Breast cancer is the most frequent cancer among women, 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. In fact, global gene-expression analyses have provided an appealing molecular classification for breast carcinomas, which is highly associated with patients’ prognosis.

The molecular classification of breast cancer established four major subtypes: the luminal A and B, the HER2-overexpressing (HER2-OE) and basal-like tumours. 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.

In the past decade, many treatments undergoing clinical trials have been developed based on breast cancer molecular profiles. 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 constitutes a heterogeneous group of tumours, due to their stem cell-like characteristics: self-renewal and differentiation. Recent evidence has demonstrated that CSC are resistant to various forms of therapies, including radio and chemotherapy. 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 and CD44+/CD24−/low. 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. CD44 being positively associated with stem cell-like characteristics and CD24 expression related to differentiated epithelial features. These in-vitro data were later demonstrated in primary breast carcinomas, but the clinical and prognostic impact of these markers in
breast cancer remains a controversial issue 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 demonstrated that normal and cancer human mammary epithelial cells with increased aldehyde dehydrogenase activity (ALDH) show stem/progenitor cell properties. Tumorigenic ALDH+ CSC are significantly more resistant to platinum treatments, are biologically aggressive, and their expression tends to be associated with a poor patient prognosis. Interestingly, CD44\(^+\)CD24\(^{-}/\text{low}\) 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. It was recently shown that ALDH1 breast CSC marker can further divide the CD44\(^+\)CD24\(^{-}/\text{low}\) cell population into fractions that are highly tumorigenic: ALDH1\(^+\)CD44\(^+\)CD24\(^{-}/\text{low}\) cells were able to generate tumours from only 20 cells, whereas ALDH1\(^+\)CD44\(^{-}/\text{low}\) and high ALDH1 expression in primary tumours is quite small (approximately 1%). 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-Cies, 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 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, 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 P-cadherin and vimentin basal markers. Immunohistochemistry was performed in 3 \(\mu\)m sections. To study CSC markers in this series, specific antibodies for CD44 (clone 156-5C11; Cell Signaling Technology, Danvers, Massachusetts, USA), CD24 (clone Ab2-SNSb; 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 diamobenzidine 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. 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, and validated by a breast cancer pathologist. The cut-off value to distinguish low from high proliferation tumours was 15.25% of Ki67 nuclear staining. The Ki67 index was based on the study published by Cheang and colleagues, 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 cytoplasmatic positivity, as previously described. Stromal expression of ALDH1 was also classified in two categories: none/weak, or moderate/strong, as previously described by Resetkova et al.
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 described42–44 (see supplementary table S4, available online only).

Flow cytometry
Cells were washed twice with phosphate-buffered saline and then harvested with versene/0.46 mM ethylenediamine 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 diethylenaminobenzaldehyde, 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 χ² 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% (257/463) 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=0.058), CK5 (p<0.0001), P-cadherin (p=0.005), 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=0.005), 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
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. In our series, 45.3% (209/461) of the tumours were considered CD44+/CD24+/low 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 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 ≥ 10% (figure 2), independently of ALDH1 expression. Luminal tumours showed a mixture between the two CD44/CD24 phenotypes: 45.0% (127/295) of luminal A tumours were CD44+/CD24+/low ≥ 10% and 57.0% (168/295) were CD44+/CD24+/low < 10%. Luminal B showed the same CSC markers distribution (41.4% CD44+/CD24+/low ≥ 10% 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

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).
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.
Analysed the immunohistochemical membrane localisation of Vimentin, P-cadherin, CK14, EGFR, Ki67, HER2, PgR, ER, and histological grade. In this study, we identified 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. 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. CD44 expression was also inversely associated with lymph node metastasis, as previously shown by Giotrmanolaki et al.,

Table 2: Associations between the combined expression of CD44/CD24 and the classic breast cancer prognostic factors and biomarkers

<table>
<thead>
<tr>
<th>Tumour size</th>
<th>n</th>
<th>CD24 -/low&lt;10%</th>
<th>CD24 -/low≥10%</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 &lt;2 cm</td>
<td>100</td>
<td>54</td>
<td>46</td>
<td>0.362</td>
</tr>
<tr>
<td>T2 2–5 cm</td>
<td>240</td>
<td>138</td>
<td>102</td>
<td>0.0001</td>
</tr>
<tr>
<td>T3 &gt;5 cm</td>
<td>63</td>
<td>30</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Lymph nodes</td>
<td>364</td>
<td>195</td>
<td>169</td>
<td>0.582</td>
</tr>
<tr>
<td>Positive</td>
<td>206</td>
<td>127</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>158</td>
<td>68</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Histological grade</td>
<td>438</td>
<td>236</td>
<td>202</td>
<td>0.008</td>
</tr>
<tr>
<td>Grade I</td>
<td>80</td>
<td>40</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Grade II</td>
<td>133</td>
<td>76</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>Grade III</td>
<td>225</td>
<td>120</td>
<td>105</td>
<td></td>
</tr>
<tr>
<td>ER</td>
<td>459</td>
<td>251</td>
<td>208</td>
<td>0.082</td>
</tr>
<tr>
<td>Positive</td>
<td>304</td>
<td>175</td>
<td>129</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>155</td>
<td>77</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>PgR</td>
<td>460</td>
<td>251</td>
<td>209</td>
<td>0.480</td>
</tr>
<tr>
<td>Positive</td>
<td>225</td>
<td>119</td>
<td>106</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>235</td>
<td>132</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>HER2</td>
<td>457</td>
<td>250</td>
<td>207</td>
<td>0.126</td>
</tr>
<tr>
<td>Positive</td>
<td>68</td>
<td>43</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>389</td>
<td>207</td>
<td>182</td>
<td></td>
</tr>
<tr>
<td>Ki67</td>
<td>441</td>
<td>241</td>
<td>200</td>
<td>0.657</td>
</tr>
<tr>
<td>High</td>
<td>29</td>
<td>17</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>412</td>
<td>224</td>
<td>188</td>
<td></td>
</tr>
<tr>
<td>EGFR</td>
<td>461</td>
<td>252</td>
<td>209</td>
<td>0.184</td>
</tr>
<tr>
<td>Positive</td>
<td>22</td>
<td>9</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>439</td>
<td>243</td>
<td>196</td>
<td></td>
</tr>
<tr>
<td>CK5</td>
<td>461</td>
<td>252</td>
<td>209</td>
<td>0.0001</td>
</tr>
<tr>
<td>Positive</td>
<td>66</td>
<td>22</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>395</td>
<td>230</td>
<td>165</td>
<td></td>
</tr>
<tr>
<td>P-cadherin</td>
<td>461</td>
<td>252</td>
<td>209</td>
<td>0.008</td>
</tr>
<tr>
<td>Positive</td>
<td>114</td>
<td>50</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>247</td>
<td>202</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td>CK14</td>
<td>461</td>
<td>252</td>
<td>209</td>
<td>0.003</td>
</tr>
<tr>
<td>Positive</td>
<td>24</td>
<td>6</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>437</td>
<td>246</td>
<td>191</td>
<td></td>
</tr>
<tr>
<td>Vimentin</td>
<td>454</td>
<td>248</td>
<td>206</td>
<td>0.0001</td>
</tr>
<tr>
<td>Positive</td>
<td>78</td>
<td>19</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>376</td>
<td>229</td>
<td>147</td>
<td></td>
</tr>
<tr>
<td>ALDH1</td>
<td>459</td>
<td>250</td>
<td>209</td>
<td>0.599</td>
</tr>
<tr>
<td>Positive</td>
<td>32</td>
<td>16</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>427</td>
<td>234</td>
<td>193</td>
<td></td>
</tr>
</tbody>
</table>

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. 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. 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. CD44 expression was also inversely associated with lymph node metastasis, as previously shown by Giatromanolaki et al.,

Table 3: Associations between the combined expression of CD44/CD24/ALDH1 and the breast cancer molecular subtypes

<table>
<thead>
<tr>
<th>Luminal A</th>
<th>Luminal B</th>
<th>HER2-OE</th>
<th>Basal-like</th>
<th>Unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD44+CD24 -/low&lt;10%</td>
<td>ALDH1+</td>
<td>9 (3.1%)</td>
<td>2 (4.9%)</td>
<td>6 (8.8%)</td>
</tr>
<tr>
<td></td>
<td>ALDH1-</td>
<td>159 (53.9%)</td>
<td>22 (53.7%)</td>
<td>10 (14.7%)</td>
</tr>
<tr>
<td>CD44+CD24 -/low≥10%</td>
<td>ALDH1+</td>
<td>6 (2.0%)</td>
<td>3 (7.3%)</td>
<td>2 (6.1%)</td>
</tr>
<tr>
<td></td>
<td>ALDH1-</td>
<td>121 (41.0%)</td>
<td>14 (34.1%)</td>
<td>7 (21.2%)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>295 (100%)</td>
<td>41 (100%)</td>
<td>33 (100%)</td>
</tr>
</tbody>
</table>

p<0.0001.
consider a CD24-positive case is also diverse among these studies. Some have been categorising the CD24 marker as CD24+ versus CD24−, whereas others compare CD24−/low versus CD24+. 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. 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. 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, 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., CD44+CD24+ cells can readily give rise to CD44+CD24−/low.

**Figure 2** Basal-like carcinoma showing CD44+CD24−/low≥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).

**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.
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. As recent studies have described that epithelial-to-mesenchymal transition generates cells with CD44+CD24low stem-cell like characteristics, 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+CD24low CSC phenotype. Previous studies have already demonstrated an association between basal-like carcinomas and the CSC phenotype CD44+CD24low. 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. ALDH1high tumour cells form visibly larger colonies and mammospheres, when compared with ALDH1low cells. Previous works also detected small percentages of ALDH1+ cases in invasive breast cancer, ranging from 4% to 19%. 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

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.
Basal-like breast cancer is the intrinsic molecular subtype harbouring the highest percentage of tumour cells with the CSC phenotype CD44⁺/CD24⁻/low and ALDH1 positivity.

Luminal and HER-2 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 GSC phenotype; basal/epithelial breast cancer cells are mainly positive for both CD44 and CD24 CSC markers.

ALDH1 activity was mainly found in HER-2 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.

The authors would like to thank Professor Raquel Seruca and the Programa Ciência 2007 for JP. Fundacão Calouste Gulbenkian and by research grants funded by the Portuguese Education and is partially supported by FCT.

The authors would like to thank Professor Raquel Seruca and for critically reading the manuscript before submission. IPATIMUP is an

Take-home messages

Basal-like breast cancer is the intrinsic molecular subtype
Luminal and HER-2 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 GSC phenotype; basal/epithelial breast cancer cells are mainly positive for both CD44 and CD24 CSC markers.

ALDH1 activity was mainly found in HER-2 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.

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 regulatory 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


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

Sara Ricardo, André Filipe Vieira, René Gerhard, Dina Leitão, Regina Pinto, Jorge F Cameselle-Teijeiro, Fernanda Milanezi, Fernando Schmitt and Joana Paredes

*J Clin Pathol* 2011 64: 937-946 originally published online June 16, 2011
doi: 10.1136/jcp.2011.090456

Updated information and services can be found at:
http://jcp.bmj.com/content/64/11/937

These include:

**References**

This article cites 59 articles, 20 of which you can access for free at:
http://jcp.bmj.com/content/64/11/937#BIBL

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**

Articles on similar topics can be found in the following collections

- Editor's choice (132)
- Breast cancer (506)
- Immunology (including allergy) (1664)

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/