S100 Family Members and Trypsinogens Are Predictors of Distant Metastasis and Survival in Early-Stage Non-Small Cell Lung Cancer

Sven Diederichs,1 Etnar Bulk,1 Björn Steffen,1 Ping Ji,1 Lara Tickenbrock,1 Kerstin Lang,2 Kurt S. Zänker,2 Ralf Metzger,3 Paul M. Schneider,4 Volker Gerke,4 Michael Thomas,1 Wolfgang E. Berdel,1 Hubert Serve,1 and Carsten Müller-Tidow1

1Department of Medicine, Hematology/Oncology, University of Münster, Münster, Germany; 2Department of Immunology, University of Witten/Herdecke, Witten, Germany; 3Department of Visceral and Vascular Surgery, University of Cologne, Cologne, Germany; and 4Center for Molecular Biology of Inflammation, Institute for Medical Biochemistry, University of Münster, Münster, Germany

Abstract

Distant metastasis is the predominant cause of death in early-stage non-small cell lung cancer (NSCLC). Currently, it is impossible to predict the occurrence of metastasis at early stages and thereby separate patients who could be cured by surgical resection alone from patients who would benefit from additional chemotherapy. In this study, we applied a comparative microarray approach to identify gene expression differences between early-stage NSCLC patients whose cancer ultimately did or did not metastasize during the course of their disease. Transcriptional profiling of 82 microarrays from two patient groups revealed differential expression of several gene families including known predictors of metastasis (e.g., matrix metalloproteinases). In addition, we found S100P, S100A2, trypsinogen C (TRY6), and trypsinogen 1Bv (PRSS3) to be overexpressed in tumors that metastasized during the course of the disease. In a third group of 42 patients, we confirmed the induction of S100 proteins and trypsinogens in metastasizing tumors and its significant correlation with survival by real-time quantitative reverse transcription-PCR. Overexpression of S100A2, S100P, or PRSS3 in NSCLC cell cultures led to increased transendothelial migration, corroborating the role of S100A2, S100P, and PRSS3 in the metastatic process. Taken together, we provide evidence that expression of S100 proteins and trypsinogens is associated with metastasis and predicts survival in early stages of NSCLC. For the first time, this implicates a role of S100 proteins and trypsinogens in the metastatic process of early-stage NSCLC.

Introduction

Non-small cell lung cancer (NSCLC) is the leading cause of tumor-related death. Distant metastasis is the most frequent reason for NSCLC lethality. Despite recent progress, the molecular mechanisms underlying metastasis have not been solved in detail. Several investigators have analyzed the involvement of individual genes (1, 2), gene expression differences in cell lines (3), and gene expression differences between primary tumors and normal lung tissue (4). These studies help to decipher biological differences in these systems but do not aid in the definition of prognostic parameters at diagnosis. Only gene expression differences in primary tumors before the onset of metastasis will allow the development of molecular biology-based therapy decisions (5). Sets of genes predictive for survival prognosis of NSCLC patients have been defined, based on microarray experiments (6). For example, microarray analyses uncovered a histologically nontidistinct tumor subgroup associated with refractory disease (7, 8).

In addition, gene expression changes associate with disease-free survival (6, 9), and transcriptional profiling studies have distinguished novel histological NSCLC subtypes (10). However, a distinct study comparing expression profiles of primary tumors at the stage of diagnosis and identifying patterns underlying metastasis is still lacking. These data could reveal deeper insights into the molecular biology of the metastatic process and would have clinical implications. It could be possible to identify patients who would benefit from adjuvant chemotherapy after resection of the primary tumor. We performed transcriptional profiling using oligonucleotide arrays of early-stage NSCLC tumors that did or did not lead to distant metastasis after prolonged follow-up. We identified several groups of genes including S100 proteins and trypsinogens that predict metastasis and survival at the time of diagnosis and are involved in the metastatic process.

Materials and Methods

Gene Expression Analysis. Primary tumor specimens were obtained at the time of initial surgery for early-stage NSCLC (2). Patient data have been published previously (5, 7). Oligonucleotide microarray (Affymetrix) hybridizations were carried out as described previously (11, 12). We hybridized cRNA obtained from 14 patients (15 arrays) with stage I or II NSCLC (adenocarcinoma and squamous cell carcinoma). In addition, we used published primary data from a second group of 52 patients (67 arrays; Ref. 7). Regulation (Fig. 1A) was calculated as the mean of the regulation of the average and of the median expression levels in both groups. In addition, S100A2, S100P, PRSS3, and TRY6 mRNA expression levels were analyzed by quantitative real-time reverse transcription-PCR (RT-PCR) (Ref. 5; see Supplementary Fig. 1 for sequences)]. Statistical analyses were carried out using SPSS 11.0. All tests were two-sided, with P = 0.05 regarded as significant.

Migration Analysis. Coding sequences of S100P, S100A2, PRSS3, and TRY6 were cloned into expression vector pCDNA3.1(+) harboring enhanced green fluorescent protein (EGFP) and stably transfected into HTB-58 (SK-MES-1). The use of bulk cultures avoided clone-specific effects. Expression was verified by quantitative RT-PCR and Western blotting [anti-EGFP (Clontech) and anti-actin (Sigma)] as well as fluorescence-activated cell-sorting (FACS) analysis. For migration analysis, 5 × 105 cells were plated into Transwell plates precoated with fibronectin. Migration was measured by counting the number of cells that migrated into the lower chamber after 24 h by FACS. Transendothelial migration (13) was detected in the same way, but 2.2 × 105 HMEC-1 endothelial cells were plated into the Transwell plates 48 h before seeding the transfected NSCLC cell cultures. Further information can be found in the supplementary data.

Results

Identification of Metastasis-Associated Genes by Comparative Microarray Analysis. We used microarray analyses to identify genes predicting the likelihood of early-stage NSCLC tumors to metastasize. For this purpose, tumors resected from patients with stage I or II...
NSCLC were shock frozen, and RNA was isolated and hybridized onto HGU95Av2 Affymetrix chips containing more than 12,000 genes. The final data set contained 82 microarrays because we combined our hybridization data with published expression profiles not analyzed for metastasis-associated genes (7). Patients were followed over time for at least 36 months.

In our analyses, we identified 39 genes with differential expression in patients whose cancer metastasized during the course of the disease (YES) and patients whose cancer did not metastasize (NO). Only two genes (collagen COL9A3 and secreted frizzled-related protein SFRP1) were down-regulated in metastasizing tumors, whereas 37 genes were up-regulated in metastasizing tumors. We provide a gene list, in which we added the new UniGene nomenclature in addition to the accession numbers given by Affymetrix and substituted all expressed sequence tags by the recently identified genes (Fig. 1A). The identified genes belong to different functional groups e.g., proteases, calcium-binding S100 proteins, extracellular matrix (ECM) proteins, or metabolic enzymes.

Some of the genes identified in our screen have previously been associated with metastasis. The matrix metalloproteinases (MMPs) degrade ECM proteins, an essential step for migration and invasion. MMPs have been linked previously to metastasis in NSCLC (3). In our analyses, we isolated MT-MMP1 (MMP14) and stromelysin-2 (MMP10). In addition, members of the keratin family of intracellular IF proteins were associated with metastasis in NSCLC (14). In contrast, the ECM protein collagen was inhibitory for cell motility and migration (15). We discovered up-regulation of keratin 16 and keratin 14, as well as down-regulation of collagen 9A3. The chemokine interleukin 8 has been associated with metastasis and tumor angiogenesis (16) as well as the metabolic enzyme pyruvate kinase 3 (PK-M2; Ref. 17). We demonstrate significant association of MT-MMP1 (MMP14), keratin 16, pyruvate kinase 3 (PK-M2), and interleukin 8 with the occurrence of metastasis based on analysis of all microarrays (n = 82; Fig. 1C).

Association of S100 Proteins with Metastasis and Survival in NSCLC:

The family of calcium-binding S100 proteins fulfills a broad range of functions in calcium-dependent stimulus response coupling and has not been implicated in the metastatic process in NSCLC. Among the S100 proteins, S100A4 (Mts1) has been linked to metastasis in mammary tumors (18). However, S100A4 expression did not correlate with melanoma progression, in which S100A2 expression was lost (19), indicating distinct roles of the S100 proteins in malignant processes.

In our analyses, two of the regulated genes belonged to the family of S100 proteins (three oligonucleotide sets, S100A2 = S100L, S100P; Fig. 2A).
We confirmed differential expression by real-time quantitative RT-PCR in a group of 42 patients that had not been included in the microarray analyses. In this group, we found 3- to 7-fold induction of the median expression in metastasizing versus nonmetastasizing tumors and thereby verified the regulation of S100A2 and S100P with a different method (Fig. 2B).

The 42 patient samples analyzed by real-time quantitative RT-PCR were divided into high (3-fold above median; n/H11005 16) and low (n/H11005 26) S100P expressers. Kaplan-Meier analysis revealed a significant benefit in overall survival for the patients with low expression of S100P (Fig. 2C; P/H11005 0.046). The same trend was observed for S100A2 but was not statistically significant.

**Association of Trypsinogens with Metastasis and Survival in NSCLC.** The serine protease trypsin and its precursor trypsinogen have been linked to tumor progression by activation of MMPs in pancreatic, gastric or colorectal cancer (20). Our microarray analysis revealed up-regulation of two trypsinogens [trypsinogen IVb (PRSS3) and trypsinogen C (TRY6)] in the metastatic process of NSCLC (Fig. 3A). The induction of PRSS3 and TRY6 in metastasizing compared with nonmetastasizing NSCLC tumors was confirmed in the additional subset of patient samples (n/H11005 42) by real-time RT-PCR (Fig. 3B). In Kaplan-Meier analyses, patients with high expression of either PRSS3 or TRY6 suffered from a substantially worse prognosis than patients expressing low levels of trypsinogens (Fig. 3C).

**Stimulation of Migration in NSCLC Cell Cultures by S100 Proteins and Trypsinogens.** To analyze the role of S100 proteins and trypsinogens in the metastatic process, we studied their effect on cell migration and invasive migration through endothelial cell layers. We established HTB-58 NSCLC cell cultures stably expressing EGFP (control vector) or EGFP fusion constructs of S100P, S100A2, PRSS3, or TRY6. To exclude clone-specific effects, bulk cultures of selected and sorted cells were used after transfection. Overexpression was verified by FACS analysis, real-time RT-PCR, and Western blotting. All cell cultures displayed at least 90% EGFP positivity (Fig. 4A). Fluorescence microscopy confirmed distinct localization patterns of the fusion proteins (data not shown). Overexpression at the mRNA level varied between the different cell cultures from 2-fold to 4000-fold, probably reflecting differences in endogenous gene expression or mRNA stability (Fig. 4B). However, consistent expression of the EGFP fusion proteins was detected at the protein level by FACS analysis. Western blotting for EGFP demonstrated expression of the EGFP fusion proteins in stable cell cultures. A nonspecific band...
(Fig. 4C, control) was used as loading control. The trypsinogen fusion proteins closely co-migrated with this nonspecific band, making it difficult to clearly distinguish them from each other (Fig. 4C).

Migration assays in Transwell plates ($n = 9$) revealed significantly increased migratory activity in cells overexpressing S100P ($P = 0.026$, Mann-Whitney $U$ test) or S100A2 ($P = 0.001$). TRY6-overexpressing cells migrated less compared with control cells transfected with EGFP alone (Fig. 4D, top panel).

To study an experimental model simulating the in vivo situation of metastasis and invasion, we analyzed the migration through endothelial cell layers seeded into Transwell chambers ($n = 9$). When compared with EGFP-expressing controls, we detected a consistent and significant increase in spontaneous transendothelial migration of cell cultures stably overexpressing S100P, S100A2, PRSS3, and TRY6 ($P = 0.012$, $P = 0.001$, $P = 0.001$, and $P = 0.034$, respectively, Mann-Whitney $U$ test; Fig. 4D, bottom panel).

Discussion

In this study, we identified differences in the gene expression patterns between primary NSCLC tumors that did or did not lead to distant metastasis. We compared microrarrays hybridized in our laboratory with previously published data (7) and verified gene regulation with an alternative method in a third independent subset of patients. Therefore, our data reached a high level of validity since the important findings were confirmed in three independent patient groups.

The quality of our data is further corroborated by the frequent...
occurrence of genes in our study that are known to be involved in metastatic processes. The isolation of different members of one protein family in the list of regulated genes extends the probability of their importance in vivo in NSCLC metastasis. Our data are in accordance with most publications describing genes involved in metastasis because we found up-regulation of proteases that degrade extracellular matrix proteins, which is essential for migration and invasion. In addition, we detected proteins that were not yet linked to metastasis but to pathways assisting in this process. For example, we found up-regulation of inhibin A, a member of the tumor growth factor β family that is known to regulate the SMAD pathway (21) and to influence the expression of the known metastasis regulator plasminogen activator inhibitor-1 (22) or collagen, which was also downregulated.

Notably, we also provide evidence for an involvement of gene families not yet assigned to NSCLC metastasis: S100 proteins and trypsinogens. Their induction in metastasizing tumors was verified in three independent patient groups using two different methods to ensure the general importance of the regulated genes for the underlying process.

S100 proteins build a family of 20 calcium-binding EF-hand proteins so far studied mainly in the immune system. They regulate
in intracellular processes such as cell growth and motility, cell cycle, transcription, and differentiation. Individual members localize to specific cellular compartments and are able to relocate upon Ca\(^{2+}\) activation, transducing the Ca\(^{2+}\) signal by interacting with specific targets (23). The involvement of S100 proteins in migratory processes has not been recognized before in NSCLC but fits well into the general picture of calcium-dependent cell movement (16). S100A4 was the only family member described to participate in metastasis (18); therefore, our study for the first time provides evidence for an important role of S100A2 and S100P in metastasis and NSCLC survival.

Trypsins are serine proteases responsible for digestion in the duodenum and activated by autocatalytic cleavage from their trypsinogen precursors. They have been implicated in the progression of malignant diseases of the pancreas, the organ in which the trypsinogens are synthesized, and in the large bowel system (20). In the lung, trypsinogen has mainly been investigated with regard to inflammatory processes. Therefore, our results link trypsin activity for the first time to tumor progression and metastasis in organs unrelated to the digestive tract. Mechanistically, trypsins could assist in invasion by degrading proteins of the ECM.

In addition, stable overexpression of S100P, S100A2, PRSS3, and TRY6 in NSCLC cells revealed significantly enhanced transendothelial migration. Because only S100 proteins induced a migratory phenotype in the absence of endothelial cells, we assume that trypsinosgens are not of major importance for migration in general; rather, they are important for specific processes such as invasion or evasion from blood vessels and, by this mechanism, are linked to metastasis.

Our study establishes a basis for the prediction of metastatic events at the time of diagnosis derived from gene expression analysis of the primary tumor. Our list of new markers for the development of metastasis, combined with published knowledge, will help to improve therapeutic decisions for early-stage NSCLC patients and create the future opportunity to apply adjuvant chemotherapy specifically to patients with the highest expected benefit from additional treatment.

Taken together, we have identified S100 proteins, trypsinosgens, and more than 30 other genes associated with metastasis in early-stage NSCLC. High expression of several of these genes might be valuable for prediction of metastasis and survival. For the first time, this implicates a role for S100 proteins and trypsinosgens in the metastatic process of early-stage NSCLC.

Acknowledgments

We thank Maria Möller and Sarah Pierschalski for excellent technical assistance and Dr. Nicole Bäumer for critical reading of the manuscript. We are grateful to Dr. James Elder (University of Michigan), Dr. Beat Schäfer (University of Zürich), and Dr. Miklos Sahin-Toth (Boston University) for providing helpful comments and DNA constructions of S100 proteins and trypsinosgens.

References

S100 Family Members and Trypsinogens Are Predictors of Distant Metastasis and Survival in Early-Stage Non-Small Cell Lung Cancer

Sven Diederichs, Etmar Bulk, Björn Steffen, et al.

Cancer Res 2004;64:5564-5569.