Abstract. The attempts to identify, isolate and characterize cancer stem cell populations are mostly dependent on cell-surface markers. In breast cancer, several putative breast cancer stem cell (BCSC) markers have already been reported, but the agreement on their phenotypic characterization is still absent. In fact, it became unfeasible to obtain a universal combination of markers that could specifically identify BCSCs in all cases of breast cancer. Breast cancer heterogeneity as reflected by various histological subtypes, with variable clinical presentations and diverse molecular signatures also contributes to major drawbacks. Indeed, intra-tumor heterogeneity leads to a single tumor to contain, at any given time, tumor cell populations displaying different molecular profiles and biological properties. As a consequence, several BCSC phenotypes were described, with some being associated with aggressive forms of breast cancer. Although the validation of the CSC model remains an ongoing task, it is important to define which BCSC phenotypes have high tumorigenic potential and ability to resist therapeutic agents. For this reason, a concise review is presented here regarding the implications of the most studied BCSC markers and phenotypes in breast cancer progression and treatment.

In the last two decades, breast cancer research has majorly focused on the identification, isolation and characterization of breast cancer stem cells (BCSCs). In order to do so, some genes with stem cell properties were studied and their corresponding proteins were subsequently validated as markers of BCSC (1, 2). As a consequence, a plethora of studies have been published describing the impact of BCSCs identified by these established markers, such as hyaluronan receptor (CD44), signal transducer CD24 (CD24) and aldehyde dehydrogenase-1 (ALDH1), as tumor-initiating cells in breast cancer progression with high propensity to metastasize and to be resistant to therapeutic treatments (2-4). With the increasing evidence for such ability, researchers have attempted to demonstrate which altered genes or dysregulated gene signaling pathways potentially contribute for the tumorigenic potential of BCSCs. In fact, NOTCH, WNT/β-catenin, or Hedgehog signaling pathways were shown to be deregulated in subpopulations of these cells (5, 6). With evidence forthcoming regarding the effects of the stroma and the microenvironment in breast tumor progression, several genes have also been reported to be associated with BCSCs (7). The phenomenon of epithelial–mesenchymal transition (EMT) and mesenchymal–epithelial transition in breast cancer cells during tumor progression was also an important discovery. Such dynamic transitions were demonstrated in BCSCs as an explanation for their ability to invade and to colonize other parts of the body, which has indeed questioned some of the tenets of the CSC model (8). With all this knowledge, targeting BCSCs for breast cancer treatment was demanded and currently, some important inhibitors targeting subpopulations of BCSCs or gene signaling pathways that regulate these subpopulations are reported to be strongly effective (9, 10).

One concern has, however, changed the definition of BCSCs: breast cancer heterogeneity (11). Due to the observations that not all BCSC markers are expressed in all breast cancer subtypes, research for different BCSC markers and different combinations of these markers that could be restricted to a specific breast cancer subtype or associated with aggressive forms of this disease is ongoing (12, 13). As a consequence, different BCSC phenotypes have been described and characterized, and in the future, other
molecules will be reported to have stem cell properties. Beyond the tenets of the CSC model, it is important to define which BCSC phenotypes have high tumorigenic potential and also great ability to resist therapeutic agents (14). Moreover, it is also crucial to determine which oncogenes or tumor-suppressor genes, other than those already described, are consistently mutated within these phenotypes being able to drive tumorigenesis.

In invasive breast cancer (IBC), several markers have been immunohistochemically characterized showing that the prevalence of stem cell-like and more differentiated markers varies according to tumor subtype and histological stage (15). For this reason, a concise review is presented here regarding the implications of the most studied markers of BCSCs and phenotypes in breast cancer progression and treatment, as well as a description of promising inhibitors able to target these cells.

**CD44+/CD24−/low Phenotype**

The combination of the BCSC markers CD44 and CD24 is by far the most extensively studied and undeniably the most contentious. The pioneering study by Al-Hajj et al. showed that as few as 100 CD44+/CD24−/low cells in patients with breast cancer were able to form tumors in mice, whereas tens of thousands of cells with alternative phenotypes failed to do so (1).

Immunohistochemically, breast cancer tissues were investigated for the prevalence of CD44+/CD24−/low tumor cells and their prognostic value. In a study including 136 patients with and without recurrence, the prevalence of CD44+/CD24−/low cells was ≤10% in 78% of cases and >10% in the other 22%. However, no significant correlation between the prevalence of this phenotype and tumor progression was noted nor were significant differences seen in recurrence, disease-free (DFS) or overall (OS) survival (16). In another study of 95 patients with IBC subjected to mastectomy, radiotherapy, chemotherapy and axillary lymph node dissection, CD44+/CD24−/low cancer cells were shown to be abundant in the basal subgroups and absent from those with human epidermal growth factor type 2 (HER2)-positive tumors (17). This phenotype was also associated with breast cancer 1 (BRCA1) mutational status, which was correlated with basal-like tumor status, and despite its association with increased poor prognostic features, it was not able to predict OS (18).

Regarding such important studies, the CD44+/CD24−/low phenotype has not a distinct prognostic value but it seems to be enriched in those with basal-like breast cancer subtype.

Gene-expression profiling of CD44+/CD24−/low breast cancer cells revealed a signature of 186 genes associated with invasion and poor prognosis (19, 20). This signature was enriched in genes related to the cell cycle, calcium-ion binding, chemotaxis, differentiation, protein transport, signal transduction and ubiquitination. Among these genes, this phenotype was observed to express high levels of interleukin-1 alpha (IL1α), IL6, ILβ and urokinase plasminogen activator (uPA), which predispose to distant metastases (Table I).

The enrichment of CD44+/CD24−/low cells demonstrated in primary breast tumors following radiation and chemotherapy has suggested an innate resistance to standard treatments (4). The presence of ATP-binding cassette transporters (which confer resistance to chemotherapeutic agents) highly expressed in a subpopulation of CSCs with these markers led to such assumption (21, 22). In fact, the ability of these cells to replicate in an in vitro model following at least four generations of xenograft transplanted mice has also suggested their significant role in tumor relapse and metastasis (23). Potential mechanisms of chemotherapy and radiation resistance associated with this phenotype were shown to include the presence of lower concentration of reactive oxygen species, cell dormancy, efficient DNA-repair mechanisms, overexpression of EMT markers, and activation of WNT/β-catenin, Hedgehog and NOTCH signaling pathways and signal transducer and activator of transcription 1 (STAT1) and STAT3 signaling (5, 24, 25-30). As a consequence, the CD44+/CD24−/low phenotype in breast cancer is currently being assessed as a therapeutic target.

One of the most promising therapeutic agents, MK0752, belongs to the class of γ-secretase inhibitors and was recently administered to patient-derived tumor xenograft in combination with docetaxel. This inhibitor was reported to improve docetaxel activity, leading to a decrease of CD44+/CD24−/low tumor cells, reduce mammosphere-forming activity, consequently leading to the inhibition of tumor formation after serial transplantations (9). With these results, a phase I clinical trial in patients with advanced breast cancer that did not respond to standard treatments was developed, culminating in a decrease of CD44+/CD24−/low tumor cells and in the reduction of the tumor bulk (9). This inhibitor is also being tested in combination with endocrine therapy (tamoxifen or letrozole in patients with early-stage breast cancer) and chemotherapy (docetaxel in patients with locally advanced or metastatic breast cancer) (31).

Another inhibitor from the same class, PF-03084014, was also administered in a phase I trial for the treatment of advanced breast cancer and was found to reduce NOTCH activity and to considerably reduce tumor cell migration and mammosphere-forming efficiency (31). Its ability to reduce self-renewal and expression of NOTCH target genes was also demonstrated in in vivo studies (32).

The major promising compound for breast cancer treatment is metformin, a drug generally used for anti-diabetic therapy. Metformin has been shown to preferentially target CD44+/CD24−/low cell subpopulations in different molecular subtypes of breast cancer cell lines and to have a synergistic
### Table I. Characteristics of the different assessed breast cancer stem cell (BCSC) phenotypes and markers.

<table>
<thead>
<tr>
<th>BCSC phenotype/markers</th>
<th>Tumorigenic potential</th>
<th>Clinicopathological features</th>
<th>Functional/mechanistic observations</th>
<th>Clinical observations</th>
<th>Inhibitors</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD44+/CD24−/low NOD/SCID mice (1)</td>
<td>Able to drive tumor formation when inoculated into NOD/SCID mice</td>
<td>Enriched in basal-like and claudin-low breast cancer subtypes (17) Associated with BRCA1 mutational status (18) Poor prognosis (16)</td>
<td>Increased expression of IL1α, IL6, ILβ and uPA (19), 20), ABC transporters, STAT1 and STAT3 (22, 28, 29), TWIST and SNAI1 (24, 25), and of WNT/β-catenin, Hedgehog and NOTCH signaling pathways (26, 31) Low levels of ROS (5) Associated with cell dormancy and efficient DNA repair mechanisms (28)</td>
<td>Tumor recurrence (25) Resistance to radiation and standard treatments (4) High metastatic propensity (23)</td>
<td>Short hairpin RNAs (102) Gamma-secretase inhibitors (9, 31, 32) Metformin (33, 34) ATRA or vorinostat (103) Niclosamide (36) Disulfiram/copper (37) Cyclophosphamide (38)</td>
</tr>
<tr>
<td>ALDH1 NOD/SCID mice (2)</td>
<td>Able to generate a stable tumor via orthotopic injection of ALDH1+ cells into NOD/SCID mice</td>
<td>Associated with poorer clinical outcomes including ER negativity, basal subtype and HER2 amplification (2, 42)</td>
<td>Increased expression of Ki-67 and EZH2 (42, 104), HIF-1/2α (47), HOXA1 and MUC4 (49) and of TGFβ, NOTCH and WNT/β-catenin signaling pathways (48, 51, 105)</td>
<td>Tumor recurrence (15) Enhanced capacity for metastatic behavior (15, 52) Resistance to sequential paclitaxel-and epirubicin-based chemotherapy (40)</td>
<td>DEAB (54) ATRA (54) Salinomycin (10) Disulfiram/copper (106)</td>
</tr>
<tr>
<td>CD44+/ALDH1+high NOD/SCID mice (14)</td>
<td>Able to drive tumor formation when inoculated into NOD/SCID mice</td>
<td>Able to identify high risk patients in breast cancer (12) Enriched in high-grade DCIS (57)</td>
<td>Increased levels of NOTCH and WNT/β-catenin signaling pathways (59) High levels of P-glycoprotein, GSTpi, and CHK1 (54)</td>
<td>Predicts distant metastasis and OS (58) High metastatic propensity (54) Resistant to standard cancer therapies (54)</td>
<td>DEAB (54) ATRA (54)</td>
</tr>
<tr>
<td>CD133 NOD/SCID mice (1)</td>
<td>Ability to form tumors in NOD/SCID mice from BRCA1-associated breast cancer cell lines (66)</td>
<td>Enriched in IBC and particularly in TNBC (67-69) CTCs detection in patients with TNBC (70, 71)</td>
<td>Increased expression of NOTCH1, ALDH1, FGFR1 and SOX1 (66) High levels of NANOG, SOX2, and BMI-1 in CD44+/CD49+high breast cancer cells (73)</td>
<td>Tumor recurrence (73) Resistant to standard cancer therapies (74)</td>
<td>Paclitaxel and surface antibody to CD133 (75)</td>
</tr>
<tr>
<td>CD29/CD49f NOD/SCID mice (79)</td>
<td>Enrichment of CD29 and CD49 in combination with CD24 for cancer-initiating cells in primary breast tumors (79)</td>
<td>CD29 associated with shorter OS and DFS (86, 87) CD49f associated with poor clinical outcomes and regarded as a prognostic factor (88, 89)</td>
<td>Increased expression of EMT markers (82, 89) EpCAM+/CD49f+ breast cancer cells proposed to be the cell origin of BRCA1-associated basal breast cancer (90)</td>
<td>Tumor relapse (107, 108) High metastatic propensity (91) Resistance to therapy (109, 110)</td>
<td>Combination of miR-9-3p with AZD6424 for CD29 (94)</td>
</tr>
</tbody>
</table>

ALDH1: Aldehyde dehydrogenase 1; ATRA: all-trans retinoic acid; BMI-1: polycomb complex protein BMI-1; BRCA1: breast cancer 1; CD44: hyaluronan receptor; CD24: signal transducer CD24; CTCs: circulating tumor cells; CHK1: checkpoint homolog 1; DCIS: ductal carcinoma in situ; DEAB: diethylaminobenzaldehyde; DFS: disease-free survival; EMT: epithelial–mesenchymal transition; EpCAM: epithelial cell adhesion molecule; ER: estrogen receptor; EZH2: Enhancer of zeste homolog 2; FGFR1: fibroblast growth factor receptor 1; GSTpi: glutathione S-transferase PI; HER2: human epidermal growth factor type 2; HIF-1/2α: hypoxia-inducible factors 1 and 2 α; HOXA1: Homeobox A1; IBC: invasive breast cancer; IL1α: Interleukin-1α; IL6: interleukin-6; ILβ: interleukin-β; MUC4: Mucin 4; NANOG: Nanog homeobox; NOD/SCID: non-obese diabetic/severe combined immunodeficiency; OS: overall survival; ROS: reactive oxygen species; SNAI1: zinc finger protein SNAI1; SOX1: SRY-box 1; SOX2: SRY-box 2; STAT1: signal transducer and activator of transcription 1; STAT3: signal transducer and activator of transcription 3; TGFβ2: transforming growth factor-beta 2; TNBC: triple-negative breast cancer; TWIST: Twist-related protein 1; uPA: urokinase plasminogen activator.
effect in eradicating CSCs upon its administration with doxorubicin (33, 34). Indeed, this compound is believed to interfere directly with the tumorigenesis of CD44+/CD24−/low tumor cells and to prevent neoplastic mammary lesions (35). Other promising compounds are being tested in order to target CD44+/CD24−/low tumor cells (Table I), although their potential still needs to be proven (36-38).

With all the experimental studies performed regarding this phenotype, it has become clear that CD44+/CD24−/low cellsurface markers are enriched for tumorigenic cells in some but not all breast cancers. Hence, the validity of the combination of these markers as being definitive of BCSCs has been called into question and additional markers have been reported, such as ALDH1 (39).

**ALDH1**

ALDH1 is broadly used as a functional marker in various types of cancer. Ginestier et al. were the first to demonstrate ALDH1 activity as a marker of stemness in normal and malignant breast cells. They were able to generate a stable tumor via orthotopic injection of 500 ALDH1-positive cells (evaluable by the ALDEFLUOR assay) into the mammary fat pads of non-obese diabetic/severe combined immunodeficiency (NOD/SCID) mice. Such tumorigenic ability of ALDEFLUOR-positive cells was also seen to be increased when shared with the CD44+/CD24−/low phenotype, since as few as 20 of such cells were sufficient to generate tumors in animals (2). Nonetheless, functional studies revealed that ALDH1+ cells were more prone to form colonies and tumors than CD44+/CD24−/low cells and also to be more chemoresistant (2, 40).

Several immunological studies have attempted to assess ALDH1 as a prognostic marker in breast cancer. The same study of Ginestier et al. with a cohort comprising 577 breast tumors from two independent tumor sets showed a prevalence of 30% for ALDH1 positivity and its correlation with high histological grade, HER2 overexpression and absence of estrogen receptor and progesterone receptor expression (2). ALDH1 was also demonstrated to be an independent prognostic factor and the same result was obtained in a group of 80 patients who underwent breast-conserving therapy. However, no association between tumor ALDH1 staining and micrometastatic disease was noted (41). In a larger cohort of more than 200 patients with primary breast cancer, Morimoto et al. reported a tendency for a worse prognosis in those with high ALDH1 expression (42). Nonetheless and regarding the type of cells expressing ALDH1, a well-designed study from Resetkova and colleagues found that ALDH1 expression in the stromal compartment of two cohorts of triple-negative breast tumors had prognostic significance, although being associated with good rather than poor DFS (43).

ALDH1 expression and clinical outcomes were also explored in inflammatory breast cancer (a particularly lethal form of breast cancer characterized by exaggerated lymphovascular invasion), revealing that ALDH1 expression was a predictive factor for early metastasis and reduced survival (44). On the contrary, in another study, no significant correlation between ALDH1 expression and clinicopathological variables was obtained, despite a trend toward association with a poorer OS (45).

With all these contradictory results, the reliability of ALDH1 expression as a clinical predictor of response to treatment is doubtful, thus enhancing the need for a standard protocol and evaluation process, as well as consideration of the differences between whole-tissue staining versus tissue microarray staining (46). Consistent findings between the reported studies are shown in Table I. Functional observations associated with ALDH1 are increased levels of NOTCH and β-catenin, which regulate the deacetylation process of ALDH1, increasing its tumorigenicity in vivo and contributing to a poor clinical outcome. Increased expression of hypoxia-inducible factors 1 and 2α was also shown to be associated with ALDH1 activity and believed to raise the metastatic propensity of ALDH1high cells (6, 47, 48). Furthermore, increased expression of homeobox A1 and mucin 4 were associated with high ALDH1 activity, also contributing to tumor relapse and metastasis. Activation of transforming grow factor-beta 2 signaling pathway was also shown to be involved in the pathological regulation of ALDH1 in breast cancer (49-51).

Using the ALDHhigh/CD44+/CD24−/low and ALDHhigh phenotypes, Croker et al. (52) and Charafe-Jauffret et al. (14), respectively, provided the first direct experimental evidence implicating ALDHhigh cells in breast cancer metastases in vivo. Cells with a CSC phenotype characterized by ALDH1 activity were shown to have an improved ability for metastatic behavior in vitro (adhesion, colony formation, migration, and invasion) and metastases in vivo, supporting the hypothesis that CSCs might act as metastasis-initiating cells (14, 49, 52-54).

ALDH1 is also involved in metabolizing chemotherapeutic drugs so its inhibition was believed to contribute to the reduction or elimination of BCSCs. Indeed, significant resistance to sequential paclitaxel- and epirubicin-based chemotherapy was found in tumor cells expressing ALDH1 (40). The inhibition of ALDH1 activity was also demonstrated to reduce stem cell-like properties and resistance to drugs and radiotherapy (54). As a consequence, these findings have emphasized the need to target ALDH1+ tumor cells in breast cancer treatment.

Even though studies evaluating the impact of pharmacological or immune targeting of ALDH on metastases in vivo are sparse, most showed a decrease of the metastatic burden. In this way, rationalized small-molecule discovery
has been proposed as a viable methodology to overcome these difficulties and such improvement has led to the current development and generation of isoform-specific ALDH inhibitors (Table I). Salinomycin (an inhibitor of the WNT signaling pathway) is currently the most promising chemotherapeutic drug demonstrated to inhibit the distinctive phenotypic properties of CSCs rather than inducing apoptosis of these cells (10). These findings underline the potential therapeutic value of targeting these properties to reduce the likelihood of tumor recurrence following chemotherapy.

Despite the enrichment of CSCs in ALDH1+ populations reported in several tissues, enzymatic activity measured by ALDEFLUOR alone is much more transient than the expression of traditional cell-surface markers. The usefulness of ALDH1 activity as a sole marker of CSCs may then be limited but can be increased if cells are stained simultaneously for ALDH1 activity and for more stable markers such as CD44 or CD133 (55).

**CD44+/ALDH1+/high Phenotype**

The importance of ALDH1 activity in breast cancer has been explored alone and in combination with the CD44+/CD24−/low phenotype. However, analysis of tumor samples revealed that only 1% of the ALDH-positive cell population had the CD44+/CD24−/low/Lineage− phenotype reported for BCSCs (2). In fact, an increase in the population of ALDH1-positive cells but not CD44+/CD24−/low cells has been observed in breast cancer tumor biopsies after neoadjuvant treatment (56). Considering the limited usefulness of ALDH1 activity as a sole marker of BCSCs, other combinations have been studied, such is the CD44+/ALDH1+/high phenotype. This phenotype was recently demonstrated to have high tumorigenic ability in breast cancer cell lines and also a highly metastatic propensity, being resistant to standard cancer therapies (54).

An interesting in situ method to define CSCs in formalin-fixed paraffin-embedded breast cancer tissues through a quantitative immunofluorescence method was designed by Neumeister and colleagues in order to measure the coexpression of CD44, ALDH1 and cytokeratin. Using a retrospective collection of 321 node-negative and 318 node-positive cases, localization of CD44+/ALDH1+ cells was shown within the epithelial (cytokeratin) compartment of breast tumor tissue. Even if this coexpression was seen in variably sized clusters and only in 6% of cases, such combination conferred a significantly worse outcome, being able to identify high-risk patients in breast cancer (12). Our research group has also studied the coexpression of CD44 and ALDH1, and remarkably, such combined expression was seen to be higher in ductal carcinomas in situ (DCIS) when compared with invasive ductal carcinomas (IDCs) of the breast, enhancing the tumorigenic potential of these putative BCSCs (57). More interestingly, in a cohort comprising 250 patients with different benign and malignant breast lesions, we also demonstrated that the CD44+/ALDH1+ phenotype was significantly increased in high-grade DCIS when compared with IDC. Moreover, this phenotype was found to be predominantly in a quiescent state (negativity for Ki-67 proliferation marker), raising some questions about the true role of dormancy in BCSCs (58).

A gene-expression analysis study revealed an increased expression of NOTCH and WNT/β-catenin signaling pathways in CD44+/ALDH1+ breast cancer cells isolated from an IBC cell line (59). However, concerning this phenotype, further genetic and even epigenetic analysis are required in order to better understand its behavior in breast cancer progression. Like CD44+/CD24−/low and ALDH1-positive tumor cells, it would be also interesting to depict the mechanisms that drive the progression of CD44+/ALDH1+ tumor cells.

Indeed, Croker and Allan directly inhibited ALDH activity with the specific ALDH inhibitor diethylaminobenzaldehyde (DEAB) and indirectly through all-trans retinoic acid (ATRA). They isolated ALDH/high/CD44+ and ALDH(low)/CD44+ populations and demonstrated that ALDH(high)/CD44+ human breast cancer cells were resistant to certain chemotherapy drugs (54). For that, they subjected both populations to treatment with doxorubicin, paclitaxel or radiation in the presence or absence of DEAB or ATRA and concluded that the reduction in cell viability was significantly greater in the ALDH(high)/CD44+ population. Furthermore, and in contrast to ALDH(low)/CD44+ cells, ALDH(high)/CD44+ cells showed increased basal activity in a series of DNA response proteins including P-glycoprotein, glutathione-S-transferase pi and checkpoint homolog 1 (CHK1) (Table I).

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

**CD133 (Prominin-1)**

CD133 has been recently included in CSC research. It is also named prominin-1 for its prominent location on the protrusion of cell membranes and was the first gene identified in those for a class of novel pentaspan transmembrane glycoproteins. Although it was initially
considered to be a marker of hematopoietic stem cells, CD133 mRNA transcript is also found in normal non-lymphoid hematopoietic tissue (60) and has been shown to play a role in SC migration and asymmetric division (61). CD133 was reported to be overexpressed in several solid tumors (62, 63), including colon cancer and glioblastoma (64, 65). In IBCs, CD133 expression was demonstrated by Liu et al. (66), where they assumed that its expression could be of help in a more accurate prediction of breast cancer aggressiveness and determination of the most suitable treatment. Actually, in BRCA1-associated breast cancer cell lines, CD133+ sorted cells were shown to have CSC properties, including a greater colony-forming efficiency, higher proliferative output and greater capability to form tumors in NOD/SCID mice (67). Moreover, CD133 was also proved to be suitable in the identification of CSCs in triple-negative breast cancers through several in vitro (68, 69) and in vivo studies (70). In addition, the recent use of CD133 to detect circulating tumor cells in patients with triple-negative breast cancer (71, 72) has increased the attention on this marker, emphasizing its role in prognosis in this breast cancer subtype. Expression of CD133 was also recently reported in 22 out of 25 cases of inflammatory breast cancer (13). Taken together, these interesting results indicate the need for more advanced research to understand the role of CD133 in BCSCs.

Expression of SC-associated genes, such as NOTCH1, ALDH1, fibroblast growth factor receptor 1 and SRY-box 1, was shown to be increased not only in CD44+/CD24− low but also in CD133+ breast cancer cells (67). Xenograft-initiating breast cancer cells enriched in CD44+/CD24high/CD133/2high cells were also shown to have elevated expression of Nanog homeobox (NANOG), SRY-box 2, and polycomb complex protein BMI-1 (73). Further extensive CD133 profiling in breast cancer needs to be performed to confirm CD133+ breast cancer cells as tumor-initiating cells.

Due to the increasing importance of CD133 expression in breast cancer progression, attempts have been made to correlate its expression with tumor relapse and resistance to chemotherapeutic agents. In fact, CD133 expression was reported to be correlated with tumor recurrence in patients with breast cancer (74). In drug-sensitive MCF-7 cells, only a small fraction of cells was found to be CD133-positive (75). In another interesting study, polymeric nanoparticles loaded with paclitaxel and surface functionalized with antibody to CD133 demonstrated efficient elimination of tumor-initiating cells in vitro and significant inhibition of tumor regrowth in vivo (76). With such results, CD133 is regarded as a potential target for anticancer therapeutics, being possible to reduce tumor recurrence in breast cancer through the elimination of CD133+ cells. Thus, additional studies investigating specific drugs that efficiently target this protein are required.

Integrins

The use of the integrins CD29 (β1) and CD49f (α6) in combination with CD24 was recently demonstrated to be able to identify mouse mammary SCs (77, 78). Since all previously described SC markers were shown not only to identify normal mammary SCs but also to isolate BCSCs, Vassilopolus et al. used CD24/CD29 and CD24/CD49f to successfully identify a subpopulation of mammary tumor cells (79). Such demonstration highlighted the importance of CD29 and CD49f in BCSCs. CD49f heterodimerizes with either the CD29 or CD104 (β4 integrin) subunits to generate the CD49fCD29 and CD49fCD104 integrins, which function primarily as laminin receptors (80). In addition, CD49f cooperates with receptor tyrosine kinases to communicate, bidirectionally, between the cell and the extracellular matrix (ECM). Interestingly, however, the CD104 subunit appears to be expressed at very low levels, if at all in CSCs when compared to non-CSCs indicating that CD49fCD29 is the dominant integrin expressed by CSCs (81, 82).

CD29 represents the predominant integrin in mammary epithelial cells in mice, and is aberrantly expressed in human breast carcinomas, contributing to diverse malignant phenotypes, including EMT, metastasis and angiogenesis (82-85). Moreover, in patients with IBC, high CD29 expression was found to be associated with significantly shorter DFS and OS (86, 87). In human breast cancer, CD49f integrin is overexpressed and was shown to be an independent prognostic factor of a poor outcome (88). CD49f+ cancer cells were also associated with a higher probability of distant metastasis after initial surgery and poor clinical outcomes with respect to both DFS and OS (89). Additionally, normal human SCs and myoepithelial progenitor cells characterized by CD49fhigh/epithelial cell adhesion molecule (EpCAM)+ cells were shown to express vimentin, a common EMT marker, suggesting that some cells may have been undergoing EMT (90). Interestingly, an aberrant luminal progenitor cell population (EpCAM+/CD49f+) was also proposed to be the cell origin of BRCA1-associated basal breast cancer (91).

Functional analysis revealed that while knockdown of CD29 or CD49f alone slightly reduced cell migration ability in BRCA1-mutant cancer cell lines, knockdown of both genes caused a profound effect, blocking migration, suggesting an overlapping, yet critical function of both genes in the migration of BCSCs (79). Such an interesting finding supports the notion that both integrins can pair with each other in order to form heterodimers for ECM components such as fibronectin and laminin (80). Consistent with the assumption that a malignant social network mediates cell-cell adhesion and communication between CSCs and their microenvironment (92), both integrins may be implicated in mediating such a network. Specifically, CD29/CD49f integrins may mediate CSC–stromal interaction, relaying
ECM signaling to cellular machinery, leading to the increased activity of CSCs in terms of viability, differentiation and metastasis (79).

Although the CD29/CD49f integrins have been implicated in the function of BCSCs and other CSCs (81, 82, 93), much needs to be learned about the contribution of these integrins to the genesis of BCSCs. It has been shown that CD49f and CD29 contribute to therapy resistance, tumor relapse and metastasis in breast cancer. As a consequence, the development of inhibitors that could potentially target these two integrins in breast cancer is required (Table 1). Interesting studies have been published with promising results in targeting these integrins such as through the use of short hairpin RNAs or micro-RNAs (94). Targeting gene signaling pathways associated with these integrins or even specific kinases such as feline sarcoma-related kinase, which controls migration and metastasis of IBC cell lines by regulating CD49f- and CD29-integrin-dependent adhesion, is also an interesting approach (95).

**Breast Cancer Stem Cells and Next-Generation Sequencing**

The continuous improvements of next-generation sequencing (NGS) technologies currently allow the analysis of hundreds of genes in just one population of cells, or even in one single cell (96, 97). Such application has opened a new window in the genomic field where a mutational, time-based lineage tree can now be delineated for a specific subtype of cancer considered to be highly aggressive. In this way and with NGS, it is possible to determine other genes from those already associated with being oncogenic, or, more importantly, to assess which mutated genes are responsible for driving tumorigenesis, considering the high levels of heterogeneity in cancer, especially in breast cancer (98). In fact, NGS has been recently used for the analysis of the molecular features of early-stage breast cancer leading to a genomic portrait of this disease. Within such a portrait, tumor suppressor p53 (TP53) and phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PI3KCA) mutations were the most frequent genomic alterations found in all breast cancer subtypes. Clinical relevance of phosphatase and tensin homolog (PTEN) mutations and deletions as well as those of v-akt murine thymoma viral oncogene homolog 1 (AKT1), BRCA1 and BRCA2 was also highlighted (99).

With this in mind, such technology would allow the definition of a mutational repertoire of each subpopulation of BCSCs presented here. Klevebring et al. undertook exome sequencing of CSCs (characterized by CD44+/CD24− and expression of ALDH1) from 12 patients with breast cancer, along with paired primary tumor samples. They found that the vast majority of mutations were shared between CSCs and the bulk primary tumor, as was the observed distribution of allelic frequencies, suggesting that a dynamic transition between cellular states (CSC and differentiated state) takes place continuously throughout the tumor development (100). Even using a small cohort, through NGS, our research group was able to detect somatic mutations in CD44+/CD24−low/CK+/CD45− breast cells isolated from non-malignant and malignant breast lesions. Mutations affecting the TP53, NOTCH1, GTPase HRas (HRAS), AKT1, PTEN, colony-stimulating factor 1 receptor (CSF1R) and ret proto-oncogene (RET) genes were detected in the malignant lesions, suggesting a heterogeneous molecular profile of these BCSCs (101). Thus, a practical example would be the application of NGS in isolated BCSCs (defined by different combinations of BCSC markers) from primary tumors and their corresponding metastases in order to determine which gene is more frequently mutated (hotspot mutations) in each subpopulation of BCSCs. Such an approach would be of great importance for the development of additional therapeutic drugs that could be promising not only for the most well-known BCSC markers (CD44, CD24 and ALDH1), but also for the discovery of new targets directly associated with other BCSC markers, such as CD133 or integrins.

In the future, this growing technology will definitely revolutionize CSC research, by providing new deregulated gene signaling pathways directly involved in the progression of tumor-initiating cells already proven to have stem cell properties, particularly in breast cancer.

**Acknowledgements**

The Authors would like to thank the Foundation for Science and Technology (Portugal). Scholarship reference: SFRH/BD/74307/2010 (ACP).

**References**


69 Brunogli F, Grassilli S, Piazzzi M, Palomba M, Nika E, Baveillon A, Capitanio S and Bertagnolo V: In triple-negative breast cancer tumors, PLC-β2 promotes the conversion of CD133high to CD133low phenotype and reduces the CD133-related invasiveness. Mol Cancer 12: 1, 2013.


75 Wang XY, Penalva LOF, Yuan H, Linolino RI, Lu J, Okano H and Glazer RI: Musashi1 regulates breast tumor cell proliferation and is a prognostic indicator of poor survival. Mol Cell 9: 221, 2010.


