Changes in keratin expression during metastatic progression of breast cancer: impact on the detection of circulating tumor cells

Simon A. Joosse1*, Juliane Hannemann1*, Julia Spötter1, Andreas Bauche1, Antje Andreas1, Volkmar Müller2, and Klaus Pantel1

1Department of Tumor Biology, Center of Experimental Medicine, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany
2Department of Gynecology, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany
*Contributed equally.

Running title: Keratin expression in circulating tumor cells
Key words: "keratin", "circulating tumor cells", "breast cancer"

Corresponding author: Klaus Pantel
Mailing address: Department of Tumor Biology, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany
E-mail: pantel@uke.de
Translational relevance
Circulating tumor cells (CTCs) found in blood might function as an early diagnostic marker for breast cancer metastasis and as a marker to monitor therapy efficacy. In this study we show that keratins, the markers currently used for the identification of CTCs in breast cancer patients, can be downregulated during metastatic progression. The keratin antibodies used to date in CTC diagnostics (e.g., C11 in the FDA-approved CellSearch System) miss CTCs expressing certain keratin proteins (e.g., K16) associated with metastatic progression in breast cancer. The results presented here point to the use of broad spectrum antibody cocktails including these biological relevant keratins for CTC detection.
Abstract

Purpose: Circulating tumor cells (CTCs) might function as early markers for breast cancer metastasis or monitoring therapy efficacy. Enrichment and identification of CTCs are based on epithelial markers that might be modulated during epithelial-mesenchymal-transition. Little is known about the expression of keratins in CTCs and whether all CTCs can be detected with antibodies directed against a limited panel of keratins.

Experimental Design: Protein expression of keratin 2, 4-10, 13-16, 18 and 19 were assessed by a cocktail of antibodies (C11, AE1, AE3, and K7) and keratin antibodies C11 and A45-B/B3 alone in 11 breast cancer cell lines and 50 primary breast carcinomas and their lymph node metastases. Furthermore, CTCs were assessed in blood of 70 metastatic breast cancer patients.

Results: Claudin-low cell lines did not show expression of normal breast epithelial keratins but were positive for K14 and K16, detected by the cocktail only. Primary breast carcinomas showed changes in keratin expression during metastatic progression to the lymph nodes. In 35/70 patients CTCs were identified of which 83%, 40%, and 57% were identified by the cocktail, C11 and A45-B/B3, respectively. Identification of CTCs by the cocktail was associated with shorter survival (p<0.01). In silico analyses revealed association between KRT16 expression and shorter relapse free survival in metastatic breast cancer.

Conclusion: Breast cancer cells show a complex pattern of keratin expression with potential biological relevance. Individual keratin antibodies recognizing only a limited set of keratins inherit the risk to miss biologically relevant CTCs in cancer patients and antibody cocktails including these keratins are therefore recommended.
Introduction
Tumor cells that are able to invade the body and enter the blood circulation may finally grow out to metastases in distant organs (1). Because cancer metastasis is the main cause of cancer-related death, early detection of (micro) metastases is of vital importance (2). Unfortunately, current imaging techniques are unable to detect micrometastases; therefore, different strategies have been explored to facilitate the detection of micrometastatic disease in early stage by identification of tumor cells in bone marrow or blood using immunocytochemistry (ICC) (3). It has been shown that the identification of disseminated tumor cells (DTCs) in the bone marrow of breast cancer patients covers prognostic information (4). Additionally, identification of DTCs could help in decision making on the need of therapy and to predict the outcome of treatment (5). Alternatively to bone marrow samples, which are difficult to obtain in a routine fashion, blood has been shown to be a valuable source for the identification of dissociated tumor cells (6-8). Detection of these circulating tumor cells (CTCs) might be used to monitor and predict therapy response in cancer treatment because repeated samples can be taken easily (9-11).

To be able to detect tumor cells in bone marrow or blood samples by monoclonal antibodies, different techniques and different markers are currently being used. The most commonly used marker to identify tumor cells from carcinomas are keratins, which have become the standard detection marker for DTCs and CTCs (1, 12-14). Twenty different keratins are described to be expressed in the cytoskeleton of epithelial cells. This family of intermediate filaments can be separated into neutral/basic (type I, K9-20) and acidic (type II, K1-8) keratins. Because hematopoietic cells rarely express keratin proteins detectable by immunostaining (15, 16), a specific detection can be achieved for cells from epithelial tumors such as breast, prostate, colon, or lung carcinomas from blood cells. However, it is thought that tumor cells undergo epithelial-mesenchymal transition (EMT) to be able to migrate and invade the body; during this process the cytoskeleton is rearranged and epithelial markers like E-cadherin, claudins and keratins are downregulated (17-19). The question, whether current detection methods using keratin antibodies are able to detect all CTCs is still an open issue to be answered. If
indeed not all CTCs are detected by the common detection methods, false-negative results will be the consequence and would hamper the clinical implementation of CTCs as diagnostic marker.

Because it has been reported that only a subset of keratins are downregulated during EMT and others might be stably expressed or even upregulated, we hypothesized that an increased detection of tumor cells could be achieved by using more antibodies against a broader range of keratins than the current standards. In the study presented here, we have therefore investigated the keratin expression pattern of eleven breast cancer cell-lines and 50 primary tumor specimens with their corresponding metastases. Moreover, we have studied the detection rate of CTCs in blood samples of 70 metastatic breast cancer patients comparing the sensitivity of different keratin antibodies and analyzed the clinical relevance of our findings by follow-up analyses of the outcome of cancer patients.
**Materials and Methods**

**Antibodies**
Protein expression of keratin 2, 4-10, 13-16, 18 and 19 was assessed by using a combination of pan-keratin antibodies C11 (Abcam cat.no.ab7753), AE1 (Millipore cat.no.MAB1612), AE3 (Millipore cat.no.MAB1611), and keratin 7 antibody (Abcam cat.no.ab9021) (Table 1), further referred to as the keratin antibody cocktail. Furthermore, we have compared the keratin antibody cocktail with two of the most commonly used keratin antibodies for the detection of epithelial tumor cells: C11 and A45-B/B3 (Micromet cat.no.R002A) (13, 14, 20). The keratins detected by these antibodies are described in Table 1. Specificity of the antibodies has been confirmed by the manufacturers and in previous studies (21, 22).

**Cell-lines**
The following breast cancer cell-lines were acquired from ATCC and cultured under the prescribed conditions: BT-20, BT-474, BT-549, GI-101, Hs 578T, MCF-7, MDA-MB-231, MDA-MB-435S, MDA-MB-468, SK-BR-3, and T-47D. Cells were grown in a 75 cm$^2$ flask until confluency was reached. Cells were harvested by using trypsin/EDTA (Gibco cat.no.R001100), washed, and resuspended in 1xPBS for either spiking experiments or cytopsin preparation for direct staining.

**Patients and blood sampling**
Seventy metastatic breast cancer patients were included into the study (patient details are described in Supplementary data 1). Patients were treated for metastatic breast cancer at the University Breast Center and received chemotherapy, endocrine treatment, or treatment with Trastuzumab alone or in combination with chemotherapy based on national and international guidelines. After a median of 3.6 years after diagnosis of metastatic disease, patients were included into our study from 2009 to 2011; there were no exclusion criteria. All patients and healthy volunteers gave informed consent and the examination of blood samples was performed blindly and was approved by the local ethics review board. Ten ml blood was drawn in an EDTA tube from all 70 metastatic breast cancer patients during their anti-cancer treatment at the hospital and from 10 healthy controls. Blood samples were stored at room
temperature and processed within 24 hours after collection as follows: full blood was
transferred to a 50 ml tube containing 40 ml HBSS (Biochrom, cat.no.L2045) and centrifuged
at 400 G for 10 minutes at 4°C. Supernatant was removed by pipetting and pellet was
resuspended in 30 ml 1xPBS. Cell suspension was added to 20 ml Ficoll (GE Healthcare,
cat.no.17-1440-03). Tube was spun at 400 G for 30 minutes at 4°C without acceleration and
deceleration. The interface and supernatant, containing the mononuclear cells (i.e., leukocytes
and tumor cells), was transferred to a new 50 ml tube. The tube was filled with 1xPBS and
centrifuged at 400 G for 10 minutes at 4°C. Supernatant was discarded and cell pellet was
resuspended in 1 ml 1xH-Lysisbuffer (R&D Systems, cat.no.WL1000) and incubated for 1 to 3
minutes with gentle shaking at room temperature. Thirty ml PBS was added and sample was
centrifuged again at 400 G for 10 minutes at 4°C. Supernatant was discarded and pellet was
resuspended in 5 to 10 ml 1x PBS for cytospin preparation.

Cytospin preparation
For human blood and cell-line cells, the following procedure was applied to prepare cytospins.
The amount of viable cells was calculated by using a Neubauer cell counting chamber. Glass
slides were placed into cytospin funnels of 220 mm² or 120 mm² in diameter and 5x10⁵ or
5x10⁴ cells were transferred for patient’s blood or cell-line experiments, respectively. Slides
were spun at 170 G for 3 minutes at room temperature and cytospins were allowed to air-dry
before staining was performed.

Cytospin staining procedure
For staining of cytospins, primary antibodies A45-B/B3 (1:100), AE1 (1:100), AE3 (1:100),
C11 (1:750), and CK7 (1:1000), or IgG1 isotype control clone MOPC21 (1:500; Sigma
cat.no.M9269) were used. Antibodies were diluted in 10% AB-Serum/PBS (Biotest
cat.no.805135). The cell section was bordered using a Dako Pen (Dako cat.no.S2002). The
following solutions were added per slide, followed by incubation, removal of solution, and 3x
washing with PBS for 3 minutes: 1) 4% PFA (pH 7.4), 10 minutes; 2) 150 µl 10% AB-
serum/PBS, 20 minutes, no washing; 3) 125 µl primary antibody or MOPC21, 45 minutes; 4)
125 µl secondary antibody (1:20; Dako cat.no.Z0259), 30 minutes; 5) 125 µl APAAP-complex
(1:100; Dako cat.no.D0651), 30 minutes; 6) 150 µl [solution 2 (415 µl NaNO₂, 166 µl 5% New Fuchsine, 3 minutes incubation in darkness) was added to solution 1 (78.8 ml H₂O, 4.2 ml 1M Tris pH9.5, 200 µl Levamisole), and mixed by pipetting. Solution 3 (42 mg Naphthol AS-BI phosphate in 500 µl dimethylformamide) was added and mixed. The mixture was filtered with 0.2 µm filter.), 20 minutes. Nuclei were counterstained with Mayer’s hemalaun blue (1:5, Merck cat.no.1.09249) for 15-20 seconds. Slides were washed in H₂O shortly, dipped 3x in 0.1% acetic acid and 3x in H₂O, washed once in 0.1% NaHCO₃ for 30 seconds and finally washed 3x in H₂O for 3 minutes.

Experimental setup
Four stainings were compared: 1) C11; 2) A45-B/B3; 3) Keratin antibody cocktail [AE1, AE3, C11, CK7]; and 4) IgG1 isotype control (see Table 1 for antibody properties). For each cell-line, 12 slides were prepared and each staining was performed three times. Of each slide, nine fields in the center of the cytospin were photographed; large cell clusters were avoided. Photographs were taken under 200 x magnifications. Three individuals (SJ, JS, and SD) each blindly reviewed three of the nine pictures per slide and in total 36 photos per cell-line. Positive and negative (as described above) cells were counted and the mean percentage of positive cells and standard deviation per cell line were calculated.

Of the blood cytospins, a total of 10 slides were stained per blood sample. Nine slides were analyzed with A45-B/B3, C11 and keratin antibody cocktail (each three times as before), and one with IgG1 isotype control MOPC21. Slides were completely reviewed using the ChromaVision ACIS II automated microscope system. All positive hits were reviewed manually.

Western blot
To confirm immunocytochemical staining, western blot was performed as described elsewhere (23). A 5% SDS polyacrylamide gel for protein collection and a 10% SDS polyacrylamide gel for protein migration were used to separate 5 µg proteins as determined with calibration series using Pierce BCA reagents (Thermo Scientific, cat.no. 23225). Proteins were transferred onto transfer membrane (Immobilon-psq, ISEQ00010), using transfer buffer
containing 50 mM Tris, 10 mM Calycin, 0.4% SDS, and 20% Methanol. Next, blocking was performed in 5% milk powder/TBST for 1 hour, followed by 3 x 5 minutes washing with TBST. Transfer membrane was incubated with primary antibodies over night as follows: A45-B/B3 (1:1000) in 5% milk powder/TBST, AE1 (1:150000) in 5% BSA/TBST, AE3 (1:150000) in 5% BSA/TBST, C11 (1:1000) in 5% BSA/TBST, and CK7 (1:1000) in 5% BSA/TBST. Next, 3 x 5 minutes washing was performed and 1:1500 secondary mouse antibody (Polyclonal Goat Anti Mouse Immunoglobulins/HRP Dako P0447) in 5% milk powder/TBST was incubated for 1 hour. After 3 x 5 minutes washing with TBST, western blots were developed using 1:50 ECL Plus (Amersham).

**Tissue microarrays**

A commercially available tissue microarray (TMA) of 100 formalin-fixed, paraffin-embedded cores (1 mm diameter) was used in this study (US Biomax Inc, cat.no.BR1005). The tissue microarray contained 100 tissue cores from 50 breast cancer patients, unrelated to the earlier mentioned patients of this study. From each patient one core originates from the primary breast carcinoma and one core originates from the corresponding carcinoma metastases in lymph nodes. Preceding the staining of the TMAs, slides were baked for 2h at 55°C and deparaffinized in Xylene for 15 minutes two times. Rehydration was done in Ethanol-series (2x 100%, 2x 96%, 2x 80%, 2x 70%) 2 minutes each and finally 1xPBS. Antigen retrieval was performed in boiling Sodium-citrate-buffer (10mM Sodium-citrate pH 6.0, 0.05% Tween 20) for 15 minutes. TMAs were washed in TBS twice and blocking was performed with ready-to-use Proteinblock (DakoCytomation, cat.no.X0909) for 20 minutes at room temperature. Incubation of primary antibodies (A45-B/B3 (1:300), AE1 (1:300), AE3 (1:300), C11 (1:2250), and CK7 (1:3000)) was performed over night at 4°C. Following staining steps were performed similarly as described for the cytospins. In total, six TMAs were stained with: the antibody cocktail, C11, and A45-B/B3, respectively, all in duplicate. Each slide was read independently by SJ and JH, blindly. Discrepancies were re-evaluated and a consensus was found. Scoring was performed as following: 0 (negative), no staining; 1+ (positive), a faint/barely perceptible
staining; 2+ (positive), a weak to moderate staining; 3+ (positive), a strong staining was observed in more than 1% of the tumor cells.

**Statistical analyses**
Statistical significance between the fractions of positive cell line cells was calculated by ANOVA using each reviewed field as a measurement per staining (total fields per cell line per staining was 27) and reviewer as random variable. Tissue microarray results were categorized into positive and negative, and analyzed using the McNemar test. Survival analysis of patients tested for CTCs was performed using the log-rank test after dividing the patient cohorts into CTC positive and CTC negative groups; hazard ratios and 95% confidence interval (CI) were calculated using Cox proportional hazards model.

**in silico analysis**
In order to examine a potential association of keratin gene expression with metastatic relapse in breast cancer, we examined the available gene expression data of two cohorts with a high rate of relapsed patients and information in metastasis-free survival (24, 25). Data were downloaded from GEO (GSE12276 and GSE2034) and all cases with relapse were used in our analysis. In order to retain the original series of the publications, there was an overlap of 12 cases in the two studies. For each individual set of keratin gene expression levels, data were sorted in decreasing order along with the patients' survival data. The upper 35% of the data was labeled as 'high keratin gene expression', whereas the lower 35% of the data was labeled as 'low keratin gene expression'; the remaining data (middle 30%) was discarded. The survival data of the low and high keratin gene expression groups were compared with each other using the log-rank test.
**Results**

Keratin protein expression was assessed in breast cancer cell lines, primary breast tumors, lymph node metastases, and circulating tumor cells (CTCs) by keratin antibodies. The used antibodies were C11 (detects K4-6, 8, 10, 13, 18), A45-B/B3 (detects K8, 18, 19), and the keratin antibody cocktail (detects K2, 4-10, 13-16, 18, 19), described in table 1. In addition, keratin mRNA expression was determined by *in silico* analyses in metastatic breast cancer.

**Keratin protein expression in breast cancer cell lines**

Cytospins of eleven breast cancer cell lines were stained with C11, A45-B/B3, keratin antibody cocktail, and IgG1 isotype control (clone MOPC21) and reviewed under a 200x magnification. In total, 108 fields per cell line were reviewed and keratin positive and negative cells were counted manually; a single field contained 80 cells on average (standard deviation $s = 25$). In three out of eleven tested cell lines, a significant higher proportion of the cells was identified as positive with the keratin antibody cocktails as compared with C11 or A45-B/B3 alone ($p<0.01$, ANOVA with Bonferroni-Holm correction) (Table 2 and Supplementary data 2). The biggest effect was seen in cell line MDA-MB-435S, in which 40% of the cells were stained using the keratin cocktail, whereas no cells were stained by C11 or A45-B/B3 alone. IgG1 isotype control was negative in all cases, demonstrating the specificity of the stainings.

Next, western blot was applied to investigate which specific keratins are expressed in cell lines BT-549, Hs 578T, and MDA-MB-435S. Protein extract from MDA-MB-231 was used as positive control since this cell line had shown to be positive in all keratin subset mixes before (Table 2). Figure 1 shows the results of five western blots, performed with each of the individual pan-keratin antibodies and K7 antibody. As expected, MDA-MB-231 was positive for all applied antibodies. Cell line MDA-MB-435S, which was positive by immunocytochemistry (ICC) in our keratin cocktail, showed clear signals using AE1. The rare staining of cell lines BT-549 and Hs 578T using the keratin cocktail could be confirmed by the presence of a weak signal by AE1. These results demonstrate that the AE1 pan-keratin antibody is important for the detection of keratins in breast cancer cells.
**Keratin gene expression in breast cancer cell lines (in silico analysis)**

According to a study of Neve and colleagues, the three cell lines BT-549, Hs 578T, and MDA-MB-435S, in which no positive cells were detected using C11 or A45-B/B3, are all of the basal-like B (claudin-low) breast cancer subtype. MDA-MD-231 has also been classified as basal-like B but was positive in all stainings. All other cell lines from our study were classified as basal-like A or luminal subtype (26). This motivated us to analyze the keratin gene expression levels of the breast cancer cell lines studied by Neve et al. to investigate whether there are differences between keratin expression levels of claudin-low, basal-like, and luminal breast cancer cell lines. Complete data of 52 breast cancer cell lines was available, of which 25 were of the luminal subtype, 12 basal-like A, and 14 basal-like B (claudin-low).

Supplementary data 3 summarizes the average gene expression levels of the different keratins per breast cancer subtype. As shown, KRT8 and KRT19 are significantly less expressed in basal-like B cell lines compared to basal-like A and luminal cell lines (p<0.0005, Bonferonni). Additionally, KRT18 is significantly lower expressed in all basal-like cell lines compared to luminal cell lines. KRT5, 6B, 14, and 16 are significantly higher expressed in basal-like A cell lines compared to luminal or basal-like B cell lines. Thus, the expression patterns of certain keratins are correlated to breast cancer subtypes.

**Keratin protein expression in primary breast carcinomas and lymph node metastases**

A tissue microarray containing cores of 50 primary breast carcinomas and their corresponding lymph node metastases was investigated to assess the expression of keratins in tumor tissue and metastatic tumor cells. Concordance in staining intensity and positive cell percentage between the independent evaluations of the individual slides was approximately 90% per slide. Using C11 and A45-B/B3, 14 (28%) and 27 cases (54%) of the primary tumors were detected to be positive for keratin, respectively. When using the keratin antibody cocktail, all primary tumors (100%) were found to be positive (defined as a positive staining in >1% of the cells). For the corresponding metastasis to lymph nodes, 19 (38%) and 28 (56%) were positively stained using C11 and A45-B/B3, respectively; however, keratin status of the primary tumor differed frequently from the corresponding lymph node metastasis: 12/50
(24%) of the C11 stained cases and 15/50 (30%) of the A45-B/B3 stained cases showed
deviation (Figure 2). Forty-nine out of fifty lymph node metastasis (98%) were stained
positive using our combination of keratin antibodies (Figure 2). These results suggest that the
keratins detected by A45-B/B3 are expressed more often in breast cancer and breast cancer
metastases in general as the keratins detected by C11 or that A45-B/B3 is more sensitive than
C11. Moreover, by adding additional (pan-)keratin antibodies, we could stain significantly
more tumors and lymph node metastases than with C11 or A45-B/B3 alone (p<0.0001,
McNemar test). Because keratin status in many of the lymph node metastases did not
correspond to that of the same patient’s primary tumor detected by C11 or A45-B/B3, these
results indicate that keratin expression changes during metastasis. This emphasizes the need
of a broader spectrum of keratin antibodies to detect metastasized cancer cells.

Keratin gene expression in metastasizing breast carcinomas (*in silico* analysis)
In order to investigate the relationship between the expression of keratins in metastatic breast
cancer and patient’s relapse free survival, we performed an *in silico* analysis on two previously
published studies (24, 25). These two studies assessed mRNA gene expression by the HG-
U133A GeneChip from Affymetrix in metastatic and non-metastatic primary breast tumors.
We extracted the expression levels of all keratins present in these studies from the cases with
metastatic relapse only. Probes for all keratins investigated in our experimental setup were
present in these data except for *KRT19*. The study of Bos *et al.* contained primary tumors of
204 patients who later developed metastatic disease and the cohort of Wang *et al.* 107
metastatic cases.
After collecting the required data, survival analysis was performed on those cases expressing
keratins high or low (see materials and methods), for each keratin separately. In both studies
of Bos *et al.* and Wang *et al.*, only expression of *KRT16* showed a significant difference in
survival (Figure 3). For the cohort from the study of Bos *et al.* log-rank test resulted in
p=0.001 and cox regression in a hazard ratio of 1.69 [95% CI 1.20-2.37]. The median relapse
free survival time after primary diagnosis related to low and high expression of *KRT16* was 23
and 14.8 months, respectively. For the cohort from the study of Wang *et al.* log-rank test
resulted in p=0.010 and cox regression in a hazard ratio of 1.79 [95% CI 1.12-2.84]. The median relapse free survival related to low and high expression of KRT16 was 28 and 17.5 months. These results suggest that a high expression of KRT16 is associated with poor survival, whereas a low expression of KRT16 is correlated with a better survival in metastasizing breast cancer.

**Keratin protein expression in circulating tumor cells of breast cancer patients**

Blood of an independent set of seventy metastatic breast cancer patients from our hospital was investigated for the presence of CTCs by capturing all mono-nuclear cells from 10 ml blood using Ficoll density gradient separation and subsequent staining with A45-B/B3, C11 and the keratin antibody cocktail. In total, CTCs were identified in 35 samples (50%) as defined by the detection of at least one positive cell, 57% of the positive patients was identified by A45-B/B2, 40% by C11, and 83% by the keratin antibody cocktail. For each individual patient, the fraction of positive cells detected with A45-B/B3, C11 and keratin antibody cocktail was calculated. Next, the relative fraction of positive cells for each staining over the whole cohort was calculated. On average, A45-B/B3 and C11 detected 26% and 14% of all the positive cells, respectively; the keratin antibody cocktail detected 61% (Supplementary data 1). Figure 4A depicts the distribution of the number of cells detected per antibody in the metastatic breast cancer patients. These results suggest that more CTCs can be detected by using our keratin antibody cocktail as compared to A45-B/B3 or C11 alone (p<0.01, two sided paired t-test).

In order to assess the rate of possible false positivity, blood of ten apparently healthy volunteers was investigated. Of each person’s blood, twenty cytospins containing each 500,000 cells were prepared and stained with keratin antibody cocktail. In all 200 slides, we have found no keratin positive cells. This result confirms that the specificity is as high as for C11 and A45-B/B3 that have already been shown in previous studies (13, 20).

**Survival analysis of CTC positive metastatic breast cancer patients**

The seventy metastatic breast cancer patients were divided by their CTC status (positive or negative) detected by each of the tested keratin antibodies individually. Next, survival analyses starting from the day of blood sampling were performed (median follow-up of 2.23 years) and
statistical significance was calculated using the log-rank test; at time of analysis, 27% (19/70) of the women had died. No difference in survival was demonstrated between the CTC positive and negative patients as defined by A45-B/B3 (p=0.8829, hazard ratio: 1.36 [95% CI 0.53-3.46], Figure 4B) or C11 (p=0.8715, hazard ratio: 1.52 [95% CI 0.55-4.20], Figure 4C). In contrast, a significant difference could be observed in the survival estimates of CTC negative and positive patients as defined by the keratin antibody cocktail (p=0.0035, hazard ratio: 4.16 [1.49-11.59], Figure 4D). Median survival time of CTC positive patients was 1.78 years, whereas for the CTC negative patients the median survival was 2.26 years. The results of this exploratory analysis indicate that detection of CTCs by the antibody cocktail is of prognostic relevance in metastatic breast cancer patients.
Discussion

Identification of CTCs in the bloodstream of cancer patients can serve as a marker for early detection of micrometastasis, but it may also be used to monitor therapy response in early and advance stage patients as a "liquid biopsy" which may allow obtaining important information on the molecular make-up of occult and/or overt metastases, respectively. The detection of CTCs has the potential to become increasingly important in the clinical management of cancer patients. Currently, the most widely used markers for the identification of CTCs are keratins; however, little is known about the changes in the keratin expression pattern in tumor cells during metastatic spread and progression. It has been described that cells from metastasizing breast cancer may undergo epithelial-mesenchymal-transition (EMT) and that they may lose the expression of several keratins during this process (19, 27, 28). Usage of pan-keratin antibodies directed against a limited number of keratins may therefore lead to a failure in detecting CTCs and finally lead to false negative results. The use of a cocktail of antibodies, directed against an even wider range of keratins that also cover the biological behavior of keratins during the metastatic process and by this also keratins differentially regulated during EMT, may substantially improve the CTC detection rate. Indeed, it has been shown recently that an increased detection of DTCs in bone marrow can be achieved by applying a broader set of keratin antibodies (29). In this study, we evaluated keratin expression of primary tumor, lymph node metastasis, and CTCs. To further explore the dynamic changes in keratin expression during metastatic progression, future studies might also include keratin expression in distant metastases. Cells that leave the blood circulation and develop to metastases might have to undergo MET and thereby restoring keratin expression that was deregulated during EMT in CTCs (12, 18).

To test our hypothesis that a combination of available pan-keratin antibodies may improve the detection of CTCs, we first investigated eleven breast cancer cell lines with two of the most commonly used pan-keratin antibodies (C11 and A45-B/B3) and a keratin antibody cocktail, which was chosen based on possible changes in keratin expression as a biological feature during metastasis. Most cell lines expressed enough keratins to be clearly stained with
either C11 or A45-B/B3 alone, except for the three cell lines BT-549, Hs 578T, and MDA-MB-435S. Using the keratin antibody cocktail, we can show that 1-2% of the BT-549 and HS 578T cells express keratin and approximately 40% of the MDA-MB-435S cells. Western blot analyses showed that these three cell lines (BT-549, Hs 578T, and MDA-MB-435S) could only be detected using AE1 and show expression of K16 and K19, although the latter could not be confirmed by A45-B/B3, possibly due to its weak expression. In addition, K5 and K6 could be detected using AE3 but not with C11. This suggests that K5 and K6 might be below the detection level of C11.

According to the dataset of Neve and colleagues, the three cell lines BT-549, Hs 578T, and MDA-MB-435S are classified as breast cancer subtype basal-like B, which is also known as the claudin-low subtype (26). Claudin-low breast tumors originate from the myoepithelial cell layer that normally express K5, 14, and 17 only (30) and are characterized by low expression of epithelial markers such as keratins (31-33). Indeed, in silico analysis confirmed the lack of KRT5/14/17 in these three claudin-low cell lines; however, ICC analyses showed heterogeneity in keratin protein expression among the cells and K14 could be detected by western blot. Because basal-like cell lines originate from luminal progenitor cells, these and luminal cell lines are expected to express K6-8/18/19 (34). The data from Neve and colleagues showed several differences between the luminal and basal-like cell lines in respect to these keratins. KRT6 was apparently downregulated in luminal cell lines, KRT18 was downregulated in basal-like cell lines, and unexpectedly, KRT16 was upregulated in basal-like breast cancer cell lines, which we could also detect by western blot. Surprisingly, cell line MDA-MB-231 has also been classified as basal-like B but showed expression of many keratins, including those that would be expected in luminal breast cancer only.

Next, we investigated a tissue microarray containing 100 core biopsies from primary tumors and corresponding lymph node metastases. We show that keratin expression is limited to only a set of keratins in primary tumors as well as in their lymph node metastases. Additionally, keratin expression seems frequently to be altered between the primary tumor and the corresponding metastasis. However, in most of the investigated core biopsies, keratin expression could be identified by using the keratin antibody cocktail. These results indicate
that keratin expression patterns are due to changes during invasion, or that keratin expression is heterogeneous within the primary tumor mass similar to our results obtained from our cell line experiments. Based on this heterogeneity in keratin expression and the results of others (19, 35, 36) it might be speculated that keratins can also be used in the future to identify the single cell subpopulation that is able to initiate metastasis. It should be noted, however, that this investigation is limited by the fact that the staining procedure and tissue conditions can influence detection rates and that tissues identified as negative in this study might express low amounts of the respective keratins. Although we are confident that the staining procedure is optimal because the results were very reproducible, we could not control for the quality of the tissue. In addition, molecular subtype of the primary tumors is not known, a factor that is also correlated to expression of certain keratins.

As a next step to investigate whether broader keratin detection also leads to improved detection of CTCs, we investigated 70 blood samples of metastatic breast cancer patients for CTCs. To identify these rare cells, we first used a Ficoll density gradient to enrich for mononuclear cells. Although different methods exist to specifically enrich for tumor cells (37), enrichment by using cellular markers like EpCAM might result in loss of cells of interest if they do not express these markers. Screening the samples of metastatic breast cancer patients for CTCs showed that an increased detection of tumor cells could be achieved with the keratin antibody cocktail. It should be noted that given the low number of detectable CTCs, it might be possible that CTCs have not been equally distributed among the investigated slides, but there is no possibility to correct for that. However, due to random staining and sufficient sample size, it seems not likely that there is any bias of the amount of CTCs among the stained slides. By performing the same screening procedure on healthy controls with no findings of positive cells, we demonstrate the high specificity of the keratin antibodies used in this study and are convinced that we did not identify false positives in the cancer samples.

Most importantly, survival analyses on the patient cohort divided by positive and negative CTC status indicated that the CTC-negative patients (as defined by the keratin antibody cocktail) have a better survival. Apparently, more CTCs are detected by the keratin antibody cocktail as compared with C11 or A45-B/B3 alone that might be biologically
relevant, resulting in a better separation between better and poor survival. A potential limitation of this investigation is that the patient cohort is not uniformly treated and that we have no untreated cohort to determine the prognostic impact of the examined factors. Therefore, it cannot be ruled out completely that also the time of blood sampling and the treatment influenced these results. Thus, to test the cocktail’s clinical significance, a larger cohort with uniform treatment and a longer follow-up will be required.

Investigating the possible association between survival and keratin expression in metastasizing breast cancer in silico, we found in published data sets from two earlier studies that patients who showed high expression of KRT16 in their primary tumor, had a shorter relapse free survival when compared to patients with KRT16 low expressing tumors (24, 25). In addition, we have seen that KRT16 upregulation is also a common phenomenon in basal-like breast cancer cell lines. These results could reflect the fact that keratin 16 expression is associated with a more aggressive course of breast cancer and would support our finding that CTCs detected by the antibody cocktail is associated with impaired survival in the metastatic setting. Our findings stress out the importance of the detection of K16, which is not recognized by either C11 or A45-B/B3. Expression of K16 in primary breast cancer and CTCs has to be further investigated for its potential clinical use.

**Conclusion**

In conclusion, our results indicate that breast tumors can alter expression of certain keratins during the process of metastatic development, which would make identification of circulating tumor cells of specifically these tumors by the currently used standard pan-keratin antibodies impossible. In this study, we show that by broaden up the spectrum of keratin detection based on the knowledge of the metastatic process compared to the use of single keratin antibodies, an increased detection of CTCs can be achieved, reducing the number of false-negative results. Our investigation also suggests that modulation of keratins due to EMT occurs frequently on CTCs in breast cancer patients and may by associated with an unfavorable outcome.
Acknowledgements
We would like to thank Stephanie Degwert and Susanne Hoppe for technical assistance.

Grant support
This work has been supported by the German Ministry for Research and Education (BMBF), grant number 01ES0724.
Reference List


Figure legends

Figure 1 - Western blot performed on cell lines BT-549, Hs 578T, MDA-MB-435S, and MDA-MB-231 (positive control). BT-549, Hs 578T, and MDA-MB-435S show signals at the height of K14, K16, and K19 as indicated by AE1.

Figure 2 - Tissue microarray containing core biopsies from 50 primary breast cancers (row A-E) and corresponding lymph node metastases (row F-J), stained by C11, A45-B/B3 and the keratin antibody cocktail. No staining (negative): white; a faint/barely perceptible staining (1+): light gray; a weak to moderate staining (2+): dark gray; a strong staining (3+): black. Percentage of positive cells as indicated or 100% (filled circles).

Figure 3 - Kaplan-Meier estimate of survival function of metastatic breast cancer patients separated on high expression of KRT16 (red - online only) and low expression of KRT16 (blue - online only). Data was taken from two published studies (24, 25). Longer survival correlated with low expression of KRT16 in both studies of Bos et al. (p=0.001, n=70) and Wang et al. (p=0.010, n=37) calculated with Logrank test.

Figure 4 - CTCs in metastatic breast cancer patients (n=70) identified by A45-B/B3, C11, and keratin antibody cocktail.
A) Total number of keratin positive cells. By using the keratin antibody cocktail, more CTCs could be identified.
B-D) Kaplan-Meier estimates of survival function of 70 metastatic breast cancer patients separated on CTC status as detected by A45-B/B3 (hazard ratio: 1.36 [95% CI 0.53-3.46]), C11 (hazard ratio: 1.52 [95% CI 0.55-4.20]) and keratin antibody cocktail (hazard ratio: 4.16 [95% CI 1.49-11.59]). Statistical significance determined by Log-rank test. Between parentheses is the median survival in years of the corresponding patient group. CTC negative groups are indicated by blue lines (online only), CTC positive (≥1 cell) groups are indicated by red lines (online only) and censored patients are indicated by vertical bars ( | ).
Bos et al.

Wang et al.
Table 1 - Keratin antibodies and detection properties. K: Keratin

<table>
<thead>
<tr>
<th>KRT</th>
<th>Size (kDa)</th>
<th>C11</th>
<th>A45-B/B3</th>
<th>AE1</th>
<th>AE3</th>
<th>K7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>65.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>59</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>58</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
</tr>
<tr>
<td>6</td>
<td>56</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
</tr>
<tr>
<td>7</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
</tr>
<tr>
<td>8</td>
<td>52.5</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
</tr>
<tr>
<td>9</td>
<td>64</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>56.5</td>
<td>•</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>53</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>50</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>49</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>48</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>45</td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>40</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2 - Mean percentage of positive cells and standard deviation (s) of eleven breast cancer cell lines. ANOVA was performed between A45-B/B3, C11 and the KRT antibody cocktail with the reviewer as random variable, to determine the significance in difference of p. P-values are corrected for multiple testing (Bonferroni-Holm).

<table>
<thead>
<tr>
<th>Cell line</th>
<th>A45-B/B3 (s)</th>
<th>C11 (s)</th>
<th>KRT cocktail (s)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT-20</td>
<td>89.2 (5.9)</td>
<td>89.3 (6.8)</td>
<td>91.1 (5.6)</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>BT-474</td>
<td>79.3 (10.6)</td>
<td>79.2 (7.0)</td>
<td>79.4 (9.0)</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>BT-549</td>
<td>0</td>
<td>0</td>
<td>1.3 (1.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GI-101</td>
<td>82.3 (8.9)</td>
<td>82.4 (6.1)</td>
<td>84.5 (5.9)</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>Hs578T</td>
<td>0</td>
<td>0</td>
<td>1.4 (1.7)</td>
<td>&lt;0.003</td>
</tr>
<tr>
<td>MCF-7</td>
<td>86.0 (7.6)</td>
<td>84.9 (6.8)</td>
<td>84.6 (7.8)</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>MDA-MB-231</td>
<td>96.7 (5.5)</td>
<td>99.1 (2.0)</td>
<td>98.4 (2.2)</td>
<td>0.38</td>
</tr>
<tr>
<td>MDA-MB-435S</td>
<td>0</td>
<td>0</td>
<td>39.2 (11.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MDA-MB-468</td>
<td>78.2 (10.2)</td>
<td>76.8 (10.1)</td>
<td>79.8 (7.8)</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>SK-BR-3</td>
<td>92.3 (4.6)</td>
<td>91.9 (3.9)</td>
<td>93.4 (4.5)</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>T-47D</td>
<td>94.5 (4.6)</td>
<td>94.0 (3.6)</td>
<td>96.1 (4.6)</td>
<td>&gt;0.99</td>
</tr>
</tbody>
</table>
ANOVA was performed between A45-B/B3, C11 and the KRT antibody cocktail with the reviewer as random variable, to determine the significance in difference of positive cells.