Genomic Profiles Associated with Early Micrometastasis in Lung Cancer: Relevance of 4q Deletion

Michaela Wrage,[1] Salla Ruosaari,[5] Paul P. Eijk,[6] Jussuf T. Kaifi,[3] Jaakko Hollmén,[4] Emre F. Yekebas,[3] Jakob R. Izbicki,[3] Ruud H. Brakenhoff,[6] Thomas Streichert,[4] Sabine Riethdorf,[1] Markus Glatzel,[2] Bauke Ylstra,[6] Klaus Pantel,[1] and Harriet Wikman[1]

Abstract

Purpose: Bone marrow is a common homing organ for early disseminated tumor cells (DTC) and their presence can predict the subsequent occurrence of overt metastasis and survival in lung cancer. It is still unclear whether the shedding of DTC from the primary tumor is a random process or a selective release driven by a specific genomic pattern.

Experimental Design: DTCs were identified in bone marrow from lung cancer patients by an immunocytochemical cytokeratin assay. Genomic aberrations and expression profiles of the respective primary tumors were assessed by microarrays and fluorescence *in situ* hybridization analyses. The most significant results were validated on an independent set of primary lung tumors and brain metastases.

Results: Combination of DNA copy number profiles (array comparative genomic hybridization) with gene expression profiles identified five chromosomal regions differentiating bone marrow-negative from bone marrow-positive patients (4q12-q32, 10p12-p11, 10q21-q22, 17q21, and 20q11-q13). Copy number changes of 4q12-q32 were the most prominent finding, containing the highest number of differentially expressed genes irrespective of chromosomal size ($P = 0.018$). Fluorescence *in situ* hybridization analyses on further primary lung tumor samples confirmed the association between loss of 4q and bone marrow-positive status. In bone marrow-positive patients, 4q was frequently lost (37% versus 7%), whereas gains could be commonly found among bone marrow-negative patients (7% versus 17%). The same loss was also found to be common in brain metastases from both small and non-small cell lung cancer patients (39%).

Conclusions: Thus, our data indicate, for the first time, that early hematogenous dissemination of tumor cells might be driven by a specific pattern of genomic changes.

Lung cancer is one of the most frequently diagnosed cancers in developed countries and the main cause of cancer-related deaths, with an overall relative 5-year survival rate of 15% (1). Approximately 40% of patients with completely resected non-small cell lung cancer (NSCLC) without lymph node metastasis (N_0) or clinical signs of overt distant metastases (M_0) at time of the primary surgery relapse within 24 months. Thus, single disseminated tumor cells (DTC), not recognized by current routine imaging methods, may be already present in distant organs at the time of resection of the primary lung tumor (2, 3).

Sensitive immunocytochemical assays have revealed that bone marrow is a common reservoir for DTCs derived from various types of malignant epithelial tumors (3). In NSCLC, several groups have reported a significant positive association between the presence of DTCs in bone marrow and an unfavorable prognosis, indicating that DTCs may give rise to distant metastases in the bone and other organs (2, 4, 5).

It is still debated whether the release of cells from the primary tumor is a selective process or rather represents a more random shedding of cells into the circulation. Recent reports have shown that early hematogenous dissemination of breast tumor cells appears to be associated with a gene expression signature of the primary tumor, suggesting that early hematogenous dissemination of DTCs is a specific process driven by a set of control genes (6, 7). However, nothing is known thus far about underlying genomic aberrations and no comparable investigations have been done in lung cancer. Genomic aberrations leading to cancer progression and metastases can be extremely useful in diagnosis and treatment decision as has been shown

Authors' Affiliations: Institutes of [1]Tumor Biology and [2]Neuropathology and Departments of [3]General, Visceral and Thoracic Surgery and [4]Clinical Chemistry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [5]Department of Information and Computer Science, Helsinki University of Technology, Espoo, Finland; and [6]Departments Pathology and Otolaryngology, Head and Neck Surgery, VU University Medical Center, Amsterdam, the Netherlands

Received 8/28/08; revised 10/21/08; accepted 10/21/08; published OnlineFirst 02/10/2009.

Grant support: European Union [Marie Curie TSFR (H. Wikman) and DISMAL (K. Pantel)], International Union Against Cancer (H. Wikman), COST B19 action WG4 (M. Glatzel), DFG grant PA341/15-2, and Academy of Finland grant 207469 (S. Ruosaari).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Requests for reprints: Harriet Wikman, Institute of Tumor Biology, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany. Phone: 49-40-7410-51913; Fax: 49-40-7410-55379; E-mail: h.wikman@uke.uni-hamburg.de.

doi:10.1158/1078-0432.CCR-08-2188

Translational Relevance

Non-small cell lung cancer is the leading cause of cancer-related deaths worldwide and the presence of early disseminated tumor cells in the bone marrow can predict the occurrence of overt metastasis and survival. Using array comparative genomic hybridization and expression profiling on primary lung tumors, five chromosomal regions were identified differentiating patients with and without disseminated tumor cells in the bone marrow at time of surgery. Loss at 4q was further validated on a larger set of both primary tumors and brain metastases. These results indicate for the first time that dissemination of tumor cells appears to be a specific process driven by a defined set of molecular changes rather than a random shedding of cells into the circulation. Identifying a genomic signature that could predict the risk of early tumor dissemination could be extremely useful in the diagnosis and treatment decision of non-small cell lung cancer.

by the utilization of commercial fluorescence *in situ* hybridization (FISH) kits for different types of cancer (8, 9).

In the present study, we investigated whether the early dissemination of tumor cells into bone marrow is associated with a specific molecular pattern in primary lung cancer. For the detection of DTCs, we used an immunocytochemical cytokeratin assay that allows the identification of one tumor cell in the background of 1 million normal bone marrow cells (10). The specificity of this assay was shown previously by the analysis of ~200 noncarcinoma control patients (10). By using high-resolution comparative genomic hybridization (CGH) and expression arrays, we compared the genomic aberration and expression profiles of lung carcinomas from patients with and without DTCs in the bone marrow. A three-step analysis with independent measurement technologies revealed tumor-specific genetic signatures associated with early hematogenous dissemination of NSCLC cells into bone marrow. Interestingly, the most prominent aberration identified was also frequently observed in brain metastases, suggesting a potential involvement in both tumor cell dissemination and metastatic progression.

Materials and Methods

Patients and tissue samples. Primary tumor samples were collected from 62 patients with adenocarcinomas, squamous cell carcinoma (SCC), or large cell lung cancer that underwent surgical resection at the Central Hospital Gauting or the University Medical Center Hamburg-Eppendorf. This study received ethics review board approval and sample donors gave written informed consent. Clinical data are summarized in Table 1. The tumor samples were divided into two groups based on their bone marrow status: bone marrow-negative (n = 30) and bone marrow-positive (n = 32) samples. Additionally, a tissue microarray containing 36 brain metastases of patients with primary lung cancer was used for FISH analyses.

Bone marrow analysis. The procedures for isolation and immunocytochemical detection of tumor cells in the bone marrow have been described in detail (11). Briefly, bone marrow was aspirated from the upper iliac crest and the isolated mononuclear cells were

Table 1. Patient characteristics

Patient characteristics	CGH (n = 30)		Affymetrix (n = 16)		FISH* (n = 43)		Brain metastases (n = 36)
	Bone marrow negative	Bone marrow positive	Bone marrow negative	Bone marrow positive	Bone marrow negative	Bone marrow positive	
Histology †							
Adenocarcinoma	16	14	9	7	14	8	14
SCC	0	0	0	0	6	14	10
Large cell lung cancer	0	0	0	0	1	0	0
SCLC	0	0	0	0	0	0	7
Gender							
Female	7	6	5	2	8	2	14
Male	9	8	4	5	13	20	22
Age (y)							
Range	49-78	51-75	49-76	55-75	37-78	48-81	44-75
Mean	61	64	65	67	64	62	57
Tumor status							
pT_1	6	4	5	1	8	5	
pT_2	9	10	3	5	11	11	
pT_3	1	0	1	1	1	5	
pT_4	0	0	0	0	1	1	
Lymph node status							
Lymph node = 0	8	6	3	2	10	10	
Lymph node > 0	8	8	6	5	11	12	
Metastatic status							
M_0	16	14	9	7	21	22	
M_1	0	0	0	0	0	0	

*Including 14 adenocarcinomas used in CGH analysis.
†Histology is missing for 5 brain metastases.

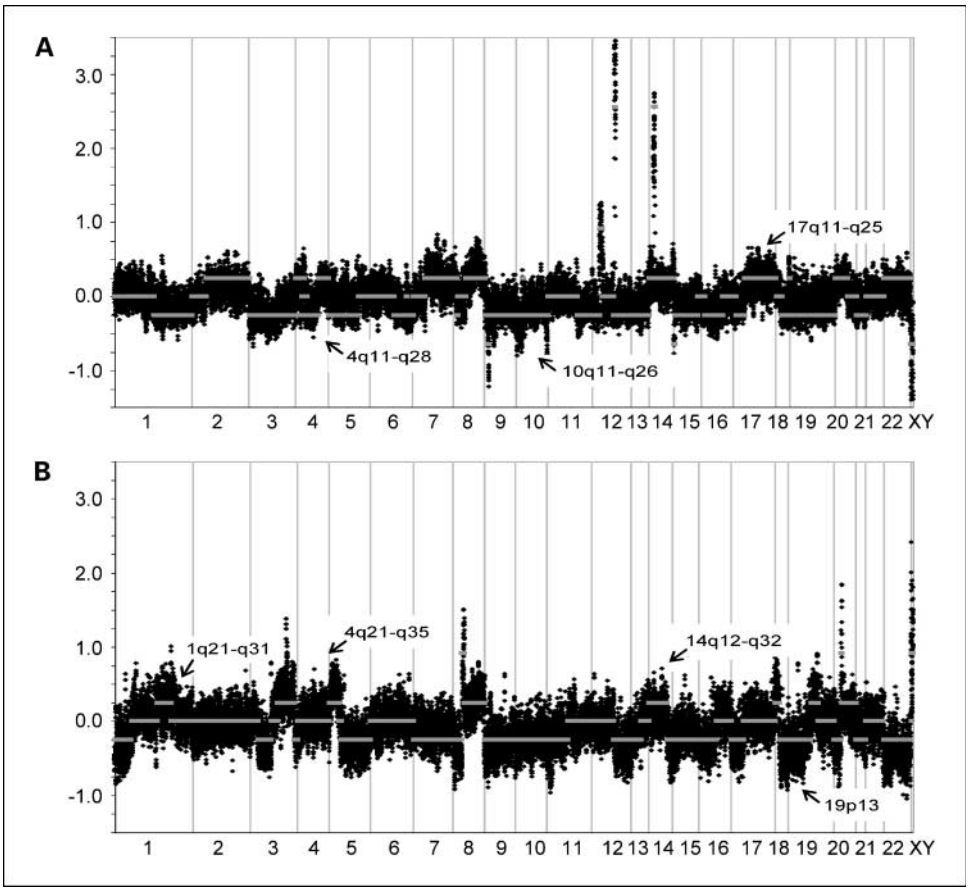

Fig. 1. Two representative array CGH copy number profiles from tumors derived from a patient with (*A*) and without (*B*) DTCs in the bone marrow. Log_2 ratios (tumor/reference using a moving average ratio of 6 adjacent clones; *Y axis*) are plotted against the chromosome number (*X axis*). *Gray lines,* gains (smooth value = 0.25), amplifications (smooth values = 0.90 and 2.50), and losses (smooth value = -0.25) calculated by aCGH-Smooth with default settings. *Arrows,* aberrations specific for the two groups; *horizontal lines,* chromosomal borders.

cytocentrifuged onto glass slides (2×10^6 per patient). DTCs were detected by immunocytochemical staining using the monoclonal antibody A45-B/B3 (Micromet). In all experiments, an isotype-matched, murine monoclonal antibody (MOPC 21, IgG1; Sigma-Aldrich) served as negative control. Screening for cytokeratin-positive cells was done in an automated fashion (ACIS System), which uses color-based imaging technology and microscopy to perform automated scanning and analysis of immunohistochemically stained slides (12).

Tumor staining and nucleic acid extraction. Fresh cryosections were mounted either on a noncoated slide or for laser pressure catapulting on a polyethylene membrane-coated slide (P.A.L.M). Either manual dissection or laser pressure catapulting using the Laser Microbeam System (P.A.L.M.) was used to obtain a tumor cell content of at least 70%.

Sections were stained with 1% toluidine blue O (Sigma-Aldrich) and 0.2% methylene blue (Fluka). Nucleic extraction was done using the RNeasy Micro Kit (Qiagen) with a DNase I treatment or the QIAmp DNA Micro Kit (Qiagen), respectively.

Array CGH. Array CGH on 30 lung adenocarcinomas was done as described (data available at GEO accession no. GSE13191; ref. 13). Briefly, 300 ng tumor DNA and 300 ng reference DNA (pooled leukocyte DNA of 10 healthy females or males) were labeled by random priming with Cy3- and Cy5-labeled dCTP (Applied Biosystems), respectively, and hybridized on oligo array CGH containing 28,830 oligonucleotides (13). To control the quality of hybridization, we performed a sex-mismatch, that is, DNA of opposite gender was used as reference. Subsequently, the X and Y chromosomes were excluded for downstream data analysis. The raw signal intensities were obtained with Bluefuse (13). The unreliable and morphologically corrupted measurement spots were removed from the data (14). The clones were first ordered based on their physical location. Thereafter, the data from each patient were scaled with global normalization to obtain equal median and variance for the middle 80 percentiles of autosomes to avoid the effect of extreme values. A Gaussian filter with the SD equal to 150,000 bp was applied to filter the data.

The differential region finding analysis method was used to identify chromosomal regions associated with the presence of DTCs in the bone marrow (15). In this nonparametric method, the array CGH data were analyzed at group level detecting the differences in DNA copy numbers between bone marrow-positive and bone marrow-negative patients. The chromosomes were divided in overlapping segments of 0.5 to 1.0 Mbp and each segment was tested for differences in copy numbers between the two groups. As a test statistic, the average number of correctly classified patients by the clones of each segment was used. The areas harboring bone marrow status-associated copy number differences were identified by means of hypothesis testing using 10,000 permutations to form the empirical probability distribution for the test statistic. To determine the *P* value of the segments, the proportion of permuted values at least as significant as the observed test statistic was calculated. The regions for which *P* values < 0.05 were obtained were assumed to be potentially associated with the bone marrow status. The regions closer than 10 Mbp were combined and the calculation of the *P* value was repeated to gain statistical power.

aCGH-Smooth was used to define losses and gains (16). The zeroline was determined based on the performed sex mismatch. In cases with dubious zerolines, FISH validation was used with different centromere probes (CEP 10, 15, and 17; Vysis).

Gene expression profiling. Gene expression arrays were done on 16 lung adenocarcinomas using the Affymetrix HG U133 Plus 2.0 GeneChips (data available at GEO accession no. GSE10799) according to MIAME standards. Thirteen of these samples also were analyzed by array CGH.

cDNA synthesis using 50 ng total RNA (Two-Cycle Target Labeling Kit; Affymetrix) was carried out according to the GeneChip Expression

Table 2. Chromosomal regions differing in copy number and expression between patients with and without DTCs in their bone marrows achieved with array CGH and gene expression profiles

Chromosomal region	Chromosomal position (bp)*		Size (Mbp)	Type of aberration	No. patients with aberration †				No. significant differentially expressed genes (all genes) ‡
	Start	Stop			Bone marrow negative (n = 16)		Bone marrow positive (n = 14)		
					Gain (%)	Loss (%)	Gain (%)	Loss (%)	
4q12-q32	53303625	160419296	107.1	Loss in bone marrow-positive Gain in bone marrow-negative	63	13	14	29	73 (361)
10p12-p11	21110362	33526584	12.4	Loss in bone marrow-positive Gain in bone marrow-negative	38	13	14	29	10 (51)
10q21-q22	69852278	81513171	11.7	Loss in bone marrow-positive Gain in bone marrow-negative	25	19	7	36	5 (86)
17q21	35426343	46398134	11.0	Gain in bone marrow-positive	31	13	36	0	5 (205)
20q11-q13	30571212	56368778	25.8	Gain in bone marrow-positive	38	13	64	0	11 (209)

*Base pair obtained by blasting array probe sequence in the University of California at Santa Cruz Blat.
† Defining gains and losses by using aCGH-Smooth; aberration < 5 Mbp were excluded.
‡ Number of genes within the region (present on the array) with different expression based on DNA copy number alterations. Significance of genes is based on ROC analysis (ROC value < 0.2 or > 0.2 and P < 0.05).

Analysis Technical Manual. The gene expression array data were preprocessed using the GC robust multi-array average method (15). The data were then $\log_2$-transformed and the median signal of the three normal patients was used as reference.

To detect chromosomal regions harboring gene expression changes associated with the presence of DTCs, we analyzed the data using the same region finding method (differential region finding) that was used for the array CGH. In addition, receiver operating characteristic (ROC) curves were used to identify individual genes that were differentially expressed between bone marrow-positive and bone marrow-negative patients independent of chromosomal localization (17, 18).

The area under the curve (ROC) was calculated to assess the diagnostic value of each gene. This area corresponds to the probability of correctly diagnosing the status of a patient knowing the value of the gene expression. This measure gives an estimate of the diagnostic value of the individual genes. Additionally, empirical P values corresponding to the ROC values were obtained by means of permutation testing using 10,000 permutations.

To identify regions with both expression and copy number changes associated with the presence of DTCs in the bone marrow, the regions detected using array CGH and gene expression data sets were compared. The common regions of the two analyses were then detected. To test whether these regions contained more differentially expressed genes than what is detected on average, the genes of the region with P values < 0.05 were counted. The significance of the finding was assessed by taking 10,000 samples of regions with equal amount of genes as located on the test region and determining the number of genes with P values < 0.05. The P value describing the significance of the finding was estimated by determining the proportion of the regions that contained at least as many differentially expressed genes.

The Cluster 3.0 program was used for supervised hierarchical clustering the differently expressed probe-set IDs and genes between the two groups using the complete linkage method and the Euclidian distance (19). For clustering the 104 significant genes within the five chromosomal regions, the mean was calculated of the significant probe-set IDs corresponding to the same gene. Classification of these genes in functional categories was done by the functional annotation chart for biological processes of the database DAVID (20).[7]

FISH. FISH analyses detecting copy number changes at 4q21 were done on 14 adenocarcinomas already used in the CGH analysis as well as 29 other fresh frozen tissue samples including 8 adenocarcinomas, 20 SCC, and 1 large cell lung cancer (Table 1).

For estimating DNA copy number alterations in metastatic tumors, a tissue microarray containing 36 brain metastasis (triplicate spots) from primary lung cancer patients was also analyzed, including 14 adenocarcinomas, 10 SCC, and 7 SCLC samples.

The BAC probe for 4q21 (Sanger clone IDs: RP11-570L13) was obtained from the German Resource Center for Genome Research. The BAC DNA was isolated using the Large Construct Kit (Qiagen) and labeled by random priming with dUTP labeled with spectrum orange using the BioPrime Labeling System (Invitrogen; ref. 21). Centromere probes (CEP 15 and 17) were used as reference (Vysis).

FISH analyses were done as described before (21). On average, 116 tumor cells of each sample were analyzed to create the average ratio. Additionally, 100 normal cells of six samples each were counted as negative control to measure the experimental bias and define the ratio cutoffs for copy number changes. Specimens containing a signal-to-centromere ratio $\geq$ 1.5 were considered to carry a gain and a ratio of <0.75 carry a loss.

Results

DNA copy number profiles in primary lung adenocarcinomas. The DNA copy number profiles of 30 lung adenocarcinomas were analyzed on a 30K high-resolution oligo array CGH. In general, all of the analyzed tumor samples showed changes

[7] http://david.abcc.ncifcrf.gov

in the copy number of DNA sequences, including common aberrations for lung cancer (22). In total, 27 different chromosomal regions harbored gains or losses in >30% of all cases irrespective of bone marrow status. The most common gains in all patients were 5pter-q11 (60%), 1q43-qter (40%), 11pter-p15.5 (39%), 8p12-p11.21 (39%), 8q12.1-qter (36%), 20pter-q11.21 (36%), 6p22.1-p21.31 (36%), and 7p13-q11.21 (36%). The most frequent losses were seen at 13q14.11-q22.1 (50%), 15qcen-q13.3 (43%), 9p13.1-q21.11 (42%), and 13q32.3-33.3 (33%).

DNA copy number aberrations associated with DTCs in the bone marrow. Patients were separated into two groups based on their bone marrow status to find distinct patterns of aberrations associated with early occult hematogenous dissemination as indicated by the presence of DTCs in the bone marrow.

Two methods were applied to identify aberrations associated with the bone marrow status. The bioinformatic tool aCGH-Smooth was used to detect potential breakpoints and define gains and losses (16). Additionally, differential region finding, which aims to identify regions harboring copy number differences at group level between two conditions, was used to detect DTC-related loci. Previously, this method has been successfully used to identify specific regions associated with asbestos-related lung cancer (15, 23).

In general, no differences could be found in the total number of aberrant clones, neither between bone marrow-positive and bone marrow-negative cases nor between lymph node-positive and lymph node-negative cases (Supplementary Table S1). However, we could detect 32 regions spanning 1.2 to 107 Mbp (median size, 2.45 Mbp) in which the DNA copy number between bone marrow-positive and bone marrow-negative groups differed ($P < 0.05$; Supplementary Table S2). All aberrations were single-allele gains or losses; thus, no high copy number aberrations were found specific for DTC status. Overall, eight chromosomal regions were found more commonly gained and three were lost in the bone marrow-positive group. None of the differentiating regions were exclusively altered in only one group.

The most significant finding was the aberration at 4q spanning a region with 107 Mbp. The chromosomal region 4q12-q32 was more often gained in bone marrow-negative (63%) and lost in bone marrow-positive cases (29%). The whole 4q-arm was lost in bone marrow-positive cases in two cases and gained in bone marrow-negative cases in three cases. Figure 1 illustrates two array CGH plots for one bone marrow-positive case and one bone marrow-negative case with aberrations representative of the bone marrow groups.

Expression signatures associated with genomic aberrations. To validate and increase the specificity of the identified genomic differences between the two groups, Affymetrix expression arrays were done. Differentially expressed genes between bone marrow-positive and bone marrow-negative groups were identified, and results were compared with regions harboring copy number differences between the two groups. Analysis was done with two different independent strategies. Differential region finding was used for the expression data in the same way as for array CGH. In addition, ROC analyses, which do not take chromosome position into account, were done to find possible target genes (17, 18).

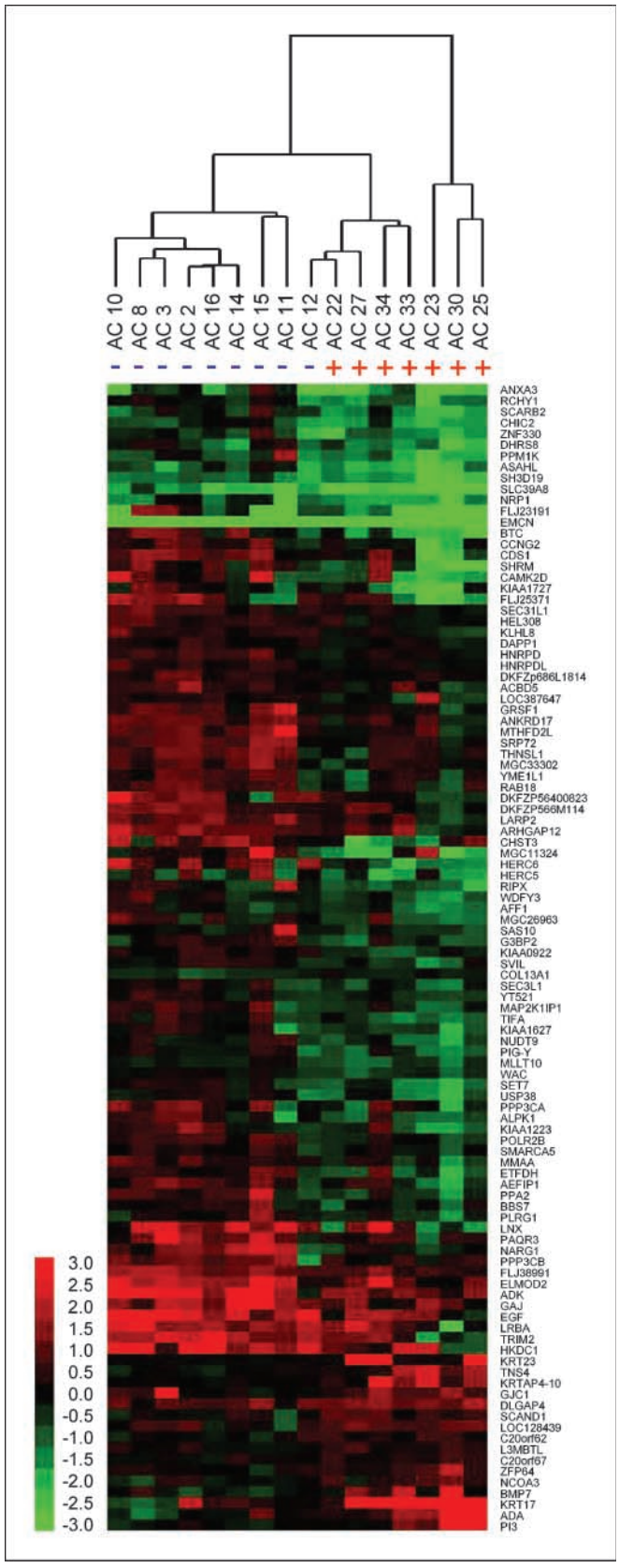

Fig. 2. Hierarchical clustering of the mean of the probe-set IDs corresponding to the 104 significant genes found in the five significant chromosomal regions separating the bone marrow-positive patients (+, AC22-AC34) from the bone marrow-negative patients (-, AC2-AC16).

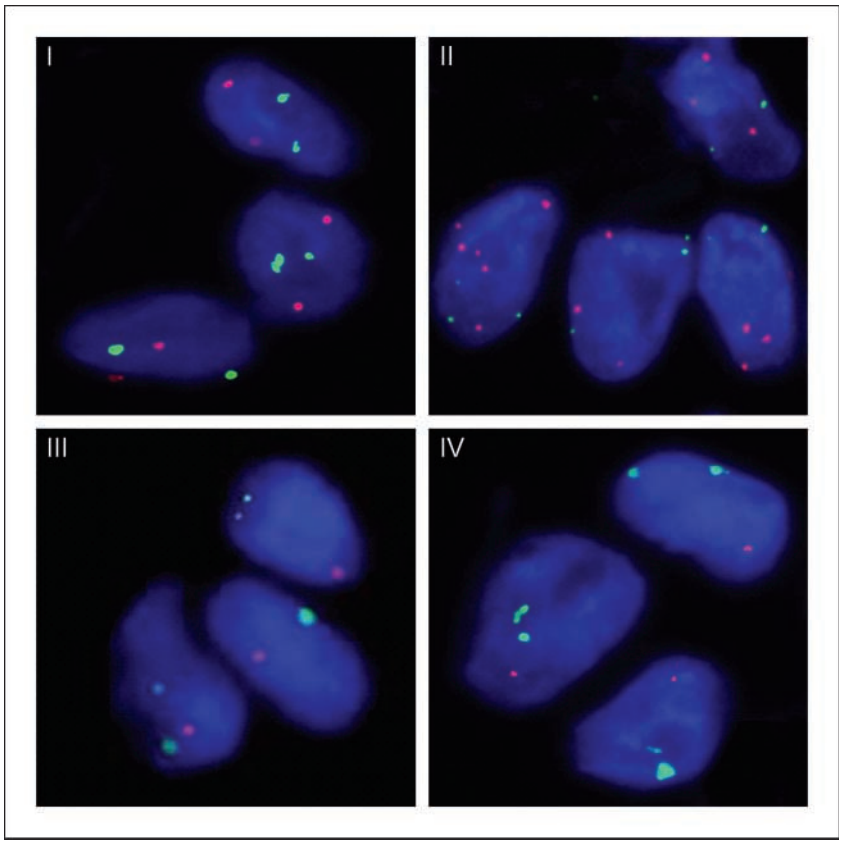

Fig. 3. Genomic aberrations of 4q21 in lung adenocarcinomas detected by FISH. *I*, tumor cells of a bone marrow-positive sample (AC26) with normal copy numbers of reference (centromere probe 15, spectrum aqua displayed in the pseudo-color green) and spectrum orange-labeled 4q21. *II*, tumor cells of a bone marrow-negative sample (AC7) with a low-level gain of 4q21. *III* and *IV*, tumor cells of bone marrow-positive tumors (AC30 and AC23) with losses of 4q21.

Overall, 5.2% of the Affymetrix probe-set IDs (n = 1733), corresponding to 1,491 different genes, were found to be significantly differentially expressed (ROC < 0.2 or > 0.8 and P < 0.05). Hierarchical clustering of these clones divided the patients into two separate groups according to their bone marrow status (Supplementary Fig. S1). Using gene ontology classification, we detected 19 subcategories (P < 0.05, 15 biological processes and 4 molecular functions), which were enriched in these 100 most significant genes compared with all genes on the array (Supplementary Table S3). These 19 subcategories could be summarized under three main functional categories: general regulation of cellular processes (n = 4), metabolism (n = 7), and transcription (n = 8).

Five chromosomal regions consistently showed an altered gene expression pattern. Table 2 shows the five regions at chromosome 4q, 10p, 10q, 17q, and 20q found to have both DNA copy number alterations combined with gene expression changes. All these regions, except 20q, were specific for the bone marrow status and not associated with the lymph node status (data not shown). Interestingly, 20q was also significantly associated with lymph node positivity and showed gains especially in double-positive patients (lymph node-positive and bone marrow-positive).

Differentially expressed probe-set IDs were found either scattered over the chromosomal regions of 10q21-q22 and 20q11-q13 or clustered together into some "hot spots" in 4q, 17q, and 10p. In these five significant chromosomal regions, 104 genes were found to be differentially expressed according to the ROC analyses (Supplementary Table S4). This number was irrespective of chromosomal size significantly higher on average than in the other regions (P < 0.006). Figure 2 shows the hierarchical clustering results of these 104 genes. The genes clustered into three separate branches with one branch containing all except one bone marrow-negative patient.

The five main functional categories that were enriched in these 104 genes compared with all genes on the array were general enzyme activation (n = 58), biopolymer metabolism (n = 27), cell organization and biogenesis (n = 16), chromosome organization and biogenesis (n = 11), and positive regulation of cell proliferation (n = 4; Supplementary Table S3).

FISH analysis of 4q. To initially confirm the array CGH results, FISH analysis was done for 4q21 on fresh frozen tissue from available tumor material used in the array experiments. Figure 3 shows representative FISH results for decreased, increased, and normal copy numbers at 4q21 in lung cancer tissues. All changes could be detected, except one loss, perhaps due to the heterogeneous nature of the tumor tissue.

For verification of the association between loss of 4q and the presence of DTCs in the bone marrow, FISH analysis for 4q21 was done on 29 additional, fresh frozen primary lung cancer tissues. Similar to the initial sample set, loss of 4q21 was significantly more frequent in bone marrow-positive patients (50%) compared with bone marrow-negative patients (6%; P = 0.039). Only one single gain was detected within the bone marrow-positive group. Combination of the FISH and the array CGH results for copy number changes in 4q21 for all 59 patients showed a significant association between losses of 4q and the presence of DTC in the bone marrow (37% versus 7%; P = 0.011; Table 3). In univariate analysis (Fisher's exact test), bone marrow status, tumor stage, and histology were associated with 4q loss. In multivariate analysis (binary logistical

regression), bone marrow status (P = 0.016) and tumor stage (P = 0.037), however, remained as only independent factors associated with loss of 4q. Using different centromere probes, the ploidy of the tumor cells was determined. Although there was no significant difference in the ploidy between the two groups, tumors of bone marrow-positive patients showed a higher tendency for polyploidy (35% versus 23%).

Additionally, FISH analyses for 4q21 were done using a tissue microarray with 36 brain metastases of the lung. Thirty-nine percent of the samples showed one allele loss of 4q, whereas gains were only found in 6% of the patients. Similar to the bone marrow-positive primary tumors, 4q loss was more commonly found in SCC (brain, 50%; primary tumors, 57%) than in adenocarcinomas (36% and 19%). Univariate analysis showed no significant association among age, gender, and 4q loss. Brain metastases showed significant more losses of 4q21 compared with primary lung tumors overall (P = 0.024).

Discussion

Cancer development is known as a multistep genomic process including accumulation of numerous genetic alterations leading to an activation of oncogenes and inactivation of tumor suppressor genes (24, 25). The present array CGH and expression results showed very complex DNA aberration and altered expressional patterns, which are typical for even quite early-stage lung tumors (26, 27). Nevertheless, by combining the expression array with the array CGH data, we could identify five chromosomal regions associated with the DTC status. Two regions were associated with gains (17q and 20q) and three with losses (4q, 10p and 10q) in the bone marrow-positive group. Gains in chromosomes 4 and 10 were simultaneously found in bone marrow-negative patients.

Genomic instability, in which an increased rate of genomic aberrations permits the accumulation of genetic events resulting often in aneuploidy, was recently argued to be "the seventh hallmark of cancer" (28). The level of aneuploidy can be furthermore translated to an expressional gene set (signature of chromosomal instability) that has a very high predictive power for clinical outcome for several types of cancer, including lung, breast, and colon (29). Also, our study indicated a linkage of aneuploidy and thereby chromosomal instability with the micrometastatic status of a patient. In addition, one of the main functional categories of the 104 genes separating the bone marrow-positive from the bone marrow-negative patients was chromosome organization and biogenesis, showing the association between chromosomal instability and specific patterns of chromosomal changes for early hematogenous dissemination of tumor cells.

The loss of chromosome 4q12-q32 in bone marrow-positive patients and the gain in bone marrow-negative patients were clearly the most prominent findings. Therefore, for validation of these array results and as an independent measurement technique, FISH analyses were done showing a highly significant association between 4q loss and a bone marrow-positive status (P = 0.039). The combined results from array CGH and FISH showed loss of a single allele of 4q21 in 37% of bone marrow-positive compared with 7% of bone marrow-negative samples (P = 0.011), indicating an important and specific role of 4q in the hematogenous spread of tumors independent of lymphatic dissemination.

In support of our results, a previous study on primary lung adenocarcinomas showed an association between loss of 4q

Table 3. Summary of 4q aberrations in primary lung carcinomas

	Primary tumor			Univariate P*	Multivariate P †
	4q loss	4q gain	4q normal		
Bone marrow status					
Positive	11	2	17	0.011	0.016
Negative	2	5	22		
Lymph node status					
N_0	6	4	19	1.000	NS
$>N_1$	7	3	20		
Histology					
All	13	7	39	0.007	NS
Adenocarcinoma	4	6	28		
SCC	9	1	10		
Large cell lung cancer	0	0	1		
Gender					
Female	1	4	12	0.082	NS
Male	12	3	27		
Tumor stage					
$T_1 + T_2$	8	6	36	0.020	0.037
$T_3 + T_4$	5	1	3		
Grade					
1-2	6	3	22	0.748	NS
3	6	3	15		
Polyploidy					
2n	7	4	20	0.295	NS
>3n	5	0	6		

*P value calculated for numbers of 4q losses versus others using Fisher's exact test.
† Binary logistical regression using bone marrow status, lymph node status, histology, gender, and tumor stage.

and distant metastases and gain with the nonmetastatic phenotype (30). In addition, in primary hepatocellular, oral SCC, and brain metastases, deletion of 4q has been found to be significantly associated with progression (31–33), making this locus highly interesting with respect to tumor dissemination.

We detected loss of 4q in brain metastases (39%) from lung cancer patients at a frequency similar to that of bone marrow-positive primary tumors (37%), whereas gains were seldom found that were typical for bone marrow-negative group (6%). We found that, in both primary tumors and brain metastases, loss of 4q was not only restricted to adenocarcinomas but also found with a higher frequency in SCC and SCLC, indicating an important role of 4q in both NSCLC and SCLC. This and previous findings thus indicate that loss of 4q might not only determine early steps of the metastatic cascade of all types of lung cancer but also seems to influence the outgrowth of distant metastases, including those to the brain.

Region 4q was the largest chromosomal region found containing the highest number of differentially expressed genes irrespective of chromosomal size ($P = 0.020$). The differentially expressed genes clustered within this region into five separate hot spots, indicating the existence of more than one putative tumor suppressor gene. Also in SCLC, esophageal adenocarcinoma, and malignant mesothelioma, several separate small deletions have been described (34–37), although the target genes remain unidentified. In SCLC, a homozygous deletion was detected at 4q21 containing the *MAPK10* gene. However, in our NSCLC data set, this gene was neither down-regulated nor homozygously deleted (37).

Altogether, 104 genes were found significantly associated with the bone marrow status in the five chromosomal regions. Not surprisingly, clustering of these 104 genes showed an almost complete separation into two groups. However, none of the 104 genes differentiating the bone marrow-positive from the bone marrow-negative group has been described as a metastatic suppressor gene (38, 39). Our micrometastatic predictor gene set did not overlap with the previously identified gene set associated with micrometastatic breast cancer, indicating either a specific metastatic signature for each tumor type or merely the very complex nature of the metastatic cascade (6, 7). Therefore, our finding of specific chromosomal patterns is an additional step in understanding the highly complex processes related to metastasis and might help identifying new important target genes.

In summary, this is the first study showing a specific pattern of gene expression and DNA alterations that are associated with the early hematogenous dissemination of lung tumor cells. Our findings indicate that early hematogenous dissemination of tumor cells appears to be a specific process driven by a set of molecular changes. Especially the large deletion on chromosome 4q in bone marrow-positive patients as well as in brain metastases suggests that important metastasis suppressing genes are located in this region. Our results thus provide new insights into the genetics of lung tumor dissemination and may promote future studies identifying the genes underlying early dissemination and growth of lung tumor cells at distant sites. However, the clinical relevance of these aberrations remains to be further elucidated in further studies before declaring them as prognostic markers for lung tumors that might spread mainly to the bone marrow.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Malgorzata Stoupiec, Kathrin Eylmann, Jolanta Kropidlowski, Kristin Reeck, Petra Merkert, and Drs. Serge Smeets and Paul van den Ijssel for excellent technical assistance, Dr. Natalie Reimers for helpful discussion comments, and Dr. Alexander Nakeff (Henry Ford Hospital) for critical reading of the article.

References

1. Beasley MB, Brambilla E, Travis WD. The 2004 World Health Organization classification of lung tumors. Semin Roentgenol 2005;40:90–7.
2. Pantel K, Izbicki J, Passlick B, et al. Frequency and prognostic significance of isolated tumour cells in bone marrow of patients with non-small-cell lung cancer without overt metastases. Lancet 1996;347:649–53.
3. Pantel K, Brakenhoff RH. Dissecting the metastatic cascade. Nat Rev Cancer 2004;4:448–56.
4. Cote RJ, Beattie EJ, Chaiwun B, et al. Detection of occult bone marrow micrometastases in patients with operable lung carcinoma. Ann Surg 1995;222:415–23; discussion 23–5.
5. Kubuschok B, Passlick B, Izbicki JR, Thetter O, Pantel K. Disseminated tumor cells in lymph nodes as a determinant for survival in surgically resected non-small-cell lung cancer. J Clin Oncol 1999;17:19–24.
6. Naume B, Zhao X, Synnestvedt M, et al. Presence of bone marrow micrometastasis is associated with different recurrence risk within molecular subtypes of breast cancer. Mol Oncol 2007;1:160–71.
7. Woelfle U, Cloos J, Sauter G, et al. Molecular signature associated with bone marrow micrometastasis in human breast cancer. Cancer Res 2003;63:5679–84.
8. Bubendorf L, Muller P, Joos L, et al. Multitarget FISH analysis in the diagnosis of lung cancer. Am J Clin Pathol 2005;123:516–23.
9. Halling KC, Rickman OB, Kipp BR, Harwood AR, Doerr CH, Jett JR. A comparison of cytology and fluorescence *in situ* hybridization for the detection of lung cancer in bronchoscopic specimens. Chest 2006;130:694–701.
10. Braun S, Pantel K, Muller P, et al. Cytokeratin-positive cells in the bone marrow and survival of patients with stage I, II, or III breast cancer. N Engl J Med 2000;342:525–33.
11. Pantel K, Schlimok G, Angstwurm M, et al. Methodological analysis of immunocytochemical screening for disseminated epithelial tumor cells in bone marrow. J Hematother 1994;3:165–73.
12. Bauer KD, de la Torre-Bueno J, Diel IJ, et al. Reliable and sensitive analysis of occult bone marrow metastases using automated cellular imaging. Clin Cancer Res 2000;6:3552–9.
13. van den Ijssel P, Tijssen M, Chin SF, et al. Human and mouse oligonucleotide-based array CGH. Nucleic Acids Res 2005;33:e192.
14. Ruosaari S, Hollmén J. Image analysis for detecting faulty spots from microarray images. Lecture Notes Comput Sci 2002;2534:259–66.
15. Nymark P, Wikman H, Ruosaari S, et al. Identification of specific gene copy number changes in asbestos-related lung cancer. Cancer Res 2006;66:5737–43.
16. Jong K, Marchiori E, Meijer G, Vaart AV, Ylstra B. Breakpoint identification and smoothing of array comparative genomic hybridization data. Bioinformatics 2004;20:3636–7.
17. Kettunen E, Anttila S, Seppanen JK, et al. Differentially expressed genes in nonsmall cell lung cancer: expression profiling of cancer-related genes in squamous cell lung cancer. Cancer Genet Cytogenet 2004;149:98–106.
18. Swets JA. Measuring the accuracy of diagnostic systems. Science 1988;240:1285–93.
19. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863–8.
20. Dennis G, Jr., Sherman BT, Hosack DA, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003;4:P3.
21. Wikman H, Nymark P, Vayrynen A, et al. CDK4 is a probable target gene in a novel amplicon at 12q13.3-q14.1 in lung cancer. Genes Chromosomes Cancer 2005;42:193–9.
22. Balsara BR, Testa JR. Chromosomal imbalances in human lung cancer. Oncogene 2002;21:6877–83.
23. Wikman H, Ruosaari S, Nymark P, et al. Gene expression and copy number profiling suggests the importance of allelic imbalance in 19p in asbestos-associated lung cancer. Oncogene 2007;26:4730–37.
24. Albertson DG, Pinkel D. Genomic microarrays in

human genetic disease and cancer. Hum Mol Genet 2003;12:145R–52.

25. Knuutila S, Autio K, Aalto Y. Online access to CGH data of DNA sequence copy number changes. Am J Pathol 2000;157:689.

26. Sy SM-H, Wong N, Lee T-W, et al. Distinct patterns of genetic alterations in adenocarcinoma and squamous cell carcinoma of the lung. Eur J Cancer 2004; 40:1082–94.

27. Yakut T, Schulten HJ, Demir A, et al. Assessment of molecular events in squamous and non-squamous cell lung carcinoma. Lung Cancer 2006;54:293–301.

28. Pollack JR. Chromosome instability leaves its mark. Nat Genet 2006;38:973–4.

29. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet 2006;38: 1043–8.

30. Goeze A, Schluns K, Wolf G, Thasler Z, Petersen S, Petersen I. Chromosomal imbalances of primary and metastatic lung adenocarcinomas. J Pathol 2002; 196:8–16.

31. Bando K, Nagai H, Matsumoto S, et al. Identification of a 1-cM region of common deletion on 4q35 associated with progression of hepatocellular carcinoma. Genes Chromosomes Cancer 1999;25:284–9.

32. Lin SC, Chen YJ, Kao SY, et al. Chromosomal changes in betel-associated oral squamous cell carcinomas and their relationship to clinical parameters. Oral Oncol 2002;38:266–73.

33. Petersen I, Hidalgo A, Petersen S, et al. Chromosomal imbalances in brain metastases of solid tumors. Brain Pathol 2000;10:395–401.

34. Cho ES, Chang J, Chung KY, Shin DH, Kim YS, Kim SK. Identification of tumor suppressor loci on the long arm of chromosome 4 in primary small cell lung cancers. Yonsei Med J 2002;43:145–51.

35. Shivapurkar N, Virmani AK, Wistuba II, et al. Deletions of chromosome 4 at multiple sites are frequent in malignant mesothelioma and small cell lung carcinoma. Clin Cancer Res 1999;5:17–23.

36. Sterian A, Kan T, Berki AT, et al. Mutational and LOH analyses of the chromosome 4q region in esophageal adenocarcinoma. Oncology 2006;70:168–72.

37. Kim YH, Girard L, Giacomini CP, et al. Combined microarray analysis of small cell lung cancer reveals altered apoptotic balance and distinct expression signatures of MYC family gene amplification. Oncogene 2006;25:130–8.

38. Vaidya KS, Welch DR. Metastasis suppressors and their roles in breast carcinoma. J Mammary Gland Biol Neoplasia 2007;12:175–90.

39. Berger JC, Vander Griend DJ, Robinson VL, Hickson JA, Rinker-Schaeffer CW. Metastasis suppressor genes: from gene identification to protein function and regulation. Cancer Biol Ther 2005;4:805–12.