Basal carcinoma of the breast revisited: an old entity with new interpretations

E Korsching,¹ S S Jeffrey,² W Meinerz,⁴ T Decker,¹ W Boecker,¹ H Buerger³

ABSTRACT
The introduction of global gene expression analysis in breast cancer research has focused attention onto a repeatedly described subgroup of invasive breast cancer, the basal-like carcinomas. This subgroup is characterised by the expression of high-molecular weight cytokeratin 5, 14 and 17; using immunohistochemical diagnosis, it represents approximately 7–20% of invasive breast cancers. Some of these tumours fulfil the criteria of grade 3 invasive ductal carcinoma, the so-called triple negative carcinomas. However, other rare subgroups of morphological, medullary and myoepithelial carcinomas also belong to this entity. Even though the initial clinical prognostic relevance of basal-like breast cancers may have been overestimated, its distinctive biology generates many questions regarding the pathogenesis, chemosensitivity and optimal clinical management of this subgroup. Physiological progenitor cells within the normal female breast share essential immunohistochemical features with basal-like breast cancers. Although the exact relationship between subgroups of normal breast cells and their respective malignant counterparts is still under investigation, the major hallmarks of physiological progenitor cells are either maintained or reactivated by distinct genetic changes in basal breast cancer cells. This review will discuss the impact of these findings on our global understanding of breast cancer pathogenesis, especially from the perspective of its potential histogenesis. Clinical consequences and potential future research directions driven by the definition of basal breast cancers will also be discussed.

CLASSIFICATION OF BREAST CANCER ACCORDING TO GENOTYPE AND IMMUNOPHENOTYPE
The classification of invasive breast cancer for surgical, prognostic and predictive purposes is part of the basic workup of every breast cancer specimen, and usually involves molecular and immunohistochemical assays.¹ These conventional, mainly morphology-based approaches have been clinically validated and serve as the gold-standard for prognosticating patient outcome. Other molecular markers, such as the oestrogen receptor (ER), progesterone receptor (PR), and expression of the growth factor receptor HER2, provide therapeutic predictive value, and are of central importance for the clinical management of breast cancer. However, many other genes have limited clinical value, and the subclassification of breast cancers based on the expression of these single genes offers incomplete prognostic and predictive information. Thus, investigators are continually searching for new breast cancer biomarkers, especially those that predict therapeutic efficacy of conventional and molecular-based chemotherapeutic treatments.

Novel classification systems based on the definition of cytogenetic alteration patterns have given rise to the postulation of multiple, parallel pathogenetic pathways in breast cancer.² For example, the distribution of 16q-losses in different grades of breast cancers shows that invasive breast cancer is comprised of multiple distinct subgroups that do not necessarily arise from sequential genetic and morphological dedifferentiation steps.³ Global gene expression profiling using microarray analysis extended our current knowledge and promises to further improve predictive and prognostic power in breast cancer pathology. Gene expression studies have used different microarray platforms to address slightly different questions. One research approach used supervised classification techniques to define gene sets for prognostic purposes, identifying genes that were differentially expressed in patients who, within a finite period of time, did and did not develop metastatic disease (e.g. “prognostic 70-gene set”).⁴ Others used different methods to determine gene signatures that would classify breast cancers into distinct molecular subtypes (the Stanford/Norway “intrinsic” gene set”).⁵ The first approach was further refined by associating the “prognostic 70-gene set” with a “core serum response” gene set—considered representative of wound healing processes—to generate a better predictor of clinical outcome in patients with earlier stage breast cancer.⁶ Classification of breast cancer using the “intrinsic” gene set defined by Perou and colleagues has stimulated recent intense interest in the pathophysiology and best treatments for diverse tumour subtypes. However, it has also generated confusion regarding appropriate subtype nomenclature.

Two decades ago, Dairkee et al described a small subgroup of breast cancers that, similar to cells from normal basal epithelium, exclusively expressed a high molecular weight basal cytokeratin, and that had poor prognostic characteristics, postulating that these cancers may originate in “basally located precursor cells often referred to as ‘stem cells’ and may represent tumours of the undifferentiated stem cell”.¹¹ A more recent study¹² analysed a group of 295 breast cancers using five diverse gene expression signature sets, including classification gene sets and prognostic gene sets; despite each signature set containing different genes, there was good concordance of patient-specific outcome predictions for the majority of the breast cancer patients. In other words, individual tumours appear to have a good or bad prognostic phenotype, regardless of the gene set used to define...
that phenotype. This means that multiple feature patterns can define and distinguish specific tumour subtypes, but no one pattern can adequately explain the causal development of the different phenotypes.

**BASAL BREAST CANCER**

Basal cytokeratins (Ck) represent a large number of high-molecular weight (HMW) cytokeratins which, according to the catalogue of human cytokeratin expression, are mainly seen in the basal cell layers of stratified epithelium. In the human breast, these cytokeratins are also expressed in the basally-located myoepithelial cell layer and in a small proportion of luminal, glandular cells. The dual meaning of the term “basal” in the breast induced confusion, as “basalness” was often interpreted as a sign of a myoepithelial origin. However, as shown below and discussed previously, this explanation might not be sufficient.

A subgroup of invasive breast cancers expressing basal cytokeratins was first described in 1987 and 1996. More recent studies have shown that the incidence among Caucasian women ranges around 16%, with significant differences among different ethnic subgroups. Further studies gave additional insights into the characteristics of this tumour subgroup, such as expression of hyperproliferation-associated cytokeratins and the coexpression of basement membrane material in these tumours. Recent work studying Ck 5 expression showed that this group is comprised of not only grade 3 invasive ductal carcinomas (NOS), but also other histological breast cancer subgroups, such as metaplastic and medullary carcinomas, with or without associated \( \text{BRCA1} \)

---

**Figure 1** Cluster analysis of 527 invasive breast cancers according to their immunohistochemical expression profile. The methods, including the source, dilution and pretreatment for the respective antibodies have been published previously. Missing immunohistochemical data (9% of all data) were replaced by the median of that specific score. Hierarchical cluster analysis (method: average linkage, agglomerative coefficient (ac, cluster quality measure) from 0 [poor] to 1 [good]) based on the correlation was applied. Evaluations were performed with the statistical platform S-Plus V.6.r2 using the function “agnes” based on published algorithms. The results were visualised in a dendrogram showing graphically the similarity of the vectors with the respective immunohistochemical observations. Four different quantification schemes for the expression rates of Ck 5 were analysed ([A] is based on all score levels; [B] to [D] are derivatives of [A]). [A] Four levels: negative, weak positive (>10%), positive (>25%) and strong positive (>50% tumour cells); (B) Two levels: threshold 10% (negative level // all positive levels); (C) Two levels: threshold 25% (negative + weak positive level // positive + strong positive level). (D) Two levels: threshold 50% (negative + weak positive positive level // strong positive level). It is noteworthy that the different thresholds have hardly any influence on the cluster analysis, indicating that no definite threshold defines the “basalness” of breast carcinomas.

---

Table 1 Overview of the design and most important results of studies focusing on the prognostic and/or predictive value of basal cytokeratin expression in invasive breast cancer

<table>
<thead>
<tr>
<th>Authors</th>
<th>Source of tissue</th>
<th>Clinical data</th>
<th>Type of basal cytokeratin</th>
<th>Threshold for the expression of basal cytokeratins</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rakha et al.</td>
<td>TMA</td>
<td>Not specified in detail</td>
<td>Ck 5/6, Ck 14</td>
<td>10%</td>
<td>Expression of Ck 5/6 and/or Ck 14 is sufficient for the definition of basal breast cancer</td>
</tr>
<tr>
<td>Jumpannen et al.</td>
<td>TMA</td>
<td>No adjuvant treatment</td>
<td>Ck 5/Ck 14/p63 antibody cocktail</td>
<td>20%</td>
<td>No impact on overall survival in ER-negative breast cancers</td>
</tr>
<tr>
<td>Abd El-Rehim et al.</td>
<td>TMA</td>
<td>Not specified in detail</td>
<td>Ck 5/6, Ck 14</td>
<td>Any positive cell</td>
<td>Expression of basal cytokeratins is an independent prognostic marker At least four different immunohistochemical phenotypes can be defined</td>
</tr>
<tr>
<td>Laakso et al.</td>
<td>Tissue sections</td>
<td>(a) without adjuvant treatment (b) dose-escalated adjuvant chemotherapy</td>
<td>Ck 5/Ck 14/p63 antibody cocktail</td>
<td>5% 70% for the “true” basal subtype The true basal breast cancer subtype in contrast to the basoluminal subtype is not associated with HER2-amplifications and showed a significantly different prognosis No prognostic value of the expression of basal cytokeratins in high-risk patients</td>
<td></td>
</tr>
<tr>
<td>Van de Rijn et al.</td>
<td>TMA</td>
<td>Patients with and without adjuvant chemotherapy, and anti-hormone therapy, respectively</td>
<td>Ck 17 and/or Ck 5/6</td>
<td>Invasive tumour cells with strong staining Basal cytokeratin expression was an independent prognostic marker in node-negative breast cancer</td>
<td></td>
</tr>
<tr>
<td>Nielsen et al.</td>
<td>TMA</td>
<td>Patients with adjuvant and neoadjuvant chemotherapy, anti-hormone therapy, respectively</td>
<td>Ck 17 and/or Ck 5/6</td>
<td>Invasive tumour cells with &lt;20% of weak staining A panel of 4 antibodies (ER, Ck 5/6, HER2 and EGFR) can accurately identify basal-like tumours</td>
<td></td>
</tr>
<tr>
<td>Jones et al.</td>
<td>Tissue sections of grade 3 breast cancers</td>
<td>Not specified in detail</td>
<td>Ck 14</td>
<td>Any positive cell Basal carcinomas have a decreased overall survival and show a distinct chromosomal alteration pattern</td>
<td></td>
</tr>
<tr>
<td>Fulford et al.</td>
<td>Tissue sections of grade 3 breast cancers</td>
<td>Not specified in detail</td>
<td>Ck 14</td>
<td>At least 1% of invasive tumour cells Basal carcinomas develop less bone and liver metastasis. Overall survival was improved in basal grade 3 carcinomas compared to non-basal grade 3 carcinomas. In metastatic disease the prognosis was decreased</td>
<td></td>
</tr>
<tr>
<td>Banerjee et al.</td>
<td>Tissue sections of grade 3 breast cancers</td>
<td>Not specified in detail</td>
<td>Ck 5/6, Ck 14, Ck 17</td>
<td>Any positive cell Significantly increased recurrence rate and overall survival Adjuvant anthracycline-based chemotherapy seems less effective compared to grade matched controls</td>
<td></td>
</tr>
</tbody>
</table>

TMA, tissue microarray.

It is noteworthy that the vast majority of these tumours always coexpress a multitude of other cytokeratins.27,28 These tumours also show specific genetic features. In contrast to all other breast cancer subgroups, these tumours, especially the metaplastic subgroup, harbour EGFR-amplifications,29 otherwise rarely seen in breast cancer.30,31 A more recent study characterised an increased expression of genes located on 6q21 and 12p13 as hallmarks of medullary carcinomas.32 A significant proportion of basal carcinomas are associated with a BRCA1 germ line mutation27,33 or sporadic p53-mutations and/or p53-protein overexpression. Compared with Ck 5/14 negative breast cancers, a significantly lower rate of 16q-losses has been described.34-37 However, this significance vanished if the same comparison was made with invasive ductal grade 3 carcinomas only.38

In general, the initially described prognostic relevance has now been challenged by various other reports. Although some initial studies evaluated prognosis without consideration of tumour grade, it has since been shown that the prognostic significance of the basal subtype disappears when evaluated alongside other high risk variables in a representative series of breast cancer patients,39 in ER-negative,40 and in grade 3 breast cancers.41 Moreover, a subgroup of basal carcinomas has been associated with an improved prognosis.42 More recent gene expression analysis also revealed that a prognostic difference could not be elaborated between the basal and the ER negative/HER2-overexpressing breast cancer subgroups.12 However, the criteria used to define basal breast cancer, as shown below, differ significantly and many accumulated results have this potential bias when compared. Another significant source of bias is that most studies grouped patients together who were undergoing different treatment protocols, again cautioning us not to over-interpret these results. To answer future questions about prognosis and appropriate therapies, controlled prospective clinical trials with uniformly treated patients are mandatory.

MORPHOLOGICAL AND IMMUNOHISTOCHEMICAL CHARACTERISATION OF BASAL BREAST CANCER

Morphologically, basal carcinomas are mainly grade 3 invasive ductal carcinomas and share a multitude of features definable
Basal breast cancer requires a set of diagnostic markers, and has been defined differently in different studies. Whereas Nielsen et al defined basal breast cancer as negative for ER and HER2 as well as positive for basal cytokeratins, EGFR and/or c-kit, other groups used the combination of negativity for ER and HER2 with positivity for Ck 5, P-cadherin and p63, or positivity for vimentin, EGFR and Ck 5/6. It may be that, table 1 shows, these different technical approaches in combination with widely differing patient cohorts might explain aspects of the widely differing experimental results for the mentioned markers. Rakha et al showed, in a multivariate analysis, a relationship between the expression of basal cytokeratins and poor prognosis. Interestingly, the addition of other markers as described above did not increase the detection rate of cases with differing outcome compared with the use of basal cytokeratins alone.

**Figure 2** Associations of immunohistochemical expression patterns in physiological cellular subgroups within the normal breast and distinct subgroups of invasive breast cancer.

Basal breast cancer requires a set of diagnostic markers, and has been defined differently in different studies. Whereas Nielsen et al defined basal breast cancer as negative for ER and HER2 as well as positive for basal cytokeratins, EGFR and/or c-kit, other groups used the combination of negativity for ER and HER2 with positivity for Ck 5, P-cadherin and p63, or positivity for vimentin, EGFR and Ck 5/6. It may be that, table 1 shows, these different technical approaches in combination with widely differing patient cohorts might explain aspects of the widely differing experimental results for the mentioned markers. Rakha et al showed, in a multivariate analysis, a relationship between the expression of basal cytokeratins and poor prognosis. Interestingly, the addition of other markers as described above did not increase the detection rate of cases with differing outcome compared with the use of basal cytokeratins alone.

**Basal Cytokeratins, Physiological Breast Progenitor Cells and Genetic Alterations**

In recent years, basal cytokeratins have been the focus of breast cancer research from another perspective. In 1988, Jarrasch et al discussed the possible existence of Ck 5 positive cells as potential progenitor cells in the female breast that could give rise to the glandular and myoepithelial cell lineage. However, the data from that time were indirect and derived from simple immunohistochemical experiments. Almost a decade later, studies showed that cells with the potential to differentiate into luminal and myoepithelial cells—one requirement for the definition of progenitor or stem cells—express basal cytokeratins. Most of these studies were based on mouse or immortalised human mammary cell models. Recently, it has also been shown that putative mammary stem cells in the mouse typically show the expression of high molecular cytokeratins and weak or almost absent expression of Ck 18 as well as lack of expression for ER. Furthermore, these cells were EGFR-positive and HER2-negative. Interestingly, identical results concerning cytokeratin expression patterns were obtained by Dontu et al, who also noted a high proliferation potential of Ck 5+, Ck 8/18+, ER− cells within primary mammospheres. The differentiation of these cells into differentiated glandular cells was paralleled by a complete phenotypic change of these cells into Ck 5−, Ck 8/18+, ER−
Moreover, in most of these experiments these cells seemed to depend on an intact EGF/EGFR system. Further characterisations also showed positivity for vimentin and $\beta$-cadherin and reduced expression of $p21$ and $p27$ in these progenitor/stem cells.

The similarities of these protein expression patterns in physiological progenitor cells of the breast and the basal breast cancer subgroup are striking. As fig 2 shows, the present state of knowledge defines a spectrum of physiological cellular compartments within the female breast. The extreme ends of this spectrum (progenitor/stem cells and luminal cells) can easily be defined by a variety of proteins. In parallel, the distribution of these protein expression patterns is similar in the spectrum of invasive breast cancer.

This spectrum is not only defined by different protein expression patterns but also by the presence of distinct genetic alterations in different cellular compartments. Figures 3 A and B illustrate this point. In figure 3 A the EGFR amplifications are characteristic for invasive breast cancer with expression of Ck 5, whereas 16q-losses are rare in breast cancer types with expression of Ck 5. In contrast, $p53$-mutations and defects/losses of BRCA1 are almost exclusively seen in high grade carcinomas that show Ck 5 expression. HER2-amplifications are not seen in carcinomas with $\geq 50\%$ Ck 5 positive cells. (B) "Stem cell hypothesis". Evidence is growing that physiological, organ-specific stem cells are the major target in the pathogenesis of different cancer entities, including breast cancer. As stem cells by definition are slowly proliferating, early progenitor cells or so called transit amplifying cells 1st order with a high proliferation capacity are the major target cells. A multitude of genetic alterations can take place in these cells. The respective genetic alteration would, in consequence, start distinct cellular expression programmes, including the change of cytokeratin expression pattern, characterising distinct subgroups of invasive breast cancer.
(cyto)genetic alteration patterns. Whereas basal carcinomas have been defined by the presence of \( \text{EGFR} \)-amplifications, \( \text{BRCA}^1 \) and/or \( p53 \)-mutations, chromosomal 16q-losses were rare in these tumours. This was in contrast to ER-positive tumours, with 16q-losses as a hallmark.

These associations suggest at least two explanations. Against the background of the above mentioned subsets of progenitor cells, it may be that distinct cytogenetic alterations can show their effect only in distinct precursor cells. 16q-losses in this regard may predominantly occur in more mature progenitor cells, whereas “basal” changes may arise in early, more stem cell-like, progenitor cells. In other words, the tumour would reflect or maintain cellular properties of its progenitor or more differentiated cell of origin (fig 3A). An alternative explanation, taking into account the ongoing discussion about organ-specific stem cells as target cells, would instead point towards direct changes in the stem cell itself and, as a result, protein expression patterns in the tumour that reflect specific genetic alterations in the progression from stem cell to tumour cell. This would then suggest that, for example, \( \text{EGFR} \)-amplifications or the loss of \( \text{BRCA}^1 \) would start a “basal” expression programme, whereas in contrast 16q-losses would lead to the expression of luminal markers as shown in fig 3B and recently shown by microarray analysis.

Alternative theories speculate about a myoepithelial origin of these tumours, as many “basal” markers are also expressed by myoepithelial cells in vitro or in vivo. Another proposed origin of basal tumours has been deduced from morphological and immunohistochemical disdifferentiation, defined as “epithelial-mesenchymal transition” (EMT), with the expression of vimentin as a major morphological hallmark of EMT. Even though all above listed theories have their shortcomings, the EMT and myoepithelial theories are flawed by some pure histopathological observations. Ductal carcinoma in situ (DCIS) is accepted as the ultimate precursor of invasive breast cancer; it is seen in the majority of basal breast cancer cases. As long as no evidence exists that myoepithelial cells are the cells of origin of DCIS, the presence of DCIS, associated with the immunohistochemical staining pattern, including the expression of vimentin, suggests a glandular, non-myoepithelial origin of these tumours (excluding the myoepithelial carcinomas). Moreover, an association of benign, myoepithelial tumours with the most common basal breast cancers has not been described to date. More recently a subgroup of DCIS with expression of HMW-cytokeratins has been described, pointing to the existence of a real precursor lesion for basal breast cancer. EMT as a general mechanism is also disputable, based on the fact that vimentin, considered the most reliable marker of EMT, is expressed in similar degrees in DCIS and the adjacent invasive carcinomas—pointing to a general characteristic of the respective, individual tumour.

However, most of these results are associative and a definite proof would require experiments targeting different subgroups of physiological progenitor cells as has been previously done. A mouse model of the basal breast cancer subgroup has been recently described. However, most studies focus on the basal, progenitor compartment as cellular targets in oncogenesis. Experiments targeting Ck 18+/Ck 5/14—cells have yet to be performed.

**CLINICAL CONSEQUENCES**

What are the current clinical consequences of our knowledge about basal breast cancers? As mentioned above, the initial overall prognostic significance of basal carcinoma is now discussed intensively. However, from a practical point of view it seems that the existing guidelines for the workup of a breast cancer specimen with the determination of the ER, PR and HER2-expression status provides a suboptimal definition of basal tumours. The recently introduced term “triple-negative breast cancer” (ER negative, PR negative, HER2 negative) roughly describes the majority of basal breast cancers; however, the overlap between basal and triple negative breast cancer is imperfect, with basal cancers comprising perhaps 60–90% of triple negative breast cancers, depending on whether they are defined immunohistochemically or by expression profiling; also, basal tumours may themselves be molecularly heterogeneous. The terms, although occasionally used interchangeably, should be considered distinct for research and clinical purposes as not all basal cancers are triple negative. Therefore, results from future studies focused on triple negative breast cancers will not provide a clear characterisation for all basal breast cancers. However, keeping these discrepancies in mind, preliminary studies on triple negative breast cancers showed that their locoregional relapse rate is identical to that of other molecular subgroups, suggesting that conservative surgical treatment options are still available to patients with basal tumours. However, one has to be aware that the basal subgroup is more frequently detected in interval breast cancers compared to screening-detected breast cancers. In vitro studies suggested different mechanisms of chemo-resistance in basal and luminal breast carcinomas, even though the relation between an increased sensitivity and a poorer outcome cannot yet be explained. It has also been hypothesised that platinum salts based therapy regimens might be of advantage in the treatment of basal breast cancer as they seem to interfere with the \( \text{BRCA}^1 \) pathway. However, currently no specific systemic therapy is recommended for the treatment of triple negative/basal breast cancers and only limited data is available to support the selection of appropriate treatment. Data about chemo-sensitivity in \( p53 \)-mutated breast cancers is controversial. In a prospective study, a complete pathological response to primary chemotherapy was reported in 45% of basal breast cancers and 45% of HER2-positive breast cancers, but only 6% of luminal breast cancers. Current recommendations regarding adjuvant chemotherapy for this tumour type are in flux and an interdisciplinary approach for individualisation of patient care remains the best approach.
REFERENCES


