Genomic analysis of single cytokeratin-positive cells from bone marrow reveals early mutational events in breast cancer

Julian A. Schardt,[1,4] Manfred Meyer,[1,4] Claudia H. Hartmann,[1,4] Falk Schubert,[2] Oleg Schmidt-Kittler,[1] Christine Fuhrmann,[1] Bernhard Polzer,[1] Marco Petronio,[1] Roland Eils,[2,3] and Christoph A. Klein[1,*]

[1]Institut für Immunologie, Ludwig-Maximilians Universität München, Goethestr. 31, 80336 München, Germany
[2]Division of Theoretical Bioinformatics, Deutsches Krebsforschungszentrum, Heidelberg, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
[3]Department of Bioinformatics and Functional Genomics, Institute of Pharmacy and Molecular Biotechnology, Universität Heidelberg, Im Neuenheimer Feld 364, 69120 Heidelberg, Germany
[4]These authors contributed equally to this work.
*Correspondence: christoph.klein@med.uni-muenchen.de

Summary

Chromosomal instability in human breast cancer is known to take place before mammary neoplasias display morphological signs of invasion. We describe here the unexpected finding of a tumor cell population with normal karyotypes isolated from bone marrow of breast cancer patients. By analyzing the same single cells for chromosomal aberrations, subchromosomal allelic losses, and gene amplifications, we confirmed their malignant origin and delineated the sequence of genomic events during breast cancer progression. On this trajectory of genomic progression, we identified a subpopulation of patients with very early *HER2* amplification. Because early changes have the highest probability of being shared by genetically unstable tumor cells, the genetic characterization of disseminated tumor cells provides a novel rationale for selecting patients for targeted therapies.

Introduction

Currently, many cancers can only be cured by timely surgery, whereas systemic therapies often fail. The major cause for the poor results of systemic therapies can be traced to the inherent genetic instability (Loeb, 2001) that constantly generates therapy-resistant, variant cells. While advances in molecular cell biology may translate into improved treatment, few therapeutic concepts today address the enormous cellular heterogeneity of systemic cancer. Importantly, it has recently been shown that therapeutic responses require first, the widespread presence of the target (e.g., protein tyrosine kinases such as EGFR or HER2) and second, the cellular dependence on the target, which often results from its genetic change (Arteaga and Baselga, 2004; Lynch et al., 2004; Paez et al., 2004; Sordella et al., 2004). According to these findings, the development of efficient *adjuvant* systemic therapies should be based on detailed knowledge about aberrations that are shared among and essential for tumor cells that disseminate before surgery and later give rise to lethal metastasis.

Representatives of this highly relevant cell population can be detected in bone marrow as single cytokeratin-positive cells long before clinical metastases manifest themselves. In approximately one-third of cancer patients who are clinically free of metastasis, 1–2 cytokeratin-positive cells are detected among 2×10^6 bone marrow cells, while approximately 2% of samples from patients without known carcinoma are cytokeratin positive (Braun et al., 2000b; Klein, 2003). Despite this rareness, cytokeratin-positive tumor cells present in bone marrow represent an important risk factor for reduced disease-free and overall survival for many carcinoma patients without overt metastases (Klein, 2003). In particular, an increasing number of clinical studies on breast cancer confirm the prognostic impact of these cells for overall survival, disease-free survival, and skeletal metastasis (Braun et al., 2000b; Gebauer et al., 2001; Gerber et al., 2001; Wiedswang et al., 2003). In addition, their detection after adjuvant therapy has been interpreted as a surrogate marker for treatment failure (Braun et al., 2000a), and their persistence in bone marrow after surgery and adjuvant therapy has been associated with particularly poor prognosis (Janni et al., 2005). Together, the clinical data suggest that cytokeratin-positive cells in bone marrow of breast cancer patients represent important cellular targets for adjuvant therapies.

SIGNIFICANCE

The hallmark of cancer is genetic instability, which results in a constant clonal divergence of the tumor cell population and gives rise to therapy-resistant, variant cells. Because only the early changes will be shared among all tumor cells, knowledge about initiating events may be key to successful treatment. To date, early aberrations have been identified by analysis of morphologically defined lesions at the primary site. Our finding of disseminated breast cancer cells in bone marrow that show less progressed genomic changes than preinvasive primary lesions provides a mechanism to uncover cancer-initiating and -promoting genetic or epigenetic alterations.

Unfortunately, knowledge about the molecular characteristics of cytokeratin-positive cells has remained elusive. Due to their extreme rareness, double labeling has been the preferred way to characterize these cells for many years. A novel amplification technique (Klein et al., 1999) has then enabled single-cell genomic analysis and uncovered that cytokeratin-positive cells of an individual patient are genetically extremely heterogeneous before metastasis is clinically detected (Klein et al., 2002). Moreover, disseminated breast cancer cells often display chromosomal abnormalities very different from their matched primary tumors (Schmidt-Kittler et al., 2003). The latter suggests that primary tumors and cell lines derived thereof may not be the most optimal surrogate tissues to detect therapy targets present on disseminated cancer cells.

On the other hand, the chromosomal heterogeneity among disseminated cancer cells of an individual patient before and the clonal expansion after manifestation of metastases suggest that chromosomal aberrations may be important for the outgrowth of disseminated cells, while they are not necessary for systemic spread. The outgrowth of cells that have homed to different tissues may occur at different rates and result in different genomic rearrangements. In contrast, all disseminated tumor cells should share the genetic changes that either caused dissemination or initiated cancer. Identification of such changes could then be very helpful for the prevention of metachronous metastases. The search for tumor-initiating changes is usually performed using samples of small tumors or morphologically defined preinvasive stages of neoplastic development. However, it has been found that chromosomal imbalances are among the earliest changes of morphologically defined lesions and are associated with hyperproliferation in breast cancer (Allred et al., 2001; Aubele et al., 2000; Buerger et al., 1999; Waldman et al., 2000). In particular, chromosomal instability was shown to occur during the transition from ductal hyperplasia to in situ carcinoma, i.e., before the cancer becomes morphologically invasive (Chin et al., 2004). We recently found that 57% of cytokeratin-positive cells isolated from the bone marrow of breast cancer patients without clinically evident metastasis do not display chromosomal aberrations, whereas the matched primary tumors of the patients regularly harbored multiple chromosomal alterations (Schmidt-Kittler et al., 2003). Thus, we reasoned that karyotypically normal (as defined by comparative genomic hybridization [CGH; Kallioniemi et al., 1992]) cytokeratin-positive cells in bone marrow may represent genetically very early stages of breast cancer development—if they are indeed tumor cells. Their analysis may provide information about genetic events that occur before breast neoplasias become morphologically detectable and chromosomally unstable.

Results

Detection of allelic losses in single cells

Cytokeratin-positive cells with normal karyotypes in bone marrow could either be nonmalignant cells that are stained for unknown reasons (e.g., false positive normal bone marrow cells [Borgen et al., 1998]) or tumor cells that have genetic changes below the resolution of comparative genomic hybridization (10–20 Mb). To differentiate between the two possibilities, we evaluated whether cytokeratin-positive cells display DNA damage when analyzed with higher resolution. Therefore, we tested whether DNA from single normal cells that had been globally amplified by adaptor-linker PCR (Klein et al., 1999) is suited for loss of heterozygosity (LOH) analysis. Because of the possibility that some DNA fragments fail to be amplified in the initial whole-genome amplification and thus cannot be detected in subsequent sequence-specific PCR, amplification of single-copy sequences in a single-cell genome had to be carefully controlled.

Single control cells were isolated from two groups of donors (Table S1 in the Supplemental Data available with this article online). Twenty-one single blood cells were isolated from four healthy donors after staining with an antibody against mesenchymal vimentin (control group 1) and 52 single cytokeratin-negative cells from bone marrow of patients with malignant epithelial cancer (control group 2). For the cytokeratin-positive cells (see below) and control cells, we performed single-cell CGH to assess the quality of the global amplification and included only those cells that could be successfully hybridized to metaphase chromosomes. All control cells displayed normal CGH profiles. Control group 1 cells served to uncover a potentially negative influence of the precipitated color substrate or the antibody binding on the PCR result of cytokeratin-positive cells. Cells from control group 2 had been incubated with the same substrate solution during immunocytochemistry and had not bound the cytokeratin antibody but were isolated from carcinoma patients. These patients suffered from breast, prostate, or various gastrointestinal cancers in clinical stage M0 and were similar in age to the breast cancer patients from whom the cytokeratin-positive cells were isolated. All 73 cells were analyzed for 29 polymorphic markers (see below), and the loss of PCR bands was assessed. As expected, loss of PCR bands was occasionally observed in cells from perfectly healthy individuals (Figure 1B), indicating either random DNA damage in individual cells or—more likely—random PCR failure. Rarely, we observed loss of both PCR bands of a specific marker, which could indicate homozygous loss of a marker, loss of the remaining allele by PCR failure in a region with true LOH, or complete PCR failure. Due to this uncertainty, we excluded homozygous losses for the assessment of LOH. To our surprise, we found that cells from control group 1 displayed significantly fewer allelic losses than control group 2 (2.9% [95% confidence interval (CI), 0.1%–5.7%] versus 9.8% [95% CI, 6.8%–12.8%]; $p = 0.006$, exact Wilcoxon rank test; Figures 1A–1C), excluding a negative influence of precipitated color or bound antibody on the PCR result. The two groups did not differ with regard to homozygous losses ($p = 0.2$). Since both groups consisted of randomly picked hematopoietic cells with a diploid genome, the major differences lay in the younger age of the donors from group 1 (median age 28 years versus 59 years in control group 2) and the fact that control group 2 consisted of patients with cancer. Comparing each marker separately, we found no significant differences for any marker between control group 1 and control group 2 (Table 1), suggesting random DNA loss throughout the genome. The control experiments demonstrated that LOH analysis with single cells is feasible but requires statistical analysis due to random PCR failure or—possibly—age-dependent DNA loss. For statistical comparison with cytokeratin-positive cells, we therefore used only the samples from the age-matched control group 2.

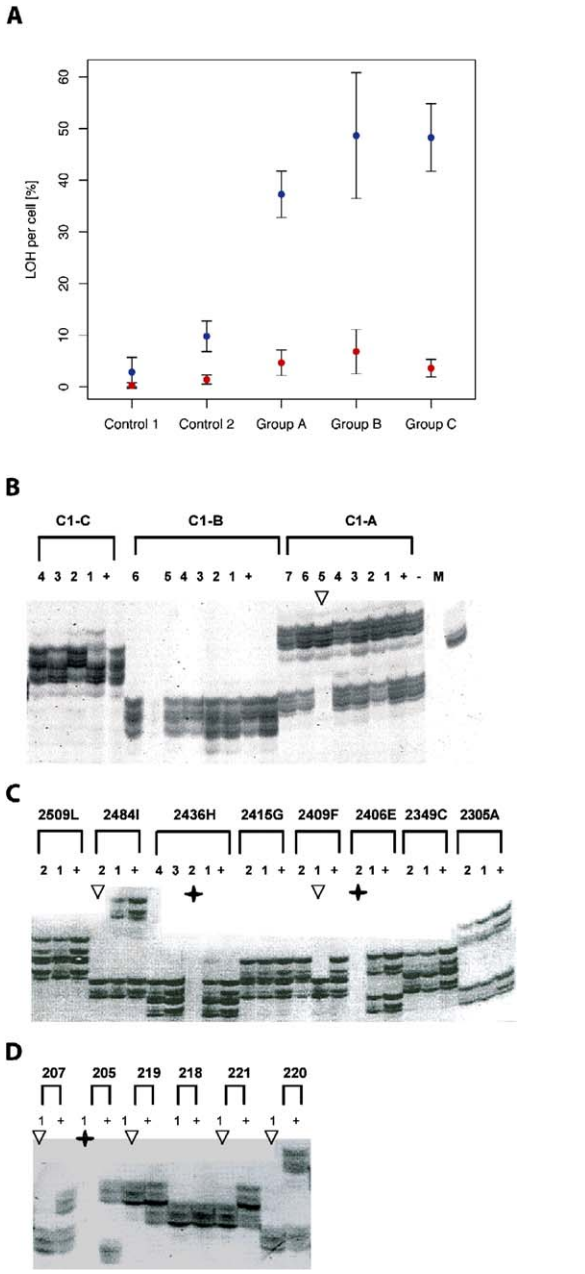


Figure 1. Cumulative analysis of allelic loss and microsatellite analysis of D16S485

A: Allelic loss per cell of all tested markers evaluated for control groups 1 and 2 and cytokeratin-positive cells from groups A–C (groups A and B were from M0 stage patients, group C was from M1 stage patients, and groups B and C were from patients with aberrant CGH profiles). Blue dots indicate heterozygous loss, and red dots indicate homozygous deletions. Error bars represent 95% confidence interval.

B–D: Polyacrylamide gel run of PCR products (D16S485) generated from single cells of control group 1 (median age 28 yrs) (**B**) and control group 2 (median age 58 yrs; **C**), and cells from patient group A (**D**). Arrows indicate loss of one band and asterisks indicate loss of two bands for a particular cell.

LOH analysis of single cytokeratin-positive cells from bone marrow

To investigate the genome of cytokeratin-positive tumor cells that were isolated from the bone marrow of breast cancer patients at higher resolution, we used 27 microsatellite and 2 PCR-RFLP (restriction fragment length polymorphism) markers. These markers were located on four chromosomes, spanning regions from 16 to 44 Mb, around genes whose products regulate cell adhesion, E-cadherin (*CDH1*), α-catenin (*CTNNA1*), β-catenin (*CTNNB1*), plakoglobin (*JUP*), and APC protein (*APC*) (Table 1). From 189 single cytokeratin-positive cells that had been isolated from the bone marrow of 371 breast cancer patients and studied by CGH (Schmidt-Kittler et al., 2003), we selected a total of 97 cells from 47 patients for LOH analysis, which included all cytokeratin-positive cells that displayed normal CGH profiles and randomly selected cells with chromosomal aberrations from patients with metastasis (clinical stage M1) and without metastasis (clinical stage M0) (Table S1). We separated these samples into three groups. Cells of groups A (n = 37) and B (n = 15) were isolated from M0 breast cancer patients that had normal and abnormal CGH profiles, respectively. Group C (n = 45) cells were isolated from M1 breast cancer patients, all of which had abnormal CGH profiles, and served to uncover genomic changes associated with advanced disease.

On average, cells from groups A, B, and C displayed 37.3% (95% CI, 32.8%–41.8%), 48.7% (95% CI, 36.5%–60.8%), and 48.2% (95% CI, 41.7%–54.8%) loss of informative markers, respectively. For statistical comparison of DNA losses, we used only the control cells from the age-matched group (control group 2; Figures 1A–1D). Strikingly, all three groups showed a significantly ($p < 0.001$) higher percentage of LOH than control group 2 cells when evaluated by exact Wilcoxon rank test (Figure 1A). Cytokeratin-positive cells that harbored chromosomal abnormalities (groups B and C) displayed a slightly higher rate of DNA loss than cytokeratin-positive cells with normal karyotypes, most likely reflecting that some markers map to regions that are frequently involved in large chromosomal rearrangements. Cells from group A and B suffered more frequently from homozygous losses than control cells ($p = 0.02$ and 0.006, respectively), while this was not the case for cells from group C ($p = 0.09$). A possible explanation for this finding could be that a complete PCR failure of a marker is more likely if one allele has been lost during tumor progression. The reduced incidence of homozygous losses in cells from group C may therefore indicate a hyperploid genome, since cytokeratin-positive cells from metastatic patients are often larger than cytokeratin-positive cells from patients without manifest metastasis (Figures 2A–2D). The significantly higher number of LOH in cytokeratin-positive cells without chromosomal imbalances (group A) compared to age-matched, bone marrow-derived cytokeratin-negative cells suggests that cytokeratin antibodies identify a subpopulation within the bone marrow that displays significant genetic instability.

Apoptosis does not account for frequent LOH in single cytokeratin-positive cells

Although the isolated cytokeratin-positive cells looked perfectly intact (Figures 2A–2D), we tested whether the high frequency of LOH might result from apoptosis induced in ectopically homing epithelial cells. We cultured normal mammary epithelial cells from a sample obtained after breast reduction surgery (undertaken for a nonneoplastic condition) and selected epithelial cells by generation of mammospheres, which allow the propagation of mammary precursor cells (Dontu et

Table 1. LOH analysis of cytokeratin-positive cells versus age-matched control cells

	Marker and genes[a]	Localization (bp)[a]	p value*			
			Control group 2 versus control group 1	Control group 2 versus group A	Control group 2 versus group B	Control group 2 versus group C
Chromosome 3	D3S3518	33662359	1.0	0.007**	0.166	0.153
	β-catenin marker[b] (intron 7)	41231926	1.0	0.0003**	0.088	0.023**
	D3S3624	44588821	1.0	0.049**	1.0	<0.0001**
	D3S32	46515925	1.0	0.129	1.0	0.0005**
	D3S1514	58193152	1.0	0.343	0.169	0.016**
Chromosome 5	D5S299	101675106	1.0	1.0	1.0	0.0003**
	rs42427 (intragenic of *APC*)	112204224	1.0	0.048**	0.025**	0.007**
	D5S346	112241523	1.0	0.132	0.010**	0.001**
	D5S471	119076934	1.0	0.088	0.329	0.034**
	D5S592	119129568	1.0	0.369	0.298	0.328
	D5S615	125191192	1.0	0.129	0.276	0.101
	D5S2117	133065027	1.0	0.228	0.088	0.088
	D5S816	135329409	1.0	0.066	0.329	0.009**
	D5S399	135991437	1.0	1.0	1.0	1.0
	D5S500	137874849	1.0	0.088	0.568	0.011**
	D5S1360	144877829	1.0	0.129	0.025**	0.427
Chromosome 16	D16S3019	64686681	1.0	0.696	1.0	0.009**
	D16S3095	68503731	1.0	0.215	0.01**	0.0008**
	D16S485	70292497	1.0	0.008**	0.052	0.0009**
	D16S3066	71887336	1.0	0.049**	0.025**	<0.0001**
	D16S3138	76161268	1.0	0.088	0.039**	0.0005**
	D16S3040	78209120	1.0	0.132	0.003**	0.0008**
	D16S511	80258795	1.0	0.129	0.023**	<0.0001**
Chromosome 17	D17S800	36309949	1.0	0.257	0.657	0.028**
	D17S855 (*BRCA1* intron 20)	38458270	1.0	0.370	0.298	0.199
	D17S1322 (*BRCA1* intron 19)	38468875	1.0	0.654	1.0	0.046**
	D17S1868	44539852	1.0	0.049**	0.011**	0.0008**
	D17S943	45195357	1.0	0.055	0.040**	0.091
	D17S1161	53686418	1.0	0.129	0.123	0.013**

*p values of Fisher's exact test between control group 2 cells and control group 1, group A, group B, or group C cells for each marker corrected for multiple testing by controlling the FDR. **Significant values.

[a] Markers are ordered according to their chromosomal localization in bp (ensembl database [July 2005]; www.ensembl.org).

[b] Intragenic polymorphism obtained as personal communication from F. van Roy; see also Nollet et al. (1996).

al., 2003). Subsequently, the cells were kept under differentiating conditions. We induced apoptosis by either camptothecin (Morris and Geller, 1996; incubated 1.5 and 3 hr) or serum starvation for 24 hr. Apoptosis was monitored by double staining using a GFP-annexin V fusion protein and propidium iodide (PI; Egger et al., 2003). GFP-annexin V single-positive cells (Figures 2E and 2F) were considered to undergo early, double-positive cells (Figures 2G–2I) to undergo advanced apoptosis. From the three conditions (1.5 hr camptothecin, 3 hr camptothecin, and serum starvation), we isolated individual unstained cells, GFP-annexin V single-positive, and GFP-annexin V-PI double-positive cells (each n = 10; Figures 2E–2L). Unstained cells were regarded as nonapoptotic and served as control. After global amplification, we performed LOH analysis using all markers (n = 16) for which the donor DNA was informative (Figure S1). Unlike with the cytokeratin-positive cells from bone marrow, we did not perform CGH analysis as quality control but included cells that were positive in a gene-specific PCR using primers binding to the cytokeratin 19 sequence.

Compared with unstained control cells, GFP-annexin V single-positive and GFP-annexin V-PI double-positive cells showed a nonsignificant increase of LOH ($p = 0.4$; Figure 2M and Figure S1). The frequency of LOH did not differ between the groups for any marker when analyzed individually. Markers displaying complete PCR failure were excluded. However, in contrast to the cytokeratin-positive and control group 2 cells, where homozygous losses occurred randomly, complete PCR failure occurred in four cells that showed "homozygous deletion" of 88%–100% of markers. Since such DNA samples are unsuited for CGH, they would not have been included in our study of cytokeratin-positive cells. Of note, the almost complete loss of DNA occurred exclusively in camptothecin-induced apoptosis. We therefore compared the different modes of apoptosis induction for their effect on the rate of LOH. Interestingly, LOH slightly increased with exposure to camptothecin, while apoptotic serum-starved cells did not show an increase of LOH as compared to the unstained controls (Figure 2N). Comparing all GFP-annexin V single-positive and GFP-annexin V-PI double-positive cells (n = 16, excluding the four cells with 88%–100% loss of DNA) with the cytokeratin-positive cells from group A, the latter displayed significantly higher frequency of LOH than apoptotic cells ($p < 0.001$).

We then tested two cancer cell lines (the breast cancer cell line SKBR3 and the colon carcinoma cell line SW480) and iso-

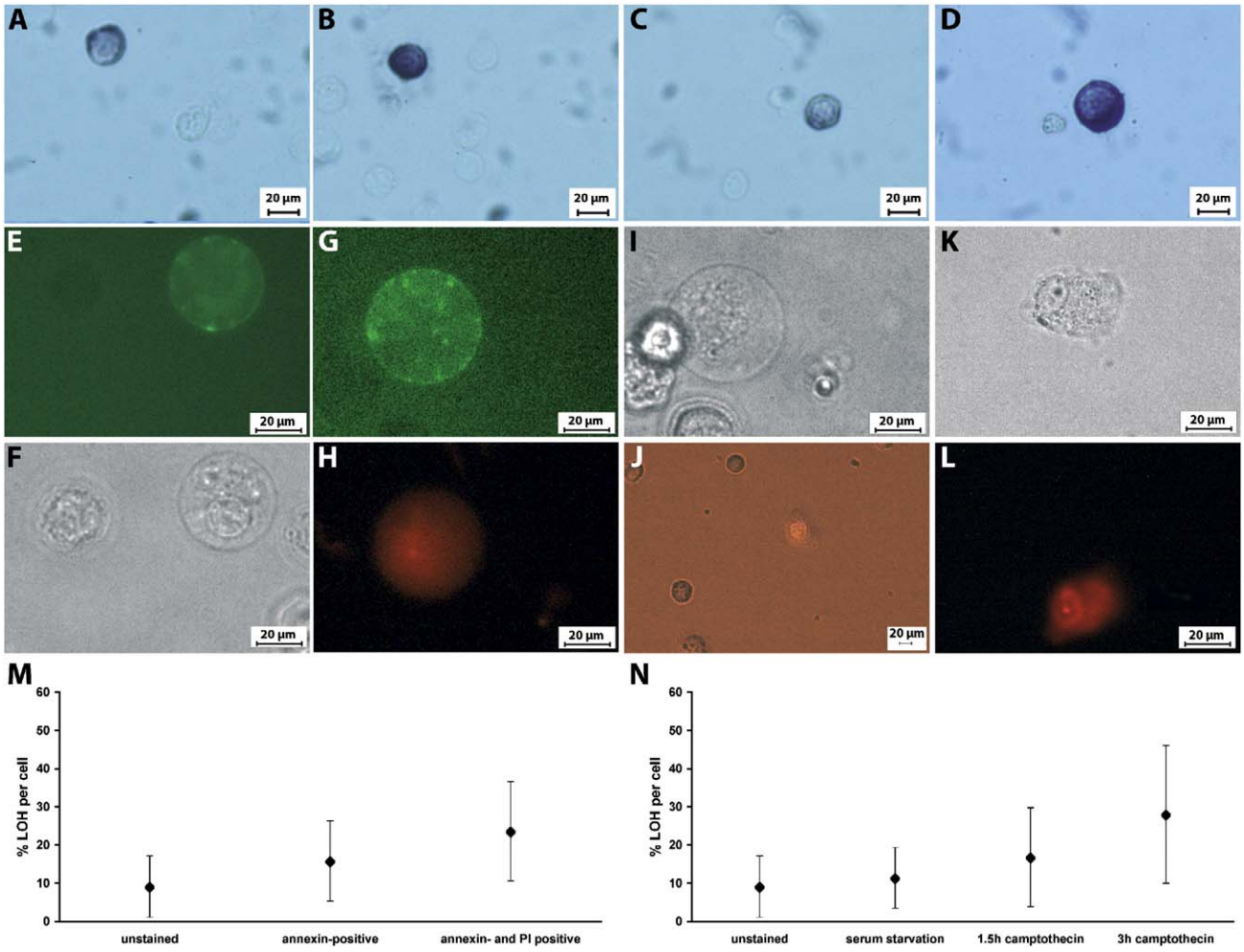


Figure 2. Apoptosis does not account for frequent LOH in cytokeratin-positive cells

A–D: Single cytokeratin-positive cells from bone marrow of breast cancer patients between unstained cytokeratin-negative bone marrow cells. The cells shown in **A**–**C** were isolated from patients in stage M0 and displayed normal CGH profiles (**A**, A-201-1; **B**, A-213-1; **C**, A-202-1; for LOH data of these cells see Figure 5), while the cell shown in **D** (C-111-2) is isolated from a patient in stage M1. Note its larger size in comparison with the cells in **A**–**C**, suggesting a polyploid genome.

E and F: GFP-annexin V single-positive cell showing the same cell in the GFP filter (**E**) and in bright-field microscopy (**F**).

G–I: GFP-annexin V and propidium iodide (PI) double-positive cell. GFP staining is shown in **G**, PI is shown in **H**, and bright field is shown in **I**.

J: Round cellular morphology is lost soon after cells become PI positive. Note the pale, flattened PI-positive cell giving low contrast, as compared with the round, PI-negative cells with high contrast (bright-field microscopy with open PI filter).

K and L: Amorphous, PI-positive cell at later stages of cell death in bright-field microscopy (**K**) and PI filter (**L**).

M: LOH of unstained, GFP-annexin V, and GFP-annexin V-PI double-positive cells. Error bars indicate 95% confidence interval.

N: The mode of apoptosis induction influences the rate of LOH. Error bars indicate 95% confidence interval.

lated PI-positive cells after induction of apoptosis using camptothecin (1.5 and 3 hr). Here, even in camptothecin-induced apoptosis, no difference was observed for LOH of nonapoptotic, coisolated cells (n = 12; mean 9% and 95% CI, 0%–19%) and PI-positive SKBR3 (n = 5; mean 14%, 95% CI, 0%–30%) or SW480 cells (n = 6; no loss observed).

Finally, we reasoned that isolation of apoptotic cells on the basis of annexin V and PI double labeling might still be too early to detect significant DNA degradation in apoptotic cells. However, once cells were PI positive, disintegration of cellular morphology occurred very rapidly. During microscopic inspection, the cells became shiny, pale, and amorphous (Figures 2J–2L)—clearly different from the cytokeratin-positive cells that we isolated from bone marrow of cancer patients (Figures 2A–2D). We were unable to amplify DNA from cells with amorphous appearance, indicating rapid DNA degradation once this stage is reached. Taken together, these experiments rule out the possibility that apoptosis accounts for the high frequency of LOH observed in CK-positive cells from bone marrow.

Identification of genomic regions frequently affected by LOH

We then compared the frequency of LOH for control group 2 cells and cytokeratin-positive cells from groups A, B, and C

for each marker using Fisher's exact test. While we found no characteristic differences between cells from control group 1 and 2, specific markers were significantly more often deleted even in cytokeratin-positive cells without chromosomal abnormalities (group A) than in control group 2 cells (Table 1). These markers may thus enable identification of DNA changes that occur in cytokeratin-positive cells before the onset of chromosomal instability. The number of significantly altered markers was similar for group A and group B cells and increased for cells from group C, which often harbored chromosomal aberrations within the tested regions. Since the markers spanned relatively large genomic regions, this analysis provided only an overview of where frequent LOH may occur. However, in two of the five selected chromosomal regions, the markers with highest statistical significance for LOH in cells from group A mapped within genes functionally involved in adhesion and cellular differentiation, i.e., *CTNNB1* and *APC* (Table 1). On chromosome 17, the gene encoding plakoglobin (*JUP*; positioned at 37.2 Mbp) was clearly not affected by LOH in cytokeratin-positive cells without CGH abnormalities, since the markers with highest significance mapped 7–8 Mbp apart. The region of α-catenin (*CTNNA1*; positioned at 138.1 Mbp on chromosome 5 between D5S500 and D5S1360) was not affected in group A cells but may occasionally be deleted in groups B and C. On chromosome 16, the marker most frequently deleted in cells from group A (D16S485) mapped only 3 Mbp from the E-cadherin gene (*CDH1*; positioned at 67.3 Mbp) toward the telomere. While D16S3095 is more closely located to *CDH1*, it was rarely informative in control cells and may have reduced the statistical significance.

Therefore, we attempted to better define the regions harboring LOH and to validate the findings. We selected the two markers displaying the highest statistical significance of LOH in cells from group A (D16S485 and β-catenin marker) and added eight markers for intragenic single-nucleotide polymorphisms (SNP) mapping proximally and distally to the microsatellite marker D16S485 and 11 SNPs in distal or proximal location of the β-catenin marker on chromosome 3. For these markers, we tested all cells from group A. Since the samples were not as often informative for the SNP markers as for the microsatellite markers, we used a sliding window approach of three adjacent markers for statistical evaluation of the regions. Thus, the region on chromosome 16 (from 67.3 Mbp to 71.9 Mbp) was analyzed by nine windows, which covered 0.9 Mbp on average. The larger region on chromosome 3 (33.4–44.6 Mbp) was analyzed by 12 windows covering 1.7 Mbp on average. For all triplet markers, the rate of LOH was significantly higher than that in control cells (Figure 3). Thus, the regions comprising *CDH1* (E-cadherin) and *CTNNB1* (β-catenin) are frequently lost in cytokeratin-positive cells without CGH abnormalities, strongly supporting the origin of the cells from transformed breast tissue.

Comparison with the matched primary tumor

The SNP microsatellite analysis allowed the definition of regions of frequent allelic losses in single disseminated breast cancer cells. We then asked whether we could detect the same LOH in the corresponding primary tumors. To this end, we obtained tissue blocks of formalin-fixed, paraffin-embedded tumors from 16 patients of group A. We performed laser microdissection to isolate the tumor cells (Figures 4A and 4B), and whenever possible (n = 8) we isolated DNA from several regions of the primary tumors. Since archived laser-microdissected, formalin-fixed, paraffin-embedded tissue frequently contains degraded or damaged DNA, we used conditions that we had previously defined for successful analysis (Stoecklein et al., 2002). Sufficient material was isolated and globally amplified similarly to single cells, and only those markers mapped on MseI fragments smaller than 400 bp were selected for analysis (Stoecklein et al., 2002). Single disseminated cancer cells and their matched primary tumors were compared for allelic losses of 25 polymorphic markers. In general, the comparison revealed various scenarios, i.e., that all areas from a primary tumor share the same genetic lesion with the disseminated cell, that only one or few areas from a primary tumor match the disseminated cell(s), or that only one of several disseminated cells matches an area of a primary tumor (Figures 4C and 4D). However, in 9/16 cases we could detect at least one area within the matched primary tumor that shared at least one LOH with the disseminated cell (Figure 4D).

Classifying a cytokeratin-positive cell as a tumor cell

Our analyses provided substantial evidence that the group of cytokeratin-positive cells differs from the group of control cells. Comparisons with primary tumors revealed a common genetic change for some cells, most likely reflecting a shared descent. However, the absence of such a shared change does not exclude the possibility that the particular cell is indeed a tumor cell. We therefore attempted to define criteria by which an individual cytokeratin-positive cell with a normal CGH profile can be identified as a tumor cell and constructed a classifier that was trained to differentiate between group 2 control cells and group A cells. Best classification was obtained with two markers only, D16S485 and the microsatellite marker that maps within the β-catenin gene (β-catenin marker in Table S2 [primer sequences]). Control cells from group 2 could be separated from group A cells with an accuracy of 74% (the 95% CI being 64%–84%; Figure 5). Generally, this classifier can be used to select breast tumor cells or control cells on the basis of LOH markers. We tested this approach and were able to correctly identify control cells from group 1 and tumor cells from group B/C cells with a high specificity (88% control group 1, 90% group B/C).

Identification of a subgroup of breast cancer patients with early HER2 amplification

Since analysis of cells from group A appeared to uncover very early genetic events in disseminated breast cancer, we asked when the therapeutically important amplification of the *HER2* gene takes place during genomic progression. We found that *HER2* gene amplifications emerge relatively late in cellular evolution in most cases, although a subset of cytokeratin-positive, CGH-normal cells displayed *HER2* gene amplifications. To study *HER2* gene amplifications at 17q12, which are observed in 25%–30% of primary breast cancers (Slamon et al., 1989), we designed a quantitative (q) PCR approach and tested 92 of the cytokeratin-positive cells previously analyzed for LOH and CGH (Figures 6 and 7). Gain of *HER2* was assigned to a cytokeratin-positive cell when the qPCR result differed significantly from that obtained from control cells (Figures 6 and 7). Since gene amplifications are defined as an increased ratio between *HER2* sequences and centromere copy number of chromo-

	SNP	Location (Mbp)	corrected p-value/triplet	
chr. 3	rs2293250	33409835		
	rs2049281	33527008	<0.0001	
	D3S3518	33662359		<0.0001
	rs969818	35700254	0.003	
	rs11710546	39281223		0.009
	rs1714414	40179182	0.002	
	rs1047855	40443900		<0.0001
	beta	41231926	<0.0001	
	rs3816945	41472119		<0.0001
	rs1565957	41582705	0.001	
	rs2290134	42226267		0.002
	rs1046910	42228668	0.02	
	rs12638282	43566146		0.01
	D3S3624	44588821		
chr. 16	rs2296408	67271324		
	rs1801552	67414942	0.001	
	D16S3095	68503731		0.0005
	rs14214	68752392	0.0007	
	rs3762171	69103735		<0.0001
	rs8051212	70053025	<0.0001	
	D16S485	70292497		<0.0001
	rs2278031	70326001	0.0004	
	rs699444	71385259		0.002
	rs740178	71389636	0.0005	
	D16S3066	71887336		

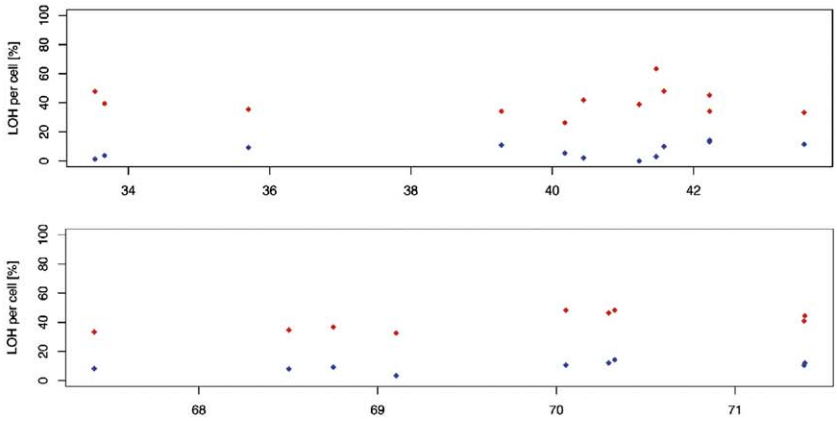


Figure 3. High-resolution analysis of markers in the proximity of *CTNNB1* on chromosome 3 and D16S485 on chromosome 16

The table depicts markers, their chromosomal location (in Mbp), and the p values (corrected for multiple testing by a FDR approach) for a region covered by three consecutive markers. Each p value is written next to the marker in the center of a triplet. Red and blue dots indicate LOH per cell for each marker triplet of cells from group A and cells from the controls, respectively. Scale on x axis is in Mbp. "beta" stands for "β-catenin marker" (see Table 1).

some 17, they can be assigned to cells with significantly elevated *HER2* measurements when the centromere copy number of chromosome 17 is normal. Consequently, an increased value for the HER2 PCR in a karyotypically normal cell by definition identifies a *HER2* gene amplification. We found *HER2* amplification or gain in 9/52 (17%) cases (group A 6/37 [16%] and group B 3/15 [20%]) in stage M0 and in 22/40 (55%) cells from metastatic patients (Figure 7B). Thus, the frequency of

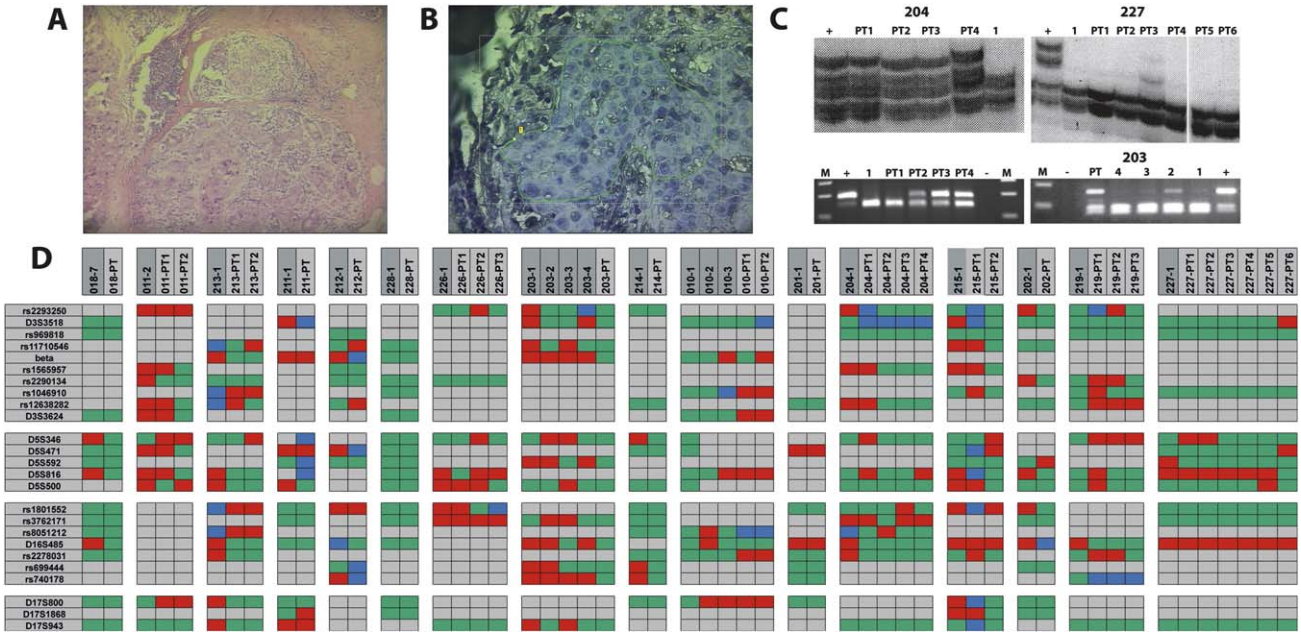


Figure 4. Comparison of disseminated tumor cells and matched primary tumors

A: Hematoxylin-eosin staining of primary breast cancer with considerable stromal component and infiltrating cells.

B: Hematoxylin staining of consecutive tissue section for laser microdissection at higher magnification. A green line marks the area selected for microdissection. Several such areas within a region were collected and pooled. Several pools were prepared from various regions of primary tumors.

C: Examples of LOH analysis. In case 204, loss of D16S485 was observed only in the tumor cells (upper panel), while PT1 and the disseminated cancer cell share loss of the same allele of rs1565957 (lower panel). D16S485 is lost in all samples of the primary tumor (PT1-PT6) and in the disseminated cancer cell of case 227. In case 203, all disseminated cancer cells display loss of rs740178, while the primary tumor harbors both alleles. (The upper weak band in cell #2 represents an incomplete restriction digest).

D: Overview of LOH analysis of group A cells and their matched primary tumors. Green color indicates informative markers, red indicates heterozygous loss, and blue indicates homozygous loss. "beta" stands for "β-catenin marker" (see Table 1).

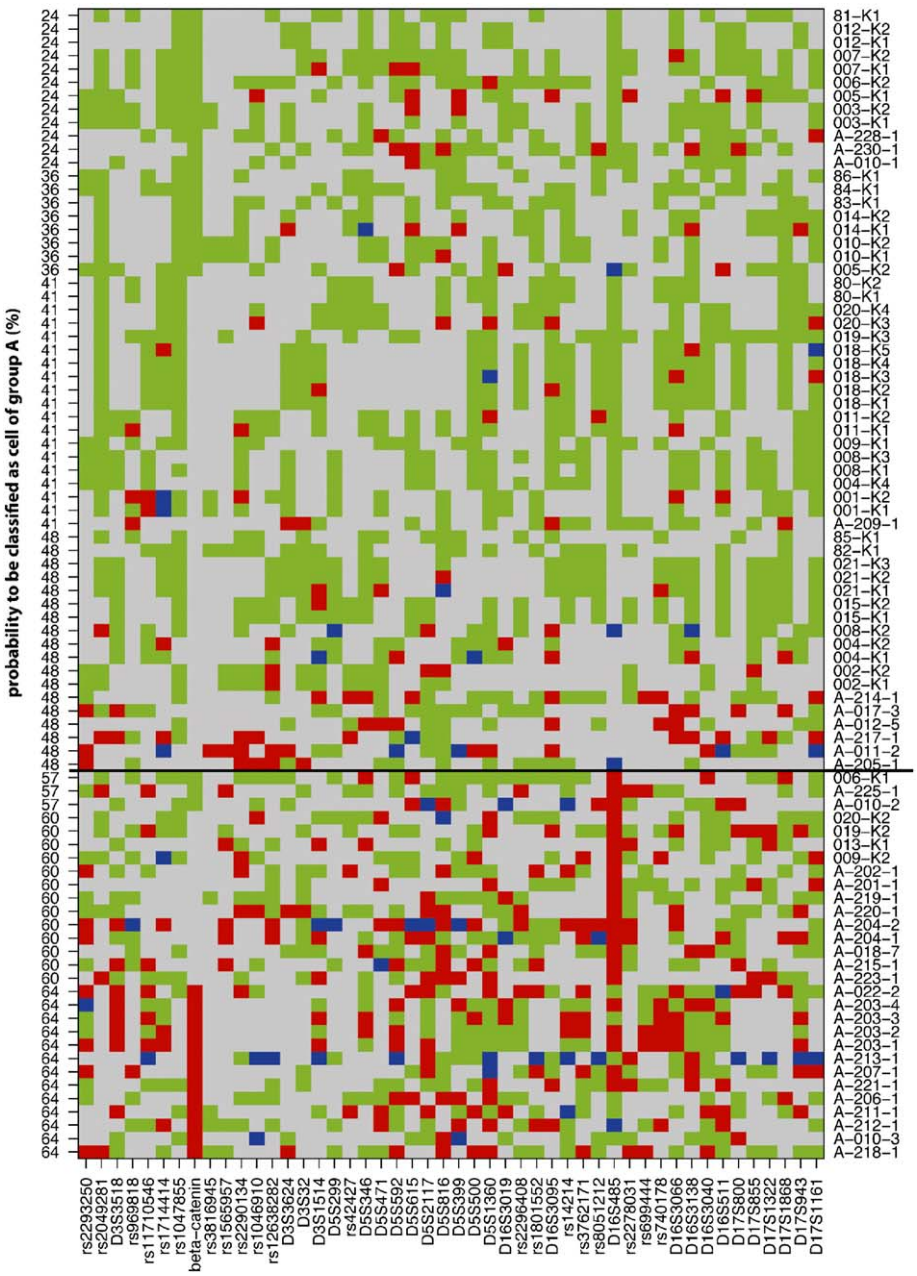
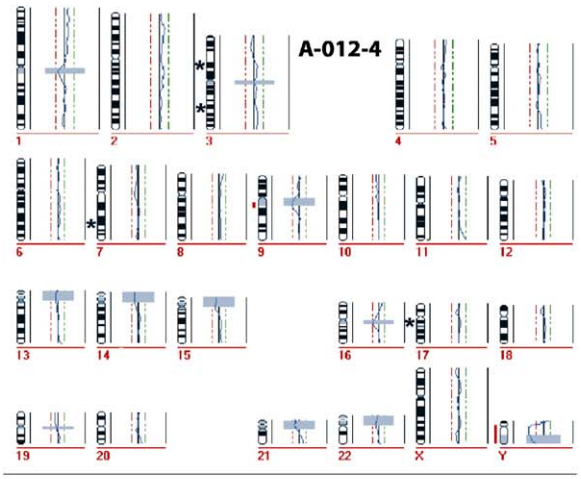
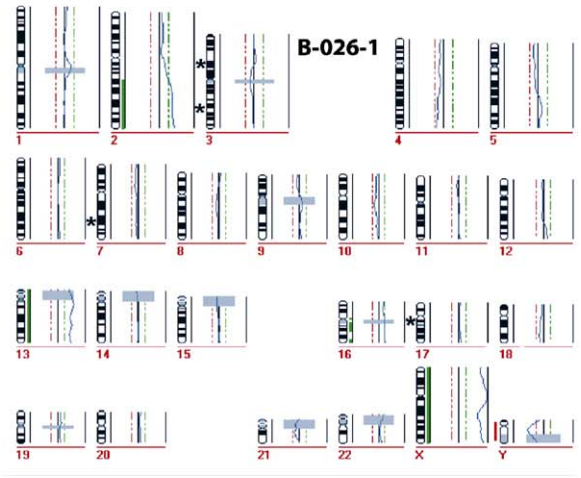


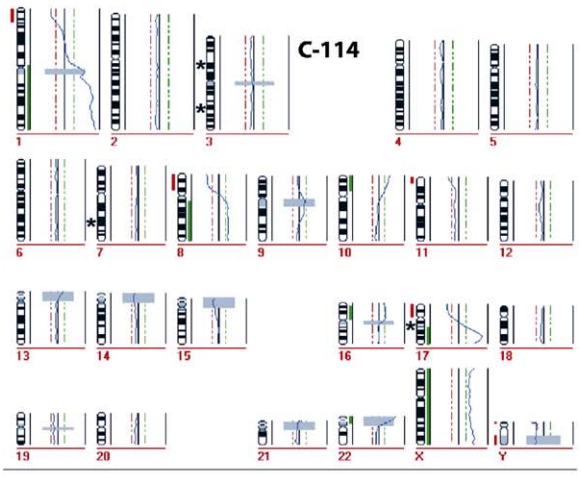
Figure 5. Classification of cells from group A based on the markers β-catenin and D16S485

A classifier was trained to distinguish between control cells and cytokeratin-positive cells without CGH aberrations. All markers (shown at the bottom of the figure) are ordered according to their chromosomal location on chromosomes 3, 5, 16, and 17. Cell identifiers are given on the right side (with letter A indicating cells from group A and letter K indicating control group 2 cells). Cells are ordered according to their probability of being classified as cells from group A. Green color indicates informative markers, red indicates heterozygous loss, and blue indicates homozygous loss.

cytokeratin-positive cells with *HER2* amplification increased significantly during systemic progression ($p < 0.001$ for group A compared with group C; Figures 7B and 7C). Cells from groups A and B did not differ significantly with regard to *HER2* amplification. Since we had previously noted that clinical metastasis concurs with the emergence of chromosomal breaks throughout the genome (Schmidt-Kittler et al., 2003) of cytokeratin-positive cells, we tested whether *HER2* gains are associated with chromosomal breaks as well. Indeed, in the majority of breast cancers *HER2* is gained after the onset of large genomic rearrangements in disseminated tumor cells (Figure 7D). When we analyzed the frequency of *HER2* gains in patients from whom we had detected and isolated more than one cell, we found *HER2* amplified in all isolated cells ("homogeneous gain" in Figure 7E) in 1/9 (11%) M0 patients. During disease progression, the percentage of mixed populations ("heterogeneous gain") and cell populations with homogeneous *HER2* gain increased (Figure 7E). We also compared the reports of HER2 immunohistochemistry of the matched primary tumors with the qPCR results of single disseminated cancer cells. Reports were available for 27 primary tumors, and *HER2* amplification was assumed for strong staining (3+). Concordant amplifications of primary tumors and disseminated tumor cells were found in 1/27 cases, while *HER2* was amplified in primary tumors but not in the disseminated cells in 10/27 matched pairs, and in 5/27 cases only the disseminated cells displayed a *HER2* amplification. The remaining pairs were double negative. Thus, these data suggest that *HER2* amplifi-







Sample	Mean rank test	Mean rank control	P-value
A-012-4	318.00	209.18	0.008
B-026-1	282.75	208.94	0.138
C-114	414.11	207.08	<0.0001

Figure 6. Analysis of HER2 amplification and karyotypic abnormalities

CGH profiles of three single cells selected from M0 stage patients with CGH abnormalities (sample B-026-1) and without CGH abnormalities (sample A-012-4) and M1 stage patients (sample C-114) and statistical evaluation of qPCR of *HER2* versus control primers for the respective cell (mean rank test) and the control cells (mean rank control). Asterisks in the CGH profiles indicate the location of the *HER2* primers on chromosome 17 and the control primers on chromosomes 3 and 7. Red and green dashed lines indicate thresholds of significance for chromosomal losses and gains, respectively, and likewise, red and green bars next to the chromosome ideogram mark regions of loss and gain. p value indicates whether or not values of mean rank test and mean rank control differed significantly (here significant for samples A-012-0 and C-114).

cation is a very early mutational event only in a subset of patients, while in most cases its gain occurs late in local and systemic tumor progression. Tumor cells with *HER2* amplification apparently become selected during systemic disease progression.

Discussion

We describe here the finding of a cell population in bone marrow of breast cancer patients that is characterized by the expression of epithelial cytokeratins and lack of chromosomal abnormalities, but general genetic instability and subchromosomal DNA changes. Due to their characteristics, the cells may represent examples of the early stages of metastasis and thus could be used to identify early genetic changes in breast cancer.

We could exclude the possibility that apoptosis accounts for these findings. While camptothecin-induced apoptosis resulted in marginally increased LOH, serum starvation did not. Genomic DNA could be globally amplified from single cells undergoing apoptosis as long as their cellular appearance was not obviously destroyed. Regardless of how apoptosis was induced, cytokeratin-positive cells with normal CGH profiles from bone marrow displayed a significantly higher rate of LOH than apoptotic cells. Since only morphologically intact cytokeratin-positive cells were isolated for this study, and their DNA had previously been successfully used in CGH analysis, and since none of the patients in stage M0 underwent chemotherapy prior to bone marrow sampling, the LOH of cytokeratin-positive cells apparently reflects their biology.

The regions frequently affected by significant LOH included genes with a potential role in the regulation of adhesion and differentiation, encoding E-cadherin, β-catenin, and APC. Although the functional consequences of LOH of these genes for tumor promotion and systemic progression are largely unknown, there is increasing evidence that a reduction in gene dosage by loss of one allele (haploinsufficiency) may severely compromise genome function (Kaern et al., 2005). For example, in a mouse model of combined APC and E-cadherin mutation haploinsufficiency of E-cadherin resulted in an increased number of tumors (Smits et al., 2000). For many tumor suppressor genes, the importance of reduced gene dosage is being recognized as a contributing mechanism of tumor progression (Santarosa and Ashworth, 2004). Generally, haploinsufficiency has been linked to increased stochasticity of gene expression. Stochastic simulations of a model of gene expression have shown that diploid cells have a higher probability than haploid cells of maintaining the abundance of an expressed gene product above a low threshold value (Cook et al., 1998). It was concluded that loss of one allele reducing expression of an essential factor below a crucial threshold might contribute to the onset of disease. Although such events

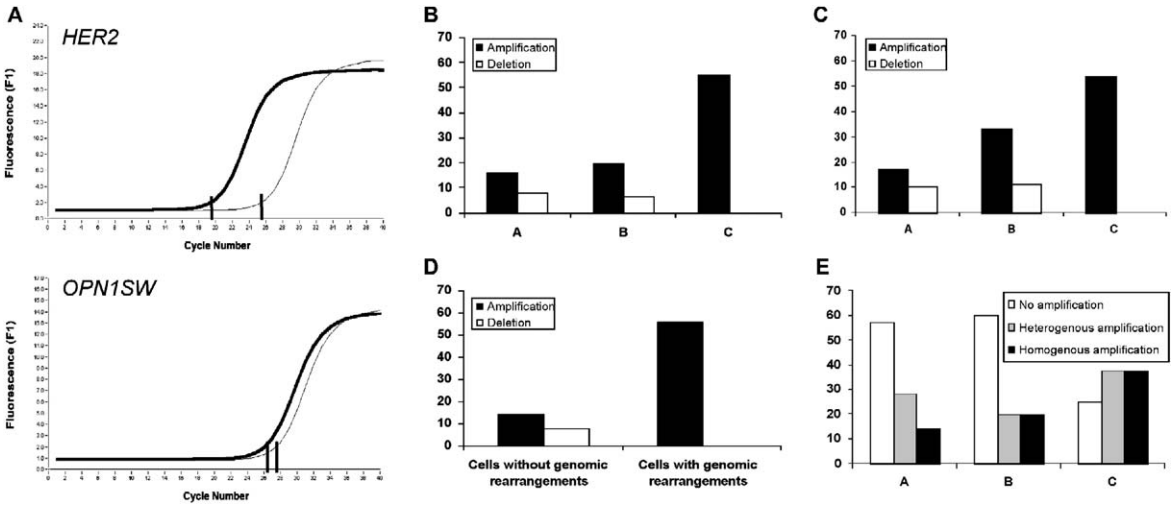


Figure 7. Quantitative PCR analysis of the *HER2* gene in single disseminated cancer cells isolated from patients with and without metastasis

A: Amplification of *HER2* (top) and the reference gene *OPN1SW* of two cytokeratin-positive cells with (solid line) and without (thin dashed line) *HER2* amplification. Note the ratios of the crossing points for HER2/*OPN1SW* (20/26.5 versus 28/27.6), indicating a higher *HER2* copy number for the first sample.
B: Percentage of *HER2* gains and *HER2* amplification and deletion in cells from groups A–C. In only one cell from group B, the higher values for the *HER2* amplicons originated from an additional chromosome 17.
C: Analysis of *HER2* amplification and deletion based on patients (i.e., from some patients several cells were isolated; in percent).
D: Percentage of *HER2* amplification and deletion in cells with or without genomic rearrangements.
E: Distribution of patients from whom more than a single cell was isolated with or without *HER2* amplification (in percent). Note that during disease progression the number of patients increases that have a mixed population of disseminated tumor cells (heterogeneous amplification) and that harbor a homogeneous cell population with respect to *HER2* amplification.
Groups A–C were defined by clinical stage (groups A and B, M0; group C, M1) and CGH profile (group A, normal profile; groups B and C, aberrant profile).

would occur only rarely, the probability would gradually accumulate, making the onset of disease more likely later in the life of the organism (Kaern et al., 2005).

In this context, our findings suggest a model in which sporadic cancers may result from genome-wide double-strand breaks, subchromosomal DNA losses, and resulting increased stochasticity of gene expression in individual cells, which are then subject to natural selection. So far, mutations in specific genes (Hahn et al., 1999; Hanahan and Weinberg, 2000), aneuploidy (Duesberg et al., 2004; Li et al., 2000), and telomere shortening leading to genomic rearrangements (Artandi et al., 2000; DePinho, 2000) have been implicated in initiating local and systemic malignant disease. We delineated the genomic aberrations in human breast cancer by multiple genetic analyses (i.e., CGH, LOH, and qPCR) of single disseminated cancer cells and determined the sequence of genomic changes. Relative to a starting point of a normal diploid genome and an end point that is represented by a metastatic cancer cell (characterized by CGH-defined genomic rearrangements such as intrachromosomal gains or losses or DNA breaks that include the telomeres [Schmidt-Kittler et al., 2003]), cytokeratin-positive cells without CGH abnormalities are located close to the starting point. They differ from normal cells by subchromosomal DNA losses and in some cases gene amplifications, such as *HER2* gains. The next genomic change could be the emergence of aneuploidy and later on large CGH-defined chromosomal rearrangements (Schmidt-Kittler et al., 2003).

Such a model is consistent with two recent reports indicating that replication stress in precancerous lesions leads to double-strand breaks not detected by CGH and activates the DNA damage checkpoint (Bartkova et al., 2005; Gorgoulis et al., 2005). Replication stress was very prominent at the stage of hyperplasia, and chromosomal aberrations arose only after the breakdown of the DNA damage checkpoint. In light of the fact that 95% of primary tumors and preinvasive stages of mammary neoplasias display chromosomal changes detectable by CGH, cytokeratin-positive tumor cells without CGH aberrations must be representatives of a very early stage of genomic progression and possibly disseminated at the stage of hyperplasia. Therefore, the finding that at the time point of surgery more than 50% of disseminated breast cancer cells display normal karyotypes (Schmidt-Kittler et al., 2003) remains puzzling and is at variance with the linear progression model deduced from colorectal cancer. It could indicate the existence of a particular state of cellular transformation at which the cells are prone to disseminate, whereas late, chromosomally aberrant primary tumors seed fewer cells—at least relative to the total number of cells in the primary tumor. It also suggests that the disseminated cells will have to overcome several thresholds to full malignancy, first the DNA damage checkpoint and second telomere crisis. The need to overcome these thresholds might explain the relatively long latency periods (on average 6 years) for small breast cancers until metastases become apparent and suggests that the probability to grow into metastasis is small for an individual disseminated tumor cell.

Although detection of cytokeratin-positive cells in bone marrow is relevant for prognosis (Braun et al., 2000b; Gebauer et al., 2001; Gerber et al., 2001; Wiedswang et al., 2003) and prediction of therapy response (Braun et al., 2000a; Janni et al., 2005), these studies did not differentiate between CK-positive

cells with and without abnormal CGH profiles. Survival analysis of our cohort within the next years should provide helpful information. However, irrespective of whether disseminated tumor cells with normal karyotypes will eventually form metastases, they might be interesting for therapy target research—not necessarily as target cells themselves but as tools to define relevant target genes. Early dissemination and independent progression at ectopic sites is likely to result in malignant genotypes and phenotypes that differ not only between disseminated cells and their primary tumors but also between colonies residing at different ectopic sites. Thus, the DNA changes that are the best candidates for being shared among all tumor cells are those preceding or causing dissemination. This rationale is exemplified by the finding that early amplification of *HER2*, i.e., prior to large CGH-defined genomic rearrangements, defines a particular group of M0 stage breast cancer patients. In most cases, *HER2* amplification will occur late in cancer progression and may not be present in disseminated cancer cells when adjuvant treatment is applied after surgery. Therefore, patients with early *HER2* amplification might benefit more from adjuvant therapy with HER2-directed drugs than those that amplify the gene later in disease progression.

For all other patients, similar targets are yet to be defined. Comparison between primary tumors and disseminated cancer cells identified a common deletion in 50% of analyzed cases. While this supports the origin of the cells from the primary site, it also indicates that more work is necessary to identify changes that caused the growth of sporadic cancers. Therefore, a genome-wide, comparative survey of disseminated cytokeratin-positive cells and primary tumors might be helpful for the identification of genetic lesions that initiate sporadic cancer and metastatic spread and useful for the detection of novel therapy targets.

Experimental procedures

Cell isolation and template preparation

Cytokeratin-positive cells of patients were isolated, and their DNA was amplified and used for CGH as described (Klein et al., 1999, 2002). Samples were taken from the study cohort published (Schmidt-Kittler et al., 2003). Informed consent was obtained from all patients, and the local ethics committees in Frankfurt and Augsburg approved the study. For LOH analysis, the primary PCR products were reamplified for all samples using 0.5 μl as template. Reamplification from primary PCR products was performed in 50 μl containing 5 μl Expand Long Template Puffer 1 (Roche), 5 μl LIB 1 or Mse 21 primer (10 μM), 1.75 μl dNTP (10 mM), 1.25 μl BSA (Roche), and 0.5 μl Taq-Polymerase (5 U/μl; Roche) for 30 cycles. No substantial differences in the amplification efficiency of single-copy sequences were observed between primary PCR products and reamplified PCR products. The positive control for each patient was generated by incubating the slide from which single cytokeratin-positive cells had been isolated with 80 μl of solution A (100 mM KCl, 10 mM Tris/HCl [pH 8.3], 2.5 mM $MgCl_2$) and 80 μl solution B (10 mM Tris/HCl [pH 8.3], 2.5 mM $MgCl_2$, 1% Tween 20, 1% NP40, Proteinase K 120 μg/ml) for 10 hr at 42°C, followed by inactivation of Proteinase K at 80°C for 10 min. After precipitation, the DNA was globally amplified like the DNA of the single cells.

LOH analysis

A volume of 0.5 μl of reamplified primary PCR product was used as template for sequence-specific PCR. The primer sequences for the microsatellite and the PCR-RFLP markers are provided as Supplemental Data. Sequence-specific PCR using the microsatellite or PCR-RFLP markers was performed in a volume of 10 μl containing 1 μl 10× PCR buffer (10 mM $MgCl_2$, 100 mM Tris [pH 8.5], 500 mM KCl, 1 mM dNTPs), 0.5 μl forward primer (8 μM), 0.5 μl reverse primer (8 μM), 0.25 μl BSA (Roche), 0.1 μl Taq-Polymerase (5 U/μl; Roche). Microsatellite markers were run on a vertical, 7% polyacrylamide gel. The DNA was visualized by incubating the gel for 10 min with SYBR Green (1× TBE, SYBR Green 1:10,000) and analyzed on a fluorimager. Detailed information about the restriction digest of SNP markers can be obtained upon request.

Quantitative PCR

Real-time PCR was performed using a LightCycler (Roche, Mannheim, Germany) and Fast Start Master SYBR Green Kits (Roche) using 1 μl of primary PCR products from the whole genome amplification diluted 1:20 in H_2O. Analysis was done using the RelQuant software (Roche) with PCR efficiency normalization and a reference sample included for every run. Three primers within the *HER2* locus were selected, and three primers on two chromosomes (chromosome 3 and chromosome 7) served as loading control. Measurements showing unspecific products in the melting curve analysis or CGH aberrations on the control loci were discarded from further statistics. All relative expression ratios *HER2*/reference from duplicate measurements were compared to those from 55 normal diploid cells (from control groups 1 and 2). Samples showing significant difference ($p < 0.05$) from the diploid cells in a Mann-Whitney test were classified as either "amplified" or "deleted" for *HER2* depending on their mean rank value. Primer sequences are provided in Table S2.

Cell culture of mammary epithelial cells and apoptosis assay

Normal breast tissue after reduction mammoplasty was dissociated to single cells by mechanical and collagenase/hyaluronidase dissociation as described (Stingl et al., 1998). Single cells were cultivated in serum-free medium conditioned with B27, 20 ng/ml EGF, and 20 ng/ml bFGF under nonadhering conditions (Dontu et al., 2003).

After 7 days of culture, mammospheres of 50–150 μm diameter had grown out and were harvested by centrifugation at 38 × g and disintegrated by consecutive cycles of incubation with trypsin/EDTA and mechanical dissociation. The single-cell suspension was plated onto collagen-coated eight-well chamber slides in differentiation medium complemented with 5% serum and 10 ng/ml EGF. After 2–3 days of incubation, growth medium was removed and replaced by either normal growth medium as positive control, serum-free medium, or medium containing 2 μg/ml camptothecin. Apoptosis was detected after trypsination using 5 μg/ml PI and GFP-annexin V (1:500) in annexin binding buffer (10 mM HEPES, 140 mM NaCl, 2.5 mM $CaCl_2$) for 15 min (Egger et al., 2003).

Statistical analysis

The rate of LOH of all subgroups was compared using a two-sided exact Wilcoxon signed rank test in the software package R. To exclude a patient effect, a simulation study by randomly choosing one cell per patient was performed. All markers were compared using Fisher's exact test. Noninformative cells and homozygous losses were treated as missing values. The false discovery rate (FDR)-controlling procedure by Benjamini and Hochberg was used to account for multiple testing. Adjacent markers were combined in a sliding window to enhance the reliability in the result. All measurements of three adjacent markers were merged and compared by Fisher's exact test.

Machine learning

The decision tree algorithm C5.0 was applied as classifier using a 10-fold crossvalidation (CV) and a balanced design. The classification accuracy was estimated from 10 runs (10 × 10 CV). Noninformative markers and homozygous losses were encoded as missing values. All calculations were performed with the software package Clementine (http://www.spss.com/clementine/), version 8.5. The standard errors for the accuracy values and the classification probabilities were calculated as described by SPSS. For the control group 1, group C, and group B cells, the specificity was estimated from class assignments with a classification probability >60%.

Supplemental data

The Supplemental Data include two tables and one figure and can be found with this article online at http://www.cancercell.org/cgi/content/full/8/3/227/DC1/.

Acknowledgments

We are indebted to Frans van Roy for the sequence of the β-catenin polymorphism and very grateful to Peter Nelson, Judith Johnson, and George Miklos for critical reading of the manuscript; Nicole Sänger and Daniel Oruzio for helpful discussion of clinical data; and Christoph Borner for the GFP-annexin V protein. We also thank the "dreamteam" Isabell Blochberger, Elke Burghart, and Barbara Lindner for excellent technical assistance. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (SFB 456); BioFuture grants 0311884 and 0311880 from the German Federal Ministry for Education and Science; the Bavarian State Ministry of Sciences, Research, and the Arts; and NGFN grant 01GR0101.

Received: December 2, 2004
Revised: May 17, 2005
Accepted: August 24, 2005
Published: September 19, 2005

References

Allred, D.C., Mohsin, S.K., and Fuqua, S.A. (2001). Histological and biological evolution of human premalignant breast disease. Endocr. Relat. Cancer *8*, 47–61.

Artandi, S.E., Chang, S., Lee, S.L., Alson, S., Gottlieb, G.J., Chin, L., and DePinho, R.A. (2000). Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. Nature *406*, 641–645.

Arteaga, C.L., and Baselga, J. (2004). Tyrosine kinase inhibitors: Why does the current process of clinical development not apply to them? Cancer Cell *5*, 525–531.

Aubele, M., Mattis, A., Zitzelsberger, H., Walch, A., Kremer, M., Welzl, G., Hofler, H., and Werner, M. (2000). Extensive ductal carcinoma in situ with small foci of invasive ductal carcinoma: evidence of genetic resemblance by CGH. Int. J. Cancer *85*, 82–86.

Bartkova, J., Horejsi, Z., Koed, K., Kramer, A., Tort, F., Zieger, K., Guldberg, P., Sehested, M., Nesland, J.M., Lukas, C., et al. (2005). DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature *434*, 864–870.

Borgen, E., Beiske, K., Trachsel, S., Nesland, J.M., Kvalheim, G., Herstad, T.K., Schlichting, E., Qvist, H., and Naume, B. (1998). Immunocytochemical detection of isolated epithelial cells in bone marrow: non-specific staining and contribution by plasma cells directly reactive to alkaline phosphatase. J. Pathol. *185*, 427–434.

Braun, S., Kentenich, C., Janni, W., Hepp, F., de Waal, J., Willgeroth, F., Sommer, H., and Pantel, K. (2000a). Lack of effect of adjuvant chemotherapy on the elimination of single dormant tumor cells in bone marrow of high-risk breast cancer patients. J. Clin. Oncol. *18*, 80–86.

Braun, S., Pantel, K., Muller, P., Janni, W., Hepp, F., Kentenich, C.R., Gastroph, S., Wischnik, A., Dimpfl, T., Kindermann, G., et al. (2000b). Cytokeratin-positive cells in the bone marrow and survival of patients with stage I, II, or III breast cancer. N. Engl. J. Med. *342*, 525–533.

Buerger, H., Otterbach, F., Simon, R., Poremba, C., Diallo, R., Decker, T., Riethdorf, L., Brinkschmidt, C., Dockhorn-Dworniczak, B., and Boecker, W. (1999). Comparative genomic hybridization of ductal carcinoma in situ of the breast—evidence of multiple genetic pathways. J. Pathol. *187*, 396–402.

Chin, K., de Solorzano, C.O., Knowles, D., Jones, A., Chou, W., Rodriguez, E.G., Kuo, W.L., Ljung, B.M., Chew, K., Myambo, K., et al. (2004). In situ analyses of genome instability in breast cancer. Nat. Genet. *36*, 984–988.

Cook, D.L., Gerber, A.N., and Tapscott, S.J. (1998). Modeling stochastic gene expression: implications for haploinsufficiency. Proc. Natl. Acad. Sci. USA *95*, 15641–15646.

DePinho, R.A. (2000). The age of cancer. Nature *408*, 248–254.

Dontu, G., Abdallah, W.M., Foley, J.M., Jackson, K.W., Clarke, M.F., Kawamura, M.J., and Wicha, M.S. (2003). In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes Dev. *17*, 1253–1270.

Duesberg, P., Fabarius, A., and Hehlmann, R. (2004). Aneuploidy, the primary cause of the multilateral genomic instability of neoplastic and preneoplastic cells. IUBMB Life *56*, 65–81.

Egger, L., Schneider, J., Rheme, C., Tapernoux, M., Hacki, J., and Borner, C. (2003). Serine proteases mediate apoptosis-like cell death and phagocytosis under caspase-inhibiting conditions. Cell Death Differ. *10*, 1188–1203.

Gebauer, G., Fehm, T., Merkle, E., Beck, E.P., Lang, N., and Jager, W. (2001). Epithelial cells in bone marrow of breast cancer patients at time of primary surgery: clinical outcome during long-term follow-up. J. Clin. Oncol. *19*, 3669–3674.

Gerber, B., Krause, A., Muller, H., Richter, D., Reimer, T., Makovitzky, J., Herrnring, C., Jeschke, U., Kundt, G., and Friese, K. (2001). Simultaneous immunohistochemical detection of tumor cells in lymph nodes and bone marrow aspirates in breast cancer and its correlation with other prognostic factors. J. Clin. Oncol. *19*, 960–971.

Gorgoulis, V.G., Vassiliou, L.V., Karakaidos, P., Zacharatos, P., Kotsinas, A., Liloglou, T., Venere, M., Ditullio, R.A., Jr., Kastrinakis, N.G., Levy, B., et al. (2005). Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. Nature *434*, 907–913.

Hahn, W.C., Counter, C.M., Lundberg, A.S., Beijersbergen, R.L., Brooks, M.W., and Weinberg, R.A. (1999). Creation of human tumour cells with defined genetic elements. Nature *400*, 464–468.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57–70.

Janni, W., Rack, B., Schindlbeck, C., Strobl, B., Rjosk, D., Braun, S., Sommer, H., Pantel, K., Gerber, B., and Friese, K. (2005). The persistence of isolated tumor cells in bone marrow from patients with breast carcinoma predicts an increased risk for recurrence. Cancer *103*, 884–891.

Kaern, M., Elston, T.C., Blake, W.J., and Collins, J.J. (2005). Stochasticity in gene expression: from theories to phenotypes. Nat. Rev. Genet. *6*, 451–464.

Kallioniemi, A., Kallioniemi, O.P., Sudar, D., Rutovitz, D., Gray, J.W., Waldman, F., and Pinkel, D. (1992). Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science *258*, 818–821.

Klein, C.A. (2003). The systemic progression of human cancer: A focus on the individual disseminated cancer cell—the unit of selection. Adv. Cancer Res. *89*, 35–67.

Klein, C.A., Schmidt-Kittler, O., Schardt, J.A., Pantel, K., Speicher, M.R., and Riethmuller, G. (1999). Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. Proc. Natl. Acad. Sci. USA *96*, 4494–4499.

Klein, C.A., Blankenstein, T.J., Schmidt-Kittler, O., Petronio, M., Polzer, B., Stoecklein, N.H., and Riethmuller, G. (2002). Genetic heterogeneity of single disseminated tumour cells in minimal residual cancer. Lancet *360*, 683–689.

Li, R., Sonik, A., Stindl, R., Rasnick, D., and Duesberg, P. (2000). Aneuploidy vs. gene mutation hypothesis of cancer: recent study claims mutation but is found to support aneuploidy. Proc. Natl. Acad. Sci. USA *97*, 3236–3241.

Loeb, L.A. (2001). A mutator phenotype in cancer. Cancer Res. *61*, 3230–3239.

Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., et al. (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N. Engl. J. Med. *350*, 2129–2139.

Morris, E.J., and Geller, H.M. (1996). Induction of neuronal apoptosis by camptothecin, an inhibitor of DNA topoisomerase-I: evidence for cell cycle-independent toxicity. J. Cell Biol. *134*, 757–770.

Nollet, F., Berx, G., Molemans, F., and van Roy, F. (1996). Genomic organization of the human β-catenin gene (CTNNB1). Genomics *32*, 413–424.

Paez, J.G., Janne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., et al. (2004). EGFR muta-

tions in lung cancer: correlation with clinical response to gefitinib therapy. Science *304*, 1497–1500.

Santarosa, M., and Ashworth, A. (2004). Haploinsufficiency for tumour suppressor genes: when you don't need to go all the way. Biochim. Biophys. Acta *1654*, 105–122.

Schmidt-Kittler, O., Ragg, T., Daskalakis, A., Granzow, M., Ahr, A., Blankenstein, T.J., Kaufmann, M., Diebold, J., Arnholdt, H., Muller, P., et al. (2003). From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. Proc. Natl. Acad. Sci. USA *100*, 7737–7742.

Slamon, D.J., Godolphin, W., Jones, L.A., Holt, J.A., Wong, S.G., Keith, D.E., Levin, W.J., Stuart, S.G., Udove, J., Ullrich, A., et al. (1989). Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science *244*, 707–712.

Smits, R., Ruiz, P., Diaz-Cano, S., Luz, A., Jagmohan-Changur, S., Breukel, C., Birchmeier, C., Birchmeier, W., and Fodde, R. (2000). E-cadherin and adenomatous polyposis coli mutations are synergistic in intestinal tumor initiation in mice. Gastroenterology *119*, 1045–1053.

Sordella, R., Bell, D.W., Haber, D.A., and Settleman, J. (2004). Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. Science *305*, 1163–1167.

Stingl, J., Eaves, C.J., Kuusk, U., and Emerman, J.T. (1998). Phenotypic and functional characterization in vitro of a multipotent epithelial cell present in the normal adult human breast. Differentiation *63*, 201–213.

Stoecklein, N.H., Erbersdobler, A., Schmidt-Kittler, O., Diebold, J., Schardt, J.A., Izbicki, J.R., and Klein, C.A. (2002). SCOMP is superior to degenerated oligonucleotide primed-polymerase chain reaction for global amplification of minute amounts of DNA from microdissected archival tissue samples. Am. J. Pathol. *161*, 43–51.

Waldman, F.M., DeVries, S., Chew, K.L., Moore, D.H., II, Kerlikowske, K., and Ljung, B.M. (2000). Chromosomal alterations in ductal carcinomas in situ and their in situ recurrences. J. Natl. Cancer Inst. *92*, 313–320.

Wiedswang, G., Borgen, E., Karesen, R., Kvalheim, G., Nesland, J.M., Qvist, H., Schlichting, E., Sauer, T., Janbu, J., Harbitz, T., and Naume, B. (2003). Detection of isolated tumor cells in bone marrow is an independent prognostic factor in breast cancer. J. Clin. Oncol. *21*, 3469–3478.