Modern Pathology (2017) 30, 698–709

P-cadherin: a useful biomarker for axillary-based breast cancer decisions in the clinical practice

André Filipe Vieira[1,2,8], Maria Rita Dionísio[1,2,3,8], Madalena Gomes[1,2], Jorge F Cameselle-Teijeiro[4], Manuela Lacerda[2], Isabel Amendoeira[5], Fernando Schmitt[6] and Joana Paredes[1,2,7]

[1]Instituto de Investigação e Inovação em Saúde (i3s), Universidade do Porto, Porto, Portugal; [2]Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), Porto, Portugal; [3]Centro Hospitalar de Lisboa Norte (CHLN), Lisboa, Portugal; [4]Complejo Hospitalario Universitário de Vigo (CHUVI), Vigo, Spain; [5]Centro Hospitalar de São João (CHSJ), Department of Pathology, Porto, Portugal; [6]Laboratoire National de Santé, Department of Pathology and Medicine, Luxembourg and [7]Faculdade de Medicina da Universidade do Porto, Porto, Portugal

Axillary lymph node metastases represent the most powerful breast cancer prognostic factor, dictating disease staging and clinical therapeutic decisions. Nonetheless, breast cancer patients with positive lymph nodes still exhibit a heterogeneous behavior regarding disease progression. Stem-like subpopulations of cancer cells show high migratory and metastatic capacity, thus we hypothesize that breast cancer stem cell markers evaluation in metastasized lymph nodes could provide a more accurate prediction of patient's prognosis. Therefore, the expression profile of P-cadherin, CD44, and CD49f, which have been already associated to stem cell properties in breast cancer, has been evaluated by immunohistochemistry in a series of 135 primary tumors and matched axillary lymph node metastases from 135 breast cancer patients. Taking in consideration the expression of the stem cell markers only in axillary nodes, P-cadherin was the only biomarker significantly associated with poor disease-free and overall patient's survival. Moreover, although a concordant expression between primary tumors and matched lymph nodes has been found in the majority of the cases, a small but significant percentage displayed divergent expression (18.2–26.2%). Remarkably, although CD44 and CD49f changes between primary tumors and lymph node metastasis did not impact survival, the cases that were positive for P-cadherin in lymph node metastases being negative in the primary tumor, presented the worst disease-free and overall survival of the whole series. Accordingly, negative cases for this marker in the lymph nodes with positive expression in the matched breast carcinoma demonstrated a better prognosis, which overlapped with tumors that were negative in both sites. P-cadherin and CD49f gain of expression was mainly found in triple-negative carcinomas. Our results indicate for the first time that the evaluation of P-cadherin expression in lymph node metastases is an important predictor of disease outcome, being a putative valuable marker for axillary-based breast cancer decisions in the clinical practice.
Modern Pathology (2017) **30,** 698–709; doi:10.1038/modpathol.2016.232; published online 13 January 2017

Breast cancer is among the most deadly malignancies in developed countries, with metastatic spread being the major reason behind this fatal outcome.[1] According to the European Society for Medical Oncology guidelines, the number of metastasized lymph nodes is a fundamental prognostic factor in breast cancer, constituting a cornerstone for multimodal and combined treatment options, namely chemotherapy, immunotherapy, radiotherapy, and surgery.[2–4] Notably, the biology within the lymph node itself has been completely undervalued regarding clinical decisions. In fact, the molecular characteristics of primary tumors form the basis for the classification into one of four molecular subtypes (Luminal A, Luminal B, epidermal growth factor receptor (HER) 2-overexpressing carcinomas and basal-like carcinomas), which is highly correlated with patient

Correspondence: Dr J Paredes, PhD, I3S, EPIC group, IPATIMUP, Rua Alfredo Allen, No 208, Porto 4200-135, Portugal.
E-mail: jparedes@ipatimup.pt
[8]These authors contributed equally to this work.
Received 26 July 2016; revised 29 November 2016; accepted 29 November 2016; published online 13 January 2017

prognosis and set the course for adjuvant treatment options.[5–10] Nevertheless, despite that many countries in the world today depend on this classification, independently of the molecular profile of the primary tumor, ~6–10% of the patients show metastatic disease at diagnosis and systemic recurrence develops in 25–30% of the patients treated with curative intent.[11–13] In recent years, it is becoming clear that having axillary biomarkers with prognostic value would be crucial to support clinical decisions, as there is still a very heterogeneous behavior regarding disease progression in patients within the same molecular subtype or in patients with lymph node involvement, some of which never develop distant metastasis.[14–16] Several studies recognize that metastases exhibit a gene expression profile that differs from the originating tumor, with variations already accumulating in the lymph nodes.[17–23] Gene copy number profiling and gene expression analysis comparing primary tumors and metastases from the same patient have revealed interesting genetic differences, pointing for new putative prognostic biomarkers.[19–23] Understanding this heterogeneity constitutes a potential critical point for predicting patient prognosis. In fact, some immunohistochemical studies have addressed the expression of biomarkers in axillary lymph nodes and in distant metastasis, focusing mainly in the classical markers that define treatment options and allow carcinoma subtype classification, such as ER, progesterone receptor (PR), Ki67, epidermal growth factor receptor 1 (EGFR), and 2 (HER2).[4,24–28] Importantly, most studies lack prognostic information for individual patients or are limited to a small cohort.

Accumulating evidence indicates that local recurrences and/or distant metastases originate from rare tumor cells that may function like stem cells in their ability to initiate, propagate and hierarchically organize secondary tumors.[29] These breast cancer stem cells are shown to exhibit unique growth abilities including self-renewal, differentiation potential and resistance to most anti-cancer agents, including chemo- and/or radiotherapy, which contribute to the overall aggressiveness of recurrent or metastatic lesions.[30,31] Thus, efforts are needed to define sensitive and reliable breast cancer stem cell biomarkers for metastatic lesions.

In pathology studies, CD44 and CD49f are among the most extensively used breast cancer stem cell markers. CD44 is the transmembrane receptor for hyaluronic acid and is highly involved in cell–cell and cell–matrix interactions.[32] A subpopulation of tumor cells strongly expressing CD44, but low levels of CD24 (the $CD44^{+}CD24^{-/D24b}$ phenotype), was identified for the first time as a population of breast CSCs by Al-Hajj *et al*.[33] and this evidence was subsequently confirmed by several other authors. Specifically, the CD44+ phenotype is correlated positively with colon, breast, prostate, and pancreatic cancer initiating cells.[33–36] Specifically in breast cancer, CD44 has been demonstrated as a marker of cancer cells with a basal-like phenotype, being present in poor prognosis tumors, and associated with more motile and invasive cancer cells.[37] Concerning CD49f (also known as α6 integrin, involved in the cell interaction with laminin at the hemidesmosome), it has been extensively used to purify mouse and human mammary stem cells,[38–40] as well as to characterize a population of human cancer cells with stem-like behavior.[41,42] Recently, Ghebeh *et al*.[43] showed that the combination of $CD44^{high}/CD24^{low}/CD49f^{+}$ might significantly improve and strengthen the measurement of breast cancer stem cells with significantly higher stem/progenitor ability and the expression of both CD44 and CD49f in breast primary carcinomas was associated with worse clinical outcome and tumor aggressive features.[41,44]

In addition, our group showed that P-cadherin also behaves as a breast cancer stem cell marker in human breast carcinomas. P-cadherin is a calcium-dependent cell–cell adhesion molecule, which expression was found to be significantly associated with CD49f and CD44 expression,[45,46] therapy resistance,[45] and cancer cell invasion.[47,48] Furthermore, P-cadherin mediates stem cell properties, namely by increasing the cell growth in anchorage-independent conditions, by activating $\alpha6\beta4$ integrin signaling, and by inducing tumorigenic ability in nude mice.[45,46] In fact, we have identified P-cadherin as an independent indicator of poor prognosis in primary breast tumors,[49] being associated with high histological grade tumors and negative hormonal receptors status.[49–53] P-cadherin overexpression is predominantly found in basal-like breast carcinomas[49,54] and is strongly linked to *BRCA1* mutations.[55] The tumor-promoting properties mentioned for P-cadherin in breast cancer, namely its invasive ability, are found in several other tissue contexts, suggesting a putative role for this molecule in the metastatic process.[56]

In this study, we addressed, for the first time, the expression of P-cadherin, CD44, and CD49f in primary breast carcinomas and matched axillary loco-regional metastases, in order to evaluate their impact in patients' survival. P-cadherin arises as an independent indicator of prognosis in the metastatic setting and as a putative useful biomarker for axillary-based breast cancer decisions in clinical practice.

Materials and methods

Patient Selection and Material Characterization

A series of 135 formalin-fixed paraffin-embedded cases of invasive breast carcinomas, including primary tumors and matched lymph node metastases, was obtained from the Pathology Departments of Hospital Xeral-Cíes (Vigo, Spain) and Hospital São João (Porto, Portugal), under patient informed consent and with ethical approval by both hospitals

Ethical Commissions. Clinical and pathological features were retrieved for this study.

Patient follow-up information was available for 118 patients, which were diagnosed between 1978 and 1992, with a maximum follow-up of 120 months after diagnosis. These breast carcinoma patients followed the adequate protocols for chemotherapy, radiotherapy, and hormone therapy given at that time. All patients have been treated with adjuvant chemotherapy, which consisted of a protocol of six cycles of cyclophosphamide, methotrexate, and fluorouracil. Patients with ER positive tumors have been treated with hormonal therapy, which was carried out exclusively with tamoxifen. Patients with HER2-overexpressing carcinomas were not treated with specific targeted therapy (trastuzumab). No neo-adjuvant treatment has been used in the patients included in this series.

The disease-free survival interval was defined as the time from the diagnosis to the date of breast cancer-derived relapse, whereas overall survival was considered as the number of months from the diagnosis to the disease-related death. For each primary tumor, the axillary lymph nodes were embedded in paraffin and carefully evaluated by a pathologist, which identified the lymph nodes with metastatic disease. Each breast cancer case was matched with the lymph node with the highest metastatic burden (macrometastasis).

The series included 26.7% of cases from women diagnosed before 50 years old and 73.3% from women with more than 50 years old. Concerning tumor size, 10.9% of the cases were classified as T1 (<2 cm), 72.8% as T2 (2–5 cm), and 16.3% as T3 (>5 cm). All cases were positive for lymph node metastasis, as this was the criterion to select the samples. Regarding histological grade, 8.9% of the cases were Grade I, 25.4% Grade II, and 65.7% Grade III. Breast carcinomas were classified considering European Society for Medical Oncology guidelines into luminal A carcinomas (ER/PR positive, HER2 negative, and Ki67 low), luminal B carcinomas (ER/PR positive, HER2 and/or Ki67 high), HER2-overexpressing carcinomas (ER/PR negative, HER2 positive), and triple-negative carcinomas (ER/PR/HER2 negative), according to the molecular classification used in the clinical setting.[57] The Ki67 cutoff value to distinguish low from high proliferative tumors was 13.25%, based on a study from Cheang and colleagues.[58] From the 135 breast cancer cases, 57.0% (77/135) were luminal A, 10.4% (14/135) were luminal B, 14.1% (19/135) were HER2-over-expressing, and 18.5% (25/135) were triple negative. Although the series included only lymph node positive patients, patient follow-up information revealed that clinical outcome varied considerably. Specifically, 10 years (120 months) after diagnosis, 48/118 (40.7%) patients had no relapse of the disease and 52/118 (44.1%) patients were alive (Supplementary Table S1).

The present study was conducted under the national regulative law for the usage of biological specimens from tumor banks, where the samples are exclusively available for research purposes in the case of retrospective studies. All analyses were performed according to the REporting recommendations for tumor MARKer prognostic studies (REMARK) recommendations for prognostic and tumor marker studies.

Immunohistochemistry

Immunohistochemistry was performed in 3 μm sections. The following antibodies were used: P-cadherin (BD Transduction Biosciences, USA; clone 56, 1:50), CD44 (Cell Signalling Technology, USA; clone 156-3C11; 1:100), and CD49f (Sigma-Aldrich, USA; HPA012696; 1:50). Epitope exposure was performed for 30 min using high-temperature antigen retrieval (95 °C) with Tris-EDTA buffer pH = 9.0 (Novocastra, UK) for P-cadherin and CD49f, or citrate buffer pH = 6.0 (Thermo Scientific, USA) for CD44. Primary antibodies were detected using the horseradish peroxidase polymer (Cytomation Envision System HRP; DAKO, USA) (for P-cadherin and CD49f) or the labeled biotin-streptavidin method (DAKO, USA) (for CD44), according to manufacturer's instructions. Diaminobenzidine was used as chromogen.

Immunohistochemical Evaluation

The expression of P-cadherin, CD44, and CD49f was evaluated according to the grading system already described.[45,59–62] In brief, staining was scored only when detected at the membrane of tumor cells and the staining extension was considered as follows: 0, 0–10% of positive tumor cells; 1+, 10–25% of positive tumor cells; 2+, 25–50% of positive tumor cells; 3+, >50% of positive tumor cells. The cases classified as (0) were considered negative, whereas (1+), (2+), and (3+) were scored as positive.

Statistical Analysis

Associations between clinicopathological parameters and the expression of the markers evaluated in primary tumors and loco-regional metastases were assessed by Pearson's χ^2- and Fisher's exact tests. Survival analyses were estimated using the Kaplan–Meier method and compared using the log-rank test. For uni/multivariate survival analysis, Cox regression models were fitted to estimate hazard ratios (HRs) and the corresponding 95% confidence interval. Statistical analyses were carried out using SPSS statistics V.17.0 software (SPSS Inc., Chicago, IL), and a significance level of 5% was considered statistically significant.

Results

P-Cadherin is Associated with CD44 and CD49f Expression in Metastatic Lymph Nodes

A breast carcinoma series comprising 135 invasive carcinomas and matched lymph node metastasis was evaluated for the expression of the breast cancer stem cell markers P-cadherin, CD44, and CD49f by immunohistochemistry (Figure 1). P-cadherin membrane staining was found in 26.7% (36/135) of primary breast carcinomas and 23.0% (31/135) of metastasized lymph nodes. Regarding CD44, 47.4% (63/133) of primary tumors and 48.5% (64/132) of lymph node metastasis had a positive staining pattern. Considering CD49f, 14.3% (19/133) of primary tumors and 31.7% (40/126) of metastatic lymph nodes were scored as positive (Supplementary Table S2). Interestingly, CD49f displayed the major difference in expression comparing primary tumors and metastatic lymph nodes, which was found to be statistically significant ($P=0.001$).

Further, the expression of breast cancer stem cell markers was correlated with classic prognostic factors and with the primary tumor molecular subtype. P-cadherin expression was able to significantly discriminate primary breast carcinomas with high histological grade ($P=0.023$) (Supplementary Table S3), confirming our previously reported data.[51] Triple-negative breast carcinomas were significantly enriched in the expression of CD44 ($P=0.009$) and CD49f ($P=0.007$). P-cadherin expression was also enriched within the triple-negative molecular subtype, but the association was not statistically significant ($P=0.105$) (Supplementary Table S3). Notably, the expression of all three breast cancer stem cell markers in lymph node metastases was significantly associated with the triple-negative phenotype as well ($P<0.001$ for P-cadherin, $P=0.007$ for CD44 and $P<0.001$ for CD49f).

Interestingly, we found that P-cadherin expression in the primary tumor was also significantly associated with an increased number of metastasized lymph nodes: the majority of cases with a low number of metastasized lymph nodes (1–3 involved lymph nodes) were P-cadherin negative (89.4% (42/47)), whereas only 10.6% (5/47) were positive for P-cadherin. In contrast, a significant increase of P-cadherin was found in cases with 4–9 metastasized lymph nodes (36.5% (19/52)), as well as >10 metastasized lymph nodes (33.2% (10/31)) ($P=0.009$). We did not find this association regarding CD44 or CD49f (Supplementary Table S4).

Examining the association between the three markers in lymph node metastases, we observed that P-cadherin expression was significantly associated with CD44 expression (73.3% of P-cadherin positive cases were CD44 positive, $P=0.003$) and CD49f expression (59.3% of P-cadherin positive cases were CD49f positive, $P=0.001$) (Table 1). In contrast, in primary breast carcinomas, no statistically significant association was found between P-cadherin and CD44 expression, or between P-cadherin and CD49f. However, CD44 and CD49f were significantly associated in both primary tumors ($P=0.001$) and lymph node metastasis ($P=0.002$) (Table 1).

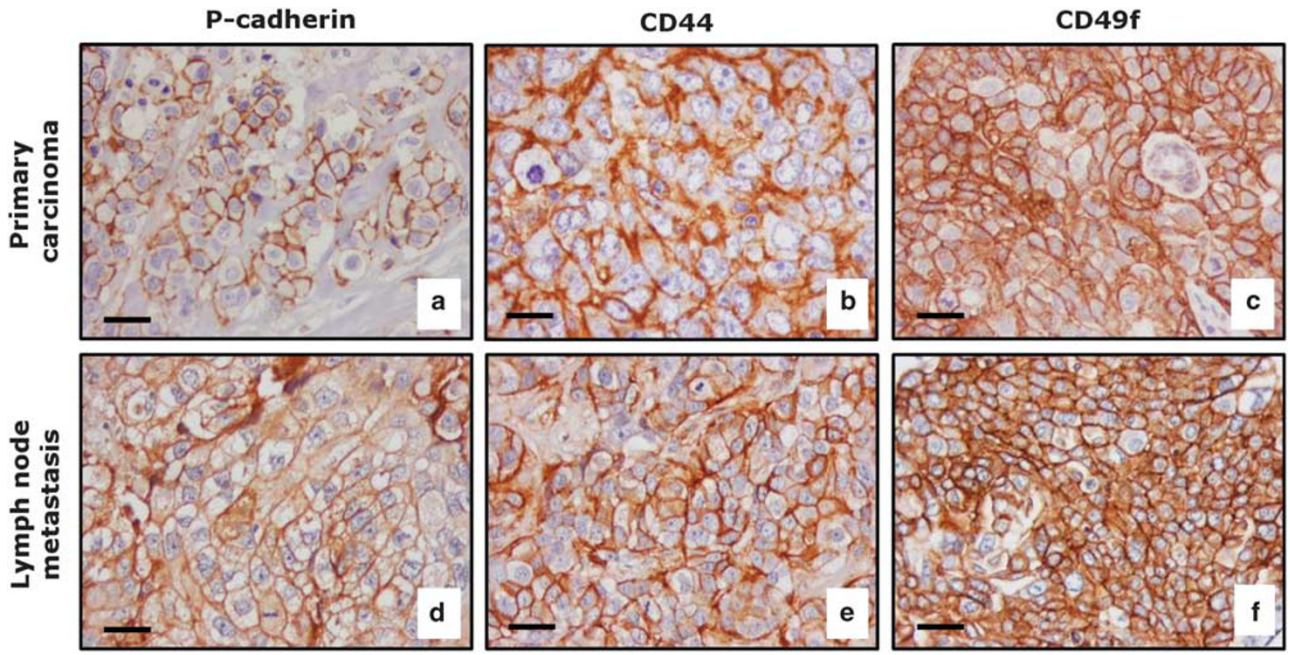

Figure 1 Breast cancer stem cell markers staining patterns. Representative images showing clear positive immunohistochemical staining for P-cadherin, CD44 and CD49f in primary tumors (**a–c**, respectively) and in lymph nodes metastases (**d–f**, respectively). Scale = 50 μm.

Table 1 Associations between the expression of P-cadherin, CD44, and CD49f in primary tumors and in metastatic lymph nodes

	P-cadherin			*CD44*		
	Positive	*Negative*	P-*value*	*Positive*	*Negative*	P-*value*
Primary tumor						
CD49f						
Positive	5 (14.3%)	14 (14.3%)	1.000	16 (25.8%)	3 (4.3%)	0.001
Negative	30 (85.7%)	84 (85.7%)		46 (74.2%)	67 (95.7%)	
CD44						
Positive	17 (48.6%)	46 (46.9%)	1.000			
Negative	18 (51.4%)	52 (53.1%)				
Lymph node metastasis						
CD49f						
Positive	16 (59.3%)	24 (24.2%)	0.001	27 (45.0%)	12 (18.8%)	0.002
Negative	11 (40.7%)	75 (75.8%)		33 (55.0%)	52 (81.2%)	
CD44						
Positive	22 (73.3%)	42 (41.1%)	0.003			
Negative	8 (26.7%)	60 (58.9%)				

P-Cadherin Expression in Metastatic Lymph Nodes is Associated with Poor Clinical Outcome

We set out to explore whether evaluating the expression of breast cancer stem cell markers at diagnosis in patients with lymph node metastatic disease could allow the prediction of clinical outcome. Regarding the expression of P-cadherin, CD44, or CD49f exclusively in primary carcinomas, analysis of survival showed no significant association with overall survival or disease-free survival (Supplementary Figure S1). However, P-cadherin expression in tumor-associated lymph nodes impacted in the patient outcome, denoting a significantly worse disease-free survival and overall survival (Figure 2a, overall survival $P=0.002$, disease-free survival $P=0.004$).

Considering CD44 and CD49f expression in metastatic lymph nodes, no significant impact was observed for the survival functions (Figure 2b and Figure 2c, overall survival $P=0.796$ and disease-free survival $P=0.704$ for CD44, and overall survival $P=0.763$ and disease-free survival $P=0.303$ for CD49f).

The Gain of P-Cadherin and CD49f Expression in Metastatic Lymph Nodes is Associated with the Triple-Negative Molecular Subtype

Breast cancer stem cell markers presented a clear distinction of expression (gain or loss) between primary tumors and matched lymph nodes metastasis, with 19.2% of the cases showing a discordant result for P-cadherin, 21.5% for CD44 and 26.2% for CD49f (Table 2).

Concerning P-cadherin expression, the large number of cases that remain negative in both primary tumor and node metastases were classified as luminal A (61/88, 69.3.%). Interestingly, the majority of the cases with P-cadherin loss were also luminal A, representing 60.0% (9/15) of the cases. Regarding the ones that gained P-cadherin expression in the lymph nodes, these were mainly from the triple-negative molecular subtype, corresponding to 63.6% (7/11) of the cases (Supplementary Table S4) ($P<0.001$). Accordingly, the triple-negative molecular subtype was significantly enriched in the gain of P-cadherin expression: 38.0% (7/25) of triple-negative carcinomas gained P-cadherin expression, in comparison with only 8.14% (11/135) in the whole series (Figure 2d and Supplementary Table S5).

Regarding CD44, the cases that were negative in both matched neoplastic lesions were frequently classified as luminal A (36/54, 66.7%) ($P=0.020$) (Supplementary Table S5). We could observe that a high percentage of HER2-overexpressing tumors showed CD44 loss in the lymph nodes (22.2% (4/18)), in comparison with the whole series (10% (13/130)) (Figure 2d and Supplementary Table S5).

Regarding CD49f expression, 71.2% (57/80) of the cases that remained negative were scored as luminal A tumors. The ones where CD49f positivity was maintained were similarly distributed along the four molecular subtypes. As previously noticed, the most frequent alteration was the gain of CD49f expression in metastatic lymph nodes (Table 2 and Figure 2d) and, notably, this phenotype was significantly enriched in the triple-negative molecular subtype: 47.8% (11/23) of triple-negative carcinomas gained CD49f expression, in comparison with only 21.4% (27/126) in the whole series. None of the HER2-overexpressing tumors showed CD49f loss ($P<0.001$) (Figure 2d and Supplementary Table S5).

The Change in P-Cadherin Expression between Primary Carcinomas and Synchronous Metastatic Lymph Nodes Predicts Patients' Survival

As the pattern of expression of breast cancer stem cell markers was discordant between primary breast carcinomas and metastatic lymph nodes, we set out

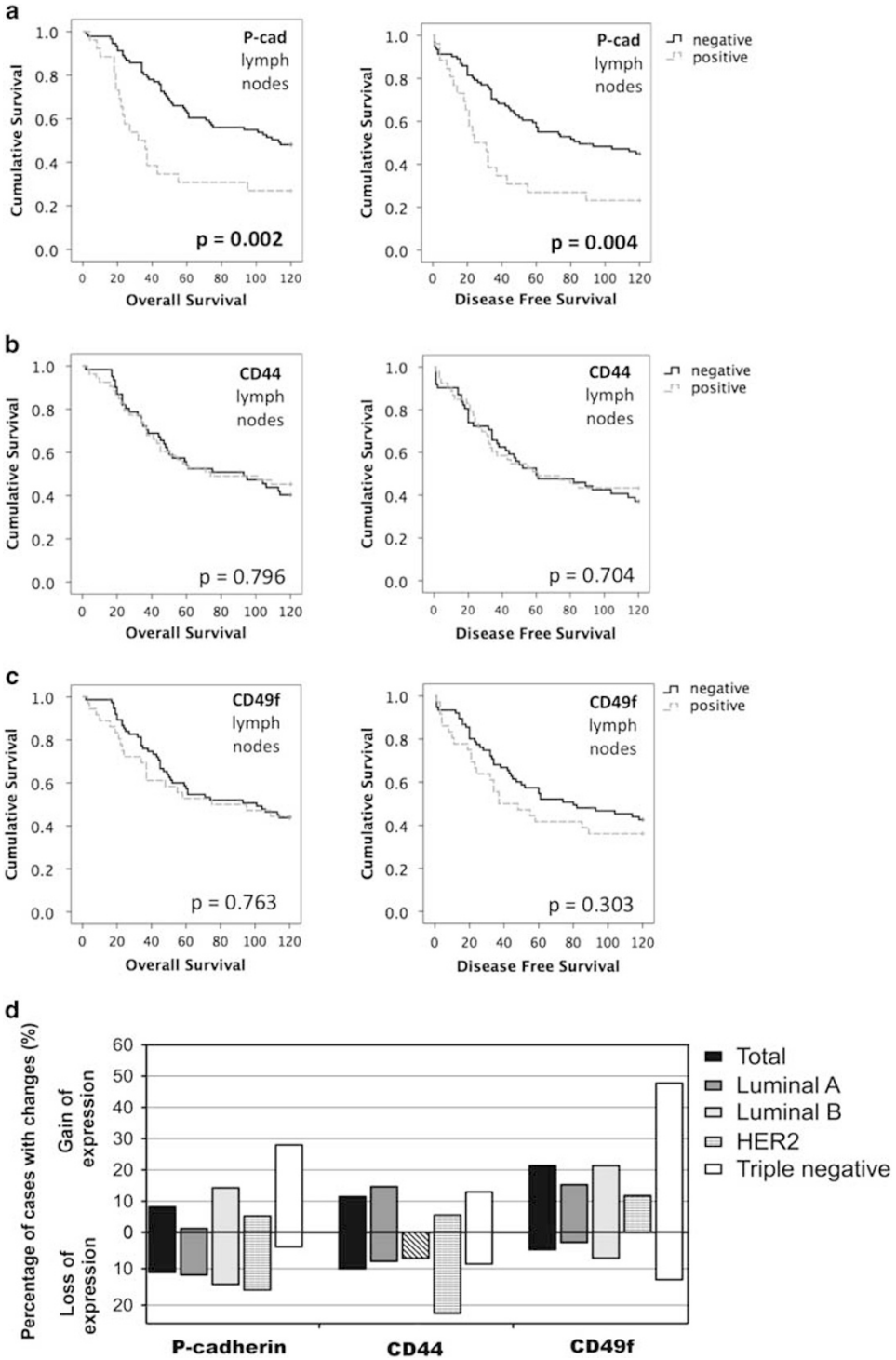

Figure 2 Analysis of breast carcinoma patients according to breast cancer stem cell markers expression in lymph nodes metastasis. Kaplan–Meier survival plots of overall survival and disease-free survival considering the expression of P-cadherin (**a**), CD44 (**b**), and CD49f (**c**) in metastasized lymph nodes in the series of invasive breast carcinomas. (**d**) Graphical representation of the percentage of cases with gain and loss of P-cadherin, CD44 and CD49f in metastatic lymph nodes *vs* primary carcinomas within each molecular subtype. For detailed expression, see Supplementary Table S3.

to investigate whether the variation in P-cadherin, CD44, and CD49f expression allowed the stratification of patients regarding clinical behavior. Remarkably, we found that breast cancer cases without P-cadherin expression in the primary tumors, but with *de novo* expression in the corresponding metastatic lymph nodes, were from patients exhibiting the worst overall survival and disease-free

Table 2 Frequency of cases according to the change in expression of breast cancer stem cell markers comparing invasive primary breast carcinomas with axillary lymph node metastasis

	P-cadherin (total = 135)	*CD44 (total = 130) (missing* n = *5)*	*CD49f (total = 126) (missing* n = *9)*
Remains positive	21 (15.6%)	48 (36.9%)	13 (10.3%)
Remains negative	88 (65.2%)	54 (41.6%)	80 (63.5%)
Gain of expression	11 (8.2%)	15 (11.5%)	27 (21.4%)
Loss of expression	15 (11.0%)	13 (10.0%)	6 (4.8%)

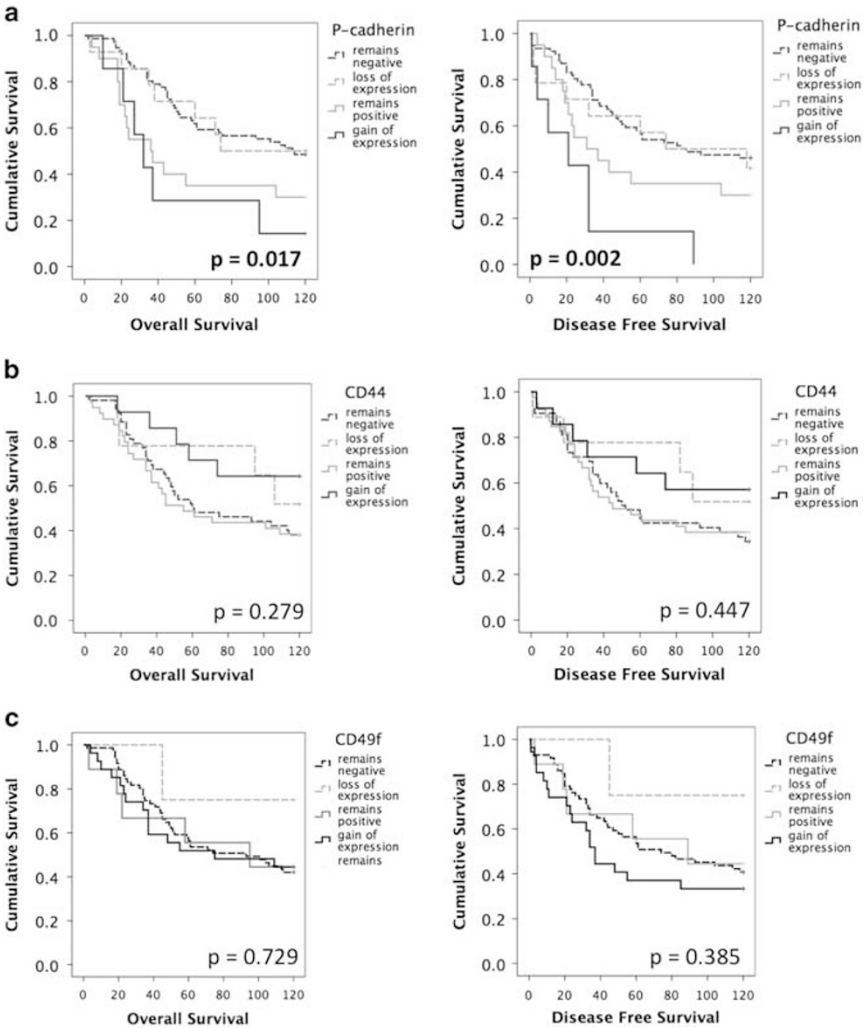

Figure 3 Survival analysis of breast carcinoma patients according to the change of breast cancer stem cell markers expression between primary tumor and lymph nodes metastasis. Kaplan–Meier survival plots of overall survival and disease-free survival of breast cancer patients considering the alterations in the expression of P-cadherin (**a**), CD44 (**b**), and CD49f (**c**) in primary tumors and matched lymph node metastasis.

survival, in comparison with other combinatorial possibilities. Accordingly, cases scored as P-cadherin positive in the primary tumor, but for which the matched lymph node was classified as negative, showed a better prognosis (Figure 3a, $P = 0.017$ for overall survival and $P = 0.002$ for disease-free survival). Univariate Cox regression analysis showed that P-cadherin gain of expression in the lymph nodes was the most relevant factor in predicting poor patients' prognosis, with a HR of 4.108 for disease-free survival, when compared with tumors remaining negative in the metastatic axillary site ($P = 0.001$). The same significant result was observed for OS (HR = 2.843, $P = 0.018$) (Table 3). Importantly, multivariate analysis showed that this effect was independent of primary tumor size, histological grade

Table 3 COX regression survival analysis

	Disease-free survival			*Overall survival*		
	HR	*95% CI*	P-*value*	*HR*	*95% CI*	P-*value*
Univariate analysis						
P-cadherin expression primary tumor/lymph node						
Remains negative ($n=77$, ref.)	1		—	1		—
Loss of expression ($n=14$)	1.146	0.537–2.446	0.724	1.008	0.451–2.255	0.984
Remains positive ($n=20$)	1.7	0.926–3.124	0.087	2.032	1.102–3.749	0.023
Gain of expression ($n=7$)	4.108	1.819–9.277	0.001	2.843	1.198–6.75	0.018
Multivariate analysis						
P-cadherin expression primary tumor/lymph node						
Remains negative ($n=73$, ref.)	1		—	1		—
Loss of expression ($n=14$)	1.157	0.516–2.596	0.723	1.020	0.430–2.418	0.964
Remains positive ($n=19$)	2.080	0.977–4.426	0.057	2.643	1.220–5.726	0.014
Gain of expression ($n=5$)	6.711	2.097–21.477	0.001	4.091	1.124–14.898	0.033
Histological grade						
Grade I ($n=11$, ref.)	1		—	1		—
Grade II ($n=29$)	2.206	0.628–7.753	0.217	2.092	0.583–7.503	0.257
Grade III ($n=71$)	2.444	0.729–8.196	0.148	2.153	0.636–7.294	0.218
Tumor size						
T1: < 2 cm ($n=12$, ref.)	1		—	1		—
T2: 2–5 cm ($n=82$)	2.130	0.757–5.997	0.152	2.032	0.720–5.736	0.181
T3: > 5 cm ($n=17$)	2.190	0.667–7.195	0.196	1.716	0.498–5.909	0.392
Molecular Subtype						
Luminal A ($n=70$, ref.)	1		—	1		—
Luminal B ($n=13$)	0.855	0.369–1.980	0.715	1.259	0.554–2.860	0.582
HER2 OE ($n=9$)	1.366	0.556–3.353	0.497	1.559	0.631–3.854	0.336
Triple negative ($n=19$)	0.635	0.266–1.514	0.305	0.598	0.236–1.515	0.278

Survival HR were evaluated according with the change in P-cadherin expression in primary tumors and paired metastatic lymph nodes (univariate analysis). The multivariate COX regression analysis included the effects histological grade, tumor size, and molecular subtypes. Missing cases were not considered for statistical analysis.

and breast cancer molecular subtype (HR = 6.711 for disease-free survival, $P=0.001$; HR = 4.091 for overall survival, $P=0.033$, Table 3).

Concerning CD44 and CD49f expression, the variation found between the primary tumor and the matched metastatic lymph nodes had no impact in patients' clinical outcome (Figure 3b and Figure 3c).

Discussion

In breast cancer patients, the presence of cancer cells in lymph nodes represents a powerful prognostic factor and the number of metastasized lymph nodes has a profound impact in the recommended adjuvant treatment by clinicians.[63] Notably, axillary positive carcinoma patients may present a heterogeneous clinical behavior.[14–16] Still, the biology of metastasized cells is not contemplated in clinical decisions, mainly owing to the assumption that their phenotype is the same as in the primary tumor.

In this study, we focused on studying the expression of breast cancer stem cell markers in primary tumors and matched lymph node metastases, as cancer cells with stem cell properties are considered crucial mediators of tumor heterogeneity, disease aggressiveness and potential seeds of metastasis.[29,31] Thus, we demonstrated that the sole analysis of the expression of the breast cancer stem cell marker P-cadherin exclusively in lymph nodes metastases is enough to predict poor clinical outcome of breast cancer patients (disease-free survival and overall survival). Although CD44 and CD49f were significantly associated with P-cadherin expression in lymph node metastasis, the expression of these markers did not impact patient survival.

As previously demonstrated by other groups, concerning the expression of hormone receptors and biomarkers involved in cell proliferation, differentiation, and apoptosis, we also found a high degree of concordance between breast cancer stem cell markers expression in primary carcinomas and matched node metastasis (80.8% for P-cadherin, 88.5% for CD44 and 73.8% for CD49f).[14,24–26,64] However, we highlight that a small, but relevant, difference was found comparing primary breast carcinomas and matched loco-regional disseminated tissue regarding their expression. Actually, we observed that this difference in the expression of P-cadherin showed a statistically significant impact in patient disease-free survival and overall survival (univariate and multivariate survival analysis). Specifically, we demonstrated that negative cases for P-cadherin in primary carcinomas, whereas being positive in lymph node metastases, were significantly associated with the highest probability of

patient relapse and death, possibly representing a high-risk group. In contrast, carcinomas that remain negative, or carcinomas that loose the expression of P-cadherin in metastatic lymph nodes showed a better patient prognosis. Thus, the evaluation of P-cadherin in the lymph nodes may explain why some breast carcinoma patients develop rapidly aggressive disease, whereas others never develop distant metastasis.[14–16] This specific result was striking when compared with the classical stem cell markers CD44 or CD49f, for which a comparable variation in expression did not impact the survival of patients. Considering that lymph nodes assessment of stem cell markers is not routinely performed in pathology laboratories, it would be interesting to further validate these results in a larger and independent series. This is particularly relevant, considering that the few existing immunohistochemical studies that report patient outcome comparing primary carcinomas and lymph node metastatic disease are limited to the classical prognostic markers, namely HER2,[24] and ER.[27,65]

Based on the results obtained, we thought that it would be interesting to cross this data with the information concerning distant dissemination sites for each one of these cases. Unfortunately, we were not able to obtain these records; however, it is reasonable to postulate that the tumors that correlated with earlier relapse and death were the ones that led to distant metastatic dissemination.

Notably, Adamczyk *et al.*[66,67] reported a dissimilar P-cadherin expression between primary tumor and synchronous lymph node metastases in up to 33.8% of invasive ductal breast carcinomas; however, the impact of this marker in patients' survival was not clarified in these studies. In that sense, our study brings an important novelty, including a 10-year follow-up of invasive breast carcinoma patients. Although additional studies will be needed to explain the poor prognostic phenotype conferred by the gain of P-cadherin expression within metastatic lymph nodes, these observations correlate with our previous results, which demonstrated that P-cadherin potentiates the invasive and colonizing potential of cancer cells and possibly contributes to the metastatic capacity of breast cancer cells.[46,47,49] Two independent studies have also analyzed P-cadherin as a putative predictive biomarker of distant metastatic risk in distinct human patient data sets, where its expression in the primary tumor was significantly associated with earlier occurrence of distant metastasis.[68,69] In our study, P-cadherin expression in the primary tumor was unable to stratify patients regarding disease-free survival or overall survival, in contrast to previously published studies.[45,49,52] However, it is worth mentioning that our series included only patients with lymph node positive disease at diagnosis, suggesting that the assessment of biomarkers solely in the primary lesions is not useful in stratifying these already poor prognosis patients.

In clinical terms, the evaluation of molecular status in residual tumors or in metastatic disease can potentially define better treatment options and estimate the actual patient prognosis. As CD44 or CD49f expression alone was not associated with patient survival in metastatic lymph nodes or primary tumors, our study points that P-cadherin is a stronger candidate biomarker for axillary-based breast cancer decisions in the clinical practice than the classical cancer stem cell markers. Future mechanistic studies should clarify the distinct P-cadherin expression profile found between primary carcinomas and matched lymph node metastasis. This may be due to the selection and expansion of a cell clone with specific advantageous features that originated in the primary tumor or the result of genetic and/or epigenetic changes promoted by the node microenvironment itself.

Finally, it is worth mentioning that patients included this study were not treated with neoadjuvant therapy and were considered treatment naïve regarding lymph node biomarker assessment. Furthermore, these patients were diagnosed from 1978 up to 1992, being subjected to the same standard adjuvant treatment following the protocols at that time. Patients with HER2+ tumors were not treated with specific targeted therapy (trastuzumab was not used in the clinic practice then) and patients with ER+ tumors were exclusively treated with tamoxifen. Therefore, although the adjuvant treatment could have influenced the results of the study, it is important to mention that an identical therapy protocol was administered to all patients, which means that under these conditions P-cadherin evaluation in lymph nodes metastases after surgery is an important predictor of disease progression.

Conclusion

The immunohistochemical analysis of biomarkers in lymph node malignancy is not routinely performed to ascertain disease outcome. In this report, we show that the breast cancer stem cell markers CD44, CD49f, and P-cadherin show a modest, but meaningful change in expression between primary lesions and paired regional lymph node metastatic disease. P-cadherin emerged as the only independent prognostic indicator. Despite only 8.2% of breast cancer patients gained P-cadherin expression, these cases presented the worst survival of the whole series, denoting a high-risk group of patients. This phenotype was mainly found in the triple-negative breast cancer subtype. Altogether, our study highlights that, although a more expansive study and independent validation will be required, the immunohistochemical detection of P-cadherin in primary carcinomas and matched metastatic nodes, can potentially help clinicians to better stratify patients according to prognosis and to improve the treatment options for breast cancer patients.

Acknowledgments

This work was funded by FEDER (Fundo Europeu de Desenvolvimento Regional) funds through COMPETE 2020—Operacional Programme for Competitiveness and Internationalisation—Portugal 2020 (POCI-01-0145-FEDER-007274), by Novartis and by FCT (Fundação para a Ciência e a Tecnologia) under the project PTDC/SAU-GMG/120049/2010. JP is a FCT Investigator (Development Grant) and AFV is a FCT fellow (SFRH/BPD/90303/2012). IPATIMUP integrates the i3s Research Unit, which is partially supported by FCT.

Author contributions

Conception and design: AFV, MRD, JP; development of methodology: AFV, MRD, MG, ML; acquisition of data (provided and managed breast carcinoma cases): AFV, MRD, MG, JFC, IA, FS; analysis and interpretation of data: AFV, MRD, MG, JP; writing, review and/or revision of the manuscript: AFV, MRD, JP; administrative, technical, or material support (reporting and organizing data, constructing databases): AFV, MRD, MG; study supervision: JP. All authors read and approved the final manuscript.

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

1 Ferlay J, Soerjomataram I, Ervik M, *et al.* GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. International Agency for Research on Cancer, 2013.
2 Cardoso F, Costa A, Norton L, *et al.* ESO-ESMO 2nd international consensus guidelines for advanced breast cancer (ABC2). Breast 2014;23:489–502.
3 ESMO. The European Society of Medical Oncology, 2015 (available from http://www.esmo.org).
4 Zagouri F, Liakou P, Bartsch R, *et al.* Discrepancies between ESMO and NCCN breast cancer guidelines: an appraisal. Breast 2015;24:513–523.
5 Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, *et al.* Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. Cancer Res 2009;69:4116–4124.
6 Herschkowitz JI, Simin K, Weigman VJ, *et al.* Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol 2007;8:R76.
7 Perou CM, Jeffrey SS, van de Rijn M, *et al.* Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc Natl Acad Sci USA 1999;96:9212–9217.
8 Perou CM, Sorlie T, Eisen MB, *et al.* Molecular portraits of human breast tumours. Nature 2000;406:747–752.
9 Sorlie T, Perou CM, Tibshirani R, *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 2001;98:10869–10874.
10 Sorlie T, Tibshirani R, Parker J, *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci USA 2003;100:8418–8423.
11 O'Shaughnessy J. Extending survival with chemotherapy in metastatic breast cancer. Oncologist 2005;10 Suppl 3:20–29.
12 Surveillance, Epidemiology, and End Results (SEER) Program Research Data (1973-2013), National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2016, based on the November 2015 submission. (available from http://www.seer.cancer.gov).
13 Cardoso F, Harbeck N, Fallowfield L, *et al.* Locally recurrent or metastatic breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2012;23 Suppl 7:vii11–vii19.
14 D'Andrea MR, Limiti MR, Bari M, *et al.* Correlation between genetic and biological aspects in primary non-metastatic breast cancers and corresponding synchronous axillary lymph node metastasis. Breast Cancer Res Treat 2007;101:279–284.
15 Rosen PP, Groshen S, Saigo PE, *et al.* Pathological prognostic factors in stage I (T1N0M0) and stage II (T1N1M0) breast carcinoma: a study of 644 patients with median follow-up of 18 years. J Clin Oncol 1989;7: 1239–1251.
16 Hellman S. Karnofsky Memorial Lecture. Natural history of small breast cancers. J Clin Oncol 1994;12: 2229–2234.
17 Riker AI, Enkemann SA, Fodstad O, *et al.* The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis. BMC Med Genomics 2008;1:13.
18 Maruschke M, Hakenberg OW, Koczan D, *et al.* Expression profiling of metastatic renal cell carcinoma using gene set enrichment analysis. Int J Urol 2013;21: 46–51.
19 Minn AJ, Kang Y, Serganova I, *et al.* Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors. J Clin Invest 2005;115:44–55.
20 Ellsworth RE, Seebach J, Field LA, *et al.* A gene expression signature that defines breast cancer metastases. Clin Exp Metastasis 2009;26:205–213.
21 Poplawski AB, Jankowski M, Erickson SW, *et al.* Frequent genetic differences between matched primary and metastatic breast cancer provide an approach to identification of biomarkers for disease progression. Eur J Hum Genet 2010;18:560–568.
22 Weigelt B, Glas AM, Wessels LF, *et al.* Gene expression profiles of primary breast tumors maintained in distant metastases. Proc Natl Acad Sci USA 2003;100: 15901–15905.
23 Weigelt B, Wessels LF, Bosma AJ, *et al.* No common denominator for breast cancer lymph node metastasis. Br J Cancer 2005;93:924–932.
24 Falck AK, Ferno M, Bendahl PO, *et al.* Does analysis of biomarkers in tumor cells in lymph node metastases give additional prognostic information in primary breast cancer? World J Surg 2010;34:1434–1441.
25 Ieni A, Barresi V, Caltabiano R, *et al.* Discordance rate of HER2 status in primary breast carcinomas versus synchronous axillary lymph node metastases: a

multicenter retrospective investigation. Onco Targets Ther 2014;7:1267–1272.
26 Jensen JD, Knoop A, Ewertz M, *et al.* ER, HER2, and TOP2A expression in primary tumor, synchronous axillary nodes, and asynchronous metastases in breast cancer. Breast Cancer Res Treat 2012;132: 511–521.
27 Rossi S, Basso M, Strippoli A, *et al.* Hormone receptor status and HER2 expression in primary breast cancer compared with synchronous axillary metastases or recurrent metastatic disease. Clin Breast Cancer 2015;15:307–312.
28 Zhao S, Xu L, Liu W, *et al.* Comparison of the expression of prognostic biomarkers between primary tumor and axillary lymph node metastases in breast cancer. Int J Clin Exp Pathol 2015;8:5744–5748.
29 Oskarsson T, Batlle E, Massague J. Metastatic stem cells: sources, niches, and vital pathways. Cell Stem Cell 2014;14:306–321.
30 Geng SQ, Alexandrou AT, Li JJ. Breast cancer stem cells: multiple capacities in tumor metastasis. Cancer Lett 2014;349:1–7.
31 Reya T, Morrison SJ, Clarke MF, *et al.* Stem cells, cancer, and cancer stem cells. Nature 2001;414: 105–111.
32 Ponta H, Sherman L, Herrlich PA. CD44: from adhesion molecules to signalling regulators. Nat Rev Mol Cell Biol 2003;4:33–45.
33 Al-Hajj M, Wicha MS, Benito-Hernandez A, *et al.* Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci USA 2003;100: 3983–3988.
34 Collins AT, Berry PA, Hyde C, *et al.* Prospective identification of tumorigenic prostate cancer stem cells. Cancer Res 2005;65:10946–10951.
35 Li C, Heidt DG, Dalerba P, *et al.* Identification of pancreatic cancer stem cells. Cancer Res 2007;67: 1030–1037.
36 O'Brien CA, Pollett A, Gallinger S, *et al.* A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature 2007;445: 106–110.
37 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;89:857–866.
38 Eirew P, Stingl J, Raouf A, *et al.* A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. Nat Med 2008;14:1384–1389.
39 Stingl J, Eirew P, Ricketson I, *et al.* Purification and unique properties of mammary epithelial stem cells. Nature 2006;439:993–997.
40 Villadsen R, Fridriksdottir AJ, Ronnov-Jessen L, *et al.* Evidence for a stem cell hierarchy in the adult human breast. J Cell Biol 2007;177:87–101.
41 Cariati M, Naderi A, Brown JP, *et al.* Alpha-6 integrin is necessary for the tumourigenicity of a stem cell-like subpopulation within the MCF7 breast cancer cell line. Int J Cancer 2008;122:298–304.
42 Friedrichs K, Ruiz P, Franke F, *et al.* High expression level of alpha 6 integrin in human breast carcinoma is correlated with reduced survival. Cancer Res 1995;55: 901–906.
43 Ghebeh H, Sleiman GM, Manogaran PS, *et al.* Profiling of normal and malignant breast tissue show CD44high/CD24low phenotype as a predominant stem/progenitor marker when used in combination with Ep-CAM/CD49f markers. BMC Cancer 2013;13:289.
44 Shipitsin M, Campbell LL, Argani P, *et al.* Molecular definition of breast tumor heterogeneity. Cancer Cell 2007;11:259–273.
45 Vieira AF, Ricardo S, Ablett MP, *et al.* P-cadherin is co-expressed with Cd44 and Cd49f and mediates stem cell properties in basal-like breast cancer. Stem Cells 2012;30:854–864.
46 Vieira AF, Ribeiro AS, Dionisio MR, *et al.* P-cadherin signals through the laminin receptor alpha6beta4 integrin to induce stem cell and invasive properties in basal-like breast cancer cells. Oncotarget 2014;5: 679–692.
47 Ribeiro AS, Albergaria A, Sousa B, *et al.* Extracellular cleavage and shedding of P-cadherin: a mechanism underlying the invasive behaviour of breast cancer cells. Oncogene 2010;29:392–402.
48 Ribeiro AS, Sousa B, Carreto L, *et al.* P-cadherin functional role is dependent on E-cadherin cellular context: a proof of concept using the breast cancer model. J Pathol 2013;229:705–718.
49 Paredes J, Albergaria A, Oliveira JT, *et al.* P-cadherin overexpression is an indicator of clinical outcome in invasive breast carcinomas and is associated with CDH3 promoter hypomethylation. Clin Cancer Res 2005;11:5869–5877.
50 Gamallo C, Moreno-Bueno G, Sarrio D, *et al.* The prognostic significance of P-cadherin in infiltrating ductal breast carcinoma. Mod Pathol 2001;14: 650–654.
51 Paredes J, Milanezi F, Viegas L, *et al.* P-cadherin expression is associated with high-grade ductal carcinoma in situ of the breast. Virchows Arch 2002;440: 16–21.
52 Peralta Soler A, Knudsen KA, Salazar H, *et al.* P-cadherin expression in breast carcinoma indicates poor survival. Cancer 1999;86:1263–1272.
53 Turashvili G, McKinney SE, Goktepe O, *et al.* P-cadherin expression as a prognostic biomarker in a 3992 case tissue microarray series of breast cancer. Mod Pathol 2011;24:64–81.
54 Paredes J, Lopes N, Milanezi F, *et al.* P-cadherin and cytokeratin 5: useful adjunct markers to distinguish basal-like ductal carcinomas in situ. Virchows Arch 2007;450:73–80.
55 Arnes JB, Brunet JS, Stefansson I, *et al.* Placental cadherin and the basal epithelial phenotype of BRCA1-related breast cancer. Clin Cancer Res 2005;11: 4003–4011.
56 Vieira AF, Paredes J. P-cadherin and the journey to cancer metastasis. Mol Cancer 2015;14:178.
57 Senkus E, Kyriakides S, Ohno S, *et al.* Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2015;26 Suppl 5:v8–30.
58 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;101:736–750.
59 Palacios J, Benito N, Pizarro A, *et al.* Anomalous expression of P-cadherin in breast carcinoma. Correlation with E-cadherin expression and pathological features. Am J Pathol 1995;146:605–612.
60 Palacios J, Honrado E, Osorio A, *et al.* Immunohistochemical characteristics defined by tissue microarray of hereditary breast cancer not attributable to BRCA1 or BRCA2 mutations: differences from breast carcinomas

arising in BRCA1 and BRCA2 mutation carriers. Clin Cancer Res 2003;9:3606–3614.
61 Ricardo S, Vieira AF, Gerhard R, *et al.* Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. J Clin Pathol 2011;64:937–946.
62 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;25:963–974.
63 Haybittle JL, Blamey RW, Elston CW, *et al.* A prognostic index in primary breast cancer. Br J Cancer 1982;45:361–366.
64 Tsutsui S, Ohno S, Murakami S, *et al.* EGFR c-erbB2 and p53 protein in the primary lesions and paired metastatic regional lymph nodes in breast cancer. Eur J Surg Oncol 2002;28:383–387.
65 Ba JL, Liu CG, Jin F. Alterations in hormonal receptor expression and HER2 status between primary breast tumors and paired nodal metastases: discordance rates and prognosis. Asian Pac J Cancer Prev 2014;15: 9233–9239.
66 Adamczyk A, Niemiec J, Ambicka A, *et al.* Expression of ER/PR/HER2, basal markers and adhesion molecules in primary breast cancer and in lymph nodes metastases: a comparative immunohistochemical analysis. Pol J Pathol 2012;63:228–234.
67 Adamczyk A, Niemiec J, Ambicka A, *et al.* Survival of breast cancer patients according to changes in expression of selected markers between primary tumor and lymph node metastases. Biomark Med 2016;10:219–228.
68 Aleskandarany MA, Soria D, Green AR, *et al.* Markers of progression in early-stage invasive breast cancer: a predictive immunohistochemical panel algorithm for distant recurrence risk stratification. Breast Cancer Res Treat 2015;151:325–333.
69 Gyorffy B, Lanczky A, Eklund AC, *et al.* An online survival analysis tool to rapidly assess the effect of 22 277 genes on breast cancer prognosis using microarray data of 1809 patients. Breast Cancer Res Treat 2010;123:725–731.

Supplementary Information accompanies the paper on Modern Pathology website (http://www.nature.com/modpathol)