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

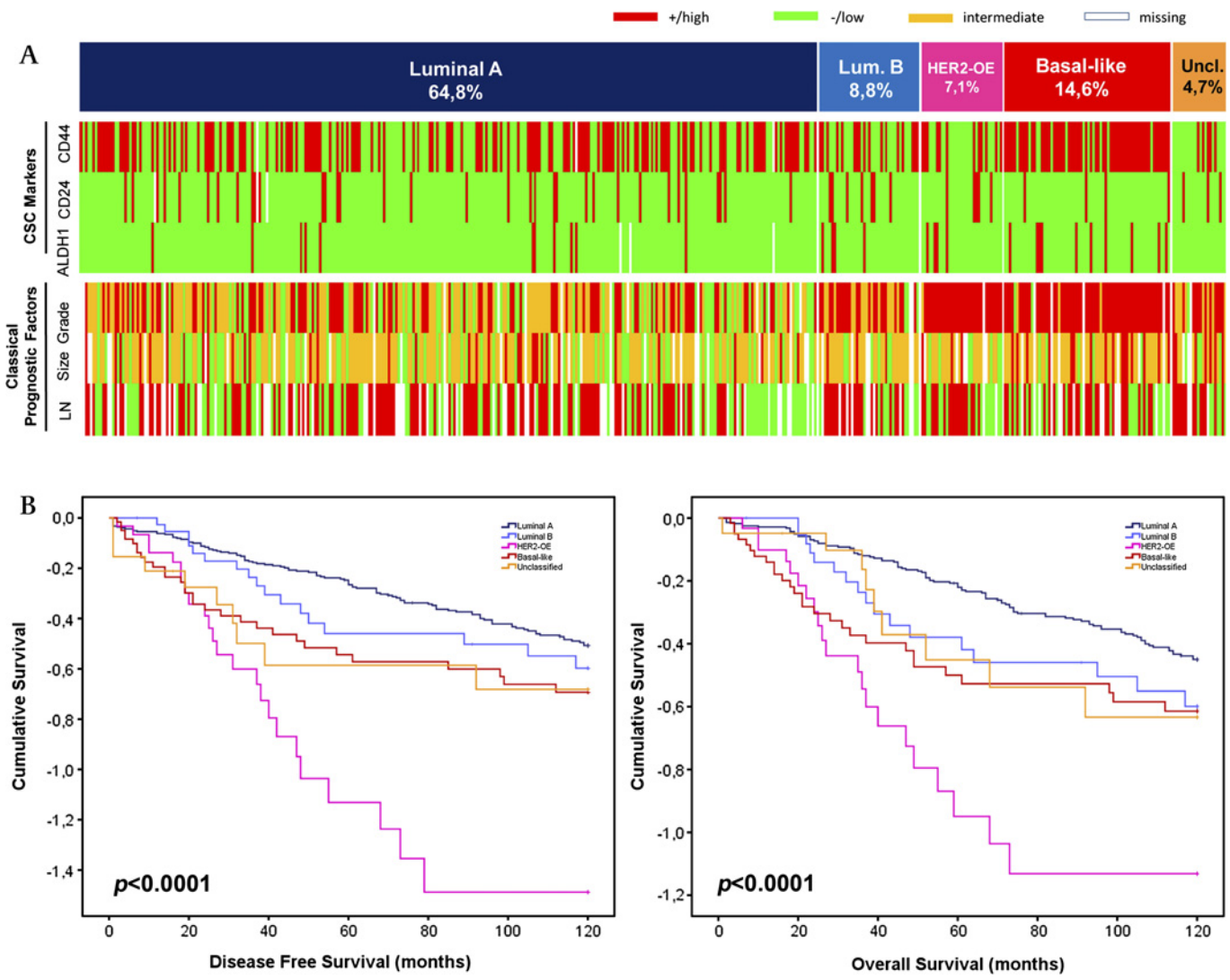

Figure 1 Breast tumour signature represented by immunohistochemistry array showing the protein expression of cancer stem cell markers (CD24, CD44 and ALDH1) and classic prognostic factors (tumour size, histological grade and lymph node metastasis) in the series of 466 invasive breast carcinomas analysed 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.

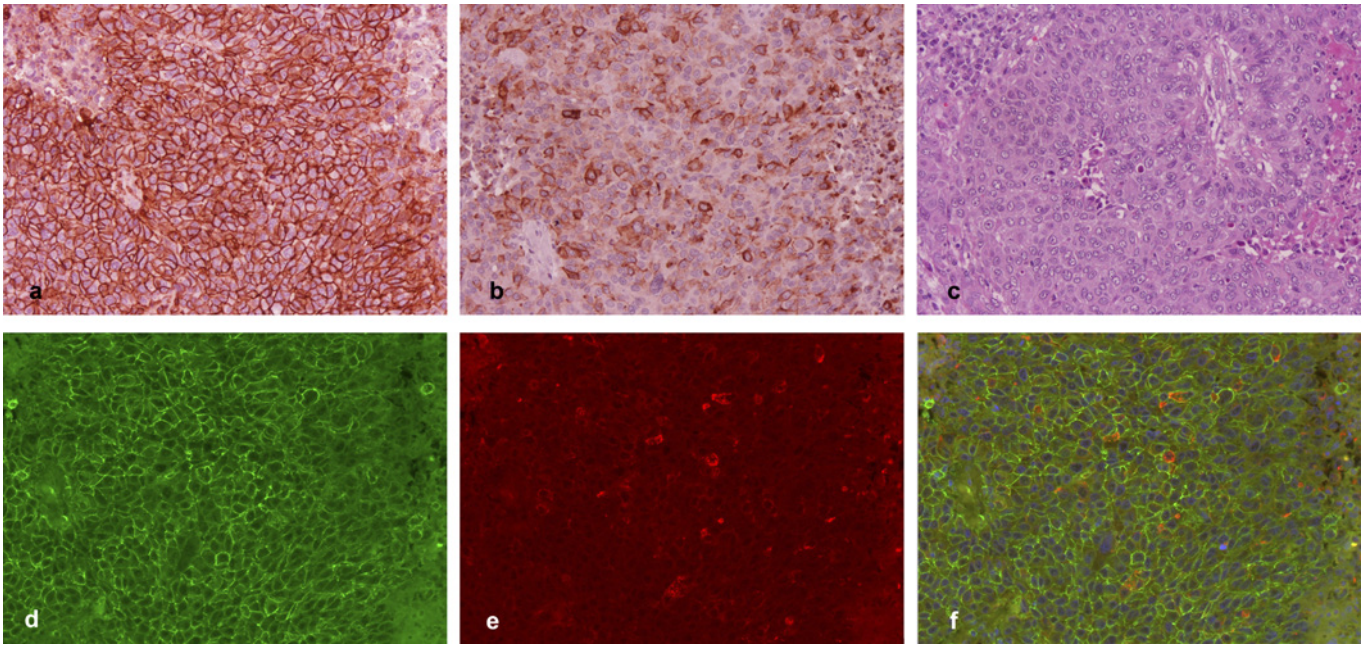

Figure 2 Basal-like carcinoma showing $CD44^{+}CD24^{-/low} \geq 10\%$ phenotype. Single-staining immunohistochemistry for $CD44^{+}$ (A), $CD24^{-/low}$ (B) and haematoxylin–eosin (C); and double-staining immunofluorescence for $CD44^{+}$ (D), $CD24^{-/low}$ (E) and merged image (F) (magnification ×200).

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 $CD44^{+}CD24^{-/low}$

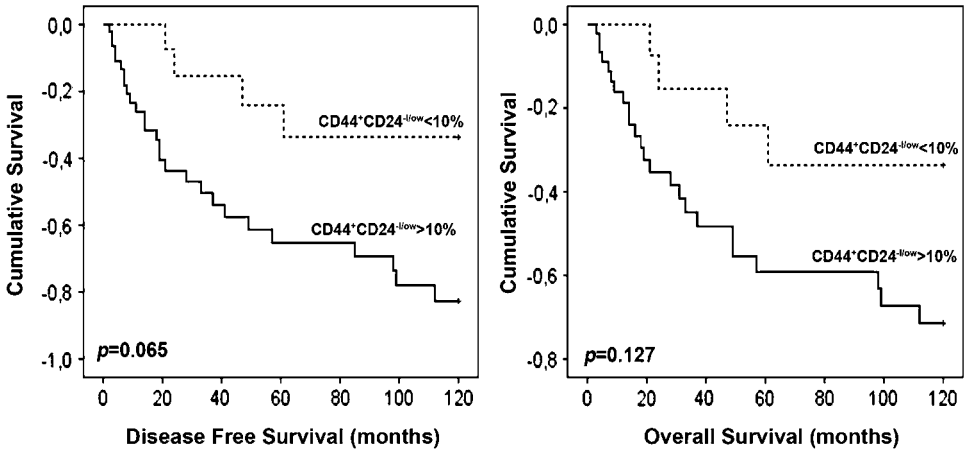

Figure 3 Kaplan–Meier plots of disease-free survival (DFS, p=0.065) and overall survival (OS, p=0.127) in the basal-like tumours defined according to the CD44/CD24 pattern of expression.

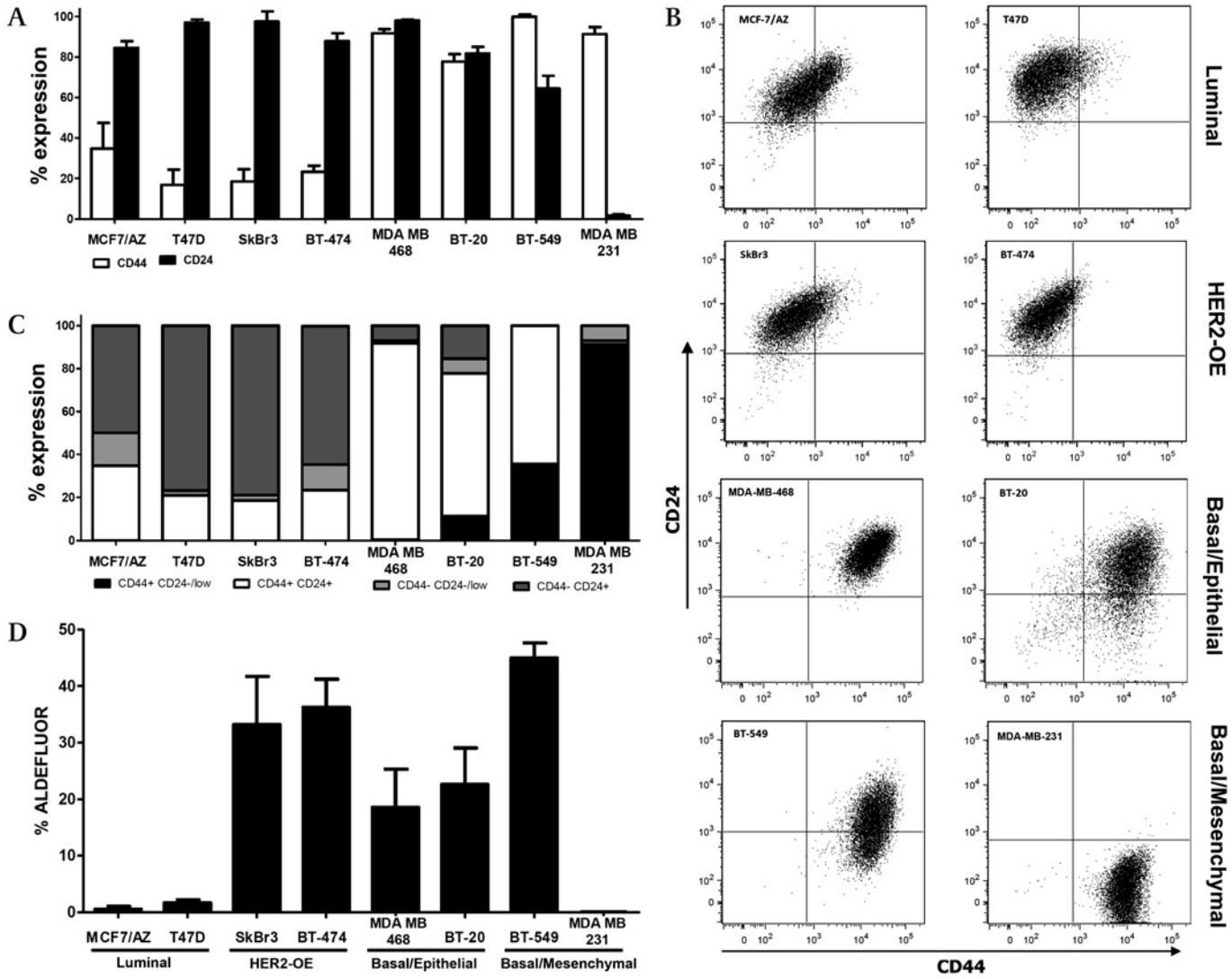

Figure 4 (A) Percentages of the subpopulations defined by the combination of the stem cell markers CD44 and CD24 in a panel of breast cancer cell lines representative of distinct molecular subtypes of the disease. Indicated is the mean±SEM of three independent experiments. (B) and (C) Subpopulations 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.