability (Wright *et al.*, 2008). Furthermore, for the human basal-like breast cancer cell line MDA-MB-231 (which is mainly CD44$^+$CD24$^{-/low}$), the phenotype ESA$^+$PROCR$^+$ could be used to highly enrich for stem/progenitor properties, including increased colony forming ability, *in vitro* regenerative capacity, and increased mesenchymal markers (Hwang-Verslues *et al.*, 2009). Also, in human ER-negative breast cancers, the phenotype CD44$^+$CD49f^{high}CD133/2^{high} defined a population of tumor initiating cells with self-renewal capacity (Meyer *et al.*, 2010).

From an immunohystochemical point of view, basal-like breast cancers share a profile similar to the physiological stem/progenitor cells of the normal female breast, including c-kit, α6-integrin, K5, K14 and prion protein. It is believed that the stem cell at the top of the hierarchy is, like the basal cell of the breast, triple negative (ER$^-$/PgR$^-$/HER2$^-$), EGFR$^+$, p63$^+$ and p21$^-$. Actually, basal-like cancers show a mixed basal/luminal phenotype: they are positive for basal K5 and K14, and luminal K8 and K18. This suggests that a stem cell or a putative bi-potential progenitor cell, that gives rise to both

luminal and myoepithelial cells, may be in the origin of these cancers (Korsching *et al.*, 2008, Yehiely *et al.*, 2006). In fact, recent data from several groups indicate that the cell of origin for basal-like breast cancers may be the luminal progenitor cell (Lim *et al.*, 2009, Molyneux *et al.*, 2010, Proia *et al.*, 2011). In fact, human *BRCA1* gene inactivation, which is often associated with this phenotype of breast cancer, causes a defect in progenitor cell lineage commitment, with an increase in the immature $K14^+/19^+$ cells (Proia *et al.*, 2011), as well as an increase in ALDH1 activity with loss of mature luminal differentiation markers (Liu *et al.*, 2008). Furthermore, although the studies performed are not completely clear, *BRCA1* mutation carriers seem to possess an expansion of the luminal progenitor cell population (Lim *et al.*, 2009, Proia *et al.*, 2011) and this cell population is most likely the target of oncogenic transformation in basal-like breast tumours (Molyneux *et al.*, 2010, Proia *et al.*, 2011).

4. Cadherin molecules in normal and malignant tissues

Cadherins are a large family of molecules that mediate cell-cell adhesion. Cadherins can be subdivided into six subfamilies: the type I and type II classical cadherins, the desmosomal cadherins, the seven-pass transmembrane cadherins, the large cadherins of the fat and dachsous group, and the group of protocadherins (Nollet *et al.*, 2000, Stemmler, 2008). The type I classical cadherins, which include *CDH1*/E-cadherin (epithelial), *CDH2*/N-cadherin (neuronal), *CDH3*/P-cadherin (placental) and *CDH4*/R-cadherin (retinal), constitutes the best characterized subgroup (Nollet *et al.*, 2000). These cadherins are composed of three components: 1) an extracellular portion responsible for calcium-dependent homotypic cadherin-cadherin interaction (which has 5 repeated cadherin domains); 2) a single pass transmembrane domain; and 3) a highly conserved cytoplasmic domain that binds to the intracellular catenins p120-catenin (p120ctn) and β-catenin (β-ctn). Catenins have a dual role, acting as signalling mediators or as adaptor molecules that stabilize the cadherin complex at the membrane and link the cadherin molecule to the actin filaments of the cytoskeleton (Wheelock *et al.*, 2001) (**Figure 7**). Cadherin mediated cell-cell adhesion is accomplished by homophilic interactions between two cadherin molecules at the surface of the respective cells (Cavallaro & Dejana, 2011) and the cadherin-catenin complex constitutes the main building block of the adherens-type junctions. These complexes, not only regulate cell shape and polarity, but also represent a major regulatory mechanism that guides cell fate decisions, influencing cell growth, differentiation, cell motility and survival (Cavallaro & Dejana, 2011).

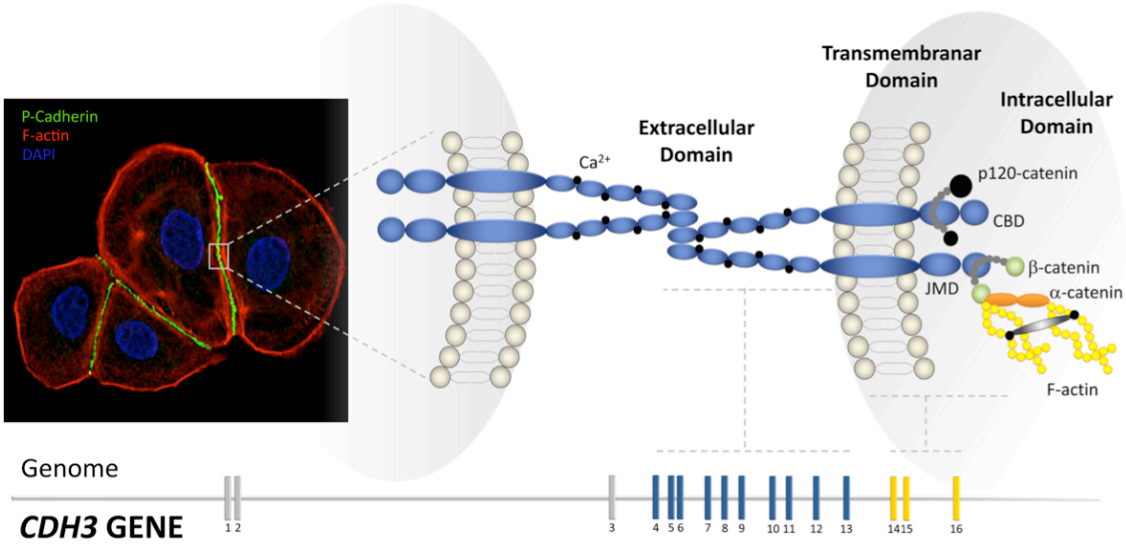


Figure 7 – Representation of the structural components of a classical cadherin adhesive structure. In the figure, P-cadherin (placental) is represented. P-cadherin extracellular domain interacts with the extracellular domain of an adjacent cell. Intracellular catenins bind to the cytoplasmic tail of P-cadherin: p120-catenin binds at the juxtamembrane domain (JMD), whereas β-catenin to the distal catenin binding domain (CBD). α-catenin associates with β-catenin and is directly linked to the actin cytoskeleton. The lower panel shows the genomic structure of the *CDH3*/P-cadherin gene, which is constituted by 16 exons: the extracellular part of P-cadherin is encoded by 10 exons (exons 4-13), whereas the transmembrane and the intracellular domains are determined by the information coded by the last 3 exons (exons 14-16). Adapted from Albergaria A, *et al., International Journal of Development Biology*, 2011 (Albergaria *et al.*, 2011).

The maintenance of normal epithelial cellular architecture is frequently altered during tumour progression. For example, lobular breast carcinoma, in which cells actively invade the mammary stroma, is characterized by E-cadherin loss (De Leeuw *et al.*, 1997). In fact, it is well known that E-cadherin acts as a tumour suppressor, negatively regulating the invasion and metastasis of tumour cells in most, if not all, cancers of epithelial origin (Yilmaz & Christofori, Cavallaro & Christofori, 2004). The loss of E-cadherin function during tumour progression can be caused by various genetic or epigenetic mechanisms. In patients with lobular breast cancer and diffuse gastric cancer, the E-cadherin gene is mutated leading to the expression of a non-functional protein (Guilford *et al.*, 1998). E-cadherin expression can also be downregulated at the transcriptional level, by the repressors Snail (Snai1), Slug (Snai2), Sip1 (Zeb2) and Twist that bind to E2 boxes in the promoter of the *CDH1*/E-cadherin gene (Batlle *et al.*, 2000, Comijn *et al.*, 2001, Hajra *et al.*, 2002). The E-cadherin gene locus can also be epigenetically silenced by hypermethylation (Di Croce & Pelicci, 2003) and proteolytic cleavage of E-cadherin by matrix metalloproteases (MMPs) is another mechanism by which E-cadherin-mediated tumour cell-cell adhesion can be ablated (Nawrocki-Raby *et al.*, 2003). The mechanisms of E-cadherin mediated signalling involve a crosstalk with receptor tyrosine kinases, such

as EGFR, c-Met and IGF1R, the activation of the Wnt signalling pathway and modulation of the available pool of β-ctn, and the signalling through RHO GTPases (Cavallaro & Christofori, 2004).

However, not all classical cadherins have a tumour suppressor function. In a process termed epithelial to mesenchymal transition (EMT), which is normally present in embryonic development and tissue repair, tumour cells progressively down-regulate E-cadherin and express *de novo* N-cadherin and other mesenchymal molecules (Thiery, 2002). This cadherin switch leads to the inhibition of cell-cell contacts and elicits signals that support tumour cell migration, invasion and metastasis (Frixen *et al.*, 1991). This type of transition is found in several human cancer types, including melanoma, prostate and breast cancer (Cavallaro & Christofori, 2004).

Importantly, P-cadherin molecule was also shown to promote cell motility, migration and cell invasion. The aggressive properties mediated by P-cadherin will be addressed in detail in the following section.

Described for the first time in 1986, as "a novel class of cadherin that appeared in developing mouse embryos", this adhesion molecule was found in the structures that gave rise to its name, the placenta (Nose & Takeichi, 1986). P-cadherin is a calcium-dependent cell-cell adhesion glycoprotein, which plays a role in many cellular processes such as embryonic development, differentiation, cell polarity, growth and migration (Larue *et al.*, 1996). In the adult, it is only expressed certain tissues, usually co-expressed with E-cadherin, such as the basal layer of the epidermis and breast, the mesothelium, the ovary, the prostate, the hair follicle, and the corneal endothelium (Nose & Takeichi, 1986, Imai *et al.*, 2008). Sharing about 67% of homology with the *CDH1*/E-cadherin gene, P-cadherin differs mainly in the extracellular portion; it is far less characterized, and it has a considerably opposing effect regarding mammary cancer biology (Albergaria *et al.*, 2011, Hulpiau & van Roy, 2009). In normal breast tissue, P-cadherin expression is restricted to myoepithelial cells as a linear cell membrane staining, whereas E-cadherin is expressed in both myoepithelial and epithelial cell layers of the breast (Paredes *et al.*, 2002a).

The *CDH3*/P-cadherin gene harbours 16 exons (**Figure 7**) and maps to chromosome 16q22.1, a region that contains a cluster of several cadherin genes, just 32 kilobases upstream of the gene encoding E-cadherin (Bussemakers *et al.*, 1994).

Concerning *CDH3*/P-cadherin gene regulation, the literature is not as extensive as regarding *CDH1*/E-cadherin gene. The main transcriptional activators described for the *CDH3*/P-cadherin gene promoter are β-catenin (Faraldo *et al.*, 2007), p63 (Shimomura *et al.*, 2008) and C/EBPβ (Albergaria *et al.*, 2010). In contrast, BRCA1/c-Myc/Sp1 complex acts as transcriptional repressor of the *CDH3* promoter (Gorski *et al.*, 2009). It was also

demonstrated that ER can indirectly repress P-cadherin expression by controlling epigenetic changes in CDH3 gene promoter (Albergaria *et al.*, 2010) (**Figure 8**).

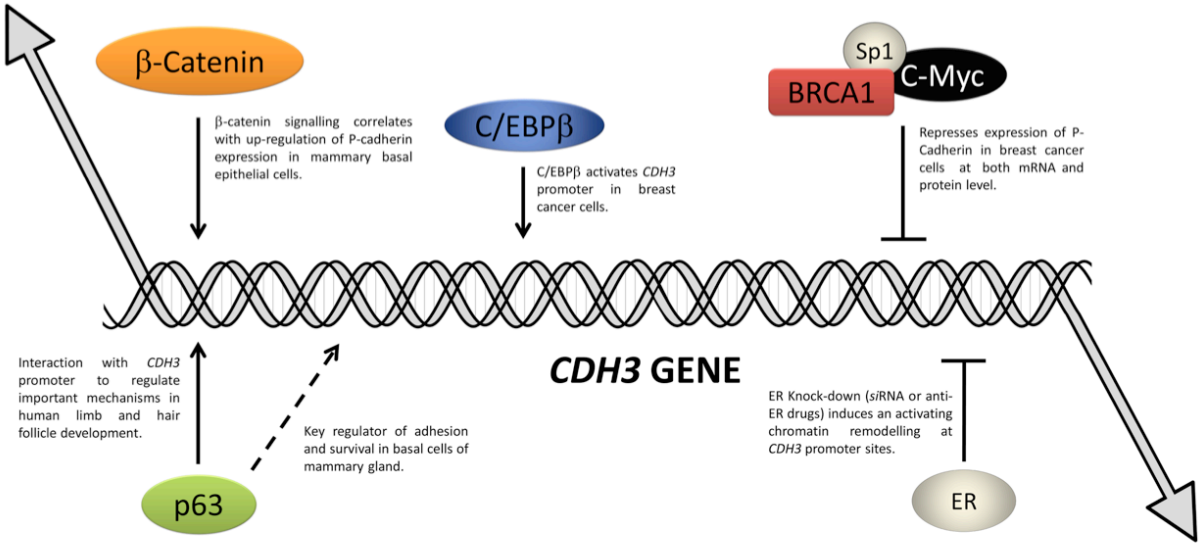


Figure 8 – Transcriptional regulators of *CDH3*/P-cadherin gene promoter. It has been shown that β-catenin, p63 and C/EBPβ are transcriptional activators of the *CDH3* promoter. In contrast, BRCA-1/c-Myc/Sp1 complex acts as transcriptional repressor of the CDH3 promoter. It was also demonstrated that oestrogen receptor (ER) can indirectly repress P-cadherin expression by controlling epigenetic changes in *CDH3* gene promoter. Adapted from Albergaria A, *et al., International Journal of Development Biology*, 2011 (Albergaria *et al.*, 2011).

4.1. P-cadherin promoting effects in breast cancer

P-cadherin aberrant expression is found in 20% to 40% of invasive breast carcinomas, as well as in 25% of pre-invasive (*in situ*) ductal carcinomas. Aberrant P-cadherin expression was shown to be associated with tumours of high histological grade, as well as with well established markers of poor prognosis, like Ki-67, EGFR, CK5, vimentin, p53 and HER-2 expression, and negatively associated with age at prognosis and hormonal receptors (ER and PgR) expression. Interestingly, none of these reports showed a significant association with tumour size and lymph node metastasis (Paredes *et al.*, 2005, Paredes *et al.*, 2002b, Gamallo *et al.*, 2001, Peralta Soler *et al.*, 1999, Turashvili *et al.*).

P-cadherin up-regulation is predominantly found in the basal-like subgroup of breast cancers (Matos *et al.*, 2005, Paredes *et al.*, 2007a, Paredes *et al.*, 2007b) and it is strongly associated with the presence of *BRCA1* mutation (Arnes *et al.*, 2005) and poor clinical outcome (Paredes *et al.*, 2005, Turashvili *et al.*, 2011).

P-cadherin is altered in various human tumours, but its effective role in the carcinogenesis process remains object of debate, since it can behave differently depending on the studied tumour cell model and context. For example, in melanoma, P-cadherin seems to have a tumor suppressive function, exactly as E-cadherin (Van Marck *et al.*, 2005). In breast and several other models, P-cadherin is often reported to exhibit tumour promoting effects.

Using breast cancer cell lines, we found that overexpression of P-cadherin promotes cell motility, cell migration, as well as invasion capacity in matrigel matrix (Ribeiro *et al.*, 2010). A similar aggressive phenotype was also observed in bladder and pancreatic cancer cell lines (Mandeville *et al.*, 2008, Taniuchi *et al.*, 2005, Van Marck *et al.*, 2011). Curiously, we have noticed that P-cadherin is able to induce invasion only in cell systems which already express an endogenous and functional E-cadherin in breast cancer cells (Paredes *et al.*, 2004, Ribeiro *et al.*, 2010, Taniuchi *et al.*, 2005). In fact, breast carcinomas that co-express E- and P-cadherin actually show a patient survival that is even worse than carcinomas that express only one of the cadherins or that do not express any of these adhesion molecules (Paredes *et al.*, 2008). These tumors have a decrease in membrane staining of p120ctn and an increase in the cytoplasmic localization for this catenin. In pancreatic and ovarian cancer, it was shown that p120ctn, once in the cytoplasm can activate Rho-GTPases, Rac1 and Cdc42, altering the actin cytoskeleton polymerization and promoting cell motility (Taniuchi *et al.*, 2005, Cheung *et al.*, 2010). Based on this, we have recently proved that P-cadherin is able to interact with E-cadherin in breast tumours and cancer cells, promoting cancer cell invasion by disrupting the interaction between E-cadherin and both p120ctn and β-ctn. In the absence of E-cadherin expression, in the same cancer model, P-cadherin is able to suppress invasion by its strong interaction with catenins, surrogating the role of E-cadherin in cell-cell adhesion (Ribeiro *et al.*, unpublished data). It is also possible that signaling pathways, such as Wnt/β-ctn, may play an important role in mediating the oncogenic effects derived from P- and E-cadherin aberrant expression. Notably, this signaling pathway is responsible for the self-renewal and pluripotency of mammary stem cells (Beachy *et al.*, 2004). Transient or irreversible inactivation of the cadherin-catenins complexes by factors secreted by tumor cells and/or stromal cells cannot be ruled out. An example of such factors could be metalloproteases. These post-translational mechanisms could also be related to accumulation of catenins in the cytoplasm. *In vitro* studies from our laboratory have shown that, in E-cadherin positive breast cancer cell lines, P-cadherin overexpression promotes cell invasion, migration and motility accompanied by the secretion of MMP-1 and MMP-2, which lead to P-cadherin ectodomain cleavage. This soluble P-cadherin fragment is able to induce *in vitro* invasion of breast cancer cells (Ribeiro *et al.*, 2010).

Importantly, in addition to breast cancer, P-cadherin upregulation is also found in other malignancies such as gastric, endometrial, colorectal and pancreatic carcinomas (Hardy *et al.*, 2002, Imai *et al.*, 2008, Stefansson *et al.*, 2004, Taniuchi *et al.*, 2005). Targeting P-cadherin in cancer may be a good therapeutical approach, since normal tissues usually express very low levels of this cadherin (Imai *et al.*, 2008). In fact, a novel and highly selective human monoclonal antibody against P-cadherin (PF-03732010) demonstrated anti-tumour and anti-metastatic activity in a panel of P-cadherin overexpressing tumour models, without significant secondary effects in mice (Park *et al.*, 2011, Zhang *et al.*, 2010). PF-03732010 has just completed a Phase I clinical trial.

4.2. P-cadherin role in breast cell differentiation, development and stem cell biology

P-cadherin deficient female mice present abnormal mammary gland morphology and development. Although no tumors were detected, these animals exhibit precocious mammary gland differentiation in the virgin state, and breast hyperplasia and dysplasia with age (Radice *et al.*, 1997). These observations implicate P-cadherin cell-cell interactions and signaling as regulatory determinants of the negative growth of the luminal epithelium, being important for the maintenance of an undifferentiated state of the normal mammary gland. Noteworthy, this mouse model was never thoroughly studied regarding cancer, namely upon oncogenic transformation of breast cells.

In humans, the loss of P-cadherin induces characteristic genetic syndromes. *CDH3* gene mutations have been shown to cause P-cadherin functional inactivation, leading to developmental defects associated with two inherited diseases in humans: 1) hypotrichosis with juvenile macular dystrophy (HJMD) and 2) ectodermal dysplasia, ectrodactyly, and macular dystrophy (EEM syndrome). The common features of both diseases are sparse hair and macular dystrophy of the retina, while only EEM syndrome shows the additional finding of split hand/foot malformation (SHFM) (Kjaer *et al.*, 2005, Sprecher *et al.*, 2001). No defects were described for these conditions, concerning the human mammary development, or other epithelial bud structures.

The role of P-cadherin in development and differentiation is clear during embryonic histogenesis, as this cadherin molecule is specifically found in extra-embryonic ectoderm and visceral endoderm, originating the placenta in mice (Hirai *et al.*, 1989b). Also in humans, this protein is detected in the placental tissue and its embryonic precursors, but with a lower expression level (Aplin *et al.*, 2009, Shimoyama *et al.*, 1989).

In a recent study, it has been shown that *CDH3* is one of the genes that encode a surface protein that identifies the pluripotent population of human embryonic stem cells. This expression is concomitant with E-cadherin (Kolle *et al.*, 2009). Interestingly, stem cell transcription factors, namely β-catenin and p63, were found to induce P-cadherin expression in (Faraldo *et al.*, 2007, Shimomura *et al.*, 2008). Additionally, in the growing hair follicle, the early hair progenitor cells from hair germs and small hair placodes, were isolated as P-cadhein^{+} (and K14^{+}/α6-integrin^{+}) cells (Rhee *et al.*, 2006). In fact, P-cadherin expression is mainly found in the basal and proliferating regions of the tissues, such as of the epidermis and hair follicles, where stem cells are known to be located (Fujita *et al.*, 1992). Like hair follicles, sweat glands and mammary glands develop also from the same discrete accumulation of stem cells resting in the primitive epidermis, the outermost cell layer of an embryo, and there is strong evidence that dynamic changes in the composition of adherens junctions are important for the development of skin appendages (Fujita *et al.*, 1992).

P-cadherin is extremely important for the establishment of the correct architecture of the mammary epithelial tissue, as recently demonstrated by Chanson *et al.*. P-cadherin contributes specifically to the organization of the myoepithelial cell layer of the breast, since when an antibody that blocks P-cadherin function was used in an *in vitro* self-organizing assay of the human mammary bilayer, the migration of mammary epithelial cells, occurring during normal sorting of both layers, was compromised (Chanson *et al.*, 2011).

Interestingly, the expression of this adhesion molecule is activated in human mammary luminal cells during late pregnancy and lactation. In human milk a soluble fragment of P-cadherin (sP-cad) with 80KDa was found to be present, corresponding to the extracellular domain of the molecule (**Figure 9**) (Soler *et al.*, 2002). Recently, Mannello and collaborators showed that the highest concentration of sP-cad is detected in milk collected during the first trimester of lactation (Mannello *et al.*, 2008). Still, it is not clear the biological and physiological role attributed to this fragment in the normal function of the breast. Some authors suggest a role for sP-cad in alveolar differentiation during lactation, or in the immune response of the mother or the baby, or as a signalling protein between epithelial and myoepithelial cells (Albergaria *et al.*, 2011).

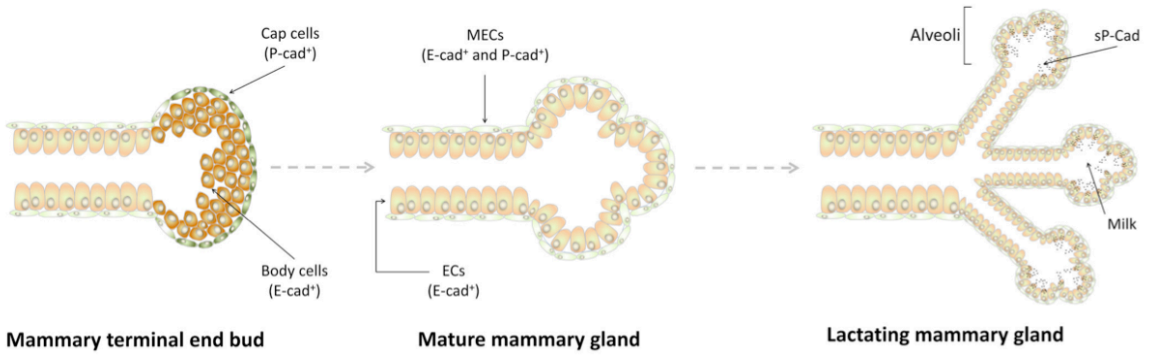


Figure 9 – Representation of cadherins expression during mammary gland development. During embryoniogensis and puberty, the growth of the ductal tree in the developing mammary gland is sustained by the terminal end bud (TEB), where cells with stem cell activity are located. Cap cells at the tip of the TEB are P-cadherin+ and E-cadherin+ and generate cells of the basal lineage, namely myoepithelial cells (MECs). The cells in the central TEB mass are known as body cells and will differentiate into the luminal epithelial lineage, producing the luminal epithelial cells (ECs) (E-cadherin+). Although it is still unclear how the TEB progresses through the fat pad, proteases and extracellular matrix remodeling enzymes are involved in the process. During lactation, luminal secretory cells present in the breast lobules express P-cadherin and secrete a soluble form of this protein (sP-cad). Adapted from Albergaria A, *et al., International Journal of Development Biology*, 2011 (Albergaria *et al.*, 2011).

5. Integrins and the extracellular matrix in breast cancer

During the past few years, results from the cellular and molecular dissection of tumour progression have led to the idea that, besides the cellular processes and molecular pathways that exist in tumor cells themselves, an equally important contribution to malignant tumor progression comes from components of the tumor microenvironment. These include endothelial cells, tumour fibroblasts, infiltrating cells of the immune system and the tumor ECM (Christofori, 2006, Bissell & Labarge, 2005). Many tumors can lie dormant or develop slowly for decades before manifesting as a clinical outcome and it is believed that the tumor microenvironment has a fundamental role in the putative niche where CSCs are located (Bissell & Labarge, 2005). Importantly, the microenvironment can function both as a tumor promoter for precancerous or even apparently normal cells and as a powerful tumor suppressor (Bissell & Labarge, 2005).

Both stromal and epithelial cells are responsible for producing many of the connective tissue ECM and the ECM degrading enzymes. Regarding ECM components, laminin is a major constituent of the basement membrane (basal lamina or lamina propria) of the breast. This is a 100 nm thick sheet of glycoproteins and proteoglycans, which is constructed around an assembled polymer of laminins and a cross-linked network of collagen IV fibrils (Yurchenco & Patton, 2009). The basement membrane has a gate-

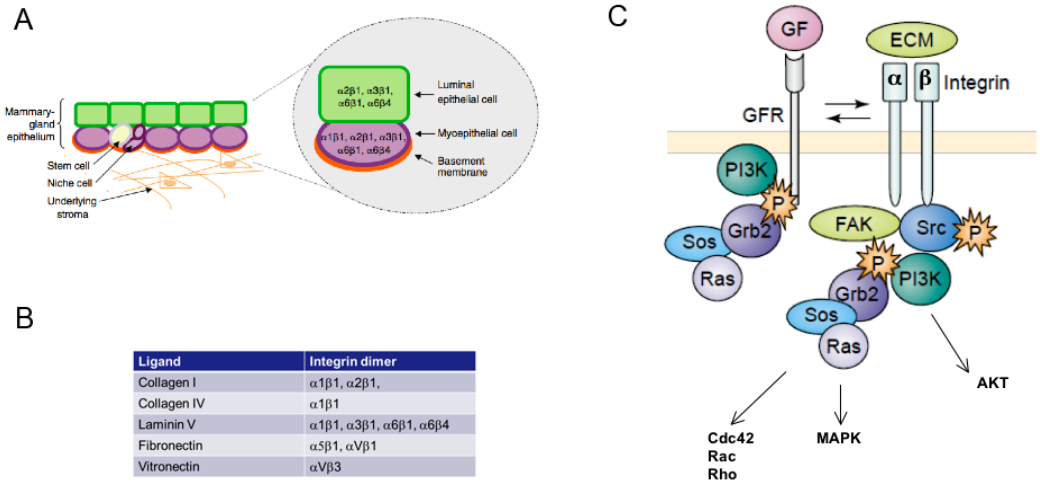
keeper function, since it maintains tissue polarity in the normal epithelium and it retains nascent *in situ* carcinomas within its boundaries, as a benign lesion (Allred *et al.*, 2008, Chin *et al.*, 2004, Hwang *et al.*, 2004). In cancer, tumor cells increase the production of proteases, occurring basement membrane remodeling at the invasive front (Rowe & Weiss, 2008) and, simultaneously, cancer associated myoepithelial cells loose the production of laminin production (Allinen *et al.*, 2004, Gudjonsson *et al.*, 2002). Upon stromal invasion, the tumor cell becomes exposed to a distinct array of matrix molecules, such as collagen type-I and hyaluronan, which influence cell invasion and metastasis (Itano & Kimata, 2008, Provenzano *et al.*, 2008). Importantly, increases in collagen synthesis and cross-linking produce a stiffer matrix, which can foster malignancy. Notably, breast density is a significant risk factor for cancer progression.

In the metastatic setting, ECM molecules have also a central role, namely in the preparation of the pre-metastatic niche and in the quiescence, survival and growth of metastasis (Kaplan *et al.*, 2006, Psaila & Lyden, 2009). In fact, changes in cell-ECM interactions are present in the progression of cancers at every stage, from premalignancy to invasion, and the survival and growth of metastasis (Bissell & Radisky, 2001, Muschler & Streuli, 2011). The dynamics of cell-ECM interactions and ECM turnover remain to be studied in detail. However, efficient therapies aimed at reverting the ECM of the tumor stroma, or disrupting the metastatic niche could improve cancer therapy (Muschler & Streuli, 2011). The manipulation of the ECM receptors (integrins) and their signalling (Pontier & Muller, 2009), the manipulation of ECM modifying enzymes (Overall & Kleifeld, 2006), or specific ECM antagonists (Tsuruta *et al.*, 2008) are possible targets for the tumour microenvironment.

Integrins are the major ECM receptors, which can also serve for some cell-cell interactions. Their ability to promote cell anchorage, proliferation, survival, migration and the induction of active ECM-degrading enzymes suggests that they play an essential role in normal mammary morphogenesis, as well as in promoting tumour progression. Integrins are heterodimers composed of two transmembrane proteins, the α-subunit and the β-subunit. The binding of ECM ligands to integrins is followed by recruitment of several signalling and adaptor proteins to the β-integrin cytoplasmic domain and the initiation of signalling cascades (**Figure 10**).

Immunohistochemical studies in human and mouse mammary gland revealed that most integrins are on sites of cell-ECM interaction, with the presence of α2, α3, α6, β1 and β4 subunits in the luminal and myoepithelial cells, and of α1, α5 and αv chains exclusively in the myoepithelial cells. Some integrins have also been found on the lateral

surface of luminal cells at sites of cell-cell interaction, where ECM proteins are not detected (Taddei *et al.*, 2003).



Ligand	Integrin dimer
Collagen I	α1β1, α2β1,
Collagen IV	α1β1
Laminin V	α1β1, α3β1, α6β1, α6β4
Fibronectin	α5β1, αVβ1
Vitronectin	αVβ3

Figure 10 – (A) Schematic representation of the structure of the mammary epithelium and the different integrin heterodimers expressed in luminal and myoepithelial cells (adapted from Pontier et al., *Journal of Cell Science*, 2009 (Pontier & Muller, 2009)); **(B)** 19 α-subunits and 8 β-subunits have been reported in vertebrates. They form at least 25 αβ heterodimers that bind to different ligands present in the ECM (adapted from Taddei *et al.*, *Journal of Mammary Gland Biology and Neoplasia,* 2003 (Taddei *et al.*, 2003)). **(C)** Integrin downstream signalling is often coupled to signals triggered by soluble growth factors (GF) and growth factor receptors (GFR) (adapted from Comoglio et al., *Current Opinion in Cell Biology,* 2003 (Comoglio *et al.*, 2003)).

Implicated in signal transduction by integrins, as well as other cell surface receptors, is a cytoplasmic tyrosine kinase, Focal Adhesion Kinase (FAK). Upon its activation and auto-phosphorylation at Tyr-397, eg., by integrin mediated cell adhesion to the ECM, FAK associates to a number of other kinases and adaptor molecules, including Src family kinases, PI3K, Grb7, Shc, paxilin and p130CAS. FAK is recruited as a participant in focal adhesion dynamics between cells, and the interaction with these signalling molecules has been shown to regulate cell spreading, migration, cell survival and cell cycle progression (Parsons, 2003). FAK expression plays an important role in mammary gland development: although the complete deletion of FAK is embryonic lethal, mammary epithelial cell-specific FAK conditional knock-out leads to reduced cyclin D1, as well as reduced Erk and STAT-5 signalling. As consequence, FAK knock-out females have severe lobulo-alveolar hypoplasia and secretory immaturity during pregnancy and lactation (Nagy *et al.*, 2007). This conditional knock-out model also proved that FAK expression in the mammary epithelium has important oncogenic effects (Luo *et al.*, 2009). Ablation of FAK leads to the depletion of cancer stem/progenitor cells and impaired self-renewal and migration *in vitro*, as well as the tumourigenic capacity *in vivo*. Analysis of human tumour samples and cell

lines derived from tumors reveals elevated expression of FAK, compared to normal tissues (Owens *et al.*, 1995). In such cells, increased FAK expression has been correlated with increased cancer cell motility, invasiveness and proliferation (Owens *et al.*, 1995, Slack *et al.*, 2001, Wang *et al.*, 2000).

An important kinase with the ability to phosphorylate FAK, which is also found in the focal adhesion structures, is Src. This is a non-receptor tyrosine kinase with several substrates that was found to be over-expressed and highly activated in a wide variety of human cancers (Irby & Yeatman, 2000). In human mammary carcinomas and in breast cancer cell lines, Src kinase activity was increased compared with normal tissues or normal cells (Egan *et al.*, 1999, Jacobs & Rubsamen, 1983). Similarly to FAK, the effect of Src over-expression and activation appears to be pleiotropic, regulating cell spreading, migration and cell survival. The essential role of Src in the breast in seen in the src-null mice, which displayed a dramatic mammary developmental delay characterized by the inability of mammary epithelial cells to activate a number of signalling pathways in response to exogenous oestrogen stimulation (Kim *et al.*, 2005).

Src has also been shown to play a significant role in cancer progression and the metastatic phenotype, regulating angiogenesis, tumour cell invasion and tumour growth (Irby & Yeatman, 2000). The specific activity of Src protein kinase is not only dependent in integrin receptor signaling, but it may also be increased by direct or indirect interaction with receptor tyrosine kinases, such as EGFR, platelet derived growth factor receptor, fibroblast growth factor receptor, colony simulating factor-1 receptor, HER2, and hepatocyte growth factor receptor (c-Met) (Irby & Yeatman, 2000). The extensive presence of activated Src in human cancers and its potential role in their development and progression, makes Src an appealing target for drug discovery efforts.

5.1. The α6β4 and α6β1 integrins in the breast

Integrins have been implicated in several aspects of tumor progression. One of the most studied integrin dimmer in breast is the laminin receptor, the α6β4 integrin, which has been implicated in tumor cell survival, migration and invasive potential (Mercurio *et al.*, 2001, O'Connor & Mercurio, 2001, Shaw, 1999). In normal cells, α6β4 forms a hemidesmosomal integrin, which is linked to the cytokeratin intermediate filament network, forming strong adhesive structures to the laminin substratum. In tumor cells, α6β4-integrin is associated with actin microfilaments and does not form hemidesmosomes. In this way, α6β4 activates PI3K, the small GTP-ases Rho and Rac, and the protein kinase A, which are molecules essential for cell migration and invasion (Mercurio *et al.*, 2001, O'Connor &

Mercurio, 2001), and tumor cell survival (Mercurio *et al.*, 2001). α6β4 also cooperates with other molecules such as HER-2 or tetraspanin, in order to induce migration (Mercurio *et al.*, 2001).

The role of α6β4 in stem/progenitor cell activity has been partly addressed. Normal breast architecture is not particularly affected by perturbation of either α6 or β4 integrin subunits (Gardner *et al.*, 1996, Klinowska *et al.*, 2001). α6 integrin (also known as CD49f) has been recently used as a marker to purify mouse and human mammary stem cells (Eirew *et al.*, 2008, Raouf *et al.*, 2008, Stingl *et al.*, 2001, Stingl *et al.*, 2006, Villadsen *et al.*, 2007) and characterize a population of human cancer cells with aggressive behavior (Cariati *et al.*, 2008). Another example where integrins were described as stem cell markers is β1 integrin (also known as CD29), which was used by Shackleton and collaborators to purify mouse mammary stem cells (Shackleton *et al.*, 2006). Interestingly, however, deletion of β1 integrin clearly affected mammary morphogenesis and the ability of stem cells to self-renew (Taddei *et al.*, 2008), causing a reduction in the TEB number, growth defects in pregnancy and lactation, and precocious dedifferentiation in involution (Taddei *et al.*, 2008). β1 integrin is able to form a heterodimer with α6 integrin, constituting another laminin receptor, the α6β1 heterodimer. It is thought that the attachment of stem/progenitor cells through α6β1, or even α6β4, to interstitial matrix components, namely laminins, might support their stem cell activity, or contribute to a cancer stem cell phenotype. The actual molecular role of integrins in these cells and the signals released from the normal and the cancer stem cell niche remain to be determined (Pontier & Muller, 2009).

AIMS

CHAPTER II

Breast cancer is a type of solid carcinoma for which tumour initiating cells exhibiting stem-like properties have already been isolated and described. However, the cellular origin of these cancer stem cells, as well as their phenotype is not completely clear among all the molecular subtypes described for this disease.

The general goal of this work is to characterize the stem cell features in the basal-like subtype of breast carcinomas with a focus in P-cadherin, a classical adhesion molecule with a poor prognostic signature and whose oncogenic signalling pathways could potentially be targeted for therapy.

Thus, comprising a series of human breast cancer cell lines and invasive breast carcinoma samples, studies were performed in order to assess the following specific aims:

i. TO EVALUATE THE EXPRESSION DISTRIBUTION OF BREAST CANCER STEM CELL MARKERS WITHIN THE DIFFERENT MOLECULAR SUBTYPES OF BREAST CANCER

Despite several methods have been described for the characterization of stem cell properties, further studies are needed in order to support the characterization of the cancer stem model in both human tissue cases and in human breast cell lines, with a comparison between both models. Furthermore, it is imperative to improve CSCs identification into routine formalin-fixed and paraffin-embedded tissue samples. Moreover, there is often inconsistency in the published literature. Therefore, the aim of this work is to characterize stem cell properties in our series of human breast cell lines and compare it with our series of human breast cancer tissues and the published work from other groups.

ii. TO EVALUATE IF P-CADHERIN EXPRESSION CAN MEDIATE STEM CELL PROPERTIES IN BASAL-LIKE BREAST CANCER

The role of the adhesion molecule P-cadherin in breast cancer aggressiveness is well described by our group and others, especially in the basal-like subtype, for which this adhesion molecule was found to be overexpressed and correlated with poor survival. This adhesion molecule has also a role in different contexts of stem cell biology. Thus, considering the Cancer Stem Cell hypothesis of tumour progression, the second aim of this work was to explore the potential role of P-cadherin in mediating cancer stem cell phenotype/activity in this special group of human breast carcinomas.

iii. TO EVALUATE IF P-CADHERIN SIGNALLING IS DEPENDENT ON INTEGRIN ACTIVATION TO INDUCE BREAST CANCER STEM CELL AND INVASIVE PROPERTIES

The specific signalling pathways mediated by P-cadherin are only recently being clarified, namely in the context of cell-cell adhesion. Importantly, studies addressing the role of this adhesion molecule in the context of cancer cell-extracellular matrix (ECM) adhesion and signalling are lacking. Thus, clarifying the importance of P-cadherin in the tumour microenvironment is crucial, as this would potentially provide new insights on the 'cancer stem cell niche' signalling. Hence, the third aim of this work was to test weather P-cadherin affects the adhesion of cancer cells to different ECM substrates, as well as to clarify the signalling pathways mediated by this adhesion molecule in response to microenvironmental stimuli, namely integrin activation.

MATERIALS AND METHODS

CHAPTER III

This chapter describes the materials and the methods that were used to generate all the data presented in Chapters IV, V and VI.

Human breast cell lines and cell culture

Different human breast cancer cell lines were used during the work, as stated in the different chapters. The breast cancer cell lines T47D, SkBr3, BT-474, MDA-MB-468, BT-20, MDA-MB-231, BT-549 and were obtained from ATCC (American Type Culture Collection, Manassas, VA). The human breast cancer cell line MCF-7/AZ was obtained from a collection developed in the laboratory of Prof. Marc Mareel (Ghent University Hospital, Belgium), which was genetically manipulated to overexpress P-cadherin (MCF-7/AZ.Pcad). The control cell line (MCF-7/AZ.mock) shows low P-cadherin levels, identical to the parental cell line (Paredes et al., 2004).

All these cell lines were grown in DMEM supplemented with 10% FBS (Invitrogen, Carlsbad, CA) and with 1% antibiotic solution (penicillin–streptomycin, Invitrogen).

The normal human MCF10A cells was also obtained from ATCC and cultured in DMEM:F12, supplemented with 5% heat inactivated horse serum (Invitrogen), 10 µg/ml insulin (Sigma-Aldrich, St Louis, MO), 0.5 µg/ml hydrocortisone (Sigma-Aldrich), 20 ng/ml EGF (Sigma-Aldrich), 100 ng/ml cholera toxin (Sigma-Aldrich).

All cell lines were routinely cultured in a humidified atmosphere with 5% CO_2 and at 37ºC. Cells were used in experiments when reached 70–80% confluence.

Flow cytometry analysis

This technique was used to investigate the cell surface expression of the stem cell markers CD44, CD49/α6-integrin and CD24 in a series of breast cancer cell lines (in Chapter IV). Moreover, P-cadherin cell surface expression was also evaluated, in Chapter V. Finally, the expression of the markers CD49f/α6-integrin and CD104/β4-integrin, CD29/β1-integrin was investigated in Chapter VI.

Cells were washed twice with PBS and then harvested with versene/0.48mM EDTA (Invitrogen). Detached cells were washed with PBS supplemented with 0.5% FBS (stain buffer), and re-suspended in the stain buffer ($1x10^6$ cells/100µl). Cells were passed through a 25G needle in order to obtain a single cell suspension. This cell suspension was labeled by fluorescence-conjugated antibodies at a concentration of 1 to 10 in stain buffer: FITC-conjugated CD44, PE-conjugated CD24, FITC-conjugated CD49f/α6-integrin, PE-

conjugated CD104/β4-integrin and PE-Cy5-conjugated CD29/β1-integrin. These antibodies were obtained from BD Biosciences (San Diego, CA). P-cadherin monoclonal antibody APC-conjugated was obtained from R&D (Minneapolis, MN) and used at the same concentration as above.

Primary antibodies or the respective isotype controls (BD Biosciences) were incubated at 4°C in the dark for 20 min. A cell viability marker was included (violet fluorescent reactive dye, Invitrogen) in order to remove dead cells. The labeled cells were washed in the stain buffer and then analysed on a FACS Calibur (BD Biosciences) or LSR-II (BD Biosciences).

In multicolor experiments, fluorescent minus one samples were used to determine the gating strategy.

ALDEFLUOR assay

ALDH activity was assessed in several breast cancer cell lines representing the main molecular subtypes of human breast cancer (Chapter IV and V). The ALDEFLUOR kit (Stem Cell Technologies) was used to characterize the population with high ALDH enzymatic activity using a FACS Calibur (BD Biosciences) or LSR-II (BD Biosciences), according to the manufacturer instructions. Briefly, cells were incubated in ALDEFLUOR assay buffer containing the ALDH substrate, BAAA (1 μmol/L per 1 x 10^6 cells). In each experiment, a sample of cells was stained, under identical conditions, with 50 mmol/L of diethylaminobenzaldehyde (DEAB), a specific ALDH inhibitor, as a negative control. A cell viability marker was included (violet fluorescent reactive dye, Invitrogen) in order to remove dead cells. Cells were washed in assay buffer and left at 4ºC, until measurements and sorting had been performed.

In the experiments performed in Chapter V, P-cadherin monoclonal antibody (APC-conjugated, R&D) was incubated after the reaction of ALDEFLUOR cells, in assay buffer, for 20 minutes on ice. Incubation with the isotype antibody (BD Biosciences) was used as a control. Cells were washed in assay buffer and analyzed in the FACS analyzer LSR-II (BD Biosciences).

Immunoblot analysis

Used to monitor protein expression levels, as shown in Chapters V and VI.

Cells grown in monolayer were lysed with phosphate-buffered saline (PBS) containing 1% Nonidet-P40 (Sigma-Aldrich), 1% Triton X100 (Sigma-Aldrich) and protease inhibitors cocktail (Roche Diagnostics Gmbh, Mannheim, Germany) and phosphatase inhibitors cocktail (Sigma-Aldrich).

In some experiments, the cell lysis was done after performing the adhesion assay over a laminin coated surface (6 wells plate, BD Biosciences) as described below.

Protein concentration was determined by Bio-Rad protein assay (Bio-Rad, Richmond, CA) and equal amounts were resolved on a denaturating polyacrylamide gel (Bio-Rad) and transferred onto a nitrocellulose membrane (Amersham Pharmacia Biotech, Piscataway, NJ). After blocking nonspecific binding with 5% non-fat dry milk (for non-phosphorylated protein detection) or 5% BSA (for phosphorylated protein detection) in PBS containing 0.5% Tween 20, each membrane was incubated for 1 hour at room temperature with each of the following primary antibodies: anti-P-cadherin (1:500, BD Transduction), anti-CD49f/α6-integrin (1:000, HPA012696, Sigma-Aldrich), anti-CD44 (1:500, Cell Signalling Technology, Danvers, MA), anti-CD44v6 (1:500, VFF-7, Abcam, Cambridge, UK), and anti-CD104/β4-integrin (1:2000, Santa Cruz Biotechnologies). The primary antibodies used to study signalling activation were: anti-pSrc Tyr416 (1:1000, Cell Signalling, Danver, MA), anti-total Src (1:1000, Cell Signalling), anti-pFAK Tyr397 (1:1000, Cell Signalling), anti-total FAK (1:500, BD Transduction), anti-pAKT Ser473 (1:2000, Cell Signalling) and anti-AKT1/2 (1:500, Santa Cruz Biotechnology).

β-actin (1:1000, Santa Cruz Biotechnology, Inc., Heidelberg, Germany) or α-tubulin (1:10 000, DM1A, Sigma-Aldrich) were detected as loading controls.

After washing four times with PBS for 5 minutes, the membranes were incubated with peroxidase-conjugated secondary antibodies from Santa Cruz Biotechnology (1:2000) for 45 minutes and washed six times for 5 minutes. Detection was detection was performed using the Amersham Hyperfilm and Amersham ECL Detection substrate (GE Healthcare, Chalfont St. Giles, UK).

Mammosphere assay

This assay was used as a measure of stem cell activity in Chapters IV, V and VI.

Breast cancer cell lines were harvested with trypsin (Invitrogen) or versene/0.48mM EDTA, neutralized with 10% FBS containing medium, centrifuged 1200 rpm for 2 min at

room temperature and re-suspended in cold PBS. This solution was passed three times through a 25-gauge needle, using a syringe, to separate cells into a single cell suspension. Cells were plated at 500 cells/ml in low attachment plates coated with 1.2% poly(2-hydroxyethylmethacrylate)/95%ethanol (Sigma-Aldrich). Cells were grown for 5 days, in DMEM/F12 containing B27 supplement (Invitrogen), and 500 ng/ml hydrochortisone (Sigma-Aldrich), 40 ng/ml insulin (Sigma-Aldrich), 20 ng/ml EGF (Sigma-Aldrich) and maintained in a humidified incubator at 37°C and 5% (v/v) CO_2. Mammosphere forming efficiency (MFE) was calculated as the number of mammospheres formed (≥50 μm, determined by using an eyepiece graticule with crossed scales), divided by the cell number plated, being expressed as a percentage.

Patient selection and TMA construction

A series of 466 primary invasive breast carcinomas was retrieved from the Pathology Department, Hospital Xeral- Cíes, Vigo, Spain, diagnosed in 1978 - 1992.

Representative tumor areas were carefully selected and at least two tissue cores (0.6 mm in diameter) were deposited into a tissue microarray. Non-neoplastic tissue cores were included as controls. Patient ages ranged from 28 to 92 years old. The formalin-fixed paraffin-embedded histological sections were reviewed and the diagnosis confirmed by a pathologist. The tumours were characterised for clinical and pathological features, namely age, tumour size, lymph nodes status and histological grade. Patient follow-up information was available for 455 cases, ranging from a minimum of one to a maximum of 120 months after the diagnosis. The disease-free survival (DFS) interval was defined as the time from the diagnosis to the date of breast-cancer-derived relapse/metastasis, whereas overall survival (OS) was considered as the number of months from the diagnosis to the disease-related death. This study was conducted under the national regulative law for the handling of biological specimens from tumour banks, being the samples exclusively available for research purposes in retrospective studies.

Immunohistochemistry and Immunohistochemical evaluation

This technique was used to characterize our series of human carcinomas into molecular subtypes and to investigate the expression and subcellular localization of the stem cell markers (CD44, CD49f, CD24 and ALDH1) presented in Chapters IV and V.

In order to classify all breast cancer tumours molecularly, we evaluated the

expression of commonly used breast cancer biomarkers for this purpose, namely the hormonal receptors ER and PgR, the proliferation marker Ki67, the tyrosine kinase receptors HER-2 and EGFR, the basal cytokeratins CK5 and CK14, and P-cadherin and vimentin as basal markers. Immunohistochemistry (IHC) was performed in 3 mm sections. **Table 1** presents the primary antibodies used to characterize this series of human breast carcinomas and the conditions used in the immunohistochemical reactions. High temperature (98ºC) antigenic retrieval with Tris-EDTA or Citrate buffer was performed before primary antibody incubation.

Antigen	Primary antibodies				Antigen retrieval	Detection method
	Clone	Origin	Incubation time	dilution		
ER	SP1	Neomarkers	30’	1:200	Citrate	HRP polymer
PgR	NCL-L-PgR-312	Novocastra	30’	1:300	Citrate	HRP polymer
HER-2	CB11	Novocastra	30’	1:400	Citrate	HRP polymer
Ki67	SP6	Neomarkers	30’	1:300	Citrate	HRP polymer
K5	XM26	Neomarkers	60’	1:50	Citrate	HRP polymer
K14	NCL-L-LL002	Novocastra	60’	1:400	Tris-EDTA	HRP polymer
EGFR	31G7	Zymed	60’	1:100	Pepsin	HRP polymer
P-cadherin	C56	BD Transduction	60’	1:50	Tris-EDTA	HRP polymer
Vimentin	V9	Dako	30’	1:150	Citrate	HRP polymer
CD44	156-3C11	Cell Signalling	30’	1:100	Citrate	Streptavidin-biotin peroxidase complex
CD49f	HPA012696	Sigma-Aldrich	120’	1:10	Citrate	HRP polymer
CD24	SN3b	Neomarkers	30’	1:100	Citrate	HRP polymer
ALDH1	EP1933Y	Abcam	60’	1:100	Citrate	HRP polymer

Table 1 – Antibodies and immunohistochemical reaction conditions used to characterize se series of invasive breast carcinomas.

The primary antibodies were detected using a secondary antibody with horseradish peroxidase polymer (Cytomation Envision System HRP; DAKO, Carpinteria, California, USA), or a biotinylated goat anti-polyvalent as secondary antibody, followed by the streptavidin-peroxidase complex (Thermo Fisher Scientific, Fremont, California, USA), according to the manufacturer’s instructions. Both methods used diaminobenzidine as chromogen.

The expression of the breast cancer biomarkers ER, PgR, HER2, EGFR, CK5, CK14, P-cadherin and vimentin was evaluated according to the grading systems already described (Sousa *et al.*, 2010). These immunohistochemical results were used to classify the tumours in the different molecular breast cancer subtypes, namely in luminal A,

luminal B, HER-2 OE, basal-like and unclassified group, according to the definition presented in **Table 2**.

	Luminal A	Luminal B	HER-2 OE	Triple Negative	
				Basal-like	Unclassified
ER	+	+	-	-	-
PgR			-	-	-
HER-2	-	+	+	-	-
Ki67	-		+ or -	+ or -	+ or -
K5	+ or -	+ or -	+ or -	+	-
K14					-
EGFR					-
Vimentin					-
P-cadherin					-

Table 2 – Immunohistochemical classification of invasive breast carcinomas in distinct molecular subtypes.

P-cadherin, CD44, CD24, CD49f staining were detected mainly at the membrane of tumor cells and the scoring was considered as follows: (0), 0-10% of positive tumor cells; (1+), 10-25% of positive tumor cells; (2+), 25-50% of positive tumor cells; (3+), >50% of positive tumor cells. For P-cadherin, CD44 and CD49f, the cases which were classified as (0) were considered negative, whereas (1+), (2+) and (3+) were established as positive cases. For CD24, the cases were divided in negative/low (-/low), when considered (0) or (1+), or in positive cases when classified as (2+) or (3+). Immunohistochemical staining of ALDH1 was classified as positive when more than 1% of tumor cells showed clear cytoplasmic positivity (Deng *et al.*, 2010, Ginestier *et al.*, 2007). Since the immunohistochemical result was not interpretable for some of these markers, the statistical analyses were performed using only the breast tumor cases with available data. The present study was conducted under the national regulative law for the usage of biological specimens from tumor banks, where the samples are exclusively available for research purposes in the case of retrospective studies.

P-cadherin and stem cell markers knock-down

Gene silencing presented in Chapters V and VI was conducted by siRNA, using sequences specifically targeting P-cadherin and the stem cell genes.

P-cadherin (*CDH3* gene) expression was silenced using the siRNA target sequence: AAGCCTCTTACCTGCCGTAAA. Inhibition of P-cadherin was maintained for at least 72h

after cell transfection, confirmed by western-blot. Inhibition of the expression of CD49f (*ITGA6* gene), target sequence: CAGGGTAATAAACTTAGGTAA, β4-integrin (*ITGB4* gene) target sequence: GTGGATGAGTTCCGGAATAAA and CD44 (*CD44* gene), target sequence AACTCCATCTGTGCAGCAAAC was also performed.

All transfections were carried out using HiPerFect transfection reagent (Qiagen, Hilden, Germany) in a final concentration of 2 or 5 nM siRNA (Qiagen), according to manufacturer instructions. Optimal inhibition of the target genes was achieved at 48h, which was confirmed by immunoblot analysis (see protocol and antibodies above). A siRNA scrambled sequence was included as a control (Qiagen).

Protein analysis was performed by western blot, as described above.

Fluorescence activated cell sorting (FACS) for P-cadherin

For the sorting experiments presented in Chapter V, two luminal cell lines (MCF7/AZ and T47D) and four basal cell lines (MDA-MB-468, BT-20, BT-549, MCF10A) were sorted according to P-cadherin expression (highest 20% expressing cells *Vs.* lowest 20% expressing cells). Briefly, cells were stained for APC-conjugated anti-P-cadherin (R&D) and a live-dead dye (Violet dye, Invitrogen) in stain buffer, as described above. Cells were then passed through a 50μm mesh to remove clumps and re-suspended in stain buffer prior to sorting. Cells were sorted using BD Influx (BD Biosciences) or FACS ARIA-II (BD Biosciences) and collected into 10% Hanks buffered solution (Invitrogen). The purity of sorted populations was checked and the sorted populations contained 80-95% purified cells. In addition, a sample of cells was also collected that passed through the laser under pressure, but not sorted, to act as a control for the effect of the pressure on the cells. No differences in cell behavior were detected.

3D cultures

This assay was used to measure the ability of selected populations of breast cancer cells to growth in differentiating conditions. The 3D on-top method was used. Briefly, single-cell suspensions were seeded at a density of 250 or 1000 cells per well into 8-well glass chamber slides containing 50 μL of 100% growth factor–reduced matrigel per well [a biologically active matrix material resembling the mammalian cellular basement membrane (BD Biosciences)]. Cells were plated in growth medium containing H14 medium (DMEM/F12 with insulin 250 ng/ml, transferrin 10 μg/ml, sodium selenite 2.6

ng/ml, estradiol 10^{-10} M, hydrocortisone $1.4x10^{-6}$M, prolactin 5 µg/ml, EGF 10 ng/ml, according to Kenny PA, *et al.,* 2007 (Kenny *et al.*, 2007) with 5% growth factor–reduced matrigel and 1% FBS).

Cells were incubated at 37 °C, with replacement of the growth medium containing 2% growth factor–reduced matrigel every 2–3 days, to allow 3D structures to form. The size and number of the structures formed was assessed microscopically after 21 days.

Cell Cycle

To evaluate the cell cycle profile in the basal breast cancer cell lines presented in Chapter V, cells grown in monolayer were washed 2 times in PBS and detached using trypsin for 5 min, 37ºC (Invitrogen). Cells were counted and resuspended to $1\text{-}2x10^6$ cells/ml. 3 ml of cold absolute ethanol was added to 1 ml of cell suspension and the cells were fixed for at least 1h at 4ºC. After washing the cells with PBS, the Hoescht-33342 dye (Invitrogen) was added to a final concentration of 2 µg/ml and incubated for 15 min at 37ºC. Cells were analysed for cell cycle fluorescence in BD LSR-II (BD Biosciences).

Cell X-ray irradiation

The role of P-cadherin in radioresistance was tested in Chapter V. Normal and cancer cells were plated in mammosphere culture conditions and immediately irradiated with 2Gy or 4Gy, respectively. Irradiations were performed using a 320 kV x-ray system (Gulmay Medical Ltd, Camberley, UK). The machine was operated at 300 kV, 10 mA, with filtration fitted in the x-ray beam to give a radiation quality of 2.3 mm Cu half-value layer. Samples were positioned at a distance of 500 mm from the x-ray focus, where the dose rate was determined to be 1.37 Gy/min.

***In vivo* assessment of P-cadherin tumourigenic capacity**

In Chapter V, the basal-like cell line MDA-MB-468 was sorted according to P-cadherin expression into two subpopulations: P-cad^{high} and P-cad^{low} fractions, as described above. The sorted cells were xeno-transplanted at varying dilutions (10^6, 10^5, or $5x10^4$ cells in 100 µl DMEM cell suspension) into the subcutaneous region, under the left abdominal mammary fat pad of 4-5 weeks old female N:NIH(s)II:nu/nu nude mice, using

a 25-gauge needle. Mice were maintained and housed at IPATIMUP Animal House, sited at the Medical Faculty of the University of Porto, in a pathogen-free environment, under controlled conditions of light and humidity. Animal experiments were carried out in accordance with the European Guidelines for the Care and Use of Laboratory Animals, directive 2010/63/UE. Mice (4 per group) were weighted, and tumor width and length were measured with calipers every week. Tumor volume was estimated by using the equation, $V = 0.5 \times a \times b^2$, where V is volume, a is the length of the major axis of the tumor, and b is the length of its minor axis. Mice were euthanized 3 months after tumor cell inoculation.

Adhesion assay to ECM substrates

Cell adhesion assay was performed (Chapter VI) in 96-well microtiter plates coated with laminin 332 (Sigma, St. Louis, MO), fibronectin (Sigma), vitronectin (BD Biosciences, San Diego, CA), type-I or -IV collagen (Sigma) (5µg/ml) overnight at 4ºC. Subsequently, plates were washed three times in PBS and non-specific-binding sites were blocked by adding 0.5% BSA (w/v) in PBS containing Pen/Strep (Invitrogen) for 2h at 37ºC. Once washed again with PBS, 100µl of cells (10^6 cells/ml) were seeded in serum-free medium for 20 min (for MDA-MB-468 cell line) or 30 min (for BT-20 cell line). Thereafter, the plates were washed with PBS to remove non-adherent cells, and the attached cells were fixed with acetone:methanol (1:1) for 10 minutes at 4ºC. Cell adhesion was determined following the colorimetric method described by Busk (Busk *et al.*, 1992). The absorbance was measured at 570nm with a microplate reader. The attachment of cells to wells coated with 1mg/ml of poly-L-Lys (Sigma) and fixed with 4% paraformaldehyde before aspiration was defined as 100% of adhesion.

Invasion assay

Matrigel invasion assay was performed in Chapter VI according to manufacturer's instructions (BD Biosciences). Briefly, transwell chambers with polycarbonate membrane filters (6.5 mm diameter, 8 µm pore size) were coated with 20 µL of a Matrigel solution. 30,000 BT-20 cells or 50,000 MCF7/AZ cells were added to the upper compartment of the chamber. The lower compartment was filled with DMEM medium supplemented with 10% FBS and 1% antibiotic solution (penicillin–streptomycin) (Invitrogen). After 24 or 48 hours of incubation (BT20 or MCF7/AZ, respectively) at 37°C, 5% CO_2, the upper surface of the filter was washed with serum-free DMEM and cleared from nonmigratory cells with a

cotton swab. The remaining (invasive) cells at the lower surface of the filter were fixed with cold methanol and stained with 4′, 6-diamidino-2-phenylindole (Sigma, 0.4 mg/mL). Invasive cells were scored by counting the whole filter with a fluorescence microscope, at 200x magnification.

Immunofluorescence microscopy

In Chapter IV, double immunofluorescence staining with CD44 and CD24 antibodies was performed to evaluate the number of CSCs (CD44+CD24-/low) present in the human carcinoma cases. Using the same primary antibodies used in the IHC analysis (see above), we performed double immunofluorescence in 10% of all breast cancer cases present in the TMAs, as well as in whole tissues. Detection of the primary antibody anti-CD44 was performed using the secondary antibody goat anti-mouse IgG (H+L) conjugated with Alexa Fluor 488 (Invitrogen) and the detection of the anti-CD24 was done using the secondary antibody goat anti-mouse IgM (μ chain) Alexa Fluor 594 (Invitrogen).

The immunofluorescence technique was also used to evaluate integrins expression, focal contact formation and actin filaments distribution in the basal cell lines used in Chapter VI. Thus, the cell lines BT-20 and MDA-MB-468 were seeded on top of glass coverslips coated with laminin-332 (Sigma). Cells were fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X-100 and blocked with 2% BSA before staining. The following primary antibodies were used for immunofluorescence: FITC-conjugated α6 integrin (CD49f) (1:10, BD Biosciences), PE-conjugated β4 integrin (1:10, CD104) (BD Biosciences) and p-FAK tyr397 (1:200 dilution, Cell Signalling). To visualize p-FAK, anti-rabbit Alexa-488 (1:1000 dilution, invitrogen) was incubated on slides for 30 minutes. F-actin was detected by staining with phalloidin conjugated to rhodamine (Invitrogen) at a dilution of 1:1000. Cells were visualized using a Zeiss Imager Z.1 microscope (Zeiss, Welwyn Garden City, UK). Representative photos were acquired using the associated software, and edited in Photoshop and Illustrator (both CS4; Adobe).

Statistical analysis

Immunohistochemical association between different categorical variables (IHC markers, clinicopathological parameters, or the different molecular subtypes) were

assessed by Pearson correlation and chi-squared tests. Survival analyses were estimated using the Kaplan-Meier method and compared using the log-rank test. A maximum cut-off value of 120 months (10 years) was considered. For multivariate survival analysis, Cox regression models were fitted to estimate hazard ratios (HR) and the corresponding 95% confidence interval.

Mammosphere forming ability and growth in 3D cultures (size and number) was compared using two-tailed unpaired t-test.

Statistical analyses were carried out using SPSS statistics 17.0 software (SPSS Inc., Chicago, IL, USA) or Prism GraphPad (La Jolla, CA) and a significant level of 5% was considered.

Flow cytometry data was analysed with the Flowjo software package (TreeStar, Ashland, OR, USA).

CHARACTERIZATION OF CANCER STEM CELL MARKERS AND STEM CELL PROPERTIES IN A SERIES OF HUMAN CANCER CELL LINES AND INVASIVE CARCINOMAS

CHAPTER IV

Published article related to this chapter:

Sara Ricardo, André Filipe Vieira, Renê Gerhard, Dina Leitão, Regina Pinto, Jorge F. Cameselle-Teijeiro, Fernanda Milanezi, Fernando Schmitt, Joana Paredes. **Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtypes**. *Journal of Clinical Pathology*, 64:937-46, 2011.

1. Introduction

In breast cancer, CSCs have been identified in freshly dissociated cell preparations by means of certain cell-surface markers that are often related to adhesion processes. The first report identifying and isolating tumourigenic CSCs from non-tumourigenic breast cancer cells used the combined expression of two cell surface markers: CD44 and CD24 (Al-Hajj *et al.*, 2003). CD44 is a glicosylated type 1 transmembrane protein, constituting the receptor for hyaluronic acid being involved in cell-cell and cell-matrix interactions. The *CD44* gene is highly polymorphic, possessing numerous alternative splice variants, with its protein undergoing extensive post-translational modifications. In addition to the standard isoform (CD44s), the isoform v6 (CD44v6) is expressed in various human carcinomas, including breast cancer (Snyder *et al.*, 2009). The CD24 antigen (heat stable antigen) is a cell surface sialoglycoprotein, anchored via a glycosyl phosphatidylinositol (GPI) to the cell membrane. It is an adhesion molecule that binds P-selectin and L1. In the pioneering study by Al-Hajj and colleagues, CSCs were identified within the mammary cell population depleted from the hematopoietic and endothelial markers (Lin-), having the phenotype $ESA^{+}CD44^{+}CD24^{-/low}$ (Al-Hajj *et al.*, 2003). Surface expression of CD44 and CD24 have also been used in normal breast tissue to fractionate epithelial subpopulations and to identify and characterize differentiation states and lineages (Shipitsin *et al.*, 2007). $CD24^{+}$ cells have luminal differentiation traits, whereas $CD44^{+}$ cells were enriched in progenitor/stem cell features.

Interestingly, *in vitro* studies revealed an enrichment of the cell populations defined as $CD44^{+}/CD24^{-/low}$ (CSC phenotype) and as $CD44^{-}/CD24^{+}$ in basal and luminal breast cancer cell lines, respectively (Fillmore & Kuperwasser, 2008, Sheridan *et al.*, 2006), being CD44 positively associated with stem cell-like characteristics and CD24 expression related with differentiated epithelial features (Park *et al.*, 2010). Fillmore and collaborators, examined $CD44^{+}/CD24^{-}$ cells in 8 breast cell lines (including normal breast epithelia and tumourigenic cell lines) and found that this subpopulation had certain stem-like features like self-renewal, pluripotency and chemotherapy resistance (Fillmore & Kuperwasser, 2008). Furthermore, Mani and collaborators showed that normal human breast cells have a subpopulation with $CD44^{+}CD24^{-/low}$ phenotype, with mesenchymal traits and increased stem cell activity (Mani *et al.*, 2008). Taken together, there are accumulating studies supporting the notion that $CD44^{+}CD24^{-}$ could be regarded as a stem-like marker in normal breast, breast cancer and cell lines.

Later, an attempt has been made in order to translate this *in vitro* data into primary breast carcinomas (Honeth *et al.*, 2008); however, the clinical and prognostic impact of these markers in breast cancer remains a controversial issue (Abraham *et al.*, 2005,

Shipitsin *et al.*, 2007), demanding additional efforts to better characterize the prognostic value of the CSC phenotype , as well as to find other CSC markers that could better predict breast cancer patient survival.

Aldehyde dehydrogenase-1 (also known as RALDH1, ALDH1 or AHD2) is an enzyme involved in the oxidation of a wide variety of aliphatic aldehydes to the corresponding carboxylic acids, being highly expressed in undifferentiated cells of the hematopoietic system (Kastan *et al.*, 1990). This enzyme converts retinol to retinoic acid during vertebrate embryogenesis and it is also involved in the detoxification of intracellular aldehydes. A flow cytometry method, using the detection of the enzymatic activity of ALDH1, has been successfully applied by Ginestier *et al.* to human breast cancer cell lines, as well as to human tissue samples, to isolate stem and progenitor cell populations (Ginestier *et al.*, 2007). In this method, designated ALDEFLUOR assay, the ALDEFLUOR-positive cells isolated from both normal and tumour human breast had the phenotypic and functional characteristics of mammary stem cells. Furthermore, in breast carcinomas, ALDH1 activity identified the tumorigenic cell fraction capable to self-renew and to differentiate, and serial passages of ALDEFLUOR-positive cells originated tumours recapitulating the phenotypical diversity of the initial tumour (Ginestier *et al.*, 2007). In the same study, the expression of ALDH1, detected by immunostaining, correlated with poor prognosis (Ginestier *et al.*, 2007). Additionally, ALDEFLUOR-positive tumour cells have increased mammosphere forming efficiency with visibly larger colonies, when compared with ALDEFLUOR-negative cells (Deng *et al.*, 2010).

Interestingly, $CD44^{+}CD24^{-/low}$ cells and $ALDH1^{+}$ cells are more frequently found in basal-like than in luminal tumours; however, $ALDH1^{+}$ cells are also commonly found in the HER-2 OE subtype (Ginestier *et al.*, 2007, Korkaya *et al.*, 2008). It was recently shown that ALDH1 can further divide the $CD44^{+}CD24^{-/low}$ cell population into fractions that are highly tumourigenic: $ALDH1^{+}CD44^{+}CD24^{-/low}$ cells were able to generate tumours from only 20 cells, whereas $ALDH1^{-}CD44^{+}CD24^{-/low}$ were not tumourigenic in this same cell density (Croker *et al.*, 2009, Ginestier *et al.*, 2007). Based on this current knowledge, there is evidence to support the idea that the use of CD44 and CD24 cell surface markers in combination with ALDH1 activity is the most accurate method to identify and isolate CSC-like cells within breast cancer populations. However, the overlap between $CD44^{+}CD24^{-/low}$ and high ALDH1 expression in primary tumours is quite small (about 1%) (Ginestier *et al.*, 2007).

In addition to the previously described stem cell markers, integrin molecules have also been used to define cells in the mouse breast with colony-forming ability *in vitro* and mammary repopulating capacity *in vivo*. In this way, α6 integrin (CD49f) and β1 integrin (CD29) have been used in combination with CD24 to define a single cell with the

phenotype $CD24^{med}CD49f^{+}CD29^{+}$ that has stem cell capacity in the murine mammary gland (Shackleton *et al.*, 2006, Stingl *et al.*, 2006). Furthermore, the phenotype $CD24^{+}CD49f^{+}$ has also been associated with cancer stem cell features in *BRCA1* deficient cells, exhibiting *in vitro* regeneration potential (Vassilopoulos *et al.*, 2008).

In the present study, we analysed the expression of the main established breast CSC markers - CD44, CD49f and CD24 and ALDH1 in breast cancer cell lines from distinct molecular subtypes, where the different cancer cell populations expressing these CSC markers were selected by flow cytometry. These features were compared with the results obtained in a large series of invasive breast carcinomas, in order to evaluate their distribution among the different molecular subtypes. In addition, we investigated the correlation between the presence of these markers and breast cancer patient survival.

Breast cancer cell lines are representative of the biological heterogeneity found in human breast cancers

In general, breast cancer cell lines exhibit the substantial genomic, transcriptional and biological heterogeneity found in primary tumours, and can also be divided in molecular subtypes. In fact, in most cases the various molecular classifications published for breast cell lines are fairly consistent and fall within the classifications derived for the tumour samples (Perou *et al.*, 1999, Perou *et al.*, 2000, Sorlie *et al.*, 2001, Sorlie *et al.*, 2003).

Table 1 summarizes the main molecular and phenotypic classifications attributed to the most commonly studied breast cancer lines.

Cell line	Neve *et al.*, 2006	Charafe-Jaufrett *et al.*, 2006	Prat *et al.*, 2010	Sieuwerts *et al.*, 2009	Keller *et al.*, 2011	Tissue source	Tumour type
MCF-7/AZ	Luminal	Luminal	Luminal	Luminal	Luminal 1	Pleural effusion	IDAC
T47D	Luminal	Luminal	Luminal	-	Luminal 2	Pleural effusion	IDAC
SkBr3	Luminal	Luminal	Luminal	HER-2	-	Pleural effusion	AC
BT-474	Luminal	-	Luminal	-	-	Primary tumour	IDAC
MDA-MB 468	Basal A	-	-	Basal-like	-	Pleural effusion	AC
BT-20	Basal A	Basal	Basal	Basal-like	Luminal 2	Primary tumour	AC
BT-549	Basal B	-	Claudin-low	Normal-like	-	Primary tumour	IDAC
MDA-MB 231	Basal B	Mesenchymal	Claudin-low	Normal-like	Mesenchymal	Pleural effusion	AC

Table 1 – Classification of the panel of breast cancer cell lines selected for this study, according to their profile of gene expression defined by Neve et al., Charafe-Jaufrett et al., Prat et al. and Sieuwerts et al. (Charafe-Jauffret *et al.*, 2006, Neve *et al.*, 2006, Prat *et al.*, 2010, Sieuwerts *et al.*, 2009). In a recent study, Keller et al. provided a phenotypic classification according to the cell-surface differentiation state (Mesenchymal – EpCAM-CD49f+, Luminal 1 – $EpCAM^{+}CD49f^{-}$, Luminal 2 – $EpCAM^{+}CD49f^{+}CD24^{+}$, Basal – $EpCAM^{+}CD49f^{+}CD24^{-}$) (Keller *et al.*, 2010). AC, adenocarcinoma; IADC – infiltrating ductal adenocarcinoma;

Throughout this work, we have mainly adopted a classification based in the morphology of the cells and the Charafe-Jauffret *et al.* or Neve *et al.* classification for tumourigenic and normal cell lines. In the classification by Charafe-Jauffrett *et al.,* cells are grouped according to their profile of gene expression into Luminal, Basal and

Mesenchymal, (Charafe-Jauffret *et al.*, 2006), whereas in the Neve *et al.* classification cell lines are grouped as Luminal, Basal B and Basal A (Neve *et al.*, 2006). In the latter classification, the Basal B subtype roughly corresponds to the Mesenchymal subtype attributed by Charaffe-Jaufret *et al.*

Thus, in the present work, tumourigenic cell lines that express the steroid receptors - ER and PgR – and present a luminal phenotype were grouped into the luminal group (MCF-7/AZ and T47D). Cell lines in which the most significant genetic alteration is the amplification of the *ERBB2* gene were included in the HER-2 OE group (SkBr3 and BT-474). Most of the steroid receptor negative cell lines form the basal group. This group was further divided in a basal/epithelial subgroup (MDA-MB-468 and BT-20) and a basal/mesenchymal subgroup (BT-549 and MDA-MB-231), according to their morphology and E-cadherin expression (**Table 2**).

Cell line	Classification used in this work	SR status	E-cad	Invasiveness	Morphology
MCF-7/AZ	Luminal	+	++	Low	Epithelial
T47D	Luminal	+	++	Low	Epithelial
SkBr3	HER-2 OE	-	-	Low	Epithelial
BT-474	HER-2 OE	+	++	Low	Epithelial
MDA-MB 468	Basal / Epithelial	-	++	Moderate	Epithelial
BT-20	Basal / Epithelial	-	++	Moderate	Epithelial
BT-549	Basal / Mesenchymal	-	-	High	Mesenchymal
MDA-MB 231	Basal / Mesenchymal	-	-	High	Mesenchymal

Table 2 – Cell lines selected in this study, classified according to our classification, based in the molecular and phenotypic profiles described in Table 1, invasiveness potential and morphology profile. MCF-7/AZ is a luminal cell line derived from the original MCF-7 cell line, retaining most of its molecular characteristics. The luminal cell lines BT-474 and SkBr3 have amplification of the oncogene Her-2. The basal A cell lines MDA-MB-468 and BT-20 have amplification of the oncogene EGFR; SR - steroid receptor; Her-2 OE – HER-2 overexpressing.

Human breast cancer cell lines have distinct expression of the cell surface biomarkers CD44, CD24 and CD49f

To address if breast cancer cell lines representing the different molecular subtypes of breast cancer could be distinguished phenotypically by the cell surface expression of stem cell markers, we measured by flow cytometry the expression of three biomarkers commonly used to define stem cell and cancer stem cell populations, both in the normal breast and in breast carcinoma. The markers used were CD24, CD44 and CD49f (**Figure 1**).

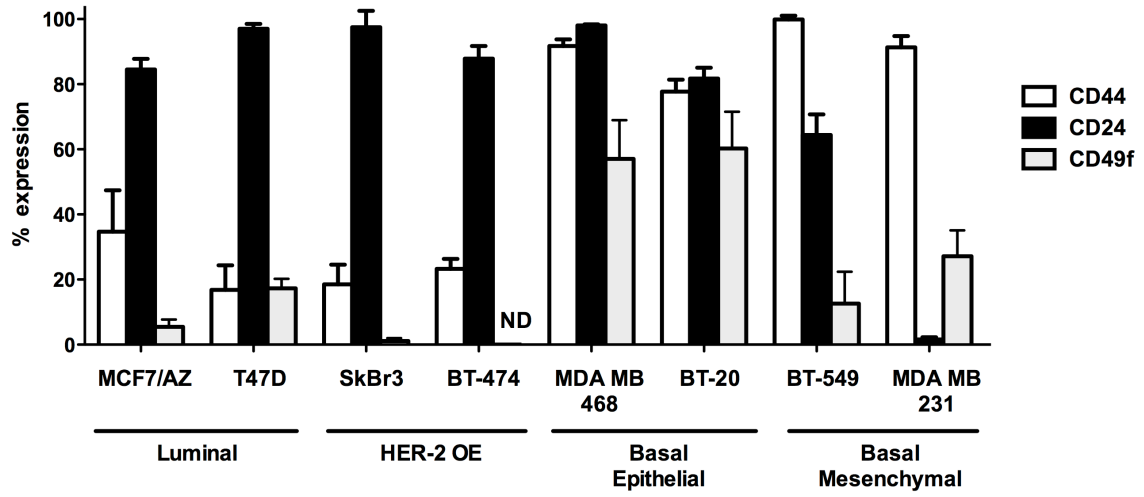


Figure 1 – Expression of commonly used cell surface biomarkers of normal/cancer stem cells of the breast. For each cell line, the % expression of CD44, CD24 or CD49f is represented ± SEM; ND – not determined.

Expression of the marker CD24 was very high (>70%) in all epithelial cell lines, independently of the molecular subtype (luminal, HER2-OE and basal/epithelial). However, in mesenchymal cell lines (basal/mesenchymal) the expression of CD24 was reduced. Notably, CD24 is regarded as a marker of mature luminal cells.

CD44 and CD49f were differentially expressed amongst the panel of cell lines analysed. Expression of CD44 was higher in the basal cell lines (epithelial and mesenchymal), compared to luminal cells. Similarly to the marker CD44, CD49f expression was mainly found in basal cells, particularly within the epithelial phenotype. This result is in accordance with the phenotype found in the normal human breast, in which CD49f is typically seen as a basal marker.

We subsequently asked weather any of the subgroups of cell lines was enriched for the most commonly used CSC phenotype, the $CD44^{+}CD24^{-/low}$ phenotype defined by Al-Hajj and colleagues (Al-Hajj *et al.*, 2003). Therefore, the combined expression of CD44/CD24 was determined by flow cytometry in the breast cancer cell lines. Results are

summarized in **Figures 2 and 3**. In fact, flow cytometry analysis allowed us to divide the populations according to different levels of CD24 and CD44 expression. The size of the subpopulations obtained were in agreement with previously published data (Sheridan *et al.*, 2006). Luminal cells (MCF-7/AZ, T47D) and HER2-OE cells (SkBr3 and BT-474) are mainly constituted by a $CD44^{-}CD24^{+}$ population. Basal cell lines (MDA-MB-468 and BT-20) are mainly double positive $CD44^{+}CD24^{+}$. The mesenchymal cell line MDA-MB-231 is essentially $CD44^{+}CD24^{-}$, which is the putative cancer stem cell phenotype (**Figure 2**).

The CSC phenotype, $CD44^{+}CD24^{-/low}$, has been related to a higher invasion rate in human breast cancer cell lines (Sheridan *et al.*, 2006). In our study, we defined a threshold for CD24 expression according to the levels of the unstained/isotype control and we did not define the "CD24 low" subpopulation. That said, MDA-MB-231 cells show the highest levels of $CD44^{+}/CD24^{-}$ phenotype and this cell line is highly invasive. Basal cell lines (MDA-MB-468 and BT-20), which are moderately invasive, have a small subpopulation defined by the CSC phenotype. Finally, luminal cells (MCF-7/AZ, T47D) and HER-2 OE cell (BT474 and SkBr3) are negative for this phenotype and demonstrate the lowest invasiveness potential (**Figure 2**). Furthermore, there was a striking relationship between the proportion of $CD44^{+}CD24^{-}$ cells and spindle cell morphology, corroborating previous studies (Fillmore & Kuperwasser, 2008, Keller *et al.*, 2010, Sheridan *et al.*, 2006).

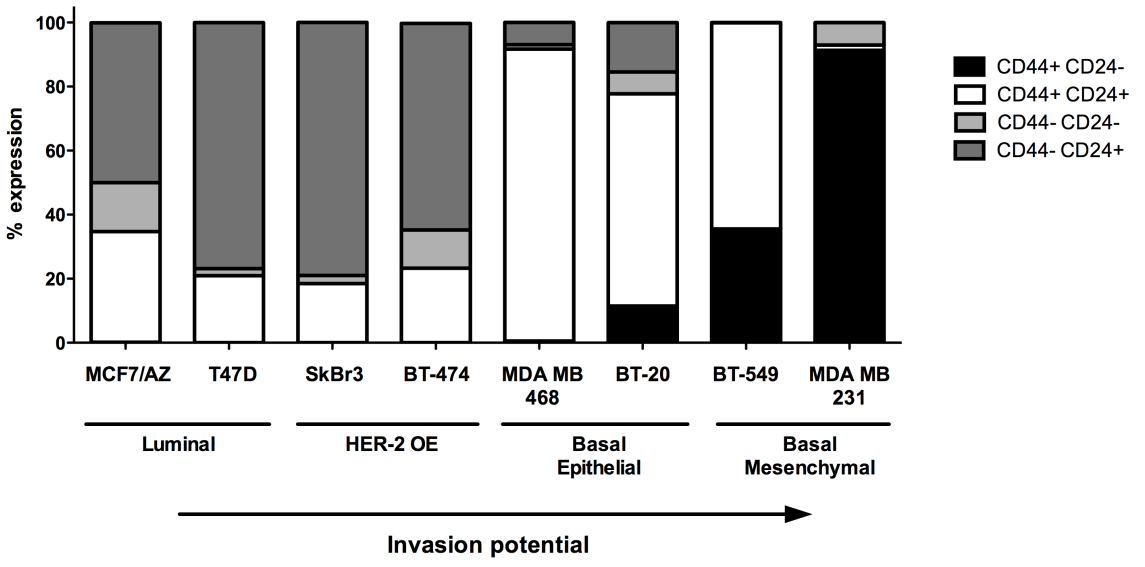
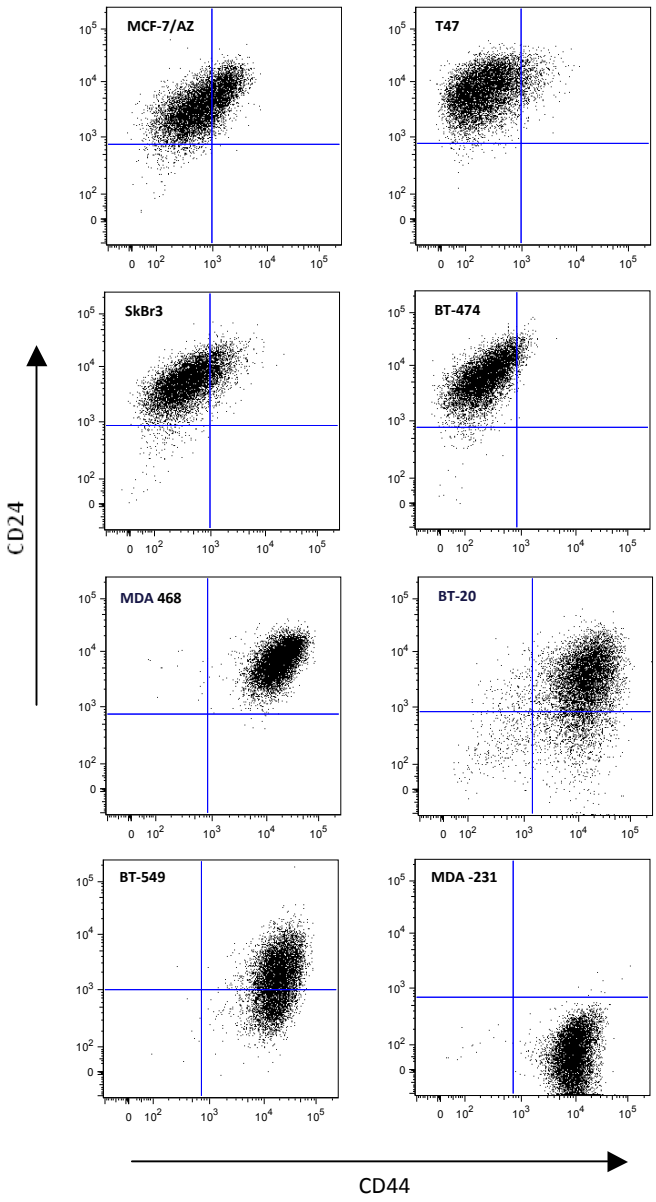


Figure 2 – Graphical representation of the expression of the subpopulations defined by the combination of CD44 / CD24 markers. There is as association of the CSC phenotype (i.e., CD44+CD24-/low) with the mesenchymal phenotype and the invasion potential. Basal epithelial cell lines (MDA-MB-468 and BT-20) are enriched for the phenotype CD44+CD24+.



Cell line	$CD44^{-}CD24^{+}$	$CD44^{+}CD24^{+}$	$CD44^{+}CD24^{-}$	$CD44^{-}CD24^{-}$	Cell type classification
MCF-7/AZ	49.93 ± 9.77	34.57 ± 12.64	0.14 ± 0.05	15.30 ± 3.33	Luminal
T47D	85.50 ± 2.40	10.00 ± 1.22	0	4.42	Luminal
SkBr3	79.03 ± 5.27	18.50 ± 6.07	0	2.50 ± 1.04	HER-2 OE
BT-474	64.52 ± 2.32	23.31 ± 5.99	0	11.91 ± 0.75	HER-2 OE
MDA MB 468	6.92 ± 2.05	91.17 ± 1.88	0.56 ± 0.21	1.38 ± 0.14	Basal / Epithelial
BT-20	15.41 ± 6.20	66.30 ± 2.10	11.45 ± 2.35	6.84 ± 1.77	Basal / Epithelial
BT-549	0	64.3 ± 5.23	35.4 ± 1.99	0	Basal / Mesenchymal
MDA MB 231	0	1.60 ± 0.65	91.35 ± 3.47	7.04 ± 2.83	Basal / Mesenchymal

Figure 3 - Identification of the subpopulations defined by CD44-FITC and CD24-PE expression in breast cancer cell lines by flow cytometry. Isotype controls for IgG2b-FITC and IgG2a-PE were performed (not shown). The percentages of the subpopulations defined by CD44 and CD24 are shown in the table. Indicated is the mean ± SEM of up to three independent experiments

As described above, we characterized our panel of breast cell lines using an additional lineage marker, CD49f, also known as α6 integrin. This marker has been used previously with EpCAM in order to define and characterize cells within the luminal and basal lineages present in normal human breast tissue (Eirew *et al.*, 2008, Keller *et al.*, 2011, Raouf *et al.*, 2008, Stingl *et al.*, 2001, Villadsen *et al.*, 2007). Additionally, some studies have included CD24 to the previous markers to better fractionate the human stem/progenitor compartments (Keller *et al.*, 2010, Lim *et al.*, 2009). Also, in the mouse breast, the combination of CD24 with CD49f was used to define the cell surface phenotype of a single cell that exhibited mammary repopulating ability (Stingl *et al.*, 2006).

To determine which differentiation states were retained in our series of breast cancer cell lines, we analysed by flow cytometry the combined expression of the markers CD49f and CD24. EpCAM expression was not included in this study. Breast cell lines have high levels of EpCAM surface expression, except for MDA-MB-231 cells, which are classified in the basal mesenchymal subgroup (Keller *et al.*, 2010).

The results of flow cytometry analysis of the double staining with CD49f and CD24 are represented in **Figure 4 and 5**. Luminal cell lines have mainly $CD49f^{-}CD24^{+}$ phenotype, basal mesenchymal cell lines have a $CD49f^{-}CD24^{-}$ and basal epithelial cell lines show a prominent $CD49f^{+}CD24^{+}$ phenotype (**Figure 4**).

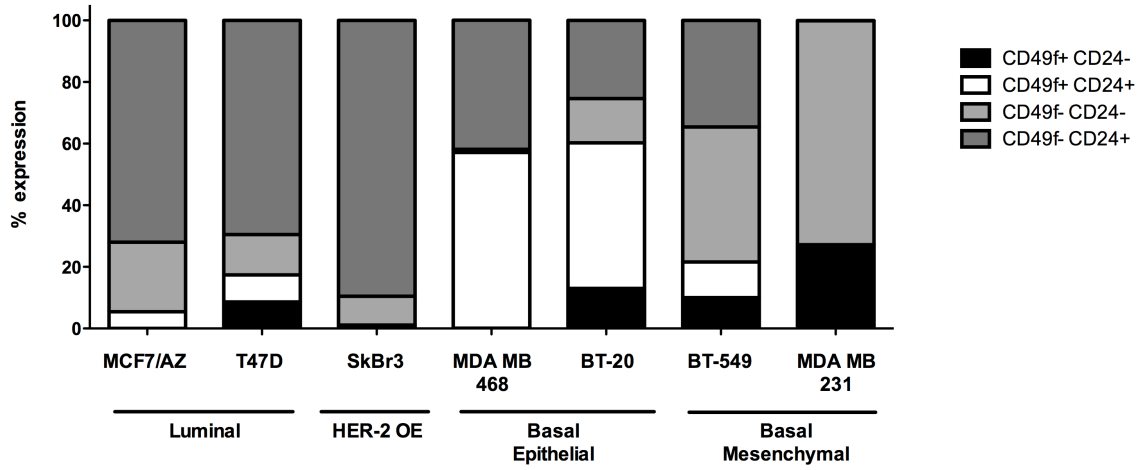
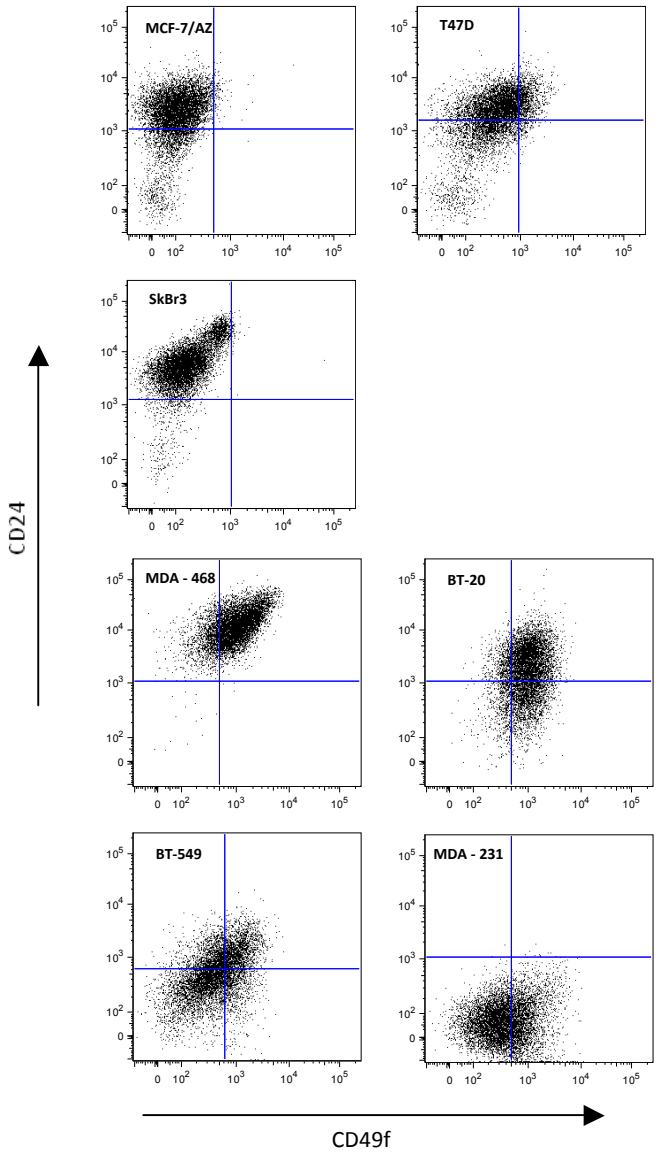


Figure 4 - Graphical representation of the expression of the subpopulations defined by the combination of CD49f / CD24 markers. Basal epithelial cell lines (MDA-MB-468 and BT-20) have mainly the phenotype $CD49f^{+}CD24^{+}$.



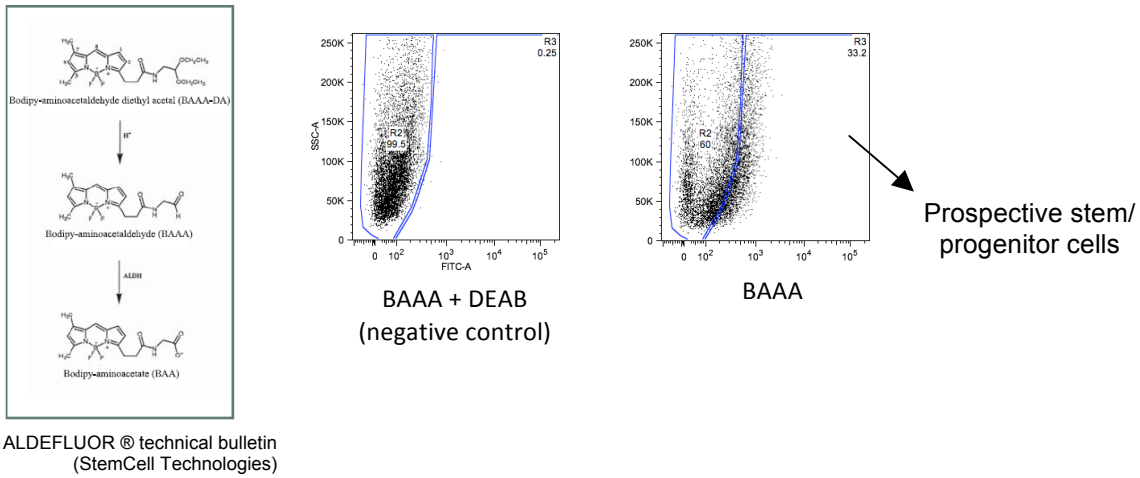
Cell line	$CD49f^{-}CD24^{+}$	$CD49f^{+}CD24^{+}$	$CD49f^{+}CD24^{-}$	$CD49f^{-}CD24^{-}$	Cell type classification
MCF-7/AZ	72.00 ± 3,00	5.46 ± 2,27	0	22.53 ± 2.13	Luminal
T47D	69.5 ± 2.99	8.8 ± 1.52	8.6 ± 1.23	13.1 ± 1.86	Luminal
SkBr3	89.55 ± 4.85	1.09 ± 0.65	0	9.37 ± 4.24	HER-2 OE
MDA MB 468	41.90 ± 11.39	57.08 ± 11,88	0	1.00 ± 0.61	Basal / Epithelial
BT-20	25.38 ± 8.62	47.20 ± 4.35	13.03 ± 7.31	14.39 ± 3.82	Basal / Epithelial
BT-549	34.60 ± 1.20	11.59 ± 3.47	10.01 ± 2.55	43.8 ± 6.75	Mesenchymal
MDA MB 231	0.03 ± 0.01	0.65 ± 0.39	26.53 ± 4.93	72.73 ± 4.59	Mesenchymal

Figure 5 – Identification of the subpopulations defined by CD49f-FITC and CD24-PE expression in breast cancer cell lines by flow cytometry. Isotype controls for IgG2b-FITC and IgG2a-PE were performed (not shown). The percentages of the subpopulations defined by CD44 and CD24 are shown in the table. Indicated is the mean ± SEM of up to three independent experiments.

Aldehyde dehydrogenase (ALDH) activity is increased in HER-2 OE and basal epithelial cell lines

In this study, the activity of the ALDH enzyme was evaluated by the ALDEFLUOR assay in a panel of breast cancer cell lines representative of the different molecular subtypes of breast cancer. A summary of the percentage of the putative stem cell fraction obtained by the ALDEFLUOR assay is presented in **Figure 7**.

A



In: ALDEFLUOR ® technical bulletin (StemCell Technologies)

B

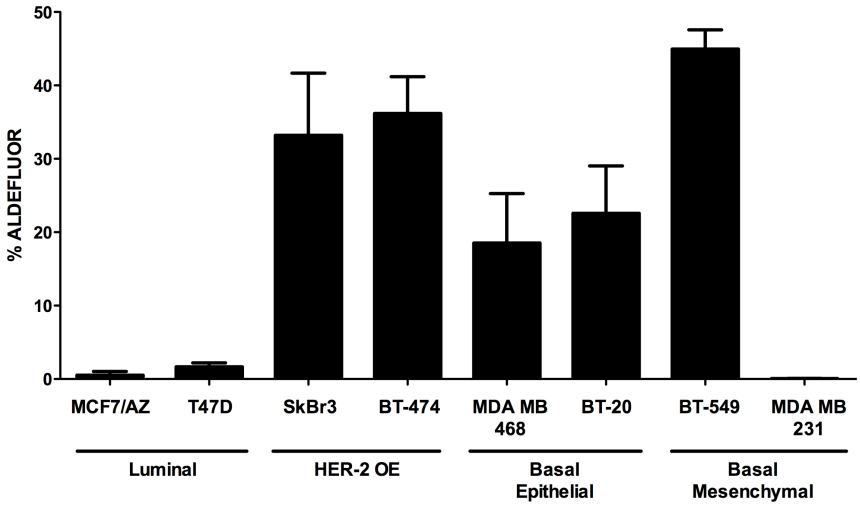


Figure 7 – (A) ALEDFLUOR assay principle and gating strategy to identify the ALDEFLUOR+ subpopulation in breast cancer cell lines (e.g., SkBr3). BAAA – ALDH1 substrate, DEAB – ALDH1 inhibitor; **(B)** Percentages of the ALDEFLUOR positive subpopulation defined by the ALDEFLUOR-assay (StemCell Technologies) in a panel of breast cancer cell lines representative of distinct molecular subtypes of the disease. Plotted is the mean ± SEM of three independent experiments.

The mesenchymal cell line MDA-MB-231, as well as the luminal cell lines MCF-7/AZ and T47D have very low, nearly absent, levels of ADEFLUOR activity. In contrast, the highest levels of the ALDEFLUOR-positive fraction are present in the cell lines SkBr3 and BT-474, which present overexpression of the HER-2 receptor due to gene amplification. Interestingly, Korkaya *et al.* showed that HER-2 overexpression expands the ALDEFLUOR-positive cell population of normal mammary epithelial cells, as well as in tumorigenic breast cancer cell lines, in a mechanism involving the up-regulation of stem cell related genes (Korkaya *et al.*, 2008). The basal epithelial breast cancer cell lines (MDA-MB-468 and BT-20) also exhibit a significant putative stem cell population, present in 18.5% and 22.5% respectively. Furthermore, the cell line BT-549 is classified as ALDEFLUOR-positive, despite being included into the basal mesenchymal subgroup and not having HER-2 amplification. In fact, the behaviour of this cell line regarding the ALDEFLUOR activity as well as the cell surface phenotype shown in the previous section indicated that BT-549 cells are probably in the transition between of the basal epithelial subtype and the basal mesenchymal state.

The survival in anchorage-independent conditions measured by the mammosphere assay is not related to the molecular subtype of breast cancer cell lines

The mammosphere technique has been used to isolate normal human mammary stem and progenitor cells (Dontu *et al.*, 2003). Under non-adherent mammosphere culture conditions, mammary stem and progenitor cells proliferate in an undifferentiated state, whereas differentiated cells die by anoikis. The final stage is a multicellular structure formed by the stem/progenitor cell and the surrounding differentiated cells, which can be dissociated and passaged again. Thus, this technique provides a measure of the self-renewal and differentiation capacity of cells with stem cell activity *in vitro*.

We used this technique to determine the mammosphere forming efficiency (MFE) of our panel of 8 breast cancer cell lines (**Figure 9**). We observed that stem/progenitor cells in breast cell lines represent a small fraction of the whole population, with the MFE ranging from 0.3 to 3%. All cell lines have a variable MFE that does not correlate with the classification into groups considered for our panel of breast cancer cell lines (**Figure 9**).

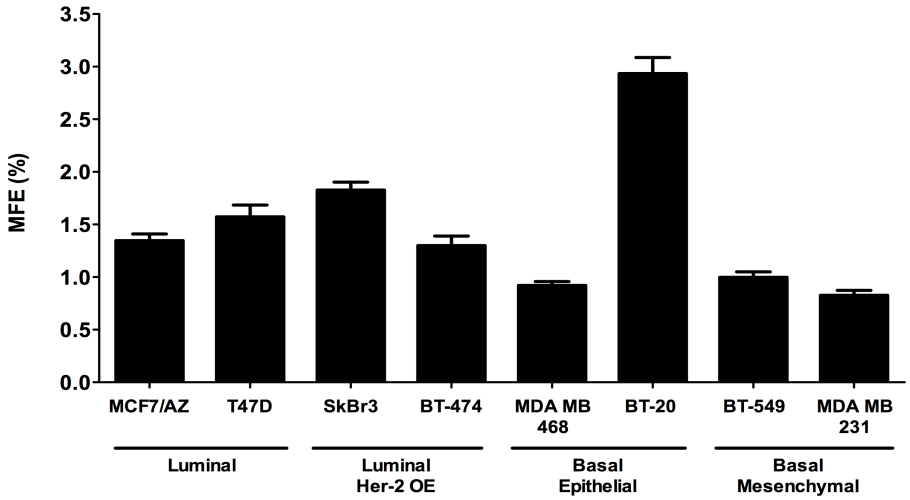


Figure 9 – Percentage of mammosfere forming efficiency (MFE) of eight breast cell lines. Plotted is the % mean ± SEM of up to three independent experiments.

Mammospheres derived from breast cancer cell lines had the capacity to form new generations of mammospheres when passaged as single cells and re-seeded in non-adherent culture conditions. Although a small decrease in MFE was found in the secondary passage, this assay revealed that primary mammospheres contained stem/progenitor cells with capacity for self-renewal.

Mammospheres are multicellular structures that can encompass different morphologies. An example of the mammospheres obtained is depicted in **Figure 10**. Our microscopical analysis showed that mammospheres could be roughly divided into two main groups: a compact spherical structure (MCF7/AZ, T47D, BT-474, BT-549, BT-20) that forms a sphere-like shape that contains in the inside the stem/progenitor cells. This central sphere can be surrounded by differentiated cells that either resisted death by anoikis or that derived from the central structure. On the other hand, some mammospheres have a loose grape-like structure and no central compact aggregate can be distinguished (SkBr3, MDA-MB-468 and MDA-MB-231). From this analysis, we concluded that mammosphere morphology does not correlate with the group classification used in our panel of breast cancer cell lines.

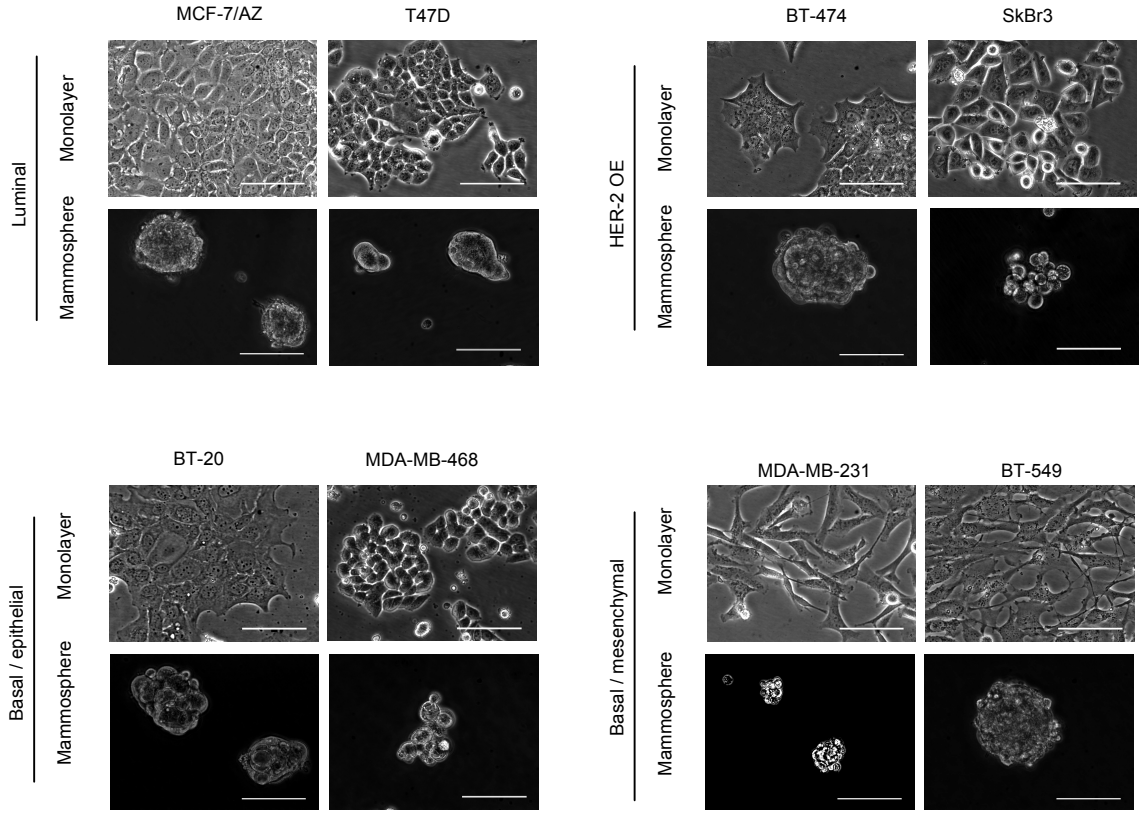


Figure 10 - Morphological characteristics of 2D (monolayer) and mammospheres formed in a panel of breast cell lines. Scale bar=100μm.

Expression of breast cancer cell markers in a series of invasive human carcinomas comprising different molecular subtypes

A series of 466 invasive breast carcinomas were classified into molecular subtypes by immunohistochemical evaluation: 64.8% (302/466) were luminal A, 8.8% (41/466) luminal B, 7.1% (33/466) HER-2 OE, 14.6% (68/466) basal-like tumours and 4.7% (22/466) were unclassified tumours (for details, see Chapter III. Materials and Methods, and Ricardo S *et al* in the appendix section (Ricardo *et al.*, 2011)). This series of human breast tumours has been completely characterized for clinical and pathological features, namely age, tumour size, lymph node status and histological grade and, as expected, the majority of basal-like and HER-2 OE tumours were grade III, highly proliferative and presented the worse patient survival, as seen in the Kaplan-Meier curves (**Figure 11**), demonstrating the validity and power provided by this series of invasive breast carcinomas.

A

Variable (n=466)	Data	
	Frequency	Percentage (%)
Age at Diagnosis		
<50 years	115	26.2
>=50 years	324	73.8
Missing (n= 27)		
Tumour Size		
T1:<2cm	101	24.8
T2: 2-5 cm	244	59.8
T3: >5cm	63	15.4
Missing (n=58)		
Lymph Nodes		
Positive	206	56.4
Negative	159	43.6
Missing(n=101)		
Histological Grade		
Grade I	81	18.3
Grade II	135	30.5
Grade III	227	51.2
Missing (n=23)		
Molecular Subtype		
Luminal A	302	64.8
Luminal B	41	8.8
HER-2 OE	33	7.1
Basal-like	68	14.6
Unclassified	22	4.7

B

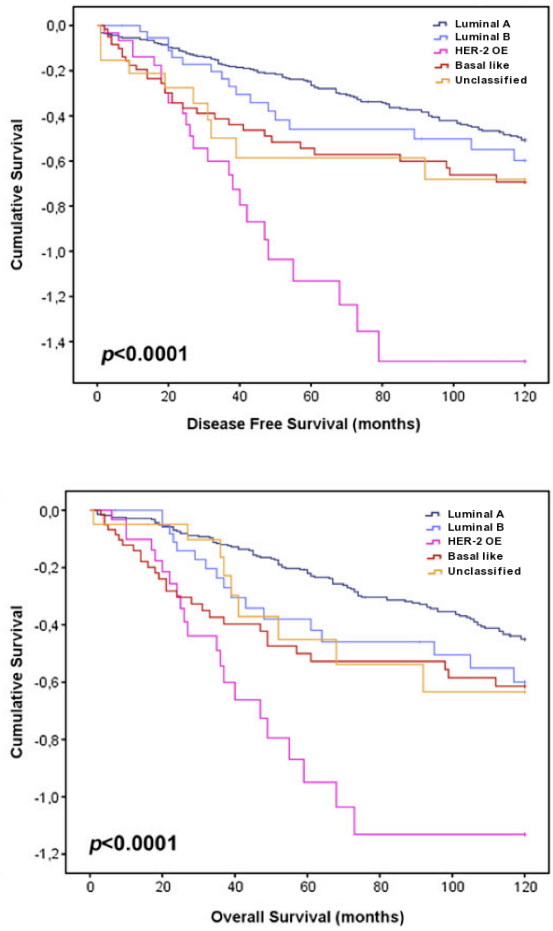


Figure 11 - Patient and tumour parameters of the series of 466 primary invasive breast carcinomas used in this study. The breast tumour signature was determined by immunohistochemistry according to the definition described in Chapter III. *Materials and Methods* **(A).** The disease-free survival and overall-survival of the different breast subtypes is shown **(B).**

This series was studied for the expression of the stem cell markers CD44, CD49f, CD24 and ALDH1 by immunohistochemistry. A representative positive stain for these markers is shown in **Figure 12**.

Concerning CD44 membranous staining, 51.2% (237/463) of the cases were positive. CD49f membranous staining was positive in only 11.5% (49/427) of the invasive breast carcinomas. Regarding the marker CD24, the majority of the cases were classified as negative/low (88.6% - 410/463), and only 11.4% (53/463) of the tumours had clear membrane staining. For ALDH1, a minority of cases were classified as positive (7.1% - 33/464), showing a clear cytoplasmic expression in tumour cells (**Table 3**).

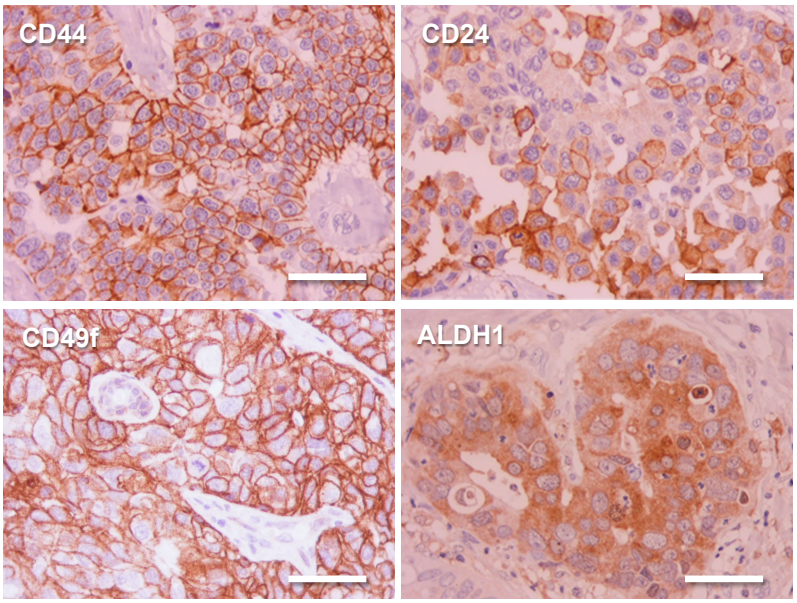


Figure 12 – The expression of the stem cell markers CD44, CD49f, CD24 and ALDH1 was evaluated by IHC in a series of 466 invasive human breast carcinomas. This figure represents a positive case for each of the markers analysed. CD44, CD49f and CD24 were evaluated according to the membranous stain, whereas ALDH1 presents a clear cytoplasmic stain. For the classification criteria of breast carcinomas see Chapter III. *Materials and Methods* (scale bar=150 μm).

Classical prognostic factors	Subtypes					p
	Luminal A (n=302)	Luminal B (n=41)	HER2-OE (n=33)	Basal-like (n=68)	Unclassified (n=22)	
CD44						
Positive (n=237)	147 (49.2%)	17 (41.5%)	12 (36.4%)	56 (82.4%)	5 (22.7%)	***<0.0001***
Negative (n=226)	152 (50.8%)	24 (58.5%)	21 (63.6%)	12 (17.6%)	17 (77.3%)	
Missing (n=3)						
CD24						
Positive (n=53)	34 (11.4%)	7 (17.1%)	5 (15.2%)	4 (5.9%)	3 (13.6%)	
Negative/Low (n=410)	265 (88.6%)	34 (82.9%)	28 (84.8%)	64 (94.1%)	19 (86.4%)	0.418
Missing (n=3)						
CD49f						
Positive (n=49)	11 (4.0%)	6 (15.8%)	2 (6.9%)	27 (42.2%)	3 (13.6%)	***<0.0001***
Negative (n=378)	263 (96.0%)	32 (84.2%)	27 (93.1%)	37 (57.8%)	19 (86.4%)	
Missing (n=39)						
ALDH1						
Positive (n=33)	12 (4.0%)	4 (9.8%)	4 (12.1%)	13 (19.1%)	0 (0%)	***<0.0001***
Negative (n=433)	287 (96.0%)	37 (90.2%)	29 (87.9%)	55 (80.9%)	22 (100%)	
Missing (n=3)						

Table 3 – Association between the expression of P-cadherin, CD44, CD24, CD49f and ALDH1 and the major breast cancer molecular subtypes. There is a statistically significant association between positive expression for P-cadherin, CD44, CD49f and ALDH1 and the basal-like molecular subtype (chi-squared test).

CD44, CD49f and ALDH1 expression were significantly associated with the classification of the breast tumours into molecular subtypes *(p<0.0001)*, whereas CD24 expression was not *(p=0.418)* (**Table 3**). The majority of basal-like carcinomas were considered $CD44^+$ (82.4% - 56/68) and CD49f+ (42,2% - 27/64) (**Table 3**). Furthermore, concerning ALDH1 cytoplasmic expression, 39.4% (13/33) of the cases were classified as basal-like carcinomas (**Table 3**). In addition, almost all basal-like tumours were $CD24^{-/low}$ (94.1% - 64/68) and, amongst the $CD24^+$ cases, 64.2% (34/53) were luminal A (**Table 3**).

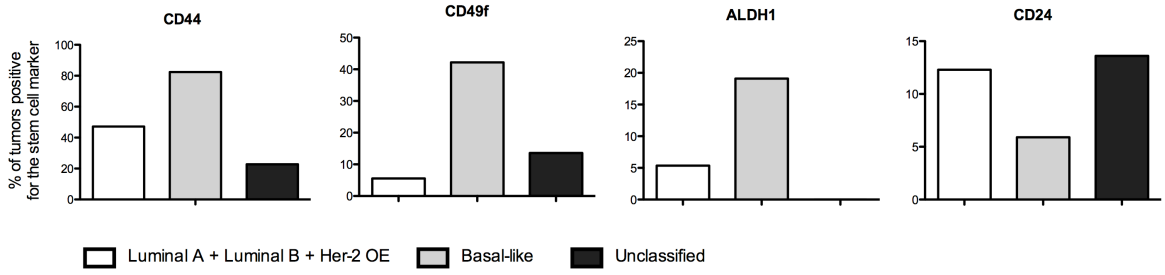


Figure 13 – Graphical representation of the percentage of tumours positive for each of the CSC markers CD44, CD49f, ALDH1 and CD24 in a series of invasive breast cancer cases. Positivity was assessed by IHC as described in the Materials and Methods section. The Luminal A, Luminal B and Her-2 OE tumors were clustered together and compared to the basal-like and unclassified subgroups.

To compare the expression of the stem cell markers in human carcinomas with the expression obtained for the breast cancer cell lines, the basal-like subgroup of breast tumours was put side by side against all the remaining luminal subgroups (luminal A + luminal B + HER-2 OE) and the unclassified group of breast tumours (**Figure 13**). Basal-like carcinomas have a higher percentage of cases positive for CD44, CD49f and ALDH1, and a lower expression of CD24, in comparison with the remaining subgroups. These results are in accordance with the observations found in the breast cancer cell lines analysed. We showed that basal/epithelial cell lines expressed higher levels of the stem cell markers CD44 and CD49f, in comparison with the luminal cell lines (**Figure 1**). Furthermore, ALDH1 activity measured by the ALDEFLUOR assay was significantly increased in the basal/epithelial cells and nearly absent in luminal cell line group (with no HER-2 amplification) (**Figure 7**). Concerning CD24 expression, human breast carcinomas seem to resemble the basal/mesenchymal cells MDA-MB231, in which the expression of this stem cell marker was very low, compared with the remaining breast cancer cell lines.

To explore the effect of the CSC phenotype $CD44^+CD24^{-/low}$ on the clinical outcome and its prevalence in the molecular subtypes of breast cancer, our group evaluated the combined expression of the markers CD44/CD24 by double immunofluorescence staining

in 45 cases of the whole breast cancer series (i.e., corresponding to approximately 10% of all the cases present in the series). We considered a tumour with a CSC phenotype when the frequency of $CD44^{+}CD24^{-/low}$ cells were more than 10%, as previously described in other studies (Abraham *et al.*, 2005, Mylona *et al.*, 2008).

	Subtypes				
	Luminal A (100%)	Luminal B (100%)	Her-2 OE (100%)	Basal-like (100%)	Unclassified (100%)
$CD44^{+}CD24^{-/low}$ ≥ 10%	43.0	41.4	27.3	76.5	18.2
$CD44^{+}CD24^{-/low}$ < 10%	57.0	58.6	72.8	23.5	81.8

p<0001, qui-square test

Table 4 – Distribution of the CSC phenotype ($CD44^{+}CD24^{-/low}$ > 10%) within human breast carcinoma subtypes present in our invasive breast cancer series. Basal-like breast cancers have the highest percentage of cases with more than 10% $CD44^{+}CD24^{-/low}$ cells.

Table 4 shows that the CSC phenotype ($CD44^{+}CD24^{-/low}$) was significantly associated with the breast cancer molecular subtype classifications (p<0.001). Most of the basal-like tumours were classified as $CD44^{+}CD24^{-/low}$ > 10% (76.5%). Both luminal A and luminal B tumours presented similar levels of the CSC-phenotype (43.0% and 41.4%, respectively), and only 27.3% of the HER-2 OE tumours presented more than 10% $CD44^{+}CD24^{-/low}$ cells. Univariate survival analysis (log-rank test) was performed for the four stem cell markers, namely CD44, CD49f, CD24 and ALDH1, as well as the combined expression of CD44/CD24. CD24 expression showed a tendency for a better prognosis in Kaplan-Meier analysis (p=0.073, log-rank test), but CD44, CD49f and ALDH1 expression failed to reach statistically significant levels (log-rank test, p=0.272, p=0.447, p=0.511, respectively), meaning that these markers were not significant predictors of disease-free survival or overall survival in this series of breast carcinomas (data not shown).

Interestingly, our series of breast tumours confirmed the results of other groups, showing that there was no association of the CSC phenotype $CD44^{+}CD24^{-/low}$ with the overall survival of cancer patients (p>0.05, log-rank test) (Mylona *et al.*, 2008). However, when we addressed the expression of CD44/CD24 pattern within the poor prognosis basal-like carcinomas, we found that tumours with a CSC phenotype (≥10% $CD44^{+}CD24^{-/low}$ cells) presented a trend towards a worse disease-free survival (log-rank test, p=0.065) (**Figure 14**).

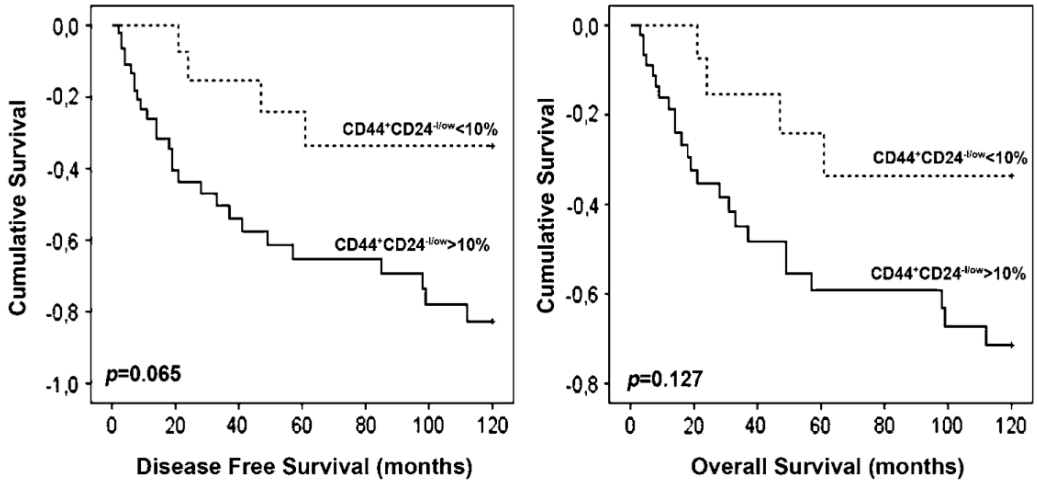


Figure 14 – Kaplan-Meier survival curves representing the disease free survival (log rank test, p=0.065) and overall survival (log rank test, p=0.127) in the basal-like subtype of human carcinomas, according to the CD44/CD24 pattern of expression.

3. Discussion

One of the recent priorities in breast cancer research is the CSC identification/isolation, since it is well accepted that tumours are essentially driven by a cellular pool with stem-like properties, which are responsible for tumour invasiveness, heterogeneity, metastasis capacity, and therapy resistance (Reya *et al.*, 2001). In this study, we analysed a series of breast cancer cell lines from distinct molecular subtypes for the expression of the breast CSC markers CD44, CD49f and CD24, as well as the activity of ALDH1 enzyme, and the survival of the cells in anchorage independent conditions (mammosphere assay). These results were compared with the ones obtained by the immunohistochemical membrane localization of the same panel CSC markers (CD44, CD49f and CD24, as well as the presence of intracellular ALDH1), in a large and well-characterized series of invasive breast carcinomas.

Regarding the characterization of the stem cell markers in cell lines, it is important to establish a comparison with similar studies described in the literature. Keller and collaborators have performed an extensive phenotypic characterization of stem cell surface markers in a panel of breast cancer cell lines which included some of the cell lines used in this study, namely in MCF7, T47D and BT-20 cells. Like in our study, the stem cell markers CD44, CD49f and CD24 were determined by flow cytometry (Keller *et al.*, 2010). Concerning CD24 staining, the results observed in the cell lines are in accordance with the published literature. In fact, all cell lines that maintain an epithelial phenotype showed enrichment in $CD24^+$ cells, whereas the mesenchymal cell lines BT-549 and MDA-MB-231 sowed lower levels or absent expression by flow cytometry of this marker. CD24 is a luminal marker illustrating a clear stain in the luminal cell lines.

Concerning the expression of the stem cell markers CD44 and CD49f, our study shows that in the selected panel of cell lines, there is a correlation between the above mentioned stem cell markers and the attributed subgroups: CD44 and CD49f were significantly more expressed in the aggressive basal epithelial cell lines. In contrast CD44 and CD49f expression was lower in luminal cells, compared to the remaining subgroups. This result contrasts to the study by Keller *et al*, in which no correlation between CD44 or CD49f expression and the subgroups of cell lines was found. This discrepancy could be due to the fact that our study presents a slightly lower expression of the markers CD44 and CD49f. Our study used a different set of fluorochromes to evaluate cell surface expression of these markers, possibly with different gating strategies. Moreover, the study performed by *Keller et al.* is not representative of our panel of cell lines, since it included only 3 out of the 8 cell lines used in this work (Keller *et al.*, 2010). Notably, recent studies

have revealed despite CD44 being expressed on the majority of cells in both the basal and luminal lineages (Raouf *et al.*, 2008), CD44 expression is lower in the luminal compartment (Chaffer *et al.*, 2011, Mani *et al.*, 2008) and in luminal cell lines (Sheridan *et al.*, 2006). Our results for human breast invasive carcinomas show that the expression of the stem cell markers CD44 and CD49f were increased in the poor prognosis basal-like tumours. It has already been demonstrated that $CD44^+$ cells show a mesenchymal stem cell-like profile, enriched for genes involved in cell motility, proliferation and angiogenesis (Shipitsin *et al.*, 2007). CD49f stem cell properties have also been implicated in the aggressive behaviour of the human breast cancer cells, namely in the cell line MCF7/AZ (Cariati *et al.*, 2008). Furthermore, despite our study showing no association of these biomarkers with disease outcome, other groups have shown that breast tumours positive for the stem cell marker CD44 have decreased patient survival (Shipitsin *et al.*, 2007). In our series of primary carcinomas, CD24 membranous staining in breast tumours was not concordant with the previous literature data. Only a small percentage of tumours showed a clear cut membrane positivity. Distinct grading systems have been used to classify CD24 immunohistochemically (Fogel *et al.*, 1999, Surowiak *et al.*, 2006), and consequently, different percentages of CD24 expression have been observed in other series of invasive breast carcinomas. For example, Mylona *et al.* considered mainly membranous CD24 staining, whereas Honeth *et al* considered CD24 staining at the cytoplasm, possibly explaining why different conclusions were drawn by both studies (Honeth *et al.*, 2008, Mylona *et al.*, 2008). Indeed, cytoplasmic expression can reflect aberrant protein overexpression, with consequent disturbance of its membrane distribution and degradation in neoplastic cells (Bircan *et al.*, 2006). Thus its significance to the most appropriate CD24 classification is still ambiguous and needs to be discussed further. Nevertheless, we found that $CD24^+$ cases were enriched in luminal A + luminal B + HER-2 OE tumours (39/53, 73.6%), while the majority of the basal-like tumours were classified as $CD24^{-/low}$ (64/68, 94.1%).

The combinatorial evaluation of CD44/CD24 for the identification of CSC population in breast cancer cell lines was analysed by flow cytometry. The main phenotype found in the basal/mesenchymal cells (of which MDA-MB-231 cells are an example) was the CSC phenotype $CD44^+CD24^-$, whereas the remaining basal cell lines were positive for both markers, $CD44^+CD24^+$. In the basal/mesenchymal cells, the CSC phenotype reflects their metaplastic or claudin low features (Prat *et al.*, 2010) and in the basal/epithelial cell lines, a more differentiated morphology, probably represents the basal-like human carcinomas. Noteworthy, in contrast to the results found in breast cell lines, the majority of basal-like tumours were significantly associated to the CSC phenotype $CD44^+CD24^{-/low}$, as shown by our study and others (Honeth *et al.*, 2008, Park *et al.*, 2010). These studies highlight the

biological heterogeneity of breast cancer and an enrichment of putative tumour-initiating cells in the aggressive basal-like tumour subtype.

In spite of the classical CSC phenotype, the stem cell properties of the $CD44^+CD24^+$ population have also been addressed. These cells were shown to present tumourigenic ability, as well as to have a dynamic switch originating the $CD44^+CD24^{-/low}$ cells (Meyer *et al.*, 2009). It is therefore important to point out that, CD24 expression presents a dynamic regulation. Thus, $CD44^+CD24^+$ cells can readily give rise to $CD44^+CD24^{-/low}$ cells and *vice versa* (Meyer *et al.*, 2009). This may explain why the phenotype of the basal cell lines is not consistent with the phenotype found for human carcinomas, namely concerning CD24 expression.

Similarly, the analysis of the combination of markers CD49f/CD24 revealed that luminal cell lines are enriched in the phenotype $CD49f^-CD24^+$ (MCF-7/AZ, T47D and SkBr3), with the $CD49f^+CD24^+$ cells being the dominant population in basal/epithelial cell lines (BT-20 and MDA-MB-468). The phenotype of the basal epithelial cell lines thus reflects the cell of origin that has been proposed for this malignancy (Lim *et al.*, 2009, Molyneux *et al.*, 2010). Interestingly, the phenotype $CD49f^+CD24^+$ has also been associated with breast cancer stem cell features in *BRCA1* deficient mice, presenting a high regeneration potential (Vassilopoulos *et al.*, 2008).

ALDH1 is an enzyme involved in retinoic acid synthesis and its activity is a biomarker of mammary luminal cells with stem/progenitor activity, constituting a predictor of poor clinical outcome in breast cancer (Eirew *et al.*, 2011, Ginestier *et al.*, 2007). The activity of this enzyme was nearly absent from our set of luminal cell lines. In contrast, the measured activity of ALDH1 enzyme was higher in basal cell lines, with the exception for MDA-MB-231, which showed undetectable ALDH1 activity, as already pointed out by Deng *et al.* (Deng *et al.*, 2010). Furthermore, the $ALDEFLUOR^+$ subpopulation was increased in the highly aggressive cell lines over-expressing the HER-2 oncogene, confirming the results of Korkaya and colleagues (Korkaya *et al.*, 2008). In what concerns the immunohistochemical evaluation of human breast invasive carcinomas, we found that ALDH1 is expressed in 7.1% of all cases. Previous works also detected small percentages of $ALDH1^+$ cases in invasive breast cancer, ranging from 4% to 19% (Deng *et al.*, 2010, Morimoto *et al.*, 2009, Park *et al.*, 2010, Resetkova *et al.*, 2010). Importantly, only the isoform ALDH1a1 was evaluated by IHC. Remarkably, the majority of the cases showing a predominant ALDH1-positive population were significantly associated with basal-like tumours but the survival rate of ALDH1-positive cases did not significantly correlate with poor clinical outcome, like stated in previous studies (Deng *et al.*, 2010, Ginestier *et al.*, 2007, Honeth *et al.*, 2008).

The mammospheres formed from breast cancer cell lines show distinct morphologies that are not associated with molecular subtype. Furthermore, comparing the different methods of assessing the stem cell population in our series of breast cancer cell lines, we could not find an association between the size of the stem cell population measured by the mammosphere assay and the CSC phenotype $CD44^{+}CD24^{-}$ or the cancer stem cell activity determined by the ALDEFLUOR assay.

The identification of stem-like cells in breast cancer cell lines shows that the identified stem cell subpopulations differ according to the cell line analysed and the method used does not necessarily identify the same population of cells. For example, the highly metastatic MDA-MB-231 cells have the phenotypic profile of the CSC, being mainly $CD44^{+}CD24^{-/low}$, however, no $ALDEFULOR^{+}$ population was identified for this cell line. Notably, this indicates that these sets of markers do not represent the same population. Another example is the HER-2 OE cell lines, SkBr3 and BT474, which showed predominance of the $CD44^{-}CD24^{+}$ luminal phenotype, but presented high ALDH activity. Thus, it seems that the different stem cell methods identify cells in different stages of differentiation. For each cell line a set of markers should be explored to better define a CSC phenotype. In certain contexts, a partial overlap of the various subpopulations defined by different methods may be possible. In fact, this has proven to enrich for the cancer stem cell population (Ginestier *et al.*, 2007).

Even though the stem-like features of the ALDEFLUOR + cells and the $CD44^{+}CD24^{-/low}$ cells are increasingly accepted, their location in the stem cell compartment hierarchy is still unclear. We believe that the ALDEFLUOR assay allows the identification of cells in a lower position of the mammary hierarchy, probably late progenitors with a proliferative epithelial phenotype. In contrast, the $CD44^{+}CD24^{-/low}$ phenotype probably identifies the most primitive stem cell or an early progenitor stem cell of the mammary hierarchy, with unlimited self-renewal ability and very low proliferation, where ALDH1 is not active Alternatively, the $CD44^{+}CD24^{-/low}$ phenotype may represent a progenitor-like cell, with a much more active proliferation rate. Stingl mentions that 'the $ESA^{+}CD44^{+}CD24^{-/low}$ phenotype is remarkably similar to the human mammary repopulating unit and bipotent progenitor cell phenotype $ESA^{+}CD49f^{+}MUC1^{-}$, since both CD44 and CD24 have a very similar distribution to CD49f and MUC1, respectively (Stingl, 2009)'. **Figure 15** represents a model for the expression of the stem cell markers CD44, CD24, CD49f and ALDH1 along the breast stem cell hierarchy and the possible association with breast cancer molecular subtypes and breast cancer cell lines.

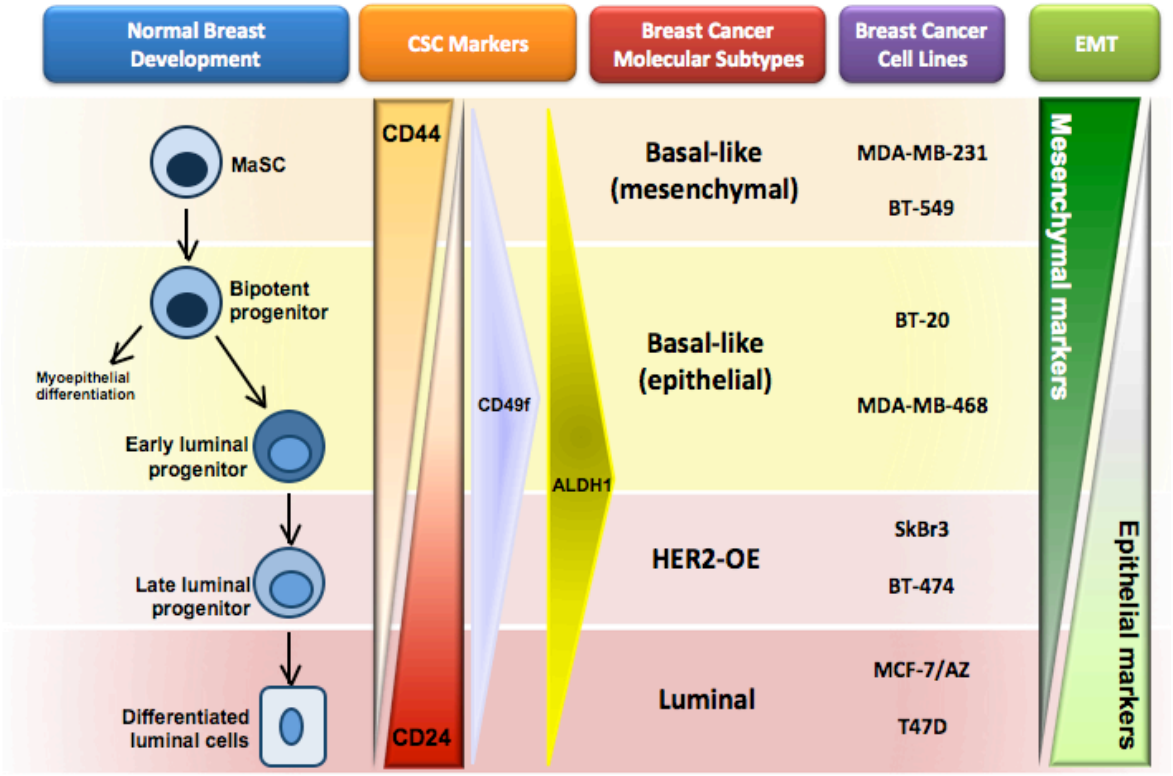


Figure 15 – Expression of the stem cell markers CD44, CD24, CD49f and ALDH1 along the breast stem cell hierarchy and the possible association with breast cancer molecular subtypes and breast cancer cell lines. The phenotype CD44+CD24-/low represents the primitive stem cell and is probably the target of transformation for basal-like mesenchymal tumors (i.e., claudin low tumours). Basal-like epithelial tumours have a luminal progenitor origin with a putative expression of ALDH1, CD24 and CD49f. The targets of transformation of HER2-OE and Luminal breast cancers have not been clearly described.

In summary, the described $CD44^{+}CD24^{-/low}$ and the $ALDH1^{+}$ stem-like phenotypes seem to identify CSCs with distinct levels of differentiation, being the former profile more related with triple carcinomas that most probably originate from the most primitive mammary stem cells, whereas the latter is a marker of basal-like and HER-2 OE tumours, putatively originated from luminal committed progenitors. With this hypothesis in mind, it seems that the paramount method and biomarkers that identify breast CSCs within the distinct molecular subtypes needs to be better explored, since it is pivotal to translate the CSC concept to the clinical practice. In the near future, the recognition of reliable markers to distinguish the CSC pool in each molecular subtype will be decisive for the development of specific target therapies.

P-CADHERIN IS CO-EXPRESSED WITH CD44 AND CD49f AND MEDIATES STEM CELL PROPERTIES IN BASAL-LIKE BREAST CANCER

CHAPTER V

Published article related to this chapter:

André Filipe Vieira, Sara Ricardo, Matthew Paul Ablett, Maria Rita Dionísio, Nuno Mendes, André Albergaria, Gillian Farnie, Renê Gerhard, Jorge F. Cameselle-Teijeiro, Raquel Seruca, Fernando Schmitt, Robert B. Clarke, Joana Paredes.
P-cadherin is co-expressed with CD44 and CD49f and mediates stem cell properties in basal-like breast cancer. *Stem Cells,* 30:854-864, 2012.

1. Introduction

Several studies have shown that solid tumors may contain a small subpopulation of cancer cells that are tumorigenic and have the ability to self-renew and generate all the diverse cancer cells present within the tumor mass. This experimental evidence supports the *cancer stem cell hypothesis* which proposes that a hierarchy exists in the solid tumors comparable to the one found in normal tissue differentiation (Reya *et al.*, 2001). In the breast, these cells are named breast cancer stem cells (CSC). Breast CSCs share important properties with mammary stem cells, namely the ability to proliferate and resist to radiation- and chemotherapy-induced cell death, allowing them to survive and to cause tumor recurrence (Li *et al.*, 2008, Phillips *et al.*, 2006). However, the identification of breast CSCs has been a hard task due to the current technical constraints and the high inter- and intra-tumor heterogeneity observed in breast cancer (Charafe-Jauffret *et al.*, 2009). Most authors make use of cell surface proteins, usually adhesion-related molecules, in an attempt to define a subpopulation of cells that represents the breast CSC population. In 2003, Michael Clarke's group isolated a subset of breast cancer cells with the phenotype $ESA^{+}/CD44^{+}/CD24^{-/low}$, which were able to self-renew and were highly tumorigenic at a low cell inoculum (Al-Hajj *et al.*, 2003). Since then, several other phenotypes/markers to isolate breast CSCs have been described.

For the basal-like breast cancer molecular subtype, in particular, which constitute 10% of all breast cancer cases, few descriptions exist concerning the isolation of their breast CSCs. Meyer *at al*. isolated $CD44^{+}/CD49f^{high}/CD133\text{-}2^{high}$ cancer cells from ER-negative patient tumors, which showed increased tumorigenic ability (Meyer *et al.*, 2010). Wright *et al.* described the phenotype $CD133^{+}$ as able to isolate CSCs from tumors developed in $BRCA1^{-/-}$ mice (Wright *et al.*, 2008). Hwang-Verslues *et al.* characterized the CSC phenotype $PROCR^{+}/ESA^{+}$ for the human basal MDA-MB-231 breast cancer cell line (Hwang-Verslues *et al.*, 2009).

It was recently demonstrated that basal-like breast cancers have a molecular phenotype comparable to the luminal progenitor of the normal breast (Lim *et al.*, 2009, Molyneux *et al.*, 2010), which raised the hypothesis that markers of luminal progenitors would be good CSC markers for basal-like lesions. Additionally, it has been shown that inactivation of BRCA1 gene in the luminal layer of the normal breast originates breast carcinomas in mice that resemble basal-like carcinomas in humans (Molyneux *et al.*, 2010). In fact, *BRCA1* is a major regulator of normal luminal maturation (Liu *et al.*, 2008) and it is essential for the repression of a panel of genes which are typically expressed in basal-like carcinomas of the breast, such as the *CDH3*/P-cadherin gene (Gorski *et al.*,

2009). These evidences suggest that P-cadherin can be an important CSC marker for this type of tumor lesions.

P-cadherin, a classical type I adhesion molecule, is normally expressed in the myoepithelial/basal layer of the breast and is frequently overexpressed in basal-like breast carcinomas (Matos *et al.*, 2005, Paredes *et al.*, 2005, Paredes *et al.*, 2002b). We have found that P-cadherin expression is linked to aggressive tumor behavior, increasing the production of MMPs by cancer cells to the extracellular matrix, as well as inducing cancer cell invasion, migration and motility, due to a mechanism involving alterations in the actin cytoskeleton and signaling through small GTPase-binding proteins (Ribeiro *et al.*, 2010, Albergaria *et al.*, 2011). However, P-cadherin is also involved in homeostatic processes, such as cell differentiation, development, and embryogenesis, illustrating an indirect effect of this adhesion molecule in stem cell biology. P-cadherin deficient female mice present abnormal mammary gland morphology, showing premature differentiation of the breast and increased risk of developing pre-neoplastic lesions, such as alveolar hyperplasia and ductal dysplasia (Radice *et al.*, 1997). In fact, P-cadherin seems to be important in the maintenance of an undifferentiated state in the malignant setting, as breast tumors with P-cadherin expression show loss of cell polarity (Paredes *et al.*, 2002b). The role of P-cadherin in development and differentiation is also seen during embryonic histogenesis, since this cadherin is present in the extra-embryonic ectoderm and visceral endoderm, structures originating the placenta (Hirai *et al.*, 1989a). P-cadherin also has a direct effect in normal stem cells, since it was identified as a stem cell surface marker in human embryonic stem cells (Kolle *et al.*, 2009). Early hair progenitor cells were also isolated as P-cadherin^{+} (and K14^{+}/α6-integrin (CD49f^{+}) cells) (Rhee *et al.*, 2006). Furthermore, the stem cell related transcription factors β-catenin, p63 and C/EBP-β were shown to induce P-cadherin promoter activation (Faraldo *et al.*, 2007, Shimomura *et al.*, 2008, Albergaria *et al.*, 2010). Interestingly, in breast, P-cadherin is found in the cap cells, characteristic stem cells that are the precursors of myoepithelial cells (Daniel *et al.*, 1995, Knudsen & Wheelock, 2005), and in the myoepithelial layer (Paredes *et al.*, 2002a), eventually contributing to the suprabasal stem cell niche.

Herein, we used human mammary cell lines (normal and malignant), as well as a series of invasive breast carcinomas, to provide evidence that P-cadherin expression is important in the cancer stem cell context, not only as a biomarker that better defines the basal-like breast cancer stem cell phenotype, but also as a protein with direct relevance in stem cell activity in this specific molecular subtype.

2. Results

Normal and tumorigenic breast cell lines, with a basal-like phenotype, are enriched for the expression of P-cadherin, as well as for the stem cell markers CD44, CD49f and show an increased ALDEFLUOR-positive subpopulation

P-cadherin is normally expressed in the basal layer of the mammary epithelium (Paredes *et al.*, 2002a) and is frequently up-regulated in basal-like breast carcinomas (Matos *et al.*, 2005, Paredes *et al.*, 2005, Paredes *et al.*, 2007b, Paredes *et al.*, 2002b, Paredes *et al.*, 2002a). In order to study if this protein was associated with the expression of cancer stem cell markers *in vitro*, a panel of breast cancer cell lines was used. This series of cell lines comprised two main subgroups: the estrogen-receptor positive luminal cell lines (MCF-7/AZ and T47D) and the estrogen-receptor-negative basal-like cell lines (MDA-MB-468, BT-20, and BT-549) (Charafe-Jauffret *et al.*, 2006, Neve *et al.*, 2006). In order to demonstrate that this putative association was not restricted to cancer cells, a normal mammary cell line was also included in the study (MCF10A), which was previously described as harboring a basal gene expression signature (Charafe-Jauffret *et al.*, 2006).

In accordance to what is seen in human breast carcinomas (Matos *et al.*, 2005, Paredes *et al.*, 2007b, Paredes *et al.*, 2002a), high/moderate cell surface expression of P-cadherin (>50% of positive cells by FACS) was preferentially observed in basal-like cell lines, in contrast to luminal cell lines, which show lower levels of this protein (**Figure 1A**). The same association was observed for CD44 and CD49f, with increased cell surface expression of these markers in the basal-like phenotype (P-cadherin high/moderate cells), in contrast to the luminal phenotype (P-cadherin low cells) (**Figure 1B**). Other studies are in accordance with our data indicating that, in breast cell lines, luminal cells usually express lower levels of CD44 (Fillmore & Kuperwasser, 2008, Olsson *et al.*, 2011, Sheridan *et al.*, 2006) and CD49f (Neve *et al.*, 2006, Yoon *et al.*, 2001) in comparison with basal cells. Importantly, the later molecules are already well established as cancer stem cell markers of the basal phenotype (Meyer *et al.*, 2010, Wright *et al.*, 2008), as well as putative mammary gland stem cell biomarkers (Lim *et al.*, 2009, Shackleton *et al.*, 2006, Stingl *et al.*, 2006). All the cell lines expressed high levels of CD24, independently of the level of P-cadherin expression.

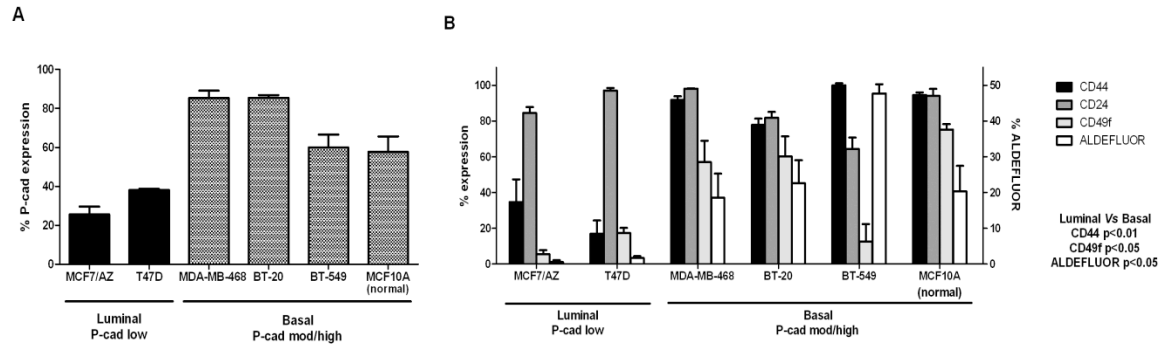
In addition to the cell surface markers, it was previously shown that normal and cancer human mammary epithelial cells with high aldehyde dehydrogenase (ALDH) enzyme activity have stem/progenitor properties (Ginestier *et al.*, 2007). The analysis of ALDH activity in this panel of cell lines revealed that, the P-cadherin$^{mod/high}$ basal cells

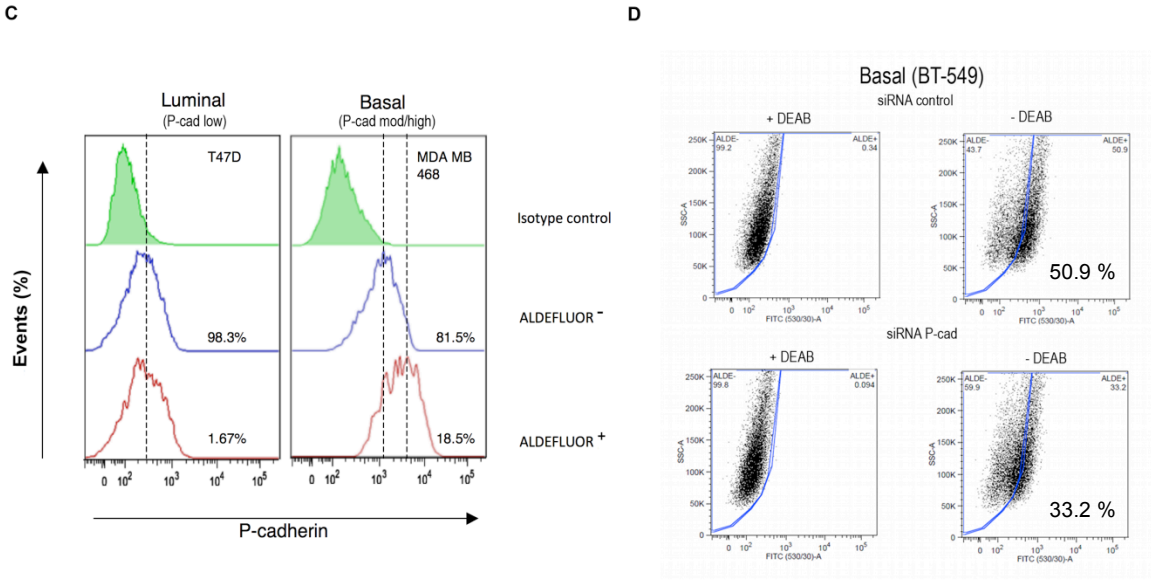
showed a significant subpopulation with this putative stem cell profile (>15%), in contrast to P-cadherin^{low} luminal cells (**Figure 1B**). Again, these results are in accordance with other studies that have shown that basal cells have an increased ALDEFLUOR$^+$ subpopulation, compared with the luminal ones (Ferlay *et al.*, 2010, Marcato *et al.*, 2011).

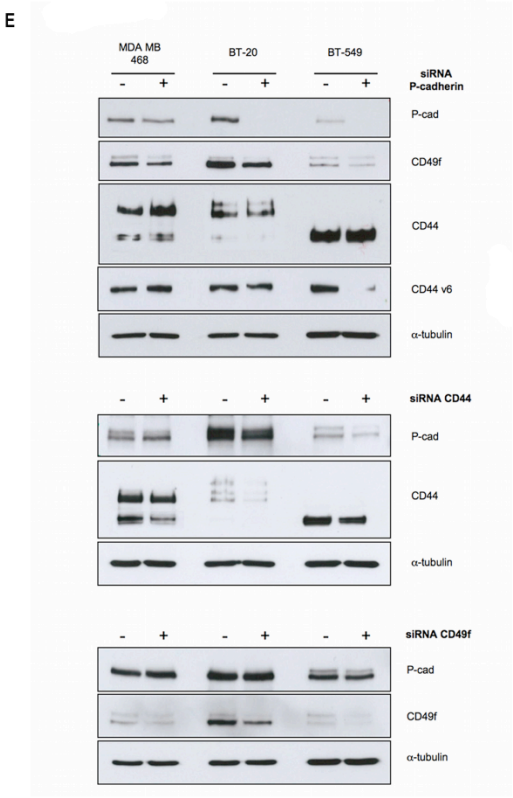
FACS analysis of the ALDEFLUOR$^+$ and ALDEFLUOR$^-$ cell compartments of cell lines from the two main groups, showed that there was an enrichment of P-cadherin expression in the ALDEFLUOR$^+$ stem cell compartment, pointing for a direct association between these two markers. Interestingly, this association was only found in the basal-like group of cell lines, as shown in MDA-MB-468 (**Figure 1C**). In order to more clearly demonstrate the link between P-cadherin and the ALDEFLUOR$^+$ subpopulation, we decided to perform transient inhibition of P-cadherin in the basal cells MDA-MB-468, BT-20, BT-549 and MCF10A. Measurement of ALDH1 activity showed that the inhibition of P-cadherin leads to a decrease in the number of ALDEFLUOR$^+$ cells in all cell lines, being more significant in MCF10A and BT549 (in the cell line BT-549, the ALDEFLUOR$^+$ subpopulation significantly decreased from 50.9% to 33.2%). These results clearly show an association between P-cadherin and the stem cell pool (**Figure 1D**).

The inhibition of P-cadherin also clearly impacted in the expression of the stem cell marker CD49f. In all basal-like cell lines analyzed, P-cadherin knock-down was accompanied by a reduction of CD49f expression (**Figure 1E**). Expression of CD44 was also affected, namely the CD44v6 isoform, with a reduction found in the cell line BT-549 after P-cadherin inhibition (**Figure 1E**). We also performed transient knock-down of CD44 or CD49f and studied the impact of these stem cell markers on P-cadherin expression. Inhibition of CD44 caused a slight reduction of P-cadherin expression in the BT-549 cell line and inhibition of CD49f did not affect P-cadherin expression in the three basal cell lines (**Figure 1E**).

Figure 1 - FACS analysis of cell surface markers and ALDEFLUOR in breast cell lines. P-cadherin is highly expressed in mammary basal-like cell lines (A). In these cells expression of the stem cell markers CD44 and CD49f was found to be increased (*Vs.* luminal cells). Furthermore, the stem cell population, defined by the ALDEFLUOR assay, is also expanded in the basal-like cells (B). In addition, basal cells have an ALDEFLUOR-positive population enriched in P-cadherin expression (C). P-cadherin levels were measured by flow cytometry within the ALDEFLUOR$^+$ and ALDEFLUOR$^-$ subpopulations (percentages of the ALDEFLUOR subpopulations are represented). The basal-like cell line MDA-MB-468 showed a stem cell population (ALDEFLUOR$^+$) enriched in P-cadherin expression, whereas the same does not occur in the luminal cell line T47D. (D) ALDEFLUOR assay was performed after P-cadherin inhibition (siRNA P-cad) in the breast basal cell lines MCF10A, MDA-MB-468, BT-20 and BT-549 showing a decrease in the ALDEFLUOR$^+$ fraction (for which BT-549 is shown as an example). (E) Using transient knock-down of P-cadherin in a panel of basal-like cell lines, the expression of the stem cell markers was evaluated by western-blot. Conversely, the impact on P-cadherin expression was evaluated after transient inhibition of the stem cell markers CD44 and CD49f.


A
% P-cad expression
MCF7/AZ
T47D
MDA-MB-468
BT-20
BT-549
MCF10A (normal)
Luminal P-cad low
Basal P-cad mod/high
B
% expression
% ALDEFLUOR
CD44
CD24
CD49f
ALDEFLUOR
Luminal Vs Basal
CD44 p<0.01
CD49f p<0.05
ALDEFLUOR p<0.05


C
Luminal (P-cad low)
Basal (P-cad mod/high)
T47D
MDA MB 468
Isotype control
ALDEFLUOR -
ALDEFLUOR +
98.3%
1.67%
81.5%
18.5%
Events (%)
P-cadherin
D
Basal (BT-549)
siRNA control
+ DEAB
- DEAB
50.9 %
siRNA P-cad
33.2 %


E
MDA MB 468
BT-20
BT-549
siRNA P-cadherin
P-cad
CD49f
CD44
CD44 v6
α-tubulin
siRNA CD44
siRNA CD49f

P-cadherin expression is associated with the phenotype of the luminal progenitor from the normal breast differentiation hierarchy: $CD49f^{+}CD24^{+}$

Combinations of the markers CD44, CD24 and CD49f have been used in the literature in order to define subpopulations within cell lines or tissues that have stem or cancer stem cell properties (Cariati *et al.*, 2008, Hutvagner & Zamore, 2002, Al-Hajj *et al.*, 2003, Keller *et al.*, 2011, Keller *et al.*, 2010, Mani *et al.*, 2008, Raouf *et al.*, 2008, Stingl *et al.*, 2006). Specifically, in the cancer stem cell field, the phenotype $CD44^{+}CD24^{-/low}$ is believed to exhibit CSC properties (Al-Hajj *et al.*, 2003) and some authors also suggest that the $CD44^{+}CD24^{+}$ phenotype similarly harbors stem-like properties (Meyer *et al.*, 2009, Rappa & Lorico, 2010). Importantly, CD44 expression in the normal human hierarchy is reported in both luminal and basal lineages by Visvader *et al.* (Visvader, 2009) and Raouf *et al.* (Raouf *et al.*, 2008), possibly with a lower expression in the luminal compartment as seen in breast cell lines. Studies performed by Shipitsin *et al.* with cancerous and normal breast tissue, indicate that cells enriched for CD44 expression represent a more basal estrogen-receptor negative phenotype, with mammary epithelial progenitor-like properties(Shipitsin *et al.*, 2007). Additionally, in the normal human breast, the luminal progenitor is reported to be $CD49f^{+}CD24^{+}$ (Molyneux *et al.*, 2010, Lim *et al.*, 2009, Stingl *et al.*, 2006, Keller *et al.*, 2010) (also $EpCAM^{+}MUC1^{+}CD133^{+}Thy1^{-}CD10^{-}$) (Eirew *et al.*, 2008, Raouf *et al.*, 2008, Stingl *et al.*, 2001, Keller *et al.*, 2011).

In this work, cells were stained with a combination of target proteins, which include P-cadherin, CD44, CD24 and CD49f. Stringent cell sorting was applied to separate and analyze the cell subpopulations with stem-like properties described above within each of the breast cell lines, after gating for P-cadherin expression. Separation of the highest 20% P-cadherin ($P\text{-}cad^{high}$) expressing cell subpopulation from the lowest 20% P-cadherin ($P\text{-}cad^{low}$) expressing cell subpopulation revealed that, in all the cell lines studied, the $P\text{-}cad^{high}$ cell subpopulation was enriched for the stem cell-associated markers CD44, CD49f and CD24, in comparison with the $P\text{-}cad^{low}$ subpopulation, as shown in **Figures 2A and Figure 3A**. Conversely, when parental cells were separated into the four possible subpopulations by CD44/CD24 and CD49f/CD24 expression, the $CD44^{+}CD24^{+}$ and $CD49f^{+}CD24^{+}$ regions were the ones enriched for P-cadherin expression (**Figure 2B and Figure 3B**).

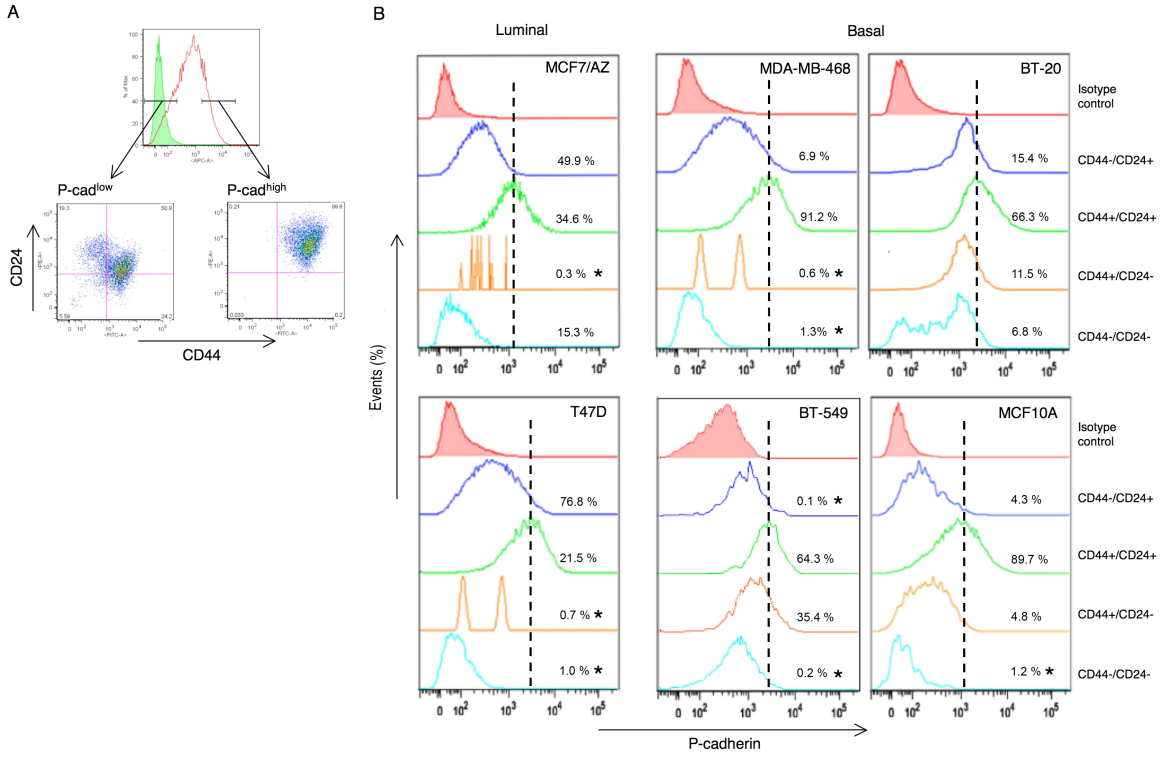


Figure 2 – FACS measurement of the combined expression of CD44/CD24 in a panel of luminal and basal breast cell lines. In all cell lines (MCF10A is represented as an example), P-cadherin^{high} subpopulation (top 20% expressing cells) expressed higher levels of CD24, CD44 and than P-cadherin^{low} subpopulation (lowest 20% expressing cells) (A). Conversely, when analyzing the different quadrants defined by the stem cell markers CD44/CD24, P-cadherin expression was enriched in the CD44^{+}/CD24^{+} subpopulation (B). * indicates ≤ 1,5% of the total cell population.

The population described by Al Hajj *et al.* as the cancer stem cell phenotype, CD44^{+}CD24$^{-/low}$, is decreased in the P-cad^{high} fraction (0.20% *Vs.* 24.2% in the P-cad^{low} fraction, in the example shown in **Figure 2A**). However, the enrichment of the P-cad^{high} fraction for the phenotype CD44^{+}/CD24^{+} (from 50.9% in the P-cad^{low} to 99.6% in the P-cad^{high}) indicates that this adhesion molecule is associated with stem-like properties (Meyer *et al.*, 2009, Rappa & Lorico, 2010). Furthermore, the P-cad^{high} cells are also enriched in CD49f^{+}/CD24^{+} cells (from 57.4% in the P-cad^{low} fraction to 99.9% in the P-cad^{high}) which is evocative of an association with the luminal progenitor profile of the breast (Molyneux *et al.*, 2010, Lim *et al.*, 2009, Keller *et al.*, 2010).

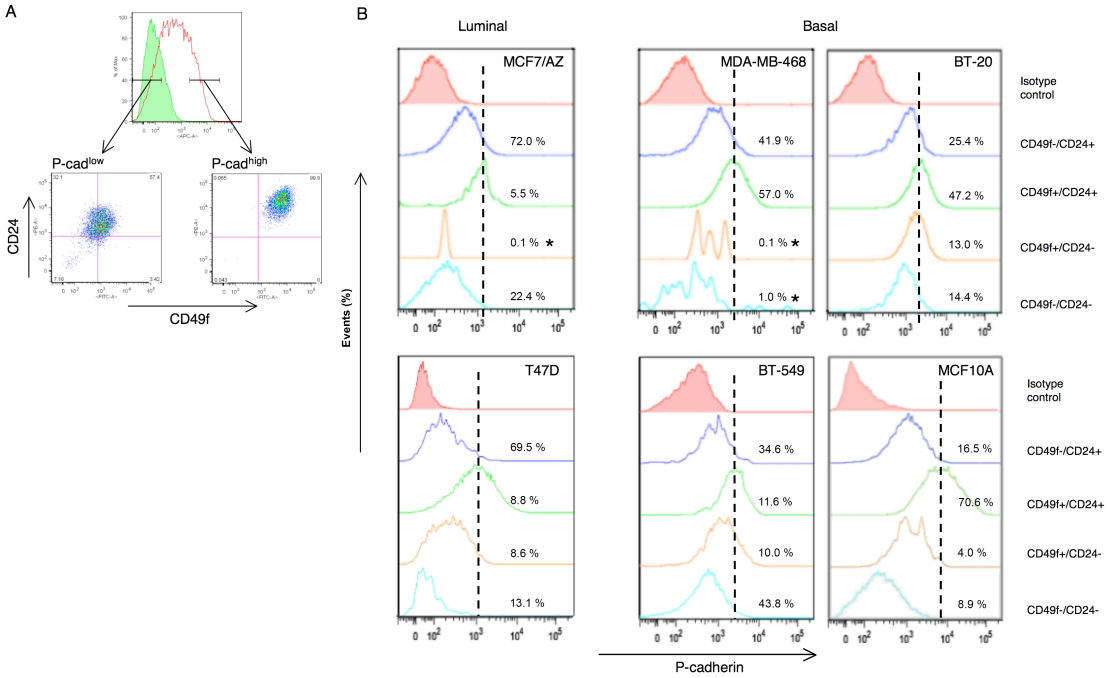


Figure 3 – FACS measurement of the combined expression of CD49f/CD24 in a panel of luminal and basal breast cell lines. In all cell lines (MCF10A is represented as an example), P-cadherin^{high} subpopulation (top 20% expressing cells) expressed higher levels of CD24, CD44 and than P-cadherin^{low} subpopulation (lowest 20% expressing cells) (A). Conversely, when analyzing the different quadrants defined by the stem cell markers CD49f/CD24, P-cadherin expression was higher in the CD49f^{+}/CD24^{+} subpopulation In all cell lines indicated (except BT-20) (B). * indicates ≤ 1,5% of the total cell population.

In human breast carcinomas, P-cadherin expression is associated with poor patient outcome, as well as with the expression of the stem cell markers CD44, CD49f and ALDH1

The expression of P-cadherin and the breast stem cell markers CD44, CD24, CD49f and the isoform ALDH1 were analyzed by immunohistochemistry in a large series of 466 invasive human breast carcinomas. This analysis was performed in order to validate in primary breast carcinomas the association of P-cadherin expression with the stem cell markers previously observed *in vitro*. P-cadherin membrane expression was found in 24.5% of the cases (114/466), 63.2% of them were basal-like carcinomas (**Table 1**). CD44 membrane staining was present in 51.2% (237/463) of the cases. CD49f membrane staining was positive in only 11.5% (49/427) of the invasive breast carcinomas. Regarding the membranous stain of CD24, the majority of the cases (88.6% - 410/463) were classified as negative/low, and only 11.4% (53/463) of the tumors had clear membrane staining. Concerning ALDH1 expression, a minority of cases (7.1% - 33/463) was classified as positive, showing a clear cytoplasmic expression in tumor cells.

Classical prognostic factors	Subtypes					p
	Luminal A (n=302)	Luminal B (n=41)	HER2-OE (n=33)	Basal-like (n=68)	Unclassified (n=22)	
P-cadherin **Positive (n=114)** **Negative (n=352)** **Missing (n=0)**	42 (13.9%) 260 (86.1%)	14 (34.1%) 27 (65.9%)	15 (45.5%) 18 (54.5%)	43 (63.2%) 25 (36.8%)	0 (0%) 22 (100%)	***<0.0001***
CD44 **Positive (n=237)** **Negative (n=226)** **Missing (n=3)**	147 (49.2%) 152 (50.8%)	17 (41.5%) 24 (58.5%)	12 (36.4%) 21 (63.6%)	56 (82.4%) 12 (17.6%)	5 (22.7%) 17 (77.3%)	***<0.0001***
CD24 **Positive (n=53)** **Negative/Low (n=410)** **Missing (n=3)**	34 (11.4%) 265 (88.6%)	7 (17.1%) 34 (82.9%)	5 (15.2%) 28 (84.8%)	4 (5.9%) 64 (94.1%)	3 (13.6%) 19 (86.4%)	0.418
CD49f **Positive (n=49)** **Negative (n=378)** **Missing (n=39)**	11 (4.0%) 263 (96.0%)	6 (15.8%) 32 (84.2%)	2 (6.9%) 27 (93.1%)	27 (42.2%) 37 (57.8%)	3 (13.6%) 19 (86.4%)	***<0.0001***
ALDH1 **Positive (n=33)** **Negative (n=433)** **Missing (n=3)**	12 (4.0%) 287 (96.0%)	4 (9.8%) 37 (90.2%)	4 (12.1%) 29 (87.9%)	13 (19.1%) 55 (80.9%)	0 (0%) 22 (100%)	***<0.0001***

Table 1 – Association between the expression of P-cadherin, CD44, CD24, CD49f and ALDH1 and the major breast cancer molecular subtypes. There is a statistically significant association between positive expression for P-cadherin, CD44, CD49f and ALDH1 and the basal-like molecular subtype.

Variable (n=466)	P-cadherin^{+} (n=114, 100%)	P-cadherin^{-} (n=352, 100%)	p value (Pearson chi-square)
CD44^{+} (n=237) CD44^{-} (n=226) **Missing (n=3)**	72 (63.2%) 42 (36.8%)	165 (47.3%) 184 (52.7%)	***0.003***
CD24^{+} (n=53) CD24$^{-/low}$ (n=410) **Missing (n=3)**	14 (12.3%) 100 (87.7%)	39 (11.2%) 310 (88.8%)	*0.747*
CD49f^{+} (n=49) CD49f^{-} (n=378) **Missing (n=39)**	29 (27.4%) 77 (72.6%)	20 (6.2%) 301 (93.8%)	***<0.001***
ALDH1^{+} (n=33) ALDH1^{-} (n=430) **Missing (n=3)**	17 (14.9%) 97 (85.1%)	16 (4.6%) 333 (95.4%)	***<0.001***

Table 2 – Association between P-cadherin expression and the expression of the stem cell markers CD44, CD24, CD49f and ALDH1, analyzed by immunohistochemistry in a series of 466 primary invasive breast carcinomas. P-cadherin positive cases are enriched in CD44, CD49f and ALDH1 expression (*Vs.* P-cadherin negative cases). No statistically significant association was seen between P-cadherin and CD24 expression.

Importantly, P-cadherin positive cases were significantly enriched for the stem cell markers CD44 (*p=0.003*), CD49f (*p<0.001*) and ALDH1 (*p<0.0001*) and no significant correlation was found between P-cadherin and CD24 expression (*p=0.747*) (**Table 2**). Further, the statistical analysis demonstrated that CD44, CD49f and ALDH1 expression were also significantly enriched in the basal-like carcinomas, in contrast to CD24 (**Table 1**).

Kaplan-Meier survival curves showed that P-cadherin positive tumors were significantly associated with poor overall survival (log-rank, *p=0.023*) (**Figure 4A**). Furthermore, when taken into account the classical prognostic markers tumor size, histological grade and node involvement, the expression of P-cadherin appeared as an independent factor, demonstrating the importance of this protein as a poor prognostic marker in breast cancer (HR=1.486, *p= 0.037*, by multivariate Cox analysis) (**Table 3**).

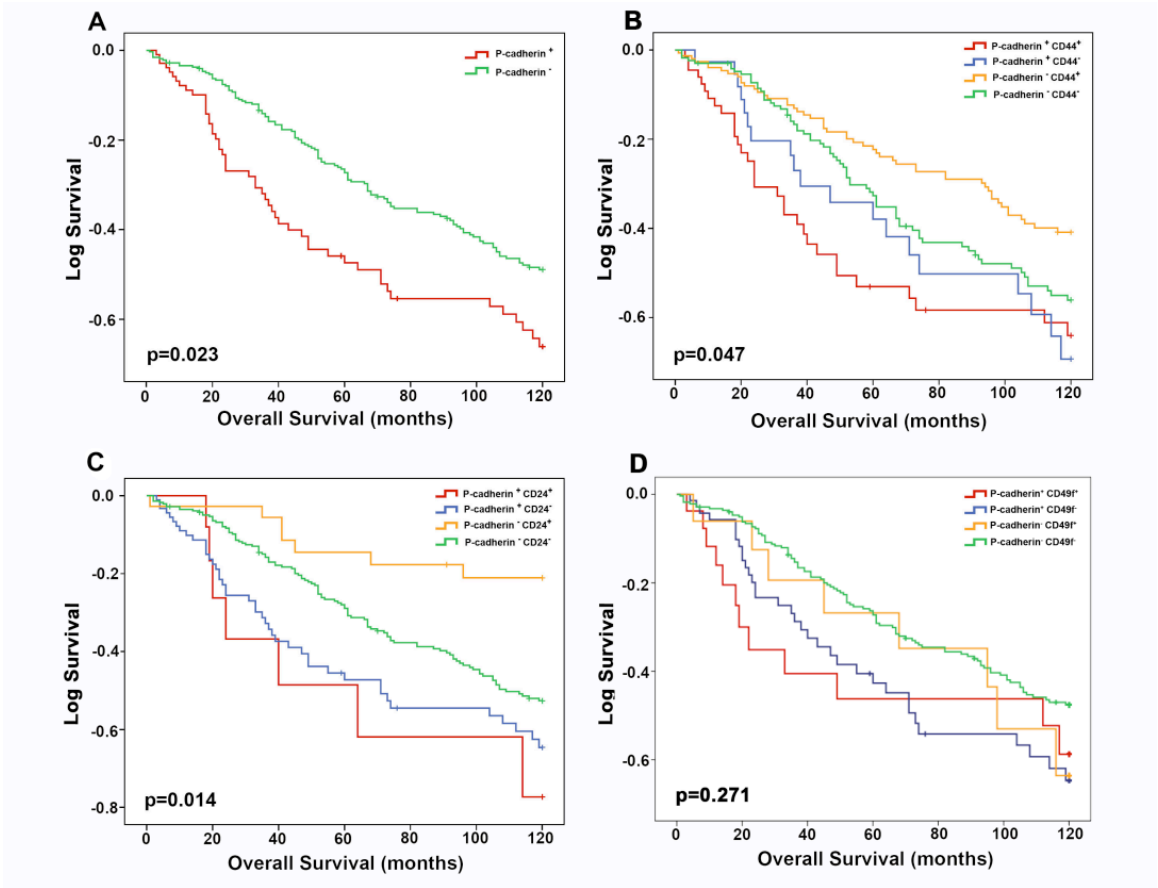


Figure 4 – Overall survival of patients with breast carcinomas (log-rank test), classified according to the expression of P-cadherin (A) and the combined expression of P-cadherin/CD44 (B), P-cadherin/CD24 (C) and P-cadherin/CD49f (D). Expression of CD44, CD24 or CD49f alone has no prognostic value (supplementary figure 2); however, when P-cadherin is combined with these markers, P-cadherin^{+}/CD44^{+} and P-cadherin^{+}/CD24^{+} cases have a worst overall patient survival.

In contrast, the expression of the stem cell markers CD44, CD24, CD49f or ALDH1 alone was not significantly associated with the clinical outcome (**Figure 5**). By multivariate Cox analysis, the expression of CD24 was the only significant stem cell marker influencing survival, in this case related to a better prognosis (HR=0.439, *p= 0.014*) (**Table 3**). Interestingly, however, when P-cadherin and CD24 were combined, the positive expression of both markers was highly associated with the worst patient overall survival (log rank, *p=0.014*) (**Figure 4C**), whereas tumors with a P-cadherin$^-$CD24$^+$ phenotype were associated with a good prognosis. The tumors with combined expression of P-cadherin and CD44 were also associated with a worst patient overall survival (**Figure 4B**) (log-rank, *p=0.047*). Although a strong association between P-cadherin and CD49f expression was found, the combined expression of both these markers were not significantly associated with survival (log-rank, *p=0.271*) (**Figure 4D**).

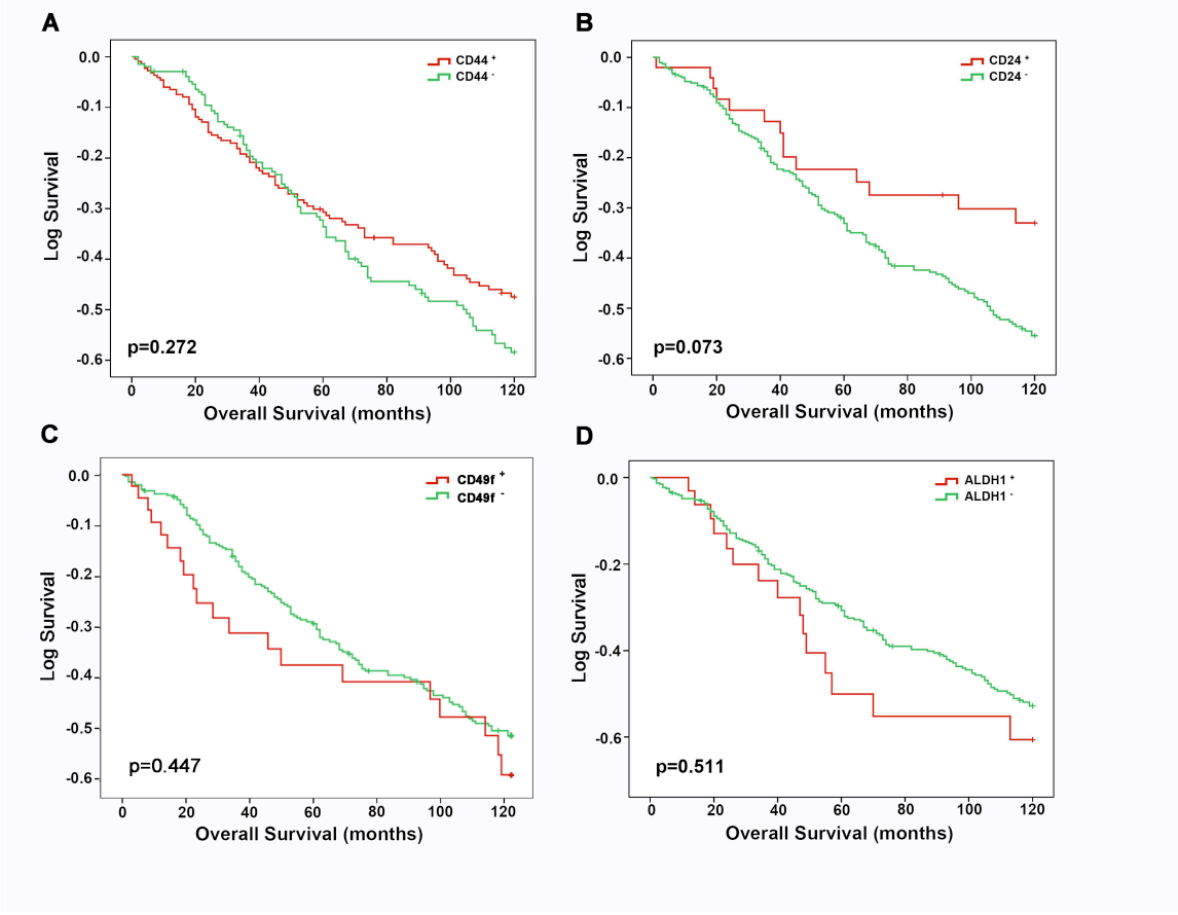


Figure 5 – Overall survival and univariate analysis (log-rank test) of patients with breast carcinomas, according to the expression of CD44 (A), CD24 (B), CD49f (C) and ALDH1 (D). Expression of CD44, CD49f or ALDH1 alone do not show any prognostic value. However, although expression of CD24 alone has not reached significance, there is a tendency for CD24$^-$ cases to have a worse prognosis than CD24$^+$ cases (*p=0.073*).

Variable	Overall survival		
	HR	(95% CI)	p
P-cadherin (positive *vs* negative, ref)	1.486	(1.024-2.155)	***0.037***
CD44 (positive *vs* negative, ref)	1.012	(0.706-1.450)	*0.949*
CD24 (positive *vs* negative, ref)	0.439	(0.228-0.844)	***0.014***
CD49f (positive *vs* negative, ref)	0.819	(0.455-1.474)	*0.506*
ALDH1 (positive *vs* negative, ref)	1.237	(0.536-2.856)	*0.618*

Table 3 – Multivariate Cox proportional hazard analysis (for overall patient survival), with models including the classical prognostic factors in breast cancer (tumor size, nodal status and histological grade), for P-cadherin, CD44, CD24, CD49f and ALDH1 expression. P-cadherin expression appears as an independent factor of worse prognosis, whereas expression of CD24 is an independent factor of good breast cancer patient prognosis.

P-cadherin expression confers increased self-renewal ability, improved cell growth in 3D cultures and radioresistance in breast cancer cell lines

The previous *in vitro* and *in vivo* indications, showing that P-cadherin is associated with the expression of stem cell makers in breast cancer, led us to study the cellular effects mediated by the expression of P-cadherin, namely in stem cell activity.

Stem cell activity can be measured by means of the mammosphere forming ability of a specific cell line/subpopulation. In different cell lines, the heterogeneous parental cell populations were separated by sorting according to P-cadherin levels of expression (high 20% *Vs.* low 20%) and they were studied for stem cell activity. The estimated purity of the sorted fractions was on average 80-95%. A representative flow analysis of the isolated subpopulations is shown in **Figure 6**.

We found that P-cad^{high} cell fractions from basal-like cell lines showed a significant increased mammosphere forming efficiency (MFE) when compared with the P-cad^{low} cell fraction (**Figure 7A**). The MFE from the basal P-cad^{high} fractions reached up to two times the levels of the negative fractions. This result was seen both in the basal tumorigenic cells, as well as in the normal MCF10A breast cell line, suggesting that P-cadherin contributes to the stem cell activity in both normal and malignant contexts.

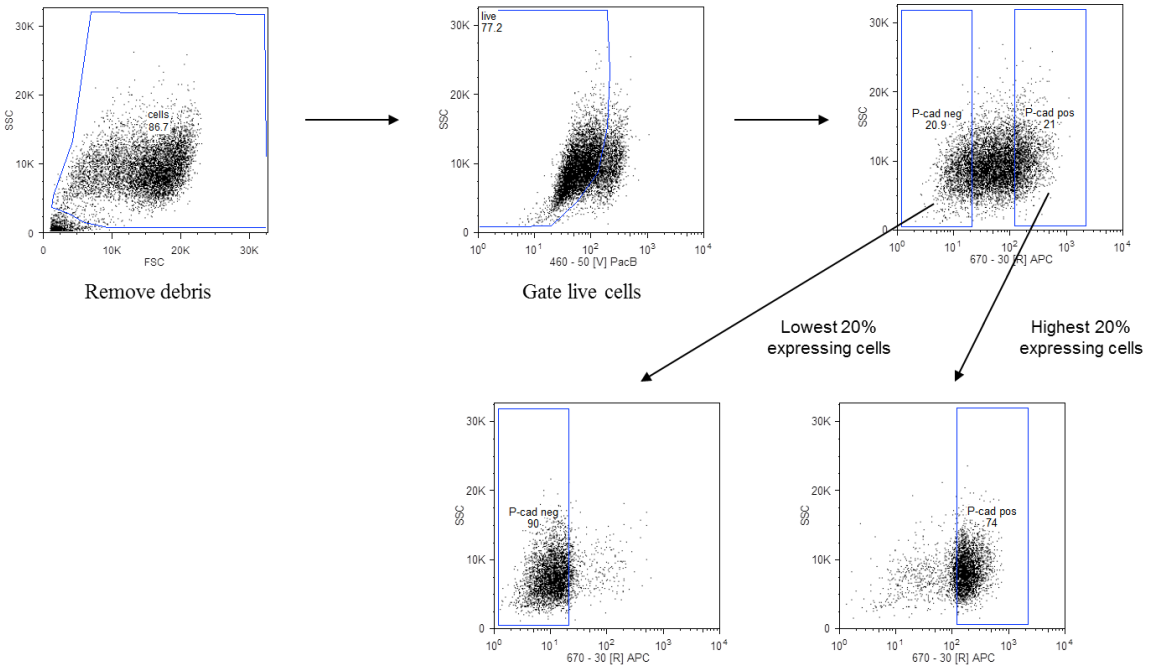


Figure 6 – Sorting strategy of the Pcad^{high} and P-cad^{low} subpopulations using BD Influx or BD FACS Aria-II cell sorters. After sorting, the purity of sorted populations was checked (this diagram represents one sorting experiment performed for the cell line BT-549).

We also studied the clonogenic capacity of P-cadherin positivity in a 3D proliferation permissive environment containing matrigel, a matrix resembling the basal lamina of the normal breast. Using the same sorting procedure, in the luminal MCF7/AZ and the basal BT-549 cells, we found that there was an increase in the number of 3D structures formed by the P-cad^{high} cell subpopulation, compared to the P-cad^{low} subpopulation in the basal-like cell line (**Figure 7B**). Furthermore, we found that the grape-like structures formed by P-cad^{high} cells were bigger than the 3D structures formed by the P-cad^{low} subpopulation (95.8µm *Vs.* 50.1 µm, *p=0.003*, data not shown). The same does not hold true for luminal cells for both parameters analyzed.

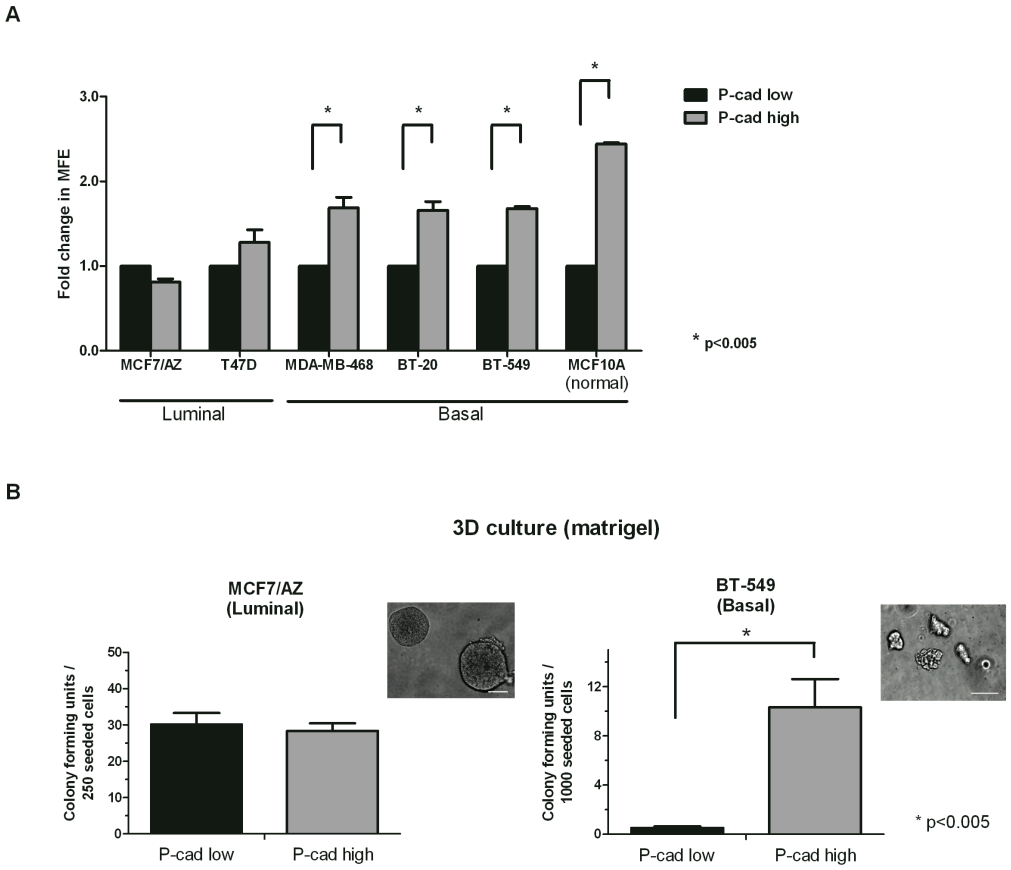


Figure 7 – Mammosphere forming efficiency (MFE) was measured in the subpopulations isolated by FACS according to P-cadherin expression in a panel of human breast cell lines. The isolated subpopulations with higher amounts of P-cadherin (top 20% expressing cells), within the basal-like cell lines, displayed highest MFE **(A)**; 3D cell cultures, in laminin-rich matrix (matrigel), revealed that the P-cadherin^{high} subpopulation from the basal BT-549 cell line has increased clonogenic capacity, whereas the same result is not obtained with luminal cells (MCF7/AZ) (scale bar = 100 μm) **(B).**

To verify the specific role of P-cadherin in stem cell activity, we used genetically manipulated cell lines, by in vitro transduction or silencing of P-cadherin. We showed that stable transduction of P-cadherin in MCF-7/AZ cells led to an increase of mammosphere formation (>25% of MFE) when compared to the mock cell line (**Figure 8A and Figure 8B**). This difference in mammosphere formation persists for at least two passages (data not shown), suggesting a role of P-cadherin in mediating self-renewal or survival. Likewise, transient inhibition of P-cadherin expression by siRNA assays, showed that mammosphere forming ability was negatively affected, though not completely abrogated in most of the cell lines studied (**Figure 8A and Figure 8C**). These results indicate that P-cadherin contributes to the stem cell activity of these cells.

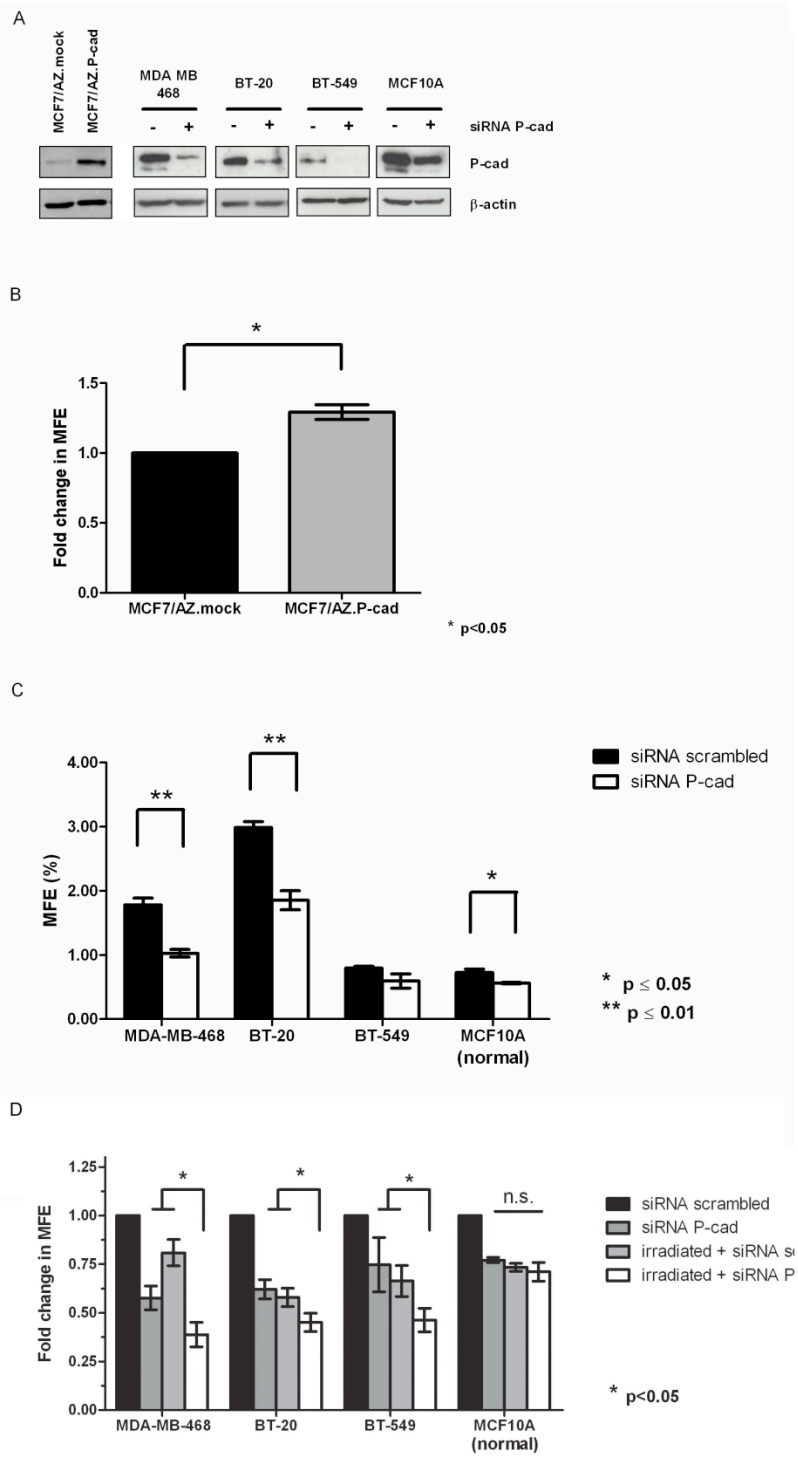


Figure 8 – P-cadherin cDNA was stably transduced and over-expressed in the luminal cell line MCF7/AZ (MCF7/AZ.P-cad) or transiently silenced in the several basal cells lines (A). Measurement of stem cell activity by mammosphere forming efficiency (MFE) indicates a direct association between P-cadherin expression and this stem cell property (B and C); X-ray irradiation was administered to the normal cell line MCF10A (2Gy) or the indicated tumorigenic cells (4Gy) and mammosphere forming efficiency (MFE) was assessed. The decrease in the number of stem cells mediated by X-ray irradiation is potentiated after P-cadherin silencing in tumorigenic cells, but not in the normal MCF10A cell line (D).

The increase in clonogenic capacity in 3D mediated by P-cadherin, and the increase in stem cell activity measured by MFE, can be the result of either increased survival of cells in the particular conditions of these assays or/and due to alterations in the growth rate that affects proliferation/quiescence of the prospective P-cad^{high} stem cells. In an attempt to understand if P-cadherin effects were mediated by changes in proliferative capacity, we evaluated the cell cycle profile in our cell lines, using Hoechst-33342 stain. Analysis of the cell cycle revealed that P-cadherin does not affect the subpopulations S and G_2M (**Figure 9**).

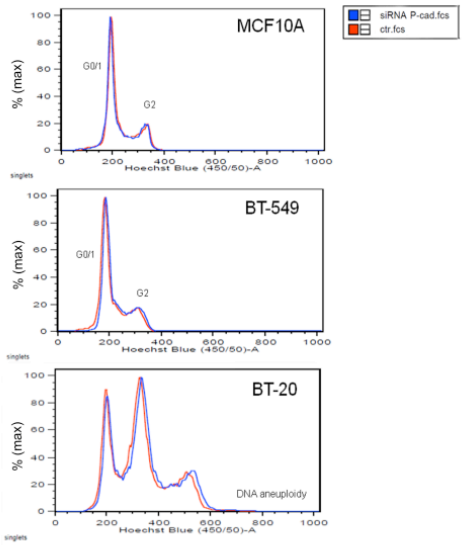


Figure 9 – Hoechst-33342 stain of the cell cycle profile in the basal cell lines MCF10A, BT-20 and BT-549 revealed that inhibition of P-cadherin did not affect the proliferative phase of the cell cycle (G_2M).

One of the features attributed to breast cancer stem cells is the increased resistance to irradiation, which allows them to survive and persist in tissues after treatment, contributing to disease relapse. We investigated the effect of X-ray irradiation alone and in combination with P-cadherin inhibition on the stem cell activity measured by MFE. We found that inhibition of P-cadherin in the tumorigenic basal-like cell lines potentiated the effect of irradiation induced cell death, decreasing the number of mammospheres formed. This same effect was not observed in the normal breast cell line MCF10A, where irradiation treatment or P-cadherin silencing alone had a negative effect in MFE; however, when the two treatments were combined, no additive or synergistic effect was observed concerning stem cell activity (**Figure 8D**).

P-cadherin expression confers increased tumorigenic ability to breast cancer cells

The indications given by *in vitro* functional assays, showing that P-cadherin expression has a role in the maintenance of stem cell properties and poor overall survival of breast cancer patients, led us to test if cancer cell populations enriched for P-cadherin have a higher capacity to promote tumour growth in nude mice. The tumorigenicity of the P-cadherin high and low subpopulations was addressed using the basal-like MDA-MB-468 breast cancer cells, which were inoculated in different dilutions, and the results obtained are shown in **Table 4**. Interestingly, we could observe that cancer cells enriched for P-cadherin expression have a higher capacity to promote tumor growth, since the number of tumors formed was higher in the P-cad^{high} group, compared to the control group, in which the same number of parental cells was injected. Furthermore, tumours derived from the P-cad^{high} cells were larger in size than the tumours derived from the P-cad^{low} fraction: by week 6 of tumour growth the average size of the P-cad^{high} tumors was 73.76 mm^3 *Vs.* 50.01 mm^3 in the P-cad^{low} group, considering the animals in which 5x10^4 cells were injected (**Figure 10A**).

Cell population	Tumors / injection		
	10^6 cells	10^5 cells	5x10^4 cells
unsorted	4/4	2/4	2/4
P-cad ^{high}		2/3	4/4
P-cad ^{low}		0/3	2/4

Table 4 – P-cadherin capacity to promote tumors was evaluated by xenografting P-cad^{high} and P-cad^{low} cells sorted from the basal-like MDA-MB-468 breast cancer cell line in the nude mouse model. The appearance of tumors, as well as the tumor mass volume was measured during time.

The tumours formed had often a central core filled with a mucous substance (**Figure 10B**). A pathologist will better explore the histopathologic features of the tumours derived from the MDA-MB-468 cell line in the future. Interestingly, all animals that developed a tumour had enlargement of the inguinal and peritoneal lymph nodes, indicating that there was regional involvement and raising the need to search for possible metastatic foci in distant sites.

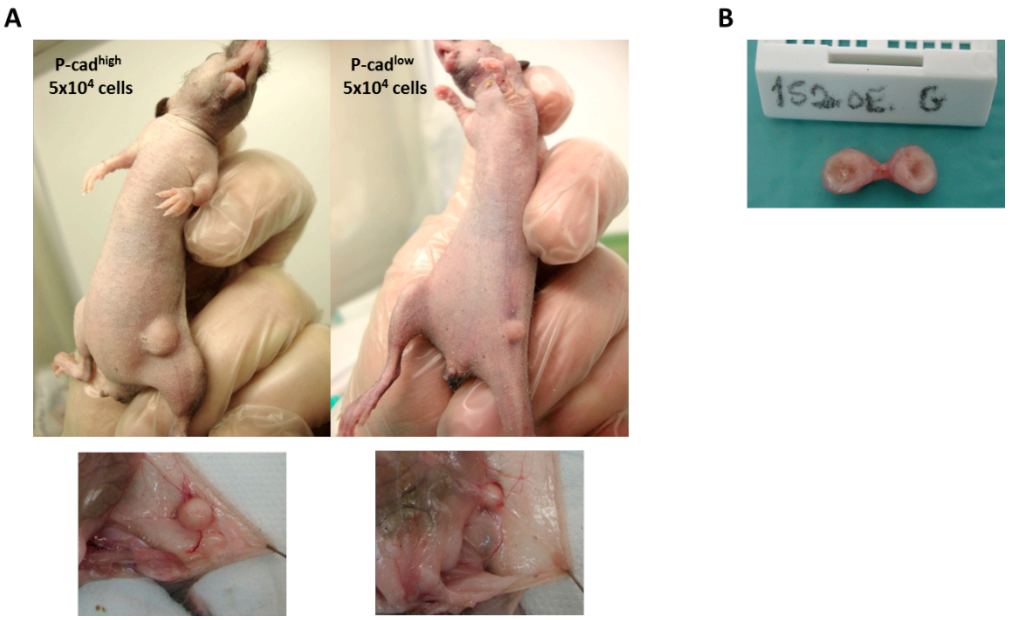


Figure 10 – Macroscopic evaluation of the tumours revealed that tumours derived from the xenotransplantation of P-cad^{high} cells had increased size when compared with the tumours derived in the P-cad^{low} group **(A).** Recruitment of blood vessels was clearly observed in all tumours and the interior of the tumours was often filled with a mucous substance **(B).**

3. Discussion

Increasing evidence supports that cancers are propagated by a small cell population, the CSCs, which originate both tumorigenic and non-tumorigenic cells and are responsible for tumor heterogeneity, therapy resistance and disease recurrence. The identification and analysis of CSCs is mandatory in carcinomas with high patient mortality rate, early relapses and lack of a targeted therapy. One of such poor-prognosis cancer is the basal-like subtype of breast cancer.

Basal-like breast carcinomas include tumors that are characterized by an expression signature similar to that of the basal cells of the breast (Rakha & Reis-Filho, 2009). However, features of the luminal gene expression profile are often also attributed to these cancers (Gusterson, 2009); therefore, basal-like breast carcinomas are often considered to have a mixed phenotype. More recently, the cell of origin for basal-like breast cancers was defined as the luminal progenitor cell from the normal breast (Lim *et al.*, 2009, Molyneux *et al.*, 2010), but the phenotype for the cancer stem cell has proven a tough task.

In this work we established a positive association between P-cadherin, a basal marker, and the stem cell markers CD44, CD24 and CD49f in human breast cell lines (normal and cancer cells). We found that all the stem cell markers analysed (CD44, CD49f and CD24) segregated together with P-cadherin in the same subset of cells, suggesting that P-cadherin could be contributing to a stem-like aggressive phenotype in the breast. In accordance to this, the expression of all of these stem cell markers have been linked to the acquisition of a malignant phenotype or decreased patient survival (Baumann et al., 2005, Cariati et al., 2008, Kristiansen et al., 2003, Marhaba & Zoller, 2004). Significantly, our study shows that breast cancer cases that express P-cadherin and CD44, or P-cadherin and CD24, have the worst prognosis. Notably, P-cadherin could be an important regulator of stem cell markers in tumor cells, since when this adhesion molecule was knocked-down a reduction was found in the expression of CD49f and CD44v6.

Based on these results, we believe that we got additional experimental evidence that P-cadherin can be classified as a stem cell marker in breast cancer. Although the simultaneous expression of all the three stem markers in the same cell was never clearly described, studies with CD49f and CD24 have shown that these markers combined in the same cell are able to identify cells with progenitor like properties in mammary epithelial cells in the mouse (Molyneux *et al.*, 2010, Stingl *et al.*, 2006) and in the human (Lim *et al.*, 2009, Keller *et al.*, 2010). Specifically, Keller *et al* showed that the normal human breast tissue contains an epithelial subtype with the phenotype $EpCAM^{high}CD49f^{+}CD24^{+}$,

consistent with a definition of luminal progenitor cells, according to the characterization of lineage markers (Keller *et al.*, 2010). CD44 has also been used in combination with CD24 to show that the phenotype $CD44^{+}CD24^{+}$ identifies cells with stem-like properties, including tumourigenicity, in breast cell lines (Meyer *et al.*, 2009, Rappa & Lorico, 2010) and in ER-negative human breast tumors (Meyer *et al.*, 2010). In this last study Meyer *et al* also show that $CD44^{pos}CD49f^{high}$ cells are enriched for tumor initiating ability in ER-negative breast cancers (Meyer *et al.*, 2010).

the association between the expression of P-cadherin and the phenotype $CD49f^{+}CD24^{+}$ in several cancer and normal cell lines supports the notion that P-cadherin could potentially be used together with other markers in the identification of the luminal progenitor of the normal breast. In fact, luminal progenitor cells have been described as the cell of origin for basal-like cancers (Lim *et al.*, 2009, Molyneux *et al.*, 2010) in which *BRCA1* inactivation is a common feature and is considered to halt the maturation towards the luminal phenotype (Liu *et al.*, 2008). This would explain the up-regulation of P-cadherin found in basal-like cancers, since BRCA1 is a P-cadherin transcriptional repressor (Gorski *et al.*, 2009). Notably, Sarrio *at al* recently reported that in a normal non-tumorigenic human breast cell line with a basal-like profile the subpopulation defined as $EpCAM^{+}CD49f^{+}$ exhibits progenitor like properties and, interestingly, this subpopulation of cells specifically expressed P-cadherin (Sarrio *et al.*, 2011).

An enrichment of P-cadherin expression in cells that have the stem/progenitor phenotype $CD44^{+}CD24^{+}$ was also found in this study. Importantly, cells with the phenotype $CD44^{+}CD24^{+}$ have been described as having tumorigenic ability and represent a dynamic population which can originate the CSC phenotype (Meyer *et al.*, 2009, Pece *et al.*, 2010, Rappa & Lorico, 2010). In fact, Pece *et al,* showed that cells with cancer stem cell activity are $CD24^{+}$ (Pece *et al.*, 2010) and the phenotype $CD44^{+}CD24^{+}$, isolated from several human breast cancer cell lines, shows tumorigenic activity in murine xenograft models (Meyer *et al.*, 2009). This indicates that targeting cells with the phenotype $CD44^{+}CD24^{+}$, which have increased P-cadherin expression, could help to eradicate the CSCs.

Our data still show that P-cadherin is not a simple bystander present in the cells throughout the breast hierarchy, but it can also be a regulator of the stem cell function. First, we showed that P-cadherin has higher capacity in promoting tumor growth in nude mice. Second, we found that $P\text{-}cad^{high}$ cell lines have increased ALDH activity and that P-cadherin expression is increased in the ALDEFLUOR+ subpopulation. Previous studies have shown that ALDH activity and expression is linked to increased stem cell activity, tumorigenicity and poor prognosis in breast carcinomas (Croker *et al.*, 2009, Ginestier *et al.*, 2007). Furthermore, P-cadherin gene expression manipulation (either by silencing with

siRNA or by P-cadherin/*CDH3* overexpression), as well as separation of cell populations by sorting, showed that this adhesion molecule mediates stem cell activity and the self-renewal of mammospheres in basal breast cell lines. The growth in 3D matrigel matrix was also increased in the P-cad^{high} subpopulation relative to the P-cad^{low} subpopulation of basal-like breast cancer cells. It is still not clear whether the increase in clonogenic capacity in 3D mediated by P-cadherin, and the increase in stem cell activity measured by MFE, is the result of increased survival or alterations in the growth rate that affects proliferation/quiescence of the prospective P-cad^{high} stem cells. Analysis of the cell cycle revealed that P-cadherin does not affect the subpopulations S and G_2M (**Figure 9**). Importantly, the fact that cells with lower expression of P-cadherin still show some tumorigenic ability and a certain level of growth in anchorage independent conditions and in matrigel 3D cultures indicates that stem cell properties are still present in the P-cadherin depleted fraction. Although P-cadherin does not seem to be an exclusive marker of CSCs, our results show that a further enrichment of the stem cell population could potentially be achieved by a combination of P-cadherin with other biomarkers. For basal-like carcinomas, several studies have used combinations of the following biomarkers to define the CSC subpopulation: ESA, CD44, CD24, CD49f, CD133-2, PROCR and ALDEFLUOR(Al-Hajj *et al.*, 2003, Croker *et al.*, 2009, Hwang-Verslues *et al.*, 2009, Meyer *et al.*, 2010, Wright *et al.*, 2008). Not surprisingly, expression of P-cadherin alone in invasive breast cancer confers bad prognosis and we observed that this marker is associated with the stem cell markers CD44, CD49f and ALDH1, and all these markers with the basal-like molecular subtype. The isolation of P-cadherin positive subpopulations within the CD44$^+$ or within CD24$^+$ fractions is of potential interest, since in our series of human breast carcinomas the tumors that exhibit P-cadherin$^+$CD44$^+$ or P-cadherin$^+$CD24$^+$ phenotype seem to be particularly aggressive. Interestingly, the phenotype P-cadherin$^+$CD44$^+$CD49f$^+$CD24$^+$ could also support the isolation of progenitor cells from the normal breast, as mentioned above.

Additionally, it is known that basal-like breast cancers are particularly aggressive because they resist to current therapeutic strategies, usually recurring in a short time frame. In particular, resistance to radiotherapy has been reported for breast cancer, and relapses often occur. Although poorly understood, this type of resistance could be mediated by several mechanisms controlling the survival and DNA repair mechanisms, allowing CSCs to survive and to give rise to a new tumor (Phillips *et al.*, 2006). In this study, we show that, when given X-ray irradiation, the silencing of P-cadherin contributes to decreased survival of the stem cell population in the tumorigenic cells, but not in the normal. Hence, inhibition of P-cadherin could be an approach to increase sensitization of tumorigenic cells to radiotherapy, potentially allowing the reduction of the current doses of

radiation administered to the malignant tissue and causing less damage to the normal cells.

In conclusion, we found, for the first time, that P-cadherin confers stem cell features to breast tumorigenic cells that could be linked to the aggressive behavior of basal-like breast cancers. We show that this molecule is associated with increased stem cell activity (tumorigenicity in athymic nude mice, mammosphere formation and growth in 3D matrix) in basal-like cell lines, but not in luminal cells. P-cadherin is associated with already described stem cell markers which define the luminal progenitor phenotype and which, in our series of primary invasive breast cancers, is associated with the shortest overall patient survival. We also found that inhibition of P-cadherin sensitizes cancer cells to X-ray induced cell death.

In the future, CSC therapies for the aggressive basal-like breast carcinomas can eventually involve the targeting of P-cadherin cell surface protein. In fact, phase-I clinical trials are underway, which will help to define if P-cadherin constitutes a good therapeutic target in breast cancer (Zhang *et al.*, 2010). Importantly, our results show that anti-P-cadherin treatment could improve radiation therapy in patients.

P-CADHERIN INTRACELLULAR SIGNALLING IS DEPENDENT ON $\alpha 6\beta 4$ INTEGRIN ACTIVATION TO INDUCE BREAST CANCER STEM CELL AND INVASIVE PROPERTIES

CHAPTER VI

Submitted article related to this chapter:

André Filipe Vieira, Ana Sofia Ribeiro, Maria Rita Dionísio, Bárbara Sousa, Ana Rita Nobre, André Albergaria, Angélica Santiago-Gómez, Fernando Schmitt, Robert B. Clarke, Joana Paredes. **P-cadherin intracellular signalling is dependent on α6β4 integrin activation to induce breast cancer stem cell and invasive properties.** 2012 (*submitted for publication*)

1. Introduction

Cadherin molecules have a major role in tumour progression. A significant example is the E-cadherin/*CDH1* gene, for which a tumour suppressor function is clearly observed in the large majority of human carcinomas. In fact, one of the first steps in the metastatic cascade is the loss of E-cadherin expression by cancer cells. Mutations of the E-cadherin/*CDH1* gene are also involved in increased risk to develop certain types of breast and gastric cancers (Paredes *et al.*, 2012). P-cadherin, on the other hand, has a tumour promoting effect in several solid tumours, including the pancreatic, prostate, colorectal and breast cancer (Albergaria *et al.*, 2011, Hardy *et al.*, 2002, Imai *et al.*, 2008, Paredes *et al.*, 2005, Taniuchi *et al.*, 2005) (Paredes *et al.*, 2012). Indeed, we have previously demonstrated that P-cadherin is a poor prognosis factor in mammary invasive ductal carcinomas and is associated with lack of differentiation and high grade carcinomas (Paredes *et al.*, 2005, Paredes *et al.*, 2002b). Its expression was found to be up-regulated in a particularly aggressive subtype of breast cancers, specifically in the basal-like subgroup (Matos *et al.*, 2005, Paredes *et al.*, 2005, Paredes *et al.*, 2002a). *In vitro* studies from our group have shown that P-cadherin mediated aggressive behaviour in breast cancer involves increased cell invasion and motility (Paredes *et al.*, 2004), increased production and activation of metalloproteinases (Ribeiro *et al.*, 2010) and increased cancer stem/progenitor cell properties (Vieira *et al.*, 2012).

It is well known that the maintenance of the stem cell activity requires signalling mediated by the ECM and by ECM receptors, also known as integrins (Bissell & Hines, 2011). Integrins are major signalling receptors that integrate external ECM signals controling the cell internal milieu. Conversely, they also signal from the inside of the cell to modulate the extracellular microenvironment. The basal/myoepithelial cells of the breast are in direct contact with the basement membrane which is composed of a complex mixture of ECM molecules that contribute to the survival and adhesion signalling of epithelial cells and to the maintenance of the stem cell niche within the normal breast. Interestingly, P-cadherin is expressed in the basal/myoepithelial cells and it was shown to be co-expressed with another adhesion molecule, the α6 integrin receptor (or CD49f), in a population of cells that mediate stem-like properties (Vieira *et al.*, 2012).

Alterations in the ECM or in integrin expression mediate major oncogenic effects in the initiation and progression of breast cancer (Bissell & Hines, 2011, Goss *et al.*, 2008). For example, ECM remodelling and integrin activation assist in the malignant transformation of cells in the primary site, as well as in the activation of quiescent cells in distant metastatic sites, such as the bone, liver, lung and brain (Barkan *et al.*, 2010b,

Pontier & Muller, 2009, Barkan *et al.*, 2010a, Psaila & Lyden, 2009). In the normal breast, the basement membrane has a crucial role in limiting tumour progression, being composed mainly by collagen type-IV and laminin-332 (formerly termed laminin 5) (Tsuruta *et al.*, 2008). In cancer, elevated expression of laminin in breast carcinomas is considered a poor prognosis factor (Tagliabue *et al.*, 1998, Tsuruta *et al.*, 2008). In fact, abnormal overexpression of laminin-332 is present in the migrating edge of the tumour mass and the expression of laminin receptors are believed to promote invasion of breast cancer cells (Kim *et al.*, 2011, Tsuruta *et al.*, 2008). Although several integrins recognize laminin substrates, the α6 integrins ($\alpha6\beta1$ and $\alpha6\beta4$) are the major receptors that contribute to breast cancer progression and have captured the interest of several cancer researchers (Mercurio *et al.*, 2001, Soung *et al.*, 2011). Thus, the role of the heterodimmer $\alpha6\beta4$ in tumour progression has been extensively investigated. Aberrant activation of the α6β4 receptor is implicated in cell survival, migration and invasive potential (Shaw *et al.*, 1997, Bon *et al.*, 2007, Mercurio *et al.*, 2001, Soung *et al.*, 2011). Interestingly, the expression of the $\beta4$ integrin subunit is associated with poor breast cancer patient prognosis (Lu *et al.*, 2008, Tagliabue *et al.*, 1998) and specifically with the basal-like molecular subtype (Lu *et al.*, 2008). Although mice in which $\beta4$ integrin was inactivated in the mammary gland have a normal breast development (Klinowska *et al.*, 2001), this integrin subunit was found to be crucial for breast cancer progression (Guo *et al.*, 2006). Furthermore, overexpression of the $\alpha6$ integrin subunit was found in invasive breast carcinomas correlating with decreased overall patient survival (Friedrichs *et al.*, 1995), being an important breast stem cell marker in both mice and humans (Eirew *et al.*, 2008, Stingl *et al.*, 2001, Stingl *et al.*, 2006, Villadsen *et al.*, 2007). A major role has been also proposed for $\beta1$ integrin subunit in the normal development of the murine breast, regulating the ability of the stem cells to self-renew and properly differentiate (Taddei *et al.*, 2008). This integrin molecule is also known as CD29 and it represents an important marker of normal murine stem cells (Shackleton *et al.*, 2006). $\beta1$ integrin has also an important role in tumourigenesis, since the disruption of this integrin in the mammary gland of a transgenic mouse model completely blocked tumour formation driven by the polyomavirus middle T antigen (White *et al.*, 2004).

Thus, the crosstalk between cell-cell and cell-ECM adhesion complexes reflects a highly integrated network. Although often spatially distinct, integrin and cadherin adhesions activate many of the signalling pathways and elicit similar cellular functions that are part of a larger adhesive structure. In cancer, an association of cadherins and integrins can originate complexes that mediate important oncogenic responses, often through interaction with other transmembrane proteins, such as growth factor receptors.

Several reports focus on the association of E-cadherin with integrin molecules (Canonici *et al.*, 2008, Chartier *et al.*, 2006, Chattopadhyay *et al.*, 2003, Wang *et al.*, 2004), but no interaction between P-cadherin and integrin molecules was ever described. P-cadherin role is well described in cell-cell interaction, however its role in the cell-ECM interaction is not known.

The aim of this study was to reveal whether P-cadherin-induced stem cell features and invasive properties in breast cancer cells are dependent on ECM components and integrin receptors signalling. We used basal-like breast cancer cell lines and demonstrated that P-cadherin affects the adhesion of cells to different ECM substrates. Furthermore, we found that, mechanistically, the signalling pathway triggered by P-cadherin in response to laminin is dependent, at least partially, on α6β4 integrin expression, and activates Src, FAK and AKT kinases in breast cancer cells.

P-cadherin dependent adhesion of breast cancer cells to ECM components

The role of P-cadherin as a cell-cell adhesion molecule is well documented; however, its role in cell - ECM adhesion is unknown. In this study, the adhesion of cancer cells to several ECM components typically implicated in tumour progression was assessed. The cell lines tested were MDA-MB-468 and BT-20, which have a basal-like epithelial phenotype characterized by high expression of E-cadherin, negativity for hormone-receptors, lack of HER-2 amplification and high levels of basal markers, including EGFR amplification and high expression of P-cadherin (>80% P-cadherin positive). Transient knock-down of P-cadherin was performed by siRNA (60% inhibition in MDA-MB-468 and 82% inhibition in BT-20) and adhesion to collagen type-I, collagen type-IV, laminin-332, vitronectin and fibronectin was measured by the crystal-violet assay (**Figure 1**).

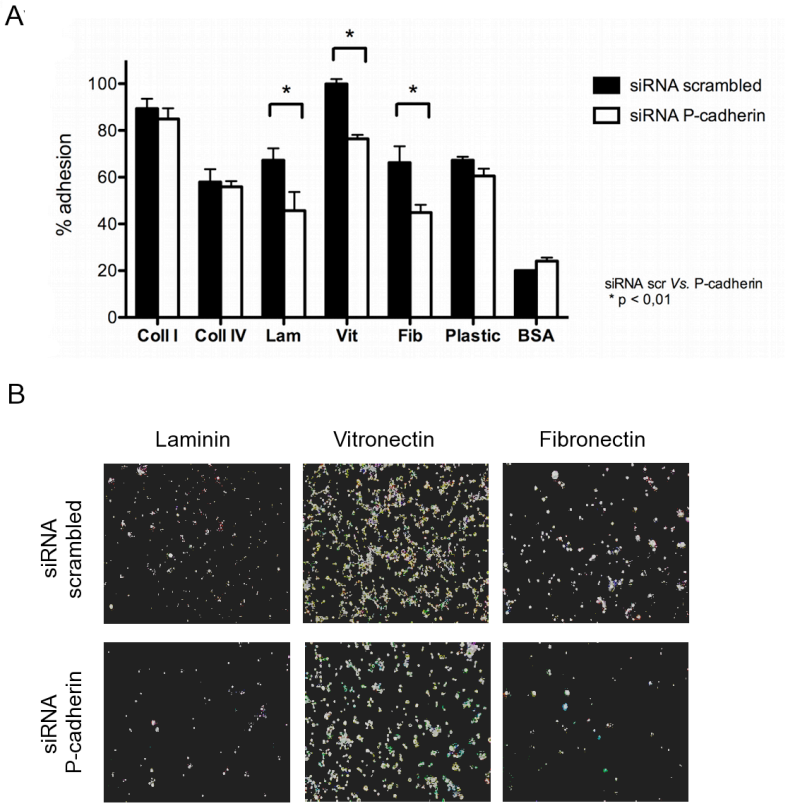


Figure 1 – Adhesion of the breast cancer cell lines MDA-MB-468 and BT-20 to extracellular matrix (ECM) components is dependent on P-cadherin. (A) In the example shown, inhibition of P-cadherin expression in MDA-MB-468 cells decreased % adhesion to laminin-332, vitronectin and fibronectin (adhesion time = 20 min). The same trend was observed for BT-20 cell line (adhesion time = 30 min). (B) Bright field images of MDA-MB-468 cells in the tissue culture plate coated with ECM substrates after the adhesion assay. Cells were fixed and the nuclei stained with crystal-violet. (Coll I – collagen I, Coll IV – collagen IV, Lam – laminin 332, Vit – vitronectin, Fib – fibronectin, BSA – bovine serum albumin, negative control).

Both cell lines behaved similarly: control transfected cells preferentially adhered to collagen type-I and vitronectin (>90% adhesion), followed by a moderate adhesion to collagen type-IV, laminin-332 and fibronectin (50-70%). Adhesion of both cell lines to plastic was approximately 70% (**Figure 1A**). When P-cadherin was inhibited, adhesion to laminin-332, vitronectin and fibronectin was significantly reduced by about 20%, whereas adhesion to the collagen molecules (type I and IV) was not affected (**Figure 1A and 1B**). Adhesion to plastic was slightly reduced, although not significantly.

P-cadherin regulates the expression of the laminin receptor α6β4 integrin in breast cancer cells

Since P-cadherin interfered with the adhesion of cancer cells to laminin-332, vitronectin and fibronectin, we set out to investigate whether this was mediated by any alteration in the function or expression of integrins, the major receptor molecules that connect cells to the ECM and which are known to be directly implicated in cancer aggressiveness. Thus, we analysed the surface expression of β1, β4 and α6 integrins by flow cytometry.

β1 integrin is a major component of most integrin heterodimers recognizing most ECM components, including laminin-332, vitronectin and fibronectin. α6 and β4 subunits bind exclusively to laminin and have important tumour promoting effects in breast cancer (Bon *et al.*, 2007, Mercurio *et al.*, 2001, Shaw *et al.*, 1997, Soung *et al.*, 2011). As shown in **Figure 2**, P-cadherin inhibition had no effect in the expression of β1 integrin. However P-cadherin knock-down caused a reduction in the cell surface expression of α6 and β4 integrins in both cell lines MDA-MB-468 and BT-20, as evaluated by flow cytometry (**Figure 2A**). Furthermore, the expression of α6 and β4 subunits was also evaluated by immunofluorescence and immunoblot, confirming a decrease in the total amount of these integrins in breast cancer cells (**Figure 2B and 2C**).

Interestingly, α6 and β4 form a heterodimer (also known as hemidesmosome, in normal cells) that recognizes the major component of the basement membrane, laminin-332, for which we demonstrated that adhesion was impaired upon P-cadherin knock-down (**Figure 1**).

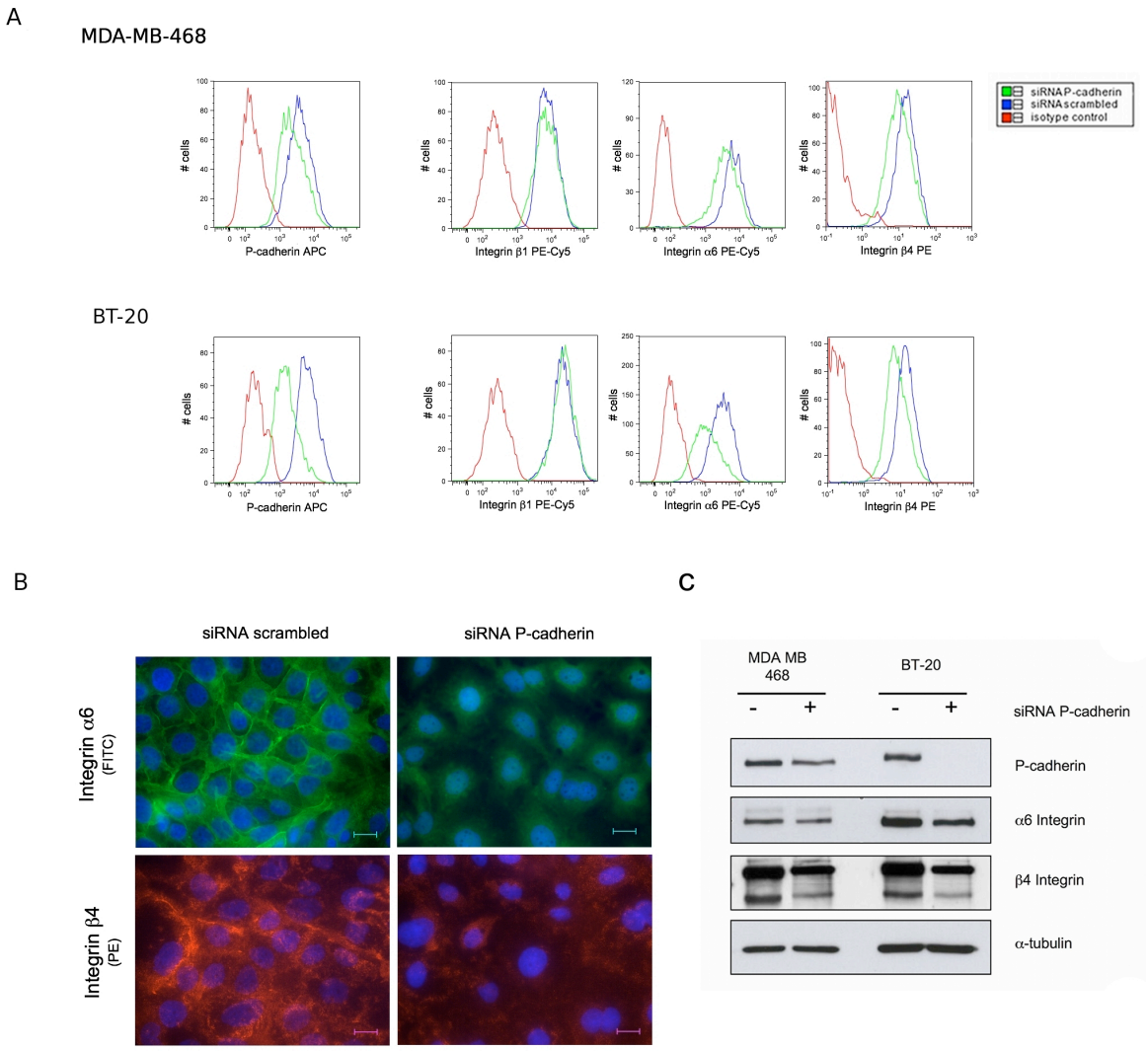


Figure 2 – P-cadherin knock-down reduces integrin α6 and β4 expression in MDA-MB-468 and in BT-20 cells. Cell surface expression of P-cadherin and integrin molecules was analysed by Fluorescence Activated Cell analysis. The median intensity of integrins α6 and β4 stain was decreased upon P-cadherin knock-down. No effect was observed in integrin β1 expression. (A). Similarly, a decrease in integrin α6 and β4 was also found by Immunofluorescence (the cell line BT-20 is represented (scale bar=20μm) and similar results were obtained for MDA-MB-468 cell line) (B) and by Immunoblot (C).

P-cadherin and the α6 integrin confer stem cell properties and invasive features to breast cancer cells

Since P-cadherin expression shows an important effect on cell-ECM adhesion and clearly modifies integrin α6β4 expression in breast cancer cells, we set out to study if this integrin heterodimer was also implicated in important aggressive properties that have been previously ascribed to P-cadherin, namely, the invasive capacity and the cancer stem cell activity. Furthermore, to clarify the crosstalk between P-cadherin and α6β4

integrin, the effect of both α6 and β4 integrin subunits in the expression levels of P-cadherin were also studied by immunoblot.

Inhibition of α6β4 in breast cancer cells decreased the mammosphere forming efficiency (MFE), as well as the invasion capacity in the same magnitude as the inhibition of P-cadherin (**Figure 3A and 3B**). Importantly, α6 integrin inhibition alone showed the same impact in MFE and in the invasion potential as the inhibition of P-cadherin or the repression of the α6β4 heterodimer. However, inhibition of the β4 integrin subunit in breast cancer cells did not show a statistically significant impact in these functional properties (**Figure 3A and 3B**). These results indicate that P-cadherin downstream signalling effects could be primarily dependent on the integrin α6 subunit function.

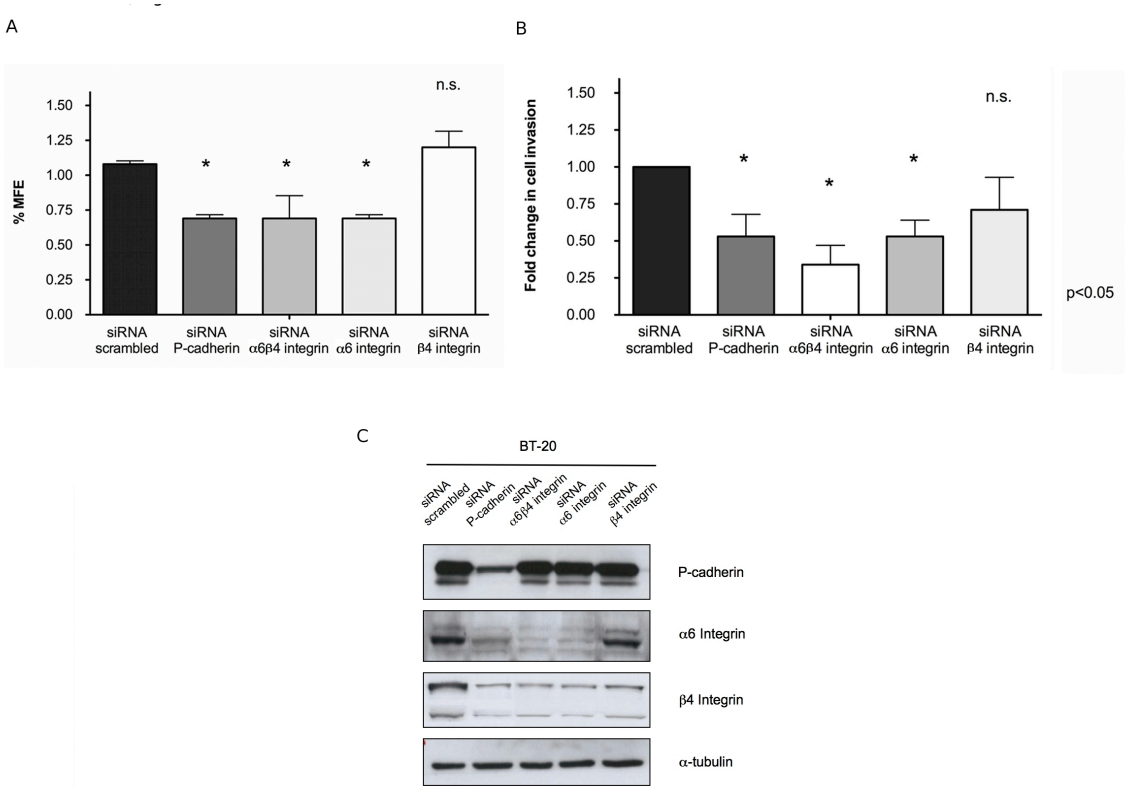


Figure 3 – Inhibition of α6β4 integrin decreased the stem cell activity in breast cancer cells to the same extent as inhibition of P-cadherin, measured by the mammosphere forming efficiency (MFE) (BT-20 cell line) (A). In the same way, the invasion capacity of these cells in matrigel was severely compromised when P-cadherin or α6β4 integrin were knocked-down (B). As P-cadherin clearly impacted in the expression of α6β4 heterodimer, these results indicated a possible association between both pathways; although the expression of the α6 and β4 subunits is decreased upon P-cadherin knock-down, the expression of P-cadherin is not affected after the inhibition of either α6 integrin or β4 integrin or both integrins at the same time (C).

It is interesting to note that, while P-cadherin knock-down caused a reduction in α6 and β4 integrin subunits, the opposite was not true (**Figure 3C**). The inhibition of α6 and/or β4 integrins showed no effect in P-cadherin expression. Nonetheless, α6 integrin

knock-down led to a decrease in the expression of its partner, the β4 integrin subunit, again raising the possibility that P-cadherin may in fact be controlling mainly the α6 subunit expression which in turn controls the β4 subunit. In summary, only when the α6 integrin subunit or the α6β4 integrin heterodimer were inhibited were the functional properties affected; the inhibition of β4 integrin subunit had no effect in MFE and invasion.

P-cadherin overexpressing cells have increased adhesion to laminin as well as increased stem cell activity and these properties are dependent on α6β4 integrin expression

The previous results raised the possibility that there could be a crosstalk between two adhesion molecules: P-cadherin and α6 integrin. Thus, to further explore the role of α6 integrin and its partner, β4 integrin, in the functional properties mediated by P-cadherin, we analysed the cell-laminin adhesion capacity and the MFE of a breast cancer cell line constitutively overexpressing P-cadherin (MCF7/AZ.P-cad) and compared these properties with control cells, which have low levels of P-cadherin (MCF7/AZ.mock). P-cadherin expression was accompanied by an increase in the expression of both the α6 integrin subunit as well as the β4 integrin subunit (**Figure 4A**). Importantly, P-cadherin upregulation lead to an increase in the adhesion of MCF7/AZ.P-cad cells on top of a laminin coated surface (**Figure 4B**) and increased the stem cell activity of these cells, measured by the MFE (**Figure 4C**). These effects were mediated, at least partially, by α6β4 integrin expression, since this integrin heterodimer was increased in P-cadherin overexpressing cells (**Figure 4C**); when both integrin subunits were simultaneously knocked-down in MCF7/AZ.P-cad cells, these functional properties were significantly reduced (**Figure 4**). Once more, P-cadherin levels were not affected by α6β4 integrin knock down, indicating that these integrin molecules are most likely acting downstream of P-cadherin activation (**Figure 4C**).

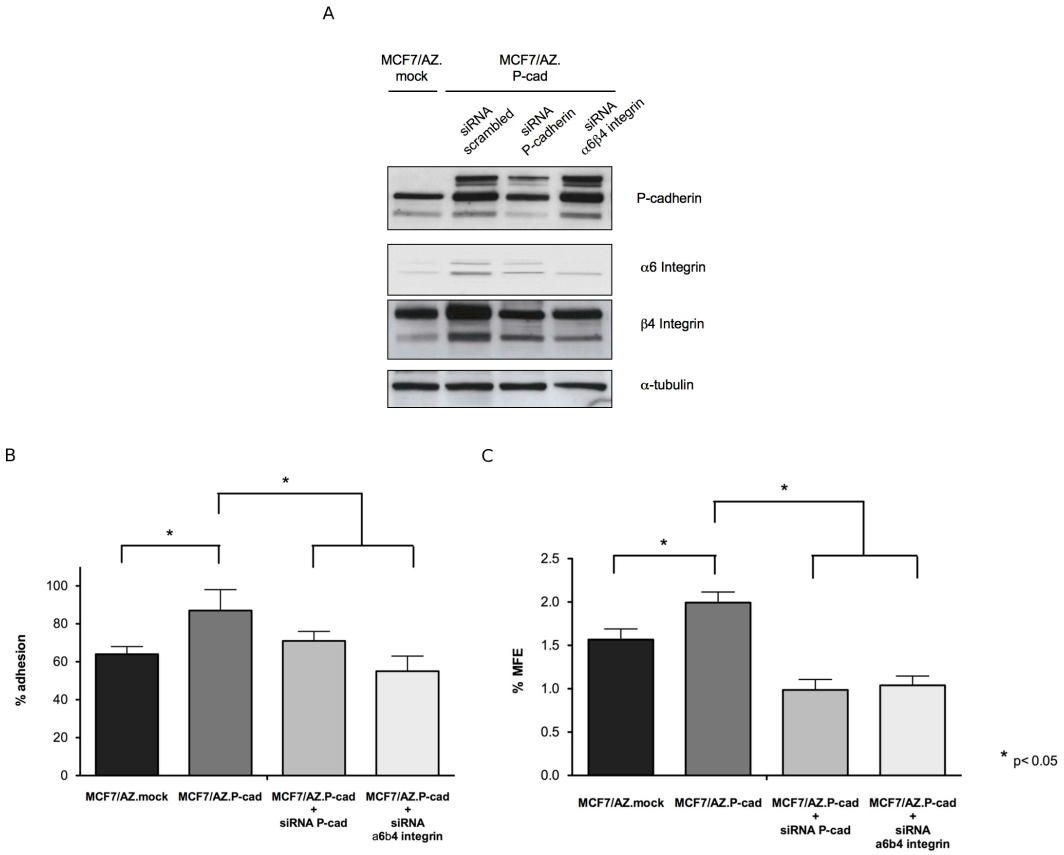


Figure 4 – The increased adhesion to laminin and the increased stem cell activity observed in P-cadherin overexpressing cells is dependent on α6β4 integrin expression. MCF7/AZ cells overexpressing P-cadherin (MCF7/AZ.P-cad) presented increased capacity to adhere to laminin (adhesion time = 40 min) and increased mammosphere forming efficiency (MFE) (*Vs.* MCF7/AZ.mock cells) (A) and (B). These effects are accompanied by an increase in the expression of the integrin subunits α6 and β4. Inhibition of the α6β4 integrin heterodimer in the P-cadherin overexpressing cells restored the levels of adhesion to the control levels, and strongly inhibited the MFE in these cells. P-cadherin expression was not affected by α6β4 integrin knock down (C).

Integrin signalling in response to laminin-332 is dependent on P-cadherin expression

The previous data established a cooperation between P-cadherin and $\alpha 6\beta 4$ integrin regarding signalling. We therefore studied whether P-cadherin could affect the main signalling molecules downstream of the $\alpha 6\beta 4$ integrin receptor in cancer cells grown on top of a laminin substrate. The activation of the integrin related kinases FAK and Src was studied by immunoblot after cell adhesion to this substrate. We found that P-cadherin inhibition in breast cancer cells reduced p-FAK Tyr397 and p-Src Tyr416 levels (**Figure**

5A). Notably, the p-FAK reduction was also detected by immunofluorescence in both cell lines studied (**Figure 5B**). Furthermore, activation of AKT was also affected, shown by the level of p-AKT Ser473 being reduced (**Figure 5A**). Altogether, these results indicate that FAK and Src activation in response to laminin is dependent on P-cadherin expression in breast cancer cells. Although lateral integrin-cadherin associations are known to occur (Canonici *et al.*, 2008, Chattopadhyay *et al.*, 2003, Weber *et al.*, 2011), we were unable to demonstrate an interaction between P-cadherin and α6 or β4 by co-immunoprecipitation (data not shown).

We also investigated if the cancer cell phenotype was affected in cells grown on top of ECM substrates. We analysed the cell actin microfilaments by phalloidin staining and visualized F-actin by fluorescence microscopy in breast cancer cells adhered to laminin coated coverslips (**Figure 5B**). We found that control cells (scrambled transfected) had more stress fibres and appeared more flattened than cancer cells with P-cadherin inhibition. The stress fibres provide the cytoskeletal tension which is required for focal adhesion formation in laminin, indicating a strong adhesion to the ECM substrate. Staining with an antibody for p-FAK Tyr397 allowed the identification of focal adhesions and sites of cell-to-cell contacts. Both focal adhesions and cell-cell contacts were decreased in P-cadherin depleted cells (**Figure 5B**). In summary, P-cadherin has a role in eliciting cell shape changes associated with adhesion to the ECM.

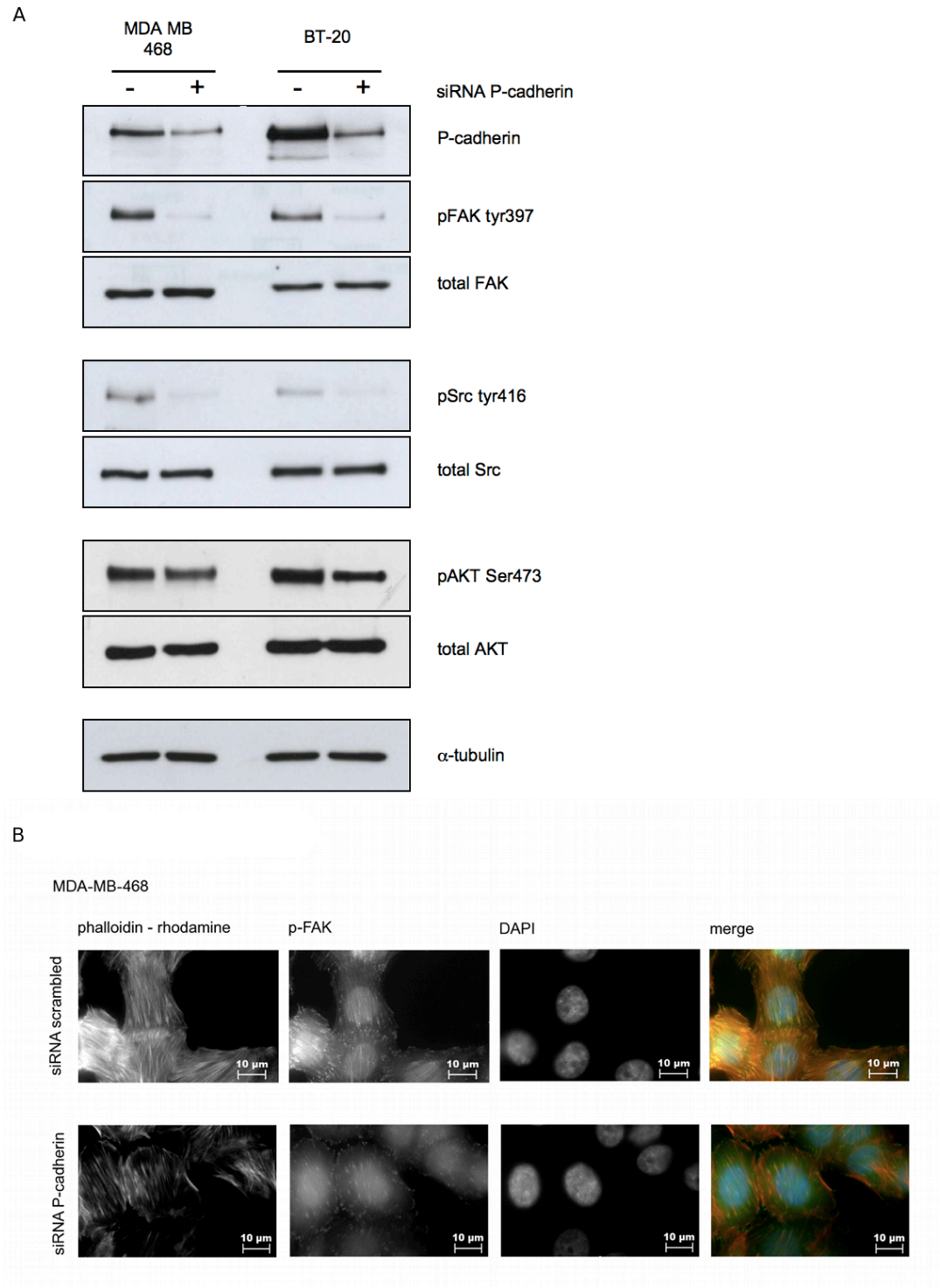


Figure 5 – P-cadherin signalling in response to laminin is dependent on FAK and Src activation. Analysis of integrin downstream signaling molecules in breast cancer cell lines was performed after adhesion to laminin-332 (20 min for MDA-MB-468 and 30 min for BT-20) (A); The number of stress fibers (F-actin was stained with phalloidin-rhodamine) and focal adhesions/contacts (stained with pFAK Tyr397 – Alexa 488) is reduced by P-cadherin knockdown in MDAB-MB-468 cells grown on top of laminin. The same result was found for BT-20 cell line (B).

3. Discussion

Cadherins are classically seen as molecules that make a major contribution for cell-to-cell adhesion. Specifically in breast, P-cadherin expression is found in the myoepithelial cell layer, where it is thought to contribute to the self-organization of these cells as well as to cell-cell adhesion (Chanson *et al.*, 2011). Notably, P-cadherin specific signalling pathways are far less well described compared to other classical adhesion molecules. Furthermore, the basal layer of the mammary epithelium is enriched in molecules involved in the adhesion of epithelial cells to the ECM, namely integrin molecules, such as $\alpha6\beta1$ and $\alpha6\beta4$. In normal cells, the later heterodimer is known as hemidesmosome and it is the receptor for laminin, the major component of the basement membrane.

In breast cancer, P-cadherin molecule appears upregulated in 30% of all breast cancer patients and it is associated with poor patient prognosis (Paredes *et al.*, 2005, Turashvili *et al.*, 2011). Importantly, cancer progression in the breast involves modifications of the normal ECM, as well as oncogenic activation of integrin signalling in both the primary, as well as in the metastatic site (Muschler & Streuli, 2011, Pontier & Muller, 2009). Here, we establish that P-cadherin is involved in the attachment of cells to ECM substrates, since silencing of P-cadherin expression rendered the cancer cells significantly less able to adhere to vitronectin, fibronectin and laminin. When integrins expression was investigated, we found that P-cadherin was necessary for the correct expression of the integrin subunits $\alpha6$ and $\beta4$. Importantly, the recognition of laminin by cancer cells has significant tumour promoting effects. For example, laminin-332 induces motility in MCF-7 cell line (Carpenter *et al.*, 2009). Furthermore, IHC analysis of laminin-332 in human carcinomas *in situ* showed that this ECM substrate is located in the myoepithelium adjacent to preinvasive cells (Carpenter *et al.*, 2009, Kim *et al.*, 2011) that could potentially contribute to early steps of stromal invasion. The interface zone between the tumour cells and the stroma is enriched in laminin and in $\alpha6\beta4$ integrin (Kim *et al.*, 2011).

P-cadherin is a major player in inducing invasion and migration of breast cancer cells (Ribeiro *et al.*, 2010). The signalling pathways that contribute to this aggressive behaviour are poorly understood, involving to some extent the activation of metalloproteinases and the consequent release of a pro-invasive P-cadherin fragment and/or the activation of small GTPases (Ribeiro *et al.*, 2010, Taniuchi *et al.*, 2005). Recently, we have shown that P-cadherin has been implicated in the maintenance of stem and progenitor properties in basal-like breast cancer cells, including the self-renewal capacity and the tumourigenic ability in nude mice (Vieira *et al.*, 2012). We also found that

P-cadherin is co-expressed with α6 integrin in breast cancer cells (Vieira *et al.*, 2012), a marker of the stem/progenitor phenotype present in the mouse and human breast (Eirew *et al.*, 2008, Stingl *et al.*, 2001, Stingl *et al.*, 2006, Villadsen *et al.*, 2007). In the present work, we explored further this association, showing that there is a crosstalk between both adhesion molecules. P-cadherin is acting upstream of a major signalling pathway that involves the activation of α6 integrin and its partner, the β4 subunit. As a consequence of the adhesion of cancer cells to laminin surface, the activation of the α6β4 heterodimer would lead to Src and FAK activation in a P-cadherin dependent manner. In fact, P-cadherin knock-down reduces FAK and Src activation. Importantly, it has been reported that α6β4 integrin promotes survival and invasion by activating the PI3K/Akt pathway (Bachelder *et al.*, 1999, Shaw *et al.*, 1997) and, notably in our work, a reduction was also found in AKT activation in P-cadherin silenced cells.

Interestingly, it was previously found that α6 integrin induces P-cadherin transcription (Deugnier *et al.*, 1999), further supporting the idea that P-cadherin could cooperate with α6 integrin signalling. However, our work revealed that P-cadherin and α6 integrin do not directly interact (by co-immunoprecipitation, data not shown).

β1 integrin is also a partner of α6 integrin subunit recognizing laminin, being also essential for the correct development of the mammary epithelium, and regulating the ability of the mammary stem cells to self-renew and differentiate properly (Taddei *et al.*, 2008). Despite not having found any alteration in the β1 integrin levels upon P-cadherin inhibition, we do not exclude the possibility that β1 integrin subunit is also implicated in the maintenance/acquisition of cancer stem cell and invasive properties, as this is the other major partner of α6 integrin, constituting an important laminin receptor.

Additionally, although the cell shape was not severely affected by P-cadherin knock-down and cells clearly maintained an epithelial phenotype, we found that the number of cell-to-cell contacts and the number or focal adhesions to laminin was clearly reduced upon P-cadherin inhibition. It is possible that α6β4 integrin and the implicated FAK/Src kinase activity may also be contributing to the stem/progenitor characteristics. It was shown that FAK deletion in the murine mammary gland suppressed tumourigenesis by decreasing the number of cancer stem cells ($CD24^+CD29^+CD61^+$ and $ALDEFLUOR^+$ populations) (Luo *et al.*, 2009). Notably, FAK activation allows for the survival of cells in anchorage-independent conditions (Xu *et al.*, 2000), which may explain why integrin knock-down, as well as P-cadherin knock-down, reduced survival of cells growing as suspension colonies in which the ECM is present within the mammosphere.

Thus, the poor patient prognosis found in P-cadherin overexpressing breast cancer cases (Paredes *et al.*, 2004) may be related, at least partially, to the fact that this cadherin

enables cells to respond to integrin signalling and promotes an oncogenic response. Importantly, strategies to inhibit P-cadherin could lead to a decrease in integrin activation and potentially oppose the oncogenic signalling mediated by laminin and its receptor. Since P-cadherin up-regulation is also found in *in situ* stage of breast cancer development (Paredes *et al.*, 2002b), it is possible that it may already be contributing to the changes in integrin signalling in the early stages of breast cancer development.

In conclusion, our results show that the cell-to-cell adhesion mediated by P-cadherin can modulate the expression or activity of another type of adhesions, in this case integrin adhesive contacts. P-cadherin and integrins have oncogenic signalling pathways that cooperate and cross-talk. This is particularly relevant because it helps to understand tumour cell-ECM interactions and the mechanisms by which ECM and ECM receptors regulate cancer progression.

GENERAL DISCUSSION AND CONCLUSIONS

CHAPTER VII

1. General Discussion

The results described in this Doctoral Thesis highlight the importance of dissecting novel CSC biomarkers, with a focus in the basal-like subtype of breast carcinomas. Given the high heterogeneity of breast cancer and the fact that no universal phenotype for the CSC has been identified, new and better ways to isolate this important population of cells are needed. In this context, a putative new CSC biomarker, P-cadherin, is presented. P-cadherin aggressive features were associated with CSC characteristics in basal-like breast carcinomas and a new signalling mechanism with implications in the stem cell niche signalling is presented.

According to the hierarchical model of tumour progression (Bonnet & Dick, 1997, Reya *et al.*, 2001), it is assumed that CSCs generate cells with aberrant and limited differentiation, which through mechanisms as yet unknown translate into the distinct breast cancer molecular subtypes. In this context, each molecular subtype is putatively composed of cancer cells with a different level of differentiation (Schmitt *et al.*, 2012).

First, in Chapter IV we established three different methods to characterize the stem cell properties present in human breast cancer cell lines: cell surface marker expression, the ALDEFLUOR assay and mammosphere formation. We showed that these methods identify different subpopulations of cells in our panel of breast cancer cell lines. We demonstrated that the expression of the stem cell markers CD44, CD49f and ALDH1 were increased in the basal/epithelial group of cell lines. Also, this subgroup of cells presented a $CD44^{+}CD24^{+}$ phenotype, as well as $CD49f^{+}CD24^{+}$ phenotype, indicating an association with the luminal progenitor of the normal breast, the cell-of-origin for basal-like breast carcinomas (Lim *et al.*, 2009, Molyneux *et al.*, 2010).

Since it is pivotal to translate the CSC concept into the clinical practice, we also analysed the expression of stem cell markers in a series of human breast carcinomas by IHC. We demonstrated that of the five molecular subtypes, the basal-like subtype harbours the highest percentage of tumour cells with a $CD44^{+}CD24^{-/low}$ and $ALDH^{+}$ phenotype. This is corroborated by data published by Honeth and collaborators, who showed that the $CD44^{+}CD24^{-/low}$ CSC phenotype is significantly associated with basal-like breast cancers in human patients and in particular with *BRCA1*-inherited cancers (Honeth *et al.*, 2008). This was also confirmed in earlier studies using human breast cancer cell lines, which reported an enrichment of $CD44^{+}CD24^{-/low}$ and $CD44^{-}CD24^{+}$ cell populations in basal-like and luminal molecular subtypes, respectively (Fillmore & Kuperwasser, 2008, Sheridan *et al.*, 2006). In fact the CSC phenotype $CD44^{+}CD24^{-/low}$ was present in our

panel of basal/mesenchymal cell lines, being associated to a highly invasive *in vitro* pattern. Even though $CD44^{+}CD24^{-/low}$ cells and $ALDH1^{+}$ cells are more frequently found in basal-like than in luminal tumours, we also noticed that ALDH1+ cells were also found in the HER2-OE subtype (Chapter IV). This may be because HER2-OE tumours present a stem-like transcriptional programme as a result of their specific transforming genetic alteration. In line with this notion, induced overexpression of HER2 in breast cancer cell lines increased the number of ALDEFLUOR-positive cells (Korkaya et al., 2008). Since the basal-like subtype of breast cancers has an increased $CD44^{+}CD24^{-/low}$ CSC population in comparison with other molecular subtypes, it has been hypothesised that in these tumours, a differentiation block exists, as the majority of the cells constituting the bulk of the tumour do not display a differentiated phenotype (Charafe-Jauffret et al., 2006, Charafe-Jauffret et al., 2008, Shipitsin et al., 2007). The transformation of a basal-like mammary stem cell to a basal-like cancer cell may either be due to (a) LOH as a second hit in *BRCA1* mutation carriers or (b) to the downregulation of *BRCA1* expression in ER-negative stem cells. This might result in the arrest of the luminal differentiation process, which fixes the cells in the undifferentiated phenotype of basal-like carcinomas with its unique expression of basal cytokeratins (K5/6) (Liu *et al.*, 2008). It is important to refer that recently, a tumour initiating cell signature, derived from $CD44^{+}CD24^{-/low}$-sorted cells and mammospheres obtain from primary human breast tumours, was found to be exclusively enriched in a new molecular subtype referred to as claudin-low (Hennessy *et al.*, 2009). This new molecular subtype is characterised by the low expression of genes involved in tight junctions and cell–cell adhesion, including claudins 3, 4 and 7, Occludin and E-cadherin and it also displays EMT features. These tumour characteristics reflect a low level of tumour differentiation and fit with the CSC and EMT phenotypes.

Thus, our results show that the basal-like cancer subtype is enriched in several stem cell features, but the recognition of reliable markers to distinguish the breast CSC pool is still necessary and it will be decisive for the development of specific target therapies.

Second, it is important to mention that, to date, there is no universal cell surface antigen, or combination of antigens, for the purification of the CSCs for each intrinsic breast cancer subtype. The identification of a CSC marker for the basal-like subtype of breast cancer is of particular importance, due to its high mortality rate, fast relapses and lack of target therapy (Rakha *et al.*, 2009). Thus, in Chapter V, we explored the use of another cell surface protein to characterize the population of cells with stem cell properties *in vivo* and *in vitro*. We showed that P-cadherin is an adhesion molecule that may be a valuable biomarker, allowing the isolation and the study of CSCs in basal-like breast

cancers. The association of P-cadherin with the phenotypes $CD24^{+}CD44^{+}$ and $CD24^{+}CD49f^{+}$, as well as with the $ALDEFLUOR^{+}$ subpopulation in breast cancer cells is highly suggestive that P-cadherin could be a marker of the luminal progenitor cell of the breast epithelial hierarchy. **Figure 1** represents the hierarchy of the normal mammary gland, with a putative enrichment of P-cadherin expression in the luminal progenitor compartment. Furthermore, as already mentioned, this cell has been proposed as the cell-of-origin for basal-like carcinomas (Lim *et al.*, 2009, Molyneux *et al.*, 2010), which are characterized by *BRCA1* inactivation (Foulkes *et al.*, 2003, Lakhani *et al.*, 2005, Turner & Reis-Filho, 2006). In fact, BRCA1 is a transcriptional repressor of the *CDH3*/P-cadherin gene (Gorski *et al.*, 2009), thus explaining the up-regulation of this protein in basal-like tumours and the association we found with the luminal progenitor cell. Further studies that include lineage tracing experiments and the use of additional sets of markers will clarify this hypothesis.

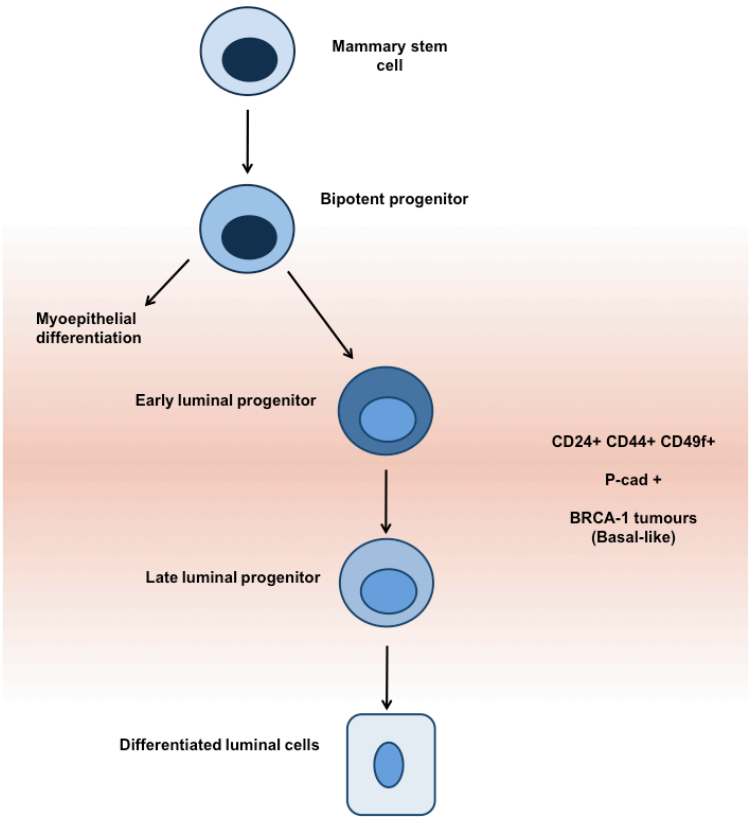


Figure 1 - Schematic representation of the hierarchy of the normal mammary gland. P-cadherin is putatively linked with *BRCA1* mutation in basal-like breast tumours and with the $CD44^{+}CD49f^{+}CD24^{+}$ phenotype. This phenotype encompasses the luminal progenitor compartment, which contains the cell-of-origin for basal-like breast cancers.

Furthermore, we showed that P-cadherin is not a simple bystander present in a subpopulation of cells, but it is in fact involved in basal-like breast CSC function. Cells enriched in P-cadherin expression exhibited increased self-renewal ability and were able to grow faster in complex 3D structures in clonal assays. Additionally, we showed that P-cadherin was able to confer increased tumorigenic ability to cancer cells in xenotransplantation experiments using immune compromised mice. With these *in vivo* and *in vitro* experiments we proved that this adhesion molecule has the capacity to mediate the three hallmarks of CSCs.

In the present work, we observed that P-cadherin enriched cells were more efficient to form mammospheres in both primary formation and in the secondary passage. We also showed that P-cadherin silencing reduced the mammosphere formation in the presence of X-ray irradiation, pointing to a role for this molecule in mediating radioresistance, another CSC property. In an attempt to understand the mechanisms behind the radioresistance, we have performed analysis of molecules involved in the DNA damage machinery. In fact, we have preliminary results that show that P-cadherin can interfere with the DNA repair machinery in non-irradiated cells: transient knock-down of P-cadherin in the basal cell lines MDA-MB-468, BT-20 and BT-549 caused an increase in the basal levels of activated Checkpoint kinase (Chk)1 and Chk2. This may indicate that cells without P-cadherin have increased basal levels of DNA damage which then leads to increased DNA repair mechanisms, therefore supporting the idea of P-cadherin in the resistance to DNA injury. Furthermore, Chk1 and Chk2 proteins have also been implicated in mediating cell death (de Lange *et al.*, 2012, Ouyang *et al.*, 2009, Wang *et al.*, 2009, Sahu *et al.*, 2009) and the observed activation of these proteins may support the role of P-cadherin in tumour cell survival. Still, more future studies need to be performed in order to unravel the mechanisms of radioresistance mediated by P-cadherin in breast cancer cells.

Another major feature of cancer stem cells is their tumourigenic ability. P-cadherin enriched cells ($P\text{-}cad^{high}$), isolated from the basal cell line MDA-MB-468, could develop tumours in immune compromised mice which were higher in number and in size, compared to the P-cadherin depleted fraction ($P\text{-}cad^{low}$). The role of P-cadherin in the initiation and/or progression of breast cancer will be clarified in the future by assessing the tumourigenic ability of MDA-MB-468 cells stably expressing the shRNA for P-cadherin, under the control of a DOX inducible system. In fact, a *CDH3* shRNA approach was used in a mouse metastatic breast cancer cell model (4T1) showing that the knock-down of P-cadherin produced smaller tumours than the control cells, as well as lower numbers of metastatic foci in the lungs (preliminary results). Additionally, in the human BT20 cell line, the inhibition of P-cadherin using a siRNA strategy rendered the cells less tumourigenic (unpublished data).

In the third part of this work, we hypothesised that P-cadherin could have a role in the stem cell niche and in the response to the ECM. Thus, in Chapter VI, we dissected the molecular mechanisms that allow P-cadherin to induce cancer cell invasion and mammosphere formation. We showed that the self-renewal ability and the invasion capacity mediated by P-cadherin expression are, at least partially, dependent on α6 integrin expression. Furthermore, the self-renewal ability and the adhesion to laminin-332 were shown to be dependent in α6β4 integrin expression in a P-cadherin overexpressing cellular system. The α6β4 integrin heterodimer constitutes a major laminin receptor. Interestingly, laminin expression is often upregulated in basal-like tumours (Rodriguez-Pinilla *et al.*, 2007, Kwon *et al.*, 2012), it is an indicator of poor prognosis (Molino *et al.*, 2003) and its presence in the interface zone between the tumour and the stroma is thought to stimulate cancer cell migration (Kim *et al.*, 2011). Thus, the cross-talk between P-cadherin and the α6β4 integrin receptor is indicative that P-cadherin has a role in the response of breast cells to ECM signals. In this signalling network, we believe that P-cadherin is acting upstream of the α6β4 integrin heterodimer. Specifically, P-cadherin is impacting on α6 integrin expression directly, which then leads to a decrease in the expression of β4 integrin. In fact, it is described that the knock down of α6 integrin in the breast leads to a decrease in the expression of its partner, β4 integrin (Klinowska *et al.*, 2001). Conversely, when β4 integrin is knocked-down specifically in the breast, the expression of α6 integrin is not affected, most likely because α6 integrin can bind to its other partner, β1 integrin (Klinowska *et al.*, 2001, Klinowska *et al.*, 1999). Actually, in our cell models, the inhibition of β4 integrin alone did not have any effect in the expression of either P-cadherin or α6 integrin and no effect was observed in the functional properties defined by the mammosphere assay and cell invasion. A schematic representation of the hypothetical signalling cooperation between the two adhesion molecules is represented in **Figure 2.** We do not exclude the possibility that P-cadherin may additionally be regulating α6β1 integrin function, despite the fact that no effect in β1 integrin subunit expression was seen upon P-cadherin knock-down. The signalling molecules activated in response to α6β4 engagement (adhesion to laminin-332) include the phosphorylation and the consequent activation of Fak, Src and AKT. Importantly, it is conceivable that this activation cooperates with another signalling network described for P-cadherin in which Src phosphorylates the p120ctn that binds to the juxtamembrane domain of P-cadherin. Activated p120ctn is dislocated to the cytoplasm, where it can subsequently recruit and modulate the activity of the small GTPases Rac1, Cdc42 and RhoA, which control the actin cytoskeleton dynamics (unpublished data and (Taniuchi *et al.*, 2005)). Specifically, RhoA activation would be an interesting downstream molecule of the P-cadherin signalling

pathway, since we have observed that inhibiting P-cadherin expression affects stress fibre formation. The role of α6β4 integrin activation in stress fibres formation and small GTPases activation was not addressed yet.

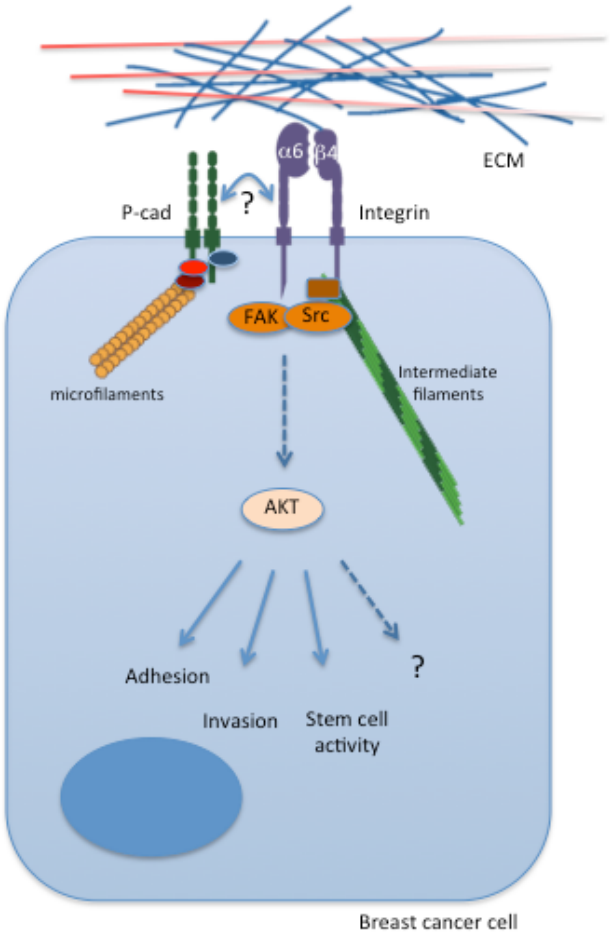


Figure 2 – Crosstalk between P-cadherin and α6β4 integrin. These signalling molecules cooperate leading to the phosphorylation and activation of FAK, Src and AKT, mediating important aggressive cancer properties *in vitro,* such as cell invasion and stem cell activity (mammosphere forming efficiency).

The relevance of P-cadherin in integrin signalling could be also attested by the fact that a subpopulation of cells depleted of P-cadherin expression had difficulties to grow in the 3D culture assays, as shown in Chapter V. This assay was performed with the cells growing on top of matrigel, which is a mixture of basement membrane ECM, being laminin a major constituent. The lack or the small size of multicellular structures formed in this assay indicated that P-cad^{low} cells had a decreased survival/growth. This observation is conceivably due to a lower integrin activity in these cells. However, the expression and activity of integrins was not yet assessed in this assay.

The current standard therapy to manage breast cancer is based on targeting the bulk of tumour cells, reducing tumour size. However this approach often does not

eliminate CSCs, which eventually may lead to tumour recurrence (**Figure 3**). It is anticipated that CSC-targeted therapies, in combination with conventional therapies, will provide a more effective treatment strategy (Ablett *et al.*, 2012). One of the potential ways to target breast CSCs is by the inhibition of the self-renewal signalling pathways thereby inducing differentiation or apoptosis and, as pointed out by our *in vitro* and *in vivo* assays, targeting the signalling axis P-cadherin/α6β4 integrin would allow a reduction in the stem cell activity of cancer cells. Thus, in the future, therapies directed at CSCs or the tumour microenvironment in the aggressive basal-like breast carcinomas can eventually involve the targeting of P-cadherin cell surface protein. Although the true clinical relevance of the CSC is yet to be revealed, there are tantalizing reports that the CSC can be selectively targeted without ablating normal stem cells. Interestingly, in 2008, Imai and collaborators have suggested *CDH3*/P-cadherin as a possible target for immunotherapy of pancreatic, gastric, and colorectal cancers, since it was identified as a novel tumour-associated antigen, meaning that was strongly expressed in tumour cells, but not in normal cells (Imai *et al.*, 2008). Thus, since P-cadherin is not frequently expressed in adult tissues, an anti-P-cadherin approach would be a valuable non-toxic therapy, as only malignant cells would be eliminated and the vital organs and normal tissues would not be critically affected.

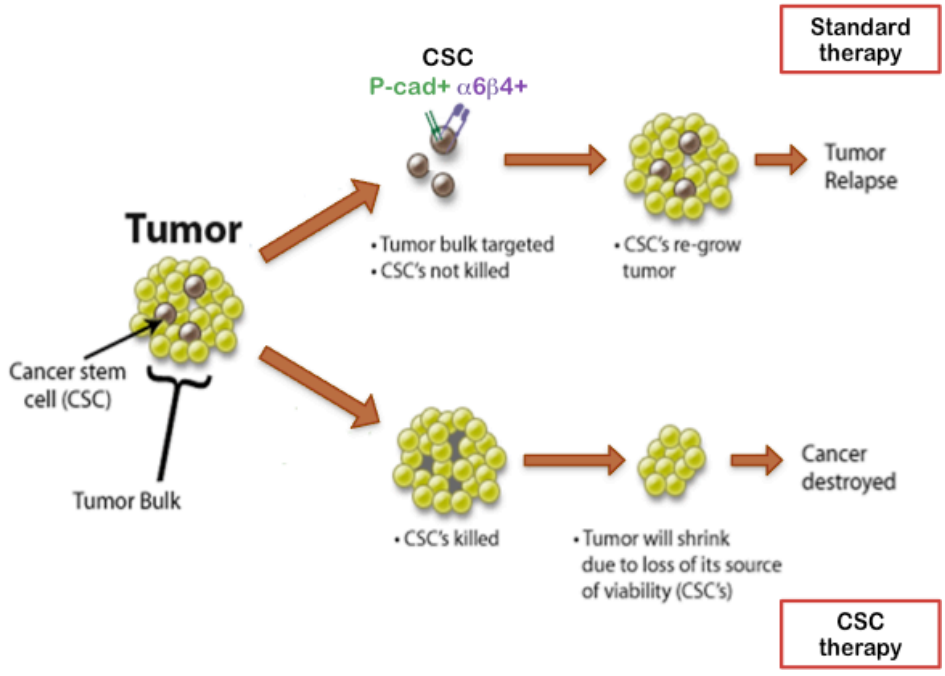


Figure 3 - Inhibiting the signalling pathway P-cadherin/a6b4 integrin may be a valuable CSC therapeutic target. Standard therapies are based in the use of drugs that reduce tumour size. Typical chemo, radio and hormone therapy target the tumour bulk and leave the CSCs unaffected, that have intrinsic molecular mechanisms that allow therapy resistance and tumour relapse.

Recently, a novel and highly selective human monoclonal antibody against P-cadherin (PF-03732010) was produced, demonstrating anti-tumour and anti-metastatic activity in a diverse panel of P-cadherin–overexpressing tumour models, without introducing any adverse secondary effects in mice (Zhang *et al.*, 2010). This antibody failed to bind to the most closely target-related family members, including E-cadherin, N-cadherin, and VE-cadherin. PF-03732010 also reduced lymph node metastases and lowered the levels of circulating tumour cells (CTC) in whole blood of P- cadherin^{+} tumour bearing mice. The anti-metastatic property of the antibody was remarkable, since it significantly inhibited tumour cell infiltration into the lungs. PF-03732010 still suppressed β-ctn, cyclin D1, vimentin, Bcl-2, Ki67 and survivin expression, and increased caspase-3 expression (Zhang *et al.*, 2010). In fact, in an attempt to reach the full potential for clinical development of the antibody, PF-03732010 has just completed a Phase I clinical trial development. The results from this first clinical trial will improve the knowledge of the antibody effective therapeutic dose; its toxicity and the pharmacokinetic and pharmacodynamic studies in humans. Preliminary reports provided by Pfizer (http://clinicaltrials.gov/ct2/show/results/NCT00557505) show that, in spite of the anti-tumour effects of PF-03732010 were not clear, the participants did not experience toxicity, even for the maximum administrated a weekly dose of 15 mg/kg.

2. Conclusions

Considering our initial aims and the data presented and discussed herein we can conclude the following:

1. The *in vitro* methods used to characterize stem cell properties identify cells in different stages of differentiation.

2. The basal-like phenotype is enriched in the stem cell markers CD44, CD49f and ALDH1 expression/activity in both *in vitro* and in a series of invasive human breast carcinomas. Basal mesenchymal cell lines were associated with the CSC phenotype $CD44^{+}CD24^{-/low}$, whereas basal epithelial cell lines were associated with the $CD44^{+}CD24^{+}$ and $CD49f^{+}CD24^{+}$ phenotypes.

3. The adhesion molecule P-cadherin is associated with increased stem cell activity in basal-like breast cancer cell lines, including the self-renewal capacity, growth in 3D matrix and the tumourigenicity in athymic nude mice.

4. P-cadherin is associated with already described stem cell markers (CD44, CD49f and CD24), namely with the phenotypes $CD44^{+}CD24^{+}$ and $CD49f^{+}CD24^{+}$.

5. P-cadherin inhibition sensitizes stem cells to X-ray-induced cell death.

6. The aggressive phenotype induced by P-cadherin is dependent, at least partially, on $\alpha6\beta4$ integrin activation.

7. P-cadherin oncogenic signalling cascade involves $\alpha6\beta4$ integrin engagement and the activation and phosphorylation of FAK, Src and AKT.

In conclusion, this study contributed to identify P-cadherin as cancer stem cell marker. Moreover, it unravelled a new molecular mechanism and the associated CSC effects mediated by this protein. Importantly, our results show that anti-P-cadherin treatment could improve radiation therapy in patients. This work reinforced the importance of P-cadherin expression as a prognostic factor for breast cancer patients, and supports the development of new therapeutics to control aggressive carcinomas that express this cadherin molecule.

FUTURE PERSPECTIVES

CHAPTER VIII

In this study, we described, for the first time, that P-cadherin mediates CSC properties linking several stem cell features to the aggressive behaviour of basal-like breast cancers. We also unravelled a novel oncogenic signalling pathway in which P-cadherin and $\alpha6\beta4$ integrin are closely involved, with a focus in the adhesion to the laminin substrate, the invasion capacity and mammosphere formation.

In this context, many interesting questions and challenges remain to be elucidated. In order to clarify P-cadherin functional and clinical relevance, we aim to gather knowledge in future studies, addressing the following topics:

1. We would like to clarify the role of P-cadherin as a marker of the luminal lineage in the normal breast hierarchy. Expanding this knowledge is of fundamental value and will enrich our notion of the physiology of the normal breast as well as breast cancer. P-cadherin will be combined with other markers already described in the literature to fractionate cell populations.

2. Our work is lacking studies using primary human mammary epithelial cells from normal breast tissue, *BRCA1* mutation carriers and basal-like breast tumours. This would help clarify the role of P-cadherin in the *in vivo* human setting.

3. Another important aspect is to clarify the role of P-cadherin in the EMT progression. This is a process through which an epithelial cell can give rise to stem cells (Mani *et al.*, 2008) and it may explain the plasticity observed in the hierarchy of breast differentiation. The role of EMT inducing agents (eg., microenvironmental factors, such as TGFβ) or the E-box transcriptional repressors Twist, Snail and Slug are very well described in the regulation of E-cadherin and N-cadherin, however, very little is known regarding P-cadherin transcription.

4. We described a major signalling pathway mediated by P-cadherin in cancer cells. This is of foremost importance if we are thinking of ways to target basal-like tumours, especially because no directed therapies exist for this entity of breast carcinomas. However, it is important to understand whether this signalling pathway is also present in normal mammary cells. If it is, does it mediate a tumour suppressive function or, on the other hand, was the P-cadherin/$\alpha6\beta4$ integrin signalling selected during tumour progression?

5. To dissect the molecular mechanisms that explain P-cadherin mediated resistance to X-ray induced DNA damage is fundamental to comprehend therapy resistance and, ideally, to improve radiotherapy regimens.

6. The optimization of a shRNA approach for the *CDH3*/P-cadherin gene, in a DOX inducible system, will provide a valuable cellular model to evaluate the role of P-cadherin in the origin and progression of basal-like breast cancer *in vivo*.

REFERENCES

Ablett, M.P., Singh, J.K. & Clarke, R.B. (2012) Stem cells in breast tumours: Are they ready for the clinic? *Eur J Cancer*.

Abraham, B.K., Fritz, P., Mcclellan, M., Hauptvogel, P., Athelogou, M. & Brauch, H. (2005) 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**,** 1154-9.

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J. & Clarke, M.F. (2003) Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A,* 100**,** 3983-8.

Albergaria, A., Ribeiro, A.S., Pinho, S., Milanezi, F., Carneiro, V., Sousa, B., Sousa, S., Oliveira, C., Machado, J.C., Seruca, R., Paredes, J. & Schmitt, F. (2010) ICI 182,780 induces P-cadherin overexpression in breast cancer cells through chromatin remodelling at the promoter level: a role for C/EBPbeta in CDH3 gene activation. *Hum Mol Genet,* 19**,** 2554-66.

Albergaria, A., Ribeiro, A.S., Vieira, A.F., Sousa, B., Nobre, A.R., Seruca, R., Schmitt, F. & Paredes, J. (2011) P-cadherin role in normal breast development and cancer. *The International journal of developmental biology,* 55**,** 811-822.

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

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

Allred, D.C., Wu, Y., Mao, S., Nagtegaal, I.D., Lee, S., Perou, C.M., Mohsin, S.K., O'connell, P., Tsimelzon, A. & Medina, D. (2008) Ductal carcinoma in situ and the emergence of diversity during breast cancer evolution. *Clin Cancer Res,* 14**,** 370-8.

Aplin, J.D., Jones, C.J. & Harris, L.K. (2009) Adhesion molecules in human trophoblast - a review. I. Villous trophoblast. *Placenta,* 30**,** 293-8.

Arnes, J.B., Brunet, J.S., Stefansson, I., Begin, L.R., Wong, N., Chappuis, P.O., Akslen, L.A. & Foulkes, W.D. (2005) Placental cadherin and the basal epithelial phenotype of BRCA1-related breast cancer. *Clin Cancer Res,* 11**,** 4003-11.

Asselin-Labat, M.L., Shackleton, M., Stingl, J., Vaillant, F., Forrest, N.C., Eaves, C.J., Visvader, J.E. & Lindeman, G.J. (2006) Steroid hormone receptor status of mouse mammary stem cells. *J Natl Cancer Inst,* 98**,** 1011-4.

Bachelder, R.E., Ribick, M.J., Marchetti, A., Falcioni, R., Soddu, S., Davis, K.R. & Mercurio, A.M. (1999) p53 inhibits alpha 6 beta 4 integrin survival signaling by promoting the caspase 3-dependent cleavage of AKT/PKB. *J Cell Biol,* 147**,** 1063-72.

Ball, S.M. (1998) The development of the terminal end bud in the prepubertal-pubertal mouse mammary gland. *The Anatomical record,* 250**,** 459-64.

Barkan, D., El Touny, L.H., Michalowski, A.M., Smith, J.A., Chu, I., Davis, A.S., Webster, J.D., Hoover, S., Simpson, R.M., Gauldie, J. & Green, J.E. (2010a) Metastatic growth from dormant cells induced by a col-I-enriched fibrotic environment. *Cancer Res,* 70**,** 5706-16.

Barkan, D., Green, J.E. & Chambers, A.F. (2010b) Extracellular matrix: a gatekeeper in the transition from dormancy to metastatic growth. *Eur J Cancer,* 46**,** 1181-8.

Batlle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J. & Garcia De Herreros, A. (2000) The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol,* 2**,** 84-9.

Baumann, P., Cremers, N., Kroese, F., Orend, G., Chiquet-Ehrismann, R., Uede, T., Yagita, H. & Sleeman, J.P. (2005) CD24 expression causes the acquisition of multiple cellular properties associated with tumor growth and metastasis. *Cancer Res,* 65**,** 10783-93.

Beachy, P.A., Karhadkar, S.S. & Berman, D.M. (2004) Tissue repair and stem cell renewal in carcinogenesis. *Nature,* 432**,** 324-31.

Beger, C., Pierce, L.N., Kruger, M., Marcusson, E.G., Robbins, J.M., Welcsh, P., Welch, P.J., Welte, K., King, M.C., Barber, J.R. & Wong-Staal, F. (2001) Identification of Id4 as a regulator of BRCA1 expression by using a ribozyme-library-based inverse genomics approach. *Proc Natl Acad Sci U S A,* 98**,** 130-5.

Bircan, S., Kapucuoglu, N., Baspinar, S., Inan, G. & Candir, O. (2006) CD24 expression in ductal carcinoma in situ and invasive ductal carcinoma of breast: an immunohistochemistry-based pilot study. *Pathol Res Pract,* 202**,** 569-76.

Bissell, M.J. & Hines, W.C. (2011) Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat Med,* 17**,** 320-9.

Bissell, M.J. & Labarge, M.A. (2005) Context, tissue plasticity, and cancer: are tumor stem cells also regulated by the microenvironment? *Cancer Cell,* 7**,** 17-23.

Bissell, M.J. & Radisky, D. (2001) Putting tumours in context. *Nat Rev Cancer,* 1**,** 46-54.

Bon, G., Folgiero, V., Di Carlo, S., Sacchi, A. & Falcioni, R. (2007) Involvement of alpha6beta4 integrin in the mechanisms that regulate breast cancer progression. *Breast Cancer Res,* 9**,** 203.

Bonnet, D. & Dick, J.E. (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med,* 3**,** 730-7.

Busk, M., Pytela, R. & Sheppard, D. (1992) Characterization of the integrin alpha v beta 6 as a fibronectin-binding protein. *J Biol Chem,* 267**,** 5790-6.

Bussemakers, M.J., Van Bokhoven, A., Voller, M., Smit, F.P. & Schalken, J.A. (1994) The genes for the calcium-dependent cell adhesion molecules P- and E-cadherin are tandemly arranged in the human genome. *Biochem Biophys Res Commun,* 203**,** 1291-4.

Cairns, J. (1975) The cancer problem. *Sci Am,* 233**,** 64-72, 77-8.

Cairns, J. (2002) Somatic stem cells and the kinetics of mutagenesis and carcinogenesis. *Proc Natl Acad Sci U S A,* 99**,** 10567-70.

Campbell, L.L. & Polyak, K. (2007) Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell Cycle,* 6**,** 2332-8.

Canonici, A., Steelant, W., Rigot, V., Khomitch-Baud, A., Boutaghou-Cherid, H., Bruyneel, E., Van Roy, F., Garrouste, F., Pommier, G. & Andre, F. (2008) Insulin-like growth factor-I receptor, E-cadherin and alpha v integrin form a dynamic complex under the control of alpha-catenin. *Int J Cancer,* 122**,** 572-82.

Cariati, M., Naderi, A., Brown, J.P., Smalley, M.J., Pinder, S.E., Caldas, C. & Purushotham, A.D. (2008) 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**,** 298-304.

Carpenter, P.M., Dao, A.V., Arain, Z.S., Chang, M.K., Nguyen, H.P., Arain, S., Wang-Rodriguez, J., Kwon, S.Y. & Wilczynski, S.P. (2009) Motility induction in breast carcinoma by mammary epithelial laminin 332 (laminin 5). *Mol Cancer Res,* 7**,** 462-75.

Cavallaro, U. & Christofori, G. (2004) Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. *Nat Rev Cancer,* 4**,** 118-32.

Cavallaro, U. & Dejana, E. (2011) Adhesion molecule signalling: not always a sticky business. *Nat Rev Mol Cell Biol,* 12**,** 189-97.

Chaffer, C.L., Brueckmann, I., Scheel, C., Kaestli, A.J., Wiggins, P.A., Rodrigues, L.O., Brooks, M., Reinhardt, F., Su, Y., Polyak, K., Arendt, L.M., Kuperwasser, C., Bierie, B. & Weinberg, R.A. (2011) Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proc Natl Acad Sci U S A,* 108**,** 7950-5.

Chanson, L., Brownfield, D., Garbe, J.C., Kuhn, I., Stampfer, M.R., Bissell, M.J. & Labarge, M.A. (2011) Self-organization is a dynamic and lineage-intrinsic property of mammary epithelial cells. *Proc Natl Acad Sci U S A,* 108**,** 3264-9.

Charafe-Jauffret, E., Ginestier, C. & Birnbaum, D. (2009) Breast cancer stem cells: tools and models one can rely on. *BMC Cancer,* 9**,** 202.

Charafe-Jauffret, E., Ginestier, C., Monville, F., Finetti, P., Adelaide, J., Cervera, N., Fekairi, S., Xerri, L., Jacquemier, J., Birnbaum, D. & Bertucci, F. (2006) Gene expression profiling of breast cell lines identifies potential new basal markers. *Oncogene,* 25**,** 2273-84.

Charafe-Jauffret, E., Monville, F., Ginestier, C., Dontu, G., Birnbaum, D. & Wicha, M.S. (2008) Cancer stem cells in breast: current opinion and future challenges. *Pathobiology,* 75**,** 75-84.

Chartier, N.T., Laine, M., Gout, S., Pawlak, G., Marie, C.A., Matos, P., Block, M.R. & Jacquier-Sarlin, M.R. (2006) Laminin-5-integrin interaction signals through PI 3-kinase and Rac1b to promote assembly of adherens junctions in HT-29 cells. *J Cell Sci,* 119**,** 31-46.

Chattopadhyay, N., Wang, Z., Ashman, L.K., Brady-Kalnay, S.M. & Kreidberg, J.A. (2003) alpha3beta1 integrin-CD151, a component of the cadherin-catenin complex, regulates PTPmu expression and cell-cell adhesion. *J Cell Biol,* 163**,** 1351-62.

Chen, M.S., Woodward, W.A., Behbod, F., Peddibhotla, S., Alfaro, M.P., Buchholz, T.A. & Rosen, J.M. (2007) Wnt/beta-catenin mediates radiation resistance of Sca1+ progenitors in an immortalized mammary gland cell line. *J Cell Sci,* 120**,** 468-77.

Chepko, G. & Dickson, R.B. (2003) Ultrastructure of the putative stem cell niche in rat mammary epithelium. *Tissue Cell,* 35**,** 83-93.

Cheung, L.W., Leung, P.C. & Wong, A.S. (2010) Cadherin switching and activation of p120 catenin signaling are mediators of gonadotropin-releasing hormone to promote tumor cell migration and invasion in ovarian cancer. *Oncogene,* 29**,** 2427-40.

Chin, K., De Solorzano, C.O., Knowles, D., Jones, A., Chou, W., Rodriguez, E.G., Kuo, W.L., Ljung, B.M., Chew, K., Myambo, K., Miranda, M., Krig, S., Garbe, J., Stampfer, M., Yaswen, P., Gray, J.W. & Lockett, S.J. (2004) In situ analyses of genome instability in breast cancer. *Nat Genet,* 36**,** 984-8.

Christofori, G. (2006) New signals from the invasive front. *Nature,* 441**,** 444-50.

Comijn, J., Berx, G., Vermassen, P., Verschueren, K., Van Grunsven, L., Bruyneel, E., Mareel, M., Huylebroeck, D. & Van Roy, F. (2001) The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Molecular cell,* 7**,** 1267-78.

Comoglio, P.M., Boccaccio, C. & Trusolino, L. (2003) Interactions between growth factor receptors and adhesion molecules: breaking the rules. *Curr Opin Cell Biol,* 15**,** 565-71.

Croker, A.K., Goodale, D., Chu, J., Postenka, C., Hedley, B.D., Hess, D.A. & Allan, A.L. (2009) 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**,** 2236-52.

Da Silva, L., Clarke, C. & Lakhani, S.R. (2007) Demystifying basal-like breast carcinomas. *J Clin Pathol,* 60**,** 1328-32.

Daniel, C.W., Strickland, P. & Friedmann, Y. (1995) Expression and functional role of E- and P-cadherins in mouse mammary ductal morphogenesis and growth. *Dev Biol,* 169**,** 511-9.

De Lange, J., Verlaan-De Vries, M., Teunisse, A.F. & Jochemsen, A.G. (2012) Chk2 mediates RITA-induced apoptosis. *Cell Death Differ,* 19**,** 980-9.

De Leeuw, W.J., Berx, G., Vos, C.B., Peterse, J.L., Van De Vijver, M.J., Litvinov, S., Van Roy, F., Cornelisse, C.J. & Cleton-Jansen, A.M. (1997) Simultaneous loss of E-cadherin and catenins in invasive lobular breast cancer and lobular carcinoma in situ. *J Pathol,* 183**,** 404-11.

Deng, S., Yang, X., Lassus, H., Liang, S., Kaur, S., Ye, Q., Li, C., Wang, L.P., Roby, K.F., Orsulic, S., Connolly, D.C., Zhang, Y., Montone, K., Butzow, R., Coukos, G. & Zhang, L. (2010) Distinct expression levels and patterns of stem cell marker, aldehyde dehydrogenase isoform 1 (ALDH1), in human epithelial cancers. *PLoS One,* 5**,** e10277.

Deugnier, M.A., Faraldo, M.M., Rousselle, P., Thiery, J.P. & Glukhova, M.A. (1999) Cell-extracellular matrix interactions and EGF are important regulators of the basal mammary epithelial cell phenotype. *J Cell Sci,* 112 (Pt 7)**,** 1035-44.

Di Croce, L. & Pelicci, P.G. (2003) Tumour-associated hypermethylation: silencing E-cadherin expression enhances invasion and metastasis. *Eur J Cancer,* 39**,** 413-4.

Dontu, G., Abdallah, W.M., Foley, J.M., Jackson, K.W., Clarke, M.F., Kawamura, M.J. & Wicha, M.S. (2003) In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev,* 17**,** 1253-70.

Edwards, S.L., Brough, R., Lord, C.J., Natrajan, R., Vatcheva, R., Levine, D.A., Boyd, J., Reis-Filho, J.S. & Ashworth, A. (2008) Resistance to therapy caused by intragenic deletion in BRCA2. *Nature,* 451**,** 1111-5.

Egan, C., Pang, A., Durda, D., Cheng, H.C., Wang, J.H. & Fujita, D.J. (1999) Activation of Src in human breast tumor cell lines: elevated levels of phosphotyrosine phosphatase activity that preferentially recognizes the Src carboxy terminal negative regulatory tyrosine 530. *Oncogene,* 18**,** 1227-37.

Eirew, P., Kannan, N., Knapp, D.J., Vaillant, F., Emerman, J.T., Lindeman, G.J., Visvader, J.E. & Eaves, C.J. (2011) Aldehyde Dehydrogenase Activity is a Biomarker of Primitive Normal Human Mammary Luminal Cells. *Stem cells (Dayton, Ohio)*.

Eirew, P., Stingl, J., Raouf, A., Turashvili, G., Aparicio, S., Emerman, J.T. & Eaves, C.J. (2008) A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. *Nat Med,* 14**,** 1384-9.

Faraldo, M.M., Teuliere, J., Deugnier, M.A., Birchmeier, W., Huelsken, J., Thiery, J.P., Cano, A. & Glukhova, M.A. (2007) beta-Catenin regulates P-cadherin expression in mammary basal epithelial cells. *FEBS letters,* 581**,** 831-6.

Farmer, H., Mccabe, N., Lord, C.J., Tutt, A.N., Johnson, D.A., Richardson, T.B., Santarosa, M., Dillon, K.J., Hickson, I., Knights, C., Martin, N.M., Jackson, S.P., Smith, G.C. & Ashworth, A. (2005) Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature,* 434**,** 917-21.

Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C. & Parkin, D.M. (2010) GLOBOCAN 2008 v1.2, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet]. *International Agency for Research on Cancer*.

Fillmore, C.M. & Kuperwasser, C. (2008) 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**,** R25.

Fogel, M., Friederichs, J., Zeller, Y., Husar, M., Smirnov, A., Roitman, L., Altevogt, P. & Sthoeger, Z.M. (1999) CD24 is a marker for human breast carcinoma. *Cancer letters,* 143**,** 87-94.

Foulkes, W.D. (2004) BRCA1 functions as a breast stem cell regulator. *J Med Genet,* 41**,** 1-5.

Foulkes, W.D., Stefansson, I.M., Chappuis, P.O., Begin, L.R., Goffin, J.R., Wong, N., Trudel, M. & Akslen, L.A. (2003) Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. *J Natl Cancer Inst,* 95**,** 1482-5.

Friedrichs, K., Ruiz, P., Franke, F., Gille, I., Terpe, H.J. & Imhof, B.A. (1995) High expression level of alpha 6 integrin in human breast carcinoma is correlated with reduced survival. *Cancer Res,* 55**,** 901-6.

Frixen, U.H., Behrens, J., Sachs, M., Eberle, G., Voss, B., Warda, A., Lochner, D. & Birchmeier, W. (1991) E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. *J Cell Biol,* 113**,** 173-85.

Fujita, M., Furukawa, F., Fujii, K., Horiguchi, Y., Takeichi, M. & Imamura, S. (1992) Expression of cadherin cell adhesion molecules during human skin development: morphogenesis of epidermis, hair follicles and eccrine sweat ducts. *Archives of dermatological research,* 284**,** 159-66.

Fulford, L.G., Reis-Filho, J.S., Ryder, K., Jones, C., Gillett, C.E., Hanby, A., Easton, D. & Lakhani, S.R. (2007) Basal-like grade III invasive ductal carcinoma of the breast: patterns of metastasis and long-term survival. *Breast Cancer Res,* 9**,** R4.

Gamallo, C., Moreno-Bueno, G., Sarrio, D., Calero, F., Hardisson, D. & Palacios, J. (2001) The prognostic significance of P-cadherin in infiltrating ductal breast carcinoma. *Mod Pathol,* 14**,** 650-4.

Gardner, H., Kreidberg, J., Koteliansky, V. & Jaenisch, R. (1996) Deletion of integrin alpha 1 by homologous recombination permits normal murine development but gives rise to a specific deficit in cell adhesion. *Dev Biol,* 175**,** 301-13.

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

Gorski, J.J., James, C.R., Quinn, J.E., Stewart, G.E., Staunton, K.C., Buckley, N.E., Mcdyer, F.A., Kennedy, R.D., Wilson, R.H., Mullan, P.B. & Harkin, D.P. (2009) BRCA1 transcriptionally regulates genes associated with the basal-like phenotype in breast cancer. *Breast Cancer Res Treat*.

Goss, P., Allan, A.L., Rodenhiser, D.I., Foster, P.J. & Chambers, A.F. (2008) New clinical and experimental approaches for studying tumor dormancy: does tumor dormancy offer a therapeutic target? *Apmis,* 116**,** 552-68.

Gudjonsson, T., Ronnov-Jessen, L., Villadsen, R., Rank, F., Bissell, M.J. & Petersen, O.W. (2002) Normal and tumor-derived myoepithelial cells differ in their ability to interact with luminal breast epithelial cells for polarity and basement membrane deposition. *J Cell Sci,* 115**,** 39-50.

Guilford, P., Hopkins, J., Harraway, J., Mcleod, M., Mcleod, N., Harawira, P., Taite, H., Scoular, R., Miller, A. & Reeve, A.E. (1998) E-cadherin germline mutations in familial gastric cancer. *Nature,* 392**,** 402-5.

Guo, W., Pylayeva, Y., Pepe, A., Yoshioka, T., Muller, W.J., Inghirami, G. & Giancotti, F.G. (2006) Beta 4 integrin amplifies ErbB2 signaling to promote mammary tumorigenesis. *Cell,* 126**,** 489-502.

Gupta, P.B., Fillmore, C.M., Jiang, G., Shapira, S.D., Tao, K., Kuperwasser, C. & Lander, E.S. (2011) Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell,* 146**,** 633-44.

Gusterson, B. (2009) Do 'basal-like' breast cancers really exist? *Nat Rev Cancer,* 9**,** 128-34.

Hajra, K.M., Chen, D.Y. & Fearon, E.R. (2002) The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res,* 62**,** 1613-8.

Hardy, R.G., Tselepis, C., Hoyland, J., Wallis, Y., Pretlow, T.P., Talbot, I., Sanders, D.S., Matthews, G., Morton, D. & Jankowski, J.A. (2002) Aberrant P-cadherin expression is an early event in hyperplastic and dysplastic transformation in the colon. *Gut,* 50**,** 513-9.

Hennessy, B.T., Gonzalez-Angulo, A.M., Stemke-Hale, K., Gilcrease, M.Z., Krishnamurthy, S., Lee, J.S., Fridlyand, J., Sahin, A., Agarwal, R., Joy, C., Liu, W., Stivers, D., Baggerly, K., Carey, M., Lluch, A., Monteagudo, C., He, X., Weigman, V., Fan, C., Palazzo, J., Hortobagyi, G.N., Nolden, L.K., Wang, N.J., Valero, V., Gray, J.W., Perou, C.M. & Mills, G.B. (2009) Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res,* 69**,** 4116-24.

Hennighausen, L. & Robinson, G.W. (2001) Signaling pathways in mammary gland development. *Developmental cell,* 1**,** 467-75.

Herschkowitz, J.I., Simin, K., Weigman, V.J., Mikaelian, I., Usary, J., Hu, Z., Rasmussen, K.E., Jones, L.P., Assefnia, S., Chandrasekharan, S., Backlund, M.G., Yin, Y., Khramtsov, A.I., Bastein, R., Quackenbush, J., Glazer, R.I., Brown, P.H., Green, J.E., Kopelovich, L., Furth, P.A., Palazzo, J.P., Olopade, O.I., Bernard, P.S., Churchill, G.A., Van Dyke, T. & Perou, C.M. (2007) Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome biology,* 8**,** R76.

Hirai, Y., Nose, A., Kobayashi, S. & Takeichi, M. (1989a) Expression and role of E- and P-cadherin adhesion molecules in embryonic histogenesis. I. Lung epithelial morphogenesis. *Development,* 105**,** 263-70.

Hirai, Y., Nose, A., Kobayashi, S. & Takeichi, M. (1989b) Expression and role of E- and P-cadherin adhesion molecules in embryonic histogenesis. II. Skin morphogenesis. *Development,* 105**,** 271-7.

Honeth, G., Bendahl, P.O., Ringner, M., Saal, L.H., Gruvberger-Saal, S.K., Lovgren, K., Grabau, D., Ferno, M., Borg, A. & Hegardt, C. (2008) The CD44+/CD24- phenotype is enriched in basal-like breast tumors. *Breast Cancer Res,* 10**,** R53.

Honrado, E., Osorio, A., Milne, R.L., Paz, M.F., Melchor, L., Cascon, A., Urioste, M., Cazorla, A., Diez, O., Lerma, E., Esteller, M., Palacios, J. & Benitez, J. (2007) Immunohistochemical classification of non-BRCA1/2 tumors identifies different groups that demonstrate the heterogeneity of BRCAX families. *Mod Pathol,* 20**,** 1298-306.

Hulpiau, P. & Van Roy, F. (2009) Molecular evolution of the cadherin superfamily. *The international journal of biochemistry & cell biology,* 41**,** 349-69.

Hutvagner, G. & Zamore, P.D. (2002) RNAi: nature abhors a double-strand. *Curr Opin Genet Dev,* 12**,** 225-32.

Hwang, E.S., Devries, S., Chew, K.L., Moore, D.H., 2nd, Kerlikowske, K., Thor, A., Ljung, B.M. & Waldman, F.M. (2004) Patterns of chromosomal alterations in breast ductal carcinoma in situ. *Clin Cancer Res,* 10**,** 5160-7.

Hwang-Verslues, W.W., Chang, K.J., Lee, E.Y. & Lee, W.H. (2008) Breast cancer stem cells and tumor suppressor genes. *J Formos Med Assoc,* 107**,** 751-66.

Hwang-Verslues, W.W., Kuo, W.H., Chang, P.H., Pan, C.C., Wang, H.H., Tsai, S.T., Jeng, Y.M., Shew, J.Y., Kung, J.T., Chen, C.H., Lee, E.Y., Chang, K.J. & Lee, W.H. (2009) Multiple lineages of human breast cancer stem/progenitor cells identified by profiling with stem cell markers. *PLoS One,* 4**,** e8377.

Illmensee, K. & Mintz, B. (1976) Totipotency and normal differentiation of single teratocarcinoma cells cloned by injection into blastocysts. *Proc Natl Acad Sci U S A,* 73**,** 549-53.

Imai, K., Hirata, S., Irie, A., Senju, S., Ikuta, Y., Yokomine, K., Harao, M., Inoue, M., Tsunoda, T., Nakatsuru, S., Nakagawa, H., Nakamura, Y., Baba, H. & Nishimura, Y. (2008) Identification of a novel tumor-associated antigen, cadherin 3/P-cadherin, as a possible target for immunotherapy of pancreatic, gastric, and colorectal cancers. *Clin Cancer Res,* 14**,** 6487-95.

Irby, R.B. & Yeatman, T.J. (2000) Role of Src expression and activation in human cancer. *Oncogene,* 19**,** 5636-42.

Itano, N. & Kimata, K. (2008) Altered hyaluronan biosynthesis in cancer progression. *Seminars in cancer biology,* 18**,** 268-74.

Jacobs, C. & Rubsamen, H. (1983) Expression of pp60c-src protein kinase in adult and fetal human tissue: high activities in some sarcomas and mammary carcinomas. *Cancer Res,* 43**,** 1696-702.

Kaplan, R.N., Psaila, B. & Lyden, D. (2006) Bone marrow cells in the 'pre-metastatic niche': within bone and beyond. *Cancer metastasis reviews,* 25**,** 521-9.

Kastan, M.B., Schlaffer, E., Russo, J.E., Colvin, O.M., Civin, C.I. & Hilton, J. (1990) Direct demonstration of elevated aldehyde dehydrogenase in human hematopoietic progenitor cells. *Blood,* 75**,** 1947-50.

Keller, P.J., Arendt, L.M., Skibinski, A., Logvinenko, T., Klebba, I., Dong, S., Smith, A.E., Prat, A., Perou, C.M., Gilmore, H., Schnitt, S., Naber, S.P., Garlick, J.A. & Kuperwasser, C. (2011) Defining the cellular precursors to human breast cancer. *Proc Natl Acad Sci U S A*.

Keller, P.J., Lin, A.F., Arendt, L.M., Klebba, I., Jones, A.D., Rudnick, J.A., Dimeo, T.A., Gilmore, H., Jefferson, D.M., Graham, R.A., Naber, S.P., Schnitt, S. & Kuperwasser, C. (2010) Mapping the cellular and molecular heterogeneity of normal and malignant breast tissues and cultured cell lines. *Breast Cancer Res,* 12**,** R87.

Kenny, P.A., Lee, G.Y., Myers, C.A., Neve, R.M., Semeiks, J.R., Spellman, P.T., Lorenz, K., Lee, E.H., Barcellos-Hoff, M.H., Petersen, O.W., Gray, J.W. & Bissell, M.J. (2007) The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression. *Mol Oncol,* 1**,** 84-96.

Kim, B.G., An, H.J., Kang, S., Choi, Y.P., Gao, M.Q., Park, H. & Cho, N.H. (2011) Laminin-332-rich tumor microenvironment for tumor invasion in the interface zone of breast cancer. *Am J Pathol,* 178**,** 373-81.

Kim, H., Laing, M. & Muller, W. (2005) c-Src-null mice exhibit defects in normal mammary gland development and ERalpha signaling. *Oncogene,* 24**,** 5629-36.

Kjaer, K.W., Hansen, L., Schwabe, G.C., Marques-De-Faria, A.P., Eiberg, H., Mundlos, S., Tommerup, N. & Rosenberg, T. (2005) Distinct CDH3 mutations cause ectodermal dysplasia, ectrodactyly, macular dystrophy (EEM syndrome). *J Med Genet,* 42**,** 292-8.

Klinowska, T.C., Alexander, C.M., Georges-Labouesse, E., Van Der Neut, R., Kreidberg, J.A., Jones, C.J., Sonnenberg, A. & Streuli, C.H. (2001) Epithelial development and differentiation in the mammary gland is not dependent on alpha 3 or alpha 6 integrin subunits. *Dev Biol,* 233**,** 449-67.

Klinowska, T.C., Soriano, J.V., Edwards, G.M., Oliver, J.M., Valentijn, A.J., Montesano, R. & Streuli, C.H. (1999) Laminin and beta1 integrins are crucial for normal mammary gland development in the mouse. *Dev Biol,* 215**,** 13-32.

Knudsen, K.A. & Wheelock, M.J. (2005) Cadherins and the mammary gland. *J Cell Biochem,* 95**,** 488-96.

Kolle, G., Ho, M., Zhou, Q., Chy, H.S., Krishnan, K., Cloonan, N., Bertoncello, I., Laslett, A.L. & Grimmond, S.M. (2009) Identification of human embryonic stem cell surface markers by combined membrane-polysome translation state array analysis and immunotranscriptional profiling. *Stem cells (Dayton, Ohio),* 27**,** 2446-56.

Korkaya, H., Paulson, A., Iovino, F. & Wicha, M.S. (2008) HER2 regulates the mammary stem/progenitor cell population driving tumorigenesis and invasion. *Oncogene,* 27**,** 6120-30.

Korsching, E., Jeffrey, S.S., Meinerz, W., Decker, T., Boecker, W. & Buerger, H. (2008) Basal carcinoma of the breast revisited: an old entity with new interpretations. *J Clin Pathol,* 61**,** 553-60.

Kristiansen, G., Winzer, K.J., Mayordomo, E., Bellach, J., Schluns, K., Denkert, C., Dahl, E., Pilarsky, C., Altevogt, P., Guski, H. & Dietel, M. (2003) CD24 expression is a new prognostic marker in breast cancer. *Clin Cancer Res,* 9**,** 4906-13.

Kufe, D.W., Pollock, R.E., Weichselbaum, R.R., Bast, R.C., Gansler, T.S., Holland, J.F. & Frei, E. (2010) *Holland-Frei Cancer Medicine.* Hamilton (ON): McGraw-Hill Medical.

Kwon, S.Y., Chae, S.W., Wilczynski, S.P., Arain, A., Carpenter & Philip, M. (2012) Laminin 332 expression in breast carcinoma. *Appl Immunohistochem Mol Morphol,* 20**,** 159-64.

Lakhani, S.R., Jacquemier, J., Sloane, J.P., Gusterson, B.A., Anderson, T.J., Van De Vijver, M.J., Farid, L.M., Venter, D., Antoniou, A., Storfer-Isser, A., Smyth, E., Steel, C.M., Haites, N., Scott, R.J., Goldgar, D., Neuhausen, S., Daly, P.A., Ormiston, W., Mcmanus, R., Scherneck, S., Ponder, B.A., Ford, D., Peto, J., Stoppa-Lyonnet, D., Bignon, Y.J., Struewing, J.P., Spurr, N.K., Bishop, D.T., Klijn, J.G., Devilee, P., Cornelisse, C.J., Lasset, C., Lenoir, G., Barkardottir, R.B., Egilsson, V., Hamann, U., Chang-Claude, J., Sobol, H., Weber, B., Stratton, M.R. & Easton, D.F. (1998) Multifactorial analysis of differences between sporadic breast cancers and cancers involving BRCA1 and BRCA2 mutations. *J Natl Cancer Inst,* 90**,** 1138-45.

Lakhani, S.R., Reis-Filho, J.S., Fulford, L., Penault-Llorca, F., Van Der Vijver, M., Parry, S., Bishop, T., Benitez, J., Rivas, C., Bignon, Y.J., Chang-Claude, J., Hamann, U., Cornelisse, C.J., Devilee, P., Beckmann, M.W., Nestle-Kramling, C., Daly, P.A.,

Haites, N., Varley, J., Lalloo, F., Evans, G., Maugard, C., Meijers-Heijboer, H., Klijn, J.G., Olah, E., Gusterson, B.A., Pilotti, S., Radice, P., Scherneck, S., Sobol, H., Jacquemier, J., Wagner, T., Peto, J., Stratton, M.R., Mcguffog, L. & Easton, D.F. (2005) Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. *Clin Cancer Res,* 11**,** 5175-80.

Lakhani, S.R., Van De Vijver, M.J., Jacquemier, J., Anderson, T.J., Osin, P.P., Mcguffog, L. & Easton, D.F. (2002) The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J Clin Oncol,* 20**,** 2310-8.

Lambie, H., Miremadi, A., Pinder, S.E., Bell, J.A., Wencyk, P., Paish, E.C., Macmillan, R.D. & Ellis, I.O. (2003) Prognostic significance of BRCA1 expression in sporadic breast carcinomas. *J Pathol,* 200**,** 207-13.

Larue, L., Antos, C., Butz, S., Huber, O., Delmas, V., Dominis, M. & Kemler, R. (1996) A role for cadherins in tissue formation. *Development,* 122**,** 3185-94.

Li, X., Lewis, M.T., Huang, J., Gutierrez, C., Osborne, C.K., Wu, M.F., Hilsenbeck, S.G., Pavlick, A., Zhang, X., Chamness, G.C., Wong, H., Rosen, J. & Chang, J.C. (2008) Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J Natl Cancer Inst,* 100**,** 672-9.

Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A., Feleppa, F., Huschtscha, L.I., Thorne, H.J., Fox, S.B., Yan, M., French, J.D., Brown, M.A., Smyth, G.K., Visvader, J.E. & Lindeman, G.J. (2009) Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med,* 15**,** 907-13.

Liu, S., Ginestier, C., Charafe-Jauffret, E., Foco, H., Kleer, C.G., Merajver, S.D., Dontu, G. & Wicha, M.S. (2008) BRCA1 regulates human mammary stem/progenitor cell fate. *Proc Natl Acad Sci U S A,* 105**,** 1680-5.

Lu, S., Simin, K., Khan, A. & Mercurio, A.M. (2008) Analysis of integrin beta4 expression in human breast cancer: association with basal-like tumors and prognostic significance. *Clin Cancer Res,* 14**,** 1050-8.

Luo, M., Fan, H., Nagy, T., Wei, H., Wang, C., Liu, S., Wicha, M.S. & Guan, J.L. (2009) Mammary epithelial-specific ablation of the focal adhesion kinase suppresses mammary tumorigenesis by affecting mammary cancer stem/progenitor cells. *Cancer Res,* 69**,** 466-74.

Macdonald, F., Ford, C.H.J. & Casson, A.G. (2004) Molecular biology of cancer. 2nd edition London; New York, N.Y. BIOS Scientific Publishers.

Mandeville, J.A., Silva Neto, B., Vanni, A.J., Smith, G.L., Rieger-Christ, K.M., Zeheb, R., Loda, M., Libertino, J.A. & Summerhayes, I.C. (2008) P-cadherin as a prognostic indicator and a modulator of migratory behaviour in bladder carcinoma cells. *BJU international,* 102**,** 1707-14.

Mani, S.A., Guo, W., Liao, M.J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., Campbell, L.L., Polyak, K., Brisken, C., Yang, J. & Weinberg, R.A. (2008) The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell,* 133**,** 704-15.

Mannello, F., Tonti, G.A., Medda, V., Pederzoli, A. & Sauter, E.R. (2008) Increased shedding of soluble fragments of P-cadherin in nipple aspirate fluids from women with breast cancer. *Cancer science,* 99**,** 2160-9.

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

Marhaba, R. & Zoller, M. (2004) CD44 in cancer progression: adhesion, migration and growth regulation. *J Mol Histol,* 35**,** 211-31.

Matos, I., Dufloth, R., Alvarenga, M., Zeferino, L.C. & Schmitt, F. (2005) p63, cytokeratin 5, and P-cadherin: three molecular markers to distinguish basal phenotype in breast carcinomas. *Virchows Arch,* 447**,** 688-94.

Mercurio, A.M., Bachelder, R.E., Chung, J., O'connor, K.L., Rabinovitz, I., Shaw, L.M. & Tani, T. (2001) Integrin laminin receptors and breast carcinoma progression. *J Mammary Gland Biol Neoplasia,* 6**,** 299-309.

Meyer, M.J., Fleming, J.M., Ali, M.A., Pesesky, M.W., Ginsburg, E. & Vonderhaar, B.K. (2009) Dynamic regulation of CD24 and the invasive, CD44posCD24neg phenotype in breast cancer cell lines. *Breast Cancer Res,* 11**,** R82.

Meyer, M.J., Fleming, J.M., Lin, A.F., Hussnain, S.A., Ginsburg, E. & Vonderhaar, B.K. (2010) CD44posCD49fhiCD133/2hi defines xenograft-initiating cells in estrogen receptor-negative breast cancer. *Cancer Res,* 70**,** 4624-33.

Molino, A., Pedersini, R., Micciolo, R., Frisinghelli, M., Giovannini, M., Pavarana, M., Santo, A., Colato, C., Piubello, Q. & Cetto, G.L. (2003) Prognostic significance of laminin, laminin receptor, and bone marrow micrometastases in breast cancer patients: are these markers of aggressive behavior and metastatic potential? *Appl Immunohistochem Mol Morphol,* 11**,** 311-8.

Molyneux, G., Geyer, F.C., Magnay, F.A., Mccarthy, A., Kendrick, H., Natrajan, R., Mackay, A., Grigoriadis, A., Tutt, A., Ashworth, A., Reis-Filho, J.S. & Smalley, M.J. (2010) BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell,* 7**,** 403-17.

Molyneux, G., Regan, J. & Smalley, M.J. (2007) Mammary stem cells and breast cancer. *Cell Mol Life Sci,* 64**,** 3248-60.

Morel, A.P., Lievre, M., Thomas, C., Hinkal, G., Ansieau, S. & Puisieux, A. (2008) Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS One,* 3**,** e2888.

Morimoto, K., Kim, S.J., Tanei, T., Shimazu, K., Tanji, Y., Taguchi, T., Tamaki, Y., Terada, N. & Noguchi, S. (2009) 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 science,* 100**,** 1062-8.

Muschler, J. & Streuli, C.H. (2011) Cell-matrix interactions in mammary gland development and breast cancer. *Cold Spring Harbor perspectives in biology,* 2**,** a003202.

Mylona, E., Giannopoulou, I., Fasomytakis, E., Nomikos, A., Magkou, C., Bakarakos, P. & Nakopoulou, L. (2008) The clinicopathologic and prognostic significance of CD44+/CD24(-/low) and CD44-/CD24+ tumor cells in invasive breast carcinomas. *Hum Pathol,* 39**,** 1096-102.

Nagy, T., Wei, H., Shen, T.L., Peng, X., Liang, C.C., Gan, B. & Guan, J.L. (2007) Mammary epithelial-specific deletion of the focal adhesion kinase gene leads to severe lobulo-alveolar hypoplasia and secretory immaturity of the murine mammary gland. *J Biol Chem,* 282**,** 31766-76.

Nawrocki-Raby, B., Gilles, C., Polette, M., Bruyneel, E., Laronze, J.Y., Bonnet, N., Foidart, J.M., Mareel, M. & Birembaut, P. (2003) Upregulation of MMPs by soluble E-cadherin in human lung tumor cells. *Int J Cancer,* 105**,** 790-5.

Neve, R.M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F.L., Fevr, T., Clark, L., Bayani, N., Coppe, J.P., Tong, F., Speed, T., Spellman, P.T., Devries, S., Lapuk, A., Wang, N.J., Kuo, W.L., Stilwell, J.L., Pinkel, D., Albertson, D.G., Waldman, F.M., Mccormick, F., Dickson, R.B., Johnson, M.D., Lippman, M., Ethier, S., Gazdar, A. & Gray, J.W. (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell,* 10**,** 515-27.

Nielsen, T.O., Hsu, F.D., Jensen, K., Cheang, M., Karaca, G., Hu, Z., Hernandez-Boussard, T., Livasy, C., Cowan, D., Dressler, L., Akslen, L.A., Ragaz, J., Gown, A.M., Gilks, C.B., Van De Rijn, M. & Perou, C.M. (2004) Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res,* 10**,** 5367-74.

Nollet, F., Kools, P. & Van Roy, F. (2000) Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. *Journal of molecular biology,* 299**,** 551-72.

Nose, A. & Takeichi, M. (1986) A novel cadherin cell adhesion molecule: its expression patterns associated with implantation and organogenesis of mouse embryos. *J Cell Biol,* 103**,** 2649-58.

Nowell, P.C. (1976) The clonal evolution of tumor cell populations. *Science,* 194**,** 23-8.

O'brien, C.A., Pollett, A., Gallinger, S. & Dick, J.E. (2007) A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature,* 445**,** 106-10.

O'brien, C.S., Farnie, G., Howell, S.J. & Clarke, R.B. (2011) Breast cancer stem cells and their role in resistance to endocrine therapy. *Hormones & cancer,* 2**,** 91-103.

O'connor, K.L. & Mercurio, A.M. (2001) Protein kinase A regulates Rac and is required for the growth factor-stimulated migration of carcinoma cells. *J Biol Chem,* 276**,** 47895-900.

Olsson, E., Honeth, G., Bendahl, P.O., Saal, L.H., Gruvberger-Saal, S., Ringner, M., Vallon-Christersson, J., Jonsson, G., Holm, K., Lovgren, K., Ferno, M., Grabau, D., Borg, A. & Hegardt, C. (2011) CD44 isoforms are heterogeneously expressed in breast cancer and correlate with tumor subtypes and cancer stem cell markers. *BMC Cancer,* 11**,** 418.

Osborne, M.P. (2000) Breast anatomy and development. *In: Harris JR, Lippman ME, Morrow M, Osborne CK, editors. Diseases of the breast. Philadelphia,* Lippincott Williams & Wilkins; P.1-14.

Ouyang, G., Yao, L., Ruan, K., Song, G., Mao, Y. & Bao, S. (2009) Genistein induces G2/M cell cycle arrest and apoptosis of human ovarian cancer cells via activation of DNA damage checkpoint pathways. *Cell Biol Int,* 33**,** 1237-44.

Overall, C.M. & Kleifeld, O. (2006) Tumour microenvironment - opinion: validating matrix metalloproteinases as drug targets and anti-targets for cancer therapy. *Nat Rev Cancer,* 6**,** 227-39.

Owens, L.V., Xu, L., Craven, R.J., Dent, G.A., Weiner, T.M., Kornberg, L., Liu, E.T. & Cance, W.G. (1995) Overexpression of the focal adhesion kinase (p125FAK) in invasive human tumors. *Cancer Res,* 55**,** 2752-5.

Paredes, J., Albergaria, A., Oliveira, J.T., Jeronimo, C., Milanezi, F. & Schmitt, F.C. (2005) P-cadherin overexpression is an indicator of clinical outcome in invasive breast carcinomas and is associated with CDH3 promoter hypomethylation. *Clin Cancer Res,* 11**,** 5869-77.

Paredes, J., Correia, A.L., Ribeiro, A.S., Albergaria, A., Milanezi, F. & Schmitt, F.C. (2007a) P-cadherin expression in breast cancer: a review. *Breast Cancer Res,* 9**,** 214.

Paredes, J., Correia, A.L., Ribeiro, A.S., Milanezi, F., Cameselle-Teijeiro, J. & Schmitt, F.C. (2008) Breast carcinomas that co-express E- and P-cadherin are associated with p120-catenin cytoplasmic localisation and poor patient survival. *J Clin Pathol,* 61**,** 856-62.

Paredes, J., Figueiredo, J., Albergaria, A., Oliveira, P., Carvalho, J., Ribeiro, A.S., Caldeira, J., Costa, A.M., Simoes-Correia, J., Oliveira, M.J., Pinheiro, H., Pinho, S.S., Mateus, R., Reis, C.A., Leite, M., Fernandes, M.S., Schmitt, F., Carneiro, F.,

Figueiredo, C., Oliveira, C. & Seruca, R. (2012) Epithelial E- and P-cadherins: Role and clinical significance in cancer. *Biochim Biophys Acta*.

Paredes, J., Lopes, N., Milanezi, F. & Schmitt, F.C. (2007b) P-cadherin and cytokeratin 5: useful adjunct markers to distinguish basal-like ductal carcinomas in situ. *Virchows Arch,* 450**,** 73-80.

Paredes, J., Milanezi, F., Reis-Filho, J.S., Leitao, D., Athanazio, D. & Schmitt, F. (2002a) Aberrant P-cadherin expression: is it associated with estrogen-independent growth in breast cancer? *Pathol Res Pract,* 198**,** 795-801.

Paredes, J., Milanezi, F., Viegas, L., Amendoeira, I. & Schmitt, F. (2002b) P-cadherin expression is associated with high-grade ductal carcinoma in situ of the breast. *Virchows Arch,* 440**,** 16-21.

Paredes, J., Stove, C., Stove, V., Milanezi, F., Van Marck, V., Derycke, L., Mareel, M., Bracke, M. & Schmitt, F. (2004) P-cadherin is up-regulated by the antiestrogen ICI 182,780 and promotes invasion of human breast cancer cells. *Cancer Res,* 64**,** 8309-17.

Park, J., Park, E., Han, S.W., Im, S.A., Kim, T.Y., Kim, W.H., Oh, D.Y. & Bang, Y.J. (2011) Down-regulation of P-cadherin with PF-03732010 inhibits cell migration and tumor growth in gastric cancer. *Investigational new drugs*.

Park, S.Y., Lee, H.E., Li, H., Shipitsin, M., Gelman, R. & Polyak, K. (2010) Heterogeneity for stem cell-related markers according to tumor subtype and histologic stage in breast cancer. *Clin Cancer Res,* 16**,** 876-87.

Park, Y. & Gerson, S.L. (2005) DNA repair defects in stem cell function and aging. *Annu Rev Med,* 56**,** 495-508.

Parsons, J.T. (2003) Focal adhesion kinase: the first ten years. *J Cell Sci,* 116**,** 1409-16.

Patanaphan, V., Salazar, O.M. & Risco, R. (1988) Breast cancer: metastatic patterns and their prognosis. *South Med J,* 81**,** 1109-12.

Pece, S., Tosoni, D., Confalonieri, S., Mazzarol, G., Vecchi, M., Ronzoni, S., Bernard, L., Viale, G., Pelicci, P.G. & Di Fiore, P.P. (2010) Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell,* 140**,** 62-73.

Peralta Soler, A., Knudsen, K.A., Salazar, H., Han, A.C. & Keshgegian, A.A. (1999) P-cadherin expression in breast carcinoma indicates poor survival. *Cancer,* 86**,** 1263-72.

Perou, C.M., Jeffrey, S.S., Van De Rijn, M., Rees, C.A., Eisen, M.B., Ross, D.T., Pergamenschikov, A., Williams, C.F., Zhu, S.X., Lee, J.C., Lashkari, D., Shalon, D., Brown, P.O. & Botstein, D. (1999) Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci U S A,* 96**,** 9212-7.

Perou, C.M., Sorlie, T., Eisen, M.B., Van De Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S.X., Lonning, P.E., Borresen-Dale, A.L., Brown, P.O. & Botstein, D. (2000) Molecular portraits of human breast tumours. *Nature,* 406**,** 747-52.

Phillips, T.M., Mcbride, W.H. & Pajonk, F. (2006) The response of CD24(-/low)/CD44+ breast cancer-initiating cells to radiation. *J Natl Cancer Inst,* 98**,** 1777-85.

Polyak, K. (2007) Breast cancer stem cells: a case of mistaken identity? *Stem Cell Rev,* 3**,** 107-9.

Pontier, S.M. & Muller, W.J. (2009) Integrins in mammary-stem-cell biology and breast-cancer progression--a role in cancer stem cells? *J Cell Sci,* 122**,** 207-14.

Potemski, P., Kusinska, R., Watala, C., Pluciennik, E., Bednarek, A.K. & Kordek, R. (2005) Prognostic relevance of basal cytokeratin expression in operable breast cancer. *Oncology,* 69**,** 478-85.

Potten, C.S., Owen, G. & Booth, D. (2002) Intestinal stem cells protect their genome by selective segregation of template DNA strands. *J Cell Sci,* 115**,** 2381-8.

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

Proia, T.A., Keller, P.J., Gupta, P.B., Klebba, I., Jones, A.D., Sedic, M., Gilmore, H., Tung, N., Naber, S.P., Schnitt, S., Lander, E.S. & Kuperwasser, C. (2011) Genetic predisposition directs breast cancer phenotype by dictating progenitor cell fate. *Cell Stem Cell,* 8**,** 149-63.

Provenzano, P.P., Inman, D.R., Eliceiri, K.W., Knittel, J.G., Yan, L., Rueden, C.T., White, J.G. & Keely, P.J. (2008) Collagen density promotes mammary tumor initiation and progression. *BMC medicine,* 6**,** 11.

Psaila, B. & Lyden, D. (2009) The metastatic niche: adapting the foreign soil. *Nat Rev Cancer,* 9**,** 285-93.

Quintana, E., Shackleton, M., Sabel, M.S., Fullen, D.R., Johnson, T.M. & Morrison, S.J. (2008) Efficient tumour formation by single human melanoma cells. *Nature,* 456**,** 593-8.

Radice, G.L., Ferreira-Cornwell, M.C., Robinson, S.D., Rayburn, H., Chodosh, L.A., Takeichi, M. & Hynes, R.O. (1997) Precocious mammary gland development in P-cadherin-deficient mice. *J Cell Biol,* 139**,** 1025-32.

Rakha, E. & Reis-Filho, J.S. (2009) Basal-like breast carcinoma: from expression profiling to routine practice. *Arch Pathol Lab Med,* 133**,** 860-8.

Rakha, E.A., El-Sayed, M.E., Reis-Filho, J. & Ellis, I.O. (2009) Patho-biological aspects of basal-like breast cancer. *Breast Cancer Res Treat,* 113**,** 411-22.

Rakha, E.A., El-Sheikh, S.E., Kandil, M.A., El-Sayed, M.E., Green, A.R. & Ellis, I.O. (2008) Expression of BRCA1 protein in breast cancer and its prognostic significance. *Hum Pathol,* 39**,** 857-65.

Rakha, E.A., Putti, T.C., Abd El-Rehim, D.M., Paish, C., Green, A.R., Powe, D.G., Lee, A.H., Robertson, J.F. & Ellis, I.O. (2006) Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation. *J Pathol,* 208**,** 495-506.

Rambhatla, L., Ram-Mohan, S., Cheng, J.J. & Sherley, J.L. (2005) Immortal DNA strand cosegregation requires p53/IMPDH-dependent asymmetric self-renewal associated with adult stem cells. *Cancer Res,* 65**,** 3155-61.

Raouf, A., Zhao, Y., To, K., Stingl, J., Delaney, A., Barbara, M., Iscove, N., Jones, S., Mckinney, S., Emerman, J., Aparicio, S., Marra, M. & Eaves, C. (2008) Transcriptome analysis of the normal human mammary cell commitment and differentiation process. *Cell Stem Cell,* 3**,** 109-18.

Rappa, G. & Lorico, A. (2010) Phenotypic characterization of mammosphere-forming cells from the human MA-11 breast carcinoma cell line. *Exp Cell Res,* 316**,** 1576-86.

Reis-Filho, J.S. & Tutt, A.N. (2008) Triple negative tumours: a critical review. *Histopathology,* 52**,** 108-18.

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

Reya, T., Morrison, S.J., Clarke, M.F. & Weissman, I.L. (2001) Stem cells, cancer, and cancer stem cells. *Nature,* 414**,** 105-11.

Rhee, H., Polak, L. & Fuchs, E. (2006) Lhx2 maintains stem cell character in hair follicles. *Science,* 312**,** 1946-9.

Ribeiro, A.S., Albergaria, A., Sousa, B., Correia, A.L., Bracke, M., Seruca, R., Schmitt, F.C. & Paredes, J. (2010) Extracellular cleavage and shedding of P-cadherin: a mechanism underlying the invasive behaviour of breast cancer cells. *Oncogene,* 29**,** 392-402.

Ricardo, S., Vieira, A.F., Gerhard, R., Leitao, D., Pinto, R., Cameselle-Teijeiro, J.F., Milanezi, F., Schmitt, F. & Paredes, J. (2011) Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. *J Clin Pathol*.

Robinson, G.W., Karpf, A.B. & Kratochwil, K. (1999) Regulation of mammary gland development by tissue interaction. *J Mammary Gland Biol Neoplasia,* 4**,** 9-19.

Rodriguez-Pinilla, S.M., Sarrio, D., Honrado, E., Moreno-Bueno, G., Hardisson, D., Calero, F., Benitez, J. & Palacios, J. (2007) Vimentin and laminin expression is associated with basal-like phenotype in both sporadic and BRCA1-associated breast carcinomas. *J Clin Pathol,* 60**,** 1006-12.

Rowe, R.G. & Weiss, S.J. (2008) Breaching the basement membrane: who, when and how? *Trends in cell biology,* 18**,** 560-74.

Sahu, R.P., Batra, S. & Srivastava, S.K. (2009) Activation of ATM/Chk1 by curcumin causes cell cycle arrest and apoptosis in human pancreatic cancer cells. *Br J Cancer,* 100**,** 1425-33.

Sapino, A., Macri, L., Gugliotta, P., Pacchioni, D., Liu, Y.J., Medina, D. & Bussolati, G. (1993) Immunophenotypic properties and estrogen dependency of budding cell structures in the developing mouse mammary gland. *Differentiation,* 55**,** 13-8.

Sarrio, D., Franklin, C.K., Mackay, A., Reis-Filho, J.S. & Isacke, C.M. (2011) Epithelial and Mesenchymal Subpopulations Within Normal Basal Breast Cell Lines Exhibit Distinct Stem Cell/Progenitor Properties. *Stem cells (Dayton, Ohio)*.

Schmitt, F., Ricardo, S., Vieira, A.F., Dionisio, M.R. & Paredes, J. (2012) Cancer stem cell markers in breast neoplasias: their relevance and distribution in distinct molecular subtypes. *Virchows Arch,* 460**,** 545-53.

Shackleton, M., Quintana, E., Fearon, E.R. & Morrison, S.J. (2009) Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell,* 138**,** 822-9.

Shackleton, M., Vaillant, F., Simpson, K.J., Stingl, J., Smyth, G.K., Asselin-Labat, M.L., Wu, L., Lindeman, G.J. & Visvader, J.E. (2006) Generation of a functional mammary gland from a single stem cell. *Nature,* 439**,** 84-8.

Shaw, L.M. (1999) Integrin function in breast carcinoma progression. *J Mammary Gland Biol Neoplasia,* 4**,** 367-76.

Shaw, L.M., Rabinovitz, I., Wang, H.H., Toker, A. & Mercurio, A.M. (1997) Activation of phosphoinositide 3-OH kinase by the alpha6beta4 integrin promotes carcinoma invasion. *Cell,* 91**,** 949-60.

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

Shimomura, Y., Wajid, M., Shapiro, L. & Christiano, A.M. (2008) P-cadherin is a p63 target gene with a crucial role in the developing human limb bud and hair follicle. *Development,* 135**,** 743-53.

Shimoyama, Y., Yoshida, T., Terada, M., Shimosato, Y., Abe, O. & Hirohashi, S. (1989) Molecular cloning of a human Ca2+-dependent cell-cell adhesion molecule

homologous to mouse placental cadherin: its low expression in human placental tissues. *J Cell Biol,* 109**,** 1787-94.

Shipitsin, M., Campbell, L.L., Argani, P., Weremowicz, S., Bloushtain-Qimron, N., Yao, J., Nikolskaya, T., Serebryiskaya, T., Beroukhim, R., Hu, M., Halushka, M.K., Sukumar, S., Parker, L.M., Anderson, K.S., Harris, L.N., Garber, J.E., Richardson, A.L., Schnitt, S.J., Nikolsky, Y., Gelman, R.S. & Polyak, K. (2007) Molecular definition of breast tumor heterogeneity. *Cancer Cell,* 11**,** 259-73.

Sieuwerts, A.M., Kraan, J., Bolt, J., Van Der Spoel, P., Elstrodt, F., Schutte, M., Martens, J.W., Gratama, J.W., Sleijfer, S. & Foekens, J.A. (2009) Anti-epithelial cell adhesion molecule antibodies and the detection of circulating normal-like breast tumor cells. *J Natl Cancer Inst,* 101**,** 61-6.

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

Slack, J.K., Adams, R.B., Rovin, J.D., Bissonette, E.A., Stoker, C.E. & Parsons, J.T. (2001) Alterations in the focal adhesion kinase/Src signal transduction pathway correlate with increased migratory capacity of prostate carcinoma cells. *Oncogene,* 20**,** 1152-63.

Snyder, E.L., Bailey, D., Shipitsin, M., Polyak, K. & Loda, M. (2009) Identification of CD44v6(+)/CD24- breast carcinoma cells in primary human tumors by quantum dot-conjugated antibodies. *Laboratory investigation; a journal of technical methods and pathology,* 89**,** 857-66.

Soler, A.P., Russo, J., Russo, I.H. & Knudsen, K.A. (2002) Soluble fragment of P-cadherin adhesion protein found in human milk. *J Cell Biochem,* 85**,** 180-4.

Sorlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., Van De Rijn, M., Jeffrey, S.S., Thorsen, T., Quist, H., Matese, J.C., Brown, P.O., Botstein, D., Eystein Lonning, P. & Borresen-Dale, A.L. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A,* 98**,** 10869-74.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J.S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., Demeter, J., Perou, C.M., Lonning, P.E., Brown, P.O., Borresen-Dale, A.L. & Botstein, D. (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A,* 100**,** 8418-23.

Soung, Y.H., Gil, H.J., Clifford, J.L. & Chung, J. (2011) Role of alpha6beta4 integrin in cell motility, invasion and metastasis of mammary tumors. *Current protein & peptide science,* 12**,** 23-9.

Sousa, B., Paredes, J., Milanezi, F., Lopes, N., Martins, D., Dufloth, R., Vieira, D., Albergaria, A., Veronese, L., Carneiro, V., Carvalho, S., Costa, J.L., Zeferino, L. & Schmitt, F. (2010) P-cadherin, vimentin and CK14 for identification of basal-like

phenotype in breast carcinomas: an immunohistochemical study. *Histol Histopathol,* 25**,** 963-74.

Sprecher, E., Bergman, R., Richard, G., Lurie, R., Shalev, S., Petronius, D., Shalata, A., Anbinder, Y., Leibu, R., Perlman, I., Cohen, N. & Szargel, R. (2001) Hypotrichosis with juvenile macular dystrophy is caused by a mutation in CDH3, encoding P-cadherin. *Nat Genet,* 29**,** 134-6.

Stefansson, I.M., Salvesen, H.B. & Akslen, L.A. (2004) Prognostic impact of alterations in P-cadherin expression and related cell adhesion markers in endometrial cancer. *J Clin Oncol,* 22**,** 1242-52.

Stemmler, M.P. (2008) Cadherins in development and cancer. *Molecular bioSystems,* 4**,** 835-50.

Stingl, J. (2009) Detection and analysis of mammary gland stem cells. *J Pathol,* 217**,** 229-41.

Stingl, J. & Caldas, C. (2007) Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. *Nat Rev Cancer,* 7**,** 791-9.

Stingl, J., Eaves, C.J., Zandieh, I. & Emerman, J.T. (2001) Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. *Breast Cancer Res Treat,* 67**,** 93-109.

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H.I. & Eaves, C.J. (2006) Purification and unique properties of mammary epithelial stem cells. *Nature,* 439**,** 993-7.

Surowiak, P., Materna, V., Gyorffy, B., Matkowski, R., Wojnar, A., Maciejczyk, A., Paluchowski, P., Dziegiel, P., Pudelko, M., Kornafel, J., Dietel, M., Kristiansen, G., Zabel, M. & Lage, H. (2006) Multivariate analysis of oestrogen receptor alpha, pS2, metallothionein and CD24 expression in invasive breast cancers. *Br J Cancer,* 95**,** 339-46.

Taddei, I., Deugnier, M.A., Faraldo, M.M., Petit, V., Bouvard, D., Medina, D., Fassler, R., Thiery, J.P. & Glukhova, M.A. (2008) Beta1 integrin deletion from the basal compartment of the mammary epithelium affects stem cells. *Nat Cell Biol,* 10**,** 716-22.

Taddei, I., Faraldo, M.M., Teuliere, J., Deugnier, M.A., Thiery, J.P. & Glukhova, M.A. (2003) Integrins in mammary gland development and differentiation of mammary epithelium. *J Mammary Gland Biol Neoplasia,* 8**,** 383-94.

Tagliabue, E., Ghirelli, C., Squicciarini, P., Aiello, P., Colnaghi, M.I. & Menard, S. (1998) Prognostic value of alpha 6 beta 4 integrin expression in breast carcinomas is affected by laminin production from tumor cells. *Clin Cancer Res,* 4**,** 407-10.

Taniuchi, K., Nakagawa, H., Hosokawa, M., Nakamura, T., Eguchi, H., Ohigashi, H., Ishikawa, O., Katagiri, T. & Nakamura, Y. (2005) Overexpressed P-cadherin/CDH3

promotes motility of pancreatic cancer cells by interacting with p120ctn and activating rho-family GTPases. *Cancer Res,* 65**,** 3092-9.

Tavassoli, F.A. & Devilee, P. (2003) WHO classification of tumors. Patholy and genetics of tumors of the breast and female genital organs. Lyon: IARC press.

Thiery, J.P. (2002) Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer,* 2**,** 442-54.

Tsuruta, D., Kobayashi, H., Imanishi, H., Sugawara, K., Ishii, M. & Jones, J.C. (2008) Laminin-332-integrin interaction: a target for cancer therapy? *Current medicinal chemistry,* 15**,** 1968-75.

Turashvili, G., Mckinney, S.E., Goktepe, O., Leung, S.C., Huntsman, D.G., Gelmon, K.A., Los, G., Rejto, P.A. & Aparicio, S.A. P-cadherin expression as a prognostic biomarker in a 3992 case tissue microarray series of breast cancer. *Mod Pathol,* 24**,** 64-81.

Turashvili, G., Mckinney, S.E., Goktepe, O., Leung, S.C., Huntsman, D.G., Gelmon, K.A., Los, G., Rejto, P.A. & Aparicio, S.A. (2011) P-cadherin expression as a prognostic biomarker in a 3992 case tissue microarray series of breast cancer. *Mod Pathol,* 24**,** 64-81.

Turner, N.C. & Reis-Filho, J.S. (2006) Basal-like breast cancer and the BRCA1 phenotype. *Oncogene,* 25**,** 5846-53.

Turner, N.C., Reis-Filho, J.S., Russell, A.M., Springall, R.J., Ryder, K., Steele, D., Savage, K., Gillett, C.E., Schmitt, F.C., Ashworth, A. & Tutt, A.N. (2007) BRCA1 dysfunction in sporadic basal-like breast cancer. *Oncogene,* 26**,** 2126-32.

Van Marck, V., Stove, C., Jacobs, K., Van Den Eynden, G. & Bracke, M. (2011) P-cadherin in adhesion and invasion: opposite roles in colon and bladder carcinoma. *Int J Cancer,* 128**,** 1031-44.

Van Marck, V., Stove, C., Van Den Bossche, K., Stove, V., Paredes, J., Vander Haeghen, Y. & Bracke, M. (2005) P-cadherin promotes cell-cell adhesion and counteracts invasion in human melanoma. *Cancer Res,* 65**,** 8774-83.

Vassilopoulos, A., Wang, R.H., Petrovas, C., Ambrozak, D., Koup, R. & Deng, C.X. (2008) Identification and characterization of cancer initiating cells from BRCA1 related mammary tumors using markers for normal mammary stem cells. *International journal of biological sciences,* 4**,** 133-42.

Venezia, T.A., Merchant, A.A., Ramos, C.A., Whitehouse, N.L., Young, A.S., Shaw, C.A. & Goodell, M.A. (2004) Molecular signatures of proliferation and quiescence in hematopoietic stem cells. *PLoS Biol,* 2**,** e301.

Vieira, A.F., Ricardo, S., Ablett, M.P., Dionisio, M.R., Mendes, N., Albergaria, A., Farnie, G., Gerhard, R., Cameselle-Teijeiro, J.F., Seruca, R., Schmitt, F., Clarke, R.B. & Paredes, J. (2012) P-Cadherin is Co-Expressed with Cd44 and Cd49f and

Mediates Stem Cell Properties in Basal-Like Breast Cancer. *Stem cells (Dayton, Ohio)*.

Villadsen, R., Fridriksdottir, A.J., Ronnov-Jessen, L., Gudjonsson, T., Rank, F., Labarge, M.A., Bissell, M.J. & Petersen, O.W. (2007) Evidence for a stem cell hierarchy in the adult human breast. *J Cell Biol,* 177**,** 87-101.

Visvader, J.E. (2009) Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. *Genes Dev,* 23**,** 2563-77.

Visvader, J.E. (2011) Cells of origin in cancer. *Nature,* 469**,** 314-22.

Visvader, J.E. & Lindeman, G.J. (2006) Mammary stem cells and mammopoiesis. *Cancer Res,* 66**,** 9798-801.

Wang, D., Grammer, J.R., Cobbs, C.S., Stewart, J.E., Jr., Liu, Z., Rhoden, R., Hecker, T.P., Ding, Q. & Gladson, C.L. (2000) p125 focal adhesion kinase promotes malignant astrocytoma cell proliferation in vivo. *J Cell Sci,* 113 Pt 23**,** 4221-30.

Wang, H., Radjendirane, V., Wary, K.K. & Chakrabarty, S. (2004) Transforming growth factor beta regulates cell-cell adhesion through extracellular matrix remodeling and activation of focal adhesion kinase in human colon carcinoma Moser cells. *Oncogene,* 23**,** 5558-61.

Wang, J., Wakeman, T.P., Lathia, J.D., Hjelmeland, A.B., Wang, X.F., White, R.R., Rich, J.N. & Sullenger, B.A. (2010) Notch promotes radioresistance of glioma stem cells. *Stem cells (Dayton, Ohio),* 28**,** 17-28.

Wang, S., Yang, D. & Lippman, M.E. (2003) Targeting Bcl-2 and Bcl-XL with nonpeptidic small-molecule antagonists. *Semin Oncol,* 30**,** 133-42.

Wang, Z., Wang, M., Kar, S. & Carr, B.I. (2009) Involvement of ATM-mediated Chk1/2 and JNK kinase signaling activation in HKH40A-induced cell growth inhibition. *J Cell Physiol,* 221**,** 213-20.

Weber, G.F., Bjerke, M.A. & Desimone, D.W. (2011) Integrins and cadherins join forces to form adhesive networks. *J Cell Sci,* 124**,** 1183-93.

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

Wellings, S.R., Jensen, H.M. & Marcum, R.G. (1975) An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. *J Natl Cancer Inst,* 55**,** 231-73.

Wheelock, M.J., Soler, A.P. & Knudsen, K.A. (2001) Cadherin junctions in mammary tumors. *J Mammary Gland Biol Neoplasia,* 6**,** 275-85.

White, D.E., Kurpios, N.A., Zuo, D., Hassell, J.A., Blaess, S., Mueller, U. & Muller, W.J. (2004) Targeted disruption of beta1-integrin in a transgenic mouse model of human breast cancer reveals an essential role in mammary tumor induction. *Cancer Cell,* 6**,** 159-70.

Wicha, M.S., Liu, S. & Dontu, G. (2006) Cancer stem cells: an old idea--a paradigm shift. *Cancer Res,* 66**,** 1883-90; discussion 1895-6.

Williams, R.T., Den Besten, W. & Sherr, C.J. (2007) Cytokine-dependent imatinib resistance in mouse BCR-ABL+, Arf-null lymphoblastic leukemia. *Genes Dev,* 21**,** 2283-7.

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

Xu, L.H., Yang, X., Bradham, C.A., Brenner, D.A., Baldwin, A.S., Jr., Craven, R.J. & Cance, W.G. (2000) The focal adhesion kinase suppresses transformation-associated, anchorage-independent apoptosis in human breast cancer cells. Involvement of death receptor-related signaling pathways. *J Biol Chem,* 275**,** 30597-604.

Yehiely, F., Moyano, J.V., Evans, J.R., Nielsen, T.O. & Cryns, V.L. (2006) Deconstructing the molecular portrait of basal-like breast cancer. *Trends in molecular medicine,* 12**,** 537-44.

Yilmaz, M. & Christofori, G. Mechanisms of motility in metastasizing cells. *Mol Cancer Res,* 8**,** 629-42.

Yoon, C.S., Kim, K.D., Park, S.N. & Cheong, S.W. (2001) alpha(6) Integrin is the main receptor of human papillomavirus type 16 VLP. *Biochem Biophys Res Commun,* 283**,** 668-73.

Yurchenco, P.D. & Patton, B.L. (2009) Developmental and pathogenic mechanisms of basement membrane assembly. *Current pharmaceutical design,* 15**,** 1277-94.

Zhang, C.C., Yan, Z., Zhang, Q., Kuszpit, K., Zasadny, K., Qiu, M., Painter, C.L., Wong, A., Kraynov, E., Arango, M.E., Mehta, P.P., Popoff, I., Casperson, G.F., Los, G., Bender, S., Anderes, K., Christensen, J.G. & Vanarsdale, T. (2010) PF-03732010: a fully human monoclonal antibody against P-cadherin with antitumor and antimetastatic activity. *Clin Cancer Res,* 16**,** 5177-88.

ESTABLISHMENT OF A STABLE *CDH3*/P-CADHERIN SHRNA DOX-INDUCIBLE MODEL IN A HUMAN BASAL-LIKE CANCER CELL LINE

APPENDIX 1

1. Introduction

Although the transient inhibition of P-cadherin is a quick way of effectively down-regulating this molecule in in distinct cell models, allowing for the study of short term or immediate effects induced by P-cadherin silencing, this method of RNA interference (RNAi) does not allow for the study of cellular effects induced by prolonged P-cadherin inhibition in breast cell lines. Thus, studies regarding the stable inhibition of P-cadherin would be useful for both in vitro and in vivo studies. Furthermore, an inducible gene silencing system would allow to switch-on / switch-off gene expression which would give a broader perspective regarding the mechanisms of the tumourigenic role of P-cadherin, either in the initiation phase or in the progression of tumour growth.

RNAi pathways are activated by various forms of double stranded RNAs (dsRNAs) that contain sequences which are homologous to the mRNA transcript of a target gene (Hammond *et al.*, 2001, Hutvagner & Zamore, 2002, Sharp, 2001). It is known that short hairpin RNAs (shRNAs) adopt a stable stem-loop structure in solution, can be easily expressed from a cloned oligonucleotide template and are a convenient and reproducible means of activating RNAi in mammalian cell lines (**Figure 1**) (Brummelkamp *et al.*, 2002, Paddison *et al.*, 2002, Paul *et al.*, 2002, Yu *et al.*, 2002).

To produce a stable human cancer cell line with inducible knock-down of *CDH3*/P-cadherin, a shRNA vector was constructed based in the small interference RNA (siRNA) known-down sequence for this gene. The vector was derived from Knock-out™ Single Vector Inducible RNAi System (Clontech, Mountain View, CA). In comparison with the *CDH3* siRNA oligos, this system constitutes a more versatile and more prolonged knock down of P-cadherin, with an inducible expression readily performed with doxycycline (DOX). The delivery method of this system is non-viral and plasmid-based.

This system allows to quickly introduce and control the expression of functional shRNA molecules for the purpose of activating gene-specific RNAi. The tight on/off regulation and the coordinate inactivation of its target gene are provided by a tetracycline-inducible system that responds to the presence of tetracycline, or its more stable derivative, DOX (Gossen *et al.*, 1993, Gossen & Bujard, 1992, Gossen & Bujard, 1995). The system features two essential components combined on its pSingle-tTS-shRNA vector: 1) CMV promoter/enhancer-controlled expression of the tetracycline-controlled regulatory protein tTS (tetracycline-controlled transcriptional silencer) and 2) a tetracycline-inducible shRNA expression cassette. The tTS protein is a powerful transcriptional repressor created by fusing the Tet repressor protein (TetR) with a KRAB transcriptional silencing domain (Freundlieb *et al.*, 1999, Witzgall *et al.*, 1994,

Wiznerowicz & Trono, 2003). A tetracycline-inducible hybrid promotor was created by linking a modified Tet-responsive element from the *tet* operon (from TRE-Tight promoter) to a minimal U6 small nuclear RNA promoter (Kunkel & Pederson, 1989).

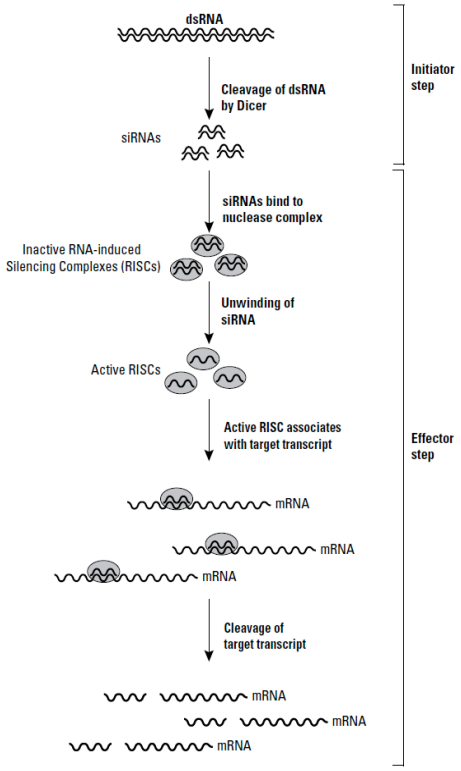


Figure 1 – Mechanism of RNA interference. RNAi is activated by introducing a double stranded RNA, whose sequence is homologous to the target gene transcript. The exogenous dsRNA is digested into 21-23 nucleotide small interfering RNAs (siRNAs), which bind a nuclease complex to form an RNA-induced silencing complex (RISC). The RISC then targets endogenous gene transcripts by base-pairing and cleaving the mRNA. Adapted from Knockout^{TM} Single Vector Inducible RNAi system User Manual.

In the absence of the inducer, DOX, the tTS protein binds tightly to the *tet*O sequences within the tetracycline responsive element (TRE) and actively silences transcription of the shRNA from the downstream minimal U6 promoter (**Figure 2**). In this basal state, i.e., in the absence of induction, background expression of the shRNA is extremely low and prevents unwanted suppression of the target gene. When DOX is added to the culture medium, tTS dissociates from the TRE relieving transcriptional repression and permits the shRNA to be transcribed from the U6 promoter. Once derepressed, the human U6 Polymerase III promoter provides high level expression in many cell types (Kunkel & Pederson, 1989) and the accumulating shRNA transcripts initiate RNAi-mediated suppression of the target gene.

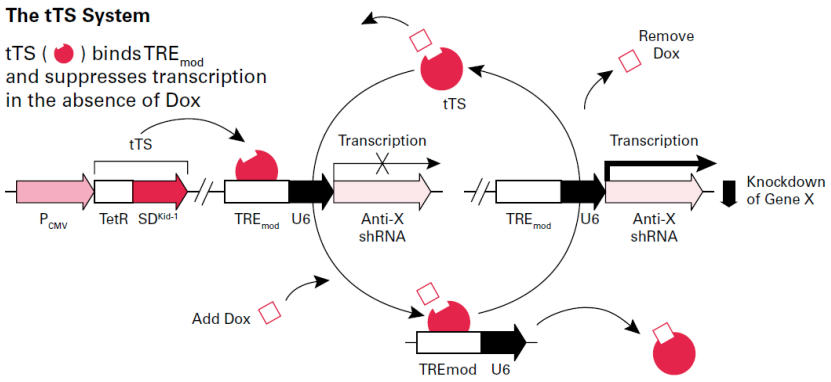


Figure 2 – The Knockout Single Vector Inducible RNAi system (Clontech, Mountain View, CA) uses a modified form of the tightly regulated, tetracycline-controlled gene expression system. In the absence of DOX, tTS binds to the tetO sequences within the TRE/U6 promoter and actively silences transcription of the shRNA. When DOX is added to the culture medium, tTS dissociates from the TRE, relieving transcriptional suppression and allowing high level transcription of the shRNA from the hybrid TRE/U6 promoter by Polymerase III in a highly dose dependent manner. Adapted from Knockout™ Single Vector Inducible RNAi system User Manual.

The pSingle-tTS-shRNA Vector Map is pictured in **Figure 3**. The shRNA sequence is under the control of the hybrid promoter TRE/thight-U6 that binds tTS or Polymerase III. The pSingle-tTS-shRNA vector also contains a bacterial origin of replication and the Amp^r gene for propagation and selection in *E. coli*. it also contains the neomycin^r gene for selection of stable transformants in mammalian cells.

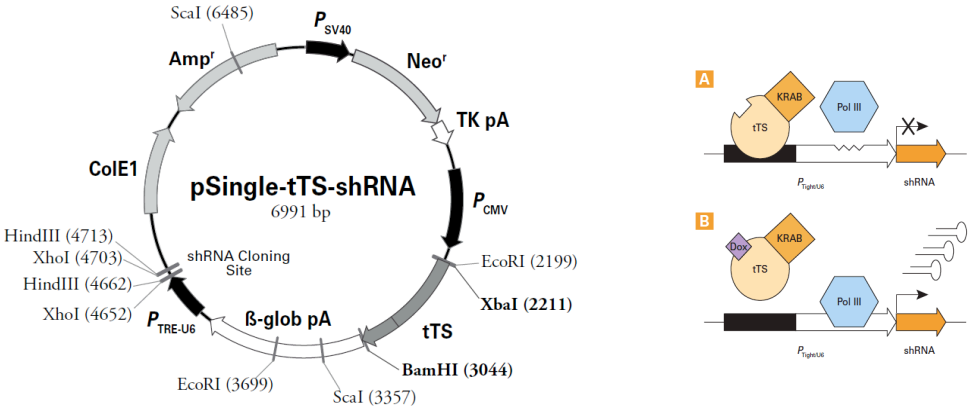


Figure 3 – pSingle-tTS-shRNA Vector Map. The pSingle-tTS-shRNA Vector expresses the tetracycline-controlled transcriptional suppressor (tTS) under the control of the CMV promoter wich, in turn, controls expression of an shRNA sequence inserted in the shRNA cloning site. The tTS protein is a fusion of the Tet repressor protein (TetR) and the KRAB- AB silencing domain of the Kid-1 protein (SD^{Kid-1}), a powerful transcriptional suppressor. The shRNA sequence is under the control of the hybrid promoter TRE/thight-U6 that binds tTS or Polymerase III. The pSingle-tTS-shRNA vector also contains a bacterial origin of replication and the Amp^r gene for propagation and selection in *E. coli*; and the neomycin^r gene for selection of stable transformants in mammalian cells. Adapted from Knockout™ Single Vector Inducible RNAi system User Manual.

2. Materials and methods

The implementation of the Knockout pSingle-tTS-shRNA Vector was performed according to the steps listed below, essentially following the protocol provided by the manufacturer (Clontech, Mountain View, CA).

Design and synthesis of the shRNA sequence

The shRNA target sequence selection to clone in the pSingle Vector was based in the already known sequence used in our transient siRNA studies. This sequence recognizes a region near the 3' of the *CDH3*/P-cadherin transcript message (exon 16) with high silencing efficiency. The shRNA oligonucleotide sequence included the 19-base target siRNA sequence, as shown in **Figure 4.** When selecting this target siRNA sequence we checked that it did not show significant homology to other genes and it did not contain a consecutive run of 3 or more thymidine (T) residues. The CG content was between 40% and 60%. This oligonucleotide sequence had minimal secondary structures, which could interfere with proper annealing. Since it is known that sequences that have at least 3 adenine or thymidine residues in positions 15-19 of the sequence appear to have increased knock-down activity, it was important to notice that our siRNA target sequence matched this requirement.

The complementary oligonucleotides were synthesized using PAGE purification by Sigma-Aldrich (St. Louis, MO). The two sequences of oligonucleotides included a 5´- XhoI restriction site overhang and 5´- HindIII overhang, that enabled directional cloning of the annealed oligonucleotides into the XhoI/HindIII-digested pSingle Vector. Furthermore, a 9-nucleotide hairpin loop sequence and a RNA Pol III terminator sequence, consisting of 6 nucleotide poly(T) tract, were also designed and included in the shRNA sequence. A unique restriction site positioned immediately downstream of the terminator sequence was also included for restriction digest analysis in order to confirm the presence of the cloned insert (**Figure 4**).

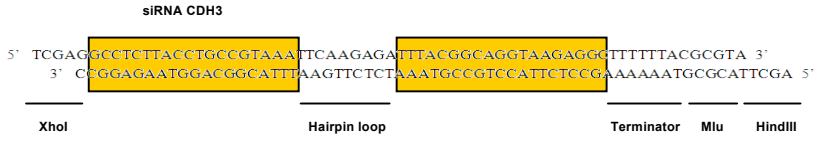


Figure 4 – shRNA sequence for *CDH3*/P-cadherin gene. Both complementary strands (upper strand and lower strands) were synthesized by Sigma-Aldrich. The hairpin loop sequence is one of the many functional loop sequences used to generate shRNAs. Termination is signalled using a poly(T) tract. The shRNA sequence includes a unique restriction site (MluI) that allows the confirmation of the cloned insert after the ligation and transformation reactions. XhoI (upper) and HindIII (lower) 5' overhangs are necessary for directional cloning into the pSingle-tTS-shRNA vetor.

The strategy was to establish the shRNA p-Single Vector in the breast cancer cell line MDA-MB-468. This cell line has a triple-negative phenotype (negative for HER-2, ER and PgR) and expresses high levels of P-cadherin, constituting an excellent model of human basal-like breast cancer *in vitro*. Furthermore, our previous experience in cloning revealed that this cell line is easily transfected using a non-viral delivery method.

Annealing the single stranded oligonucleotides to form the double stranded Oligonucleotides

Each oligonucleotide (top and lower strands) was resuspended in tris-EDTA buffer to a final concentration of 100 μM and then mixed to a 1:1 ratio. To form the double strand oligonucleotides, the following sequence of temperatures was applied to the mixture: 95ºC for 30 sec, 72ºC for 2 min, 37ºC for 2 min, 25ºC for 2 min.

Cloning shRNA oligonucleotides into pSingle-tTS-shRNA vector

1 ug of pSingle-tTS-shRNA vector (cat. Number 630933, Clonetech) was digested with XhoI and HindIII restriction enzymes (New England Biolabs, Hitchin, UK), with 1 U in endonuclease buffer NEB2 (New England Biolabs) containing 0.1 mg/ml BSA (New England Biolabs), during 1h ate 37ºC. After running the product of the endonuclease reaction in a 1% agarose gel, one band was observed at ~6900 bp. This band was cut out from the gel and the vector was extracted using QIAquick gel extraction kit (Qiagen, West Sussex, UK).

Ligation reaction

In order to ligate the annealed double stranded oligonucleotides into pSingle-tTS-shRNA, 50 ng of the digested pSingle-tTS-shRNA vector DNA was mixed in a ligation reaction containing 0.5 mM annealed oligonucleotides, T4 DNA ligase (Roche, Burgess Hill, UK) and DNA ligase buffer (Roche). The ligase reaction mixture was incubated at room temperature for 3 hours. A control ligation was assembled using 1 μl nuclease-free water instead of the annealed oligos.

Transformation of competent cells, identification of recombinant clones

Fusion–Blue competent cells (Clontech, cat. No. 636700) are an *E.coli* K12 strain that provides high transformation efficiency. The strain carries recA and endA mutations that turn them into a good host for obtaining high yields of plasmid DNA. These competent *E.coli* cells were thawed on ice and transformed with 2 μL of the ligation reactions (previously diluted 1:5 to a final concentration of 10ng/μl), according to the protocol

supplied with the cells. Briefly, after adding the ligation reactions (one containing the *CDH3* pSingle shRNA vector, and the other containing the empty vector) or water (control for transfection) to chilled test tubes containing 100 uL of *E.coli* cells, these were incubated on ice for 30 min and then heat shocked at 42ºC for 45 seconds. Next, 900 μL of S.O.C. medium (Invitrogen, cat. No. 15544-034) and the transfection mixture was shaken at 37ºC for 1 hour. Different volumes from each transformation (20-150 μL) were plated on LB agar + ampicillin plates (50 μg/ml) which were incubated overnight at 37 ºC.

The control ligation reaction (water instead of plasmid) did not originate colonies; the ligation reaction performed with the pSingle-tTS-shRNA *CDH3* vector and the one performed with the empty pSingle vector originated several colonies. Thus, well isolated colonies obtained from the ligation reaction were inoculated each into a small-scale liquid culture containing LB and ampicilin. Colonies grew overnight at 37ºC with shaking. Plasmid DNA minipreps were prepared from 4 ml of culture medium of each colony (Qiagen, West Sussex, UK) and the recombinant plasmids were identified by restriction analysis using the unique restriction site MluI within the shRNA oligonucleotide sequence. Briefly, 1 μL of the mini-prep DNA plasmid was digested with 1U of MluI and 1U of EcoRI restriction enzymes (New England Biolabs, Hitchin, UK) in endonuclease buffer NEB3 (New England Biolabs) 1,5h at 37ºC. The product of the endonuclease reaction was run in a 2% agarose gel, as shown in **Figure 5**.

According to the pattern of bands obtained after endonuclease digestion, we verified that all the selected clones (A-D) had the correct oligonucleotide insert. We selected one of the clones to continue our protocol and the insert was further verified by sequencing. Using a primer for the upper strand (TGTCGAGGTAGGCGTGTACGGT) and a primer for the lower strand (CCGCGCGTTGGCCGATTCAT).

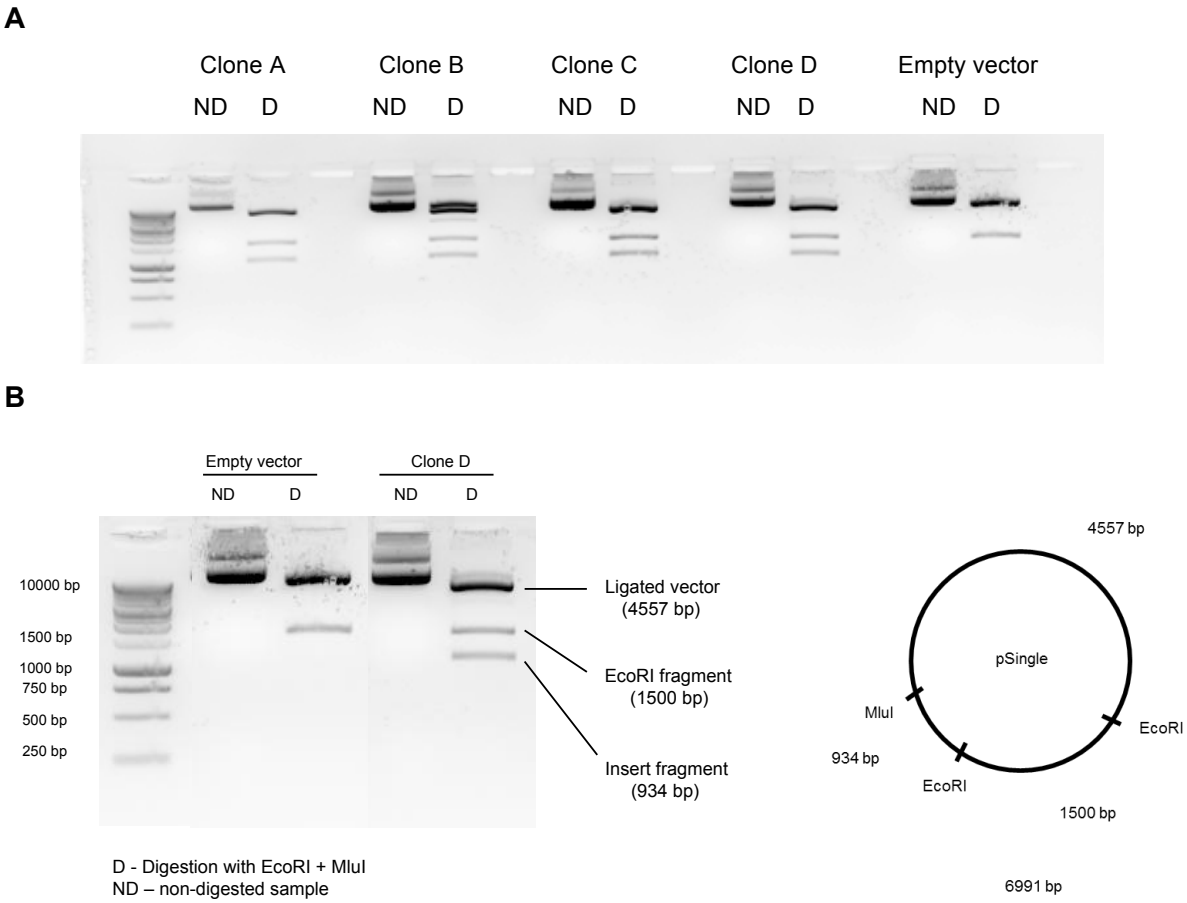


Figure 5 – (A) Verification of the presence of the insert after the ligation reaction. Four clones were selected after the ligation reaction and the transfection of competent cells (clones A-D). Control sample represent the ligation reaction product obtained without the oligonucleotide shRNA sequence (water), corresponding to the empty vector. N – no digestion; D – digestion with EcoRi + MluI. **(B).** The pattern of bands obtained after digestion in clone D is in agreement with the expected resulting bands.

Preparation of DNA for transfection (maxi-prep)

Once a positive clone was identified, a large-scale DNA prep of the recombinant pSingle-tTS-shRNA vector was performed. We performed a high purification Maxi-prep (Qiagen, West Sussex, UK) according to the manufacturer instructions. Briefly, 200 ml of LB broth was inoculated with 1 ml of a mini-prep cell culture of the selected clone. After overnight incubation at 37ºC with shaking, the culture was centrifuged and resuspended in lysis buffer. After adding the neutralizing buffer, the mixture was centrifuged and the supernatant applied to a QIAprep column. The column was washed and the DNA eluted. After DNA precipitation, the pellet was resuspendend in RNase-free water.

Transfection and selection of pSingle-tTS-shRNA *CDH3* stable cell lines

In order to transfect the breast cancer cell line MDA-MB-468, optimization experiments were performed to determine the optimal antibiotic (G418) concentration for the selection of stable clones. For this cell line, the optimal concentration of antibiotic was previously determined to be 0.5 mg/ml. This is the lowest concentration of G418 (Invitrogen, cat. No. 10131-019) that begins to result in massive cell death in ~5 days and kills all the cells in two weeks. This determination was performed in the lab using the Sulforhodamine B colorimetric assay for cytotoxicity screening (Vichai & Kirtikara, 2006Vichai & Kirtikara, 2006) that allowed the representation of a dose-response curve. We did not titrate the optimal concentration of DOX to use, as this was already established in our lab. This concentration was determined to be 1 μg/ml for the Doxycycline hyclate (Sigma-Aldrich, cat. No. D9891).

The transfection method used was electroporation. An $8x10^6$ cell suspension in PBS was mixed with 20 μg of plasmid in an electroporation cuvete kept on ice. A control electroporation tube was included (cells only). Using the Bio Rad Gene Pulser electroporator (Bio Rad, Hercules, CA), the optimal electroporation method was 800 V and 25 uF, with a time constant lower than 20 ms. After electroporation, cells were incubated 10 min on ice. Cells were then washed with 10 ml of complete medium (DMEM, 10% FBS), centrifuged 1200 rpm 5 min, and resuspended in 4 ml of complete medium. In two 10 cm culture dish, 2 ml of electroporated cells were plated. Two days after electroporation, the selection medium was added to the dishes (DMEM with 10% FBS and 0.5 mg/ml G418). The selection medium was replaced with fresh complete medium plus G418 every four days or more often if dead cells accumulated or if medium became depleted. After 3-5 weeks, isolated G418-resistant colonies begun to appear. Several colonies were observed (~5 colonies/ dish). The large healthy colonies were trypsinized using a cloning cylinder and transferred to individual wells.

Although it is recommended that as many clones as possible be isolated following transfection, at the present time we could only isolate two clones (**Figure 6**).

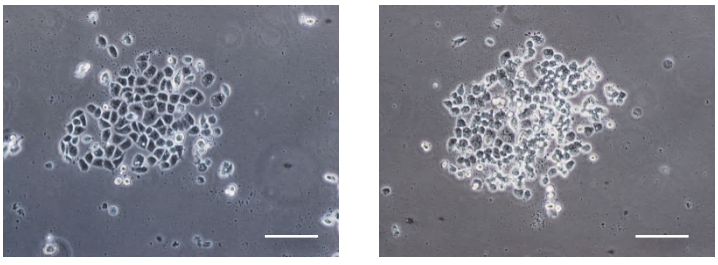


Figure 6 – Example of healthy colonies formed 1-2 months after transfection of pSingle-tTS-shRNA CDH3 in the MDA-MB-468 breast cancer cell line. Scale bar = 200 μm.

3. Cell model characterization

The clones that exhibit the highest overall shRNA expression (highest level of gene suppression) in the presence of DOX, and with little or no background suppression in the absence of DOX should be elected for propagation and further testing. However, up till now, we have obtained only one clone that has been amplified and screened for shRNA induction. This clone was routinely cultured in DMEM containing tetracyclin-free FBS (Clontech) plus 0.5 mg/ml G418 (Invitrogen). The commonly used FBS (EU approved from Lonza, Basel, Switzerland, or from Gibco, Paisley, UK) can contain residual tetracycline activity that could cause a small induction of the shRNA expression (leakage); so, in our initial studies we opted for a tetracycline free FBS.

Cells were treated with DOX 1μg/ml (Sigma-Aldrich, St. Louis, MO), with medium renewal every three days, and protein lysates were obtained in different time points. Control cells were treated with water, the vehicle. P-cadherin protein expression was detected by western-blot using anti-P-cadherin antibody (BD Biosciences, Erembodegem, Belgium). Protein band quantification was performed using Image J (http://rsbweb.nih.gov/ij/). Results are presented in **Figure 7.** Western-blot analysis of this clone revealed that the efficiency of the knock-down after treatment with DOX was very low, with less than 20% inhibition. Further confirmation will be necessary using other gene screening methods, namely RT-PCR.

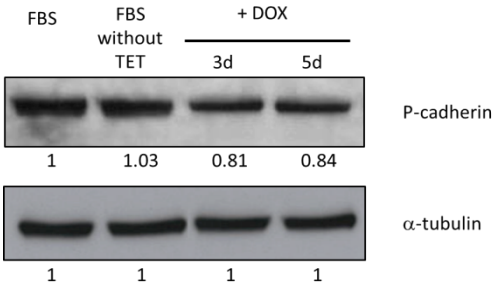


Figure 7 – A stable clone derived from the cell line MDA-MB-468 containing the plasmid pSingle-tTS-shRNA *CDH3* was screened for shRNA induction after treatment with 1 ug/ml doxycycline (DOX) for three (3d) and five days (5d). Western-blot was used to detect P-cadherin and the bands were quantified with Image J software.

A functional assay was also used to check the efficiency of this knock-down method in this clone. The mammosphere assay was performed with a continuous administration of DOX every three days. Administration of DOX started three days before the mammosphere assay was performed and lasted for the whole 5 days of mammospheres

formation, with addition of DOX to the mammosphere medium. The mammosphere formation efficiency and the mammosphere phenotype is presented in **Figure 8**.

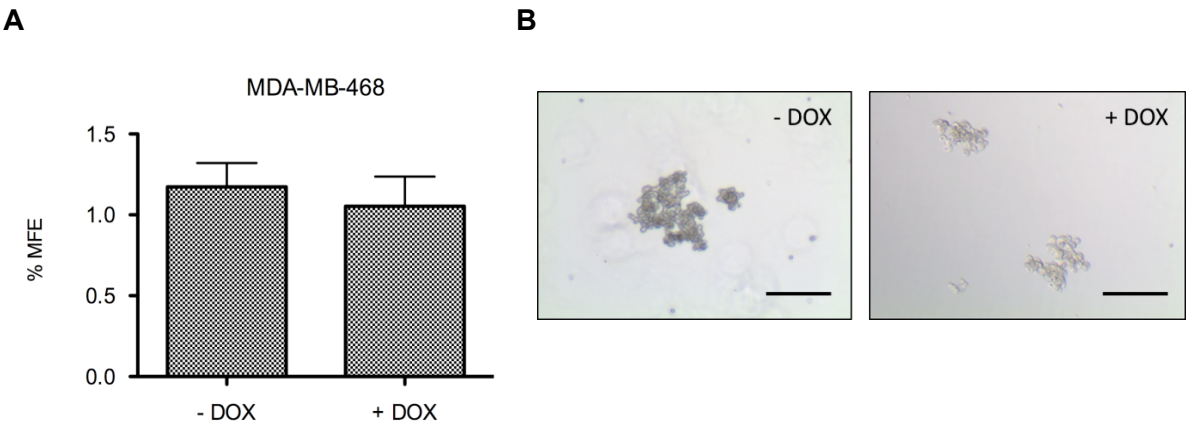


Figure 8 – (A) MFE is not affected by DOX treatment in this clone of MDA-MB-468 containing the plasmid pSingle-tTS-shRNA *CDH3*. **(B)** Mamospheres originated from DOX treated cells appeared to contain more loose spheres than the non-induced control cells. Scale bar = 250 μm.

The MFE of MDA-MB-468 was not significantly affected by the addition of DOX during the experiment, even when DOX concentration was increased from 1 μg/ml to 10 μg/ml (with no observable toxic effects). However, the grape-like shape of the mammospheres was affected by the induction procedure. DOX treatment lead to the formation of mammospheres with a less compact core and with less aggregation than the DOX untreated condition. It is difficult to say, at this point, whether this effect was caused by P-cadherin inhibition in the DOX treated cells and further experiments are needed.

It is important to say that clonal variation in expression of both the tTS silencer and the shRNA construct is affected by the genomic integration site, therefore, the screening of more clones is needed. Furthermore, we could search for more shRNA sequences to find out which has the optimal gene silencing.

In conclusion, although the protocol for the tetracycline-controlled system was established successfully in this cell line, more clones will need to be screened in order to find an optimal expressor.

4. References

Brummelkamp, T.R., Bernards, R. & Agami, R. (2002) A system for stable expression of short interfering RNAs in mammalian cells. *Science,* 296**,** 550-3.

Freundlieb, S., Schirra-Muller, C. & Bujard, H. (1999) A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. *The journal of gene medicine,* 1**,** 4-12.

Fujita, M., Furukawa, F., Fujii, K., Horiguchi, Y., Takeichi, M. & Imamura, S. (1992) Expression of cadherin cell adhesion molecules during human skin development: morphogenesis of epidermis, hair follicles and eccrine sweat ducts. *Archives of dermatological research,* 284**,** 159-66.

Gossen, M., Bonin, A.L. & Bujard, H. (1993) Control of gene activity in higher eukaryotic cells by prokaryotic regulatory elements. *Trends in biochemical sciences,* 18**,** 471-5.

Gossen, M. & Bujard, H. (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A,* 89**,** 5547-51.

Gossen, M. & Bujard, H. (1995) Efficacy of tetracycline-controlled gene expression is influenced by cell type: commentary. *BioTechniques,* 19**,** 213-6; discussion 216-7.

Hammond, S.M., Caudy, A.A. & Hannon, G.J. (2001) Post-transcriptional gene silencing by double-stranded RNA. *Nat Rev Genet,* 2**,** 110-9.

Hutvagner, G. & Zamore, P.D. (2002) RNAi: nature abhors a double-strand. *Curr Opin Genet Dev,* 12**,** 225-32.

Kunkel, G.R. & Pederson, T. (1989) Transcription of a human U6 small nuclear RNA gene in vivo withstands deletion of intragenic sequences but not of an upstream TATATA box. *Nucleic Acids Res,* 17**,** 7371-9.

Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J. & Conklin, D.S. (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev,* 16**,** 948-58.

Paul, C.P., Good, P.D., Winer, I. & Engelke, D.R. (2002) Effective expression of small interfering RNA in human cells. *Nat Biotechnol,* 20**,** 505-8.

Sharp, P.A. (2001) RNA interference--2001. *Genes Dev,* 15**,** 485-90.

Vichai, V. & Kirtikara, K. (2006) Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nature protocols,* 1**,** 1112-6.

Witzgall, R., O'leary, E., Leaf, A., Onaldi, D. & Bonventre, J.V. (1994) The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc Natl Acad Sci U S A,* 91**,** 4514-8.

Wiznerowicz, M. & Trono, D. (2003) Conditional suppression of cellular genes: lentivirus vector-mediated drug-inducible RNA interference. *Journal of virology,* 77**,** 8957-61.

Yu, J.Y., Deruiter, S.L. & Turner, D.L. (2002) RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc Natl Acad Sci U S A,* 99**,** 6047-52.

PUBLICATIONS

APPENDIX 2

EDITOR'S CHOICE

Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype

Sara Ricardo,[1,2,3] André Filipe Vieira,[1,2] Renê Gerhard,[1] Dina Leitão,[1,4] Regina Pinto,[1] Jorge F Cameselle-Teijeiro,[5] Fernanda Milanezi,[1] Fernando Schmitt,[1,4] Joana Paredes[1,4]

► Additional tables and figure are published online only. To view these files please visit the journal online (http://jcp.bmj.com).

[1]IPATIMUP—Institute of Molecular Pathology and Immunology of Porto University, Porto, Portugal
[2]ICBAS—Abel Salazar Biomedical Science Institute, Porto, Portugal
[3]Vale Polytechnic Health Institute of the North, Vale do Sousa Higher School of Health, Department of Anatomical Pathology Cytology and Tanatology, Gandra, Portugal
[4]Medical Faculty of Porto University, Porto, Portugal
[5]Complexo Hospitalar Universitario de Vigo (CHUVI), Vigo, Spain

Correspondence to
Joana Paredes, IPATIMUP, Rua Dr Roberto Frias s/n, 4200-465 Porto, Portugal;
jparedes@ipatimup.pt

Accepted 22 May 2011
Published Online First 16 June 2011

ABSTRACT

Background and Aim The study of CD44/CD24 and ALDH1 expression is the most accurate method to identify cancer stem cells (CSC) from breast cancer populations. However, the overlap between $CD44^+CD24^{-/low}$ and $ALDH1^{high}$ CSC phenotypes in breast cancer seems to be very small, as well as their distribution among intrinsic breast cancer subtypes. Due to this discrepancy, it is imperative to improve the understanding of breast CSC marker distribution.

Methods 466 invasive breast carcinomas and eight breast cancer cell lines were analysed for the expression of CD44, CD24 and ALDH1, to evaluate their distribution among the distinct molecular subtypes.

Results Basal-like tumours (76.5%) contained the higher percentage of cells with the CSC phenotype $CD44^+CD24^{-/low}$ ($p<0.0001$). From ALDH1-positive cases, 39.4% were also basal-like tumours ($p<0.0001$). The analysis of breast cancer cell lines indicated that luminal cell lines are mainly enriched in a $CD44^{-/low}CD24^+$ cell population, basal/mesenchymal breast cancer cell lines are enriched in the $CD44^+CD24^{-/low}$ phenotype, whereas the remaining basal/epithelial cell lines are mainly positive for both markers. ALDH1 activity was mainly found in HER-OE and basal/epithelial breast cancer cell.

Conclusions $CD44^+CD24^{-/low}$ and $ALDH1^+$ phenotypes seem to identify CSC with distinct levels of differentiation. It seems that the paramount method and biomarkers that identify breast CSC within the distinct molecular subtypes need to be better explored, because it is pivotal to translate the CSC concept to clinical practice. In the future, the recognition of reliable markers to distinguish the CSC pool in each molecular subtype will be decisive for the development of specific target therapies.

Breast cancer is the most frequent cancer among women,[1] being a heterogeneous disease, with distinct morphologies, metastatic behaviour and therapeutic response. It is actually known that variation in transcriptional programmes is the major reason for biological diversity among human breast cancers.[2] In fact, global gene-expression analyses have provided an appealing molecular classification for breast carcinomas, which is highly associated with patients' prognosis.[2–5]

The molecular classification of breast cancer established four major subtypes: the luminal A and B, the HER2-overexpressing (HER2-OE) and basal-like tumours.[2–4] Luminal A is the most prevalent subtype and is characterised by the expression of oestrogen and progesterone receptors (ER and PgR, respectively) in cancer cells, whereas the luminal B subtype is characterised by ER and PgR expression together with HER2 overexpression and/or high rates of cell proliferation. In contrast, HER2-OE tumours are negative for hormonal receptors and overexpress HER2 protein, which is highly associated with gene amplification. Finally, within triple-negative tumours, characterised by the absence of ER, PgR and HER2 expression, the basal-like subtype still constitutes a heterogeneous group of tumours, expressing distinct basal markers. Actually, the correct identification of basal-like breast carcinomas is clinically relevant, because these are highly associated with aggressive histological features and poor patient survival, still lacking an efficient therapy.[6–10]

In the past decade, many treatments undergoing clinical trials have been developed based on breast cancer molecular profiles.[11] However, one of the most promising therapy targets came with the identification of a pool of cancer cells with stem characteristics—cancer stem cells (CSC). The CSC model proposes that tumours, as normal tissues, are organised in a cellular hierarchy, in which CSC are the only cells with unlimited proliferation and tumorigenic potential; therefore, being capable of driving tumour growth, progression and metastasis due to their stem cell-like characteristics: self-renewal and differentiation.[12 13] Recent evidence has demonstrated that CSC are resistant to various forms of therapies, including radio and chemotherapy.[14–20] Based on these observations, the CSC model became the foundation for new preventive and therapeutic strategies in cancer.

In breast cancer, the first report identifying and isolating tumorigenic CSC from non-tumorigenic cancer cells used the combined expression of two cell surface markers: $CD44^+/CD24^{-/low}$.[21–26] Interestingly, some studies revealed an enrichment of the $CD44^+/CD24^{-/low}$ and $CD44^-/CD24^+$ cell populations in basal-like and luminal breast cancer cell lines, respectively,[27 28] CD44 being positively associated with stem cell-like characteristics and CD24 expression related to differentiated epithelial features.[29] These in-vitro data were later demonstrated in primary breast carcinomas,[30] but the clinical and prognostic impact of these markers in

breast cancer remains a controversial issue,[25 31–33] demanding additional efforts to find other CSC markers that could better predict breast cancer patient survival.

Using in-vitro and in-vivo experimental systems, Ginestier *et al*[34] demonstrated that normal and cancer human mammary epithelial cells with increased aldehyde dehydrogenase activity (ALDH) show stem/progenitor cell properties. Tumorigenic ALDH1$^+$ CSC are significantly more resistant to platinum treatments, are biologically aggressive, and their expression tends to be associated with a poor patient prognosis.[34–36] Interestingly, CD44$^+$CD24$^{-/low}$ cells and ALDH1+ cells are more frequently found in basal-like than in luminal tumours; however, ALDH1$^+$ cells are also commonly found in the HER2-OE subtype.[34] It was recently shown that ALDH1 breast CSC marker can further divide the CD44$^+$CD24$^{-/low}$ cell population into fractions that are highly tumorigenic: ALDH1$^+$CD44$^+$CD24$^{-/low}$ cells were able to generate tumours from only 20 cells, whereas ALDH1$^-$CD44$^+$CD24$^{-/low}$ were not tumorigenic in this same cell density.[34 37]

Based on this current knowledge, there is evidence to support the idea that the use of CD44 and CD24 cell surface markers in combination with ALDH1 activity is the most accurate method to identify and isolate CSC-like cells within breast cancer populations. However, the overlap between CD44$^+$CD24$^{-/low}$ and high ALDH1 expression in primary tumours is quite small (approximately 1%).[34] Due to this discrepancy, it is imperative to improve CSC identification into routine formalin-fixed and paraffin-embedded tissue samples.

In the present study, we analysed the expression of the main established breast CSC markers—CD44, CD24 and ALDH1, in a large series of invasive breast carcinomas, in order to evaluate their distribution among the different molecular subtypes. In addition, we investigated the correlation between the presence of these markers and the clinicopathological features and patient survival. Finally, these features were compared with the results obtained with breast cancer cell lines from distinct molecular subtypes, in which the different cancer cell populations, expressing these CSC markers, were selected by flow cytometry.

MATERIAL AND METHODS

Patient selection

A series of 466 primary and sporadic invasive breast carcinomas was retrieved from the Pathology Department, Hospital Xeral-Cíes, Vigo, Spain, diagnosed in 1978–1992. Patients' ages ranged from 28 to 92 years of age. The formalin-fixed paraffin-embedded histological sections were reviewed and the diagnoses confirmed. The tumours have been characterised for clinical and pathological features—namely age, tumour size, lymph nodes status and histological grade (data summarised in supplementary table S1, available online only). Patient follow-up information was available for 455 cases, ranging from a minimum of one to a maximum of 120 months after the diagnosis. The disease-free survival (DFS) interval was defined as the time from the diagnosis to the date of breast-cancer-derived relapse/metastasis, whereas overall survival (OS) was considered as the number of months from the diagnosis to the disease-related death. This study was conducted under the national regulative law for the handling of biological specimens from tumour banks, being the samples exclusively available for research purposes in retrospective studies.

TMA construction and immunohistochemistry

Representative tumour areas were selected on haematoxylin and eosin-stained sections and marked on paraffin blocks. At least two tissue cores (0.6 mm in diameter) were obtained from each selected specimen and deposited into a recipient paraffin block, using a tissue microarray (TMA) workstation (Manual Tissue Arrayer; Beecher Instruments, Inc. Sun Prairie, Wisconsin, USA). The 12 TMA blocks were designed and constructed according to rules previously described[38] and non-neoplastic tissue cores were included as controls.

In order to classify all breast cancer tumours molecularly, we evaluated the expression of the most commonly used breast cancer biomarkers,[10] namely the hormonal receptors ER and PgR, the proliferation marker Ki67, the tyrosine kinase receptors HER2 and EGFR, the basal cytokeratins CK5 and CK14 and also P-cadherin and vimentin basal markers. Immunohistochemistry was performed in 3 μm sections. To study CSC markers in this series, specific antibodies for CD44 (clone 156-3C11; Cell Signaling Technology, Danvers, Massachusetts, USA), CD24 (clone Ab2-SN3b; Neomarkers, Fremont, California, USA) and ALDH1 (clone EP1933Y; Abcam, Cambridge, Massachusetts, USA) were assessed. The primary antibodies were detected using a secondary antibody with horseradish peroxidase polymer (Cytomation Envision System HRP; DAKO, Carpinteria, California, USA), or a biotinylated goat anti-polyvalent as secondary antibody, followed by the streptavidin-peroxidase complex (Thermo Fisher Scientific, Fremont, California, USA), according to the manufacturer's instructions. Both methods used diaminobenzidine as chromogen. Detailed conditions for each antibody can be found in supplementary table S2 (available online only).

Immunohistochemical evaluation

The expression of the breast cancer biomarkers ER, PgR, HER2, EGFR, CK5, CK14, P-cadherin and vimentin was evaluated according to the grading systems already described.[10] The quantification of cell proliferation by Ki67 expression was measured using the publicly available web application software ImunoRatio, as recently described by Tuominen *et al*,[39] and validated by a breast cancer pathologist. The cut-off value to distinguish low from high proliferation tumours was 13.25% of Ki67 nuclear staining. The Ki67 index was based on the study published by Cheang and colleagues,[40] in which its expression was considered as a continuous variable and the cut point was determined by the receiver operating characteristic method, using gene expression profile as the gold standard. These immunohistochemical results were used to classify the tumours in the different molecular breast cancer subtypes, namely in luminal A, luminal B, HER2-OE and basal-like, according to supplementary table S3 (available online only).

CD44 and CD24 staining were detected mainly at the membrane of tumour cells and the scoring was considered as follows: 0, 0–10% of positive tumour cells; 1+, 10–25% of positive tumour cells; 2+, 25–50% of positive tumour cells; 3+, more than 50% of positive tumour cells. Cytoplasmic staining was not considered for any of these markers, in order to compare these results with those obtained by flow cytometry in cell lines, which selects only cells expressing these markers at the cell surface. For CD44, the cases classified as 0 were considered negative, whereas 1+, 2+ and 3+ were established as positive cases. For CD24, the cases were divided into negative/low (−/low), when considered 0 or 1+, or in positive cases, when classified as 2+ or 3+. Immunohistochemical staining of ALDH1 was classified as positive when more than 1% of tumour cells showed clear cytoplasmic positivity, as previously described.[34 36] Stromal expression of ALDH1 was also classified in two categories: none/weak, or moderate/strong, as previously described by Resetkova *et al*.[41]

Immunofluorescence

To control the reliability of the CD44 and CD24 single staining and evaluation, double staining immunofluorescence with the same primary antibodies was performed in 10% of all cases, not only in TMA, but also in the whole tissue. Detection of the primary antibody anti-CD44 was performed using a secondary antibody goat anti-mouse IgG (H+L) conjugated with Alexa Fluor 488 (Cell Signaling Technology, Inc.) and the detection of the anti-CD24 was done using a secondary antibody goat anti-mouse IgM (μ chain) Alexa Fluor 594 (Cell Signaling Technology, Inc.). The results from both techniques were exactly the same.

Cell culture

Human breast cancer cell lines MCF-7/AZ, T47D, SkBr3, BT474, BT-20, MDA-MB-468, BT-549 and MDA-MB-231 were obtained from ATCC or from collections developed at Professor Mareel's laboratory (Laboratory of Experimental Cancerology, Ghent University Hospital, Belgium). All cell lines were grown in Dulbecco's modified essential medium (Invitrogen, Carlsbad, California, USA), supplemented with 10% fetal bovine serum (Invitrogen) and with 1% antibiotic solution (penicillin–streptomycin; Invitrogen), being routinely cultured in a humidified atmosphere with 5% carbon dioxide and at 37°C. These breast cancer cell lines were selected to be studied, because they harbour distinct molecular profiles, as already described[42–44] (see supplementary table S4, available online only).

Flow cytometry

Cells were washed twice with phosphate-buffered saline and then harvested with versene/0.48 mM ethylenediammine tetraacetic acid (Gibco, Invitrogen Ltd., Paisley, UK). Detached cells were re-suspended in phosphate-buffered saline supplemented with 0.5% fetal bovine serum (1×10^6 cells/50 μl). Combinations of fluorochrome-conjugated monoclonal antibodies against human CD44 (FITC; cat. #555478) and CD24 (PE; cat. #555428) were obtained from BD Biosciences (San Diego, California, USA). Primary antibodies or the respective isotype controls (BD Biosciences) were added to the cell suspension, as recommended by the manufacturer, and incubated at 4°C in the dark for 20 min. The labelled cells were analysed on a FACS Calibur (BD Biosciences).

ALDEFLUOR assay

The ALDEFLUOR kit (Stem Cell Technologies, Grenoble, France) was used to analyse the cell population with high ALDH enzymatic activity, using a FACS Calibur (BD Biosciences), according to the manufacturer's instructions. Briefly, cells were incubated in ALDEFLUOR assay buffer containing ALDH substrate (1 μmol/l per 1×10^6 cells). In each experiment, a sample of cells was incubated, under identical conditions, with 50 mmol/l of diethylaminobenzaldehyde, a specific ALDH inhibitor, as a negative control.

Statistical analysis

Association between the CD44/CD24 phenotype and ALDH1 pattern and the different molecular subtypes, the clinicopathological parameters or the different molecular markers were assessed by Pearson correlation and χ^2 tests. Survival analyses were estimated using the Kaplan–Meier method and compared using the log-rank test. Statistical analyses were carried out using SPSS statistics V.17.0 software, and a significance level of 5% was considered statistically significant.

RESULTS

Tumour classification in breast cancer molecular subtypes

The series of invasive breast carcinomas was studied for the expression of ER, PgR, HER2, Ki67, EGFR, CK5, P-cadherin, CK14 and vimentin, in order to classify them in the different immunohistochemical molecular subtypes. The results from the different biomarkers are shown in supplementary table S5 (available online only). From the 466 invasive breast cancer cases, 64.8% (302/466) were luminal A, 8.8% (41/466) luminal B, 7.1% (33/466) HER2-OE, 14.6% (68/466) basal-like tumours and 4.7% (22/466) were unclassified tumours (figure 1A and supplementary table S1, available online only). As expected, the majority of basal-like and HER2-OE tumours were grade III, highly proliferative, with worse patient survival curves (figure 1B), demonstrating the validity and power provided by this series of invasive breast carcinomas.

Association between the expression of CD44, CD24 and ALDH1 with other breast cancer parameters

The expression of CD44, CD24 and ALDH1 was analysed in all breast cancer cases and an example of the pattern of expression of these three CSC markers is shown in supplementary figure 1 (available online only). Concerning CD44 membrane staining, 51.2% (237/463) of the cases were positive. In contrast, for membrane CD24, the majority of the cases (88.7%, 411/463) were classified as negative/low, and only 11.4% (53/463) of the tumours had clear membrane staining. For ALDH1, a minority of cases (7.1%, 33/464) was classified as positive, showing a clear cytoplasmic expression in tumour cells. Moderate/strong stromal staining for ALDH1 was also observed in 37.8% (176/466) of the cases.

When CSC markers were associated with classic prognostic factors, as well as with other biomarkers studied, CD44 expression was significantly associated with lymph node metastasis ($p=0.006$), and with the expression of basal markers: EGFR ($p=0.038$), CK5 ($p<0.0001$), P-cadherin ($p=0.003$), CK14 ($p=0.005$) and vimentin ($p<0.0001$) (table 1). In contrast, any significant correlation between single CD24 expression and the other parameters evaluated was found (table 1). Concerning ALDH1 cytoplasmic expression, it was significantly associated with ER negativity ($p=0.003$), and with basal marker expression, namely EGFR ($p=0.004$), CK5 ($p<0.0001$), P-cadherin ($p<0.0001$), CK14 ($p<0.0001$) and vimentin ($p=0.01$); no association was found with HER2 overexpression. Concerning classic prognostic factors, ALDH1 expression was significantly correlated with high grade tumours, as 78.8% (26/33) of the positive cases were grade III ($p=0.003$) (table 1). When CSC markers were associated within themselves, a significant association between $CD24^{-/low}$ tumours and ALDH1 expression ($p=0.018$) was found, 75.8% (25/33) of the ALDH1-positive cases also being $CD24^{-/low}$ (data not shown). Concerning stromal ALDH1 staining, no associations were found with the several parameters evaluated. However, there was a significant association between CD24 positivity and moderate/strong stromal ALDH1 expression ($p=0.018$) (see supplementary table S6, available online only).

CSC markers, breast cancer molecular subtypes and patient survival

CD44 expression was significantly associated with breast cancer molecular subtype ($p<0.0001$), whereas CD24 was not ($p=0.418$) (table 1). The majority of basal-like carcinomas (80.9%, 55/68) were considered $CD44^+$, in contrast to what was verified in others subtypes (table 1). In addition, almost all

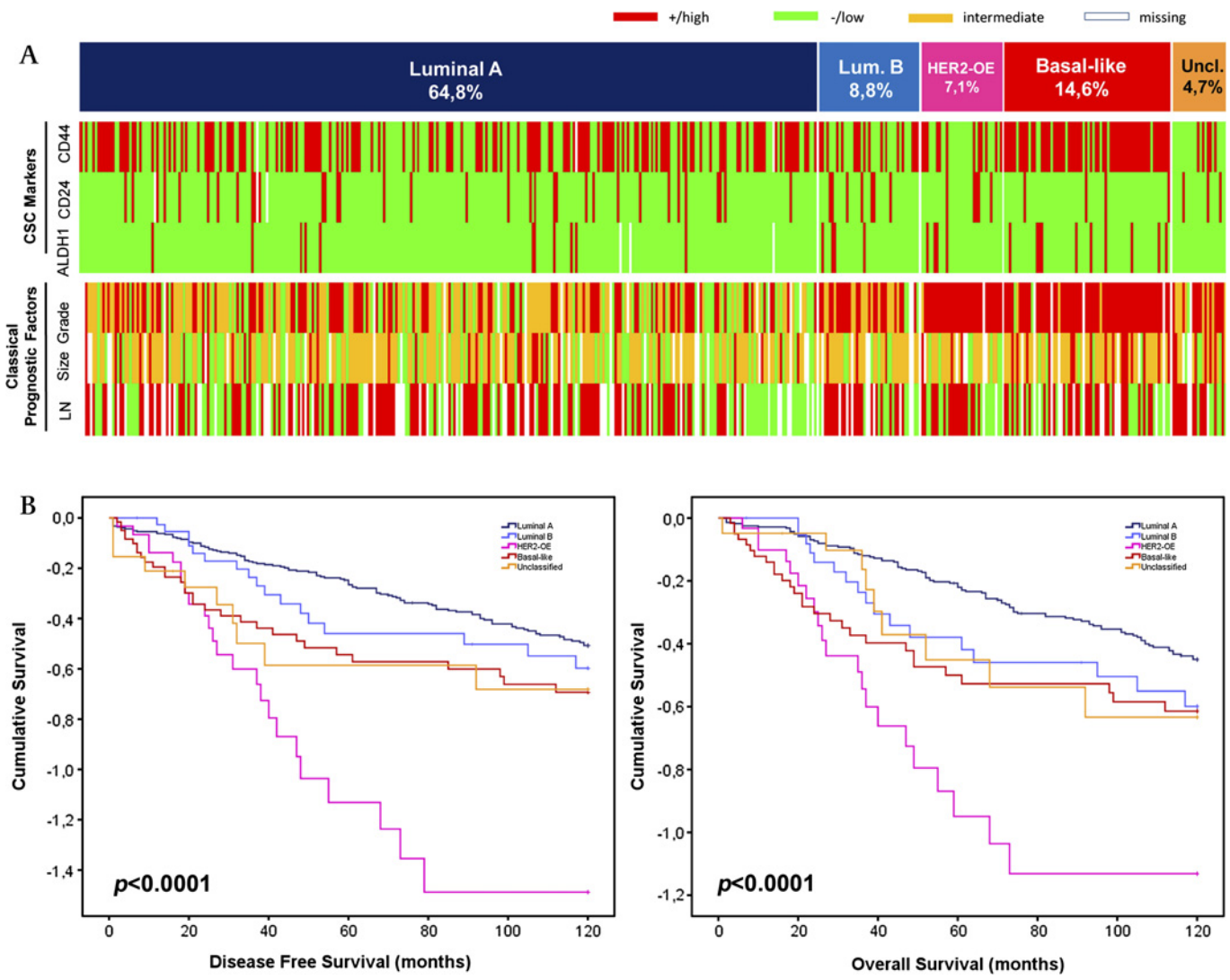


Figure 1 Breast tumour signature represented by immunohistochemistry array showing the protein expression of cancer stem cell markers (CD24, CD44 and ALDH1) and classic prognostic factors (tumour size, histological grade and lymph node metastasis) in the series of 466 invasive breast carcinomas analysed in this study (A); and disease-free survival and overall survival of the different molecular subtypes (B).

basal-like tumours were $CD24^{-/low}$ (94.1%, 64/68) and among $CD24^{+}$ cases 64.2% (34/53) were luminal A (table 1). Concerning ALDH1 cytoplasmic expression, 39.4% (13/33) were classified as basal-like carcinomas, this association being statistically significant ($p<0.0001$) (table 1). ALDH1 stromal staining did not correlate with any molecular subtype (see supplementary table S6, available online only).

To explore the effect of the CSC phenotype $CD44^{+}CD24^{-/low}$ on the clinical outcome and its prevalence within the intrinsic molecular subtypes, we decided to consider a tumour with CSC phenotype when the frequency of $CD44^{+}CD24^{-/low}$ cells were more than 10%, as previously described in other studies.[25 32] In our series, 45.3% (209/461) of the tumours were considered $CD44^{+}CD24^{-/low}\geq10\%$ and 54.7% (252/461) $CD44^{+}CD24^{-/low}<10\%$. The CSC phenotype $CD44^{+}CD24^{-/low}$ was significantly increased in node-negative tumours ($p<0.0001$) and in tumours expressing the basal markers CK5 ($p<0.0001$), P-cadherin ($p=0.008$), CK14 ($p=0.003$) and vimentin ($p<0.0001$) (table 2).

Table 3 shows that the CD44/CD24 phenotype was also significantly associated with breast cancer molecular subtypes ($p<0.0001$). Most of the basal-like tumours (76.5%, 52/68) were classified as $CD44^{+}CD24^{-/low}\geq10\%$ (figure 2), independently of ALDH1 expression. Luminal tumours showed a mixture between the two CD44/CD24 phenotypes: 43.0% (127/295) of luminal A tumours were $CD44^{+}CD24^{-/low}\geq10\%$ and 57.0% (168/295) were $CD44^{+}CD24^{-/low}<10\%$. Luminal B showed the same CSC markers distribution (41.4% $CD44^{+}CD24^{-/low}\geq10\%$ and 58.6% $CD44^{-}CD24^{-/low}<10\%$).

Univariate survival analyses were performed for the three CSC markers, namely CD44, CD24 and ALDH1, as well as for the combined expression of CD44/CD24, and all failed to reach statistically significant levels, meaning that these markers were not significant predictors of DFS or OS (data not shown). However, when we addressed the distribution of the CD44/CD24 pattern within the poor prognosis basal-like carcinomas, we found that tumours with more than 10% of the cells with the CSC phenotype showed a trend to be associated with a worse DFS ($p=0.065$) and OS ($p=0.127$) (figure 3). In accordance with the Kaplan–Meier survival curves, Cox univariate survival analysis, within basal-like carcinomas, demonstrated a tendency for tumours with more than 10% of $CD44^{+}CD24^{-/low}$ cells to present an increased risk of DFS, when compared with tumours with less than 10% of $CD44^{+}CD24^{-/low}$ cells. Nevertheless, multivariate analysis, with models including tumour

Table 1 Associations between the expression of the breast cancer stem cell markers CD44, CD24 and ALDH1 and the classic breast cancer prognostic factors, biological markers and molecular subtypes

	CD44				CD24				ALDH1			
	n	Positive	Negative	p Value	n	Positive	Neg/low	p Value	n	Positive	Negative	p Value
Tumour size	405	207	198	0.414	405	50	355	0.286	405	31	374	0.013
T1 <2 cm	100	49	51		100	13	87		101	1	100	
T2 2–5 cm	242	121	121		242	33	209		241	23	218	
T3 >5 cm	63	37	26		63	4	59		63	7	56	
Lymph nodes	364	191	173	0.006	365	46	319	0.055	362	26	337	0.263
Positive	206	95	111		206	32	174		206	12	194	
Negative	158	96	62		159	14	145		157	14	143	
Histological grade	440	228	212	0.496	440	51	389	0.065	440	33	407	0.003
Grade I	81	42	39		80	4	76		81	1	80	
Grade II	134	64	70		133	14	119		134	6	128	
Grade III	225	122	103		227	33	194		225	26	199	
ER	461	236	225	0.057	461	53	408	0.574	461	33	428	0.003
Positive	306	147	159		306	37	269		306	14	292	
Negative	155	89	66		155	16	139		155	19	136	
PgR	462	237	225	0.346	462	53	409	0.374	462	33	429	0.063
Positive	226	121	105		227	23	204		226	11	215	
Negative	236	116	120		235	30	205		236	22	214	
HER2	459	234	225	0.080	459	52	407	0.075	459	33	426	0.114
Positive	68	28	40		68	12	56		68	8	60	
Negative	391	206	185		391	40	351		391	25	366	
Ki67	442	228	214	0.988	443	53	390	0.754	441	33	408	0.181
High	29	15	14		29	4	25		29	4	25	
Low	413	200	213		414	49	365		412	29	383	
EGFR	463	237	226	0.038	463	53	410	0.741	463	33	430	0.004
Positive	22	16	6		22	3	19		22	5	17	
Negative	441	221	220		441	50	391		441	28	413	
CK5	462	237	225	<0.0001	463	53	410	0.546	462	33	429	<0.0001
Positive	66	52	14		66	9	57		66	12	54	
Negative	396	185	211		397	44	353		396	21	375	
P-cadherin	463	237	226	0.003	463	53	410	0.747	463	33	430	<0.0001
Positive	114	72	42		114	14	100		114	17	97	
Negative	349	165	184		349	39	310		349	16	333	
CK14	462	237	225	0.005	463	52	410	0.250	462	33	429	<0.0001
Positive	24	19	5		24	1	23		24	7	17	
Negative	438	218	220		439	52	387		438	26	412	
Vimentin	455	234	222	<0.0001	456	53	403	0.234	455	33	422	0.010
Positive	78	62	16		78	6	72		78	11	67	
Negative	377	172	205		378	47	331		377	22	355	
Molecular subtypes	464	237	226	<0.0001	463	53	410	0.418	463	33	430	<0.0001
Luminal A	299	147	152		299	34	265		299	12	287	
Luminal B	41	17	24		41	7	34		41	4	37	
HER2-OE	33	12	21		33	5	28		33	4	29	
Basal-like	68	56	12		68	4	64		68	13	55	
Unclassified	22	5	17		22	3	19		22	0	22	

ER, oestrogen receptor; PgR, progesterone receptor.

size, grade and lymph node involvement, showed that CD24/CD44 was not an independent factor of the prediction of patient DFS (data not shown).

$CD44^{+}CD24^{-/low}$ phenotype and ALDH1 activity in breast cancer cells

Flow cytometric analysis allows us to separate the cancer cell populations according to different levels of the surface CSC markers CD44 and CD24. As shown in figure 4A–C, luminal (MCF-7/AZ and T47D) and HER2-OE (SkBr3 and BT474) breast cancer cell lines are mainly constituted by cells with high levels of CD24 and low levels of CD44, in accordance with an epithelial luminal cell phenotype. In contrast, the basal/epithelial cell lines BT-20 and MDA-MB-468 showed enrichment in cell populations with high levels of both markers. Basal/mesenchymal BT-549 and MDA-MB-231 cells show a lower expression of CD24, reflecting their mesenchymal phenotype.

The activity of ALDH1 enzyme was also evaluated in this panel of breast cancer cells, using the ALDEFLUOR assay. Figure 4C shows the percentage of the putative CSC fraction obtained in the different cell lines. In this analysis, luminal breast cancer cell lines showed the lowest percentage of tumour cells with ALDH1 activity; in contrast, HER2-OE and basal-like breast cancer cell lines showed increased levels of ALDH1 activity; the exception was the MDA-MB-231 cell line, in which we could not detect any ALDEFLUOR-positive subpopulation.

Table 2 Associations between the combined expression of CD44/CD24 and the classic breast cancer prognostic factors and biomarkers

	n	$CD44^+$ $CD24^{-/low}$<10%	$CD44^-$ $CD24^{-/low}$≥10	p Value
Tumour size	403	222	181	0.362
T1 <2 cm	100	54	46	
T2 2–5 cm	240	138	102	
T3 >5 cm	63	30	33	
Lymph nodes	364	195	169	<0.0001
Positive	206	127	79	
Negative	158	68	90	
Histological grade	438	236	202	0.582
Grade I	80	40	40	
Grade II	133	76	57	
Grade III	225	120	105	
ER	459	251	208	0.082
Positive	304	175	129	
Negative	155	76	79	
PgR	460	251	209	0.480
Positive	225	119	106	
Negative	235	132	103	
HER2	457	250	207	0.126
Positive	68	43	25	
Negative	389	207	182	
Ki67	441	241	200	0.657
High	29	17	12	
Low	412	224	188	
EGFR	461	252	209	0.184
Positive	22	9	13	
Negative	439	243	196	
CK5	461	252	209	<0.0001
Positive	66	22	44	
Negative	395	230	165	
P-cadherin	461	252	209	0.008
Positive	114	50	64	
Negative	347	202	145	
CK14	461	252	209	0.003
Positive	24	6	18	
Negative	437	246	191	
Vimentin	454	248	206	<0.0001
Positive	78	19	59	
Negative	376	229	147	
ALDH1	459	250	209	0.599
Positive	32	16	16	
Negative	427	234	193	

ER, oestrogen receptor; PgR, progesterone receptor.

DISCUSSION

One of the recent priorities in breast cancer research is CSC identification/isolation, because it is well accepted that tumours are essentially driven by a cellular pool with stem-like properties, which are responsible for tumour invasiveness, heterogeneity, metastasis capacity and therapy resistance.[45 46] In this study, we analysed the immunohistochemical membrane localisation of the breast CSC markers CD44 and CD24, as well as the presence of intracellular ALDH1, in a large and well characterised series of invasive breast carcinomas. These results were compared with those obtained by flow cytometry in breast cancer cell lines from distinct molecular subtypes, studying the same panel of CSC markers.

As shown in figure 1, the CD44 CSC marker was commonly expressed among primary breast carcinomas (51.2% of positive cases), whereas expression of CD24 and ALDH1 was present in a minority of cases (11.4% and 7.1%, respectively). When the same CSC markers were studied in the selected breast cancer cell lines, half of them expressed high levels of CD44. However, the majority of the cell lines expressed increased amounts of membrane CD24, as well as a high percentage of ALDEFLUOR-positive cells.

Indeed, the results for CD44 were the most comparable between tumours and cell lines, and the results obtained in previous studies.[29] This agreement is probably associated with the specific and clear membrane staining observed for CD44. CD44 was significantly expressed in poor prognosis basal-like tumours and aggressive basal-like cell lines, and was highly associated with basal markers (EGFR, CK5, P-cadherin, CK14 and vimentin). It has already been demonstrated that $CD44^+$ cells show a mesenchymal stem cell-like profile, enriched for genes involved in cell motility, proliferation and angiogenesis, and its positivity has been associated with decreased patient survival.[33] CD44 expression was also inversely associated with lymph node metastasis, as previously shown by Giatromanolaki *et al*,[47] probably because basal-like tumours usually also metastasise via a haematogenic route.[48] In addition, it has already been shown that stem-like gene expression patterns, in lymph node-negative primary breast tumours, correlate with shorter distant metastasis-free survival.[33] All these results reinforce the prognostic relevance of this CSC marker and its possible use as a therapeutic target.

Concerning CD24 membrane staining, the results were not concordant between primary tumours and cell lines, or with previous literature data. In tumours, only a small percentage of the cases showed clear cut membrane positivity; however, with the exception of MDA-MB-231, all cell lines showed CD24 positivity by flow cytometry. Distinct grading systems have been used to classify CD24 immunohistochemical results[49 50] and, consequently, different percentages of CD24 expression have been observed in other series of invasive breast carcinomas. For example, Mylona *et al*[32] considered mainly membrane CD24, whereas Honeth *et al*[30] considered CD24 staining at the cytoplasm, possibly explaining why opposite conclusions were drawn by both studies. Indeed, cytoplasmic expression can reflect aberrant protein overexpression, with consequent disturbance of its membrane distribution and degradation in neoplastic cells;[51] thus, its significance to the most appropriate CD24 classification is still ambiguous, and needs to be discussed further. Moreover, the extension of staining to

Table 3 Associations between the combined expression of CD44/CD24/ALDH1 and the breast cancer molecular subtypes

		Luminal A	Luminal B	HER2-OE	Basal-like	Unclassified
$CD44^+CD24^{-/low}$<10%	$ALDH1^+$	9 (3.1%)	2 (4.9%)	2 (6.1%)	6 (8.8%)	0 (0%)
	$ALDH1^-$	159 (53.9%)	22 (53.7%)	22 (66.7%)	10 (14.7%)	18 (81.8%)
$CD44^+CD24^{-/low}$≥10%	$ALDH1^+$	6 (2.0%)	3 (7.3%)	2 (6.1%)	11 (16.2%)	0 (0%)
	$ALDH1^-$	121 (41.0%)	14 (34.1%)	7 (21.2%)	41 (60.3%)	4 (18.2%)
Total		295 (100%)	41 (100%)	33 (100%)	68 (100%)	22 (100%)

p≤0.0001.

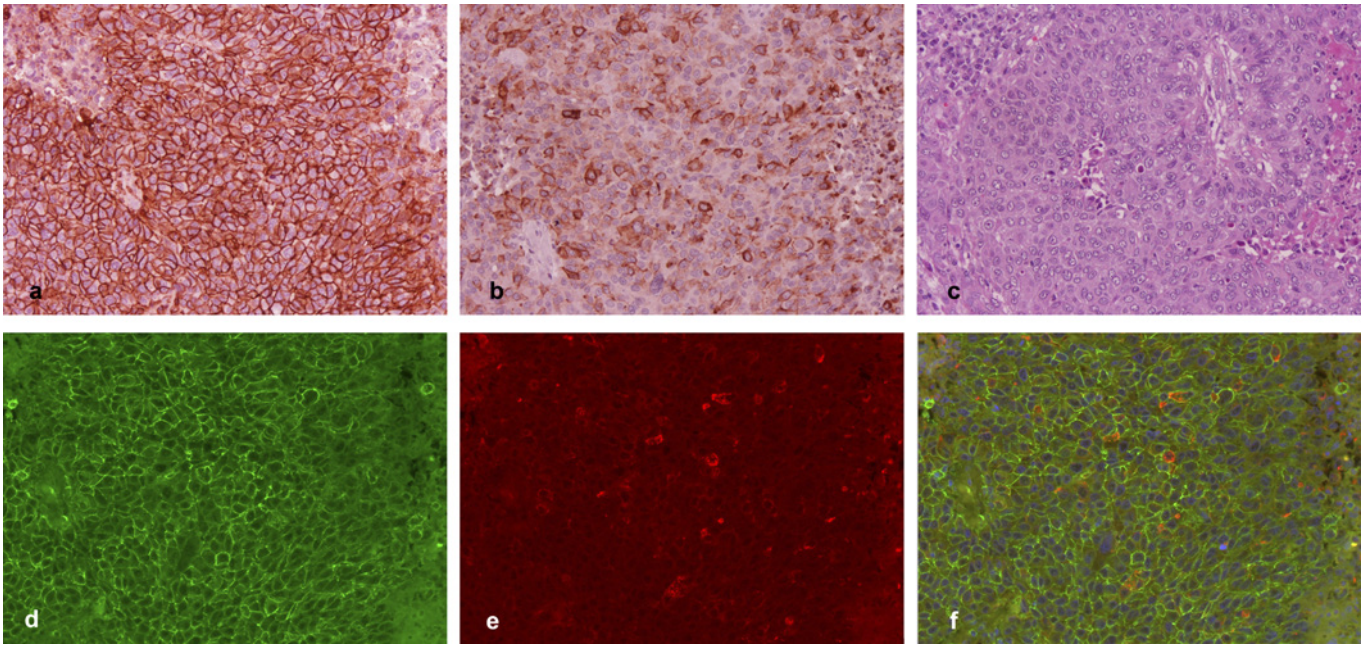


Figure 2 Basal-like carcinoma showing $CD44^+CD24^{-/low} \geq 10\%$ phenotype. Single-staining immunohistochemistry for $CD44^+$ (A), $CD24^{-/low}$ (B) and haematoxylin–eosin (C); and double-staining immunofluorescence for $CD44^+$ (D), $CD24^{-/low}$ (E) and merged image (F) (magnification ×200).

consider a CD24-positive case is also diverse among these studies:[32 35 52] some have been categorising the CD24 marker as $CD24^+$ versus $CD24^-$,[29 30 35] whereas others compare $CD24^{-/low}$ versus $CD24^+$.[25 32] These variables certainly affect the results concerning breast CSC identification in tumours and, therefore, the prognostic value of this marker. Nevertheless, we found that $CD24^+$ cases were enriched in luminal A tumours (34/53, 64.2%), while the majority of the basal-like tumours were classified as $CD24^{-/low}$ (64/68, 94.1%). Although not statistically significant, these results are in agreement with previous studies addressing the fact that $CD24^+$ cells are related to more differentiated tissues or tumours, whereas $CD24^{-/low}$ cells have stem or progenitor-like properties.[29 30 33] This same trend was observed in cell lines, because those maintaining an epithelial phenotype showed enrichment in $CD24^+$ cells, whereas the mesenchymal cell lines BT-549 and MDA-MB-231 showed lower levels or no expression of CD24, respectively.

The combinatorial evaluation of CD44/CD24 for the identification of the CSC population in breast cancer cell lines mimicked in a way the results found in primary tumours. The majority of basal-like tumours showed more than 10% of cells expressing the CSC phenotype $CD44^+CD24^{-/low}$, which was also the main phenotype found in the basal/mesenchymal MDA-MB-231 breast cancer cell line. The remaining basal cell lines were positive for both markers, which we believe are cancer cells representative of basal-like tumours with higher levels of differentiation (basal-like A, which maintain an epithelial phenotype), whereas the MDA-MB-231 cells are representative of the most poorly differentiated basal tumours (basal-like B), showing a mesenchymal phenotype and CD24 negativity.[53] This hypothesis reinforces the idea that CSC marker expression rather reflects the cell of origin of the different breast cancer lesions. It has already been reported that luminal progenitors (which are $CD24^+$) are the most probable cell of origin of the majority of basal-like carcinomas,[54] explaining the positivity for the CD24 marker in these basal-like breast cancer cell lines. It is also important to point out that, as CD24 expression presents a dynamic regulation, as recently demonstrated by Meyer *et al*,[55] $CD44^+CD24^+$ cells can readily give rise to $CD44^+CD24^{-/low}$

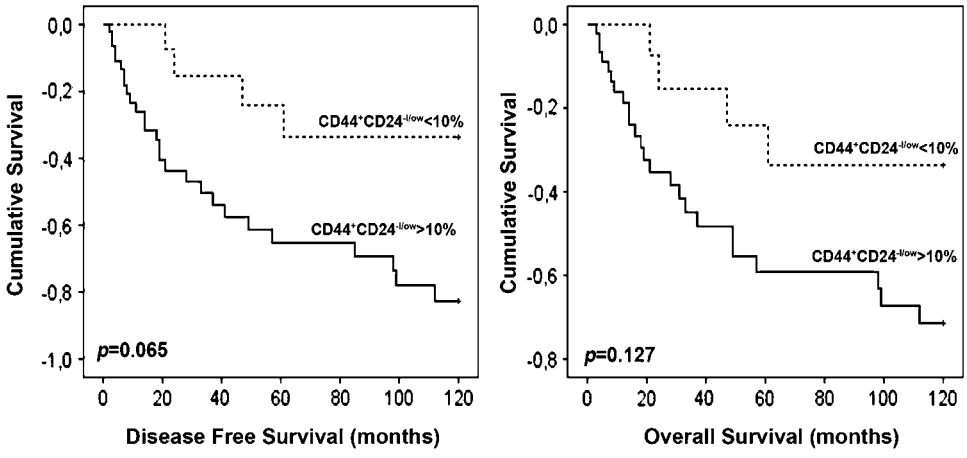


Figure 3 Kaplan–Meier plots of disease-free survival (DFS, p=0.065) and overall survival (OS, p=0.127) in the basal-like tumours defined according to the CD44/CD24 pattern of expression.

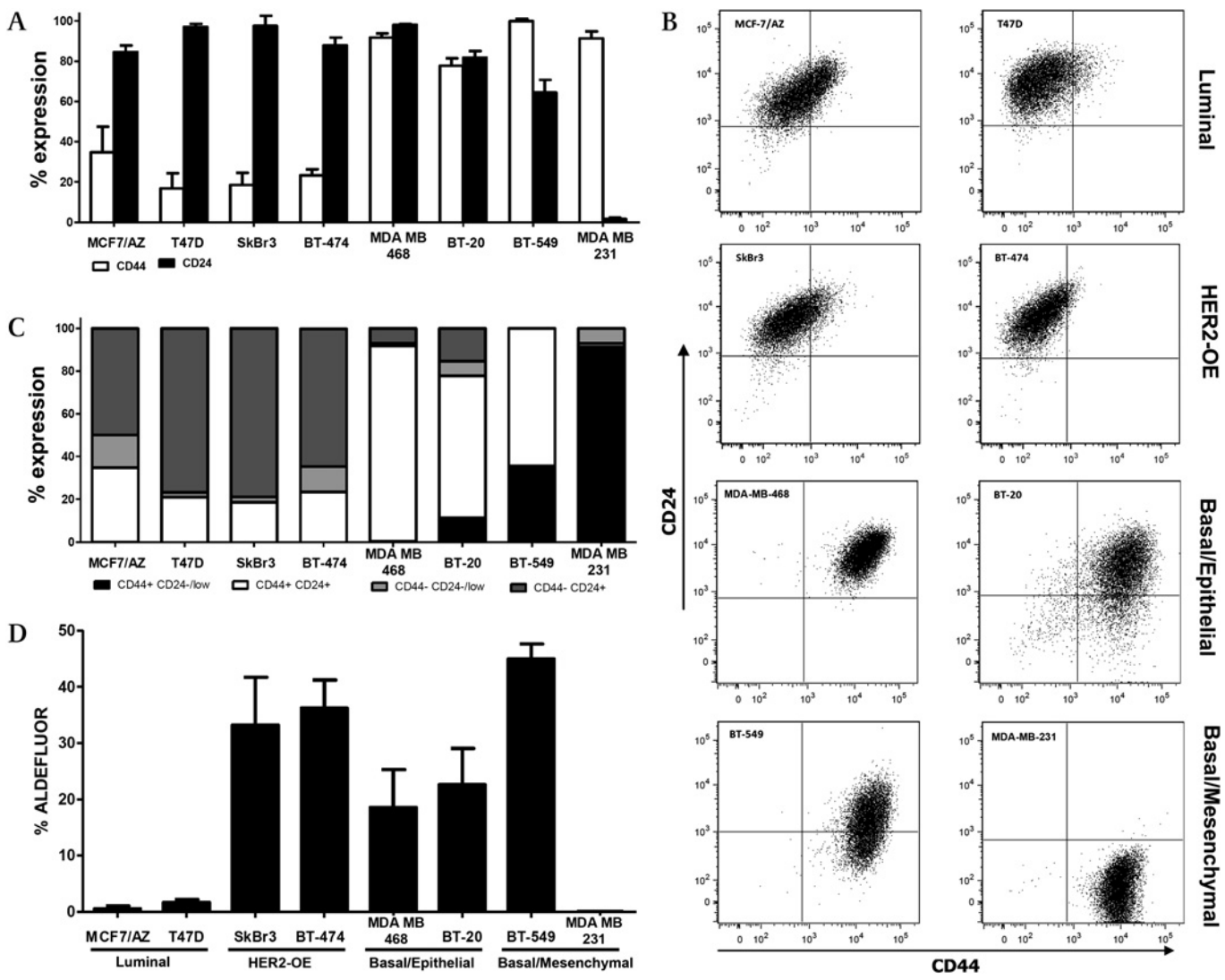


Figure 4 (A) Percentages of the subpopulations defined by the combination of the stem cell markers CD44 and CD24 in a panel of breast cancer cell lines representative of distinct molecular subtypes of the disease. Indicated is the mean±SEM of three independent experiments. (B) and (C) Subpopulations defined by expression of the stem cell markers CD44-FITC and CD24-PE in a panel of breast cancer cell lines representative of distinct molecular subtypes by flow cytometry. Isotype controls were performed (not shown). (D) Percentages of the ALDEFLUOR-positive subpopulation defined by the ALDEFLUOR assay in a panel of breast cancer cell lines representative of distinct molecular subtypes of the disease. Plotted is the mean±SEM of three independent experiments.

cells and vice versa; therefore, the main phenotype seen in basal-like tumour samples may be a consequence of CD24 loss of expression during tumour progression, whereas the cell of origin probably sustains its expression. In addition, it was also shown that distant breast cancer metastasis are enriched with luminal epithelial $CD24^+$ cells, implying a phenotypic switch or a clonal selection for cells with the $CD24^+$ phenotype.[33 56] As recent studies have described that epithelial-to-mesenchymal transition generates cells with $CD44^+CD24^{-/low}$ stem-cell like characteristics,[57] a reversion of the process seems to occur in distant metastasis.

Moreover, within basal-like tumours, a tendency towards worse patient survival (DFS and OS) was demonstrated, when carcinomas showed a predominant $CD44^+CD24^{-/low}$ CSC phenotype. Previous studies have already demonstrated an association between basal-like carcinomas and the CSC phenotype $CD44^+CD24^{-/low}$.[27–30] These results highlight the biological heterogeneity of breast cancer and an enrichment of putative tumour-initiating cells in the aggressive basal-like tumour subtype. Furthermore, it seems to reflect the fact that whenever CSC markers are present in tumours, they probably identify the tumour cell of origin more than cells harbouring a higher selective advantage for tumour progression, because highly aggressive HER2-overexpressing tumours did not show an increased expression of these markers.

The breast CSC marker ALDH1 has been described as a marker of both normal and malignant breast stem/progenitor cells.[34 36 58] $ALDH1^{hi}$ tumour cells form visibly larger colonies and mammospheres, when compared with $ALDH1^{low}$ cells.[36] Previous works also detected small percentages of $ALDH1^+$ cases in invasive breast cancer, ranging from 4% to 19%.[29 35 36 41 59] In our series, we found 7.1% of ALDH1 expression. Remarkably, the majority of cases showing a predominant ALDH1-positive population were significantly associated with basal-like tumours. Besides the low number of positive cases, ALDH1 expression was significantly associated with high histological

Take-home messages

- Basal-like breast cancer is the intrinsic molecular subtype harbouring the higher percentage of tumour cells with the CSC phenotype $CD44^{+}CD24^{-/low}$ and ALDH1 positivity.
- Luminal and HER-OE breast cancer cell lines are mainly enriched in $CD44^{-/low}CD24^{+}$ tumour cells, whereas basal/mesenchymal breast cancer cell lines are enriched in the $CD44^{+}CD24^{-/low}$ CSC phenotype; basal/epithelial breast cancer cells are mainly positive for both CD44 and CD24 CSC markers.
- ALDH1 activity was mainly found in HER-OE and basal/epithelial breast cancer cell lines.
- The described $CD44^{+}CD24^{-/low}$ and ALDH1+ CSC phenotypes seem to identify breast CSC with distinct levels of differentiation.

grade but the survival rate of ALDH1-positive cases did not significantly correlate with poor clinical outcome, as stated in previous studies.[34 30 36]

As verified in primary tumours, the measured activity of the ALDH1 enzyme was also higher in basal-like cell lines, with the exception of MDA-MB-231, which showed undetectable ALDH1 activity, as already pointed out by Deng *et al*.[36] Noteworthy was the prevalence of the $CD44^{+}CD24^{-/low}$ cell population and the absent ALDEFLUOR-positive population in these cells, indicating that these markers probably do not refer to the same cell of origin that gives rise to the CSC compartment in distinct breast tumours. Other examples are the HER2-OE SkBr3 and BT474 breast cancer cell lines, which showed predominance of the $CD44^{-}CD24^{+}$ luminal phenotype, but presented with high levels of ALDH1.

In summary, the described $CD44^{+}CD24^{-/low}$ and $ALDH1^{+}$ stem-like phenotypes seem to identify CSC with distinct levels of differentiation, the former profile being more related to basal-like carcinomas that most probably originate from the most primitive mammary stem cells, whereas the latter is a marker of basal-like and HER2-overexpressing tumours, putatively originating from luminal committed progenitors. With this hypothesis in mind, it seems that the paramount method and biomarkers that identify breast CSC within the distinct molecular subtypes need to be better explored, because it is pivotal to translate the CSC concept to clinical practice. In the near future, the recognition of reliable markers to distinguish the CSC pool in each molecular subtype will be decisive for the development of specific target therapies.

Acknowledgements The authors would like to thank Professor Raquel Seruca (coordinator from the Cancer Genetics group at IPATIMUP) for scientific assistance and for critically reading the manuscript before submission. IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and is partially supported by FCT.

Funding The work presented was mainly supported by a scientific project funded by Fundação Calouste Gulbenkian and by research grants funded by the Portuguese Science and Technology Foundation (FCT): a PhD research grant for AFV (SFRH/BD/41363/2007) and the Programa Ciência 2007 for JP.

Competing interests None to declare.

Ethics approval This study was conducted under the national regulative law for the handling of biological specimens from tumour banks, being the samples exclusively available for research purposes in retrospective studies.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Ferlay J,** Shin HR, Bray F, *et al*. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010;**127**:2893–917.
2. **Perou CM,** Sorlie T, Eisen MB, *et al*. Molecular portraits of human breast tumours. *Nature* 2000;**406**:747–52.
3. **Sorlie T,** Perou CM, Tibshirani R, *et al*. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;**98**:10869–74.
4. **Sorlie T,** Tibshirani R, Parker J, *et al*. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 2003;**100**:8418–23.
5. **Sotiriou C,** Neo SY, McShane LM, *et al*. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 2003;**100**:10393–8.
6. **Rakha E,** Reis-Filho JS. Basal-like breast carcinoma: from expression profiling to routine practice. *Arch Pathol Lab Med* 2009;**133**:860–8.
7. **Yehiely F,** Moyano JV, Evans JR, *et al*. Deconstructing the molecular portrait of basal-like breast cancer. *Trends Mol Med* 2006;**12**:537–44.
8. **Nielsen TO,** Hsu FD, Jensen K, *et al*. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* 2004;**10**:5367–74.
9. **Tang P,** Wang J, Bourne P. Molecular classifications of breast carcinoma with similar terminology and different definitions: are they the same? *Hum Pathol* 2008;**39**:506–13.
10. **Sousa B,** Paredes J, Milanezi F, *et al*. P-cadherin, vimentin and CK14 for identification of basal-like phenotype in breast carcinomas: an immunohistochemical study. *Histol Histopathol* 2010;**25**:963–74.
11. **Harris L,** Fritsche H, Mennel R, *et al*. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol* 2007;**25**:5287–312.
12. **Reya T,** Morrison SJ, Clarke MF, *et al*. Stem cells, cancer, and cancer stem cells. *Nature* 2001;**414**:105–11.
13. **Charafe-Jauffret E,** Monville F, Ginestier C, *et al*. Cancer stem cells in breast: current opinion and future challenges. *Pathobiology* 2008;**75**:75–84.
14. **Morrison BJ,** Schmidt CW, Lakhani SR, *et al*. Breast cancer stem cells: implications for therapy of breast cancer. *Breast Cancer Res* 2008;**10**:210.
15. **Chuthapisith S,** Eremin J, El-Sheemey M, *et al*. Breast cancer chemoresistance: emerging importance of cancer stem cells. *Surg Oncol* 2010;**19**:27–32.
16. **Dontu G.** Breast cancer stem cell markers—the rocky road to clinical applications. *Breast Cancer Res* 2008;**10**:110.
17. **Charafe-Jauffret E,** Ginestier C, Birnbaum D. Breast cancer stem cells: tools and models to rely on. *BMC Cancer* 2009;**9**:202.
18. **Kakarala M,** Wicha MS. Implications of the cancer stem-cell hypothesis for breast cancer prevention and therapy. *J Clin Oncol* 2008;**26**:2813–20.
19. **Phillips TM,** McBride WH, Pajonk F. The response of CD24(-/low)/CD44+ breast cancer-initiating cells to radiation. *J Natl Cancer Inst* 2006;**98**:1777–85.
20. **Li X,** Lewis MT, Huang J, *et al*. Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J Natl Cancer Inst* 2008;**100**:672–9.
21. **Al-Hajj M,** Wicha MS, Benito-Hernandez A, *et al*. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* 2003;**100**:3983–8.
22. **Kelly PN,** Dakic A, Adams JM, *et al*. Tumor growth need not be driven by rare cancer stem cells. *Science* 2007;**317**:337.
23. **Hill RP,** Perris R. "Destemming" cancer stem cells. *J Natl Cancer Inst* 2007;**99**:1435–40.
24. **Bauerschmitz GJ,** Ranki T, Kangasniemi L, *et al*. Tissue-specific promoters active in CD44+CD24-/low breast cancer cells. *Cancer Res* 2008;**68**:5533–9.
25. **Abraham BK,** Fritz P, McClellan M, *et al*. 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**:1154–9.
26. **Snyder EL,** Bailey D, Shipitsin M, *et al*. Identification of CD44v6(+)/CD24- breast carcinoma cells in primary human tumors by quantum dot-conjugated antibodies. *Lab Invest* 2009;**89**:857–66.
27. **Fillmore CM,** Kuperwasser C. Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Res* 2008;**10**:R25.
28. **Sheridan C,** Kishimoto H, Fuchs RK, *et al*. CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. *Breast Cancer Res* 2006;**8**:R59.
29. **Park SY,** Lee HE, Li H, *et al*. Heterogeneity for stem cell-related markers according to tumor subtype and histologic stage in breast cancer. *Clin Cancer Res* 2010;**16**:876–87.
30. **Honeth G,** Bendahl PO, Ringner M, *et al*. The CD44+/CD24- phenotype is enriched in basal-like breast tumors. *Breast Cancer Res* 2008;**10**:R53.
31. **Buess M,** Rajski M, Vogel-Durrer BM, *et al*. Tumor-endothelial interaction links the CD44(+)/CD24(-) phenotype with poor prognosis in early-stage breast cancer. *Neoplasia* 2009;**11**:987–1002.
32. **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**:1096–102.
33. **Shipitsin M,** Campbell LL, Argani P, *et al*. Molecular definition of breast tumor heterogeneity. *Cancer Cell* 2007;**11**:259–73.

34. **Ginestier C,** Hur 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—67.
35. **Morimoto K,** Kim SJ, Tanei T, *et al*. 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* 2009;**100**:1062—8.
36. **Deng S,** Yang X, Lassus H, *et al*. Distinct expression levels and patterns of stem cell marker, aldehyde dehydrogenase isoform 1 (ALDH1), in human epithelial cancers. *PLoS One* 2010;**5**:e10277.
37. **Croker AK,** Goodale D, Chu J, *et al*. High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability. *J Cell Mol Med* 2009;**13**:2236—52.
38. **Avninder S,** Ylaya K, Hewitt SM. Tissue microarray: a simple technology that has revolutionized research in pathology. *J Postgrad Med* 2008;**54**:158—62.
39. **Tuominen VJ,** Ruotoistenmaki S, Viitanen A, *et al*. ImmunoRatio: a publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67. *Breast Cancer Res* 2010;**12**:R56.
40. **Cheang MC,** 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—50.
41. **Resetkova E,** Reis-Filho JS, Jain RK, *et al*. Prognostic impact of ALDH1 in breast cancer: a story of stem cells and tumor microenvironment. *Breast Cancer Res Treat* 2010;**123**:97—108.
42. **Prat A,** Parker JS, Karginova O, *et al*. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* 2010;**12**:R68.
43. **Neve RM,** Chin K, Fridlyand J, *et al*. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 2006;**10**:515—27.
44. **Charafe-Jauffret E,** Ginestier C, Monville F, *et al*. Gene expression profiling of breast cell lines identifies potential new basal markers. *Oncogene* 2006;**25**:2273—84.
45. **Shackleton M,** Quintana E, Fearon ER, *et al*. Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell* 2009;**138**:822—9.
46. **Sottoriva A,** Verhoeff JJ, Borovski T, *et al*. Cancer stem cell tumor model reveals invasive morphology and increased phenotypical heterogeneity. *Cancer Res* 2010;**70**:46—56.
47. **Giatromanolaki A,** Sivridis E, Fiska A, *et al*. The CD44+/CD24- phenotype relates to 'triple-negative' state and unfavorable prognosis in breast cancer patients. *Med Oncol* 2011;**28**:745—52.
48. **Kennecke H,** Yerushalmi R, Woods R, *et al*. Metastatic behavior of breast cancer subtypes. *J Clin Oncol* 2010;**28**:3271—7.
49. **Fogel M,** Friederichs J, Zeller Y, *et al*. CD24 is a marker for human breast carcinoma. *Cancer Lett* 1999;**143**:87—94.
50. **Surowiak P,** Materna V, Gyorffy B, *et al*. Multivariate analysis of oestrogen receptor alpha, pS2, metallothionein and CD24 expression in invasive breast cancers. *Br J Cancer* 2006;**95**:339—46.
51. **Bircan S,** Kapucuoglu N, Baspinar S, *et al*. CD24 expression in ductal carcinoma in situ and invasive ductal carcinoma of breast: an immunohistochemistry-based pilot study. *Pathol Res Pract* 2006;**202**:569—76.
52. **Aulmann S,** Waldburger N, Penzel R, *et al*. Reduction of CD44(+)/CD24(-) breast cancer cells by conventional cytotoxic chemotherapy. *Hum Pathol* 2010;**41**:574—81.
53. **Prat A,** Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol* 2011;**5**:5—23.
54. **Lim E,** Vaillant F, Wu D, *et al*. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* 2009;**15**:907—13.
55. **Meyer MJ,** Fleming JM, Ali MA, *et al*. Dynamic regulation of CD24 and the invasive, CD44posCD24neg phenotype in breast cancer cell lines. *Breast Cancer Res* 2009;**11**:R82.
56. **Bloushtain-Qimron N,** Yao J, Snyder EL, *et al*. Cell type-specific DNA methylation patterns in the human breast. *Proc Natl Acad Sci U S A* 2008;**105**:14076—81.
57. **Mani SA,** Guo W, Liao MJ, *et al*. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008;**133**:704—15.
58. **Russo JE,** Hilton J. Characterization of cytosolic aldehyde dehydrogenase from cyclophosphamide resistant L1210 cells. *Cancer Res* 1988;**48**:2963—8.
59. **Tanei T,** Morimoto K, Shimazu K, *et al*. 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* 2009;**15**:4234—41.

STEM CELLS®

CANCER STEM CELLS

P-Cadherin Is Coexpressed with CD44 and CD49f and Mediates Stem Cell Properties in Basal-like Breast Cancer

ANDRÉ FILIPE VIEIRA,[a,b,c] SARA RICARDO,[a] MATTHEW PAUL ABLETT,[c] MARIA RITA DIONÍSIO,[a,d] NUNO MENDES,[a] ANDRÉ ALBERGARIA,[a] GILLIAN FARNIE,[c] RENÊ GERHARD,[a] JORGE F. CAMESELLE-TEIJEIRO,[e] RAQUEL SERUCA,[a,f] FERNANDO SCHMITT,[a,f] ROBERT B. CLARKE,[c] JOANA PAREDES[a,f]

[a]IPATIMUP—Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal; [b]ICBAS—Abel Salazar Biomedical Sciences Institute, Porto, Portugal; [c]Breast Biology Group, School of Cancer and Enabling Sciences, University of Manchester, Manchester, United Kingdom; [d]Gulbenkian Programme for Advanced Medical Education; [e]Complexo Hospitalar Universitario de Vigo (CHUVI), Vigo, Spain; [f]Faculty of Medicine of the University of Porto, Porto, Portugal

Key Words. *CDH3* gene • P-cadherin • Cancer stem cells • Biomarker • Basal-like breast cancer

ABSTRACT

Although the luminal progenitor cell of the normal mammary gland hierarchy has been proposed as the cell-of-origin for basal-like breast cancers, finding the cancer stem cell (CSC) phenotype for this malignancy has proven a difficult task, mostly due to the lack of specific markers. Recently, basal-like sporadic and familial cases of breast cancer have been linked to *BRCA1* gene inactivation, which enables the upregulation of the target-repressed *CDH3*/*P*-cadherin gene, an important biomarker of basal-like breast carcinomas. Previously, we demonstrated that P-cadherin overexpression can mediate aggressive behavior in these tumors. Thus, our aim was to test whether P-cadherin mediates stem cell properties in basal-like breast carcinomas. Using a series of breast cancer cell lines and primary tumors, we showed that P-cadherin was directly associated with the expression of the breast stem markers CD44, CD49f, and aldehyde dehydrogenase 1 in the basal subtype. Moreover, cell population enriched for P-cadherin expression comprised increased in vitro mammosphere-forming efficiency and capacity to grow colonies in three-dimensional cultures as well as greater tumorigenicity. Importantly, an association was found with stem-/progenitor-like phenotypes of the breast, including the luminal progenitor population, $CD49f^{+}CD24^{+}$. Additionally, P-cadherin expression conferred resistance to x-ray-induced cell death, sustaining a role for this molecule in another stem cell property. In summary, we demonstrated, for the first time, that P-cadherin mediates stem cell properties, which could be explored in order to better define the CSC phenotype of basal-like breast tumors and the cell-of-origin of this malignancy. STEM CELLS *2012;30:854–864*

Disclosure of potential conflicts of interest is found at the end of this article.

INTRODUCTION

Several studies have shown that solid tumors may contain a small subpopulation of cancer cells that are tumorigenic and have the ability to self-renew and generate all the diverse cancer cells present within the tumor mass. This experimental evidence supports the cancer stem cell (CSC) hypothesis which proposes that a hierarchy exists in the solid tumors comparable to the one found in normal tissue differentiation [1]. In breast, these cells are named breast CSCs. Breast CSCs share important properties with mammary stem cells, namely the ability to proliferate and resist to radiation- and chemotherapy-induced cell death, allowing them to survive and to cause tumor recurrence [2, 3]. However, the identification of breast CSCs has been a hard task due to the current technical constraints and the high intertumor and intratumor heterogeneity observed in breast cancer [4]. Most authors make use of cell surface proteins, usually adhesion-related molecules, in an attempt to define a subpopulation of cells that represents the breast CSC population. In 2003, Michael Clarke's group isolated a subset of breast cancer cells with the phenotype $ESA^{+}/CD44^{+}/CD24^{-/low}$, which were able to self-renew and were highly tumorigenic at a low cell inoculum [5]. Since then, several other phenotypes/markers to isolate breast CSCs have been described.

For the basal-like breast cancer molecular subtype, in particular, which constitute 10% of all breast cancer cases, few descriptions exist concerning the isolation of their breast CSCs.

Author contributions: A.F.V.: conception and design, collection and/or assembly of data, data analysis and interpretation, and manuscript writing; S.R.: collection and/or assembly of data and data analysis and interpretation; M.P.A., M.R.D., and N.M.: collection and/or assembly of data; A.A., G.F., R.G., R.S., and F.S.: data analysis and interpretation; J.F.C.-T.: provision of study material or patients; R.B.C.: conception and design, provision of study material or patients, and data analysis and interpretation; J.P.: conception and design, data analysis and interpretation, manuscript writing, and final approval of manuscript.

Correspondence: Joana Paredes, Ph.D., IPATIMUP—Institute of Molecular Pathology and Immunology of the University of Porto, Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal. Telephone: +351-22-5570700; Fax: +351-22-5570799; e-mail: jparedes@ipatimup.pt Received September 16, 2011; accepted for publication February 7, 2012; first published online in STEM CELLS *EXPRESS* March 2, 2012. 1066-5099/2012/$30.00/0 doi: 10.1002/stem.1075

Meyer et al. isolated $CD44^+/CD49f^{high}/CD133\text{-}2^{high}$ cancer cells from estrogen-receptor (ER)-negative patient tumors, which showed increased tumorigenic ability [6]. Wright et al. described the phenotype $CD133^+$ as able to isolate CSCs from tumors developed in *BRCA1*$^{-/-}$ mice [7]. Hwang-Verslues et al. characterized the CSC phenotype $PROCR^+/ESA^+$ for the human basal MDA-MB-231 breast cancer cell line [8].

It was recently demonstrated that basal-like breast cancers have a molecular phenotype comparable to the luminal progenitor of the normal breast [9, 10], which raised the hypothesis that markers of luminal progenitors would be good CSC markers for basal-like lesions. Additionally, it has been shown that inactivation of *BRCA1* gene in the luminal layer of the normal breast originates breast carcinomas in mice that resemble basal-like carcinomas in humans [10]. In fact, BRCA1 is a major regulator of normal luminal maturation [11] and it is essential for the repression of a panel of genes that are typically expressed in basal-like carcinomas of the breast, such as the *CDH3*/P-cadherin gene [12]. These evidences suggest that P-cadherin can be an important CSC marker for this type of tumor lesions.

P-cadherin, a classic type I adhesion molecule, is normally expressed in the myoepithelial/basal layer of the breast and is frequently overexpressed in basal-like breast carcinomas [13–15]. We have found that P-cadherin expression is linked to aggressive tumor behavior, increasing the production of metaloproteases (MMPs) by cancer cells to the extracellular matrix as well as inducing cancer cell invasion, migration, and motility, due to a mechanism involving alterations in the actin cytoskeleton and signaling through small guanosine triphosphatase (GTPase)-binding proteins [16, 17]. However, P-cadherin is also involved in homeostatic processes, such as cell differentiation, development, and embryogenesis, illustrating an indirect effect of this adhesion molecule in stem cell biology. P-cadherin deficient female mice present abnormal mammary gland morphology, showing premature differentiation of the breast and increased risk of developing preneoplastic lesions, such as alveolar hyperplasia and ductal dysplasia [18]. In fact, P-cadherin seems to be important in the maintenance of an undifferentiated state in the malignant setting, as breast tumors with P-cadherin expression show loss of cell polarity [15]. The role of P-cadherin in development and differentiation is also seen during embryonic histogenesis, since this cadherin is present in the extraembryonic ectoderm and visceral endoderm, structures originating the placenta [19]. P-cadherin also has a direct effect in normal stem cells, since it was identified as a stem cell surface marker in human embryonic stem cells [20]. Early hair progenitor cells were also isolated as P-cadherin$^+$ (and $K14^+/\alpha6$-integrin [$CD49f^+$] cells) [21]. Furthermore, the stem cell related transcription factors β-catenin, p63, and C/EBP-β were shown to induce P-cadherin promoter activation [22–24]. Interestingly, in breast, P-cadherin is found in the cap cells, characteristic stem cells that are the precursors of myoepithelial cells [25, 26], and in the myoepithelial layer [27], eventually contributing to the suprabasal stem cell niche. Herein, we used human mammary cell lines (normal and malignant) as well as a series of invasive breast carcinomas to provide evidence that P-cadherin expression is important in the CSC context not only as a biomarker that better defines the basal-like breast CSC phenotype but also as a protein with direct relevance in stem cell activity in this specific molecular subtype.

Materials and Methods

Cell Culture

Human breast cell lines T47D, MDA-MB-468, BT-20, BT-549, and MCF-10A were obtained from ATCC (American Type Culture Collection, Manassas, VA, www.atcc.org). The human breast cancer cell line MCF-7/AZ was obtained from a collection developed in the laboratory of Prof. Marc Mareel (Ghent University Hospital, Ghent, Belgium), which was genetically manipulated to overexpress P-cadherin (MCF-7/AZ.Pcad). The control cell line (MCF-7/AZ.mock) shows low P-cadherin levels, identical to the parental cell line [28]. MCF-10A cells were cultured in Dulbecco's modified Eagle's medium (DMEM): F12, supplemented with 5% heat inactivated horse serum, 10 μg/ml insulin, 0.5 μg/ml hydrocortisone, 20 ng/ml epidermal growth factor (EGF), and 100 ng/ml cholera toxin (Sigma-Aldrich, St Louis, MO, www.sigmaaldrich.com). All the other cell lines were grown in DMEM supplemented with 10% fetal bovine serum (FBS) and with 1% antibiotic solution (penicillin–streptomycin) (Invitrogen, Carlsbad, CA, www.invitrogen.com). All cell lines were routinely cultured in a humidified atmosphere with 5% CO_2 and at 37°C. Cells were used in experiments when reached 70%–80% confluence.

Fluorescence-Activated Cell Sorting Analysis and Sorting for P-Cadherin and for the Breast Stem Markers CD44, CD24, and CD49f

Cells were harvested with versene/0.48 mM EDTA (Invitrogen). Detached cells were washed with phosphate buffered saline (PBS) supplemented with 0.5% FBS and resuspended in the stain buffer (2 mM EDTA + 0.5% bovine albumin in PBS). A single-cell suspension was labeled by fluorescence-conjugated antibodies at a concentration of 1–10 in stain buffer: fluorescein isothiocyanathe (FITC)-conjugated CD44, phycoerythrin (PE)-conjugated CD24, and FITC-conjugated CD49f. These antibodies were obtained from BD Biosciences (San Diego, CA, www.bdbiosciences.com). P-cadherin monoclonal antibody APC-conjugated was obtained from R&D (Minneapolis, MN, www.rndsystems.com) and used at the same concentration as above. A live-dead stain (Invitrogen) and the primary antibodies or the respective isotype controls (BD Biosciences) were incubated at 4°C, in the dark, for 15 minutes. The labeled cells were then washed in the stain buffer and analyzed on a LSR-II (BD Biosciences). In multicolor experiments, fluorescent minus one samples were used to determine the gating strategy.

For the sorting experiments, the normal cell line MCF-10A, two luminal cell lines, and two basal cell lines were selected and subpopulations were sorted according to P-cadherin expression (highest 20% expressing cells vs. lowest 20% expressing cells). Briefly, cells were stained for P-cadherin (APC) and a live-dead dye in stain buffer, as above. Cells were then passed through a 50 μm mesh to remove clumps and resuspended in stain buffer prior to sorting. Cells were sorted using BD Influx or FACS ARIA-II (BD Biosciences) and collected into 10% Hanks buffered solution (Invitrogen). The purity of sorted population was checked and the sorted population contained 80%–95% purified cells. In addition, a further sample was also collected from cells passed through the laser under pressure, but not sorted, to act as a control for the effect of the pressure on the cells. No differences in cell behavior were detected.

Aldehyde Dehydrogenase Activity

Aldehyde dehydrogenase (ALDH) activity was assessed using the ALDEFLUOR assay (Stem Cell Technologies, Grenoble, France, www.stemcell.com). In brief, cells were incubated in ALDEFLUOR assay buffer containing ALDH substrate (1.5 μmol/l per 10^6 cells). In each experiment, a sample of cells was also incubated with 15 μmol/l of diethylaminobenzaldehyde, a specific ALDH inhibitor, to use as a negative control. P-cadherin monoclonal antibody (APC-conjugated; R&D) was incubated with the sorted ALDEFLUOR positive and negative cells in assay buffer for 20 minutes on ice. Incubation with the isotype antibody (BD Biosciences) was used as a control. Cells were washed in assay buffer and analyzed in the FACS analyzer LSR-II (BD Biosciences).

Tissue Microarray and Immunohistochemistry

A series of 466 primary invasive breast carcinomas were retrieved from the files of the Department of Pathology, Hospital Xeral-Cíes, Vigo, Spain, which were diagnosed between 1978 and 1992. Breast cancer patient follow-up information was available for 455 cases ranging from 1 to 120 months after the diagnosis. Representative tumor areas were carefully selected and at least two tissue cores (0.6 mm in diameter) were deposited into a tissue microarray. Immunohistochemistry was performed for P-cadherin (BD Biosciences) (1/50, 60 minutes, RT) and the stem cell markers CD44 (clone 156-3C11; Cell Signaling Technology, Danvers, MA, www.cellsignal.com) (1/100, 30 minutes, RT), CD24 (clone Ab2-SN3b; Neomarkers, Fremont, CA) (1/100, 60 minutes, RT), CD49f (HPA012696; Sigma-Aldrich, St. Louis, MO) (1/10, 120 minutes, RT), and ALDH1 (clone EP1933Y; Abcam, Cambridge, UK, www.abcam.com) (1/100, 60 minutes, RT).

High temperature (98°C) antigenic retrieval with Tris-EDTA (P-cadherin) or citrate buffer (CD44, CD24, CD49f, and ALDH1) was performed before primary antibody incubation. The primary antibodies were detected using a secondary antibody with horseradish peroxidase polymer (Cytomation Envision System HRP; DAKO, Carpinteria, CA, www.dako.com) using diaminobenzidine as chromogen, according to the manufacturer's instructions. This series of invasive breast carcinomas was previously characterized in terms of patient and tumor parameters (Supporting Information Table S1) [29].

P-cadherin, CD44, CD24, and CD49f staining were detected mainly at the membrane of tumor cells and the scoring was considered as follows: (0), 0%–10% of positive tumor cells; (1+), 10%–25% of positive tumor cells; (2+), 25%–50% of positive tumor cells; (3+), >50% of positive tumor cells. For P-cadherin, CD44, and CD49f, the cases which were classified as (0) were considered negative, whereas (1+), (2+), and (3+) were established as positive cases. For CD24, the cases were divided into negative/low (−/low), when considered (0) or (1+), or into positive cases when classified as (2+) or (3+). Immunohistochemical staining of ALDH1 was classified as positive when more than 1% of tumor cells showed clear cytoplasmic positivity [30, 31]. Since the immunohistochemical result was not interpretable for some of these markers, the statistical analyses were performed using only the breast tumor cases with available data. This study was conducted under the national regulative law for the usage of biological specimens from tumor banks, where the samples are exclusively available for research purposes in the case of retrospective studies.

P-Cadherin and Stem Cell Markers Knockdown and Immunoblot Analysis

P-cadherin (*CDH3* gene) expression was silenced by a specific siRNA, target sequence: AAGCCTCTTACCTGCCGTAAA. Inhibition of P-cadherin was maintained for at least 72 hours after cell transfection, confirmed by Western blot. Inhibition of the expression of CD49f (*ITGA6* gene), target sequence: CAGGGTA ATAAACTTAGGTAA, and CD44 (*CD44* gene), target sequence: AACTCCATCTGTGCAGCAAAC, was also performed.

All transfections were carried out using HiPerFect transfection reagent (Qiagen, Hilden, Germany) in a final concentration of 2 nM siRNA (Qiagen), according to manufacturer instructions. A siRNA scrambled sequence was included as a control (Qiagen).

Protein analysis was performed by Western blot using the following antibodies: anti-P-cadherin (BD Transduction, San Diego, CA, www.bdbiosciences.com), anti-CD49f (HPA012696; Sigma-Aldrich), anti-CD44 (Cell Signaling Technology, Danvers, MA), anti-CD44v6 (VFF-7; Abcam, Cambridge, U.K.), and anti-α-tubulin (DM1A; Sigma-Aldrich). Secondary antibodies were peroxidase conjugated, from Santa Cruz Biotechnology (Heidelberg, Germany, www.scbt.com) and detection was performed using the Amersham Hyperfilm and Amersham ECL Detection substrate (GE Healthcare, Chalfont St. Giles, U.K., www.gehealthcare.com).

Mammosphere Assay

Monolayer cells were enzymatically detached with 0.125% trypsin-EDTA (Sigma, St. Louis, MO), manually disaggregated with a 25-gauge needle to a single-cell suspension, and resuspended in cold PBS. Cells were plated at 500 cm^{-2} in nonadherent culture conditions, in flasks coated with 1.2% poly(2-hydroxyethylmethacrylate)/95% ethanol (Sigma). Cells were grown for 5 days, in DMEM/F12 containing B27 supplement, 500 ng/ml hydrochortisone, 40 ng/ml insulin, 20 ng/ml EGF, and maintained in a humidified incubator at 37°C and 5% (vol/vol) CO_2. Mammosphere-forming efficiency (MFE) was calculated as the number of mammospheres ($\geq$50 μm) formed, divided by the cell number plated, being expressed as a percentage.

Three-Dimensional Cultures

The three-dimensional (3D) on-top method was used. Briefly, single-cell suspensions were seeded at a density of 250 or 1,000 cells per well into eight-well glass chamber slides containing 50 μl of 100% growth factor-reduced matrigel per well (a biologically active matrix material resembling the mammalian cellular basement membrane [BD Biosciences]). Cells were plated in growth medium containing H14 medium (DMEM/F12 with insulin 250 ng/ml, transferrin 10 μg/ml, sodium selenite 2.6 ng/ml, estradiol 10^{-10} M, hydrocortisone 1.4×10^{-6} M, prolactin 5 μg/ml, and EGF 10 ng/ml, according to Kenny et al. [32] with 5% growth factor-reduced matrigel and 1% FBS).

Cells were incubated at 37°C, with replacement of the growth medium containing 2% growth factor-reduced matrigel every 2–3 days, to allow 3D structures to form. The size and number of the structures formed were assessed microscopically after 21 days.

Cell X-Ray Irradiation

Normal and cancer cells were plated in mammosphere culture conditions and immediately irradiated with 2 Gy or 4 Gy, respectively. Irradiations were performed using a 320 kV X-ray system (Gulmay Medical Ltd., Camberley, U.K.). The machine was operated at 300 kV, 10 mA, with filtration fitted in the X-ray beam to give a radiation quality of 2.3 mm Cu half-value layer. Samples were positioned at a distance of 500 mm from the X-ray focus, where the dose rate was determined to be 1.37 Gy/minute.

In Vivo Assessment of P-Cadherin Tumorigenic Capacity

The basal-like cell line MDA-MB-468 was sorted according to P-cadherin expression into two subpopulations: P-cad^{high} and P-cad^{low} fractions, as described above. The sorted cells were xenotransplanted at varying dilutions (10^6, 10^5, or 5×10^4 cells in 100 μl cell suspension) into the subcutaneous region, under the mammary fat pad of 4–5-week-old female N:NIH(s)II:nu/nu nude mice, using a 25-gauge needle. Mice were maintained and housed at IPATIMUP Animal House, sited at the Medical Faculty of the University of Porto, in a pathogen-free environment, under controlled conditions of light and humidity. Animal experiments were carried out in accordance with the European Guidelines for the Care and Use of Laboratory Animals, directive 2010/63/UE. Mice (four per group) were weighted, and tumor width and length were measured with calipers every week. Tumor volume was estimated by using the equation, $V = 0.5 \times a \times b^2$, where V is volume, a is the length of the major axis of the tumor, and b is the length of its minor axis. Mice were euthanized 3 months after tumor cell inoculation.

Statistical Analysis

Mammosphere-forming ability and growth in 3D cultures (size and number) were compared using two-tailed unpaired t test. Immunohistochemical associations between the molecular markers were assessed by Pearson's correlation and Chi-squared tests. Survival curves were estimated by the Kaplan-Meier method and compared using the log-rank test to assess significant differences

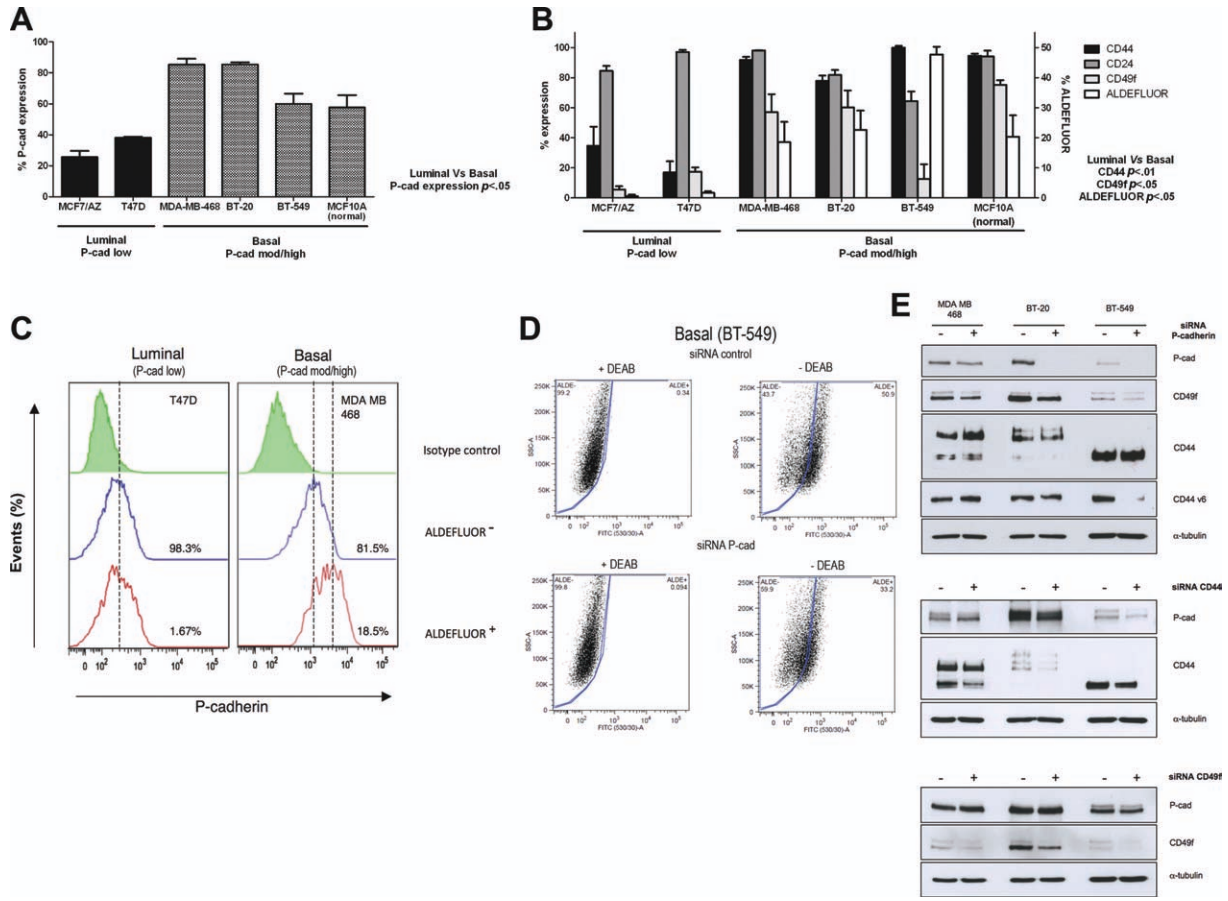


Figure 1. Fluorescence-activated cell sorting analysis of cell surface markers and ALDEFLUOR in breast cell lines. P-cadherin is highly expressed in mammary basal-like cell lines **(A)**. In these cells, expression of the stem cell markers CD44 and CD49f was found to be increased (vs. luminal cells). Furthermore, the stem cell population, defined by the ALDEFLUOR assay, is also expanded in the basal-like cells **(B)**. In addition, basal cells have an ALDEFLUOR-positive population enriched in P-cadherin expression **(C)**. P-cadherin levels were measured by flow cytometry within the ALDEFLUOR$^+$ and ALDEFLUOR$^-$ subpopulations (percentages of the ALDEFLUOR subpopulations are represented). The basal-like cell line MDA-MB-468 showed a stem cell population (ALDEFLUOR$^+$) enriched in P-cadherin expression, whereas the same does not occur in the luminal cell line T47D. **(D):** ALDEFLUOR assay was performed after P-cadherin inhibition (siRNA P-cad) in the breast basal cell lines MCF-10A, MDA-MB-468, BT-20, and BT-549 showing a decrease in the ALDEFLUOR$^+$ fraction (for which BT-549 is shown as an example). **(E):** Using transient knockdown of P-cadherin in a panel of basal-like cell lines, the expression of the stem cell markers was evaluated by Western blot. Conversely, the impact on P-cadherin expression was evaluated after transient inhibition of the stem cell markers CD44 and CD49f. Abbreviation: FITC, fluorescein isothiocyanate; DEAB, diethyl amino benzaldehyde.

for overall survival. A maximum cutoff value of 120 months (10 years) was considered. For multivariate survival analysis, Cox regression models were fitted to estimate hazard ratios (HRs) and the corresponding 95% confidence interval. Statistical analyses were carried out using SPSS statistics 17.0 software (SPSS Inc., Chicago, IL), and a significant level of 5% was considered. Flow cytometry data were analyzed with the Flowjo software package (Treestar, Ashland, OR).

Results

Normal and Tumorigenic Breast Cell Lines, with a Basal-Like Phenotype, Are Enriched for the Expression of P-Cadherin as well as for the Stem Cell Markers CD44, CD49f and Show an Increased ALDEFLUOR-Positive Subpopulation

P-cadherin is normally expressed in the basal layer of the mammary epithelium [27] and is frequently upregulated in basal-like breast carcinomas [13–15, 27, 33]. In order to study whether this protein was associated with the expression of CSC markers in vitro, a panel of breast cancer cell lines was used. This series of cell lines comprised two main subgroups: the ER-positive luminal cell lines (MCF-7/AZ and -T47D) and the ER-negative basal-like cell lines (MDA-MB-468, BT-20, and BT-549) [34, 35]. In order to demonstrate that this putative association was not restricted to cancer cells, a normal mammary cell line was also included in the study (MCF-10A), which was previously described as harboring a basal gene expression signature [34].

In accordance to what is seen in human breast carcinomas [13, 27, 33], high/moderate cell surface expression of P-cadherin (>50% of positive cells by fluorescence-activated cell sorting [FACS]) was preferentially observed in basal-like cell lines, in contrast to luminal cell lines, which show lower levels of this protein (Fig. 1A). The same association was observed for CD44 and CD49f, with increased cell surface expression of these markers in the basal-like phenotype (P-cadherin high/moderate cells), in contrast to the luminal phenotype (P-cadherin low cells) (Fig. 1B). Other studies are

in accordance with our data indicating that, in breast cell lines, luminal cells usually express lower levels of CD44 [36–38] and CD49f [35, 39] in comparison with basal cells. Importantly, the later molecules are already well established as CSC markers of the basal phenotype [6, 7] as well as putative mammary gland stem cell biomarkers [9, 40, 41]. All the cell lines expressed high levels of CD24, independently of the level of P-cadherin expression. In addition to the cell surface markers, it was previously shown that normal and cancer human mammary epithelial cells with high ALDH enzyme activity have stem/progenitor properties [31]. The analysis of ALDH activity in this panel of cell lines revealed that the P-cadherin$^{mod/high}$ basal cells showed a significant subpopulation with this putative stem cell profile (>15%), in contrast to P-cadherin^{low} luminal cells (Fig. 1B). Again, these results are in accordance with other studies that have shown that basal cells have an increased ALDEFLUOR^{+} subpopulation, compared with the luminal ones [42, 43].

FACS analysis of the ALDEFLUOR^{+} and ALDEFLUOR^{-} cell compartments of cell lines from the two main groups showed that there was an enrichment of P-cadherin expression in the ALDEFLUOR^{+} stem cell compartment, pointing for a direct association between these two markers. Interestingly, this association was only found in the basal-like group of cell lines, as shown in MDA-MB-468 (Fig. 1C). In order to more clearly demonstrate the link between P-cadherin and the ALDEFLUOR^{+} subpopulation, we decided to perform transient inhibition of P-cadherin in the basal cells MDA-MB-468, BT-20, BT-549, and MCF-10A. Measurement of ALDH1 activity showed that the inhibition of P-cadherin leads to a decrease in the number of ALDEFLUOR^{+} cells in all cell lines, being more significant in MCF-10A and BT-549. In the cell line BT-549, the ALDEFLUOR^{+} subpopulation significantly decreased from 50.9% to 33.2% (Fig. 1D). These results clearly show an association between P-cadherin and the stem cell pool.

The inhibition of P-cadherin also clearly impacted in the expression of the stem cell marker CD49f. In all basal-like cell lines analyzed, P-cadherin knockdown was accompanied by a reduction of CD49f expression (Fig. 1E). Expression of CD44 was also affected, namely the CD44v6 isoform, with a reduction found in the cell line BT-549 after P-cadherin inhibition (Fig. 1E). We also performed transient knockdown of CD44 or CD49f and studied the impact of these stem cell markers on P-cadherin expression. Inhibition of CD44 caused a slight reduction of P-cadherin expression in the BT-549 cell line and inhibition of CD49f did not affect P-cadherin expression in the three basal cell lines (Fig. 1E).

P-Cadherin Expression Is Associated with the Phenotype of the Luminal Progenitor from the Normal Breast Differentiation Hierarchy: CD49f^{+}CD24^{+}

Combinations of the markers CD44, CD24, and CD49f have been used in the literature in order to define subpopulations within cell lines or tissues that have stem or CSC properties [5, 41, 44–49]. Specifically, in the CSC field, the phenotype CD44^{+}CD24$^{-/low}$ is believed to exhibit CSC properties [5] and some authors also suggest that the CD44^{+}CD24^{+} phenotype similarly harbors stem-like properties [50, 51]. Importantly, CD44 expression in the normal human hierarchy is reported in both luminal and basal lineages by Visvader [52] and Raouf et al. [49], with a lower expression in the luminal compartment as seen in breast cell lines. Studies performed by Shipitsin et al. with cancerous and normal breast tissue indicate that cells enriched for CD44 expression represent a more basal ER-negative phenotype, with mammary epithelial progenitor-like properties [53]. Additionally, in the normal human breast, the luminal progenitor is reported to be CD49f^{+}CD24^{+} [9, 10, 41, 47] (also EpCAM^{+}MUC1^{+}CD133^{+}Thy1^{-}CD10^{-}) [46, 49, 54, 55].

In this work, cells were stained with a combination of target proteins, which include P-cadherin, CD44, CD24, and CD49f. Stringent cell sorting was applied to separate and analyze the cell subpopulations with stem-like properties described above within each of the breast cell lines, after gating for P-cadherin expression. Separation of the highest 20% P-cadherin (P-cad^{high})-expressing cell subpopulation from the lowest 20% P-cadherin (P-cad^{low})-expressing cell subpopulation revealed that, in all the cell lines studied, the P-cad^{high} cell subpopulation was enriched for the stem cell-associated markers CD44, CD49f, and CD24, in comparison with the P-cad^{low} subpopulation, as shown in Figure 2A, 2B. Conversely, when parental cells were separated into the four possible subpopulations by CD44/CD24 and CD49f/CD24 expression, the CD44^{+}CD24^{+} and CD49f^{+}CD24^{+} cell subpopulations were the ones enriched for P-cadherin expression (Fig. 2C and Supporting Information Fig. S1).

The population described by Al Hajj et al. as the CSC phenotype, CD44^{+}CD24$^{-/low}$, is decreased in the P-cad^{high} fraction (0.20% vs. 24.2% in the P-cad^{low} fraction, in the example shown in Fig. 2A). However, the enrichment of the P-cad^{high} fraction for the phenotype CD44^{+}/CD24^{+} (from 50.9% in the P-cad^{low} to 99.6% in the P-cad^{high}) indicates that this adhesion molecule is associated with stem-like properties [50, 51]. Furthermore, the P-cad^{high} cells are also enriched in CD49f^{+}/CD24^{+} cells (from 57.4% in the P-cad^{low} fraction to 99.9% in the P-cad^{high}), which are evocative of an association with the luminal progenitor profile of the breast [9, 10, 47].

In Human Breast Carcinomas, P-Cadherin Expression is Associated with Poor Patient Outcome as well as with the Expression of the Stem Cell Markers CD44, CD49f, and ALDH1

The expression of P-cadherin and the breast stem cell markers CD44, CD24, CD49f, and the isoform ALDH1 was analyzed by immunohistochemistry in a large series of 466 invasive human breast carcinomas. This analysis was performed in order to validate, in primary breast carcinomas, the association of P-cadherin expression with the stem cell markers previously observed in vitro. P-cadherin membrane expression was found in 24.5% of the cases (114/466), 63.2% of them were basal-like carcinomas (Supporting Information Table S2). CD44 membrane staining was present in 51.2% (237/463) of the cases. CD49f membrane staining was positive in only 11.5% (49/427) of the invasive breast carcinomas. Regarding the membranous stain of CD24, the majority of the cases (88.6%—410/463) were classified as negative/low, and only 11.4% (53/463) of the tumors had clear membrane staining. Concerning ALDH1 expression, a minority of cases (7.1%—33/463) were classified as positive, showing a clear cytoplasmic expression in tumor cells. Importantly, P-cadherin-positive cases were significantly enriched for the stem cell markers CD44 ($p = .003$), CD49f ($p < .001$), and ALDH1 ($p < .0001$), and no significant correlation was found between P-cadherin and CD24 expression ($p = .747$) (Table 1). Furthermore, the statistical analysis demonstrated that CD44, CD49f, and ALDH1 expressions were also significantly enriched in the basal-like carcinomas, in contrast to CD24 (Supporting Information Table S2).

Kaplan-Meier survival curves showed that P-cadherin-positive tumors were significantly associated with poor overall survival (log rank, $p = .023$) (Fig. 3A). Furthermore, when taken into account the classic prognostic markers tumor size,

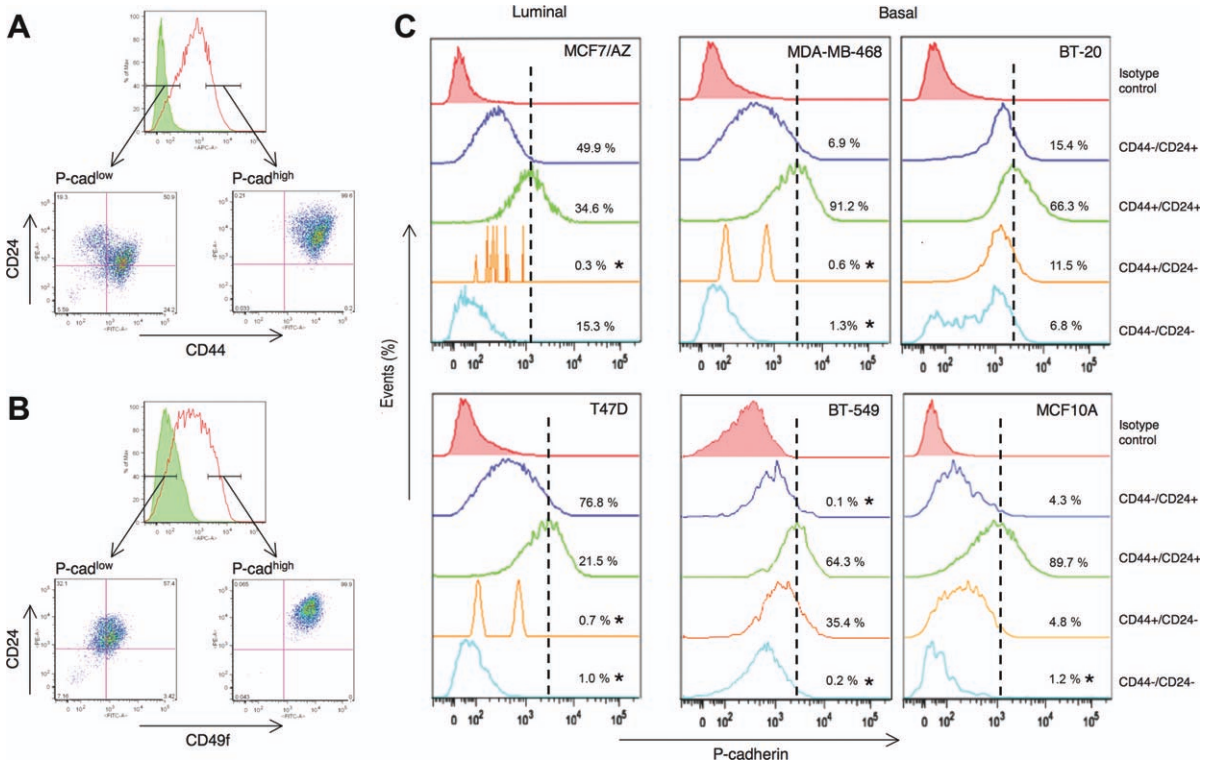


Figure 2. Fluorescence-activated cell sorting measurement of the combined expression of CD44/CD24 and CD49f/CD24 in a panel of luminal and basal breast cell lines. In all cell lines (MCF-10A is represented as an example), P-cadherin^{high} subpopulation (top 20% expressing cells) expressed higher levels of CD24, CD44, and CD49f than P-cadherin^{low} subpopulation (lowest 20% expressing cells) **(A, B)**. Conversely, when analyzing the different quadrants defined by the stem cell markers CD44/CD24, P-cadherin expression was enriched in the CD44^{+}/CD24^{+} subpopulation **(C)**. * indicates $\leq$1.5% of the total cell population.

histological grade, and node involvement, the expression of P-cadherin appeared as an independent factor, demonstrating the importance of this protein as a poor prognostic marker in breast cancer (HR = 1.486, p = .037, by multivariate Cox analysis) (Supporting Information Table S3). In contrast, the expression of the stem cell markers CD44, CD24, CD49f, or ALDH1 alone was not significantly associated with the clinical outcome (Supporting Information Fig. S2). By multivariate Cox analysis, the expression of CD24 was the only significant stem cell marker influencing survival, in this case related to a better prognosis (HR = 0.439, p = .014) (Supporting Information Table S3). Interestingly, however, when P-cadherin and CD24 were combined, the positive expression of both markers was highly associated with the worst patient overall survival (log rank, p = .014) (Fig. 3C), whereas tumors with a P-cadherin^{-}CD24^{+} phenotype were associated with a good prognosis. The tumors with combined expression of P-cadherin and CD44 were also associated with a worst patient

Table 1. Association between P-cadherin expression and the expression of the stem cell markers CD44, CD24, CD49f, and ALDH1 analyzed by immunohistochemistry in a series of 466 primary invasive breast carcinomas

Variable (n = 466)	P-cadherin+ (n = 114, 100%)	P-cadherin− (n = 352, 100%)	p value (Pearson chi-square)
CD44+ (n = 237)	72 (63.2%)	165 (47.3%)	***.003***
CD44− (n = 226)	42 (36.8%)	184 (52.7%)	
Missing (n = 3)			
CD24+ (n = 53)	14 (12.3%)	39 (11.2%)	*.747*
CD24$^{-/low}$ (n = 410)	100 (87.7%)	310 (88.8%)	
Missing (n = 3)			
CD49f+ (n = 49)	29 (27.4%)	20 (6.2%)	***<.001***
CD49f− (n = 378)	77 (72.6%)	301 (93.8%)	
Missing (n = 39)			
ALDH1+ (n = 33)	17 (14.9%)	16 (4.6%)	***<.001***
ALDH1− (n = 430)	97 (85.1%)	333 (95.4%)	
Missing (n = 3)			

P-cadherin-positive cases are enriched in CD44, CD49f, and ALDH1 expression (vs. P-cadherin-negative cases). No statistically significant association was seen between P-cadherin and CD24 expression.
Abbreviation: ALDH1, aldehyde dehydrogenase 1.
$p < .05$ was considered significant.

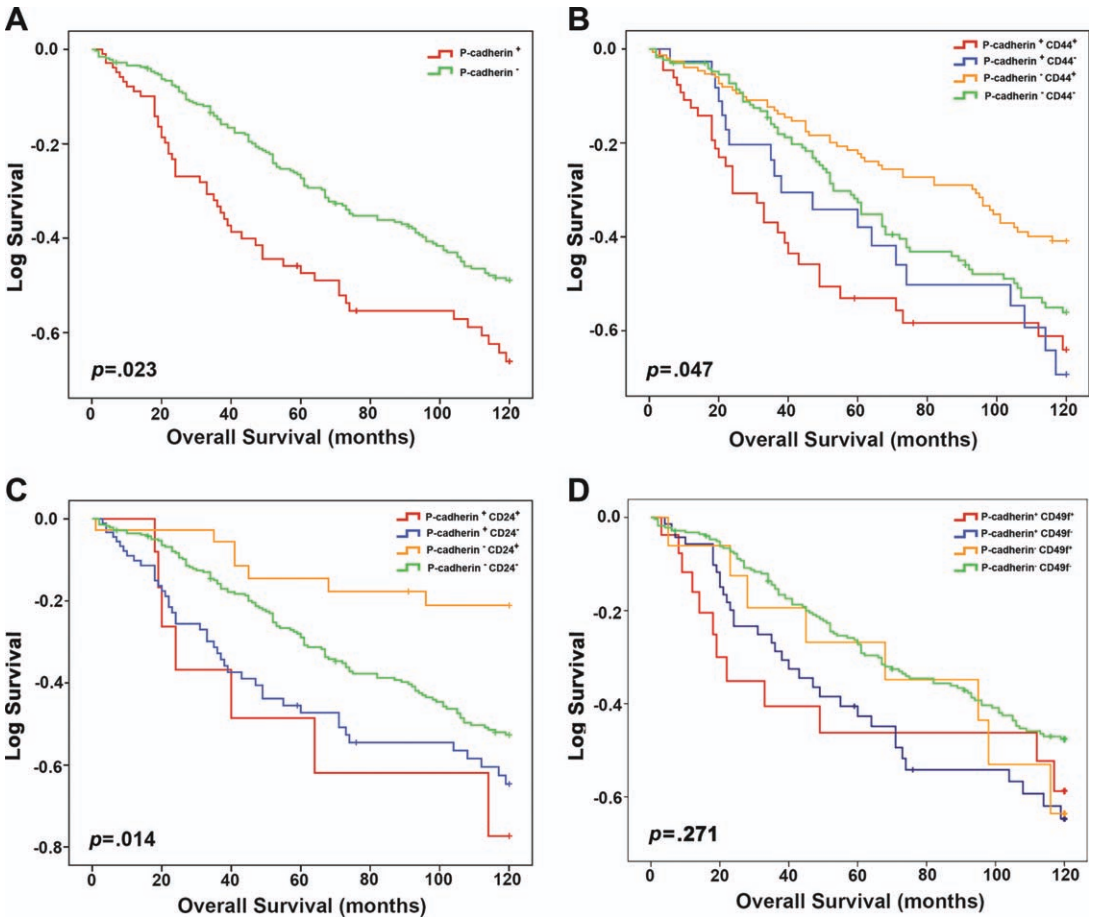


Figure 3. Overall survival of patients with breast carcinomas (log-rank test), classified according to the expression of P-cadherin **(A)** and the combined expression of P-cadherin/CD44 **(B)**, P-cadherin/CD24 **(C)**, and P-cadherin/CD49f **(D)**. Expression of CD44, CD24, or CD49f alone has no prognostic value (Supporting Information Fig. S2); however, when P-cadherin is combined with these markers, P-cadherin^{+}/CD44^{+} and P-cadherin^{+}/CD24^{+} cases have a worst overall patient survival.

overall survival (Fig. 3B) (log rank, $p = .047$). Although a strong association between P-cadherin and CD49f expression was found, the combined expression of both these markers was not significantly associated with survival (log rank, $p = .271$) (Fig. 3D).

P-Cadherin Expression Confers Increased Self-Renewal Ability, Improved Cell Growth in 3D Cultures, and Radioresistance in Breast Cancer Cell Lines

The previous in vitro and in vivo indications, showing that P-cadherin is associated with the expression of stem cell makers in breast cancer, led us to study the cellular effects mediated by the expression of P-cadherin, namely in stem cell activity.

Stem cell activity can be measured by means of the mammosphere-forming ability of a specific cell line/subpopulation. In different cell lines, the heterogeneous parental cell population were separated by sorting according to P-cadherin levels of expression (high 20% vs. low 20%) and they were studied for stem cell activity. We found that P-cad^{high} cell fractions from basal-like cell lines showed a significant increased MFE when compared with the P-cad^{low} cell fraction (Fig. 4A). The MFE from the basal P-cad^{high} fractions reached up to two times the levels of the negative fractions. This result was seen both in the basal tumorigenic cells as well as in the normal MCF-10A breast cell line, suggesting that P-cadherin contributes to the stem cell activity in both normal and malignant contexts.

We also studied the clonogenic capacity of P-cadherin positivity in a 3D proliferation permissive environment containing matrigel, a matrix resembling the basal lamina of the normal breast. Using the same sorting procedure, in the luminal MCF-7/AZ and the basal BT-549 cells, we found that there was an increase in the number of 3D structures formed by the P-cad^{high} cell subpopulation, compared with the P-cad^{low} subpopulation in the basal-like cell line (Fig. 4B). Furthermore, we found that the grape-like structures formed by P-cad^{high} cells were bigger than the 3D structures formed by the P-cad^{low} subpopulation (95.8 μm vs. 50.1 μm, $p = .003$, data not shown). The same does not hold true for luminal cells for both parameters analyzed.

To verify the specific role of P-cadherin in stem cell activity, we used genetically manipulated cell lines, by in vitro transduction or silencing of P-cadherin. We showed that stable transduction of P-cadherin in MCF-7/AZ cells led to an increase of mammosphere formation (>25% of MFE) when compared with the mock cell line (Fig. 5A, 5B). This difference in mammosphere formation persists for at least two passages (data not shown), suggesting a role of P-cadherin in mediating self-renewal or survival. Likewise, transient inhibition of P-cadherin expression by siRNA assays showed that mammosphere-forming ability was negatively affected, although not completely abrogated in most of the cell lines

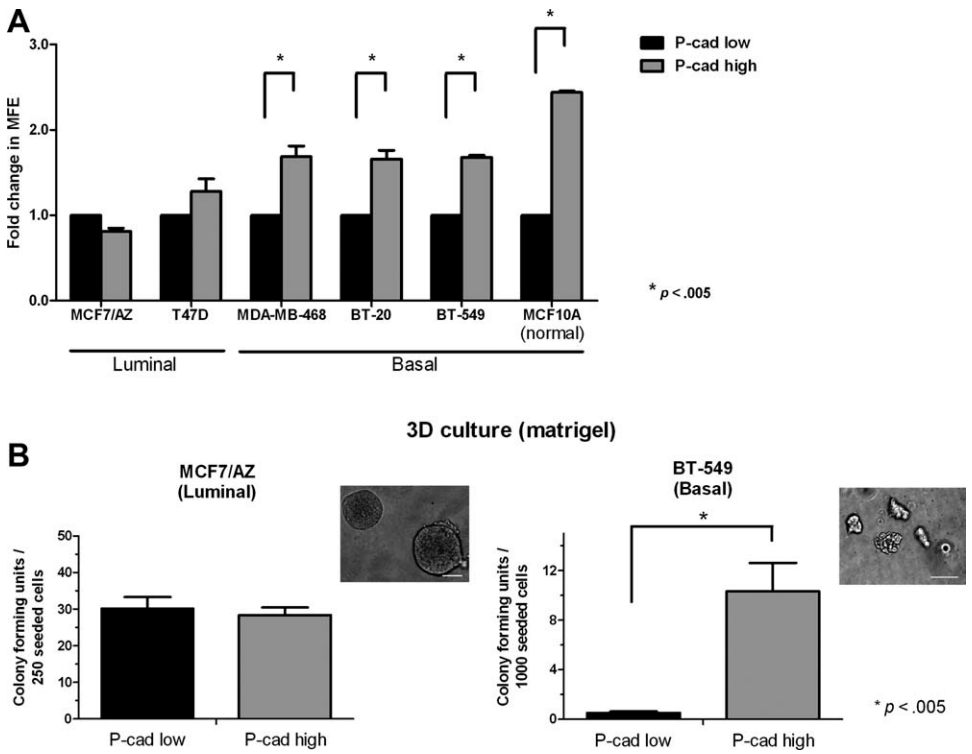


Figure 4. MFE was measured in the subpopulations isolated by fluorescence-activated cell sorting according to P-cadherin expression in a panel of human breast cell lines. The isolated subpopulations with higher amounts of P-cadherin (top 20% expressing cells), within the basal-like cell lines, displayed highest MFE **(A)**; 3D cell cultures, in laminin-rich matrix (matrigel), revealed that the P-cadherin^{high} subpopulation from the basal BT-549 cell line has increased clonogenic capacity, whereas the same result is not obtained with luminal cells (MCF-7/AZ) (scale bar = 100 μm) **(B)**. Abbreviations: 3D, three dimension; MFE, mammosphere-forming efficiency.

studied (Fig. 5A, 5C). These results indicate that P-cadherin contributes to the stem cell activity of these cells.

One of the features attributed to breast CSCs is the increased resistance to irradiation, which allows them to survive and persist in tissues after treatment, contributing to disease relapse. We investigated the effect of x-ray irradiation alone and in combination with P-cadherin inhibition on the stem cell activity measured by MFE. We found that inhibition of P-cadherin in the tumorigenic basal-like cell lines potentiated the effect of irradiation-induced cell death, decreasing the number of mammospheres formed. This same effect was not observed in the normal breast cell line MCF-10A, where irradiation treatment or P-cadherin silencing alone had a negative effect in MFE; however, when the two treatments were combined, no additive or synergistic effect was observed concerning stem cell activity (Fig. 5D).

P-Cadherin Expression Confers Increased Tumorigenic Ability to Breast Cancer Cells

The indications given by in vitro functional assays, showing that P-cadherin expression has a role in the maintenance of stem cell properties and poor overall survival of breast cancer patients, led us to test whether cancer cell population enriched for P-cadherin have a higher capacity to promote tumor growth in nude mice. The tumorigenicity of the P-cadherin high and low subpopulations was addressed using the basal-like MDA-MB-468 breast cancer cells, which were inoculated in different dilutions, and the results obtained are shown in Table 2. Interestingly, we could observe that cancer cells enriched for P-cadherin expression have a higher capacity to promote tumor growth, since the number of tumors formed was higher in the P-cad^{high} group, compared with the control group, in which the same number of parental cells was injected. Furthermore, tumors derived from the P-cad^{high} cells were larger in size than the tumors derived from the P-cad^{low} fraction.

Discussion

Increasing evidence supports that cancers are propagated by a small cell population, the CSCs, which originate both tumorigenic and nontumorigenic cells and are responsible for tumor heterogeneity, therapy resistance, and disease recurrence. The identification and analysis of CSCs are mandatory in carcinomas with high patient mortality rate, early relapses, and lack of a targeted therapy. One of such poor-prognosis cancer is the basal-like subtype of breast cancer.

Basal-like breast carcinomas include tumors that are characterized by an expression signature similar to that of the basal cells of the breast [56]. However, features of the luminal gene expression profile are often also attributed to these cancers [57]; therefore, basal-like breast carcinomas are often considered to have a mixed phenotype. More recently, the cell-of-origin for basal-like breast cancers was defined as the luminal progenitor cell from the normal breast hierarchy [9, 10], but the phenotype for the CSC has proven a tough task.

Thus, within this work, we established a positive association between P-cadherin, a basal marker, and the stem cell markers CD44, CD24, and CD49f in human breast cell lines. Indeed, we found that all the stem cell markers analyzed segregated with P-cadherin in the same subset of cells, suggesting that this adhesion molecule could be contributing to a stem-like aggressive phenotype in the breast. Accordingly, the expression of all these markers has been previously linked to

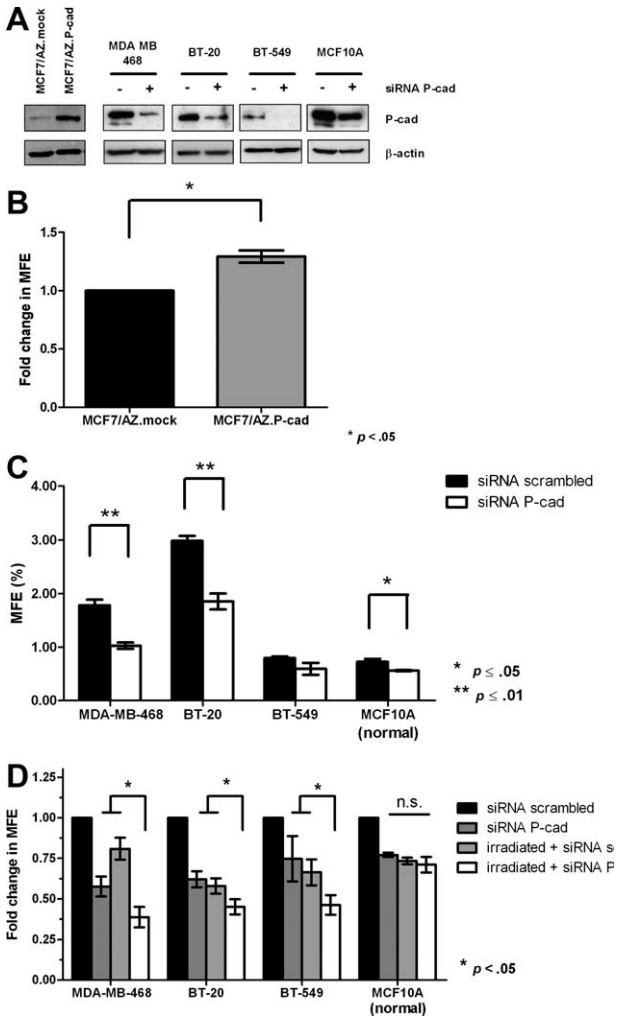


Figure 5. P-cadherin cDNA was stably transduced and overexpressed in the luminal cell line MCF-7/AZ (MCF-7/AZ.P-cad) or transiently silenced in the several basal cells lines (**A**). Measurement of stem cell activity by MFE indicates a direct association between P-cadherin expression and this stem cell property (**B, C**); x-ray irradiation was administered to the normal cell line MCF-10A (2 Gy) or the indicated tumorigenic cells (4 Gy) and MFE was assessed. The decrease in the number of stem cells mediated by x-ray irradiation is potentiated after P-cadherin silencing in tumorigenic cells but not in the normal MCF-10A cell line (**D**). Abbreviation: MFE, mammosphere-forming efficiency.

the acquisition of a malignant phenotype or to a decreased breast cancer patient survival [44, 58, 59]. Significantly, our study shows that breast cancer cases that express P-cadherin and CD44, or P-cadherin and CD24, have the worst patient prognosis. Additionally, P-cadherin could be an important regulator of these stem cell markers in tumor cells, since when this adhesion molecule was knocked down, a reduction was found in the expression of CD49f and CD44v6.

Based on these results, we believe that we got experimental evidence that P-cadherin can be classified as a stem cell marker in breast cancer. Although the simultaneous expression of all the three stem markers (CD44, CD24, and CD49f) in the same cell was never clearly described, studies with CD49f and CD24 have shown that these markers combined in the same cell are able to identify cells with progenitor-like properties in the mouse [10, 41] and human [9, 47] mammary epithelial tissue. Specifically, Keller et al. showed that the normal human breast tissue contains an epithelial subtype with the phenotype $EpCAM^{high}CD49f^{+}CD24^{+}$, consistent with a definition of luminal progenitor cells, according to the characterization of lineage markers [47]. Additionally, CD44 has also been used in combination with CD24 to show that the phenotype $CD44^{+}CD24^{+}$ identifies cells with stem-like properties, including tumorigenicity, in breast cell lines [50, 51] and in ER-negative human breast tumors [6]. In this last study, Meyer et al. also showed that $CD44^{pos}CD49f^{high}$ cells are enriched for tumor-initiating ability in ER-negative breast cancers [6].

The association between the expression of P-cadherin and the luminal progenitor phenotype $CD49f^{+}CD24^{+}$ in several cancer and normal cell lines supports the notion that P-cadherin could potentially be used together with other markers in the identification of the luminal progenitor of the normal breast. In fact, luminal progenitor cells have been described as the cell-of-origin for basal-like cancers [9, 10], in which *BRCA1* inactivation is a common feature and is considered to halt the maturation toward the luminal phenotype [11]. This would explain the upregulation of P-cadherin found in basal-like cancers, since BRCA1 is a P-cadherin transcriptional repressor [12]. Notably, Sarrio et al. recently reported that, in a normal nontumorigenic human breast cell line with a basal-like profile, the subpopulation defined as $EpCAM^{+}CD49f^{+}$ exhibits progenitor-like properties and, interestingly, this subpopulation of cells specifically expressed P-cadherin [59].

An enrichment of P-cadherin expression in cells that have the stem/progenitor phenotype $CD44^{+}CD24^{+}$ was also found in this study. Importantly, cells with the phenotype $CD44^{+}CD24^{+}$ have been described as having tumorigenic ability and represent a dynamic population that can originate the CSC phenotype $CD44^{+}CD24^{-/low}$ [50, 51, 60]. In fact, Pece et al. showed that cells with CSC activity are $CD24^{+}$ [60] and that the phenotype $CD44^{+}CD24^{+}$, isolated from several human breast cancer cell lines, shows tumorigenic activity in murine xenograft models [50]. This indicates that targeting cells with the phenotype $CD44^{+}CD24^{+}$, which have increased P-cadherin expression, could help to eradicate CSCs.

Our data still show that P-cadherin is not a simple bystander present in cells harboring a stem or progenitor phenotype, but it can also be a regulator of the stem cell function. First, we showed that high P-cadherin expression is associated with an increased capacity of cancer cells to promote tumor growth in nude mice. Second, we found that $P\text{-cad}^{high}$ cell lines have increased ALDH activity and that P-cadherin expression is increased in the $ALDEFLUOR^{+}$ subpopulation. Previous studies have shown that ALDH activity and expression is linked to increased stem cell activity, tumorigenicity, and poor prognosis in breast carcinomas [31, 61]. Furthermore, P-cadherin gene expression manipulation (either by silencing with siRNA or by P-cadherin/*CDH3* overexpression), as well as separation of cell population by sorting,

Table 2. P-cadherin capacity to promote tumors was evaluated by xenografting $P\text{-cad}^{high}$ and $P\text{-cad}^{low}$ cells sorted from the basal-like MDA-MB-468 breast cancer cell line in the nude mouse model

	Tumors/injection		
Cell population	1×10^6	1×10^5	5×10^4
Unsorted	4/4	2/4	2/4
$P\text{-cad}^{high}$		2/3	4/4
$P\text{-cad}^{low}$		0/3	2/4

The appearance of tumors as well as the tumor mass volume was measured during time.

showed that this adhesion molecule mediates stem cell activity and the self-renewal of mammospheres in basal breast cell lines. The growth in 3D matrigel matrix was also increased in the P-cad^{high} subpopulation relative to the P-cad^{low} subpopulation of basal-like breast cancer cells. It is still not clear whether the increase in clonogenic capacity in 3D cultures mediated by P-cadherin, and the increase in stem cell activity measured by MFE, is the result of increased survival or alterations in the growth rate that affects proliferation/quiescence of the prospective P-cad^{high} stem cells. Analysis of the cell cycle revealed that P-cadherin does not affect the subpopulations S and G_2M (Supporting Information Fig. S4). Importantly, the fact that cells with lower expression of P-cadherin still show some tumorigenic ability, as well as a certain level of growth in anchorage-independent conditions and in matrigel 3D cultures, indicates that stem cell properties are still present in the P-cadherin-depleted fraction. Although P-cadherin does not seem to be an exclusive marker of CSCs, our results show that a further enrichment of the stem cell population could potentially be achieved by a combination of P-cadherin with other biomarkers. For basal-like carcinomas, several studies have used combinations of the following biomarkers to define the CSC subpopulation: ESA, CD44, CD24, CD49f, CD133-2, PROCR, and ALDEFLUOR [5–8, 61]. Not surprisingly, expression of P-cadherin alone in invasive breast cancer confers bad prognosis, and we observed that this marker is associated with the stem cell markers CD44, CD49f, and ALDH1, and all these markers with the basal-like molecular subtype. The isolation of P-cadherin-positive subpopulations within the CD44$^+$ or within CD24$^+$ fractions is of potential interest, since in our series of human breast carcinomas, the tumors that exhibit P-cadherin$^+$CD44$^+$ or P-cadherin$^+$CD24$^+$ phenotype seem to be particularly aggressive. Interestingly, the phenotype P-cadherin$^+$ CD44$^+$CD49f$^+$CD24$^+$ could also support the isolation of progenitor cells from the normal breast, as mentioned above.

Additionally, it is known that basal-like breast cancers are particularly aggressive because they resist to current therapeutic strategies, usually recurring in a short time frame. In particular, resistance to radiotherapy has been reported for breast cancer, and relapses often occur. Although poorly understood, this type of resistance could be mediated by several mechanisms controlling cell survival and DNA repair mechanisms, allowing CSCs to survive and to give rise to a new tumor [3]. In this study, we show that, when given x-ray irradiation, the silencing of P-cadherin contributes to decreased survival of the stem cell population in the tumorigenic cells but not in the normal. Hence, inhibition of P-cadherin could be an approach to increase sensitization of tumorigenic cells to radiotherapy, potentially allowing the reduction of the current doses of radiation administered to the malignant tissue and causing less damage to the normal cells.

Conclusions

We found, for the first time, that P-cadherin confers stem cell features to breast tumorigenic cells that could be linked to the aggressive behavior of basal-like breast cancers. We show that this molecule is associated with increased stem cell activity (tumorigenicity in athymic nude mice, mammosphere formation, and growth in 3D matrix) in basal-like cell lines but not in luminal cells. P-cadherin is associated with already described stem cell markers that define the luminal progenitor phenotype and which, in our series of primary invasive breast cancers, is associated with the shortest overall patient survival. We also found that inhibition of P-cadherin sensitizes cancer cells to x-ray-induced cell death.

In the future, CSC therapies for the aggressive basal-like breast carcinomas can eventually involve the targeting of P-cadherin cell surface protein. In fact, phase-I clinical trials are underway, which will help to define whether P-cadherin constitutes a good therapeutic target in breast cancer [62]. Importantly, our results show that anti-P-cadherin treatment could improve radiation therapy in patients.

Acknowledgments

This work was supported by a grant (Project 96633-2010) from Fundação Gulbenkian, Portugal. We acknowledge Fundação Luso Britânica and Fundação para a Ciência e Tecnologia (FCT), Portugal, for fellowship support (A.V. was supported by SFRH/BD/41363/2007) and for the Programa Ciência 2007 (J.P.) (POPH—QREN—Tipology 4.2). We thank Dina Leitão and Regina Pinto for preparing the TMAs. The Programme for Advanced Medical Education is sponsored by Fundação Gulbenkian, Fundação Champalimaud, Portuguese Ministry of Science, Technology and Higher Education and FCT, Portugal. IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and is partially supported by FCT.

Disclosure of Potential Conflicts of Interest

The authors indicate no potential conflicts of interest.

References

1 Reya T, Morrison SJ, Clarke MF et al. Stem cells, cancer, and cancer stem cells. Nature 2001;414:105–111.
2 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.
3 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.
4 Charafe-Jauffret E, Ginestier C, Birnbaum D. Breast cancer stem cells: Tools and models to rely on. BMC Cancer 2009;9: 202.
5 Al-Hajj M, Wicha MS, Benito-Hernandez A et al. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci USA 2003;100:3983–3988.
6 Meyer MJ, Fleming JM, Lin AF et al. CD44posCD49fhiCD133/2hi defines xenograft-initiating cells in estrogen receptor-negative breast cancer. Cancer Res 2010;70:4624–4633.
7 Wright MH, Calcagno AM, Salcido CD et al. Brca1 breast tumors contain distinct CD44+/CD24− and CD133+ cells with cancer stem cell characteristics. Breast Cancer Res 2008;10:R10.
8 Hwang-Verslues WW, Kuo WH, Chang PH et al. Multiple lineages of human breast cancer stem/progenitor cells identified by profiling with stem cell markers. PLoS One 2009;4:e8377.
9 Lim E, Vaillant F, Wu D et al. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat Med 2009;15:907–913.
10 Molyneux G, Geyer FC, Magnay FA et al. BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. Cell Stem Cell 2010;7:403–417.
11 Liu S, Ginestier C, Charafe-Jauffret E et al. BRCA1 regulates human mammary stem/progenitor cell fate. Proc Natl Acad Sci USA 2008; 105:1680–1685.

12 Gorski JJ, James CR, Quinn JE et al. BRCA1 transcriptionally regulates genes associated with the basal-like phenotype in breast cancer. Breast Cancer Res Treat. 2010;122:721–731.
13 Matos I, Dufloth R, Alvarenga M et al. p63, cytokeratin 5, and P-cadherin: Three molecular markers to distinguish basal phenotype in breast carcinomas. Virchows Arch 2005;447:688–694.
14 Paredes J, Albergaria A, Oliveira JT et al. P-cadherin overexpression is an indicator of clinical outcome in invasive breast carcinomas and is associated with CDH3 promoter hypomethylation. Clin Cancer Res 2005;11:5869–5877.
15 Paredes J, Milanezi F, Viegas L et al. P-cadherin expression is associated with high-grade ductal carcinoma in situ of the breast. Virchows Arch 2002;440:16–21.
16 Ribeiro AS, Albergaria A, Sousa B et al. Extracellular cleavage and shedding of P-cadherin: A mechanism underlying the invasive behaviour of breast cancer cells. Oncogene 2010;29:392–402.
17 Albergaria A, Ribeiro AS, Vieira AF et al. P-cadherin role in normal breast development and cancer. Int J Dev Biol 2011;55:811–822.
18 Radice GL, Ferreira-Cornwell MC, Robinson SD et al. Precocious mammary gland development in P-cadherin-deficient mice. J Cell Biol 1997;139:1025–1032.
19 Hirai Y, Nose A, Kobayashi S et al. Expression and role of E- and P-cadherin adhesion molecules in embryonic histogenesis. I. Lung epithelial morphogenesis. Development 1989;105:263–270.
20 Kolle G, Ho M, Zhou Q et al. Identification of human embryonic stem cell surface markers by combined membrane-polysome translation state array analysis and immunotranscriptional profiling. Stem Cells 2009;27:2446–2456.
21 Rhee H, Polak L, Fuchs E. Lhx2 maintains stem cell character in hair follicles. Science 2006;312:1946–1949.
22 Faraldo MM, Teuliere J, Deugnier MA et al. beta-Catenin regulates P-cadherin expression in mammary basal epithelial cells. FEBS Lett 2007;581:831–836.
23 Shimomura Y, Wajid M, Shapiro L et al. P-cadherin is a p63 target gene with a crucial role in the developing human limb bud and hair follicle. Development 2008;135:743–753.
24 Albergaria A, Ribeiro AS, Pinho S et al. ICI 182,780 induces P-cadherin overexpression in breast cancer cells through chromatin remodelling at the promoter level: A role for C/EBPbeta in CDH3 gene activation. Hum Mol Genet 2010;19:2554–2566.
25 Daniel CW, Strickland P, Friedmann Y. Expression and functional role of E- and P-cadherins in mouse mammary ductal morphogenesis and growth. Dev Biol 1995;169:511–519.
26 Knudsen KA, Wheelock MJ. Cadherins and the mammary gland. J Cell Biochem 2005;95:488–496.
27 Paredes J, Milanezi F, Reis-Filho JS et al. Aberrant P-cadherin expression: Is it associated with estrogen-independent growth in breast cancer? Pathol Res Pract 2002;198:795–801.
28 Paredes J, Stove C, Stove V et al. P-cadherin is up-regulated by the antiestrogen ICI 182,780 and promotes invasion of human breast cancer cells. Cancer Res 2004;64:8309–8317.
29 Ricardo S, Vieira AF, Gerhard R et al. Breast cancer stem cell markers CD44, CD24 and ALDH1: Expression distribution within intrinsic molecular subtype. J Clin Pathol 2011;64:937–946.
30 Deng S, Yang X, Lassus H et al. Distinct expression levels and patterns of stem cell marker, aldehyde dehydrogenase isoform 1 (ALDH1), in human epithelial cancers. PLoS One 2010;5:e10277.
31 Ginestier C, Hur 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.
32 Kenny PA, Lee GY, Myers CA et al. The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression. Mol Oncol 2007;1:84–96.
33 Paredes J, Lopes N, Milanezi F et al. P-cadherin and cytokeratin 5: Useful adjunct markers to distinguish basal-like ductal carcinomas in situ. Virchows Arch 2007;450:73–80.
34 Charafe-Jauffret E, Ginestier C, Monville F et al. Gene expression profiling of breast cell lines identifies potential new basal markers. Oncogene 2006;25:2273–2284.
35 Neve RM, Chin K, Fridlyand J et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 2006;10:515–527.
36 Fillmore CM, Kuperwasser C. Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. Breast Cancer Res 2008;10:R25.
37 Olsson E, Honeth G, Bendahl PO et al. CD44 isoforms are heterogeneously expressed in breast cancer and correlate with tumor subtypes and cancer stem cell markers. BMC Cancer 2011;11:418.
38 Sheridan C, Kishimoto H, Fuchs RK et al. CD44+/CD24− breast cancer cells exhibit enhanced invasive properties: An early step necessary for metastasis. Breast Cancer Res 2006;8:R59.
39 Yoon CS, Kim KD, Park SN et al. Alpha(6) integrin is the main receptor of human papillomavirus type 16 VLP. Biochem Biophys Res Commun 2001;283:668–673.
40 Shackleton M, Vaillant F, Simpson KJ et al. Generation of a functional mammary gland from a single stem cell. Nature 2006;439:84–88.
41 Stingl J, Eirew P, Ricketson I et al. Purification and unique properties of mammary epithelial stem cells. Nature 2006;439:993–997.
42 Burga LN, Hu H, Juvekar A et al. Loss of BRCA1 leads to an increase in epidermal growth factor receptor expression in mammary epithelial cells, and epidermal growth factor receptor inhibition prevents estrogen receptor-negative cancers in BRCA1-mutant mice. Breast Cancer Res 2011;13:R30.
43 Marcato P, Dean CA, Pan D et al. Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. Stem Cells 2011;29:32–45.
44 Cariati M, Naderi A, Brown JP et al. Alpha-6 integrin is necessary for the tumourigenicity of a stem cell-like subpopulation within the MCF7 breast cancer cell line. Int J Cancer 2008;122:298–304.
45 Vassilopoulos A, Wang RH, Petrovas C et al. Identification and characterization of cancer initiating cells from BRCA1 related mammary tumors using markers for normal mammary stem cells. Int J Biol Sci 2008;4:133–142.
46 Keller PJ, Arendt LM, Skibinski A et al. Defining the cellular precursors to human breast cancer. Proc Natl Acad Sci USA. 2012;109:2772–2777.
47 Keller PJ, Lin AF, Arendt LM et al. Mapping the cellular and molecular heterogeneity of normal and malignant breast tissues and cultured cell lines. Breast Cancer Res 2010;12:R87.
48 Mani SA, Guo W, Liao MJ et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 2008;133: 704–715.
49 Raouf A, Zhao Y, To K et al. Transcriptome analysis of the normal human mammary cell commitment and differentiation process. Cell Stem Cell 2008;3:109–118.
50 Meyer MJ, Fleming JM, Ali MA et al. Dynamic regulation of CD24 and the invasive, CD44posCD24neg phenotype in breast cancer cell lines. Breast Cancer Res 2009;11:R82.
51 Rappa G, Lorico A. Phenotypic characterization of mammosphere-forming cells from the human MA-11 breast carcinoma cell line. Exp Cell Res 2010;316:1576–1586.
52 Visvader JE. Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. Genes Dev 2009;23:2563–2577.
53 Shipitsin M, Campbell LL, Argani P et al. Molecular definition of breast tumor heterogeneity. Cancer Cell 2007;11:259–273.
54 Eirew P, Stingl J, Raouf A et al. A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. Nat Med 2008;14:1384–1389.
55 Stingl J, Eaves CJ, Zandieh I et al. Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. Breast Cancer Res Treat 2001;67:93–109.
56 Rakha E, Reis-Filho JS. Basal-like breast carcinoma: From expression profiling to routine practice. Arch Pathol Lab Med 2009;133:860–868.
57 Gusterson B. Do 'basal-like' breast cancers really exist? Nat Rev Cancer 2009;9:128–134.
58 Marhaba R, Zoller M. CD44 in cancer progression: Adhesion, migration and growth regulation. J Mol Histol 2004;35:211–231.
59 Sarrio D, Franklin CK, Mackay A et al. Epithelial and mesenchymal subpopulations within normal basal breast cell lines exhibit distinct stem cell/progenitor properties. Stem Cells 2012;30:292–303.
60 Pece S, Tosoni D, Confalonieri S et al. Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. Cell 2010;140:62–73.
61 Croker AK, Goodale D, Chu J et al. High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability. J Cell Mol Med 2009;13:2236–2252.
62 Zhang CC, Yan Z, Zhang Q et al. PF-03732010: A fully human monoclonal antibody against P-cadherin with antitumor and antimetastatic activity. Clin Cancer Res 2010;16:5177–5188.


See www.StemCells.com for supporting information available online.

Int. J. Dev. Biol. 55: 811-822
doi: 10.1387/ijdb.113382aa

THE INTERNATIONAL JOURNAL OF DEVELOPMENTAL BIOLOGY
www.intjdevbiol.com

P-cadherin role in normal breast development and cancer

ANDRÉ ALBERGARIA[1], ANA-SOFIA RIBEIRO[1], ANDRÉ-FILIPE VIEIRA[1], BÁRBARA SOUSA[1], ANA-RITA NOBRE[1], RAQUEL SERUCA[1,2], FERNANDO SCHMITT[1,2] and JOANA PAREDES*[,1,2]

[1]IPATIMUP – Institute of Molecular Pathology and Immunology of the University of Porto and [2]Medical Faculty of Porto University, Porto, Portugal.

ABSTRACT P-cadherin is a cell-cell adhesion molecule, whose expression is highly associated with undifferentiated cells in normal adult epithelial tissues, as well as with poorly differentiated carcinomas. Its expression has been already reported in human embryonic stem cells and it is presumed to be a marker of stem or progenitor cells of some epithelial tissues. In normal breast, P-cadherin has an essential role during ductal mammary branching, being expressed by the monolayer of epithelial cap cells at the end buds. In mature mammary tissue, its expression is restricted to the myoepithelium; it has been postulated that it may also be present in early luminal progenitor cells. In breast cancer, P-cadherin is frequently overexpressed in high-grade tumours, being a well-established indicator of poor patient prognosis. It has been reported as an important inducer of cancer cell migration and invasion, with underlying molecular mechanisms involving the signalling mediated by its juxtamembrane domain, the secretion of matrix metalloproteases to the extracellular media, and the cleavage of a P-cadherin soluble form with pro-invasive activity. Intracellularly, this protein interferes with the endogenous cadherin/catenin complex, inducing p120-catenin delocalization to the cytoplasm, and the consequent activation of Rac1/Cdc42 and associated alterations in the actin cytoskeleton. Considering P-cadherin's role in cancer cell invasion and metastasis formation, a humanized monoclonal antibody was recently produced to antagonize P-cadherin-associated signalling pathways, which is currently under Phase I clinical trials. In this review, the most important findings about the role of P-cadherin in normal breast development and cancer will be illustrated and discussed, with emphasis on the most recent data.

KEY WORDS: *P-cadherin, CDH3 gene, mammary gland, breast cancer*

Introduction

Classical cadherins constitute a family of molecules that mediate calcium-dependent cell-cell adhesion, localized to the adherens-type junctions. The intracellular domains of cadherins bind directly to cytoplasmic catenins, which link them with the actin cytoskeleton, providing the molecular basis for stable cell interactions. The cadherin/catenin complex, as well as the signalling pathways controlled by this structure, represent a major regulatory mechanism that guide cell fate decisions, through its influence on cell growth, differentiation, motility, and survival (Cavallaro and Dejana, 2011).

Classical cadherins include *CDH1*/E-cadherin (epithelial), *CDH2*/N-cadherin (neuronal), *CDH3*/P-cadherin (placental) and *CDH4*/R-cadherin (retinal), designated by their tissue distribution. E-cadherin is the predominant cadherin family member expressed in all epithelial tissues, being extremely important to the maintenance of the cell shape and polarity; in fact, it is well known that *CDH1* acts as a tumour suppressor gene, negatively regulating the invasion and metastasis of tumour cells in several malignancies (Yilmaz and Christofori, 2010). In contrast, N-cadherin is up-regulated in several cancers and contributes to an invasive phenotype by interacting with fibroblast growth factor receptor (FGFR) and its

Abbreviations used in this paper: α–ctn, α-catenin; βctn, β-catenin; CBD, catenin-binding domain; CDH, cadherin; C/EBPβ, CCAAT/enhancer-binding protein β; CK, cytokeratin; CSC, cancer stem cell; CTC, circulating tumour cell; E-cad, E-cadherin; EC, epithelial cell; EEM, ectodermal dysplasia, ectrodactyly, and macular dystrophy; EGFR, epidermal growth factor receptor; ER, estrogen receptor; EMT, epithelial-to-mesenchymal transition; FGFR, fibroblast growth factor receptor; HDAC, histone deacetylase; HJMD, hypotrichosis with juvenile macular dystrophy; IBC, inflammatory breast cancer; JMD, juxtamembrane domain; MECs, myoepithelial cells; MFE, mammosphere forming efficiency; NAF, nipple aspirate fluid; P-cad, P-cadherin; PgR, progesterone receptor; p120ctn, p120-catenin; SHFM, split hand/foot malformation; sP-cad, soluble P-cadherin; TEB, terminal end buds; TSA, trichostatin A.

***Address correspondence to:** Joana Paredes. IPATIMUP, Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal. Tel: +35-12-2557-0700. Fax: +35-12-2557-0799. e-mail: jparedes@ipatimup.pt

Final, author-corrected PDF published online: 29 November 2011

ISSN: Online 1696-3547, Print 0214-6282

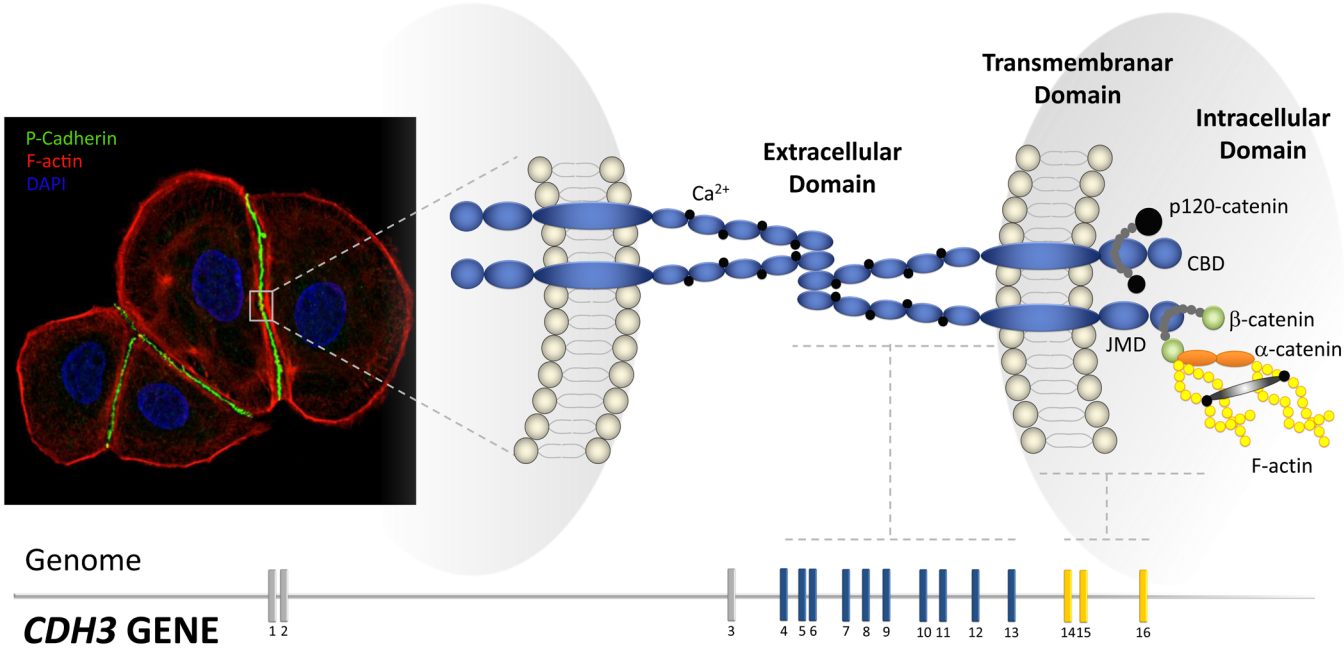


Fig. 1. Schematic representation of the structural components of the P-cadherin adhesive junction. *Lateral clustering of P-cadherin molecules is required to form stable cell-to-cell contacts in BT-20 breast cancer cells [immunofluorescence: P-cadherin (green), F-actin (red), DAPI (blue)]. In the intercellular space, P-cadherin extracellular domains interact with P-cadherin extracellular domains of adjacent cells to mediate cell adhesion. The intracellular catenins bind to the cytoplasmic tail of P-cadherin. p120-catenin binds the cadherin tail at the juxtamembrane domain (JMD), whereas* β*-catenin binds to the distal catenin binding domain (CBD).* α*-catenin associates with* β*-catenin and is directly linked to the actin cytoskeleton. The lower panel shows the genomic structure of* CDH3*/P-cadherin gene, which is constituted by 16 exons: the extracellular part of P-cadherin is encoded by 10 exons (exons 4-13), whereas the transmembrane and intracellular domains are determined only by the information included in the last 3 exons (exons 14-16).*

downstream signalling (Suyama *et al.*, 2002).

P-cadherin is also often reported to correlate with increased tumour cell motility and invasiveness when overexpressed (Cheung *et al.*, 2010, Paredes *et al.*, 2004, Ribeiro *et al.*, 2010, Taniuchi *et al.*, 2005). Although the role of P-cadherin encoding gene (*CDH3*) in cancer is far less well characterized than the one attributed to *CDH1*, the opposite effects in mammary cancer are weird, since these molecules share more than 67% of homology (Hulpiau and van Roy, 2009). The *CDH3* gene harbours 16 exons (Fig. 1) and maps to chromosome 16q22.1, a region that contains a cluster of several cadherin genes, just 32 kilobases upstream of the gene encoding human E-cadherin (Bussemakers *et al.*, 1994). The mature P-cadherin glycoprotein structure is similar to that of classical cadherins, comprising three distinct domains (extracellular, transmembrane and intracellular), in order to promote homotypic interactions. At the cell membrane, these create lateral dimers that act together in a zipper-like structure between neighbouring cells (Shapiro *et al.*, 1995) (Fig. 1).

The function and strength of P-cadherin-mediated adhesion depends on its dynamic association with catenins, which link the cadherin cytoplasmic tail to the actin cytoskeleton and facilitate clustering into the junctional structure, forming cadherin/catenin complexes. This tail comprises two main domains: the juxtamembrane domain (JMD), which has been suggested to play a critical role in cadherins stability at the cell membrane, and the catenin-binding domain (CBD), which is known to be essential for cadherin function. The p120-catenin (p120ctn), β-catenin (βctn) and α-catenin (αctn) are the major documented interaction partners that bind to cadherin intracellular domains and allow the binding to the actin cytoskeleton of the cell (Green *et al.*, 2010) (Fig. 1).

P-cadherin upregulation was frequently observed in various malignant tumours, including breast, gastric, endometrial, colorectal and pancreatic carcinomas, and is correlated with poor survival of breast cancer patients (Hardy *et al.*, 2002, Imai *et al.*, 2008, Paredes *et al.*, 2005, Stefansson *et al.*, 2004, Taniuchi *et al.*, 2005). In contrast, significantly low levels of the P-cadherin gene expression were detected in a diverse panel of normal tissues (Imai *et al.*, 2008). Thus, disruption of P-cadherin signalling represents an intriguing opportunity for the development of novel targeted therapeutic agents in cancer.

P-cadherin role in epithelial cell differentiation

Classical cadherins play important roles in maintaining the structural integrity of epithelial tissues and are mainly involved in cell differentiation during embryogenesis. There are several indications in the literature that point to the relationship between cell adhesion molecules and stem cell features, not only as biomarkers that help to isolate and characterise stem cells, but also as important mediators of stem cell activity, via modulation of signalling pathways (Raymond *et al.*, 2009). Regarding the classical cadherins, an important amount of data comes from the identification of P-cadherin as a marker of undifferentiated stem or progenitor cells (Kendrick *et al.*, 2008, Raymond *et al.*, 2009).

In a very recent study, it has been shown that *CDH3* is one of the genes that encode a surface protein that identify the pluripotent population of human embryonic stem cells (Kolle *et al.*, 2009). This expression is concomitant with the one of E-cadherin, which was

shown to be present even at the one cell stage of embryogenesis (Hyafil *et al.*, 1980) (Fig. 2A). In fact, mouse embryo implantation into the uterine epithelium involves both E- and P-cadherin. The most dramatic expression of P-cadherin was observed in the placenta, both in the embryonic and maternal regions, hence the classical denomination of placental-cadherin. The expression of P-cadherin in the uterus began with the appearance of the decidua, into which the extraembryonic cells expressing P-cadherin of implanted embryos invade to establish the embryo-maternal connection (Aplin *et al.*, 2009, Nose and Takeichi, 1986). Early reports specified low expression in human placenta (Shimoyama *et al.*, 1989), although P-cadherin is detectable where trophoblasts adjoin (cytotrophoblast-cytotrophoblast and cytotrophoblast–syncytiotrophoblast) in the first trimester villus, with some immunoreactivity still detectable at term (Aplin *et al.*, 2009) (Fig. 2A).

In contrast, E-cadherin was found expressed only in the embryonic region of placenta with a sharp boundary to the maternal region. These observations may suggest complementary roles of the two cadherins, such that P-cadherin is required for association of embryonic and maternal tissues during the late implantation stage, while E-cadherin is essential in preventing the embryonic tissues from mixing with the maternal tissues (Aplin *et al.*, 2009, Nose and Takeichi, 1986) (Fig. 2A).

It was also shown that E- and P-cadherins are both expressed in the ectoplacental cone, ectoderm, some endodermal tissues and nephric tubules, whereas both P- and N-cadherins are expressed in each cell of the lateral plate mesoderm, corneal endothelium, and pigmented retina (Nose and Takeichi, 1986) (Fig. 2A).

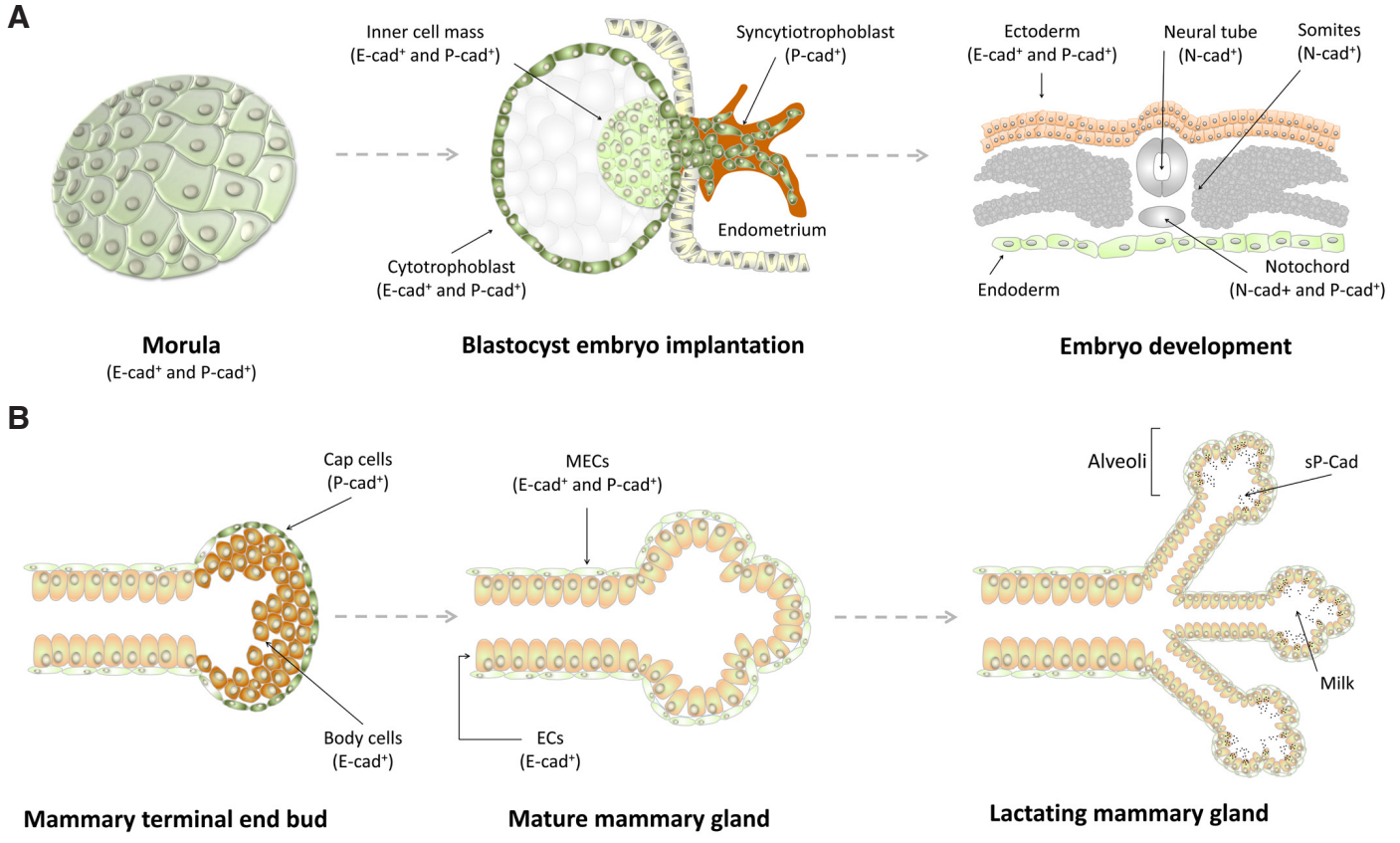


Fig. 2. Schematic representation of P-cadherin expression during embryogenesis and mammary gland development. (A) *Undifferentiated embryonic stem cells included in the morula, as well as in the inner mass of the blastocyst express E- and P-cadherin. During the blastocyst embryo implantation in the endometrial lining of the uterus, the trophectoderm differentiates into the cytotrophoblast and syncytiotrophoblast, which are key steps in placental development. There is an E-cadherin downregulation in the syncytiotrophoblast, which mainly expresses P-cadherin, while cells actively invade the uterine wall. Early in embryonic development, there is the formation of the neural tube, where a strip of specialized cells, forming the notochord, induces the cells of the ectoderm directly above it to become the primitive nervous system. Meanwhile, the ectoderm and endoderm continue to curve around and fuse beneath the embryo to create the body cavity, completing the transformation of the embryo from a flattened disk to a three–dimensional body. It is known that the ectoderm is E- and P-cadherin positive, which will give rise to the skin and its appendages. After cadherin switch, the neural tube becomes N-cadherin positive, as well as the somites. It was described that the notochord is N- and P-cadherin positive.* **(B)** *The ducts of the developing mammary gland are established, with their inner luminal epithelial cell layers (ECs) and outer myoepithelial cell layers (MECs), while the terminal end buds (TEB) move through the mammary fat pad. It is thought that cap cells at the tip of the TEB, which are P-cadherin positive, generate transit cells of the myoepithelial lineage on the outer side of the TEB (E- and P-cadehrin positive); at the same time, these cells also generate transit cells that form the central TEB mass, known as body cells, which will constitute the luminal epithelial lineage (E-cadherin positive). The ductal lumen is formed as body cells enter in apoptose and outer cells differentiate into luminal epithelial cells. Extracellular-matrix enzymes degrade the stroma in front of the TEB to enable it to move through the fat pad; however, it is unclear how the structures actually move through the gland. During lactation, secretory cells in the breast alveoli become P-cadherin positive at the cytoplasm, and secrete a soluble form of this protein (sP-cad) that is found in the milk.*

In adult tissues, the expression of P-cadherin is mainly found in the basal layer of several epithelial structures, such as skin, uterine cervix, prostate, and lung, contributing to the maintenance of the epithelial phenotype. The expression of cadherin molecules was extensively studied in mouse epidermis, in adulthood and during fetal development, where has been found that E-cadherin is expressed both in the basal and intermediate layers of epidermis, whereas P-cadherin is only expressed in the basal and proliferative layer (Pizarro *et al.*, 1995). Furthermore, loss of E-cadherin plays an important role in bud formation and in the acquisition of an invasive behaviour, whereas P-cadherin becomes predominant expressed later in development, namely in the growing hair follicle and in the early progenitor cells from hair germs and small hair placodes (Fujita *et al.*, 1992, Rhee *et al.*, 2006). Like hair follicles, sweat glands and mammary glands develop also from the same discrete accumulation of stem cells resting in the primitive epidermis, the outermost cell layer of an embryo, and there is strong evidence that dynamic changes in the composition of adherens junctions are important for the development of skin appendages (Fujita *et al.*, 1992).

The final evidence showing the importance of P-cadherin for the architecture and development of epithelial tissues was demonstrated by human genetic syndromes that are induced due to P-cadherin loss. *CDH3* gene mutations have been shown to cause P-cadherin functional inactivation, leading to developmental defects associated with two inherited diseases in humans: 1) hypotrichosis with juvenile macular dystrophy (HJMD) and 2) ectodermal dysplasia, ectrodactyly, and macular dystrophy (EEM syndrome). The common features of both diseases are sparse hair and macular dystrophy of the retina, while only EEM syndrome shows the additional finding of split hand/foot malformation (SHFM) (Kjaer *et al.*, 2005, Sprecher *et al.*, 2001). No defects were described for these conditions, concerning the human mammary development, or other epithelial bud structures. However, it is known that during bud patterning, a special arrangement occurs, where cells change their interaction with their neighbours and break their attachments to the extracellular matrix (ECM). Cells achieve this by activating specific transcriptional programs (Shimomura *et al.*, 2008).

P-cadherin role in normal breast development

Two members of the cadherin family are found to be expressed in the normal adult mature non-lactating mammary gland, usually at sites of cell-to-cell contact: E-cadherin is present in both luminal epithelial (ECs) and myoepithelial cells (MECs), whereas P-cadherin is confined to the myoepithelium (Paredes *et al.*, 2002). This type of cell localization is already found during mammary gland development, since P-cadherin expression is only found in the precursor cells of the myoepithelial compartment, the cap cells of the ductal end buds, whereas luminal cells and body cells do not show any expression of P-cadherin and are typically E-cadherin positive (Daniel *et al.*, 1995) (Fig. 2B).

Besides the restricted expression of P-cadherin in the normal breast, this protein is extremely important to the establishment of the correct architecture of the tissue, as demonstrated by functional-blocking antibody experiments *in vitro* and *in vivo*. Daniel and collaborators exposed the end buds and mature mammary glands of 5 week-old virgin mice to slow-release plastic implants liberating a monoclonal antibody for P-cadherin. No effect in the luminal layer was found, but disruption of the basally located cap cell layer was clearly observed (Daniel *et al.*, 1995). Also, more recently, Chanson *et al.,* described that P-cadherin contributes specifically to the organization of the myoepithelial cell layer of the breast, since when an antibody that blocks P-cadherin function was used in an *in vitro* self-organizing assay of the human mammary bilayer, the migration of MECs, occurring during normal sorting of both layers, was compromised (Chanson *et al.*, 2011). These experiments indicate that selective expression of P-cadherin in the basal layer is necessary for the maintenance of mammary tissue integrity.

In fact, deletion of P-cadherin affects normal mammopoiesis, since the *CDH3*-null female mice exhibit precocious mammary gland differentiation in the virgin state, and breast hyperplasia and dysplasia with age (Radice *et al.*, 1997). These observations in knockout animals indicate P-cadherin cell-cell interactions and signalling as regulatory determinants of the negative growth of the luminal epithelium, being important for the maintenance of an undifferentiated state of the normal mammary gland.

Interestingly, the expression of this adhesion molecule is activated in human mammary luminal cells during late pregnancy and lactation (Soler *et al.*, 2002). However, in these alveolar lactating cells, P-cadherin expression pattern is not restricted to the cell-cell borders, but shows a cytoplasmic staining, typical of a secreted protein. Indeed, in human milk, a soluble fragment of P-cadherin (sP-cad) with 80KDa was found to be present, corresponding to the extracellular domain of the molecule (Soler *et al.*, 2002) (Fig. 2B). Recently, Mannello and collaborators showed that the highest concentration of sP-cad is detected in milk collected during the first trimester of lactation (Mannello *et al.*, 2008). Still, it is not clear which is the biological and physiological role attributed to this fragment in the normal function of the breast. Some authors suggest a role for sP-cad in alveolar differentiation during lactation, or in the immune response of the mother or the baby, or as a signalling protein between epithelial and myoepithelial cells. Further studies are in progress to determine the sites of proteolysis of the sP-cad-secreted protein in different body fluids where it has been previously described (such as milk, serum, semen, nipple aspirate fluid (NAF), breast cyst fluid and amniotic fluid) (Mannello *et al.*, 2008, Soler *et al.*, 2002).

Prognostic relevance of P-cadherin in breast cancer

As mentioned above, P-cadherin is expressed in normal breast MECs and in MECs associated with non-invasive breast proliferations, showing no significant cross-reactivity with luminal/ECs, stromal myofibroblasts and blood vessels (Reis-Filho *et al.*, 2003). However, P-cadherin was described as being overexpressed in 20% to 40% of invasive breast carcinomas, as well as in 25% of ductal carcinomas *in situ* (DCIS) (Paredes *et al.*, 2007a, Paredes *et al.*, 2007b, Paredes *et al.*, 2002). Most important, several studies have reported P-cadherin as a marker of poor prognosis in breast cancer, since P-cadherin-positive carcinomas were significantly associated with short-term overall and disease-specific survival, as well as with distant and loco-regional relapse-free interval (Gamallo *et al.*, 2001, Paredes *et al.*, 2005, Peralta Soler *et al.*, 1999, Turashvili *et al.*, 2011).

P-cadherin expression has also been positively associated with high histological grade tumours, as well as with well-established

markers of poor prognosis, like Ki-67, epidermal growth factor receptor (EGFR), cytokeratin 5 (CK5), vimentin, p53, and HER2 expression and negatively associated with age at diagnosis, hormonal receptors (ER and PgR), and Bcl-2 expression. Interestingly, none of these reports showed a significant association with tumour size and lymph node metastasis (Gamallo *et al.*, 2001, Paredes *et al.*, 2005, Peralta Soler *et al.*, 1999, Turashvili *et al.*, 2011).

Besides the strong association between P-cadherin expression, poor patient prognosis and tumour aggressiveness, transgenic mice overexpressing *CDH3*/P-cadherin in the luminal epithelial layer of the mammary gland, under the control of the MMTV promoter, showed normal morphogenesis, architecture, lactation and involution, and no mammary tumours formed spontaneously (Radice *et al.*, 2003). Nevertheless, Mannello *et al.,* demonstrated a significant increased shedding of sP-cad in NAFs from women with breast cancer when compared with healthy subjects or with women with pre-cancer conditions, suggesting its possible release via proteolytic processing in cancer cells (Mannello *et al.*, 2008).

P-cadherin: marker of histological and molecular subtypes in breast cancer

Besides the strong association between P-cadherin expression and poor patient prognosis, no significant correlation was ever observed between this protein and a specific breast cancer histological type. The majority of positive P-cadherin tumours are invasive ductal carcinomas NOS, or carcinomas with metaplastic or medullary features (Paredes *et al.*, 2005, Reis-Filho *et al.*, 2003, Turashvili *et al.*, 2011). The observation that metaplastic and medullary breast carcinomas are consistently immunoreactive for P-cadherin supports a myoepithelial/basal transcriptomic programme for these lesions (Han *et al.*, 1999, Jacquemier *et al.*, 2005). Han and coworkers reported P-cadherin expression in almost all studied cases of medullary, carcinosarcomas, and sarcomatoid metaplastic breast carcinomas (Han *et al.*, 1999); in addition, all the metaplastic cases that we have studied were positive for at least one basal/myoepithelial marker, including P-cadherin (Reis-Filho *et al.*, 2003). We also showed that P-cadherin expression, in canine malignant tumours, was significantly related to spindle cell carcinoma, carcinosarcoma and osteosarcoma. In these lesions, both carcinomatous and sarcomatous components of carcinosarcoma expressed P-cadherin (Gama *et al.*, 2004, Gama *et al.*, 2008).

Concerning molecular profiling classification, at least five subtypes of invasive breast carcinoma were identified (Luminal A and B, Normal-like, HER2-overexpressing and Basal-like), exhibiting distinct clinical prognostic behaviour (Perou *et al.*, 2000). P-cadherin is one of the most important biomarkers to identify basal-like and HER2-overexpressing breast cancers (Arnes *et al.*, 2005, Paredes *et al.*, 2007b, Turashvili *et al.*, 2011). Basal-like breast cancer expresses genes characteristic of basal epithelial cells, which include, besides P-cadherin, high-molecular weight basal cytokeratins (CK5/6, CK14, CK17), vimentin, αB-crystalline, caveolins1/2 and EGFR (Arnes *et al.*, 2005). Until now, the most accepted criterion to identify basal-like breast carcinomas, by immunohistochemistry, is the triple negative phenotype along with CK5 and/or EGFR positivity (Nielsen *et al.*, 2004). However, we demonstrated that P-cadherin expression shows higher sensitivity to distinguish the basal phenotype of breast carcinomas, being a reliable option compared to the "gold standard" pair CK5/EGFR (Sousa *et al.*, 2010). Although this still need validation by gene expression profiles, these results can introduce the idea of using P-cadherin as an additional option in the daily workup of breast pathology laboratories to identify basal-like breast cancers.

P-cadherin is also prominently expressed in inflammatory breast cancer (IBC), which is a distinct and aggressive form of locally-advanced breast cancer, with high metastatic potential and high death rate. These tumours are characterized by frequent basal and HER2 phenotypes but, surprisingly, luminal IBC also express the basal marker P-cadherin (Ben Hamida *et al.*, 2008). This profile suggests a specificity that needs to be further investigated.

Interestingly, the expression profiling of *BRCA1*-deficient hereditary tumours has identified a pattern of gene expression similar to basal-like breast tumours (Palacios *et al.*, 2003). Very recently, Gorski *et al.* showed that *BRCA1* and *c-Myc* form a repressor complex on the promoters of specific basal genes, including *CDH3* gene, and represent a potential mechanism to explain the observed overexpression of key basal markers in *BRCA1*-deficient tumours (Gorski *et al.*, 2010). Actually, it has been shown that P-cadherin expression in breast carcinomas is strongly associated with the presence of *BRCA1* mutations (Arnes *et al.*, 2005).

P-cadherin role in adhesion, invasion and motility

Carcinomas progress by promotion of local invasion and distant metastasis. The acquisition of this invasive behaviour is one of the first steps in the metastatic process. Those cancer cells often develop alterations in their shape, as well as in their attachment to other cells and to ECM. Therefore, cell-cell and cell-matrix interactions play the most important role during tumour progression, since disruption of cell-cell adhesion during carcinogenesis is the basis for motility, invasion and metastasis of tumour cells (Yilmaz and Christofori, 2010).

P-cadherin has been detected as altered in various human tumours, but its effective role in the carcinogenesis process remains discussible, since it behaves differently depending on the studied tumour cell model and context. For instance, in a colorectal cancer cell line (HT-29), P-cadherin has been suggested to act as a pro-adhesive and anti-invasive/anti-migration molecule, exactly as E-cadherin (Van Marck *et al.*, 2011). Also, in melanomas, P-cadherin behaves as an invasion suppressor gene. Indeed, in highly invasive melanoma cell lines (that lack E-cadherin expression), P-cadherin overexpression was able to promote the formation of cell-cell contacts and counteract invasion (Van Marck *et al.*, 2005). The anti-invasive effect of P-cadherin was also recently confirmed in *in vivo* experiments, showing that its expression is refractory to invasive signals induced by myofibroblasts. Nevertheless, it was found a secreted truncated variant of P-cadherin in malignant melanomas, which negatively regulates cell-cell adhesion and induces a more motile phenotype, thus playing an important role in migration and metastasis of melanoma cells (Bauer and Bosserhoff, 2006).

On the other hand, in several other models, including breast cancer, P-cadherin behaves as an oncogene, and is often reported to correlate with increased tumour cell motility and invasiveness when aberrantly expressed (Cheung *et al.*, 2010, Mandeville *et al.*, 2008, Paredes *et al.*, 2007a, Paredes *et al.*, 2004, Taniuchi *et al.*, 2005, Van Marck *et al.*, 2011). Using *in vitro* breast cancer cell models, we found that overexpression of P-cadherin promotes single cell motility, directional cell migration, as well as invasion

capacity through the matrigel matrix (Ribeiro *et al.*, 2010). This same migratory phenotype was observed in bladder, pancreatic and cholangiocarcinoma cancer cell lines (Baek *et al.*, 2010, Mandeville *et al.*, 2008, Taniuchi *et al.*, 2005, Van Marck *et al.*, 2011).

Curiously, we have noticed that P-cadherin is able to induce invasion only in cell systems which already express an endogenous and functional cadherin, like E-cadherin in breast cancer cells, or N-cadherin in HEK293T cells and PDAC pancreatic cancer cells (Paredes *et al.*, 2004, Ribeiro *et al.*, 2010, Taniuchi *et al.*, 2005). Based on this hypothesis, we have recently proved that P-cadherin is able to interact with E-cadherin in breast tumours and cancer cells, promoting cancer cell invasion by disrupting the interaction between E-cadherin and both p120ctn and βctn. In the absence of E-cadherin expression, in the same cancer model, P-cadherin is able to suppress invasion by its strong interaction with catenins, surrogating the role of E-cadherin in cell-cell adhesion (unpublished data).

P-cadherin role in EMT and cadherin switch

Among the cadherin families, E-cadherin and N-cadherin are the most highly characterized subgroup of adhesion proteins. E-cadherin is ubiquitously expressed throughout most epithelial tissues and serves as a negative regulator to functionally block the βctn signalling pathway and suppress tumour cell growth and invasion (Onder *et al.*, 2008). However, numerous preclinical and clinical studies have shown that the loss of E-cadherin occurs concurrently with the upregulation of N-cadherin or other cadherin family members implicated in invasive growth, like P-cadherin or cadherin-11. This process, known as cadherin switching, has been reported to promote epithelial-to-mesenchymal transition (EMT) and leads to tumour cell invasion and metastasis (Thiery *et al.*, 2009).

Indeed, the switch from E- to N-cadherin is the one better known and reported by several studies. N-cadherin overexpression, via cadherin switching, was observed in various invasive cancer cell lines and tumours, namely from the esophagus, prostate, cervix, and ovary. This specific cadherin switch leads to the inhibition of cell-cell contacts and elicits active signals that support tumour-cell migration, invasion, and metastatic dissemination (Thiery *et al.*, 2009).

The cadherin switch from E- to P-cadherin is a common event during embryo development; however, few reports describe it during tumour progression. Indeed, some invasive and aggressive epithelial tumours, namely the local advanced IBC, and some highly metastatic breast cancer cells, as the 4T1 cell model, maintain E-cadherin expression at the cell membrane and show aberrant concomitant expression of P-cadherin (Ben Hamida *et al.*, 2008, Lou *et al.*, 2008). Nevertheless, there are some reports showing a switch from these two epithelial cadherins during tumour progression, namely in ovarian, endometrial and bladder carcinoma (Bryan *et al.*, 2008, Patel *et al.*, 2003, Stefansson *et al.*, 2004). In all these studies, P-cadherin increased expression significantly correlated with decreased E-cadherin expression and, consequently, represented a key step in disease progression. However, it has been already shown that, in cholangiocarcinoma cells, the E- to P-cadherin switch does not induce EMT signalling, since does not affect the expression of mesenchymal markers, such as Snail 1 and 2, vimentin, and fibronectin (Baek *et al.*, 2010).

Recognized regulators of *CDH3*/P-cadherin transcription

Signalling pathways or other cellular mechanisms that are involved in the regulation of cadherin-mediated adhesion are thought to underlie the dynamics of the adhesive interactions between cells.

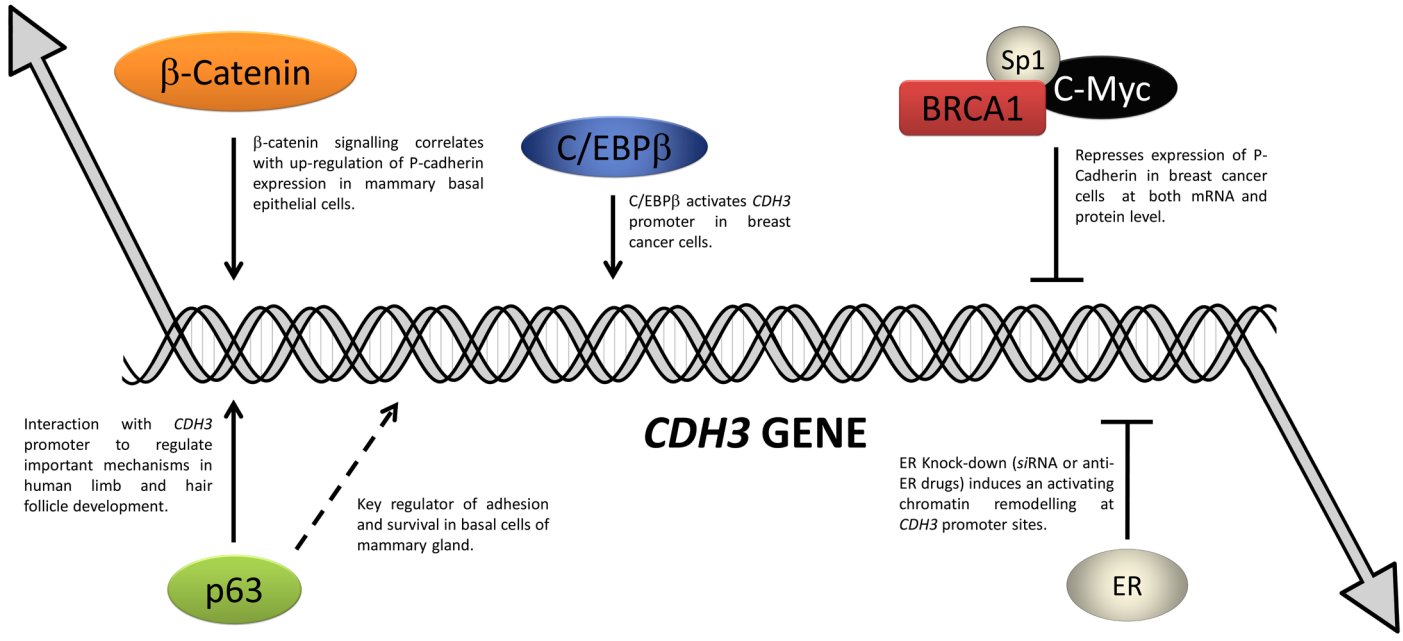


Fig. 3. Schematic representation of the described transcriptional regulators of *CDH3* /P-cadherin promoter gene. *It has been shown that β-catenin, p63 and C/EBPβ are transcriptional activators of* CDH3 *promoter, inducing its expression at the mRNA and protein level. In contrast, estrogen receptor (ER), as well as the BRCA1/c-Myc/Sp1 complex, act as transcriptional repressors of* CDH3 *promoter gene.*

Although the evidence that the expression of cadherins can result from growth factors and from changes in the promoter regions of cadherins, data concerning *CDH3* promoter regulation is still very limited.

One of the most prominent demonstrations regarding the importance of a classical transcription factor in the regulation of cell adhesion programmes in epithelial cells was demonstrated by Carroll and collaborators. This study implicated p63, a p53-family related transcription factor, as a key regulator of adhesion and survival in basal cells of the mammary gland. Importantly, the authors showed that p63 expression caused downregulation of cell adhesion-associated genes and detachment between mammary epithelial cells (Carroll *et al.*, 2006). This involvement of p63 in cell adhesion mechanisms was finally linked with *CDH3* gene, when Shimomura and colleagues demonstrated that P-cadherin is a direct p63 transcriptional target and that this interplay has a crucial role in human limb bud and hair follicle development (Shimomura *et al.*, 2008) (Fig. 3).

Furthermore, it has been shown that βctn is also associated with *CDH3* promoter activation and P-cadherin expression in basal mammary epithelial cells. Down-regulation of endogenous βctn levels inhibited *CDH3* promoter activity, while activation of βctn signalling was correlated with up-regulation of P-cadherin expression in *in vivo* mammary gland mice models, eventually contributing to the establishment of the basal phenotype (Faraldo *et al.*, 2007) (Fig. 3).

Recently, we still found that the CCAAT/enhancer-binding protein β (C/EBPβ) transcription factor was able to activate *CDH3* promoter in breast cancer cells. We showed that this novel activator of *CDH3* promoter activity exerts its activation preferably through its truncated LIP isoform, being the abundance of Sp1 sites within *CDH3* promoter a feature which potentiate the C/EBPβ-LIP activation role on *CDH3* gene (Albergaria *et al.*, 2010) (Fig. 3).

Regulation of *CDH3* gene has been also explored in terms of its transcriptional repression. In 2004, our group explored the link between ER-signalling and the regulation of P-cadherin expression in breast cancer cell lines, since we have already observed that breast tumours positive for P-cadherin expression were essentially ER negative. We verified that P-cadherin expression was induced by the pure anti-oestrogen ICI 182,780 and counteracted by 17β-oestradiol. In fact, breast cancer cells treated with ICI 182,780 showed a significant increase of P-cadherin mRNA and protein levels in a time and dose dependent manner, establishing that the lack of ER-signalling is responsible for the increase of P-cadherin, therefore, categorizing *CDH3* as an ER-repressed gene (Paredes *et al.*, 2004) (Fig. 3). Very recently, in order to deeply explore this antiestrogen-mediated mechanism, we described a cellular adaptation process where ICI 182,780 is able to induce a chromatin structural remodelling, which lead to activation of *CDH3* gene and overexpression of P-cadherin protein (Albergaria *et al.*, 2010). Such genomic de-repression effect may contribute to an augmented invasive phenotype of ER-positive breast cancer cells.

As a gene associated with the basal-like phenotype in breast cancer, *CDH3*/P-cadherin gene was recently described to be transcriptionally repressed by functional BRCA1 protein in breast cancer cell lines, at both mRNA and protein level. This same study also showed that, together with BRCA1, c-Myc form a repressor complex on the *CDH3* promoter (Fig. 3), suggesting a potential mechanism to explain the observed overexpression of key basal markers in BRCA1-deficient tumours (Gorski *et al.*, 2010).

Epigenetic regulation of P-cadherin expression

Epigenetic regulation of *CDH3*/P-cadherin has been highly reported in the last few years, with greater emphasis in cancer models. The epigenetic deregulation of P-cadherin was firstly demonstrated by Sato *et al.*, which identified *CDH3* gene promoter to be aberrantly methylated in 20% of pancreatic cancers, but not in normal pancreatic epithelia (Sato *et al.*, 2003). Similarly, *CDH3* gene was also shown to be silenced by methylation in melanoma cells (Tsutsumida *et al.*, 2004).

However, in 2005, we analysed P-cadherin promoter methylation in normal breast tissue, from which only epithelial cells were microdissected, and methylation of *CDH3* gene promoter was found in the normal epithelial/luminal cell layer from all the specimens analysed, which was associated with negative P-cadherin expression in these cells. But, in contrast to what has been verified in E-cadherin control of expression by hypermethylation of its promoter in cancer, our group found a significant correlation between P-cadherin overexpression and *CDH3* promoter hypomethylation. Using a large series of invasive breast carcinomas, we found that 71% of P-cadherin-negative breast cancer cases were methylated for the *CDH3* gene, whereas 65% of P-cadherin-positive cases were unmethylated (Paredes *et al.*, 2005).

Indeed, the genomic structure of the proximal *CDH3* gene promoter, such as the enrichment in CpG islands, as well as the attributed DNA hypersensitive sites, suggests that it is likely to be regulated by epigenetic events, others than only methylation. In fact, we observed an up-regulation of *CDH3* promoter activity and P-cadherin protein expression in cells treated with the histone deacetylases (HDAC) inhibitor Trichostatin A (TSA), showing that chromatin-activating modifications are also important in the modulation of this gene (Albergaria *et al.*, 2010). Thus, if we previously described that overexpression of P-cadherin could result from a loss of promoter methylation, we have now evidences to assume that chromatin remodelling also play an important modulator role in *CDH3* gene activity.

Reinforcing our results, *CDH3* promoter was also found hypomethylated in colonic aberrant crypt foci, in colorectal cancer, and, occasionally, in the normal epithelium adjacent to cancer (Milicic *et al.*, 2008). This hypomethylation pattern was associated with the induction of P-cadherin expression in the neoplastic colon. Finally, demethylation of the *CDH3* gene was recently detected in a large percentage of primary gastric carcinomas and was significantly associated with increasing TNM stage, suggesting that it is also a frequent event in gastric carcinomas (Kim *et al.*, 2010).

P-cadherin-downstream signalling pathways

Increasing evidences indicate that cadherins role in carcinogenesis and tumour progression do not solely lie on their adhesive function, but also depend on their interaction with other molecules (such as cytoskeletal components, integrins, and growth-factor receptors, among others) and signalling pathways (Onder *et al.*, 2008). Therefore, the stabilization of the cadherin/catenin complex represents a major regulatory mechanism for oncogenic signalling pathways, that guide cell fate decisions through the modulation of specific genes at the transcriptional level and, as a consequence, regulation of several crucial cellular processes, as proliferation, survival, polarization, differentiation, shape and migration, which

in turn affect embryogenesis, tissue formation and pathogenic events, such as cancer.

Although E-cadherin-induced signalling pathways have been extensively studied in cancer, little is known about the role of P-cadherin (Paredes *et al.*, 2004, Taniuchi *et al.*, 2005, Van Marck *et al.*, 2005). It is some kind expected that P-cadherin share common signalling pathways with other cadherins, due to its function as a cell-cell adhesion molecule; however, it is not known whether the pathways are triggered in the same way.

Sarrió and collaborators analysed microarray gene expression of a breast cancer cell line (MDA-MB-231), negative for cadherins, after expression of E- and P-cadherin. The data revealed that these molecules can activate signalling pathways leading to significant changes in gene expression. Although the expression patterns induced by E- and P-cadherin showed more similarities than differences, 40 genes were differentially modified by the expression of either cadherin type. According to data bases, these genes belonged to a wide range of biological functions, including signal transduction and growth factors (VEGFC, FGFR4), cell cycle (CCNA2), cell adhesion and ECM (CDH4, COL12A1), or cytokines and inflammation (IL24), among others (Sarrio *et al.*, 2009). This indicates that, in addition to their role in cell adhesion, E-cadherin and P-cadherin have a significant impact on the overall genetic program of breast cancer cells.

One of the molecules that have been several times referred has having a specific role in signalling related to P-cadherin is p120ctn (Fig. 4). We demonstrated that the pro-invasive activity of P-cadherin requires the JMD of its cytoplasmic tail. Transfection of HEK293T cells with several mutants of P-cadherin showed that only the ones with altered JMD were not able to induce cell invasion in *in vitro* cell models (Paredes *et al.*, 2004). Moreover, we observed that breast carcinomas co-expressing E- and P-cadherin were associated with p120ctn cytoplasmic localisation and poor patient survival (Paredes *et al.*, 2008). Since then, several other reports have been exploring that pathway. Indeed, Taniuchi *et al.,* showed that the induced cell migration by P-cadherin expression was due to activation of the Rho

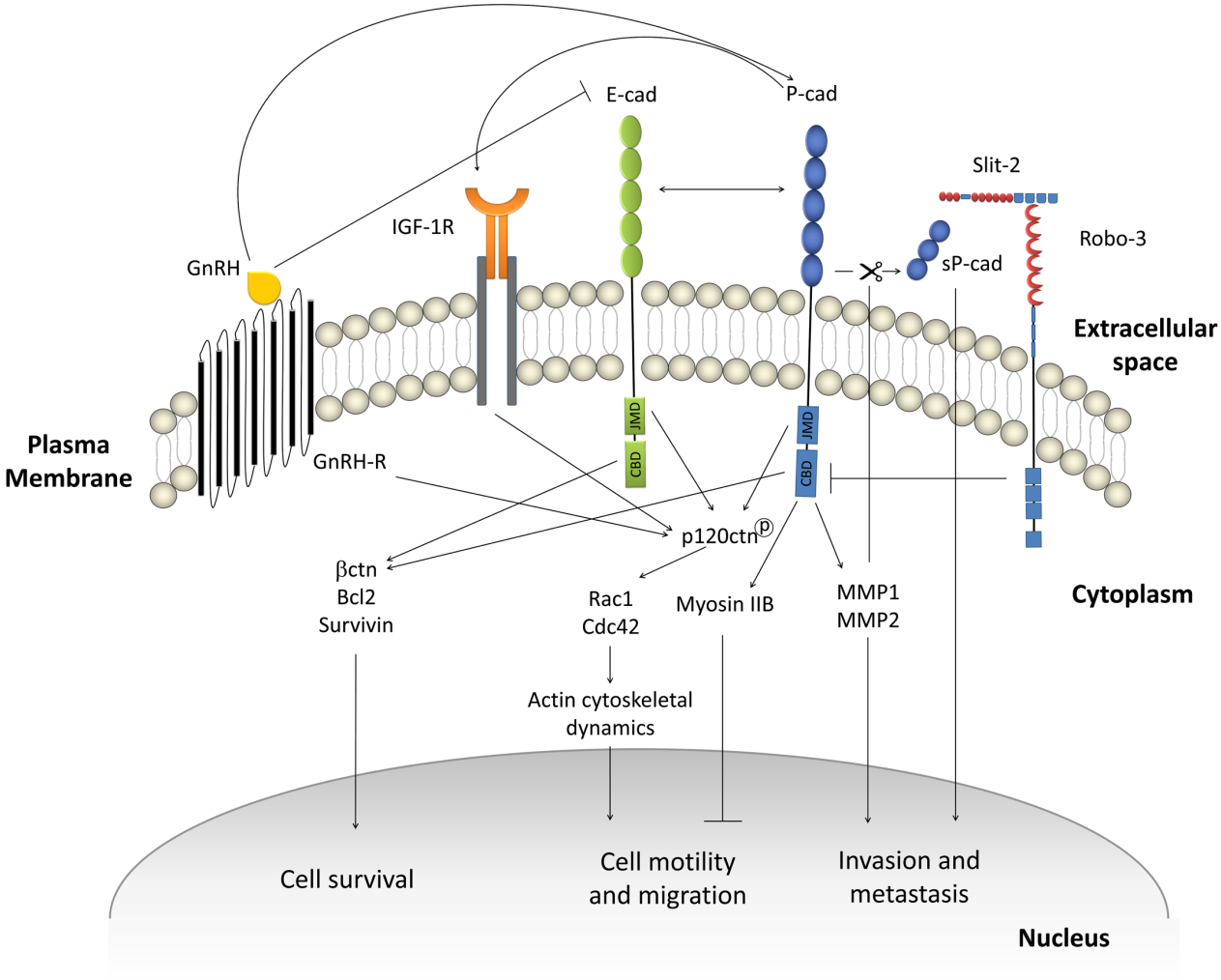


Fig. 4. Schematic representation of the signalling pathways regulated by P-cadherin expression. *P-cadherin signals are transduced by many intracellular signalling pathways, which ultimately result in alterations of the cancer cells survival, as well as cell migration and invasion capacity. For simplicity, only some of the known interactions are depicted. It should be noted that the effect of P-cadherin on the overall gene expression program of cancer cells is highly dependent on the cellular type and the biological context.*

GTPases, Rac1 and Cdc42, through accumulation of p120ctn in the cytoplasm in pancreatic cancer cell model (Taniuchi *et al.*, 2005) (Fig. 4). Very recently, P-cadherin has been also shown to cooperate with insulin-like growth factor-1 receptor to promote metastatic signalling of gonadotropin-releasing hormone in ovarian cancer via p120ctn (Cheung *et al.*, 2010). These same authors had previously shown that this p120ctn signalling mediated by P-cadherin expression, also lead to increased activity levels of Rac1 and Cdc42 (Fig. 4). Still another study has shown that p120ctn and P-cadherin, but not E-cadherin, regulate cell motility and invasion of DU145 prostate cancer cells (Kumper and Ridley, 2010).

Although binding of proteins to the JMD of P-cadherin has just been documented for p120ctn (Reynolds *et al.*, 1996), other molecules, like Hakai and presenilin-1 (PS-1), have been reported to bind to the JMD of classical cadherins. This binding is established through a sequence adjacent to, or overlapping, the p120ctn-binding domain, thereby competing with p120ctn (Baki *et al.*, 2001, Fujita *et al.*, 2002). Although the significance of these interactions is not well known, we cannot exclude the possibility that disruption of the p120ctn–binding sequence may introduce conformational changes and/or uncouples the interaction of these or other proteins, which could explain our observations. Striking examples of this were shown for E-cadherin, where functional differences have been noted between larger and minimal deletions of the JMD, with even the minimal changes disrupting binding of multiple molecules (Baki *et al.*, 2001).

Recently, it has been shown that the P-cadherin regulatory role in cell migration is also related with the expression of the non-muscle myosin II-B isoform, which is an ATP-dependent molecular motor protein that can interact with and contract filamentous actin (F-actin) (Jacobs *et al.*, 2010) (Fig. 4). These results implicate that there is a coordinated cross-talk between adhesion molecules and cellular migration-related proteins.

More recently, the role of P-cadherin was investigated in oral squamous cancer cell model, where the authors used a cell line that was deficient for classical cadherins. After P-cadherin overexpression, cells gained an epithelial-like morphology, with Snail translocation to the cytoplasm. Analysing the signalling mechanism behind it, they found glycogen-synthase-kinase-3β (GSK-3β) bound to Snail, as well as an increase in activated GSK-3β that phosphorylated Snail leading to its cytoplasmic translocation (Bauer *et al.*, 2009). These same authors also showed that Slit-2, a secreted ECM glycoprotein that acts as a molecular guidance cue in cellular migration, facilitates the interaction of P-cadherin with Robo-3, its receptor, and inhibits cell migration in oral squamous cell carcinoma cell line models (Bauer *et al.*, 2011) (Fig. 4).

In terms of breast cancer cell invasion, we found that the presence of P-cadherin, in an E-cadherin positive cellular background, is able to provoke the secretion of pro-invasive factors, such as MMP-1 and MMP-2, leading to P-cadherin ectodomain cleavage (sP-cad) which induces a pro-invasive activity by itself (Ribeiro *et al.*, 2010). This study clarified the mechanism associated to P-cadherin-induced cancer cell invasion.

Different signalling pathways should be triggered in different cell models, in order to identify new interaction partners of P-cadherin, as well as to study whether the interaction of known partner molecules differ between cadherins. Finally, it is important to highlight that the effect of cadherins on the overall gene expression program of cancer cells is highly dependent on the cellular type and the biological context. Thus, P-cadherin regulation of specific transcriptional factors may depend on the activation of other signalling pathways, or on the presence of additional molecular alterations.

P-cadherin as a breast cancer stem cell marker

An increasing body of evidence supports the notion that cancers are propagated by a small population of cells present in the malignant tissue, that possess the ability to form a hierarchy similar to the one present in normal tissues (Visvader, 2011). These cancer stem cells (CSCs) are able to proliferate, originating more stem-like cells, to exhibit resistance to current therapies and to remain quiescent during long periods of time. However, it is still not clear whether the CSC originates from the normal stem cells of the tissue that deregulate their self-renewal ability, or from normal mature cells or progenitor cells that acquired stem cell characteristics (Visvader, 2011). Importantly, attempts have been made in order to find a universal phenotype for the breast cancer stem cell; but due to the high heterogeneity of this malignancy, it is not expected that a single CSC phenotype would apply to all breast cancers.

The identification of a cancer stem cell marker for basal-like subtype of breast cancer is of particular importance, due to its high mortality rate, fast relapses and lack of target therapy (Rakha *et al.*, 2009). Recently, it has been demonstrated that the luminal progenitor of normal breast hierarchy is the cell of origin for this malignancy, since the induction of a *BRCA1* mutation in this cell was able to induce the formation of a tumour with basal phenotype (Lim *et al.*, 2009, Molyneux *et al.*, 2010). Since *CDH3* gene is repressed by BRCA1, it is likely that P-cadherin could be a good cancer stem cell marker of this specific type of tumours. In fact, using a series of breast cancer cell lines, we found that P-cadherin enriched populations (by genetic manipulation or by sorting) were enriched for mammosphere forming efficiency (MFE), as well as for the expression of CD24, CD44 and CD49f, already described as CSC markers. When compared with luminal cell lines, basal-like cell lines also showed a greater ALDEFLUOR^{bright} subpopulation and the P-cadherin positive subfraction of these cell lines was enriched in stem cell activity (MFE and 3D growth) (unpublished data). This observation linked P-cadherin expression with the luminal progenitor phenotype, which is CD44^{+}CD24^{+}CD49f^{+} (Lim *et al.*, 2009). Importantly, it has been described that the phenotype CD44^{+}CD24^{+} is tumorigenic (Meyer *et al.*, 2009). Hence, the strategy of directing therapies to the luminal progenitor phenotype, by specifically targeting P-cadherin, could potentially help to eradicate the CSCs. Interestingly, P-cadherin also conferred resistance to X-ray induced DNA damage, supporting a role for this molecule in the maintenance of yet another CSC property (unpublished data).

P-cadherin - potential therapeutic target in cancer

As clearly stated in this review, P-cadherin–mediated adhesion and the associated signalling pathways play diverse roles in the regulation of cancer cell survival, invasiveness and metastatic potential. Interestingly, in 2008, Imai and collaborators have suggested *CDH3*/P-cadherin as a possible target for immunotherapy of pancreatic, gastric, and colorectal cancers, since it was identified as a novel tumour-associated antigen, meaning that was strongly expressed in tumour cells, but not in normal cells (Imai *et al.*, 2008). Indeed, we have found that P-cadherin silencing, in breast cancer

cells inoculated in nude mice, was able to significantly inhibit *in vivo* tumour growth (unpublished data).

Recently, a novel and highly selective human monoclonal antibody against P-cadherin (PF-03732010) was produced, demonstrating anti-tumour and anti-metastatic activity in a diverse panel of P-cadherin–overexpressing tumour models, without introducing any adverse secondary effects in mice (Zhang *et al.*, 2010). This antibody failed to bind to the most closely target-related family members, including E-cadherin, N-cadherin, and VE-cadherin. PF-03732010 also reduced lymph node metastases and lowered the levels of circulating tumour cells (CTC) in whole blood of P-cadherin$^+$ tumour bearing mice. The anti-metastatic property of the antibody was remarkable, since it significantly inhibited tumour cell infiltration into the lungs. PF-03732010 still suppressed βctn, cyclin D1, Vimentin, Bcl-2, and survivin expression, decreased the Ki67 levels, and increased caspase-3 expression (Zhang *et al.*, 2010) (Fig. 4).

Taken together, these recent data highlight the critical role of P-cadherin signalling in regulating tumorigenesis and metastasis, especially because its inhibition leads to anti-tumour and anti-metastatic effects in target-associated tumour models without any adverse indication. These observations provide the rationale and guidance for the clinical development of PF-03732010, in which tumours with high P-cadherin expression will be essential criteria for patient selection. Future work is warranted to seek a reproducible method to quantify P-cadherin in human tumours and to find a reasonable cut-off of expression related with therapeutic response, in an attempt to reach the full potential for clinical development of the antibody. PF-03732010 is currently under Phase I clinical trial development.

Conclusions

Although this review is mainly focused on P-cadherin role as a poor prognostic factor, as well as a therapeutic target in breast cancer, its upregulation is also found in several other malignancies, affecting organs such as pancreas, stomach, bladder and prostate, where it is also associated with an aggressive phenotype and poor prognosis. Thus, antagonizing P-cadherin represents a novel approach for anticancer therapy, by targeting tumours with high P-cadherin expression. Interestingly, P-cadherin silencing induces significant growth inhibition in several tumour models tested; however, this anti-proliferative activity was never observed *in vitro* (Zhang *et al.*, 2010). This discrepancy suggests that fully functioning P-cadherin signalling may require the cell-cell and cell-stroma crosstalk in intact tumour architecture during tumorigenesis and metastasis, a process that may not be recapitulated under *in vitro* conditions and that should be further studied in the future.

References

ALBERGARIA, A., RIBEIRO, A.S., PINHO, S., MILANEZI, F., CARNEIRO, V., SOUSA, B., SOUSA, S., OLIVEIRA, C., MACHADO, J.C., SERUCA, R. *et al.,* (2010). ICI 182,780 induces P-cadherin overexpression in breast cancer cells through chromatin remodelling at the promoter level: a role for C/EBPb in *CDH3* gene activation. *Hum Mol Genet* 19: 2554-2566.

APLIN, J.D., JONES, C.J. and HARRIS, L.K. (2009). Adhesion molecules in human trophoblast - a review. I. Villous trophoblast. *Placenta* 30: 293-298.

ARNES, J.B., BRUNET, J.S., STEFANSSON, I., BEGIN, L.R., WONG, N., CHAPPUIS, P.O., AKSLEN, L.A. and FOULKES, W.D. (2005). Placental cadherin and the basal epithelial phenotype of BRCA1-related breast cancer. *Clin Cancer Res* 11: 4003-4011.

BAEK, S., LEE, Y.W., YOON, S., BAEK, S.Y., KIM, B.S. and OH, S.O. (2010). CDH3/P-Cadherin regulates migration of HuCCT1 cholangiocarcinoma cells. *Anat Cell Biol* 43: 110-117.

BAKI, L., MARAMBAUD, P., EFTHIMIOPOULOS, S., GEORGAKOPOULOS, A., WEN, P., CUI, W., SHIOI, J., KOO, E., OZAWA, M., FRIEDRICH, V.L., JR. *et al.,* (2001). Presenilin-1 binds cytoplasmic epithelial cadherin, inhibits cadherin/p120 association, and regulates stability and function of the cadherin/catenin adhesion complex. *Proc Natl Acad Sci USA* 98: 2381-2386.

BAUER, K., DOWEJKO, A., BOSSERHOFF, A.K., REICHERT, T.E. and BAUER, R. (2011). Slit-2 facilitates interaction of P-cadherin with Robo-3 and inhibits cell migration in an oral squamous cell carcinoma cell line. *Carcinogenesis* 32: 935-943.

BAUER, K., DOWEJKO, A., BOSSERHOFF, A.K., REICHERT, T.E. and BAUER, R.J. (2009). P-cadherin induces an epithelial-like phenotype in oral squamous cell carcinoma by GSK-3beta-mediated Snail phosphorylation. *Carcinogenesis* 30: 1781-1788.

BAUER, R. and BOSSERHOFF, A.K. (2006). Functional implication of truncated P-cadherin expression in malignant melanoma. *Exp Mol Pathol* 81: 224-230.

BEN HAMIDA, A., LABIDI, I.S., MRAD, K., CHARAFE-JAUFFRET, E., BEN ARAB, S., ESTERNI, B., XERRI, L., VIENS, P., BERTUCCI, F., BIRNBAUM, D. *et al.,* (2008). Markers of subtypes in inflammatory breast cancer studied by immunohistochemistry: prominent expression of P-cadherin. *BMC Cancer* 8: 28.

BRYAN, R.T., ATHERFOLD, P.A., YEO, Y., JONES, L.J., HARRISON, R.F., WALLACE, D.M. and JANKOWSKI, J.A. (2008). Cadherin switching dictates the biology of transitional cell carcinoma of the bladder: ex vivo and *in vitro* studies. *J Pathol* 215: 184-194.

BUSSEMAKERS, M.J., VAN BOKHOVEN, A., VOLLER, M., SMIT, F.P. and SCHALKEN, J.A. (1994). The genes for the calcium-dependent cell adhesion molecules P- and E-cadherin are tandemly arranged in the human genome. *Biochem Biophys Res Commun* 203: 1291-1294.

CARROLL, D.K., CARROLL, J.S., LEONG, C.O., CHENG, F., BROWN, M., MILLS, A.A., BRUGGE, J.S. and ELLISEN, L.W. (2006). p63 regulates an adhesion programme and cell survival in epithelial cells. *Nat Cell Biol* 8: 551-561.

CAVALLARO, U. and DEJANA, E. (2011). Adhesion molecule signalling: not always a sticky business. *Nat Rev Mol Cell Biol* 12: 189-197.

CHANSON, L., BROWNFIELD, D., GARBE, J.C., KUHN, I., STAMPFER, M.R., BISSELL, M.J. and LABARGE, M.A. (2011). Self-organization is a dynamic and lineage-intrinsic property of mammary epithelial cells. *Proc Natl Acad Sci USA* 108: 3264-3269.

CHEUNG, L.W., LEUNG, P.C. and WONG, A.S. (2010). Cadherin switching and activation of p120 catenin signaling are mediators of gonadotropin-releasing hormone to promote tumor cell migration and invasion in ovarian cancer. *Oncogene* 29: 2427-2440.

DANIEL, C.W., STRICKLAND, P. and FRIEDMANN, Y. (1995). Expression and functional role of E- and P-cadherins in mouse mammary ductal morphogenesis and growth. *Dev Biol* 169: 511-519.

FARALDO, M.M., TEULIERE, J., DEUGNIER, M.A., BIRCHMEIER, W., HUELSKEN, J., THIERY, J.P., CANO, A. and GLUKHOVA, M.A. (2007). beta-Catenin regulates P-cadherin expression in mammary basal epithelial cells. *FEBS Lett* 581: 831-836.

FUJITA, M., FURUKAWA, F., FUJII, K., HORIGUCHI, Y., TAKEICHI, M. and IMAMURA, S. (1992). Expression of cadherin cell adhesion molecules during human skin development: morphogenesis of epidermis, hair follicles and eccrine sweat ducts. *Arch Dermatol Res* 284: 159-166.

FUJITA, Y., KRAUSE, G., SCHEFFNER, M., ZECHNER, D., LEDDY, H.E., BEHRENS, J., SOMMER, T. and BIRCHMEIER, W. (2002). Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex. *Nat Cell Biol* 4: 222-231.

GAMA, A., PAREDES, J., ALBERGARIA, A., GARTNER, F. and SCHMITT, F. (2004). P-cadherin expression in canine mammary tissues. *J Comp Pathol* 130: 13-20.

GAMA, A., PAREDES, J., GARTNER, F., ALVES, A. and SCHMITT, F. (2008). Expression of E-cadherin, P-cadherin and beta-catenin in canine malignant mammary tumours in relation to clinicopathological parameters, proliferation and survival. *Vet J* 177: 45-53.

GAMALLO, C., MORENO-BUENO, G., SARRIO, D., CALERO, F., HARDISSON, D. and PALACIOS, J. (2001). The prognostic significance of P-cadherin in infiltrating ductal breast carcinoma. *Mod Pathol* 14: 650-654.

GORSKI, J.J., JAMES, C.R., QUINN, J.E., STEWART, G.E., STAUNTON, K.C., BUCKLEY, N.E., MCDYER, F.A., KENNEDY, R.D., WILSON, R.H., MULLAN, P.B. *et al.,* (2010). BRCA1 transcriptionally regulates genes associated with the basal-like phenotype in breast cancer. *Breast Cancer Res Treat* 122: 721-731.

GREEN, K.J., GETSIOS, S., TROYANOVSKY, S. and GODSEL, L.M. (2010). Intercellular junction assembly, dynamics, and homeostasis. *Cold Spring Harb Perspect Biol* 2: a000125.

HAN, A.C., SOLER, A.P., KNUDSEN, K.A. and SALAZAR, H. (1999). Distinct cadherin profiles in special variant carcinomas and other tumors of the breast. *Hum Pathol* 30: 1035-1039.

HARDY, R.G., TSELEPIS, C., HOYLAND, J., WALLIS, Y., PRETLOW, T.P., TALBOT, I., SANDERS, D.S., MATTHEWS, G., MORTON, D. and JANKOWSKI, J.A. (2002). Aberrant P-cadherin expression is an early event in hyperplastic and dysplastic transformation in the colon. *Gut* 50: 513-519.

HULPIAU, P. and VAN ROY, F. (2009). Molecular evolution of the cadherin superfamily. *Int J Biochem Cell Biol* 41: 349-369.

HYAFIL, F., MORELLO, D., BABINET, C. and JACOB, F. (1980). A cell surface glycoprotein involved in the compaction of embryonal carcinoma cells and cleavage stage embryos. *Cell* 21: 927-934.

IMAI, K., HIRATA, S., IRIE, A., SENJU, S., IKUTA, Y., YOKOMINE, K., HARAO, M., INOUE, M., TSUNODA, T., NAKATSURU, S. *et al.,* (2008). Identification of a novel tumor-associated antigen, cadherin 3/P-cadherin, as a possible target for immunotherapy of pancreatic, gastric, and colorectal cancers. *Clin Cancer Res* 14: 6487-6495.

JACOBS, K., VAN GELE, M., FORSYTH, R., BROCHEZ, L., VANHOECKE, B., DE WEVER, O. and BRACKE, M. (2010). P-cadherin counteracts myosin II-B function: implications in melanoma progression. *Mol Cancer* 9: 255.

JACQUEMIER, J., PADOVANI, L., RABAYROL, L., LAKHANI, S.R., PENAULT-LLORCA, F., DENOUX, Y., FICHE, M., FIGUEIRO, P., MAISONGROSSE, V., LEDOUSSAL, V. *et al.,* (2005). Typical medullary breast carcinomas have a basal/myoepithelial phenotype. *J Pathol* 207: 260-268.

KENDRICK, H., REGAN, J.L., MAGNAY, F.A., GRIGORIADIS, A., MITSOPOULOS, C., ZVELEBIL, M. and SMALLEY, M.J. (2008). Transcriptome analysis of mammary epithelial subpopulations identifies novel determinants of lineage commitment and cell fate. *BMC Genomics* 9: 591.

KIM, M.A., JUNG, E.J., LEE, H.S., LEE, H.E., YANG, H.K., OH, D.Y., BANG, Y.J. and KIM, W.H. (2010). P-cadherin expression in gastric carcinoma: its regulation mechanism and prognostic significance. *Hum Pathol* 41: 877-885.

KJAER, K.W., HANSEN, L., SCHWABE, G.C., MARQUES-DE-FARIA, A.P., EIBERG, H., MUNDLOS, S., TOMMERUP, N. and ROSENBERG, T. (2005). Distinct CDH3 mutations cause ectodermal dysplasia, ectrodactyly, macular dystrophy (EEM syndrome). *J Med Genet* 42: 292-298.

KOLLE, G., HO, M., ZHOU, Q., CHY, H.S., KRISHNAN, K., CLOONAN, N., BERTONCELLO, I., LASLETT, A.L. and GRIMMOND, S.M. (2009). Identification of human embryonic stem cell surface markers by combined membrane-polysome translation state array analysis and immunotranscriptional profiling. *Stem Cells* 27: 2446-2456.

KUMPER, S. and RIDLEY, A.J. (2010). p120ctn and P-cadherin but not E-cadherin regulate cell motility and invasion of DU145 prostate cancer cells. *PLoS One* 5: e11801.

LIM, E., VAILLANT, F., WU, D., FORREST, N.C., PAL, B., HART, A.H., ASSELIN-LABAT, M.L., GYORKI, D.E., WARD, T., PARTANEN, A. *et al.,* (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* 15: 907-913.

LOU, Y., PREOBRAZHENSKA, O., AUF DEM KELLER, U., SUTCLIFFE, M., BARCLAY, L., MCDONALD, P.C., ROSKELLEY, C., OVERALL, C.M. and DEDHAR, S. (2008). Epithelial-mesenchymal transition (EMT) is not sufficient for spontaneous murine breast cancer metastasis. *Dev Dyn* 237: 2755-2768.

MANDEVILLE, J.A., SILVA NETO, B., VANNI, A.J., SMITH, G.L., RIEGER-CHRIST, K.M., ZEHEB, R., LODA, M., LIBERTINO, J.A. and SUMMERHAYES, I.C. (2008). P-cadherin as a prognostic indicator and a modulator of migratory behaviour in bladder carcinoma cells. *BJU Int* 102: 1707-1714.

MANNELLO, F., TONTI, G.A., MEDDA, V., PEDERZOLI, A. and SAUTER, E.R. (2008). Increased shedding of soluble fragments of P-cadherin in nipple aspirate fluids from women with breast cancer. *Cancer Sci* 99: 2160-2169.

MEYER, M.J., FLEMING, J.M., ALI, M.A., PESESKY, M.W., GINSBURG, E. and VONDERHAAR, B.K. (2009). Dynamic regulation of CD24 and the invasive, CD-44posCD24neg phenotype in breast cancer cell lines. *Breast Cancer Res* 11: R82.

MILICIC, A., HARRISON, L.A., GOODLAD, R.A., HARDY, R.G., NICHOLSON, A.M., PRESZ, M., SIEBER, O., SANTANDER, S., PRINGLE, J.H., MANDIR, N. *et al.,* (2008). Ectopic expression of P-cadherin correlates with promoter hypomethylation early in colorectal carcinogenesis and enhanced intestinal crypt fission in vivo. *Cancer Res* 68: 7760-7768.

MOLYNEUX, G., GEYER, F.C., MAGNAY, F.A., MCCARTHY, A., KENDRICK, H., NATRAJAN, R., MACKAY, A., GRIGORIADIS, A., TUTT, A., ASHWORTH, A. *et al.,* (2010). BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell* 7: 403-417.

NIELSEN, T.O., HSU, F.D., JENSEN, K., CHEANG, M., KARACA, G., HU, Z., HERNANDEZ-BOUSSARD, T., LIVASY, C., COWAN, D., DRESSLER, L. *et al.,* (2004). Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* 10: 5367-5374.

NOSE, A. and TAKEICHI, M. (1986). A novel cadherin cell adhesion molecule: its expression patterns associated with implantation and organogenesis of mouse embryos. *J Cell Biol* 103: 2649-2658.

ONDER, T.T., GUPTA, P.B., MANI, S.A., YANG, J., LANDER, E.S. and WEINBERG, R.A. (2008). Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res* 68: 3645-3654.

PALACIOS, J., HONRADO, E., OSORIO, A., CAZORLA, A., SARRIO, D., BARROSO, A., RODRIGUEZ, S., CIGUDOSA, J.C., DIEZ, O., ALONSO, C. *et al.,* (2003). Immunohistochemical characteristics defined by tissue microarray of hereditary breast cancer not attributable to BRCA1 or BRCA2 mutations: differences from breast carcinomas arising in BRCA1 and BRCA2 mutation carriers. *Clin Cancer Res* 9: 3606-3614.

PAREDES, J., ALBERGARIA, A., OLIVEIRA, J.T., JERONIMO, C., MILANEZI, F. and SCHMITT, F.C. (2005). P-cadherin overexpression is an indicator of clinical outcome in invasive breast carcinomas and is associated with CDH3 promoter hypomethylation. *Clin Cancer Res* 11: 5869-5877.

PAREDES, J., CORREIA, A.L., RIBEIRO, A.S., ALBERGARIA, A., MILANEZI, F. and SCHMITT, F.C. (2007). P-cadherin expression in breast cancer: a review. *Breast Cancer Res* 9: 214.

PAREDES, J., CORREIA, A.L., RIBEIRO, A.S., MILANEZI, F., CAMESELLE-TEIJEIRO, J. and SCHMITT, F.C. (2008). Breast carcinomas that co-express E- and P-cadherin are associated with p120-catenin cytoplasmic localisation and poor patient survival. *J Clin Pathol* 61: 856-862.

PAREDES, J., LOPES, N., MILANEZI, F. and SCHMITT, F.C. (2007b). P-cadherin and cytokeratin 5: useful adjunct markers to distinguish basal-like ductal carcinomas in situ. *Virchows Arch* 450: 73-80.

PAREDES, J., MILANEZI, F., VIEGAS, L., AMENDOEIRA, I. and SCHMITT, F. (2002). P-cadherin expression is associated with high-grade ductal carcinoma *in situ* of the breast. *Virchows Archiv* 440: 16-21.

PAREDES, J., STOVE, C., STOVE, V., MILANEZI, F., VAN MARCK, V., DERYCKE, L., MAREEL, M., BRACKE, M. and SCHMITT, F. (2004). P-cadherin is up-regulated by the antiestrogen ICI 182,780 and promotes invasion of human breast cancer cells. *Cancer Res* 64: 8309-8317.

PATEL, I.S., MADAN, P., GETSIOS, S., BERTRAND, M.A. and MACCALMAN, C.D. (2003). Cadherin switching in ovarian cancer progression. *Int J Cancer* 106: 172-177.

PERALTA SOLER, A., KNUDSEN, K.A., SALAZAR, H., HAN, A.C. and KESHGEGIAN, A.A. (1999). P-cadherin expression in breast carcinoma indicates poor survival. *Cancer* 86: 1263-1272.

PEROU, C.M., SORLIE, T., EISEN, M.B., VAN DE RIJN, M., JEFFREY, S.S., REES, C.A., POLLACK, J.R., ROSS, D.T., JOHNSEN, H., AKSLEN, L.A. *et al.,* (2000). Molecular portraits of human breast tumours. *Nature* 406: 747-752.

PIZARRO, A., GAMALLO, C., BENITO, N., PALACIOS, J., QUINTANILLA, M., CANO, A. and CONTRERAS, F. (1995). Differential patterns of placental and epithelial cadherin expression in basal cell carcinoma and in the epidermis overlying tumours. *Br J Cancer* 72: 327-332.

RADICE, G.L., FERREIRA-CORNWELL, M.C., ROBINSON, S.D., RAYBURN, H., CHODOSH, L.A., TAKEICHI, M. and HYNES, R.O. (1997). Precocious mammary gland development in P-cadherin-deficient mice. *J Cell Biol* 139: 1025-1032.

RADICE, G.L., SAUER, C.L., KOSTETSKII, I., PERALTA SOLER, A. and KNUDSEN, K.A. (2003). Inappropriate P-cadherin expression in the mouse mammary epithelium is compatible with normal mammary gland function. *Differentiation* 71: 361-373.

RAKHA, E.A., ELSHEIKH, S.E., ALESKANDARANY, M.A., HABASHI, H.O., GREEN, A.R., POWE, D.G., EL-SAYED, M.E., BENHASOUNA, A., BRUNET, J.S., AKSLEN, L.A. *et al.,* (2009). Triple-negative breast cancer: distinguishing between basal and nonbasal subtypes. *Clin Cancer Res* 15: 2302-2310.

RAYMOND, K., DEUGNIER, M.A., FARALDO, M.M. and GLUKHOVA, M.A. (2009). Adhesion within the stem cell niches. *Curr Opin Cell Biol* 21: 623-629.

REIS-FILHO, J.S., MILANEZI, F., PAREDES, J., SILVA, P., PEREIRA, E.M., MAEDA, S.A., DE CARVALHO, L.V. and SCHMITT, F.C. (2003). Novel and classic myoepithelial/stem cell markers in metaplastic carcinomas of the breast. *Appl Immunohistochem Mol Morphol* 11: 1-8.

REYNOLDS, A.B., DANIEL, J.M., MO, Y.Y., WU, J. and ZHANG, Z. (1996). The novel catenin p120cas binds classical cadherins and induces an unusual morphological phenotype in NIH3T3 fibroblasts. *Exp Cell Res* 225: 328-337.

RHEE, H., POLAK, L. and FUCHS, E. (2006). Lhx2 maintains stem cell character in hair follicles. *Science* 312: 1946-1949.

RIBEIRO, A.S., ALBERGARIA, A., SOUSA, B., CORREIA, A.L., BRACKE, M., SERUCA, R., SCHMITT, F.C. and PAREDES, J. (2010). Extracellular cleavage and shedding of P-cadherin: a mechanism underlying the invasive behaviour of breast cancer cells. *Oncogene* 29: 392-402.

SARRIO, D., PALACIOS, J., HERGUETA-REDONDO, M., GOMEZ-LOPEZ, G., CANO, A. and MORENO-BUENO, G. (2009). Functional characterization of E- and P-cadherin in invasive breast cancer cells. *BMC Cancer* 9: 74.

SATO, N., FUKUSHIMA, N., MAITRA, A., MATSUBAYASHI, H., YEO, C.J., CAMERON, J.L., HRUBAN, R.H. and GOGGINS, M. (2003). Discovery of novel targets for aberrant methylation in pancreatic carcinoma using high-throughput microarrays. *Cancer Res* 63: 3735-3742.

SHAPIRO, L., FANNON, A.M., KWONG, P.D., THOMPSON, A., LEHMANN, M.S., GRUBEL, G., LEGRAND, J.F., ALS-NIELSEN, J., COLMAN, D.R. and HENDRICKSON, W.A. (1995). Structural basis of cell-cell adhesion by cadherins. *Nature* 374: 327-337.

SHIMOMURA, Y., WAJID, M., SHAPIRO, L. and CHRISTIANO, A.M. (2008). P-cadherin is a p63 target gene with a crucial role in the developing human limb bud and hair follicle. *Development* 135: 743-753.

SHIMOYAMA, Y., YOSHIDA, T., TERADA, M., SHIMOSATO, Y., ABE, O. and HIROHASHI, S. (1989). Molecular cloning of a human Ca2+-dependent cell-cell adhesion molecule homologous to mouse placental cadherin: its low expression in human placental tissues. *J Cell Biol* 109: 1787-1794.

SOLER, A.P., RUSSO, J., RUSSO, I.H. and KNUDSEN, K.A. (2002). Soluble fragment of P-cadherin adhesion protein found in human milk. *J Cell Biochem* 85: 180-184.

SOUSA, B., PAREDES, J., MILANEZI, F., LOPES, N., MARTINS, D., DUFLOTH, R., VIEIRA, D., ALBERGARIA, A., VERONESE, L., CARNEIRO, V. *et al.,* (2010). P-cadherin, vimentin and CK14 for identification of basal-like phenotype in breast carcinomas: an immunohistochemical study. *Histol Histopathol* 25: 963-974.

SPRECHER, E., BERGMAN, R., RICHARD, G., LURIE, R., SHALEV, S., PETRONIUS, D., SHALATA, A., ANBINDER, Y., LEIBU, R., PERLMAN, I. *et al.,* (2001). Hypotrichosis with juvenile macular dystrophy is caused by a mutation in CDH3, encoding P-cadherin. *Nat Genet* 29: 134-136.

STEFANSSON, I.M., SALVESEN, H.B. and AKSLEN, L.A. (2004). Prognostic impact of alterations in P-cadherin expression and related cell adhesion markers in endometrial cancer. *J Clin Oncol* 22: 1242-1252.

SUYAMA, K., SHAPIRO, I., GUTTMAN, M. and HAZAN, R.B. (2002). A signaling pathway leading to metastasis is controlled by N-cadherin and the FGF receptor. *Cancer Cell* 2: 301-314.

TANIUCHI, K., NAKAGAWA, H., HOSOKAWA, M., NAKAMURA, T., EGUCHI, H., OHIGASHI, H., ISHIKAWA, O., KATAGIRI, T. and NAKAMURA, Y. (2005). Overexpressed P-cadherin/CDH3 promotes motility of pancreatic cancer cells by interacting with p120ctn and activating rho-family GTPases. *Cancer Res* 65: 3092-3099.

THIERY, J.P., ACLOQUE, H., HUANG, R.Y. and NIETO, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. *Cell* 139: 871-890.

TSUTSUMIDA, A., HAMADA, J., TADA, M., AOYAMA, T., FURUUCHI, K., KAWAI, Y., YAMAMOTO, Y., SUGIHARA, T. and MORIUCHI, T. (2004). Epigenetic silencing of E- and P-cadherin gene expression in human melanoma cell lines. *Int J Oncol* 25: 1415-1421.

TURASHVILI, G., MCKINNEY, S.E., GOKTEPE, O., LEUNG, S.C., HUNTSMAN, D.G., GELMON, K.A., LOS, G., REJTO, P.A. and APARICIO, S.A. (2011). P-cadherin expression as a prognostic biomarker in a 3992 case tissue microarray series of breast cancer. *Mod Pathol* 24: 64-81.

VAN MARCK, V., STOVE, C., JACOBS, K., VAN DEN EYNDEN, G. and BRACKE, M. (2011). P-cadherin in adhesion and invasion: opposite roles in colon and bladder carcinoma. *Int J Cancer* 128: 1031-1044.

VAN MARCK, V., STOVE, C., VAN DEN BOSSCHE, K., STOVE, V., PAREDES, J., VANDER HAEGHEN, Y. and BRACKE, M. (2005). P-cadherin promotes cell-cell adhesion and counteracts invasion in human melanoma. *Cancer Res* 65: 8774-8783.

VISVADER, J.E. (2011). Cells of origin in cancer. *Nature* 469: 314-322.

YILMAZ, M. and CHRISTOFORI, G. (2010). Mechanisms of motility in metastasizing cells. *Mol Cancer Res* 8: 629-642.

ZHANG, C.C., YAN, Z., ZHANG, Q., KUSZPIT, K., ZASADNY, K., QIU, M., PAINTER, C.L., WONG, A., KRAYNOV, E., ARANGO, M.E. *et al.,* (2010). PF-03732010: a fully human monoclonal antibody against P-cadherin with antitumor and antimetastatic activity. *Clin Cancer Res* 16: 5177-5188.

Virchows Arch (2012) 460:545–553
DOI 10.1007/s00428-012-1237-8

REVIEW AND PERSPECTIVES

Cancer stem cell markers in breast neoplasias: their relevance and distribution in distinct molecular subtypes

Fernando Schmitt · Sara Ricardo · André Filipe Vieira · Maria Rita Dionísio · Joana Paredes

Received: 23 December 2011 / Revised: 29 March 2012 / Accepted: 10 April 2012 / Published online: 5 May 2012

Abstract The identification and characterization of cancer stem cells might lead to more effective control of neoplastic disease, by directing therapies to the most aggressive cells. For that reason, the identification of cancer stem cells (CSCs) in breast tumours is one of the priorities in breast cancer research, which has resulted in many studies attempting to identify their presence based on the expression of specific molecular markers. In this review, we describe the main molecular markers that have been identified as being able to recognise CSCs in breast carcinomas, the major molecular pathways that regulate CSCs and their association with the different molecular subtypes.

Keywords Breast cancer · Cancer stem cells · Stem cell markers

Abbreviations

ALDH	Aldehyde dehydrogenase
ATM	Ataxia telangiectasia mutated gene
BCSC	Breast cancer stem cell
Bmi-1	Polycomb group repressor protein
BRCA1	Breast cancer 1, early onset
CD24	Heat-stable antigen
CD29	β1-integrin
CD44	Receptor for hyaluronic acid
CD49f	α6-integrin
CD61	β3-integrin
CD133	Prominin 1
CK5/6	Cytokeratin 5/6
CSC	Cancer stem cell
DCIS	Ductal carcinoma in situ
ECM	Extracellular matrix
EMT	Epithelial-to-mesenchymal transition
ER	Estrogen receptor
ESA	Epithelial-specific antigen
GPI	Glycosylphosphatidylinositol
HER2	Human epidermal growth factor receptor 2
HER2-OE	HER2 overexpressing molecular subtype
Ki67	Cellular proliferation marker
LOH	Loss of heterozygosity
mRNA	Messenger RNA
NOD/SCID	Non-obese diabetic/severe combined immunodeficient mice
SP	Side population
TIC	Tumour-initiating cell

F. Schmitt (✉) · S. Ricardo · A. F. Vieira · M. R. Dionísio · J. Paredes
IPATIMUP–Institute of Molecular Pathology and Immunology of Porto University,
Rua Dr. Roberto Frias s/n,
4200-465 Porto, Portugal
e-mail: fernando.schmitt@ipatimup.pt

F. Schmitt · J. Paredes
Medical Faculty of Porto University,
Porto, Portugal

S. Ricardo · A. F. Vieira
ICBAS–Abel Salazar Biomedical Science Institute,
Porto, Portugal

M. R. Dionísio
Gulbenkian Programme for Advanced Medical Education,
Lisbon, Portugal

Introduction

Breast cancer is a heterogeneous disease consisting of a growing number of biologically distinct subtypes. This

heterogeneity is reflected by differences in hormone receptor and human epidermal growth factor receptor 2 (HER2) expression status, diverse histological subtypes, as well as distinct biological behaviour, response to therapy and disease outcome [1]. This heterogeneity implies that in order to plan for the optimally effective therapy of each breast cancer subtype, it has become crucial to identify those cancer cells in a tumour that drive disease progression. The prevailing concept for many years was that the majority of the cells in a tumour have the potential to extensively proliferate and metastasize and should all be eliminated in order to effectively treat the disease. However, recent studies have suggested that the ability of tumours to proliferate, progress and propagate relies on a limited subpopulation of cells with stem cell-like properties, for which the term cancer stem cells (CSCs) has become mainstream [2]. These cancer stem cells share essential characteristics with normal adult stem cells, such as the capacity to self-renew through asymmetric cell division, producing one tumorigenic CSC and one non-tumorigenic cancer cell destined for terminal differentiation. This concept implies that cells in a cancer are hierarchically organised in a manner very similar to cells in a normal tissue: the progeny of a normal stem cell differentiates into phenotypically diverse cells with limited proliferative potential, and so, the progeny of CSCs also sustains genetic/epigenetic alterations analogous to the differentiation of normal cells, resulting in a population of phenotypically diverse non-tumorigenic cancer cells that compose the bulk of the tumour (Fig. 1a). These alterations are associated with loss of tumorigenic capacity, implying that the vast majority of cells in such carcinomas no longer contribute to disease progression [3].

Cancer stem cell model: rediscovery of an old story

Stem cells are present in many different somatic tissues and are important participants in their normal physiology. The progenitor cell and the mature cell populations that derive from it are organised in a hierarchical fashion, with the stem cell residing at the apex of the differentiation pathway [2]. Stem cells have three distinctive properties: (a) self-renewal (i.e. at cell division, one or both daughter cells retain the same biological properties as the parental cell); (b) the capacity to generate multiple cell lineages; and (c) the potential for sustained proliferation. Indeed, the attribute of self-renewal is especially notable because its subversion is highly relevant to oncogenesis and malignancy [4, 5]. Therefore, malignant cells harbouring the three features that define normal stem cells have been termed "cancer stem cells" (Fig. 1a).

Even though, recently, the awareness of this concept in cancer research has greatly expanded, it is not new. In 1855, Rudolf Virchow proposed the embryonal rest hypothesis, stating that cancer arises from the activation of "dominant" cells present in mature tissues that are remaining embryonic cells [6]. This theory, based on the morphological similarities between developing foetal cells and some cancer cells, was the first description of what we now call the CSC. Later, in 1875, this theory was expanded by other pathologists such as Julius Cohnheim, who proposed that misplacement of stem cells during embryonic development could lead later on in life to the development of tumours [7].

Back in 1937, Furth and Kahn established that engrafting of a single cell from a mouse tumour could initiate a new tumour in the recipient mouse [8]. The resulting tumours typically showed the same morphologic heterogeneity as the original tumour. Later, Kleinsmith and Pierce showed that malignant teratocarcinomas contain highly tumorigenic cells that, as single cells, have the capacity to differentiate into multiple mature, non-tumorigenic cell types [9]. The conclusion was drawn that teratocarcinomas tend to follow the pathways of development in embryonal tissues. To corroborate the previous findings, the same authors showed by radioactive pulse labelling of proliferating cells that these occurred almost exclusively in undifferentiated areas. At later time points, however, the DNA label appeared also in well-differentiated areas, which had to be derived from earlier labelled undifferentiated cells. These well-differentiated cells did not form tumours when transplanted into compatible recipients [10]. These results, among others, led Pierce to formulate the following early definition of the CSC concept [11]: "carcinomas are caricatures of tissue renewal, in that they are composed of a mixture of malignant stem cells".

Finally, in 1994, the first human CSC, the leukaemia stem cell, was identified in samples from patients with acute myeloid leukaemia [12] and it was proposed that blocked differentiation ("maturation arrest") was responsible for CSC proliferation and tumour growth [13]. Three years later, Blair and collaborators as well as Bonnet and Dick showed that only a small percentage of acute myeloid leukaemia cells (0.002–1 %), which phenotypically resembled hematopoietic stem cells, could transfer disease when transplanted into rodent recipients [14, 15]. These lines of evidence supported the notion that many, if not all cancers, depend on a small population of CSCs for their sustained growth and expansion. Since then, several investigators have documented the existence of CSCs in brain [16], breast [17], prostate [18] and lung cancers [19], among others.

Breast cancer stem cell markers

Finding breast CSCs (BCSCs) has been an important goal for many breast cancer researchers, who were trying to

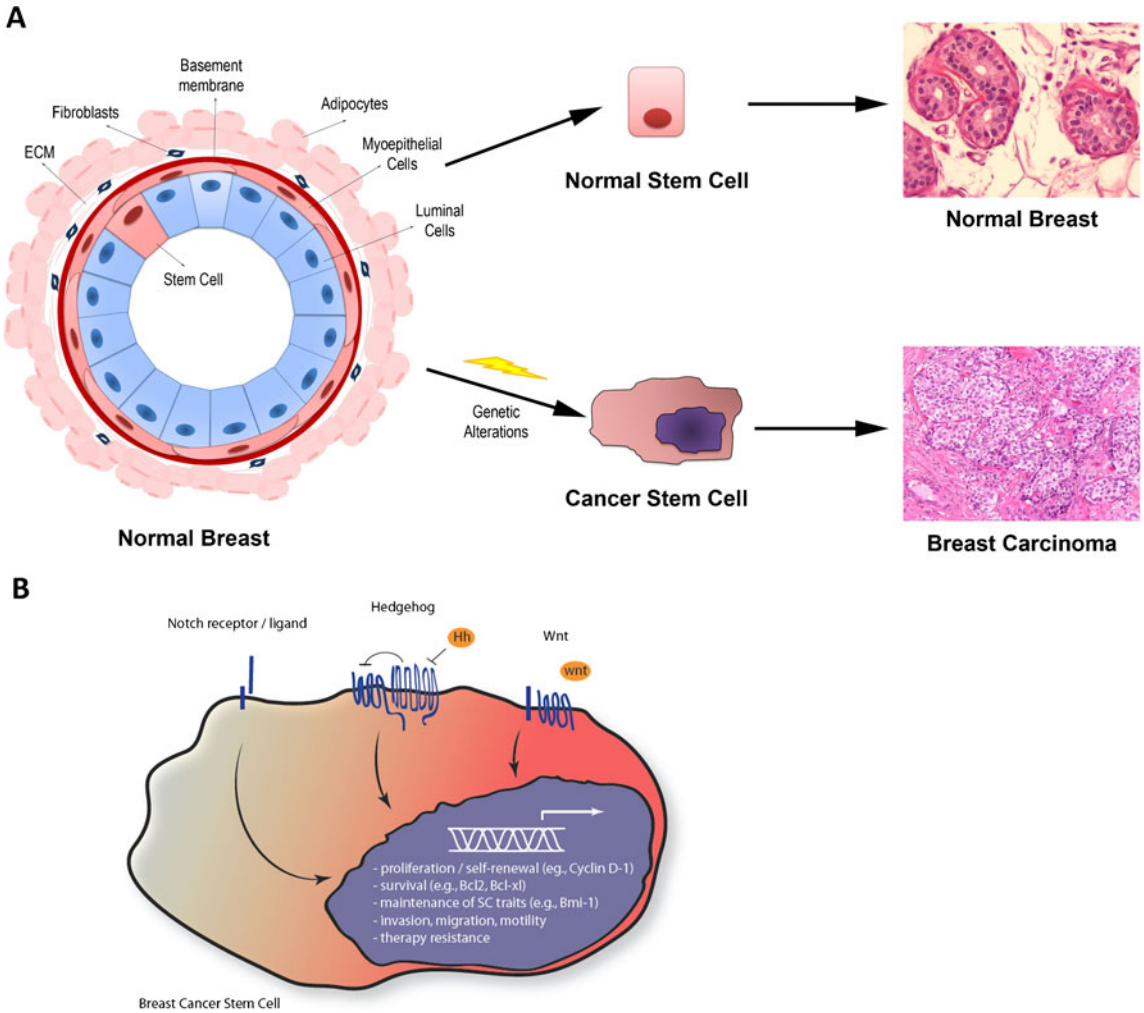


Fig. 1 **a** Schematic representation of the cellular components of a normal duct of the mammary gland. There is an outer layer of myoepithelial cells (*red*) surrounding an inner layer of luminal epithelial cells (*blue*). Stem cells of the normal mammary gland have two main distinctive properties: self-renewal and the capacity to develop into both lineages of fully breast-differentiated cells. Malignant cells harbouring the same features that define normal stem cells have been termed cancer stem cells, which originate due to the accumulation of genetic (or epigenetic) alterations that target the cellular components of the breast. **b** The three major signalling pathways involved in the maintenance of stem cell features in the breast cancer stem cell. The Notch, Wnt and Hedgehog molecules are well-known mediators of normal stem cell biology, which also mediate downstream effects in the transcription profile of cancer stem cells. These effects include the control of proliferation/self-renewal, survival, invasion and therapy resistance. Cancer stem cell therapy may encompass the modulation of these signalling pathways, which are over-activated in CSCs, without affecting normal breast physiology

define a universal marker or combination of markers able to specifically identify these cells in breast tumours and ultimately isolate them. Reviewing the literature, it appears that the most studied BCSC markers are the transmembrane proteins CD44 and CD24. A subpopulation of tumour cells that strongly express CD44 but not CD24 (the $CD44^{+}CD24^{-/low}$ phenotype) was identified as CSCs by Al-Hajj and collaborators [5], and this was subsequently confirmed by other authors. CD44 is a transmembrane glycoprotein, of which several isoforms exist, that normally regulates cell–cell adhesion and cell–matrix interactions, as well as cell migration. This glycoprotein binds mainly to hyaluronic acid, as well as to collagen, fibronectin, laminin and chondroitin sulphate—all important components of the ECM. It also binds the cytokine osteopontin [20]. Members of the CD44 family differ in their extracellular domain by the insertion of variable regions through alternative splicing [21]. The gene-encoding CD44 consists of 20 exons. In the standard form (CD44s), 10 of the 20 exons are transcribed. Multiple variant isoforms (CD44v1–v10) arise from alternative messenger RNA (mRNA) splicing of the other ten exons [22]. In contrast to the standard form of CD44, which is usually ubiquitously expressed on epithelial cells and lymphocytes, CD44 variants exhibit tissue-specific expression. Some of these variants, in particular splice variant CD44v6, are associated with aggressive tumour behaviour

in that their expression correlates with poor prognosis in a variety of human malignancies including breast cancer [23]. CD24 is a small, heavily glycosylated mucin-type protein, which is linked to the cell membrane via a GPI anchor. This molecule is involved in the regulation of cell proliferation and cell–cell interactions and it was shown to be expressed by normal pre-B lymphocytes (which lose its expression during the maturation into plasma cells), as well as in various haematological malignancies and solid tumours of some organs [24]. CD24 is the ligand of P-selectin and also an adhesion receptor expressed on activated endothelial cells and platelets, and this has led to the suggestion that it might play an important role in the metastatic process [25, 26]. Using a combination of these two cell surface markers, Al-Hajj and collaborators were the first to distinguish cancer cells that were tumorigenic in immunocompromised (non-obese diabetic/severe combined immunodeficient (NOD/SCID)] mice from non-tumorigenic cells, in a small number of breast tumours. In flow sorting experiments, they used the ESA as an epithelial cell marker with $CD44^{+}/CD24^{-/low}$ as CSC marker combination, after eliminating the non-epithelial cells that were stained for their lineage-specific markers (hematopoietic and endothelial) [5]. Since then, many more breast carcinomas have been reported to contain a subpopulation of $CD44^{+}/CD24^{-/low}$ cancer cells, which are capable of generating tumours in the NOD/SCID mice, even when implanted in very low numbers. In contrast, other cancer cell populations fail to generate tumours, even when implanted in high numbers. These reports therefore established tumourigenicity and self-renewal potential of these cells, in vitro and in vivo [27, 28].

Even though some clinical studies confirmed that $CD44^{+}/CD24^{-/low}$-expressing tumours have a poor prognosis [29–31], controversy remains concerning this issue [27, 32]. Shipitsin et al. demonstrated that genes specifically expressed in $CD44^{+}$ cells, among which many known stem cell markers, identified carcinomas with poor patient survival [33], suggesting that $CD44^{+}$ expression is prognostically relevant and justifying its consideration as a new therapeutic target for breast cancer. In contrast, Mylona et al. [32] observed that breast cancers with the opposite $CD44^{-}/CD24^{+}$ phenotype are associated with poor patient prognosis, in stark contrast with the CSC $CD44^{+}/CD24^{-/low}$ phenotype. Furthermore, Abraham et al. failed to confirm that the occurrence of $CD44^{+}/CD24^{-/low}$ tumour cells in breast cancer is associated with worse survival [27]. These contradictory data demanded additional efforts to find other markers that could complement the CSC markers CD44 and CD24, to arrive at an improved correlation with patient survival.

Along these lines, a new marker proposed for the identification of BCSCs is aldehyde dehydrogenase 1 (ALDH1). Ginestier et al. showed that normal human mammary epithelial cells and breast cancer cells with high ALDH1 activity have stem/progenitor cell properties [34]. This finding offered an important new potential molecular target for the study of breast CSCs. It was subsequently found that $ALDH1^{+}$ cells (putative CSCs) are significantly more resistant to platinum treatment and are biologically more aggressive, and that $ALDH1^{+}$ cancers are associated with poor patient prognosis [34–36]. Unlike the $CD44^{+}CD24^{-/low}$ phenotype, the $ALDH1^{+}$ phenotype consistently correlates with worse clinical outcome and with certain histological and clinical characteristics, such as high tumour grade, HER2 positivity and Ki67 proliferation status [37]. In addition, the ALDH1 marker can further divide the $CD44^{+}CD24^{-/low}$ cell population into fractions that are highly tumorigenic: while as few as 20 xenografted $ALDH1^{+}CD44^{+}CD24^{-/low}$ cells resulted in tumours, at this cell density, $ALDH1^{+}CD44^{+}CD24^{-/low}$ cells were not tumorigenic [34, 38]. Taken together, this evidence supports the notion that $ALDH1^{-}CD44^{+}CD24^{-/low}$ cells in a neoplastic population represent CSC-like cells. However, the fraction of $ALDH1^{+}$ cells in the $CD44^{+}CD24^{-/low}$ population in primary tumours is quite small (1.16 %) [34].

Another candidate marker for the BCSC phenotype is CD133, also known as Prominin-1 because of its location on cell membrane protrusions. Although initially considered a marker specific for haematopoietic stem cells, CD133 mRNA transcripts are also found in normal bone marrow and in a variety of other tissues [39]. Furthermore, CD133 is expressed in various solid tumours, including triple-negative invasive breast carcinomas, but more restricted in expression in comparison with other previously mentioned CSC markers, such as CD44 and ALDH1. In early-onset breast cancer 1 (BRCA1)-associated breast cancer cell lines, $CD133^{+}$-sorted cells have CSC properties, including greater colony-forming efficiency, higher proliferative activity and higher tumourigenicity in NOD/SCID mice [40]. Moreover, the expression of CD133 has been reported in the majority of inflammatory breast cancers, as well as in triple-negative tumours [41, 42].

Other markers, such as CD29 (β1-integrin), CD49f (α6-integrin) and CD61 (β3-integrin), have been used for the isolation of mammary CSCs. Vassilopoulos et al. [43] used CD24/CD29 or CD24/CD49f to identify a subpopulation of mammary tumour cells, while Vaillant et al. [44] found that CD61 identifies a cell subpopulation highly enriched for tumorigenic capacity in comparison with CD61-negative cells.

Complementary to the use of biomarkers to identify or isolate putative CSCs, by flow cytometry or immunolabeling using CD44 and CD24 antibodies or ALDH1 activity (using the ALDEFLUOR assay), functional assays have been extensively used for this purpose (Fig. 2). One of these methods is based on the capacity of purported stem cells to

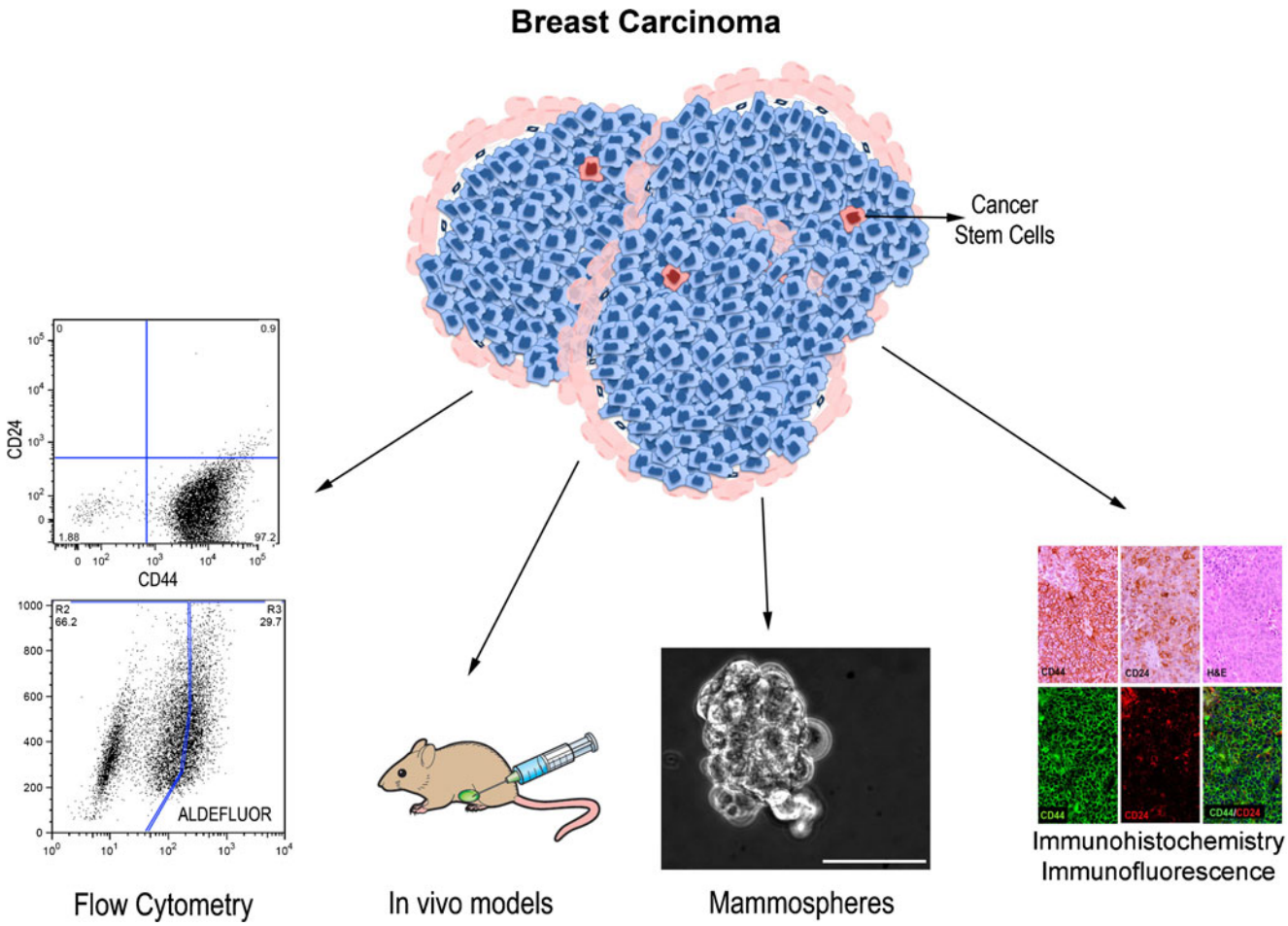


Fig. 2 Breast cancer stem cell identification and/or isolation rely in several breast CSC biomarkers and methodologies, including flow cytometry, in vivo tumorigenic models, mammospheres and histological immunostaining. Flow cytometry allows the identification and isolation of cancer cell subpopulations, using cell surface CSC markers, such as CD44/CD24, and ALDH1 activity (ALDEFLUOR assay). The mammosphere assay measures the survival of cancer cells in anchorage-independent conditions and constitutes an indicator of in vitro self-renewal capacity. Immunohistochemistry or immunofluorescence depend on the use of specific antibodies against breast CSC markers (such as CD44 and CD24), in order to detect the CSC phenotype in fixed paraffin-embedded tissue samples. The tumourigenic capability of CSCs subpopulations in immune-deficient mice allows proving the cancer stem cell properties in vivo (author's original images)

efflux a lipophilic fluorescent dye (Hoechst 33342), through the same mechanism they use to expulse drugs and which confers drug resistance. This dye excluding cell population has been called "side population" (SP). This method was initially described to isolate haematopoietic stem cells and was subsequently applied to breast cancer cell lines [45]. However, unsolved issues with potential toxicity of this dye for non-SP cells hinder the further application of this functional assay for the identification of BCSC subpopulations. Another important functional assay, used for breast epithelial stem cells, is the mammosphere-forming assay, in which mammary epithelial cells are cultured in a serum-free medium on a surface to which they do not adhere and are thus induced to form a three-dimensional cell cluster which is called mammosphere (Fig. 2). Pointi et al. employed a similar approach to derive mammospheres from human breast cancers and found CSCs within the cancer cell-derived mammospheres with the same phenotype as previously reported by Al-Hajj ($CD44^{+}CD24^{-/low}$) [46]. Since then, several other markers associated with a stem cell phenotype, such as components of the Notch and Wnt signalling pathways, have been found overexpressed in mammospheres from breast cancer cell lines [47].

Signalling pathways in breast cancer stem cells

Cancer stem cells share molecular markers with normal stem cells, but in addition, they show aberrant activation of self-renewal pathways, such as Wnt, Hedgehog and Notch, also operative in normal stem cells. Normal human mammary stem cells express high levels of the Notch-4

receptor, which are assumed to play an activating role in the stem cell compartment. Notch-4 and Notch-1 are involved in normal development of the mammary gland and mutated forms of these receptors are associated with the development of mouse mammary tumours [48]. Aberrant activation of Notch signalling is probably an early event in breast cancer, since this has been found in ductal carcinoma in situ (DCIS). High expression of Notch-1 in breast cancer is associated with worse prognosis, and a gamma-secretase inhibitor which blocks Notch signalling or a Notch-4-neutralising antibody reduced mammosphere formation from DCIS [48]. Notch-4 has been shown to be important for the self-renewal of cancer stem cells, and in xenotransplantation assays, it played an active role in tumour formation [49]. The Notch pathway activation and effects are heterogeneous and depend on the cellular context and on the in vivo model used. However, in both man and mice, the Notch pathway stimulates differentiation of mammary stem cells towards the luminal compartment. In man, this luminal cell fate determination is mediated by Notch-3, whereas in mice, Notch-1 seems to be more important [48, 50].

Stem cell self-renewal, maintenance and cell fate are also controlled by Hedgehog signalling in different tissues. In the mammary gland, strict regulation of the Hedgehog pathway is required for normal development and alterations in Hedgehog signalling cause defects during embryonic development resulting in an abnormal mammary gland [51]. Hedgehog signalling is activated in $CD44^{+}CD24^{-/low}$ human breast cancer stem cells [51], and stemness, measured by mammosphere-forming capacity, is increased upon the activation of this pathway [52]. Conversely, inhibition of Hedgehog signalling results in a reduction of the number of mammosphere-initiating cells and mammosphere size, effects that are mediated by Bmi-1 [51], which suggests that the Hedgehog pathway might provide new therapeutic targets.

Elucidation of Wnt signalling is complex, given the fact that there are 19 Wnt ligands, several Wnt receptors and involvement of this signalling pathway in embryogenesis, development, differentiation and proliferation. Inappropriate Wnt/β-catenin signalling can result in deregulated self-renewal. The activation of Wnt-1 causes expansion of the mammary stem cell compartment and an aberrant population of progenitor cells that have in vivo regenerative ability [44, 53]. Since in the presence of inappropriate Wnt signalling mammary glands express high levels of basal markers, Wnt is thought to induce dedifferentiation of mammary cells and confer a more stem-like phenotype. A role for Wnt signalling in the development of breast cancers of the basal-like subtype, and also the predictive capacity of Wnt signalling for brain metastasis, has been reported and the ligands Wnt5a/b have been identified as highly relevant for this subtype of breast cancer [54].

In addition to these main pathways, represented in Fig. 1b, many others have been implicated in the determination of breast stem cell fate. These pathways, when deregulated in cancer, might present new therapeutic targets. Molecules that interfere with these pathways should be tested in vitro with regard to their effect on the CSC compartment, and when found effective, subsequently in in vivo xenografted animal models of human cancer, before testing in patient tumours.

Cancer stem cells in molecular subtypes of breast cancer

In breast cancer, it is assumed that CSCs generate cells with aberrant and limited differentiation, which through mechanisms as yet unknown translate into the distinct breast cancer molecular subtypes. In this concept, each molecular subtype putatively is composed of cancer cells with a different level of differentiation; CSCs being most prevalent in the least differentiated molecular subtype (Fig. 3). Recently, our group demonstrated that of the molecular subtypes, the basal-like type harbours the highest percentage of tumour cells with a $CD44^{+}CD24^{-/low}$ and $ALDH1^{+}$ cancer stem cell phenotype [31]. This confirms earlier studies using breast cancer cell lines, which reported enrichment of $CD44^{+}/CD24^{-/low}$ and $CD44^{-}/CD24^{+}$ cell populations in basal-like and luminal molecular subtypes, respectively [55, 56];

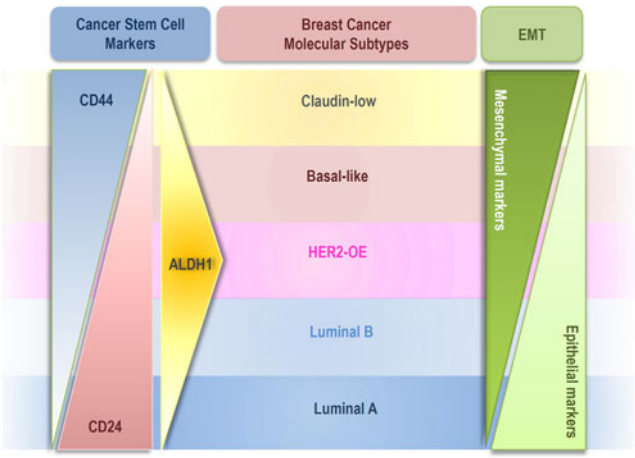


Fig. 3 Representation of the putative association between breast cancer molecular subtypes and the expression pattern of breast CSC biomarkers. In this hypothesis, each molecular subtype is enriched in cancer cells with distinct levels of differentiation, being CSCs more prevalent within the more undifferentiated molecular breast cancers, such as claudin-low tumours. Basal-like tumours have a phenotype close to the CSC phenotype but with an increase in the epithelial phenotype. This subtype of breast cancer and the Her-2 overexpressing molecular subtype have increased ALDH activity, a marker of the luminal progenitor of the breast. Luminal A and luminal B carcinomas present a mature differentiated cell profile, with the lowest expression of mesenchymal markers and an epithelial $CD24^{+}$ phenotype

the $CD44^+$ cells being more stem cell-like and $CD24^+$ cells more differentiated [37]. This is corroborated by data published by Honeth and collaborators, who showed that the $CD44^+/CD24^{-/low}$ CSC phenotype is significantly associated with basal-like breast cancers in human patients and in particular with BRCA1-inherited cancers [29]. However, ALDH1 is not associated with a particular molecular subtype of breast cancer [34, 35]. Even though $CD44^+CD24^{-/low}$ cells and $ALDH1^+$ cells are more frequently found in basal-like than in luminal tumours, $ALDH1^+$ cells are also commonly found in the HER2-overexpressing (HER2-OE) subtype [31]. This difference between basal-like and HER2-OE tumours may be due to a different cell-of-origin or to the induction of a stem-like transcriptional programme as a result of specific transforming genetic alterations (Fig. 3). In line with this notion, induced overexpression of HER2 expression in breast cancer cell lines increased the number of ALDEFLUOR-positive cells [57].

Moreover, basal-like tumours with a predominantly $CD44^+CD24^{-/low}$ CSC phenotype show a tendency towards worse prognosis, in agreement with earlier studies demonstrating an association between basal-like carcinomas and the $CD44^+CD24^{-/low}$ CSC phenotype [29, 37, 55, 56]. These results contribute to our understanding of the morphological and molecular heterogeneity of breast cancer. Furthermore, they suggest that the CSC phenotype may identify the cell-of-origin of the tumour, rather than the unique population of responsible for tumour progression, since the highly aggressive HER2-overexpressing tumours do not show increased expression of these markers.

Since basal type breast cancers have an increased stem cell-like population in comparison with other molecular subtypes, it has been hypothesised that in these tumours, a differentiation block exists, as the majority of the cells constituting the bulk of the tumour do not display a differentiated phenotype [51, 58]. The transformation of a basal-like mammary stem cell to a basal-like cancer cell may either be due to (a) LOH as a second hit in BRCA1 mutation carriers or (b) to the downregulation of BRCA1 expression in estrogen receptor (ER)-negative stem cells. This might result in the arrest of the luminal differentiation process, which fixes the cells in the undifferentiated phenotype of basal-like carcinomas with its unique expression of basal cytokeratins (CK5/6) [59]. Presumably, this goes along with an increment of telomerase activity, since BRCA1 suppresses telomerase expression, and in DNA repair defects, which generate inactivation or loss of the ataxia telangiectasia mutated (ATM) gene and as a result increased genomic instability, which may be the reason for the greater number of genomic gains and losses in sporadic and familial basal-like cancers [60].

Recently, a tumour-initiating cell (TIC) genomic signature, derived from $CD44^+/CD24^{-/low}$-sorted cells and mammospheres obtained from primary human breast tumours, was found to be exclusively enriched in a new molecular subtype referred to as claudin-low [61]. This new molecular subtype is characterised by the low expression of genes involved in tight junctions and cell–cell adhesion, including claudins 3, 4 and 7, Occludin and E-cadherin. It also displays epithelial-to-mesenchymal transition (EMT) features, such as high levels of expression of vimentin, Snail-1, Snail-2, TWIST1, TWIST2, ZEB1 and ZEB2. In addition, these tumours show low expression of luminal genes, inconsistent basal gene expression, high expression of lymphocyte and endothelial cell markers and lack of ER expression, and they were all diagnosed as grade II or III infiltrating ductal carcinomas. These tumour characteristics reflect a low level of tumour differentiation and fit with the CSC and EMT phenotypes [61].

Conclusion

The CSC model proposes that in tumours, like in normal tissues, a cellular hierarchy exists in which CSCs are the only cells endowed with an unlimited potential for proliferation and, as a consequence, capable of driving tumour growth and metastasis due to their stem cell-like characteristics [51]. The presently available markers that label cells with stem cell properties in tumours should be further tested for specificity, and additional markers, maybe in the form of specific gene signatures, are needed to further define and potentially target the CSC population of different breast cancer subtypes. However, in the design of novel effective anti-BCSC therapy strategies, the possibility that (a) marker expression is downregulated/silenced, (b) marker-negative BCSC cell clones evolve from marker-positive cell clones and (c) different BCSC subpopulations coexist in a single tumour or at different metastatic sites should be taken into consideration. Probably, in the future, it will be necessary to use a panel of BCSC markers, in order to effectively translate knowledge of CSCs into breast cancer benefit.

Acknowledgments This work was supported by a grant (project 96633–2010) from Fundação Gulbenkian, Portugal. The authors would also like to acknowledge Fundação Luso Britânica and Fundação para a Ciência e Tecnologia (FCT), Portugal, for fellowship support (A.F.V. was supported by SFRH/BD/41363/2007) and for the Programa Ciência 2007 (J.P.) [POPH-QREN-Tipology 4.2]. The Programme for Advanced Medical Education is sponsored by Fundação Gulbenkian, Fundação Champalimaud, Portuguese Ministry of Science, Technology and Higher Education and FCT, Portugal. IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and is partially supported by FCT.

Conflict of interest We declare that we have no conflict of interest.

References

1. Weigelt B, Geyer FC, Reis-Filho JS (2010) Histological types of breast cancer: how special are they? Mol Oncol 4:192–208
2. Jordan CT, Guzman ML, Noble M (2006) Cancer stem cells. N Engl J Med 355:1253–1261
3. Shackleton M, Quintana E, Fearon ER, Morrison SJ (2009) Heterogeneity in cancer: cancer stem cells versus clonal evolution. Cell 138:822–829
4. Pardal R, Clarke MF, Morrison SJ (2003) Applying the principles of stem-cell biology to cancer. Nat Rev Cancer 3:895–902
5. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF (2003) Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci USA 100:3983–3988
6. Virchow R (1855) Archiv fur Pathologische Anatomie und Physiologie fur Klinische Medizin, Volume 8.
7. Cohnheim V (1875) Congenitales, quergestreiftes Muskelsarkom der Nieren. Virchows Arch Pathol Anat Physiol Klin Med 65:64–69
8. Furth JK, Kahn MC (1937) The transmission of leukemia of mice with a single cell. Am J Cancer 31:276–282
9. Kleinsmith LJ, Pierce GB Jr (1964) Multipotentiality of single embryonal carcinoma cells. Cancer Res 24:1544–1551
10. Pierce GB, Wallace C (1971) Differentiation of malignant to benign cells. Cancer Res 31:127–134
11. Pierce GB, Speers WC (1988) Tumors as caricatures of the process of tissue renewal: prospects for therapy by directing differentiation. Cancer Res 48:1996–2004
12. Lapidot T, Sirard C, Vormoor J, Murdoch B, Hoang T, Caceres-Cortes J, Minden M, Paterson B, Caligiuri MA, Dick JE (1994) A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature 367:645–648
13. Sell S, Pierce GB (1994) Maturation arrest of stem cell differentiation is a common pathway for the cellular origin of teratocarcinomas and epithelial cancers. Lab Invest 70:6–22
14. Bonnet D, Dick JE (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med 3:730–737
15. Blair A, Hogge DE, Ailles LE, Lansdorp PM, Sutherland HJ (1997) Lack of expression of Thy-1 (CD90) on acute myeloid leukemia cells with long-term proliferative ability in vitro and in vivo. Blood 89:3104–3112
16. Singh SK, Hawkins C, Clarke ID, Squire JA, Bayani J, Hide T, Henkelman RM, Cusimano MD, Dirks PB (2004) Identification of human brain tumour initiating cells. Nature 432:396–401
17. Dontu G, Al-Hajj M, Abdallah WM, Clarke MF, Wicha MS (2003) Stem cells in normal breast development and breast cancer. Cell Prolif 36(Suppl 1):59–72
18. Xin L, Lawson DA, Witte ON (2005) The Sca-1 cell surface marker enriches for a prostate-regenerating cell subpopulation that can initiate prostate tumorigenesis. Proc Natl Acad Sci USA 102:6942–6947
19. Kim CF, Jackson EL, Woolfenden AE, Lawrence S, Babar I, Vogel S, Crowley D, Bronson RT, Jacks T (2005) Identification of bronchioalveolar stem cells in normal lung and lung cancer. Cell 121:823–835
20. Orian-Rousseau V (2010) CD44, a therapeutic target for metastasising tumours. Eur J Cancer 46:1271–1277
21. Ponta H, Sherman L, Herrlich PA (2003) CD44: from adhesion molecules to signalling regulators. Nat Rev Mol Cell Biol 4:33–45
22. Tolg C, Hofmann M, Herrlich P, Ponta H (1993) Splicing choice from ten variant exons establishes CD44 variability. Nucleic Acids Res 21:1225–1229
23. Afify A, Purnell P, Nguyen L (2009) Role of CD44s and CD44v6 on human breast cancer cell adhesion, migration, and invasion. Exp Mol Pathol 86:95–100
24. Lim SC (2005) CD24 and human carcinoma: tumor biological aspects. Biomed Pharmacother 59(Suppl 2):S351–S354
25. Aigner S, Ramos CL, Hafezi-Moghadam A, Lawrence MB, Friederichs J, Altevogt P, Ley K (1998) CD24 mediates rolling of breast carcinoma cells on P-selectin. FASEB J 12:1241–1251
26. Kim HJ, Kim JB, Lee KM, Shin I, Han W, Ko E, Bae JY, Noh DY (2007) Isolation of CD24(high) and CD24(low/-) cells from MCF-7: CD24 expression is positively related with proliferation, adhesion and invasion in MCF-7. Cancer Lett 258:98–108
27. Abraham BK, Fritz P, McClellan M, Hauptvogel P, Athelogou M, Brauch H (2005) Prevalence of CD44+/CD24-/low cells in breast cancer may not be associated with clinical outcome but may favor distant metastasis. Clin Cancer Res 11:1154–1159
28. Snyder EL, Bailey D, Shipitsin M, Polyak K, Loda M (2009) Identification of CD44v6(+)/CD24− breast carcinoma cells in primary human tumors by quantum dot-conjugated antibodies. Lab Invest 89:857–866
29. Honeth G, Bendahl PO, Ringner M, Saal LH, Gruvberger-Saal SK, Lovgren K, Grabau D, Ferno M, Borg A, Hegardt C (2008) The CD44+/CD24− phenotype is enriched in basal-like breast tumors. Breast Cancer Res 10:R53
30. Buess M, Rajski M, Vogel-Durrer BM, Herrmann R, Rochlitz C (2009) Tumor-endothelial interaction links the CD44(+)/CD24(−) phenotype with poor prognosis in early-stage breast cancer. Neoplasia 11:987–1002
31. Ricardo S, Vieira AF, Gerhard R, Leitao D, Pinto R, Cameselle-Teijeiro JF, Milanezi F, Schmitt F, Paredes J (2011) Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtypes. J Clin Pathol 64:937–946
32. Mylona E, Giannopoulou I, Fasomytakis E, Nomikos A, Magkou C, Bakarakos P, Nakopoulou L (2008) The clinicopathologic and prognostic significance of CD44+/CD24(−/low) and CD44-/CD24 + tumor cells in invasive breast carcinomas. Hum Pathol 39:1096–1102
33. 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 SJ, Nikolsky Y, Gelman RS, Polyak K (2007) Molecular definition of breast tumor heterogeneity. Cancer Cell 11:259–273
34. 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, Dontu G (2007) 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
35. Morimoto K, Kim SJ, Tanei T, Shimazu K, Tanji Y, Taguchi T, Tamaki Y, Terada N, Noguchi S (2009) 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 100:1062–1068
36. Deng S, Yang X, Lassus H, Liang S, Kaur S, Ye Q, Li C, Wang LP, Roby KF, Orsulic S, Connolly DC, Zhang Y, Montone K, Butzow R, Coukos G, Zhang L (2010) Distinct expression levels and patterns of stem cell marker, aldehyde dehydrogenase isoform 1 (ALDH1), in human epithelial cancers. PLoS One 5:e10277
37. Park SY, Lee HE, Li H, Shipitsin M, Gelman R, Polyak K (2010) Heterogeneity for stem cell-related markers according to tumor subtype and histologic stage in breast cancer. Clin Cancer Res 16:876–887
38. Croker AK, Goodale D, Chu J, Postenka C, Hedley BD, Hess DA, Allan AL (2009) 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 13:2236–2252
39. Miraglia S, Godfrey W, Yin AH, Atkins K, Warnke R, Holden JT, Bray RA, Waller EK, Buck DW (1997) A novel five-transmembrane

hematopoietic stem cell antigen: isolation, characterization, and molecular cloning. Blood 90:5013–5021
40. Wright MH, Calcagno AM, Salcido CD, Carlson MD, Ambudkar SV, Varticovski L (2008) Brca1 breast tumors contain distinct CD44+/CD24− and CD133+ cells with cancer stem cell characteristics. Breast Cancer Res 10:R10
41. Xiao Y, Ye Y, Yearsley K, Jones S, Barsky SH (2008) The lymphovascular embolus of inflammatory breast cancer expresses a stem cell-like phenotype. Am J Pathol 173:561–574
42. Zhao P, Lu Y, Jiang X, Li X (2011) Clinicopathological significance and prognostic value of CD133 expression in triple-negative breast carcinoma. Cancer Sci 102:1107–1111
43. Vassilopoulos A, Wang RH, Petrovas C, Ambrozak D, Koup R, Deng CX (2008) Identification and characterization of cancer initiating cells from BRCA1 related mammary tumors using markers for normal mammary stem cells. Int J Biol Sci 4:133–142
44. Vaillant F, Asselin-Labat ML, Shackleton M, Forrest NC, Lindeman GJ, Visvader JE (2008) The mammary progenitor marker CD61/beta3 integrin identifies cancer stem cells in mouse models of mammary tumorigenesis. Cancer Res 68:7711–7717
45. Hirschmann-Jax C, Foster AE, Wulf GG, Nuchtern JG, Jax TW, Gobel U, Goodell MA, Brenner MK (2004) A distinct "side population" of cells with high drug efflux capacity in human tumor cells. Proc Natl Acad Sci USA 101:14228–14233
46. Ponti D, Costa A, Zaffaroni N, Pratesi G, Petrangolini G, Coradini D, Pilotti S, Pierotti MA, Daidone MG (2005) Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. Cancer Res 65:5506–5511
47. Rappa G, Lorico A (2010) Phenotypic characterization of mammosphere-forming cells from the human MA-11 breast carcinoma cell line. Exp Cell Res 316:1576–1586
48. Raouf A, Zhao Y, To K, Stingl J, Delaney A, Barbara M, Iscove N, Jones S, McKinney S, Emerman J, Aparicio S, Marra M, Eaves C (2008) Transcriptome analysis of the normal human mammary cell commitment and differentiation process. Cell Stem Cell 3:109–118
49. Harrison H, Farnie G, Howell SJ, Rock RE, Stylianou S, Brennan KR, Bundred NJ, Clarke RB (2010) Regulation of breast cancer stem cell activity by signaling through the Notch4 receptor. Cancer Res 70:709–718
50. Bouras T, Pal B, Vaillant F, Harburg G, Asselin-Labat ML, Oakes SR, Lindeman GJ, Visvader JE (2008) Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. Cell Stem Cell 3:429–441
51. Charafe-Jauffret E, Monville F, Ginestier C, Dontu G, Birnbaum D, Wicha MS (2008) Cancer stem cells in breast: current opinion and future challenges. Pathobiology 75:75–84
52. Moraes RC, Zhang X, Harrington N, Fung JY, Wu MF, Hilsenbeck SG, Allred DC, Lewis MT (2007) Constitutive activation of smoothened (SMO) in mammary glands of transgenic mice leads to increased proliferation, altered differentiation and ductal dysplasia. Development 134:1231–1242
53. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat ML, Wu L, Lindeman GJ, Visvader JE (2006) Generation of a functional mammary gland from a single stem cell. Nature 439:84–88
54. Klemm F, Bleckmann A, Siam L, Chuang HN, Rietkotter E, Behme D, Schulz M, Schaffrinski M, Schindler S, Trumper L, Kramer F, Beissbarth T, Stadelmann C, Binder C, Pukrop T (2011) β-catenin-independent WNT signaling in basal-like breast cancer and brain metastasis. Carcinogenesis 32:434–442
55. Fillmore CM, Kuperwasser C (2008) 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:R25
56. Sheridan C, Kishimoto H, Fuchs RK, Mehrotra S, Bhat-Nakshatri P, Turner CH, Goulet R Jr, Badve S, Nakshatri H (2006) CD44+/CD24− breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. Breast Cancer Res 8:R59
57. Korkaya H, Paulson A, Iovino F, Wicha MS (2008) HER2 regulates the mammary stem/progenitor cell population driving tumorigenesis and invasion. Oncogene 27:6120–6130
58. Charafe-Jauffret E, Ginestier C, Monville F, Finetti P, Adelaide J, Cervera N, Fekairi S, Xerri L, Jacquemier J, Birnbaum D, Bertucci F (2006) Gene expression profiling of breast cell lines identifies potential new basal markers. Oncogene 25:2273–2284
59. Liu S, Ginestier C, Charafe-Jauffret E, Foco H, Kleer CG, Merajver SD, Dontu G, Wicha MS (2008) BRCA1 regulates human mammary stem/progenitor cell fate. Proc Natl Acad Sci USA 105:1680–1685
60. Tommiska J, Bartkova J, Heinonen M, Hautala L, Kilpivaara O, Eerola H, Aittomaki K, Hofstetter B, Lukas J, von Smitten K, Blomqvist C, Ristimaki A, Heikkila P, Bartek J, Nevanlinna H (2008) The DNA damage signalling kinase ATM is aberrantly reduced or lost in BRCA1/BRCA2-deficient and ER/PR/ERBB2-triple-negative breast cancer. Oncogene 27:2501–2506
61. Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee JS, Fridlyand J, Sahin A, Agarwal R, Joy C, Liu W, Stivers D, Baggerly K, Carey M, Lluch A, Monteagudo C, He X, Weigman V, Fan C, Palazzo J, Hortobagyi GN, Nolden LK, Wang NJ, Valero V, Gray JW, Perou CM, Mills GB (2009) Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. Cancer Res 69:4116–4124