Expression ratio of the TGFβ-inducible gene MYO10 is prognostic for overall survival of squamous cell lung cancer patients and predicts chemotherapy response

D. Dvornikov1,2,3, M. A. Schneider2,4, S. Ohse5, M. Szczzygieł1,2,3, I. Titkova1,3, M. Rosenblatt6, T. Muley2,4, A. Warth2,7, F. J. Herth2,8, H. Dienemann2,3, M. Thomas2,10, J. Timmer9,11,12, M. Schilling1, H. Busch5,13, M. Boerries5,14, M. Meister2,4 & U. Klingmüller1,2

In lung cancer a deregulation of Transforming Growth Factor-β (TGFβ) signaling has been observed. Yet, the impact of TGFβ in squamous cell carcinoma of the lung (LUSC) remained to be determined. We combined phenotypic and transcriptome-wide studies and showed that the stimulation of the LUSC cell line SK-MES1 with TGFβ results in an increase of migratory invasive properties. The analysis of the dynamics of gene expression by next-generation sequencing revealed that TGFβ stimulation orchestrates the upregulation of numerous motility- and actin cytoskeleton-related genes. Among these the non-muscle myosin 10 (MYO10) showed the highest upregulation in a LUSC patient cohort of the Cancer Genome Atlas (TCGA). Knockdown of MYO10 abrogated TGFβ-induced collagen gel invasion of SK-MES1 cells. The analysis of MYO10 mRNA expression in paired tissues of 151 LUSC patients with corresponding 80-month clinical follow-up data showed that the mRNA expression ratio of MYO10 in tumor and tumor-free tissue is prognostic for overall survival of LUSC patients and predictive for the response of these patients to adjuvant chemotherapy. Thus, MYO10 represents a new clinical biomarker for this aggressive disease and due to its role in cellular motility and invasion could serve as a potential molecular target for therapeutic interventions in patients with LUSC.

Non-small-cell lung cancer (NSCLC) is the leading cause of cancer-related mortalities. Early metastasis and therapy resistance are the main features that result in high mortality among lung cancer patients1. Adenocarcinoma of the lung (LUAD) and squamous cell carcinoma of the lung (LUSC) are the two major subtypes of NSCLC. Although the prevalence of LUSC in developed countries is declining, it still accounts for about 25% of NSCLC

1Division Systems Biology of Signal Transduction, German Cancer Research Center (DKFZ), 69120, Heidelberg, Germany. 2Translational Lung Research Center Heidelberg (TLRC), German Center for Lung Research (DZL), 69120, Heidelberg, Germany. 3Faculty of Biosciences, Heidelberg University, 69120, Heidelberg, Germany. 4Translational Research Unit, Thoraxklinik at University Hospital Heidelberg, 69126, Heidelberg, Germany. 5Institute of Molecular Medicine and Cell Research, University of Freiburg, 79104, Freiburg, Germany. 6Institute of Pathology, Heidelberg University, 69120, Heidelberg, Germany. 7Department of Physics, University of Freiburg, 79104, Freiburg, Germany. 8Department of Pneumology and Critical Care Medicine, Thoraxklinik at University Hospital Heidelberg, 69126, Heidelberg, Germany. 9Department of Surgery, Thoraxklinik at University Hospital Heidelberg, 69126, Heidelberg, Germany. 10Department of Thoracic Oncology, Thoraxklinik at University Hospital Heidelberg, 69126, Heidelberg, Germany. 11Freiburg Centre for Systems Biology, University of Freiburg, 79104, Freiburg, Germany. 12BIOSS Centre for Biological Signaling Studies, University of Freiburg, 79104, Freiburg, Germany. 13Institute of Experimental Dermatology & Institute of Cardiogenetics, University of Lübeck, 23562, Lübeck, Germany. 14German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), 69120, Heidelberg, Germany. D. Dvornikov, M. A. Schneider and S. Ohse contributed equally to this work. H. Busch, M. Boerries, M. Meister and U. Klingmüller jointly supervised this work. Correspondence and requests for materials should be addressed to U.K. (email: u.klingmueller@dkfz.de)
cases\(^2\). Despite the great progress in developing targeted approaches in LUAD, therapeutic options for LUSC remain very limited as driver oncogene mutations are uncommon\(^1\). For decades platinum-based chemotherapy has been the gold standard for first-line therapy for LUSC patients. However, in a significant proportion of patients cancer cells are resistant to chemotherapy and the disease rapidly progresses\(^1\). Thus, there is an urgent need to gain insights into mechanism contributing to LUSC in order to establish mechanism-based biomarkers that help clinicians to identify patients at the highest risk for disease progression and therapy resistance.

Both early metastasis and therapy resistance are attributed to cancer cells undergoing epithelial-to-mesenchymal transition (EMT) and acquiring a more invasive phenotype with cancer stem cell-like properties\(^2\). Tumor cells harboring EMT features were repeatedly reported to localize at the invasive front of the tumor, hence mediating cancer cell dissemination and metastasis\(^6\).

There is growing evidence that deregulated TGF\(\beta\) signaling contributes to the acquisition of an EMT phenotype by lung cancer cells. In the context of LUSC, elevated TGF\(\beta\) levels were correlated with poor patient prognosis\(^7\) and over-activation of the TGF\(\beta\) pathway was reported as a common feature in lung cancer\(^7\). Moreover, the EMT phenotype was widely observed in surgically resected specimens and associated with a worse clinical outcome and chemoresistance\(^8\). However, a mechanistic understanding of TGF\(\beta\)-induced changes and their impact on LUSC progression remained to be established. Therefore, we combined phenotypic and transcriptome-wide approaches to determine TGF\(\beta\)-induced dynamic changes in the transcriptome of a LUSC cell line and thereby derived a candidate prognostic biomarker that we validated in a clinical cohort.

Results

**TGF\(\beta\) treatment enhances pro-tumorigenic properties of LUSC cells.** To study the impact of TGF\(\beta\) on LUSC cells, we used the LUSC cell line SK-MES1 as a cellular model system. By quantitative immunoblotting we showed that TGF\(\beta\)-induced phosphorylation of Smad2 and Smad3 in SK-MES1 cells reached a maximum after 30 min and declined thereafter (Fig. 1A and Supplementary Fig. S1A). SK-MES1 cells usually grow in tight epithelial colonies, but after treatment with TGF\(\beta\) they lost cell-cell contacts and acquired an elongated spindle-shaped morphology (Fig. 1B), a feature commonly observed upon TGF\(\beta\)-induced epithelial-to-mesenchymal transition (EMT). In line with these morphological alterations, TGF\(\beta\) treatment of SK-MES1 cells induced the mRNA expression of classical EMT markers such as SNAI1, ZEB1, VIM and MMP9 (Fig. 1C and Supplementary Fig. 1B).

We explored how the activation of TGF\(\beta\) signal transduction and target gene expression as well as the morphological changes translated into altered phenotypic responses. To this aim, we established workflows to quantitatively assess at the single cell level the impact of TGF\(\beta\) on SK-MES1 cells in a 2D cell migration assay and a 3D collagen invasion assay (Fig. 2A,B upper panels). This included the development of a mathematical algorithm to correct for the concave surface of the collagen gels to improve the reliability of quantifications in the 3D collagen invasion assay. In the 2D migration assay we observed by analyzing more than 1000 of single cell tracks per condition that the TGF\(\beta\) treatment resulted in a two-fold increase in migration speed (from 4 to 8 µm/h) (Fig. 2A and Supplementary Fig. S2A). Co-treatment with a type I TGF\(\beta\) receptor inhibitor prevented this effect. In the 3D collagen invasion assay TGF\(\beta\) treatment resulted in a two-fold increase in the number of collagen-invaded SK-MES1 cells. Some of the TGF\(\beta\)-treated SK-MES1 cells invaded more than 100 µm into the dense collagen gels, while untreated cells invaded on average not more than 20 µm (Fig. 2B and Supplementary Fig. S2B). The increase in the invasion capacity was TGF\(\beta\)-specific because it was abolished by co-treatment with a type I TGF\(\beta\) receptor inhibitor.

It was reported that the EMT phenotype correlates with increased resistance to chemotherapy\(^9\). To examine the impact of TGF\(\beta\) on the resistance of SK-MES1 cells to cisplatin, a cell viability assay based on metabolic activity and an apoptosis assay based on caspase 3/7 activity was employed (Fig. 2C). We observed that pre-treatment of SK-MES1 with TGF\(\beta\) for 3 days resulted in a 4.4-fold increase of viable cells after 3 days exposure to 10 µg/ml cisplatin (Fig. 2C and Supplementary Fig. S2C). Likewise, pretreatment with TGF\(\beta\) reduced the caspase 3/7 activity across all tested doses of cisplatin by 25% (Fig. 2C and Supplementary Fig. S2D). Collectively, these data indicate that SK-MES1 cells acquire a more aggressive phenotype upon exposure to TGF\(\beta\).

**Multiple actin cytoskeleton- and motility related genes are upregulated in LUSC cells upon TGF\(\beta\) stimulation.** To elucidate mechanisms that contribute to TGF\(\beta\)-induced cancer cell invasion and increased chemotherapy resistance in LUSC cell line SK-MES1, we performed a time-resolved whole-transcriptome RNA-Seq analysis of SK-MES1 cells that were treated with TGF\(\beta\) for up to 48 hours or were left untreated. Genes were considered as differentially regulated if their overall mRNA expression dynamics in treated versus untreated cells was significantly different (multiple testing adjusted P-value < 0.01). In total expression of 2323 genes significantly changed in response to TGF\(\beta\) treatment (Fig. 3A). The resulting list of differentially regulated genes was used for Gene Set Enrichment Analysis (GSEA) to identify regulated gene ontology (GO) terms of cellular components, which were subsequently visualized with the REVIGO tool\(^10\) to establish clusters with distinct gene expression patterns. This approach revealed a preferential regulation of four gene clusters encoding actin cytoskeleton-, motility-, ECM- and secretory-related proteins (Fig. 3B). To narrow down the list of potential candidates involved in mediating the TGF\(\beta\)-induced invasive properties of LUSC cells, the five genes per cluster with the lowest multiple testing-adjusted P-values and with at least two-fold upregulation after normalization to untreated samples were selected. Because some of the genes were among the top five candidates in more than one cluster this resulted in a list of 15 TGF\(\beta\)-regulated genes (Fig. 3C,D and Supplementary Fig. S3A). Interestingly, genes identified as candidates in our approach included MYO10, SERPINE1, ITGB3, ITGA5, TGFBI, VIM and MARCKS. These TGF\(\beta\)-regulated genes were previously associated with increased cancer invasiveness, chemoresistance and worse clinical outcome in different cancer entities including breast, lung cancer (both LUSC and LUAD), invasive melanoma and prostate cancers\(^12\-17\). To determine which of these TGF\(\beta\)-regulated genes are...
Figure 1. TGFβ treatment triggers EMT in SK-MES1 cells. (A) TGFβ induces Smad2/3 phosphorylation in TGFβR1-dependent way. SK-MES1 cells were pretreated with TGFβR1 inhibitor SB-431542 or DMSO and then stimulated with 2 ng/ml TGFβ1. Data presented correspond to mean and SD, n is the number of independent experiments. Additional replicates are shown in Supplementary Fig. S1A. Full-length blots are shown in Supplementary Fig. S6. (B) Prolonged exposure of SK-MES1 cells to TGFβ1 induces acquisition of EMT-like morphology. Cells were either stimulated with 2 ng/ml TGFβ1 or left untreated for 3 days, fixed and stained for F-actin (white) and DNA (blue). Scale bar corresponds to 50 µm. (C) EMT marker genes are upregulated upon TGFβ1 treatment. Growth factor-depleted SK-MES1 cells were stimulated with 2 ng/ml TGFβ1 or left untreated. RNA was extracted and analyzed using qRT-PCR. mRNA expression was normalized to four housekeepers: GUSB, HPRT, GAPDH and G6PD. Each dot represents a biological replicate. A second independent experiment is shown in Supplementary Fig. S1B.
Figure 2. TGFβ treatment increases invasiveness and cisplatin resistance of squamous lung carcinoma cells SK-MES1. (A) Top, schematics of the 2D migration assay. Bottom, migratory properties of SK-MES1 cells are increased upon TGFβ treatment. Cells were seeded in 24-well plate, pretreated with either SB-431542 or DMSO, stimulated with 2 ng/ml TGFβ1 and imaged for 60 hours. Migration speed from each single cell track was quantified. Center lines show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend to 5th and 95th percentiles. N indicates the number of quantified single cell tracks per condition. Additional independent experiments are shown in Supplementary Fig. S2A. Statistical analysis was performed using one-way ANOVA; ***P < 0.001; n.s., not significant. (B) Top, schematics of the collagen 3D invasion assay. Bottom, TGFβ stimulation increases number of invading cells and the average invasion depth. SK-MES1 cells were seeded in 96-well plates with precast collagen gels, allowed to attach overnight, growth factor-depleted for three hours, pretreated with either SB-431542 or DMSO, stimulated with 2 ng/ml TGFβ1, allowed to invade for four days, stained with Hoechst and imaged with a confocal microscope. The number of invaded cells and invasion depth were assessed. One representative experiment is shown. Data are presented as median and SD, every dot corresponds to a biological replicate (n = 15). N indicates the number of invaded cells. Additional independent
relevant in the context of LUSC, we evaluated the alterations of mRNA levels of the selected 15 candidate genes in a cohort of 501 LUSC patients from the Cancer Genome Atlas (TCGA) (Fig. 3E). Strikingly, of those genes the MYO10 gene was the top candidate as it was upregulated in 27% of the patients, whereas an upregulation of the mRNAs of the other genes was only observed in 2–5% of the LUSC patients. Interestingly, the genes with the highest percentage of mRNA upregulation in LUSC patients belonged either to the migration or the actin cytoskeleton clusters, while genes from the ECM and secretory clusters were rarely altered in LUSC patients, although several of the genes from these clusters showed a high fold increase in SK-MES1 cells upon TGFβ treatment (Fig. 3C,D and Supplementary Fig. S3A). Given the prominent upregulation of MYO10 expression in LUSC patients and the pivotal role of non-muscle myosins in mediating cancer cell invasion in multiple cancer entities18, we examined whether other non-muscle-encoded genes scored high in our analysis but were not among the top five TGFβ-regulated genes. Indeed, in the LUSC patients of the TCGA cohort the second and third most regulated genes encoding myosins were MYH9 and MYO1E, which were previously implicated in cancer progression19,20. Both of these myosin genes were upregulated in LUSC patients of the TCGA cohort with MYH9 being overexpressed in 7% of the cases (Fig. 3E lower panel). A significant co-occurrence of an upregulation of the mRNAs of MYO10, MYH9, MYO1E and TGFB1 was observed in the LUSC patients of the TCGA cohort (Supplementary Fig. S3B), suggesting that the exposure of tumor cells to elevated levels of TGFβ might have stimulated upregulation of motility and invasion-related myosins. Therefore, all three myosin genes were included for further analysis.

**TGFβ-induced myosin motors are essential for TGFβ-mediated cancer cell invasion.** To examine the biological importance of the candidate TGFβ-induced myosins for LUSC, we validated the time-resolved RNA-Seq data (Fig. 4A) with time-resolved examinations by qRT-PCR of TGFβ-stimulated SK-MES1 cells (Fig. 4B). mRNA expression assessment by both methods significantly correlated with Spearman coefficients of $ρ = 0.92, 0.78$ and 0.54 for MYO10, MYO1E and MYH9 mRNAs, respectively (Supplementary Fig. S4A). All three candidate genes demonstrated strong mRNA induction upon TGFβ treatment, with MYO10 expression being the most pronounced and most sustained for up to 48 hours. Interestingly, cisplatin treatment alone did not alter expression of MYO10, MYH9 and MYO1E mRNAs, but slightly modulated TGFβ-induced induction when applied in combination with 2 ng/ml of TGFβ (Supplementary Fig. S4B).

Given the role of non-muscle myosins in cancer metastasis, we studied the effect of gene silencing on the ability of the SK-MES1 cells to invade 3D collagen gels in response to TGFβ stimulation. We used siRNA to knockdown MYO10, MYH9 or MYO1E and achieved for all of them a knockdown efficiency of more than 85% at the mRNA level (Fig. 4C). Whereas TGFβ treatment of SK-MES1 cells transfected with control non-targeting siRNA resulted in a two-fold increase in the number of invaded cells compared to unstimulated cells (Fig. 4D), the TGFβ-enhanced invasion of SK-MES1 cells was abrogated upon downregulation of the different myosins. These results indicate that the TGFβ-induced non-muscle myosins MYO10, MYH9 and MYO1E play a non-redundant and crucial role in mediating TGFβ-regulated invasiveness of the LUSC cells.

**MYO10 mRNA overexpression is prognostic for overall survival of patients with squamous cell carcinoma.** Actin-based protrusions and TGFβ-induced myosins are crucial for multiple phases of the metastatic cascade18. Among the identified TGFβ-regulated non-muscle myosins MYO10 showed the strongest upregulation in response to TGFβ stimulation in SK-MES1 cells and the highest mRNA overexpression in LUSC patients of the TCGA cohort. Therefore, we further assessed its clinical relevance in paired tumor and tumor-free tissue from our NSCLC cohort consisting of both LUAD and LUSC patients with a similar tumor stage distribution among LUSC patients as in the TCGA LUSC cohort (Supplementary Table S1 and Supplementary Fig. S5B). For each tumor entity, patients were divided into two subgroups, MYO10 fold change $< 1$ and MYO10 fold change $> 1$, based on the expression ratio of MYO10 mRNA in tumor versus tumor-free tissue (Fig. 5A). To investigate the prognostic value of the MYO10 mRNA expression ratio, we performed Cox regression analysis (Table 1). Univariate analysis indicated that a high MYO10 mRNA expression ratio ($P = 0.018$, gender ($P = 0.04$) as well as the pathological stages ($P = 0.005$ for stage II and $P < 0.001$ for stage III) were prognostic factors for the overall patient survival. The multivariate analysis suggested that a high MYO10 mRNA expression ratio was only prognostic for LUSC patients, but not for LUAD patients (Table 2). Using the MYO10 mRNA expression ratio to separate the patient groups, we confirmed that LUSC patients with high MYO10 mRNA expression ratio demonstrated reduced overall survival independent of the tumor stage and treatment regimen ($P = 0.008$, Fig. 5B left), which was not observed in LUAD patients ($P = 0.57$, Fig. 5B right). Next, the LUSC patients were subdivided according to those that had no further treatment after the resection (mainly stage IB and IIA patients with a median disease-free survival of 28.4 months) and those who received adjuvant chemotherapy (mainly stage IIB and IIIA patients with a median disease-free survival of 35.6 months) (Fig. 5C and Supplementary Fig. S5C,D). This analysis showed that for untreated patients MYO10 mRNA expression ratio expression was not predictive for overall survival ($P = 0.429$) (Fig. 5C, left panel). On the contrary, patients with low MYO10 mRNA expression
**Figure 3.** TGFβ treatment of LUSC cells results in upregulation of migration- and actin cytoskeleton-related genes. (A) Non-supervised hierarchical clustering of z-scored differentially regulated 2323 genes (adjusted P-value < 0.01) between TGFβ-treated and untreated conditions. SK-MES1 cells were stimulated with 2 ng/ml TGFβ or left untreated. RNA was extracted and sequenced using HiSeq 4000. (B) Clusters of significantly upregulated GO cellular component gene sets between TGFβ-treated and untreated conditions. Significantly upregulated GO terms (adjusted P-value < 0.01) were visualized using REVIGO (allowed similarity 0.5). Thickness of connecting grey lines corresponds to the similarity of the GO terms. Only clusters that consist of at least two GO terms are displayed. (C) Volcano plot of differentially regulated genes between TGFβ-treated and untreated conditions. Fold change of averaged 8–48 h time points between both conditions is displayed. Only significantly regulated genes (adjusted P-value < 0.01) with a fold change of at least two are shown. Five most regulated genes from each cluster of upregulated GO cellular component gene sets are indicated with corresponding colors. Grey circles indicate differentially regulated genes that do not belong to any of the four clusters. (D) Time-resolved dynamics of top differentially regulated candidate genes from each of the clusters. Top five genes from each of the four clusters with the lowest adjusted P-values and fold change of at least two after normalization to untreated samples were selected as candidates. In case the same gene belonged to different clusters and satisfied the inclusion criteria, it was marked as belonging to both clusters. Single gene plots are shown in the Supplementary Fig. S3A. (E) TCGA LUSC cohort RNA-Seq expression data of selected candidate genes sorted by frequency of mRNA upregulation. MYH9, TGFβ1 and MYO1E genes were additionally included.
MYO10 siRNA
MYH9 siRNA
MYO1E siRNA

Figure 4. TGFβ-inducible myosins are required for TGFβ-mediated cancer cell invasion. (A) Time-resolved RNA-Seq data of selected myosin genes upon TGFβ treatment in SK-MES1 cells. Cells were growth factor-depleted for three hours and stimulated with 2 ng/ml TGFβ1 or left untreated. mRNA was extracted and sequenced using HiSeq 4000. Data are presented in TPM (transcripts per million) values. Each dot represents a biological replicate, shaded areas correspond to standard error. (B) qPCR validation of RNA-Seq data. RNA of TGFβ-treated and untreated SK-MES1 cells that was used for RNA-Seq was also analyzed with qRT-PCR. Data presented correspond to mean and SD from three biological replicates. (C) siRNA knockdown results in 90% knockdown efficiency of TGFβ-inducible myosins. After 36 hours of siRNA transfection SK-MES1 cells were stimulated for 1 hour with 2 ng/ml TGFβ1 or left untreated. mRNA was extracted and the knockdown efficiency was analyzed using qRT-PCR. Data represent mean and SD from six biological replicates. (D) Myosins knockdowns inhibit TGFβ-mediated cancer cell invasion. SK-MES1 cells were transfected with siRNAs for 36 hours and stimulated with 2 ng/ml TGFβ1 for four days. Amount of invaded cells into the collagen gel was assessed. One representative experiment is shown. Every dot corresponds to a biological replicate (n ≥ 7), black line indicates the median. Additional independent experiment is shown in Supplementary Fig. S4C. Statistical analysis was performed using one-way ANOVA; ****P < 0.001.

ratio strongly benefited from the adjuvant chemotherapy treatment in comparison to patients with a high MYO10 mRNA expression ratio (P = 0.001) (Fig. 5C, right panel). Therefore, we conclude that MYO10 mRNA expression ratio is predictive for the outcome of adjuvant chemotherapy treatment of LUSC patients.

Because of the observed enhanced chemoresistance of LUSC cells after TGFβ treatment (Fig. 2C) and because TGFβ-induced EMT has been associated with chemotherapy resistance in patients9,21, we determined the expression of EMT markers in tissue of LUSC patients. Notably, patients with an elevated MYO10 mRNA expression ratio displayed a higher expression of EMT signature genes such as SNAI2, TWIST1 and VIM (Fig. 5D and Supplementary Fig. S5A). The fact that TGFβ is one of the most potent EMT-inducers22 and the co-occurrence of MYO10 and TGFβ1 mRNA upregulation in a substantial proportion of LUSC patients (Supplementary Fig. S3B) suggest that activation of TGFβ signaling might trigger the observed alterations in LUSC patients. Finally, a higher MYO10 mRNA expression ratio was observed in patients with stage III disease (Fig. 5E), making it prognostic for patients with a higher pathological stage and affected local or distant lymph nodes (Fig. 5F, G). Taken together, our studies suggest that the mRNA expression ratio of MYO10 can be used as a new independent prognostic biomarker for survival in patients with resected LUSC.

Discussion
We investigated TGFβ-induced changes in signal transduction, gene expression and phenotypic responses in the LUSC cell line SK-MES1 to shed light on TGFβ-induced mechanisms that might contribute to tumor progression in LUSC. Our analysis identified the TGFβ-inducible non-muscle MYO10 as an essential mediator of TGFβ-regulated cancer cell invasion. Finally, we showed that a high MYO10 mRNA expression ratio is an independent biomarker for patients at risk for a more aggressive course of the disease.

Previous reports regarding the importance of TGFβ signaling in LUSC progression are contradictory. It was shown that elevated TGFβ1 plasma levels and upregulated downstream gene targets, such as TWIST1 and SNAI2, correlated with poor patient prognosis23,24, while it was also reported that reduced TGFβRII immunostaining in LUSC patients is associated with a more aggressive tumor behavior and reduced patient survival25. Here, we observed that exposure of the LUSC cell line SK-MES1 to TGFβ triggers changes in cell morphology that were accompanied by increased chemoresistance as well as an increase in migratory and invasive properties of the cells. These observations support a pro-metastatic role of TGFβ signaling in the context of LUSC.
Figure 5. **MYO10** mRNA expression ratio is prognostic for overall survival of LUSC but not LUAD patients. (A) **MYO10** mRNA expression ratio of tumor and adjacent non-tumor tissues in LUSC and LUAD patients. RNA was isolated from fresh-frozen resected tissues and **MYO10** expression was measured using qRT-PCR. (B) Kaplan-Meier curves for overall survival using **MYO10** mRNA expression ratio in LUSC and LUAD cohorts. Significance of difference between the two groups was tested using non-parametric Mann-Whitney U test. (C) Kaplan-Meier curves for adjuvant chemotherapy response in **MYO10** low (left) and **MYO10** high (right) patients. Significance of difference between the two groups was tested using non-parametric Mann-Whitney U test. (D) Differences of **SNAI2** and **TWIST1** expression depending on **MYO10** mRNA expression ratio were tested by unpaired t-tests; ***P < 0.001; ****P < 0.0001. (E) **MYO10** expression ratio in different stages of LUSC. Differences were tested by one-way ANOVA; *P ≤ 0.05. (F) Kaplan-Meier curve for the pathological stages of LUSC patients. Significance of difference between the two groups was tested using non-parametric Mann-Whitney U test. (G) Kaplan-Meier curves for lymph node status using **MYO10** mRNA expression ratio. Significance of difference between the two groups was tested using non-parametric Mann-Whitney U test.
There is growing evidence that aberrant upregulation of actin cytoskeleton proteins plays a central role in metastatic progression in different types of cancers. Particularly, non-muscle myosins are implicated in tumor progression through their roles in cell migration and invasion. It was reported that TGF-β1 stimulation of the epithelial-like lung adenocarcinoma cell line H1437 triggered upregulation of cytoskeletal proteins and induced a more aggressive cell phenotype. By integration of transcriptome and methylome data it was shown that the enrichment of actin cytoskeleton proteins correlated with poor survival of lung adenocarcinoma patients. Others observed that high expression of the actin-binding filopodial-crosslinking protein fascin was prognostic for an aggressive metastatic disease and for poor survival of patients with breast and esophageal squamous cell tumors. Here, we showed that siRNA-mediated knockdown of either MYO10, MYH9 or MYO1E was sufficient to abrogate TGF-β-induced 3D collagen invasion of the LUSC cell line SK-MES1. These observations are in agreement with previous work demonstrating that shRNA knockdown of either MYO10, MYH9 or MYO1E was sufficient to abrogate TGFβ-induced 3D collagen invasion of the LUSC cell line SK-MES1. These observations are in agreement with previous work demonstrating that shRNA knockdown of either MYO10, MYH9 or MYO1E was sufficient to abrogate TGFβ-induced 3D collagen invasion of the LUSC cell line SK-MES1. These observations are in agreement with previous work demonstrating that shRNA knockdown of either MYO10, MYH9 or MYO1E was sufficient to abrogate TGFβ-induced 3D collagen invasion of the LUSC cell line SK-MES1. These observations are in agreement with previous work demonstrating that shRNA knockdown of either MYO10, MYH9 or MYO1E was sufficient to abrogate TGFβ-induced 3D collagen invasion of the LUSC cell line SK-MES1. These observations are in agreement with previous work demonstrating that shRNA knockdown of either MYO10, MYH9 or MYO1E was sufficient to abrogate TGFβ-induced 3D collagen invasion of the LUSC cell line SK-MES1. These observations are in agreement with previous work demonstrating that shRNA knockdown of either MYO10, MYH9 or MYO1E was sufficient to abrogate TGFβ-induced 3D collagen invasion of the LUSC cell line SK-MES1. These observations are in agreement with previous work demonstrating that shRNA knockdown of either MYO10, MYH9 or MYO1E was sufficient to abrog...
pre-existing lymph node metastases have significantly poorer overall survival ($P = 0.017$, Fig. 5G) alongside with increased chemotherapy resistance (Fig. 5C). Similar observations that tumor cells with increased metastatic potential are more resistant to chemotherapy treatment were previously made across several cancer types, including lung and breast cancers. This was attributed to a higher metastatic potential and a reduced sensitivity to chemotherapeutic agents of tumor cells undergoing EMT that has been associated with a decreased proliferation rate and upregulation of resistance-related genes. Therefore, the determination of the MYO10 mRNA expression ratio in resected lung tumors could be used to discriminate between LUSC patients that would benefit from platinum-based adjuvant chemotherapy or not. Additionally, the prognostic value of the MYO10 mRNA expression ratio is to help clinicians identify patients at higher risk of chemotherapy failure and disease progression. Those patients should not only be monitored more closely, but also potentially be considered for treatment with immune checkpoint inhibitors as an emerging treatment strategy in LUSC. Finally, as it appears that a large fraction of LUSC tumors rely on TGF-β-inducible genes that we identified in our time-resolved transcriptome-wide study, including the non-muscle myosins MYO1E and MYH9, might comprise a promising robust prognostic gene signature for patients with resected LUSC.

Concluding, non-muscle myosin MYO10 serves as a downstream effector contributing to a more aggressive clinical course through increased migratory and invasive properties of the tumors as well as enhanced chemoresistance. We showed that the MYO10 mRNA expression ratio can be used as an independent prognostic factor for survival of patients with resected lung squamous cell carcinoma and propose that MYO10 may represent a new molecular target for therapeutic intervention.

Materials and Methods

Cell lines and culture conditions. The human LUSC cell line SK-MES1 was purchased from ATCC and authenticated with a multiplexed human cell line authentication test (Multiplexion). Cells were grown in culture medium: DMEM (Lonza) supplemented with 10% FCS (Gibco), 100 U/ml penicillin and 100 μg/ml streptomycin (Gibco). For TGFβ3 stimulation (R&D Systems, #240-B-010) stimulation, growth factor-depleted medium was used: DMEM without Phenol red supplemented with 1 mg/ml BSA (Sigma), 100 U/ml penicillin, 100 μg/ml streptomycin and 2 mM L-glutamine. Cells were kept in 5% CO₂ at 37°C and 95% relative humidity. Cells were passaged to a maximum of 25 passages and controlled for mycoplasma contamination (Multiplexion).

Immunoblotting. To examine the activation of the TGFβ3/Smad pathway, 1.4 × 10⁶ of SK-MES1 cells were plated on 6-cm dishes (TPP, #93060). The next day, medium was replaced with growth factor-depleted DMEM for three hours, followed by stimulation with 2 ng/ml TGFβ3 or in combination with TGFβ3/R1 inhibitor SB-431542 (System Bioscience, #ZRD-SB-02). At the indicated time points, cells were lysed in 500 μl whole-cell lysis buffer (1% NP40, 150 mM NaCl, 50 mM Tris-HCl pH 7.4, 2.5 mM NaN₃, 1 mM EDTA, 0.5 mM Na₃VO₄, 1 mg/ml deoxycholic acid, 2 μg/ml aprotinin and 200 μg/ml AEBSF). Lysates were kept on ice for 10 min, sonicated for 25 sec (Bandelin Sonopuls) and centrifuged for 10 min at 21,000 g at 4°C. Protein concentration was determined by BCA assay (Pierce, #23225). For immunoprecipitation (IP), 700 μg of the protein was mixed with 2.5 μl of antibodies against Smad3/2/3 (BD Biosciences, #610842) and 25 μl 50% ProteinA sepharose-slurry (GE Healthcare, #17-0963-03). IP samples were rotated overnight at 4°C, washed twice with whole-cell lysis buffer, once with TNE buffer (10 mM Tris, 100 mM NaCl, 1 mM EDTA and 100 μM NaVO₃) and boiled for 5 min at 95°C. Samples were separated on 10% SDS-PAGE, transferred to nitrocellulose membrane (Amersham, #1060001), blocked for 1 h with 33% (v/v) Odyssey blocking buffer (LI-COR, #927-40010) in PBS and incubated with anti-Smad3/2/3 (BD Biosciences, #610842, 1:2,000), anti-pSmad3 (Cell Signaling Technology, #9520, 1:2,000) antibodies. Secondary antibodies coupled to IRDye infrared dyes (LI-COR, #926-32211 and #926-68070, 1:15,000) were used for detection with infrared Odyssey imager (LI-COR). Primary and secondary antibodies were diluted in 33% (v/v) Odyssey blocking buffer (LI-COR, #927-40010) in 0.2% TBS-Tween 20. Infrared Odyssey imager (LI-COR) was used for detection. Signal quantification was performed using the ImageQuant TL software (GE Healthcare). Independent replicates were scaled and averaged using methods previously described.

2D migration assay. 25,000 cells/well were seeded in 24-well plate (Zell Kontakt # 3231-20), allowed to attach overnight, growth factor-depleted for 3 h, stained with Hoechst (Santa Cruz, #sc-396575), stimulated with 2 ng/ml TGFβ3 alone or in combination with TGFβ3/R1 inhibitor SB-431542 and imaged on an environment-controlled microscope (IX81, Olympus). Nine positions per well (3 × 3 grid) were acquired in 20 min intervals for 60 h. Images were stitched with the Grid/Collection Imagej plugin. Single cell tracking was performed with the ImageJ Mtrack2 plugin.

3D collagen invasion assay. 3D collagen gels were prepared as described previously. Briefly, ice-cold 1 M HEPES buffer, 0.7 M NaOH, 10 × PBS pH 8.0 and bovine skin collagen G solution (L1613, Biochreme) were mixed in 1:1:2.16 ratio. 50 μl of the resulting solution was added per well of a 96-well plate with a flat bottom (BD #353376). Plates were kept overnight at 4°C and for 2 h at 37°C to allow gelation of the collagen. After gelation, 12,500 cells per well were seeded on top of the matrix, allowed to adhere overnight, stimulated with 2 ng/ml TGFβ3 in FCS-depleted medium and allowed to invade for 4 days. Afterwards, cells were fixed in 3.7% paraformaldehyde for 1 h and subsequently stained with Hoechst (Santa Cruz, #sc-396575). Collagen embedded cells were imaged using a LSM710 confocal microscope (Carl Zeiss) equipped with a EC Plan-Neofluar DIC 10 × /0.3 NA objective lens (Carl Zeiss). For each well, 2 × 2 tile z-stacks were acquired. Image analysis was performed using Imaris software (Bitplane). Spot detection algorithm was applied to assign coordinates for the center of each nucleus. Coordinates of all detected nuclei were exported and corrected by a conical plane function using the...
Invasion correction R package\(^{40}\). The threshold for invasion was set to 30 \(\mu\)m. Percentage of invaded cells and the invasion depth were the output.

**Cell viability and caspase activity assays.** 10,000 cells per well were seeded in 96-well plates, allowed to attach overnight and stimulated with 2 ng/ml TGF\(\beta\) in growth factor-depleted medium or left untreated. Three days later the medium was replaced with culture DMEM, with or without TGF\(\beta\), and supplemented with different concentrations of cisplatin (Teva). CellTiter-Blue Viability (Promega) and Apo-ONE Homogeneous Caspase-3/7 (Promega) assays were used to measure the viability of cells and caspase 3/7 activity.

**RNA-Seq.** SK-MES1 cells were seeded on 6-cm dishes at a density of \(10^6\) cells/dish. Two days later, cells were growth factor-depleted for 3 hours and stimulated with 2 ng/ml TGF\(\beta\) or left untreated. Total RNA was extracted at the indicated time points and sequenced using HiSeq 4000 (Illumina). The RNA sequencing results were aligned to the GRCh38 reference genome (Ensembl Release 79) and transcript quantification was performed using kallisto (v0.43.0)\(^{41}\). Subsequent transcript abundances were summarized to the gene level for analysis using tximport (v1.2.0)\(^{42}\). All ID mapping was performed using biomaRt (v2.30.0)\(^{43}\) and genes with read counts <10 were filtered out. Differential expression analysis was performed according to limma voom (v3.30.13)\(^{44}\) to obtain F-values and adjusted P-values. Adjusted P-values were calculated by multiple testing correction according to the Benjamini–Hochberg procedure. Gene set enrichment analysis (GSEA) was performed on the F-values of the differential expression analysis according to GAGE\(^{45}\). Upregulated GO cellular components terms (adjusted P-value < 0.01) were used for network analysis and visualized using REVIGO\(^{11}\) with allowed similarity of 0.5 and Cytoscape\(^{46}\).

**siRNA transfections.** A pool of four targeting siRNAs (5 nM; SMARTpool, Dharmacon) was in-solution transfected during cell seeding using Lipofectamine RNAiMAX (ThermoFisher Scientific, #13778150) according to the manufacturer’s instructions. A pool of four non-targeting siRNA was used as a negative control (5 nM, Dharmacon, #D-001810-10-20). Experiments were started 36 hours after transfection. Sequences of used siRNAs are listed in the Supplementary Table S2.

**Tissue sample collection, characterization and preparation.** Tissue samples were provided by Lung Biobank Heidelberg, a member of the accredited Tissue Bank of the National Center for Tumor Diseases (NCT) Heidelberg, the BioMaterialBank Heidelberg and the Biobank platform of the German Center for Lung Research (DZL). All patients provided written informed consent for the use of the tissue for research purpose. The study was approved by the local ethics committee of the University of Heidelberg (No. 270/2001). All research was performed in accordance with relevant guidelines/regulations. Tumor and matched distant (>5 cm) normal lung tissue samples from NSCLC patients (n = 362) who underwent resection for primary lung cancer at the Thoraxklinik at University Hospital, Heidelberg, Germany were collected. All diagnoses were made according to the 2004 WHO classification for lung cancer\(^{47}\) by at least two experienced pathologists. Tissues were snap-frozen within 30 minutes after resection and stored at \(-80^\circ\)C until the time of analysis. Only samples with a viable tumor content of \(\geq 50\%\) were used for subsequent analyses. For nucleic acid isolation 10–15 tumor cryosections (10–15 \(\mu\)m each) were prepared for each patient. The first and the last sections in each series were stained with hematoxylin and eosin (H&E) and were reviewed by an experienced lung pathologist to determine the proportions of viable tumor cells, stromal cells, normal lung cell cells, infiltrating lymphocytes and necrotic areas\(^{48}\).

**Total RNA isolation and cDNA synthesis.** Frozen tumor cryosections were homogenized with the TissueLyser mixer-mill disruptor (Qiagen). Matched normal lung tissue pieces were homogenized using a Micrcr D-8 rotor-stator homogenizer (Art- moderne Labortechnik). Total RNA was isolated from tissue or from cultured cells using an RNeasy Mini Kit (Qiagen). The quantity of RNA was measured with a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies). The quality of total RNA was assessed with an Agilent 2100 Bioanalyzer and Agilent RNA 6000 Nano Kit (Agilent Technologies). Total RNA was transcribed to ssDNA with a Transcriptor First Strand cDNA Synthesis Kit (Roche) in three independent reactions that were afterwards pooled, mixed and stored at \(-80^\circ\)C until further analyses.

**Quantitative Real-Time PCR.** qRT-PCR was performed using a LightCycler 480 (Roche) under MIQE guidelines\(^{49}\). CT values were calculated using the second derivative maximum method (Roche). To evaluate differences in patient gene expression, a relative quantification method based on the \(\Delta\Delta^PCT\) method was performed\(^{50}\). Target genes were normalized with the geometric mean of two housekeeping genes (ESD and RPS18) for patient samples or four genes (GUSB, GAPDH, G6PD and HPRT) for SK-MES1 samples. Sequences of used PCR primers are listed in the Supplementary Table S3.

**Statistical analyses.** The qRT-PCR data were statistically analyzed under REMARK criteria\(^{51}\) with SPSS 22.0 for Windows (IBM). The endpoint of the study was overall survival. Survival was calculated from the date of surgery until the last date of contact or death. Univariate analysis of survival data was performed according to Kaplan and Meier\(^{52}\) and using Cox proportional hazard models. Multivariate survival analysis was performed using the Cox proportional hazards model. The non-parametric Mann-Whitney U test was used to investigate significant differences between the patient groups.

**Data availability.** The RNA-Seq data was deposited at the Sequence Read Archive (SRA) and is accessible via the Gene Expression Omnibus ID GSE95536.
References

1. Rami-Porta, R., Crowley, J. J. & Goldstraw, P. The revised TNM staging system for lung cancer. *Ann Thorac Cardiovasc Surg* **15**, 4–9 (2009).
2. Drilon, A., Rekhtman, N., Ladanyi, M. & Paik, P. Squamous-cell carcinomas of the lung: emerging biology, controversies, and the promise of targeted therapy. *Lancet Oncol* **13**, e418–426, https://doi.org/10.1016/S1470-2045(12)70291-7 (2012).
3. Rooney, M., Devarakonda, S. & Govindan, R. Genomics of squamous cell lung cancer. *The oncologist* **18**, 707–716, https://doi.org/10.1016/j.jtho.2013.06.003 (2013).
4. Kim, E. S. et al. Gefitinib versus docetaxel in previously treated non-small-cell lung cancer (INTEREST): a randomised phase III trial. *Lancet* **372**, 1809–1818, https://doi.org/10.1016/S0140-6736(08)61578-4 (2008).
5. Yu, M. et al. Circulating breast cancer cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* **339**, 580–584, https://doi.org/10.1126/science.1228522 (2013).
6. Munn, Y. et al. Transcription factors related to epithelial mesenchymal transition in tumor margin and in invasive lung adenocarcinoma. *Int J Exp Pathol* **97**, 4095–4103 (2014).
7. Sterlacci, W. et al. High transforming growth factor beta expression represents an important prognostic parameter for surgically resected non-small cell lung cancer. *Human pathology* **43**, 339–349, https://doi.org/10.1016/j.humpath.2011.05.017 (2012).
8. Marwitz, S. et al. Downregulation of the TGFBeta Pseudoreceptor BAMBI in Non-Small Cell Lung Cancer Enhances TGFbeta Signaling and Invasion. *Cancer Res* **76**, 3785–3801, https://doi.org/10.1158/0008-5472.CAN-15-1326 (2016).
9. Shintani, Y. et al. Epithelial to mesenchymal transition is a determinant of sensitivity to chemoradiotherapy in non-small cell lung cancer. *Ann Surg Oncol* **92**, 1794–1804; discussion 1804, https://doi.org/10.1007/s10432-011-1703-2 (2011).
10. Arumugam, T. et al. Epithelial to mesenchymal transition contributes to drug resistance in pancreatic cancer. *Cancer Res* **69**, 5820–5828, https://doi.org/10.1158/0008-5472.CAN-08-2819 (2009).
11. Supek, F., Bosnjak, M., Skunca, N. & Smuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, e21800, https://doi.org/10.1371/journal.pone.0021800 (2011).
12. Makowska, K. A., Hughes, R. E., White, K. J., Wells, C. M. & Peckham, M. Specific Myosins Control Actin Organization, Cell Morphology, and Migration in ProstateCancer Cells. *Cell reports* **13**, 2118–2125, https://doi.org/10.1016/j.celrep.2015.11.012 (2015).
13. Ruiz de Garibay, G. et al. Lymphangiomyomatosis Biomarkers Linked to Lung Metastatic Potential and Cell Stemness. *PLoS One* **10**, e0132546, https://doi.org/10.1371/journal.pone.0132546 (2015).
14. Soltermann, A. et al. Prognostic significance of MYH9 expression in resected non-small cell lung cancer. *PLoS One* **10**, e0114160, https://doi.org/10.1371/journal.pone.0114160 (2015).
15. Kolon, K. et al. Molecular portraits of epithelial, mesenchymal, and hybrid States in lung adenocarcinoma and their relevance to survival. *Cancer Res* **75**, 1789–1800, https://doi.org/10.1158/0008-5472.CAN-14-2535 (2015).
16. Hashimoto, Y. et al. Prognostic significance of fascin overexpression in human esophageal squamous cell carcinoma. *Clin Cancer Res* **20**, 2597–2605, https://doi.org/10.1158/0001-1340.CAN-14-3178 (2015).
17. Boer, B. J. et al. The expression of fascin, an actin-bundling motility protein, correlates with hormone receptor-negative breast cancer and a more aggressive clinical course. *Clin Cancer Res* **11**, 186–192 (2005).
18. Arjomand, A. et al. Molecular portraits of Twist, Slug, and Foxc2 expression in stage I non-small cell lung cancer after curative resection. *Clin Lung Cancer* **13**, 280–287, https://doi.org/10.1016/j.clincr.2011.11.005 (2012).
19. Malkoš, P. et al. Loss of transforming growth factor beta type II receptor increases aggressive tumor behavior and reduces survival in lung adenocarcinoma and squamous cell carcinoma. *Clin Cancer Res* **18**, 2173–2183, https://doi.org/10.1158/1078-0432.CCR-11-2557 (2012).
20. Yilmaz, M. & Christofori, G. EMT, the cytoskeleton, and cancer cell invasion. *Cancer Metastasis Rev* **28**, 15–33, https://doi.org/10.1007/s10555-008-9699-0 (2009).
21. Schliekelman, M. J. et al. Molecular portraits of epithelial, mesenchymal, and hybrid States in lung adenocarcinoma and their relevance to survival. *Cancer Res* **75**, 1789–1800, https://doi.org/10.1158/0008-5472.CAN-14-2535 (2015).
22. Hashimoto, Y. et al. Prognostic significance of fascin overexpression in human esophageal squamous cell carcinoma. *Clin Cancer Res* **20**, 2597–2605, https://doi.org/10.1158/0001-1340.CAN-14-3178 (2015).
23. Vonder, B. J. et al. The expression of fascin, an actin-bundling motility protein, correlates with hormone receptor-negative breast cancer and a more aggressive clinical course. *Clin Cancer Res* **11**, 186–192 (2005).
24. Arjomand, A. et al. Mutant p53-associated myosin-X upregulates promote breast cancer invasion and metastasis. *J Clin Invest* **124**, 1069–1082, https://doi.org/10.1172/JCI67280 (2014).
25. Xia, Z. K. et al. Nonsmooth myosin IIA is associated with poor prognosis of esophageal squamous cancer. *Diseases of the esophagus: official journal of the International Society for Diseases of the Esophagus* **25**, 427–436, https://doi.org/10.1111/j.1442-0530.2011.01261.x (2012).
26. Ouderkerk, J. J. & Krendel, M. Myosin 1e is a component of the invadosome core that contributes to regulation of invadopod dynamics. *Exp Cell Res* **322**, 265–276, https://doi.org/10.1016/j.yexcr.2014.01.015 (2014).
27. Xiong, D. et al. Myosin II is an independent predictor of overall survival for cystectomy candidates with early-stage bladder cancer. *Oncology reports* **28**, 1625–1632, https://doi.org/10.3892/or.2012.1965 (2012).
28. Acharya, S. et al. A CXCL1 paracrine network links cancer chemoresistance and metastasis. *Cell* **150**, 165–178, https://doi.org/10.1016/j.cell.2012.04.042 (2012).
29. Xu, J. et al. Nonsmooth myosin X-A contributes to cisplatin resistance and lymphatic metastasis in non-small cell lung cancer and predicts poor prognosis. *Journal of translational medicine* **12**, 200, https://doi.org/10.1186/1479-5867-12-200 (2014).
30. Zheng, X. et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* **527**, 525–530, https://doi.org/10.1038/nature16064 (2015).
31. Braham, J. et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. *N Engl J Med* **373**, 123–135, https://doi.org/10.1056/NEJMoa1504627 (2015).
32. von der Heyde, S. et al. RPPanalyzer toolbox: an improved R package for analysis of reverse phase protein array data. *BioTechniques* **57**, 125–135, https://doi.org/10.1387/biotech.2144480 (2014).
33. Burgstaller, G., Oehrle, B., Koch, I., Lindner, M. & Eickelberg, O. Multiplex profiling of cellular invasion in 3D cell culture models. *PLoS One* **8**, e63121, https://doi.org/10.1371/journal.pone.0063121 (2013).
52. Dinse, G. E. & Lagakos, S. W. Nonparametric estimation of lifetime and disease onset distributions from incomplete observations.
51. McShane, L. M.
47. Beasley, M. B., Brambilla, E. & Travis, W. D. The 2004 World Health Organization classification of lung tumors.
48. Muley, T. R., Herth, F. J., Schnabel, P. A., Dienemann, H. & Meister, M. From tissue to molecular phenotyping: pre-analytical
49. Bustin, S. A.
40. Rosenblatt, M. In R package version 0.1. (CRAN, https://CRAN.R-project.org/package=InvasionCorrection 2017).
41. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. Nature biotechnology 34,
525–527, https://doi.org/10.1038/nbt.3519 (2016).
42. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level
inferences. F1000Research 4, 1521, https://doi.org/10.12688/f1000research.7563.2 (2015).
43. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/
Bioconductor package biomart. Nature protocol 4, 1184–1191, https://doi.org/10.1038/nprot.2009.97 (2009).
44. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts.
Genome biology 15, R29, https://doi.org/10.1186/gb-2014-15-2-r29 (2014).
45. Luo, W., Friedman, M. S., Shedden, K., Hankenson, K. D. & Woolf, P. J. GAGE: generally applicable gene set enrichment for pathway
analysis. BMC Bioinformatics 10, 161, https://doi.org/10.1186/1471-2105-10-161 (2009).
46. Cline, M. S. et al. Integration of biological networks and gene expression data using Cytoscape. Nature protocols 2, 2366–2382,
https://doi.org/10.1038/nprot.2007.324 (2007).
47. Beasley, M. B., Brambilla, E. & Travis, W. D. The 2004 World Health Organization classification of lung tumors. Seminars in
oncology 40, 90–97 (2005).
48. Muley, T. R., Herth, F. J., Schnabel, P. A., Diememann, H. & Meister, M. From tissue to molecular phenotyping: pre-analytical
requirements heidelberg experience. Translational lung cancer research 1, 111–121, https://doi.org/10.3978/j.issn.2218-6751.2011.12.07 (2012).
49. Bustin, S. A. et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin
Chem 55, 611–622, https://doi.org/10.1373/clinchem.2008.112797 (2009).
50. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2ΔΔCT Method.
Methods 25, 402–408, https://doi.org/10.1006/meth.2001.1262 (2001).
51. McShane, L. M. et al. Rreporting recommendations for tumour MARKer prognostic studies (REMARK). Eur J Cancer 41,
1690–1696, https://doi.org/10.1016/j.ejca.2005.03.032 (2005).
52. Dinse, G. E. & Lagakos, S. W. Nonparametric estimation of lifetime and disease onset distributions from incomplete observations.
Biometrics 38, 921–932 (1982).

Acknowledgements
The support of Damir Krunic from the DKFZ Light Microscopy facility is gratefully acknowledged. Authors thank Bettina Oehrle and Gerald Burgstaller from Oliver Eickelbergs lab at CPC Munich for their help with establishing the 3D collagen invasion assay. We would like to thank Martin Fallenbuechel, Chang Xu, Jessica Eschenbach and Christa Stolp for expert technical assistance. This study was funded by the German Center for Lung Research [Deutsches Zentrum für Lungenforschung (DZL), grant numbers 82DZL00404, 82DZL00402, 82DZL004A2]. Funding from the BMBF within the CancerSys network “LungSys II” (grant numbers 0316042A, 0316042B, 0316042D) and within the framework of the e:Med research and funding concept – DeCaRe is also acknowledged (grant number 01ZX1409B). Irina Titkova and Magdalena Szczygjel were supported by the IMOMESIC project (grant number 82DZL00404, 82DZL00402), the German Center for Lung Research [Deutsches Zentrum für Lungenforschung (DZL), grant numbers 82DZL00404, 82DZL00402, 82DZL004A2]. Funding from the BMBF within the CancerSys network “LungSys II” (grant numbers 0316042A, 0316042B, 0316042D) and within the framework of the e:Med research and funding concept – DeCaRe is also acknowledged (grant number 01ZX1409B). Irina Titkova and Magdalena Szczygjel were supported by the IMOMESIC project within ERASysAPP network (grant number 031A604A), Melanie Boerries was supported by the DFG CRC 850 “Control of Cell Motility in Morphogenesis, Cancer Invasion and Metastasis” (grant number SFB/2) and Hauke Busch acknowledges support by the DFG excellence cluster 306 “Inflammation at Interfaces” (grant number EXC306). Jens Timmer and Marcus Rosenblatt were funded by MS_DILI (grant number 031L0074B).

Author Contributions
Immunoblot analysis, EMT gene expression, siRNA knockdowns, microscopy and survival experiments: D. Dvornikov. qPCR analysis of patient samples and survival analysis: M. Schneider and T. Muley. RNA-Seq data analysis: D. Dvornikov and S. Ohse. 3D invasion experiments: M. Szczygjel. 2D migration experiments: I. Titkova. Development of R script for analysis of 3D invasion data: M. Rosenblatt and M. Szczygieł. Access to clinical samples and corresponding clinical information: T. Muley, A. Warth, F. J. Herth, H. Diememann, M. Thomas. Writing the manuscript: D. Dvornikov, M.A. Schneider, S. Ohse. Study supervision: J. Timmer, M. Schilling, H. Busch, M. Boerries, M. Meister, and U. Klingmüller. All authors reviewed the manuscript.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-27912-1.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018