Carcinogenesis vol.35 no.1 pp.2–13, 2014
doi:10.1093/carcin/bgt279
Advance Access publication August 16, 2013

Cancer stemness in Wnt-driven mammmary tumorigenesis

Joana Monteiro, Claudia Gaspar, Wilfrid Richer, Patrick F. Franken, Andrea Sacchetti, Rosalie Joosten, Anouar Idali, Joana Brandao, Charles Decraene and Riccardo Fodde

Department of Pathology, Josephine Nefkens Institute, Erasmus MC 3000 CA Rotterdam, The Netherlands and 1Department of Translational Research, Institut Curie, Centre de Recherche and 2CNRS, UMR144, 75248 Paris, France
Present address: Instituto Medicina Molecular, Av Prof Egas Moniz, 1649-028 Lisboa, Portugal
*To whom correspondence should be addressed. Department of Pathology, Josephine Nefkens Institute, Erasmus MC, PO Box 2040, 3000 CA Rotterdam, The Netherlands. Tel: +31 10 704 44 90; Fax: +31 10 704 44 50; Email: r.fodde@erasmusmc.nl

Wnt signaling plays a central role in mammary stem cell (MaSC) homeostasis and in breast cancer. In particular, epigenetic alterations at different members of the Wnt pathway have been identified among triple-negative, basal-like breast cancers. Previously, we developed a mouse model for metaplastic breast adenocarcinoma, a subtype of triple-negative breast cancer. By targeting a hypomorphic mutations in the endogenous Apc gene (Apc1572T/+), here, by employing the CD24 and CD29 cell surface antigens, we have identified a subpopulation of mammary cancer stem cells (MaCSCs) from Apc1572T/+ capable of self-renewal and differentiation both in vivo and in vitro. Moreover, immunohistochemical analysis of micro- and macro lung metastases and preliminary intravenous transplantation assays suggest that the MaCSCs underlie metastasis at distant organ sites. Expression profiling of the normal and tumor cell subpopulations encompassing MaSCs and CSCs revealed that the normal stem cell compartment is more similar to tumor cells than to their own differentiated progenies. Accordingly, Wnt signaling appears to be active in both the normal and cancer stem cell compartments, although at different levels. By comparing normal with cancer mammary compartments, we identified a MaCSC gene signature able to predict outcome in breast cancer in man. Overall, our data indicate that constitutive Wnt signaling activation affects self-renewal and differentiation of MaSCs leading to metaplasia and basal-like adenocarcinomas.

Introduction

The heterogeneous composition of epithelial tumors encompassing distinct differentiated lineages offers a powerful model system to assess and characterize the role played by cancer stem cells (CSCs) not only in continuously fueling the neoplastic mass but also in local invasion and distant metastasis (1). Notwithstanding the controversies arisen around the operational definition of CSCs, a more pragmatic understanding of the CSCs concept is that tumor masses are characterized by hierarchical structures similar to those found in normal adult stem cell niches where self-renewing stem-like cells give rise to progenitors and to progressively more committed cell types. This is also supported by the observation according to which signal transduction pathways known to regulate homeostasis within adult stem cell niches by fine-tuning the delicate balance between self-renewal and differentiation are often deregulated and play key roles in malignant transformation of the same tissues (2). The canonical Wnt/β-catenin pathway, e.g. known to regulate self-renewal and differentiation not only in embryonic development but also in adult intestinal, epidermal, mammary and hematopoietic stem cell niches, is often constitutively activated in malignancies arising from the very same tissues (3).

Although generally associated with colon cancer, constitutive activation of the Wnt/β-catenin signaling pathway has also been reported in breast cancer (4) and in particular in triple-negative, basal-like breast cancers (TNBC) where both nuclear and cytosolic accumulation of β-catenin is predictive of poor outcome (5,6). TNBCs represent one of the six distinct subtypes into which breast cancers have been stratified on the basis of gene expression profiling (7,8), comprises 15–20% of the total breast cancer burden (9) and is characterized by a very aggressive clinical behavior reflected by the lack or limited expression of ER, PR and HER2/ERBB2, which makes them resistant to therapeutic protocols based on the pharmacological targeting of these very receptors (9). TNBCs include a very broad spectrum of subtypes ranging from poorly differentiated ductal adenocarcinomas to the multilinage and chemoresistant metastatic breast carcinomas encompassing epithelial, basal/myoepithelial and epithelial-to-mesenchymal transition components (10). More recently, six relatively homogeneous TNBC subtypes were identified; three of which (i.e. basal-like, mesenchymal-like, a mesenchymal stem-like) earmarked by constitutive activation of the Wnt pathway (11). Wnt activation among breast cancers is mainly achieved through upregulation of potentially oncogenic receptors and coreceptors such as FZD7 and LRP6 (12,13) but also by hypermethylation of the promoter of different Wnt antagonist genes such as APC, WIF1 and SFRP1 (14,15).

Further experimental evidence for the role played by Wnt signaling in basal-like mammary cancer has been provided by mouse models. Several members of the Wnt signaling cascade, including the Wnt1 ligand and β-catenin, have been shown to result in mammary hyper- and neoplasia when overexpressed in a tissue-specific fashion in transgenic mice (16–20). Notably, these animals develop mammary tumors reminiscent of metaplastic breast carcinomas in man. More recently, our laboratory developed a novel mouse model for Wnt-driven mammary cancer by targeting a constitutive mutation, Apc1572T, in the endogenous Apc tumor suppressor gene (21). This allele results in intermediate levels of Wnt/β-catenin signaling activation and in multifocal mammary adenocarcinomas and subsequent pulmonary metastases in both genders. Notably, the histology of the primary mammary tumors in Apc1572T+ animals is highly heterogeneous with luminal, myoepithelial and squamous lineages, reminiscent of metaplastic breast carcinoma of the breast, a rare TNBC subtype characterized by an admixture of adenocarcinoma with areas of squamous, spindle cell and/or mesenchymal phenotype (21–23).

Normal stem/progenitor cells of the mammary gland [mammary stem cells (MaSCs)] of the mouse, earmarked by the Lin−CD29−CD24+ combination of cell surface markers, are able to reconstitute functional mammary glands in vivo even when transplanted as single cells in recipient animals (24). The same markers were proven useful to enrich for CSCs when applied, alone or in combination with CD61 or Thy1, to mammary tumors from the Tpr53−/− and MMTV-Wnt1 mouse models for breast cancer (25–27). Here, we have taken advantage of the heterogeneous cellular composition of mammary adenocarcinomas from the Apc1572T+ model to: (i) prospectively isolate and characterize the corresponding CD29−/CD24+ CSCs; (ii) compare them with normal stem cells of the mouse mammary gland and derive a CSC-specific signature able to predict breast cancer outcome in man and (iii) assay the CSCs’ capacity to form distant metastases, which fully recapitulate the primary lesion upon dissemination into the blood stream.

Abbreviations: CSC, cancer stem cell; FACS, fluorescence-activated cell sorting; IHC, immunohistochemistry; MaSC, mammary cancer stem cell; MaSC, mammary stem cell; mCSC, migrating cancer stem cell; NOD-SCID, non-obese diabetic-severe combined immunodeficiency; TNBC, triple-negative, basal-like breast cancer.

1These authors contributed equally to this work.

© The Author 2013. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com
Materials and methods

**Cell labeling and fluorescence-activated cell sorting analysis**

Blocking was performed with rat γ-globulin (Jackson Laboratories). Antibodies used for extracellular epitope labeling were CD24-FITC (1:500, 555428; BD Biosciences) and CD29-FITC (1:50, ab23834; Abcam). Lineage-specific (Lin⁺) cells were counterselected using the following antibodies: CD31-Biotin (1:50, 553371; BD Biosciences); Ter119-Biotin (1:500, 553672; BD Biosciences); CD45-Biotin (1:5000, 553078; BD Biosciences). A streptavidin-APC-labeled (1:2000, 554067; BD Biosciences) secondary antibody was employed for Lin⁺ detection. Incubations were performed at 4°C for 25 min. Live/dead cell discrimination was performed with Hoescht33258 (1:10 000; Invitrogen). Fluorescence-activated cell sorting (FACS) was performed with a FACSaria™ cell sorter (BD Biosciences).

**Immunohistochemistry and immunofluorescence**

Tissues were fixed in paraformaldehyde (4%) and embedded in paraffin. Sections (5 μm) were mounted on slides and stained by hematoxylin/eosin for routine histology. Antibodies employed for immunohistochemistry (IHC) analysis include: β-catenin (1:2000, 1247-1; Epitomics), Trom-1 (directed against a Ck8 epitope, 1:400; Hybridoma Bank), smooth muscle actin (Sma, 1:200, M0851; DakoCytmomation), Ck14 (1:10 000, PRB-155P; Covance) and Ck6 (1:3500, PRB-169P; Covance). The Ck6 and Ck14 antibodies are employed to detect hair follicle and skin differentiation, respectively. However, when employed at a lower dilution (1:1000), Ck6 also detects mammary epithelial cells. All IHC images were obtained with higher dilution (1:5000) aiming at the identification of squamous differentiation lineages. The same primary antibody dilutions were employed for immunofluorescence analysis, with rabbit anti-rat-FITC (Sigma) and goat anti-mouse-AP594 (Invitrogen) for signal detection.

**Transplantation assays**

Cells were sorted into DMEM:F12 medium supplemented with 10% fetal calf serum. Upon sorting, cells were resuspended in a mixture of DMEM:F12 and Matrigel (1:1) and injected into non-obese diabetic-severe combined immunodeficiency (NOD-SCID) female mice either subcutaneously or into non-cleared inguinal mammary glands in a total volume of 100 and 20 μl, respectively. For intravenous injections, cells were sorted in phosphate-buffered saline supplemented with 10% fetal calf serum and injected in the tail vein of NOD-SCID females in a maximum volume of 200 μl. These procedures were performed with animals under isoflurane/O₂ anesthesia.

**RNA preparation and quality control**

Five independent mammary adenocarcinomas from C57BL6J Apc⁰¹⁵⁷²Ţţ mice and three independently isolated pools of mammary glands from C57BL6J Apc⁰¹⁵⁷²Ţţ mice were employed to sort 10 000 cells of each of the following populations: Lin⁺CD29⁺CD24⁺, Lin⁺CD29⁻CD24⁺, Lin⁻CD29⁺CD24⁻, Lin⁻CD29⁻CD24⁻. Total RNA was isolated with RNasea Micro kit (QIAGEN) and quality controlled with RNA 6000 Pico and Nano LabChip kits (Agilent Technologies).

**Complementary RNA synthesis and microarray hybridization**

Complementary RNA synthesis was performed with total RNA using the Two-cycle Labelling kit (Affymetrix). GeneChip Mouse Genome 430 2.0 Array was hybridized, revealed and washed according to the Affymetrix protocol. Genechips were scanned using a 7G scanner (Affymetrix). The microarray (and reverse transcription–PCR) data discussed in this publication have been deposited in NCBI’s Gene Expression Omnibus (28) and are accessible through Gene Expression Omnibus Series accession numbers GSE40702 and GSE40704 (supersedes series: GSE40715; http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE40715).

**Differential gene expression analysis**

Using Partek Genomics Suite Software (Partek), after GCRMA normalization, differentially expressed genes were identified by analysis of variance using the false discovery rate method (1%). In addition, the modulations of gene expression value was up to 4.5 (corresponding to the estimated background cutoff, data not shown) and the corresponding fold change >1.5. Hierarchical clustering analysis and principal component analysis were performed with the Partek Genomics Suite Software (Partek). Hierarchical clustering analysis was carried out using the Ward’s method and Pearson’s dissimilarity for gene and sample clustering, respectively. Molecular and functional interactions were addressed using Ingenuity Pathway Analysis tools (Ingenuity® Systems).

**Survival study**

To determine the clinical significance of the newly identified mouse gene signature, we investigated whether its expressions in human tumors was correlated with clinical outcome of breast cancer patients by employing the survival package 2.35-8 in R 2.11.1 software. As described by Shpitzen et al. (29), we produced four derived gene signatures corresponding to up- and down-regulated genes to analyze previously published data sets of breast cancer expression profiling studies.

**Results**

The Lin⁺CD29⁺CD24⁺ tumor cell subpopulation from Apc⁰¹⁵⁷²Ţţ mammary metaplastic adenocarcinomas encompass CSCs capable of self-renewal and differentiation

Previously, it was shown that the cell surface antigens Lin⁺CD29⁺CD24⁺ identify a subpopulation that is highly enriched for mouse MaSCs able to reconstitute functional mammary glands in vivo. Even when transplanted as single cells in recipient animals (24). Moreover, this subpopulation was significantly expanded in premalignant but otherwise normal mammary glands of MMTV-Wnt1 mice (24). We tested whether the same surface markers could identify a subpopulation enriched in CSCs from Apc⁰¹⁵⁷²Ţţ mammary adenocarcinomas by using FACS. As depicted in Figure 1a, the ‘Lin⁺CD29⁺CD24⁺ subpopulation of Apc⁰¹⁵⁷²Ţţ mammary tumor cells represents ~2–3% of the bulk (Lin⁻) tumor cells. In agreement with a previous report (30), the ‘Lin⁺CD29⁺CD24⁺ population is almost exclusively encompassed by myoepithelial cells (smooth muscle actin positive and E-cadherin negative) (30), whereas the ‘Lin⁺CD29⁺CD24⁻ is predominantly composed by luminal (E-cadherin positive) cells, with a minority of myoepithelial/basal cells (double positive smooth muscle actin/E-cadherin) (Supplementary Figure S1, available at Carcinogenesis Online). In contrast, Lin⁺CD29⁺CD24⁺ and in particular Lin⁺CD29⁺CD24⁻ cells showed an enrichment of double positive myoepithelial/basal cells (Supplementary Figure S1, available at Carcinogenesis Online).

Next, we tested whether the sorted subpopulations were differentially enriched for tumor-initiating cells. To this aim, different multiplicities of Lin⁺ (bulk), Lin⁺CD29⁺CD24⁺ and Lin⁺CD29⁻CD24⁻ tumor cells were transplanted both subcutaneously and orthotopically (non-cleared mammary fat pad) in NOD-SCID animals (Table I). As few as 500 Lin⁺CD29⁺CD24⁺ cells sorted from primary Apc⁰¹⁵⁷²Ţţ mammary adenocarcinomas were able to form tumors in immune-deficient mice, whereas the Lin⁻ (bulk tumor cells) and Lin⁻CD29⁺CD24⁻ populations did so exclusively at higher multiplicities (120 000 and 5000, respectively; Table I). Notably, the tumors resulting from transplantation of Lin⁻CD29⁺CD24⁻ cells are highly reminiscent of the primary Apc⁰¹⁵⁷²Ţţ mammary adenocarcinomas in that they encompass luminal, myoepithelial and squamous lineages, as shown by IHC (Figure 1b, left column).

To assay self-renewal and differentiation capacity of the tumor-initiating cells from Apc⁰¹⁵⁷²Ţţ mammary adenocarcinomas, tumors obtained by transplanting limiting multiplicities of Lin⁺CD29⁺CD24⁺ cells into NOD-SCID recipients were again sorted for the CD29⁺CD24⁻ combination of surface antigens to allow serial transplantation of the mammary cancer stem cells (MaSCs). In total, two rounds of serial transplantations were carried out and the resulting tumors again analyzed by FACS and IHC. Both the FACS patterns and the heterogeneous histology (with squamous, myoepithelial and luminal cellular types) of the tumors obtained were highly reminiscent of the primary Apc⁰¹⁵⁷²Ţţ mammary adenocarcinomas (Table I and Figure 1b). To further assess the potential of the tumor-derived Lin⁺CD29⁺CD24⁺ cells to self-renew and differentiate, we tested their capacity to form organoids in 3D serum-free culture conditions when compared with bulk (Lin⁻) and Lin⁺CD29⁺CD24⁻ tumor cells. As shown in Supplementary Table S1, available at Carcinogenesis Online, organoid formation capacity was dramatically increased among the Lin⁺CD29⁺CD24⁺ cells. IHC analysis of the resulting organoids showed that the three main cell lineages characteristic of the primary Apc⁰¹⁵⁷²Ţţ mammary adenocarcinomas, namely luminal, myoepithelial and squamous, coexist in individual ex vivo organoids (Figure 2). The same organoids were passaged twice by single cell dissociation and replated in 3D culture conditions, which invariably resulted in de novo organoid formation (data not shown).
Fig. 1. (a) FACS plots representative of CD24 and CD29 (β1 integrin) expression in primary Apc\(^{1572T+}\) mammary carcinomas. The gating strategy was designed to separate the cells of interest from large aggregates and debris (initial gate on forward scatter (FSC) versus side scatter (SSC) plot), deplete dead cells (Hoechst 33258 staining), Lin\(^+\) cells (CD31\(^+\)CD45\(^+\)Ter119\(^+\), APC-labeled antibodies) and doublets/aggregates (standard gates on both FSC-width and SSC-width) (data not shown). Cells were then analyzed and sorted based on their combined CD24 and CD29 expression levels. (Left panels) FACS plots showing live, single, Lin\(^-\) cells analyzed for their staining pattern with unspecific FITC- and PE-labeled isotype control antibodies (upper left panel) and specific anti-CD24-PE and
Wnt-driven breast cancer stem cells

<table>
<thead>
<tr>
<th>Table I.</th>
<th>Overview of the transplantation assays of the Lin(^+), Lin(\text{CD29}^{+})/(\text{CD24}^{-}) and Lin(\text{CD29}^{+})/(\text{CD24}^{-}) FACSsorted subpopulations from Apc(^{1572T/+}) mammary adenocarcinomas (primary tumors) in NOD-SCID recipient mice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary tumors</td>
<td>0.5 × 10(^6)</td>
</tr>
<tr>
<td>Lin(^+)</td>
<td>1/2</td>
</tr>
<tr>
<td>Lin(\text{CD29}^{+})/(\text{CD24}^{-})</td>
<td>—</td>
</tr>
<tr>
<td>Lin(\text{CD29}^{+})/(\text{CD24}^{-})</td>
<td>—</td>
</tr>
<tr>
<td>First round</td>
<td>0.5 × 10(^{-1})</td>
</tr>
<tr>
<td>Lin(^+)</td>
<td>1/2</td>
</tr>
<tr>
<td>Lin(\text{CD29}^{+})/(\text{CD24}^{-})</td>
<td>—</td>
</tr>
<tr>
<td>Lin(\text{CD29}^{+})/(\text{CD24}^{-})</td>
<td>—</td>
</tr>
<tr>
<td>Second round</td>
<td>1.0 × 10(^{-1})</td>
</tr>
<tr>
<td>Lin(^+)</td>
<td>—</td>
</tr>
<tr>
<td>CD29(^+)/CD24(^{-})</td>
<td>—</td>
</tr>
<tr>
<td>CD29(^+)/CD24(^{-})</td>
<td>—</td>
</tr>
</tbody>
</table>

Data from orthotopic and subcutaneous assays from six independent primary tumors did not differ substantially and were therefore combined. Tumor growth was followed for a period of 6 months. Latency time varied between 31 and 140 days. The first and second transplantation rounds refer to the self-renewal assays where 5000 Lin\(\text{CD29}^{+}\)/\(\text{CD24}^{-}\) cells from the tumors grown from the primary transplantations were again transplanted in NOD-SCID recipients. *Tumors obtained from Lin\(\text{CD29}^{+}\)/\(\text{CD24}^{-}\) cells were smaller in size when compared with the corresponding tumor grown from Lin\(\text{CD29}^{+}\)/\(\text{CD24}^{-}\).* In this case, the exact amount of injected cells was 8300.

Hence, the Lin\(\text{CD29}^{+}\)/\(\text{CD24}^{-}\) subpopulation of Apc\(^{1572T}\) tumor cells encompass MaCSCs with tumor-propagating, self-renewing and differentiation capacities, as shown both in vivo and in vitro.

Gene expression analysis of normal and cancer mammary stem cells

Our results indicate that the Lin\(\text{CD29}^{+}\)/\(\text{CD24}^{-}\) cell surface antigens allow the prospective enrichment of normal and cancer mammary stem cells (MaSCs and MaCSCs) from wild-type glands and Apc\(^{1572T/+}\) adenocarcinomas, respectively. Even though the MaSCs and MaCSCs are only enriched and by no means ‘purified’ in the CD29\(^+\)/CD24\(^{-}\) -sorted subpopulation, comparison of their genome-wide expression profiles is likely to provide useful insights on the mechanisms underlying stemness in mammary homeostasis and cancer.

Single cell suspensions were prepared from wild-type mammary glands (Apc\(^{+/+}\)) and from Apc\(^{1572T/+}\) mammary adenocarcinomas. Ten thousand cells of each of the Lin\(^+\) (bulk), Lin\(\text{CD29}^{+}\)/\(\text{CD24}^{-}\) and Lin\(\text{CD29}^{+}\)/\(\text{CD24}^{-}\) subpopulations were sorted by FACS (Figure 1a and Supplementary Figure S2, available at Carcinogenesis Online) and their transcriptome analyzed using the GeneChip® mouse genome 430 2.0 Array (Affymetrix).

To address the relationships among the different cell subpopulations within the tumor and normal compartments, we performed principal component analysis (31) and unsupervised hierarchical clustering. Principal component analysis shows that the different cell subpopulations obtained from normal mammary gland and tumor samples group according to their origin with the only exception of the N-CD29\(^+\)/CD24\(^{-}\) subpopulation (enriched in normal stem/progenitor cells), which seems to constitute a separate entity (Figure 3a).

To confirm this observation, unsupervised hierarchical clustering was implemented using the most significant differentially expressed probe sets (~8500) across samples identified by estimating the statistical dispersion of the probe set expression values by the interquartile range method. As depicted in Figure 3b, all tumor-derived subpopulations cluster within the same branch of the dendrogram together with N-CD29\(^+\)/CD24\(^{-}\) (MaSCs encompassing), whereas the bulk and the more differentiated normal mammary gland cells are resolved into a distinct branch.

Statistically significant differentially expressed genes were obtained by pairwise comparison of gene expression profiles obtained from each cell subpopulation using analysis of variance and applying a false discovery rate of 1% (Supplementary Table S2, available at Carcinogenesis Online). Overall, the subpopulation encompassing the normal MaSCs (N-CD29\(^+\)/CD24\(^{-}\)) differ more from the normal differentiated lineages than the MaCSC-containing fraction differs from more committed tumor cells. Hence, the wild-type (Apc\(^{+/+}\)) mammary gland subpopulation encompassing normal stem and progenitor cells appears more similar to the Apc- mutant mammmary tumor cells than to their own differentiated progenies.

Further supervised clustering analysis revealed that signaling pathways known to play important roles in the regulation of homeostasis in stem cell niches (Tgf-β, Sonic Hedgehog and Tp53), are differentially expressed between normal- and tumor-derived mammary cells (Supplementary Figure S3, available at Carcinogenesis Online). However, supervised clustering performed with genes belonging to the canonical Wnt signaling pathway resolved the normal Apc\(^{1572T/+}\) cells (encompassing normal mammary stem/progenitor cells) together with all the tumor-derived subpopulations (Figure 3c), similar to what observed with unsupervised methods (Figure 3b) and, to a lesser degree, with the Notch pathway (Supplementary Figure S3b, available at Carcinogenesis Online). Nevertheless, when comparing the relative level of Wnt activation between the normal and tumor compartments encompassing MaSCs (N-CD29\(^+\)/CD24\(^{-}\)) and MaCSCs (T-CD29\(^+\)/CD24\(^{-}\)) by Ingenuity Pathway Analysis, Wnt signaling appears to be more active among the tumor-derived CD29\(^+\)/CD24\(^{-}\) cells when compared with their normal equivalents (Figure 3d). This was further validated by quantitative PCR analysis showing upregulation of Wnt target genes such as Axin2, C44 and Apcdd1 in T-CD29\(^+\)/CD24\(^{-}\) versus N-CD29\(^+\)/CD24\(^{-}\) cells (Figure 3e).

Hence, Wnt signaling appears to be active in both normal and tumor stem/progenitor cell compartments though at a higher level in the latter. This is likely to result from the difference in the underlying (epi)genetic mechanisms. We have shown previously that mammary tumors from the Apc\(^{1572T/+}\) model are characterized by somatic loss of the wild-type Apc allele (Apc\(^{1572T/L0H}\)) (21). On the other hand, activation of Wnt signaling in normal mammmary stem and progenitor cells is more likely to result from the epigenetic silencing of one or more of its member genes, as also suggested by the similar levels of expression of several Wnt genes between tumor and N-CD29\(^+\)/CD24\(^{-}\) cells (Figure 3f). In fact, expression profiling and subsequent quantitative PCR validation confirmed that the same Apc gene is expressed at very low levels in N-CD29\(^+\)/CD24\(^{-}\) cells, similar
to those observed in tumor cells (Figure 3e). This again points at the (epigenetic) silencing of the Apc tumor suppressor gene as one possible mechanism underlying the activation of canonical Wnt signaling in normal MaSCs. Nevertheless, bisulfite PCR analysis of the mouse Apc gene promoter in MaSCs failed to reveal any specific methylation pattern (data not shown). This indicates that other, possibly histone-mediated, epigenetic mechanisms regulate Apc gene expression levels in normal mammary stem and progenitor cells.

Nevertheless, also in agreement with the Ingenuity Pathway Analysis (Figure 3d), other well-known Wnt downstream targets were found to be upregulated in T-CD29+/CD24+ (and in tumor-derived cells at large) when compared with N-CD29+/CD24+ cells (Supplementary Figure S4a, available at Carcinogenesis Online). This is indicative of the relatively higher level of Wnt/β-catenin signaling activation in cancer cells compared with the normal stem/progenitor cell compartment, possibly due to the distinct underlying mechanisms (genetic versus epigenetic silencing). Notably, the intestinal stem cell marker Lgr5 (32) appears to earmark normal MaSCs but is not further upregulated in their malignant counterparts (Figure 3e).

To further validate the overall gene expression profiling data, a broad spectrum of genes found to be differentially expressed (n = 181) were analyzed by quantitative PCR on independently FACSorted samples from normal mammary glands (Apc+/-) and Apc1572T+/+ adenocarcinomas (Supplementary Table S3, available at Carcinogenesis Online). Overall, the results were confirmatory of the microarray data (Supplementary Figure S4, available at Carcinogenesis Online).

The mouse CSC signature from Apc1572T mammary tumors predicts clinical outcome in human breast cancer

In order to obtain a gene expression signature specific for the mouse MaCSCs and evaluate its relevance for breast cancer in women, we hypothesized that the distribution of the SC- and CSC-specific genes reflects the proportion of SCs and CSCs present in each cell subpopulation. Moreover, we postulated that MaCSCs are likely to express genes that are also active in normal stem and progenitor cells and in the bulk of the tumor mass, which should be selectively removed to increase specificity. By following this general strategy (see legend to Supplementary Table S4, available at Carcinogenesis Online), a MaCSC-enriched gene signature was obtained composed of 231 genes, 136 and 95 of which were up- and downregulated in T-Lin+/CD29+/CD24+, respectively (Supplementary Table S4a and b, available at Carcinogenesis Online). Likewise, the final MaSC-enriched (N-Lin+/CD29+/CD24+) signature is composed of 376 genes, 353 and 23 of which were up- and downregulated, respectively (Supplementary Table S4c and d, available at Carcinogenesis Online). The corresponding human orthologous signatures encompass 129 upregulated and 86 downregulated (MaCSC+ and MaCSC- signatures) genes, and 348 upregulated and 23 downregulated (MaSC+ and MaSC- signatures) genes, respectively (Supplementary Table S4, available at Carcinogenesis Online).

Next, the newly generated MaSC- and MaCSC-enriched signatures were employed to assess their capacity to predict breast cancer clinical outcome in nine independent, previously published profiling data sets from cancer patients, seven of which from breast cancer cohorts (33–39), one from lung adenocarcinoma (40) and one from gastric cancer patients (41). None of these cancer cohorts included patients who received chemotherapy or hormonal therapy. Survival analysis shows that the whole ortholog MaCSC signature (n = 231 genes) is significantly associated with poor overall survival, relapse- and distant metastasis-free survival in all breast cancer cohorts (P value ≤0.05) though not in the lung and stomach cancer cohorts. By applying the method by Shipitsin et al. (29), significantly improved results were obtained with the upregulated MaCSC+ signature genes alone (n = 136), whereas the MaCSC- signature (downregulated genes) was overall not significant. The same was true for both the MaSC-enriched signature, which did not correlate with any clinical outcomes (Figure 4 and Supplementary Figure S5, available at Carcinogenesis Online). These results indicate that the predictive ability of the MaCSC signature is encompassed by the 136 upregulated genes and is specific for breast malignancies.

Gene ontology analysis of the 136 genes encompassed within the MaCSC+ signature was performed in order to shed more light on their biological function. When compared with the MaCSC- signature (i.e. not predictive of overall survival, relapse- and distant metastasis-free survival), the upregulated genes appear to fall into very distinct gene ontology categories related to DNA metabolism, replication and repair, and to cell division (Supplementary Table S5, available at Carcinogenesis Online). Also, comparison of the MaCSC+ signature with previously published ones revealed a very limited degree of overlap (Supplementary Figure S6, available at Carcinogenesis Online).
Wnt-driven breast cancer stem cells

Hence, the newly developed signature is highly specific and is composed of a large number of new genes previously not reported as part of CSC-enriched signatures.

Do MaCSCs from Apc1572T adenocarcinomas underlie metastasis in distant organ sites?

According to the migrating CSCs (mCSCs) model, CSCs are not only responsible for tumor maintenance and growth at the primary site but also underlie local and distal metastases (42). This implies that, if caught at an early stage, metastatic lesions should encompass a relative enrichment in stem-/progenitor-like cells with limited differentiation. The availability of a mouse model for mammary cancer that spontaneously develops pulmonary metastases (21) allowed us to test this hypothesis both by IHC analysis of the early steps of the metastatic process and by direct assessment of the CSCs’ ability to form distant metastasis when compared with bulk and differentiated tumor cells.

As shown previously, Apc1572T/+ lung macrometastases display luminal, myoepithelial as well as squamous lineages, thus closely recapitulating the primary mammary adenocarcinomas (21). However, lung micrometastases found in Apc1572T/+ animals bearing primary mammary tumors revealed a more limited degree of differentiation with a relative increase in cells with intracellular β-catenin (Figure 5a). The relatively undifferentiated nature of these early lesions suggests that CSCs are likely to represent the cells of origin of the full-blown metastatic lesion.

Next, in order to provide additional support to the role played by MaCSCs in initiating metastatic lesions at distant organ sites, we resected primary mammary adenocarcinomas from F1 (C57Bl6 × 129OLA) Apc1572T/+ mice and isolated the different CD24/CD29 subpopulations by FACS to then inject them through the tail vein into recipient animals to mimic tumor cell dissemination. As shown in Supplementary Table S6, available at Carcinogenesis Online, although equal amount of cells from the different tumor subpopulations were intravenously injected (ranging from 2.6 × 10^4 up to 2.5 × 10^5), only the MaCSCs-encompassing Lin−CD24+CD29hi cells gave rise to metastatic lesions at a broad spectrum of organ sites. The number, anatomical location and time of onset of the observed lesions were largely dependent on the number of injected cells (Supplementary Table S6, available at Carcinogenesis Online). Most importantly, all metastatic lesions obtained from MaCSCs fully recapitulated the primary mammary carcinomas characteristic of Apc1572T/+ mice in that they encompassed luminal, myoepithelial as well as squamous lineages (Figure 5b). Overall, although the number of tumors employed in the tail vein transplantation assay is admittedly underpowered and do not allow us to draw more definitive conclusions on the MaCSC’s role as cell of origin of the distant metastases, the cumulative results of the IHC analyses and transplantation assays point at the capacity of these CSCs from Apc1572T/+ adenocarcinomas to underlie not only tumor maintenance and growth of the primary cancers but are also to initiate the pulmonary metastasis. Nevertheless, also in view of the complex experimental
nature of the tail vein transplantation assays that limited the number of animals employed, definitive evidence for the role of MaCSCs at cells of origin of the distant metastases will be obtained from lineage-tracing studies.

Discussion

Apart from their role in tumor maintenance and growth at the primary site, the intrinsic plasticity that defines CSCs is consistent with their alleged functions in local invasion, dissemination and metastasis formation at distant organ sites. From this perspective, tumors encompassing mixed differentiated lineages offer an excellent study model as the capacity of CSCs to give rise to a heterogeneous spectrum of tumor cell types at the primary site is possibly reflected by the cellular composition of the corresponding metastases. Depending on their degree of cell autonomy and the capacity of responding to specific secreted microenvironmental factors upon their dissemination to distant sites, these mCSCs are predicted to give rise to metastases, which will recapitulate the heterogeneous composition of the primary lesion. This yet hypothetical scenario arose from and found verification in the molecular and immunohistological analysis of Wnt-driven colon cancers where nuclear β-catenin accumulation earmarks CSCs at the leading edge of the primary tumor from where they detach, undergo epithelial-to-mesenchymal transition-like processes to then disseminate through-out stromal tissues, blood and lymphatic circulation. Thereafter, by means of the reverse phenomenon mesenchymal-to-epithelial transition, mCSCs colonize and metastasize distant organ sites (42,43).

As constitutive activation of the canonical Wnt/β-catenin signaling pathway underlies many different tumor types (3), the mCSCs model may apply to a much broader spectrum of cancer sites than those in the colon alone. In the specific case of breast cancer, ample experimental evidence is available supporting the fundamental and clinical relevance of the role played by Wnt signaling in breast cancer (4,44) and in particular in TNBC where intracellular accumulation of β-catenin is predictive of poor outcome (5,6,11). The cause–effect relationship between Wnt signaling and basal-like breast cancers is further supported by mouse models and in particular by the Apc1572T-targeted allele developed in our laboratory, where a constitutive mutation of the Apc tumor suppressor gene results in highly penetrant mammary adenocarcinomas reminiscent of metaplastic breast cancers in man, a TNBC subtype encompassing an admixture of adenocarcinoma cells with squamous, spindle cell and/or mesenchymal cell types (21–23). Notably, Apc1572T/+ mice also develop pulmonary metastases whose histological composition closely recapitulates the heterogeneity of the primary mammary cancers (21).

Here, by employing the surface antigen markers Lin−CD24+CD29hi, we were able to enrich mammary CSCs from the Apc1572T+/− model. Previously, CSCs were found to be enriched in a subpopulation of tumor cells expressing similar combinations of cell surface markers (Thy1+CD24+CD49f+CD45−) in the MMTV-Wnt1 transgenic mouse model for Wnt-driven mammary tumorigenesis (27) but also in the Tp53-null breast cancer mouse model (Lin−CD24+CD29hi) (25). In agreement with the Wnt-ON nature of the Apc1572T+/− mouse model, the heterogeneous composition of its mammary adenocarcinomas
encompassing luminal, myoepithelial and squamous lineages (21), and with the above-mentioned observations by Shackleton et al. (24), we successfully employed FACS of Lin−CD24−CD29−Apc1572T/LOH tumor cells and characterized them as encompassing bona fide CSCs in view of their tumor-propagating, self-renewal and differentiation capacities both in vitro and in vivo.

The identification of multipotent CSCs in mammary adenocarcinomas from Apc1572T/LOH mice is of interest in view of the current debate around the existence of a bipotent adult MaSC. Previous studies based on orthotopic transplantations strongly supported the existence of bipotent MaSCs in the adult mouse (24,45). However, more recent lineage-tracing work by the Blanpain laboratory indicate that long-lived, unipotent basal and luminal progenitor cells are established around the onset of puberty, which display extensive renewing capacities during morphogenesis and adult life as well as during several cycles of pregnancy (46). Hence, upon transplantation, the basal Lin−CD24−CD29− cells acquire multipotential features, which are not observed when the same cells are traced in vivo, i.e. under more physiological conditions. More recently, by means of lineage tracing of Wnt-ON (Axin2−) cells, it was confirmed that different unipotent MaSCs are specified at different developmental and adult stages, whereas Wnt-ON cells in the embryo earmark the prospective luminal lineage, prior to the onset of puberty the same cells are almost exclusively committed to the

![Image](97x414 to 577x754)

**Fig. 3.** (a) Spatial representation of the similarities within all cell subpopulations after principal component analysis (PCA) (31) and using the first three principal components. Green, N-Lin− (bulk); red, N-CD29hiCD24−; blue, N-CD29−CD24− (MaSCs); yellow, T-CD29−CD24− (MaSCs); purple, T-CD29−CD24−; azure, T-Lin− (bulk). (b) Unsupervised hierarchical clustering of all cell subpopulations using the most significant differentially expressed genes (~8500) across the samples. These genes have been identified by estimating the statistical dispersion of their expression values (i.e. interquartile range method). Each row represents a gene; each column represents a sample. The level of expression of each gene, in each sample, relative to the mean level of expression of that gene across all of the samples, is represented by using a red–green color scale as shown in the key. Red box, normal-derived cell subpopulation; blue box, tumor-derived cell subpopulation; green box, Lin−CD29−CD24+ cell subpopulation; orange box, Lin− cell population; purple box, Lin−CD29−CD24− cell subpopulation. (c) Unsupervised hierarchical clustering of genes belonging or less directly related to the canonical Wnt/β-catenin signaling pathway. Each row represents a gene; each column represents a cell sample. The level of expression of each gene, in each sample, relative to the mean level of expression of that gene across all of the samples, is represented by using a red–green color scale as shown in the key. Genes displayed in red have been validated by quantitative PCR. Of note, several Wnt genes (Fzd1, Fzd2, Fzd5, Wnt10a, Dkk1, Apc and Sox7) appear to have similar expression levels between wild-type and tumor-derived stem/progenitor cells. Red box, normal-derived cell subpopulation; blue box, tumor-derived cell subpopulation; green box, Lin−CD29−CD24+ cell subpopulation; orange box, Lin− cell population; purple box, Lin−CD29−CD24− cell subpopulation. (d) Wnt pathway comparative analysis between genes expressed in MaSCs and MaCSCs using Ingenuity Pathway Analysis. Red, genes upregulated in the MaCSCs; green, genes downregulated in the MaCSCs by comparison with the MaSCs. (e) Quantitative reverse transcription–PCR analysis of Apc and other representative members of the Wnt signaling pathway. The different cellular populations of the normal mammary gland (blue) and tumor (red) are represented by the geometrical symbols as indicated. Error bars represent standard deviation values obtained from assays performed in (biological) triplicates, with the exception of normal Lin− cells for which duplicates were employed. Statistical significance is represented by colored square brackets for three different comparisons: *(blue), normal Lin−CD29−CD24− versus normal Lin−CD29−CD24+; *(green), normal Lin−CD29−CD24+ versus tumor Lin−CD29−CD24− (P value 0.05, Mann–Whitney U-test). Note, the Wnt target and stem cell marker Lgr5 (32) was not significantly upregulated in the MaCSCs-encompassing compartment from Apc1572T/LOH mammary tumors when compared with N-CD29−CD24− (MaSCs containing) or with their more differentiated tumor counterparts. However, comparison between N-CD29−CD24− and differentiated mammary epithelial cells (N-CD29−CD24−) revealed a significant Lgr5 upregulation.
When lineage tracing is performed in the adult (virgin) mammary gland followed by multiple cycles of pregnancy and involution, the Wnt-competent cells appear to be bipotent as they give rise to both luminal and basal progenies. In view of these studies, the results here presented seem to indicate that constitutive activation of the Wnt/β-catenin signal transduction pathway, brought about by the Apc1572T allele followed by loss of heterozygosity of the wild-type Apc allele, expands and possibly enhances the subpopulation of multipotent stem/progenitor cancer cells leading to metaplastic adenocarcinomas encompassing basal, luminal and squamous lineages. Moreover, the observation according to which the distant metastases observed both in Apc1572T/+ animals and those obtained upon the preliminary tail vein injections of MaCSCs-encompassing subpopulations of tumor cells fully recapitulate the heterogeneity of the primary lesions, suggest that these cells are truly multipotent and cell autonomous.

In this study, by comparing genome-wide expression profiles derived from CD24/CD29-sorted subpopulations from wild-type mammary glands and Apc1572T/+ tumors on the same genetic background, we were able to derive a MaCSC-specific signature of 136 upregulated genes. The expression of their human orthologous genes significantly correlated with poor overall survival, relapse- and distant metastasis-free survival in seven independent, previously published profiling data sets from breast cancer patients. However, preliminary gene ontology analysis of the genes encompassed by the MaCSC+ signature did not reveal specific Wnt or epithelial-to-mesenchymal transition features, as previously reported. Instead DNA metabolism, replication and repair, and cell division were among the top five cellular and molecular functions in the MaCSC+ signature, similar to those identified in Lin−CD29hiCD24hi tumor cells from Tp53-null mice. Overall, the MaCSC+ signature derived from Apc1572T/+ tumors appears to be highly specific and composed of a large number of novel genes previously not reported as part of CSC signatures (Supplementary Figure S6, available at Carcinogenesis Online). Notably, the observed prognostic value of this and many other CSC-derived gene signatures is unlikely to be due to the recognition of stem-like cells within bulk tumor RNA preparations where CSCs are often represent a minority subpopulation and are as such diluted into the transcriptome of more committed progenitors and differentiated tumor cells. In this respect, Vincent Detours and collaborators convincingly demonstrated that 60% of the published breast cancer outcome signatures were not significantly better predictors than random signatures of identical size, whereas 23% were even worst predictors than the median random signature. This was interpreted as resulting from the general correlation of cell proliferation with the breast cancer transcriptome, which would then integrate most prognostic information in this disease. However, even by employing the method described by these authors to compare our mouse-derived signatures with randomly generated signatures of identical sizes, the MaCSC and in particular the MaCSC+ signature

![Fig. 4. Kaplan–Meier survival curves and corresponding log-rank test P value for the whole MaCSC, the positive and the negative MaCSC (CSC+, CSC−) and the positive and the negative MaSC (SC+, SC−) signatures. For two selected data sets, a correlation of the gene signatures with overall survival time (OS), relapse-free survival time (RFS) or distant metastasis-free survival times (DMFS) was calculated as described by Shipitsin et al. (29). Patients whose tumors did or did not express the genes encompassed by each signature (blue and red curves, respectively) were compared using log-rank tests (whole CSC, CSC+, CSC, SC+ and SC−) and Kaplan–Meier analysis (whole CSC and CSC+) using R. Log-rank test P values <0.05 were considered statistically significant (Kaplan–Meier curves marked by an asterisk). Accordingly, log-rank P values obtained for the lung and gastric cancer cohorts show that the MaCSC signature is specific for breast cancer. The results for all data sets are presented in Supplementary Figure S6, available at Carcinogenesis Online. Red, significant P value; green, no significant P value.](image-url)
Fig. 5. Mammary CSCs are the cell of origin of distant organ metastases in Apc\textsuperscript{1572Tfs}\textsuperscript{+} mice. (a) IHC analysis of Apc\textsuperscript{1572Tfs}\textsuperscript{+} macro- and micrometastasis. Analysis of differentiation (CK8, SMA, CK14) markers reveals a low level of differentiated lineages among micrometastases when compared with their macroscopically visible equivalents. Inversely, intracellular accumulation of $\beta$-catenin is clearly visible even at early stage of the metastasis process. Notably, pulmonary metastases were mostly found in animals with mixed background (C57Bl6/129OLA) characterized by a more aggressive manifestation of the malignant disease (data not shown). (b) IHC analysis of metastases from different organs obtained by tail vein injection of MaCSCs. HE, hematoxylin/eosin.
performed significantly better as predictor (Supplementary Figure S7, available at Carcinogenetcs Online).

Notwithstanding the mechanisms underlying its prognostic power, the notion that expression of genes differentially expressed in MaCSCs predict breast cancer outcome is indicative of the alleged role of these cells in metastasis formation at distant organ sites. Indeed, two yet preliminary lines of evidence point to the tropic, multipotent and cell autonomous nature of the mCSCs from Apc<sup>+/-</sup> mammary tumors. First, Apc<sup>+/-</sup> mice were shown to spontaneously develop pulmonary metastases, which fully recapitulate the multilineage feature of the primary adenocarcinomas with luminal, myoepithelial and squamous characteristics (21). Moreover, markers specific for the above-mentioned differentiated lineages are underrepresented among micrometastases identified in the lung of Apc<sup>+/-</sup> mice when compared with macrometastases. Inversely, β-catenin intra-cellular accumulation is more frequently observed in micro- than macrometastases, thus suggesting a metastasis-initiating role for the more undifferentiated and stem-like tumor cells (Figure 5a). In support of the IHC data, upon tail vein dissemination of the different CD24/CD29 subpopulations, only the MaCSCs-encompassing Lin<sup>-</sup>CD29<sup>+/hi</sup>CD24<sup>+</sup> cells were able to form multifocal metastases at a broad spectrum of organ sites, each fully recapitulating the histology of the primary mammary adenocarcinomas (Figure 5b). These results indicate that, under these specific experimental conditions, the MaCSCs appear to be multipotent, highly tropic and cell autonomous in their capacity to differentiate in the same lineages, independently of the distant organ site. However, as the Apc<sup>+/-</sup> mice were to date found to present with pulmonary metastases, it appears that under more physiological conditions of tumor cell dissemination from the mammary carcinomas, the migrating MaCSCs preferentially home into lung tissues. However, more definitive and indisputable evidence for the role of MaCSC as cell of origin of distant metastases will be obtained by Cre-Lox (e.g. with Lgr5 or Axin2 CreERT knock-in mice) lineage-tracing analysis.

In conclusion, we have shown that constitutive activation of Wnt signaling in the mouse mammary gland affects self-renewal and differentiation capacity of stem/progenitor cells leading to the establishment of MaCSCs, which form TNBC-like tumors at the primary site. The same MaCSCs, upon systemic dissemination, are also likely to underlie distant metastases. Further elucidation of the genetic and epigenetic alterations and mechanisms underlying the activation of Wnt and other signaling pathways in stem-like subpopulations of triple-negative breast cancer will probably offer novel therapeutic targets for this specific group of women with poor prognostic indications and limited treatment alternatives.

Supplementary material

Supplementary Materials and methods, Tables S1–S6 and Figures S1–S7 can be found at http://carcin.oxfordjournals.org/.

Funding

Dutch Cancer Society (EMCR 2001-2482); the Netherlands Organisation for Scientific Research (NWO/Vici 016.036.636); BSIK (Kennisinfrastructuur) program of the Dutch Government (03038); the Netherlands Institute for Regenerative Medicine (www.nirm.nl); EU FP6 and FP7 consortia ‘Migrating Cancer Stem Cells’ (037297; www.mescs.eu); Integrated concept of tumor metastasis (HEALTH-F2-2008-201662; http://itgmv1.fzk.de/www/tumic/tumic_main.htm); the Portuguese Foundation for Science and Technology (FCT; http://www.fct.pt/index.pt.html.cn).

Acknowledgements

The authors are grateful to Mr F. Panne for his assistance with the artwork and to Dr J. Martens for critically reading the manuscript.

Conflict of Interest Statement: None declared.

References


Received March 8, 2013; revised July 21, 2013; accepted August 1, 2013