CD74-NRG1 fusions in lung adenocarcinoma

Lynnette Fernandez-Cuesta1,3, Dennis Plenker1,3, Hirotaka Osada2, Ruping Sun3, Roopika Menon4,5, Frauke Leenders1,6, Sandra Ortiz-Cuaron1, Martin Peifer1,7, Marc Bos1, Juliane Daßler8, Florian Malchers1, Jakob Schöttle1,9, Wenzel Vogel1, Ilona Dahmen1, Mirjam Koker1, Roland T. Ullrich9,10, Gavin M. Wright11, Prudence A. Russell12, Zoe Wainier13, Benjamin Solomon13, Elisabeth Brambilla14, Hélène Nagy-Mignotte15, Denis Morosibilot15, Christian G. Brambilla15, Sylvie Lantuejoul14, Janine Altmüller16,17, Christian Becker16, Peter Nürnberg16,17, Johannes M. Heuckmann4, Erich Stoelben18, Iver Petersen19, Joachim H. Clement20, Jörg Sänger21, Lucia A. Muscarella22, Annamaria la Torre22, Vito M. Fazio22,23, Ildoy Lahortiga24, Timothy Perera25, Souichi Ogata25, Marc Parade25, Dirk Brehmer25, Martin Vingron3, Lukas C. HeuKamp26, Reinhard Buettner6,26,27, Thomas Zander1,10,27, Jürgen Wolf6,10,27, Sven Perner6, Sascha Ansén10, Stefan A. Haas3, Yasushi Yatabe28 and Roman K. Thomas1,6,26

1 Department of Translational Genomics, Center of Integrated Oncology Köln-Bonn, University of Köln, 50924 Köln, Germany
2 Division of Molecular Oncology, Aichi Cancer Center Research Institute, 464-8681 Nagoya, Japan
3 Computational Molecular Biology Department, Max Planck Institute for Molecular Genetics, D-14195 Berlin, Germany
4 Blackfield AG, Gottfried-Hagen-Str. 60, 51105 Köln, Germany
5 Department of Prostate Cancer Research, Institute of Pathology, University Hospital of Bonn, 53127 Bonn, Germany
6 Laboratory of Translational Cancer Genomics, Center of Integrated Oncology Köln-Bonn, University of Köln, 50924 Köln, Germany
7 Center for Molecular Medicine Cologne (CMMC), University of Köln, Köln, Germany
8 Institute for Clinical Chemistry and Clinical Pharmacology, University Hospital Bonn, 53127 Bonn, Germany
9 Max Planck Institute for Neurological Research, D-50931 Köln, Germany
10 Department I of Internal Medicine, Center of Integrated Oncology Köln-Bonn, University of Köln, 50924 Köln, Germany
11 Department of Surgery, St Vincent’s Hospital, Melbourne, 3065 Victoria, Australia
12 Department of Pathology, St. Vincent’s Hospital, Melbourne, 3065 Victoria, Australia
13 Department of Haematology and Medical Oncology, Peter MacCallum Cancer Centre, Melbourne, 3002 Victoria, Australia
14 Department of Pathology, CHU Grenoble INSERM U823, Institute Albert Bonniot, Grenoble-Alpes University 38043 CS10217 Grenoble, France
15 CHU Grenoble INSERM U823, Institute Albert Bonniot 38043 CS10217 Grenoble, France
16 Cologne Center for Genomics (CCG), University of Köln, Köln, 50931 Germany

Downloaded from cancerdiscovery.aacrjournals.org on July 3, 2017. © 2014 American Association for Cancer Research.
17 Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Köln, Köln, Germany
18 Thoracic Surgery, Lungenklinik Merheim, Kliniken der Stadt Köln gGmbH, 51109 Köln, Germany
19 Institute of Pathology, Jena University Hospital, Friedrich-Schiller-University, 07743 Jena, Germany
20 Department of Internal Medicine II, Jena University Hospital, Friedrich-Schiller-University, 07743 Jena, Germany
21 Institute for Pathology Bad Berka, 99438 Bad Berka, Germany
22 Laboratory of Oncology IRCCS Casa Sollievo della Sofferenza, 71013 San Giovanni Rotondo, Italy
23 Centre for Integrated Biomedical Research, University Campus Bio-Medico, 00128 Rome, Italy
24 Center for the Biology of Disease, VIB, Leuven 3000, Belgium
25 Oncology Discovery, Janssen Research & Development, a division of Janssen Pharmaceutica NV, Turnhoutseweg 30, 2340 Beerse, Belgium
26 Department of Pathology, University Hospital Medical Center, University of Köln, 50937 Köln, Germany
27 Network Genomic Medicine, University Hospital Köln, Center of Integrated Oncology Köln – Bonn, 50924 Köln, Germany
28 Department of Pathology and Molecular Diagnostics, Aichi Cancer Center, Nagoya 464-8681 Japan

* These authors contributed equally to this work

Correspondence

Roman K Thomas, Department of Translational Genomics, University of Köln, Weyertal 115b, 50931 Köln, Germany, +49-221-478-98771, email: roman.thomas@uni-koeln.de

Running title

CD74-NRG1 fusions in lung adenocarcinoma

Keywords

Transcriptome sequencing, fusion gene, lung adenocarcinoma, neuregulin

Financial support

This work was supported by the EU-Framework Programme CURELUNG (HEALTH-F2-2010-258677 to EB, JW, and RKT); by the Deutsche Forschungsgemeinschaft through TH1386/3-1 (to RKT) and through SFB832 (TP5 to LCH; and TP6 to RTU, JW, and RKT); by the German Ministry of Science and Education (BMBF) as part of the NGFNplus program (grant 01GS08100 to JW and RKT); by the Deutsche Krebshilfe as part of the Oncology Centers of Excellence funding program (RB, JW, and RKT); by a Stand Up To Cancer Innovative Research Grant,
Program of the Entertainment Industry Foundation (SU2C-AACR-IRG60109 to RKT); by funds of the DFG Excellence Cluster ImmunoSensation to JD; by the Italian Ministry of Health (Ricerca Corrente RC1303LO57 and GR Program 2010-2316264) and by the “5x1000” voluntary contributions to LAM; by the Project for Development of Innovative Research on Cancer Therapeutics (P-Direct), Ministry of Education, Culture, Sports, Science and Technology of Japan to YY; by a research project grant (IWT 110431) to DB; by the Belgium government agency for Innovation by Science and Technology (IWT) to IL, SO, MP, TP, and DB; and by Agiradom and French Health Ministry (PPHRC) to CGB.

Disclosure of Potential Conflicts of Interest

LFC and RKT are inventors on a patent application related to the findings described in this manuscript. RKT, MP, and JMH are co-founders and shareholders of Blackfield AG. LFC, FL, MP, and FM received consulting fees from Blackfield AG. RKT received consulting and lecture fees (Sanofi-Aventis, Merck, Roche, Lilly, Boehringer-Ingelheim, Astra-Zeneca, Atlas-Biolabs, Daiichi-Sankyo, MSD, Puma, Blackfield AG) as well as research support (Merck, EOS and AstraZeneca). RB is a cofounder and advisor of Targos Molecular Pathology Inc, Kassel Germany and received honoraria for lectures and advisory board meetings from Roche, Novartis, Pfizer, Lilly, Qiagen, Merck Serono. JW received consulting and lecture fees from Roche, Novartis, Boehringer-Ingelheim, AstraZeneca, Bayer, Lilly, Merck, Amgen and research support from Roche, Bayer, Novartis, and Boehringer-Ingelheim. SO, TP, MP & DB are employees of Janssen Research & Development, a division of Janssen Pharmaceutica NV.
ABSTRACT

We discovered a novel somatic gene fusion, \textit{CD74-NRG1}, by transcriptome sequencing of 25 lung adenocarcinomas of never smokers. By screening 102 lung adenocarcinomas negative for known oncogenic alterations we found four additional fusion-positive tumors, all of which were of the invasive mucinous subtype. Mechanistically, CD74-NRG1 leads to extracellular expression of the EGF-like domain of NRG1 III-\beta3, thereby providing the ligand for ERBB2-ERBB3 receptor complexes. Accordingly, ERBB2 and ERBB3 expression was high in the index case and expression of phospho-ERBB3 was specifically found in tumors bearing the fusion (p<0.0001). Ectopic expression of CD74-NRG1 in lung cancer cell lines expressing ERBB2 and ERBB3 activated ERBB3 and the PI3K-AKT pathway, and led to increased colony formation in soft agar. Thus, \textit{CD74-NRG1} gene fusions are activating genomic alterations in invasive mucinous adenocarcinomas and may offer a therapeutic opportunity for a lung tumor subtype with, so far, no effective treatment.

SIGNIFICANCE

\textit{CD74-NRG1} fusions may represent a therapeutic opportunity for invasive mucinous lung adenocarcinomas, a tumor with no effective treatment that frequently presents with multifocal unresectable disease.

INTRODUCTION

Lung adenocarcinomas of patients who have never smoked frequently bear kinase gene alterations, such as \textit{EGFR} mutations and translocations affecting \textit{ALK}, \textit{ROS1}, and \textit{RET} genes (1–6). These alterations cause “oncogene dependency” on the activated kinase and thus, sensitivity of the tumor cells to kinase inhibitors. Patients whose tumors bear kinase gene alterations can be effectively treated with an ever-growing number of kinase inhibitors; for example, patients with \textit{EGFR}-mutant lung cancer treated with \textit{EGFR} inhibitors have a significantly longer progression-free survival compared to patients treated with conventional chemotherapy (7). Similarly, \textit{ALK} and \textit{ROS1} inhibition induces clinically relevant remissions in patients bearing the respective genomic fusion (8–10). Unfortunately, despite substantive cancer genome sequencing efforts, a majority of lung tumors still lacks therapeutically tractable kinase alterations (1). We therefore sought to identify novel therapeutically relevant \textit{driver} alterations in otherwise \textit{driver}-negative lung adenocarcinomas.
RESULTS

We collected a cohort of 25 lung adenocarcinoma specimens of never smokers that lacked mutations in KRAS or EGFR, in which we performed chromosomal gene copy number analysis as well as transcriptome sequencing with the aim of identifying new oncogenic driver alterations. We applied a novel computational data analysis strategy that combines split-read and read-pair analyses with denovo assembly of candidate regions containing potential breakpoints to achieve sensitive and accurate detection of fusion transcripts (See Methods; Fernandez-Cuesta et al., will be published elsewhere). Of the 25 samples analyzed (Supplementary Table S1), ten carried a known oncogene. One sample exhibited EGFR amplification, paralleled by overexpression of the gene (Fig. 1A, Supplementary Fig. S1). We also found three cases each of ALK, ROS1 and RET fusions (Fig. 1A, Supplementary Table S2). In addition, we detected one sample carrying a novel chimeric transcript fusing the first six exons of CD74 to the exons encoding the EGF-like domain of the NRG1 III-β3 isoform (Fig. 1A and B, Supplementary Table S2). This fusion raised our interest because CD74 is part of recurrent fusions affecting the ROS1 (3) kinase in lung adenocarcinoma, and because NRG1 encodes a ligand of ERBB receptor tyrosine kinases, which are also frequently affected by genome alterations in this tumor type. NRG1 provides the ligand for ERBB3 and ERBB4 receptors (11). The NRG1 isoform present in our fusion transcript belongs to the type III and carries the EGF-like domain type β, which has higher affinity to the receptors than the α-type (12). NRG1 type III expression is mostly limited to neurons and is the only isoform displaying this degree of tissue-specific expression (13). Only the sample carrying the CD74-NRG1 fusion exhibited high expression of the NRG1 III-β3 isoform (74 fragments per kilobase per million reads, FPKM) (Fig. 1C, upper panel; Supplementary Table S3), and in this specimen, there was no expression of the wild-type allele (Fig. 1C, lower panel). In addition, NRG1 was generally not expressed in lung adenocarcinoma as shown by transcriptome sequencing data of our cohort of 25 lung adenocarcinomas of never smokers (Fig. 1C, Supplementary Table S3), and of a cohort of 15 unselected lung adenocarcinomas (Fig. 1C, Supplementary Table S4). The fusion resulted from a somatic genomic event as CD74-NRG1 fusion FISH and NRG1 break-apart FISH revealed rearrangements in the respective chromosomal regions in the tumor cells, but not in surrounding non-tumoral cells (Fig. 1D, Supplementary Fig. S2). Furthermore, by applying hybrid-capture-based massively parallel genome sequencing (Fig. 1D, Supplementary Table S5), we found 5 and 2 reads spanning and encompassing the chromosomal breakpoint (chr5:149,783,493 and chr8:32,548,502), respectively.
We next performed reverse-transcriptase polymerase chain reaction (RT-PCR) using primers specific for the chimeric transcript to identify additional tumors bearing the fusion in a set of 102 pan-negative adenocarcinomas of never smokers (wild type for *EGFR, KRAS, BRAF, ERBB2, ALK, ROS*, and *RET* genes). We identified 4 additional tumors carrying the fusion (**Supplementary Table S6**), which were also confirmed by break-apart FISH. All 5 cases (including the index case) occurred in invasive mucinous lung adenocarcinomas (IMA) of women who had never smoked (**Fig. 2A**). Invasive mucinous lung adenocarcinoma is highly associated with *KRAS* mutations (14). Indeed, out of 15 invasive mucinous lung adenocarcinoma specimens (all derived from an East Asian population), 6 carried a *KRAS* mutation (40%), and 4 carried the *CD74-NRG1* fusion (27%) (**Fig. 2B, Supplementary Table S7**). We additionally tested other lung tumor subtypes (63 cases), as well as 4 other cancer types (21 cases) and all were negative for the fusion gene (**Supplementary Table S6**) suggesting a strong link between presence of *CD74-NRG1* and invasive mucinous adenocarcinoma.

Characteristic features of type III NRG1 are cytosolic N-termini and membrane-tethered EGF-like domains (13, 15). In the case of CD74-NRG1, the part of CD74 is predicted to replace the transmembrane domain present in wild-type NRG1 III-β3, preserving the membrane-tethered EGF-like domain (**Fig. 2C**). To validate this prediction, we transduced NIH-3T3 cells with CD74-NRG1-encoding retroviruses, and performed flow cytometry analyses to determine the subcellular distribution of expression of the fusion protein. As expected, we observed a positive intracellular (but not extracellular) signal for CD74 (**Fig. 2D, left panel**), and a positive extracellular signal for NRG1 (**Fig. 2D, right panel**). Similar results were observed in H2052 cells (**Supplementary Fig. S3**). Furthermore, we were unable to detect the fusion in the supernatant of transduced cells with a polyclonal antibody raised against the EGF-like domain (data not shown). Thus, the fusion does not lead to secretion of the EGF-like domain, but probably generates a membrane-bound protein with the EGF-like domain presented on the outside of the cell.

We next analyzed the expression of ERRB receptors in the index case: *ERBB1 (EGFR)* was almost not expressed (FPKM=1.9, **Fig. 3A, Supplementary Table S8, Supplementary Fig. S4**) and neither phosphorylated (**Supplementary Fig. S4**). By contrast, *ERBB2* was expressed (FPKM=22.9, **Fig. 3A, Supplementary Table S8**) and phosphorylated (**Fig. 3B, left panel**); similar to *ERBB2, ERBB3* was also expressed at relatively high levels (FPKM=22.8, **Fig. 3A, Supplementary Table S8**) and also phosphorylated (**Fig. 3B, right panel**). *ERBB4* was not expressed in the index case (FPKM=0.2, **Fig. 3A, Supplementary Table S8**). To our surprise, expression of p-ERBB3 was almost exclusively restricted to fusion-positive cases as determined by an immunohistochemical analysis of a tissue microarray containing 241 unselected adenocarcinomas. While a positive signal was detected for p-ERBB3 in the 5 CD74-NRG1-positive invasive mucinous adenocarcinomas, only 6 of the 241 unselected...
adenocarcinomas exhibited detectable levels of p-ERBB3 (p<0.0001, Fig. 3C). Together, these observations support the notion that CD74-NRG1 might provide the ligand for ERBB2-ERBB3 heterodimers, thus activating the PI3K-AKT pathway, as previously shown for wild-type NRG1 (16).

To formally test this hypothesis we transduced different cell lines with retroviruses encoding CD74-NRG1 and performed western-blots under starving conditions. Since NIH-3T3 cells have low-to-absent expression of ERBB receptors, and NIH-3T3 cells ectopically expressing ERBB2 and ERBB3 are already oncogenic (Supplementary Fig. S5), we decided to use H322 and H1568 lung cancer cell lines expressing normal ERBB2 and ERBB3 levels instead. We transduced these cell lines either with an empty vector, a virus containing the full fusion transcript, or a virus containing a truncated version of the fusion lacking the EGF-like domain (Supplementary Fig. S6). We observed that H322 and H1568 cell lines ectopically expressing CD74-NRG1 showed increased levels of p-ERBB2, p-ERBB3, p-AKT, and p-S6K when compared with the empty vector control (Fig. 3D). Furthermore, both p-ERBB3 and p-AKT depended on the presence of the EGF-like domain of CD74-NRG1 in the fusion (Fig. 3E). In addition, co-culture of NIH-3T3 cells ectopically expressing CD74-NRG1 with Ba/F3 cells genetically engineered to express normal ERBB2 and ERBB3 levels, also led to activation of AKT (Supplementary Fig. S7). Finally, H1568 cells ectopically expressing CD74-NRG1 exhibited enhanced colony formation in soft agar assays (Fig. 3F, Supplementary Table S9). Taken together, these data suggest that CD74-NRG1 leads to overexpression of the EGF-like domain of NRG1 III-β3 that acts as a ligand for ERBB3 inducing its phosphorylation and subsequent activation of the downstream PI3K-AKT pathway.

DISCUSSION

We have discovered CD74-NRG1, a novel recurrent fusion gene in lung adenocarcinoma that arises from a somatic genomic event. Taking into account the frequencies of mutations of EGFR (11.3%), KRAS (32.2%), BRAF (7%), ERBB2 (1.7%), or fusions affecting ALK (1.3%), ROS (1.7%), and RET (0.9%), (17, 18) for which our cohort was negative, and the fact that we found 4 positive cases in our validation cohort of 102 pan-negative lung adenocarcinomas, we estimate that the frequency of CD74-NRG1 in lung adenocarcinomas is approximately 1.7%; however, it is of note that our validation cohort was from an Asian population, so this frequency might be different in Caucasians. CD74-NRG1 occured specifically in invasive mucinous lung adenocarcinomas of never smokers, a tumor type that is otherwise associated with KRAS mutations (14). In our cohort of limited size (n=15), CD74-NRG1 fusions accounted for 27% of invasive mucinous lung adenocarcinoma; together KRAS mutations and CD74-NRG1 may therefore be considered the causative oncogenes in more than 60% of the cases.
We provide evidence that CD74-NRG1 signals through induction of ERBB2-ERBB3 heterodimers thus leading to PI3K-AKT pathway activation and stimulation of oncogenic growth. In light of the multitude of available drugs targeting ERBB2, ERBB3, and their downstream pathways (19), CD74-NRG1 fusions may represent a therapeutic opportunity for invasive mucinous lung adenocarcinomas, which frequently present with multifocal and unresectable disease, and for which no effective treatment exists.

**METHODS**

**Sample preparation, DNA, RNA extraction, and Illumina Sequencing**

Sample preparation, DNA and RNA extraction was performed as previously described (20). RNAseq was performed on cDNA libraries prepared from PolyA+ RNA extracted from tumor cells using the Illumina TruSeq protocol for mRNA. The final libraries were sequenced with a paired-end 2×100 bp protocol aiming at 8.5 Gb per sample, resulting on a 30x mean coverage of the annotated transcriptome. All the sequencing was carried on an Illumina HiSeq™ 2000 sequencing instrument (Illumina, San Diego, CA, USA).

**Analysis of chromosomal gene copy number (SNP 6.0) and RNAseq data**

Hybridization of the Affymetrix SNP 6.0 arrays was carried out according to the manufacturers' instructions and analyzed using previously described method (20). For the analysis of RNAseq data, we have developed a pipeline that affords accurate and efficient mapping and downstream analysis of transcribed genes in cancer samples (Fernandez-Cuesta et al., will be published elsewhere). A brief description of the method was previously provided (20).

**Analysis of targeted enrichment genome sequencing**

Genomic DNA was isolated from fresh frozen tumor tissue and subjected to CAGE Scanner analysis. This approach involves liquid-phase hybrid capture of genomic partitions enriched for genome alterations affecting 333 known cancer-associated genes (also including CD74). Subsequent to generation of genomic libraries from tumor DNA and capture, sequencing was performed on the Illumina platform according to the manufacturers instructions. Significant genomic alterations were identified using approaches described previously (20).

**Dideoxy Sequencing**
In case of validation, sequencing primer pairs were designed to enclose the putative mutation, or to encompass the candidate rearrangement or chimeric transcript as previously described (20). Sequencing was carried out and electropherograms were analyzed by visual inspection using 4 Peaks.

Interphase FISH on FFPE sections
Two sets of probes were prepared. One was for break-apart FISH of which probes were mapped at centromeric and telomeric regions between the break point. The other was for fusion FISH that spanned the NRG1 and CD74 loci. To intensify the signals, each probe was made of two or three BAC clones as follows, and the probes were labeled with SpectrumGreen and SpectrumOrange (Abbott Molecular-Vysis). Centromeric probes for break-apart FISH: RP11-1002K11 and PR11-25D16. Telomeric probes for break-apart FISH: RP11-23A12 and PR11-715M18. NRG1 probes for fusion FISH: RP11-715H18, RP11-5713 and PR11-1002K11. CD74 probes for fusion FISH: PR11-759G10 and PR11-468K14.

Immunohistochemistry (IHC)
IHC was performed as previously described (21). In brief, the tissue samples were stained with p-ERBB2 (Tyr1221/1222, Cell Signaling Technology, USA) and total ERBB1 (EGFR) (Dako, Germany) at a dilution of 1:1000 and 1:50 respectively. The Zeiss MIRAK DESK scanner was used to digitize the stained tissue. Staining for p-EGFR (Tyr1068, Cell Signaling Technology, USA) and p-ERBB3 (Tyr1289, Cell Signaling Technology, USA) were processed with an automated stainer (Autostainer, Dako Copenhagen, Denmark), using FLEX+ detection system (Dako).

Cell culture
H2052, H322, and H1568 cells were obtained from American Type Culture Collection (ATCC) and maintained in RPMI-1640 medium (Life Technologies) supplemented with 10% fetal calf serum (FCS) (Gibco) and 1% penicillin-streptomycin (Gibco). The cells were cultured in a humidified incubator with 5% CO2 and 37°C. For western-blots experiments cells were serum starved without FCS for 24h. NIH-3T3 cells were maintained similarly but in DMEM medium (Life Technologies). The cells were confirmed to be wild type for KRAS, EGFR, ERBB2, and ERBB3 by PCR amplification followed by Sanger sequencing of the PCR products. The cell lines have been authenticated via genotyping (SNP 6.0, Affymetrix) and tested for mycoplasma contamination on a regular basis (MycolAlert, Lonza).
FACS analysis

NIH-3T3 mouse fibroblast cells were transduced with retrovirus containing empty-vector, CD74-NRG1, ERBB2, ERBB3 and ERBB2+ERBB3. H2052 cells were transduced with retrovirus containing empty-vector or CD74-NRG1. Transduced cells (200,000) were washed in FACS-Buffer (PBS, 2% FCS) and fixed in 4% PFA for 30 min at room temperature. For permeabilization, cells were washed twice in Saponin-Buffer (PBS, 0.5% Saponin, 2% FCS) and intracellular staining of CD74-NRG1 was performed with anti-human-CD74-PE (1:100) (BioLegend). Intracellular staining of ERBB2 and ERBB3 was performed with anti-ERBB2 and anti-ERBB3 antibodies (1:50) (Cell Signaling). Binding of ERBB2 or ERBB3 was detected with goat-anti-rabbit-Alexa488 (Life Technologies). Extracellular staining was performed prior permeabilization with anti-human-CD74-PE and anti-NRG1 antibody (1:20) (R&D Systems). Binding of the NRG1 part was detected with donkey-anti-goat-Alexa488 (Life Technologies). Subsequently, cells were analyzed on a BD LSR II (Beckman Coulter) and quantification was assessed with FlowJo (Treestar).

Western-blot

Immunoblotting was performed using standard procedures. The following antibodies were obtained from Cell Signaling Technology: p-AKT Ser473 (Catalog No. #9271), p-P70/S6 (Catalog No. #9205), total ERBB2 (Catalog No. #2242), p-ERBB2 (Catalog No. #2243), total ERBB3 (Catalog No. #4754), and p-ERBB3 (Catalog no. #4791). Anti human CD74 was obtained from Abcam (Catalog No. # ab22603), anti polyclonal NRG1 beta 1 was obtained from R&D Systems (Catalog No. AF396-NA). Actin-HRP antibody was obtained from SantaCruz (Catalog No. sc47778). The antibodies were diluted in 5% BSA/ TBST and incubated at 4°C overnight. Proteins were detected with HRP-conjugated anti-mouse, anti-goat or anti-rabbit antibodies (Millipore) using ECL reagent (GE Healthcare).

Colony formation assay

On a layer of bottom agar (1%) NIH-3T3 cells were suspended at low density in top agar (0.5%) containing 10% FCS, and were grown for 14 days. Subsequently pictures were taken and systematic analyses were performed with the Scanalyzer (LeuMaTec). H1568 cells were cultured under standard conditions in RPMI in 10% FCS and 1% P/S. pBabe retroviral vector inserts were confirmed via Sanger-sequencing. The cells were generated by at least two independent transducitons with retrovirus containing empty-vector, CD74-NRG1 or CD74-NRG1_AEGF. After selection for 7 days with puromycin (3µg/ml), cell lysates were taken for western blot analysis, and cells were also used for colony formation assays as follows: on a layer of bottom agar (1.2%) cells were
suspended at low density in top agar (0.6%) containing 10 % FCS (final concentration), and were grown for 14 days. Subsequently pictures were taken with a Zeiss Axiocam 40 CFL microscope at 100x magnification and colony size was assessed with ImageJ (http://rsbweb.nih.gov/ij/).

**Generation of Ba/F3_ERBB2+ERBB3 cells**

The ERBB2 and ERBB3 open reading frames were amplified by PCR and cloned into the MSCV-puroycin or MSCV-neomycin vectors respectively (ClonEtech). Ba/F3 cells expressing ERBB2 and ERBB3 were generated by retroviral transduction and subsequent puromycin or/neomycin selection. We verified the expression of the correct proteins by western blot. Ba/F3 cells were cultured in RPMI-1640 medium supplemented with 10 % FBS and 1 ng/ml mouse IL3.

**Statistical analyses**

In Fig. 3C and 3F, we used a two-tailed Fisher's exact test.

**Data accession number**

EGAS00001000653

**Acknowledgements**

We are indebted to the patients donating their tumor specimens as part of the Clinical Lung Cancer Genome Project initiative. Additional biospecimens for this study were obtained from the Victorian Cancer Biobank, Melbourne, Australia. The Institutional Review Board (IRB) of each participating institution approved collection and use of all patient specimens in this study. We thank Philipp Lorimier, Marek Franitza, Graziella Bosco, and Juan Luis Fernandez Mendez de la Vega for their technical assistance. We furthermore thank the regional computing center of the University of Köln (RRZK) for providing the CPU time on the DFG-funded supercomputer ‘CHEOPS’ as well as the support.
References


Figure Legends

Figure 1. Identification of CD74-NRG1 fusion gene. **A.** Overview of driver genes detected in a cohort of 25 EGFR- and KRAS-negative lung adenocarcinomas of never smokers. **B.** Detection of CD74-NRG1 fusion transcript by transcriptome sequencing. Schematic representation of the fusion transcript domains and some of the transcriptome sequencing reads spanning the fusion point. **C.** Expression levels of NRG1 isoforms in 15 unselected and 23 pan-negative lung adenocarcinomas (AD) (wild type for EGFR, KRAS, BRAF, ERBB2, ALK, ROS, and RET genes), and in the index case, inferred from transcriptome sequencing data. Average FPKM values are shown (upper panel). RNaseq analysis for NRG1 reads to show where the breakpoint of CD74-NRG1 occurs. The dip in exon 4 represents reads of the fusion that could not be mapped. No reads could be mapped to exons 1-3 (lower panel). **D.** The upper part shows the genomic intron/exon structure of the CD74 (in green) and the NRG1 locus (in orange) with the genomic breakpoints marked in red. Sequencing reads were obtained from hybrid-capture-based genomic sequencing of 333 genes employing genomic DNA of the index case (See Methods). The breakpoint-spanning reads are shown by means of the Integrative Genomics Viewer (www.broadinstitute.org/igv/) focused on the CD74 gene (lower part). The grey area of the read is aligned to the CD74 reference sequence. Colored area on the right indicates bases not matching the CD74 reference sequence. Sequence comparison reveals alignment to the NRG1 reference sequence. Encompassing reads whose mate pairs are mapped to the NRG1 locus on chromosome 8 are displayed in dark purple. The lower left panel shows a representative picture of NRG1 break-apart FISH. Arrows show break-apart signals.

Figure 2. Association of CD74-NRG1 with invasive mucinous adenocarcinoma (IMA), and membrane localization of the fusion protein. **A.** Clinical characteristics of the index case and the 4 additional cases found to harbor CD74-NRG1. **B.** Frequency of KRAS-mutations and CD74-NRG1 rearrangements in a cohort of 15 IMA tumors (East Asian population). **C.** Schematic representation of wild-type NRG1 III-β3 and predicted CD74-NRG1 fusion protein in the cellular membrane. **D.** Intracellular and extracellular staining of CD74 (left panel), and extracellular staining of NRG1 (right panel) in CD74-NRG1 transduced NIH-3T3 cells, detected by flow cytometry. The % of Max is the number of cells in each bin divided by the number of cells in the bin that contains the largest number of cells.

Figure 3. Functional relevance of CD74-NRG1. **A.** Expression levels of ERBB receptors in the index case inferred from transcriptome sequencing data. FPKM values are shown. **B.** Levels of p-ERBB2 and p-ERBB3 detected by immunohistochemical analysis in a CD74-NRG1 positive case using and antibody directed against...
Tyr1221/1222 and Tyr1289, respectively. C. The same p-ERBB3 antibody was used to stain a tissue microarray composed of 241 lung adenocarcinomas. The frequency of p-ERBB3 positive cases in this cohort versus the 5 CD74-NRG1 positive samples is shown (p-value<0.0001). D. Activation of PI3K-AKT pathway detected by western-blot of H322 and H1568 lung cancer cells transduced with retroviruses encoding CD74-NRG1 or the empty vector control (e.v). E. Levels of p-ERBB3 and p-AKT measured by western-blot in the presence of an empty vector (e.v.), CD74-NRG1, or a truncated version lacking the EGF-like domain (CD74-NRG1 ΔEGF). F. Anchorage independent growth of H1568 cells expressing an empty vector (e.v.), CD74-NRG1, or a truncated version lacking the EGF-like domain (CD74-NRG1 ΔEGF). The upper panel shows the average colony size for the three conditions, with error bars representing standard deviations. The experiment was performed with two independent transductions for a total of 4 times. P-values are indicated by ** if p-value<0.01, and *** if p-value<0.001. The lower panel shows representative pictures of the colony formation assay. Please note that H1568 are oncogenic and form small colonies without any manipulation.
Table 1

<table>
<thead>
<tr>
<th>Sample</th>
<th>Age</th>
<th>Sex</th>
<th>Stage</th>
<th>Driver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case-01</td>
<td>63</td>
<td>female</td>
<td>Ia</td>
<td>EGFR-amp</td>
</tr>
<tr>
<td>Case-08</td>
<td>68</td>
<td>male</td>
<td>Ia</td>
<td>EML4-ALK</td>
</tr>
<tr>
<td>Case-10</td>
<td>46</td>
<td>female</td>
<td>Ia</td>
<td>EML4-ALK</td>
</tr>
<tr>
<td>Case-17</td>
<td>39</td>
<td>male</td>
<td>IIIa</td>
<td>EML4-ALK</td>
</tr>
<tr>
<td>Case-06</td>
<td>56</td>
<td>female</td>
<td>IIIb</td>
<td>CD74-ROS1</td>
</tr>
<tr>
<td>Case-09</td>
<td>60</td>
<td>female</td>
<td>Ia</td>
<td>CD74-ROS1</td>
</tr>
<tr>
<td>Case-02</td>
<td>68</td>
<td>female</td>
<td>IIIb</td>
<td>EZR-ROS1</td>
</tr>
<tr>
<td>Case-15</td>
<td>65</td>
<td>female</td>
<td>Ila</td>
<td>KIF5B-RET</td>
</tr>
<tr>
<td>Case-25</td>
<td>75</td>
<td>male</td>
<td>la</td>
<td>KIF5B-RET</td>
</tr>
<tr>
<td>Case-23</td>
<td>66</td>
<td>female</td>
<td>IIIb</td>
<td>CCDC6-RET</td>
</tr>
<tr>
<td>Case-19*</td>
<td>64</td>
<td>female</td>
<td>lb</td>
<td>CD74-NRG1</td>
</tr>
<tr>
<td>Case-03</td>
<td>72</td>
<td>female</td>
<td>IV</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-04</td>
<td>74</td>
<td>male</td>
<td>IV</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-05</td>
<td>50</td>
<td>female</td>
<td>IV</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-07</td>
<td>70</td>
<td>female</td>
<td>la</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-11</td>
<td>79</td>
<td>female</td>
<td>Ila</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-12</td>
<td>72</td>
<td>male</td>
<td>IV</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-13</td>
<td>59</td>
<td>male</td>
<td>lb</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-14</td>
<td>63</td>
<td>female</td>
<td>IIIa</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-16</td>
<td>48</td>
<td>male</td>
<td>IIIb</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-18</td>
<td>71</td>
<td>female</td>
<td>IIIa</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-20</td>
<td>74</td>
<td>female</td>
<td>IIIa</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-21</td>
<td>73</td>
<td>male</td>
<td>lb</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-22</td>
<td>80</td>
<td>female</td>
<td>lb</td>
<td>Unknown</td>
</tr>
<tr>
<td>Case-24</td>
<td>66</td>
<td>female</td>
<td>IIIa</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

*Index case (Caucasian, Invasive Mucinous Adenocarcinoma)
Figure 2

<table>
<thead>
<tr>
<th>Sample</th>
<th>Age</th>
<th>Sex</th>
<th>Stage</th>
<th>Smoking status</th>
<th>AD subtype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index-case</td>
<td>64</td>
<td>female</td>
<td>I b</td>
<td>never</td>
<td>Invasive mucinous</td>
</tr>
<tr>
<td>Case-A</td>
<td>73</td>
<td>female</td>
<td>I a</td>
<td>never</td>
<td>Invasive mucinous</td>
</tr>
<tr>
<td>Case-B</td>
<td>72</td>
<td>female</td>
<td>I a</td>
<td>never</td>
<td>Invasive mucinous</td>
</tr>
<tr>
<td>Case-C</td>
<td>66</td>
<td>female</td>
<td>I a</td>
<td>never</td>
<td>Invasive mucinous</td>
</tr>
<tr>
<td>Case-D</td>
<td>31</td>
<td>female</td>
<td>I a</td>
<td>never</td>
<td>Invasive mucinous</td>
</tr>
</tbody>
</table>

EGFR, KRAS, BRAF, HER2, ALK, ROS, RET negative

B

Unknown 33%
KRAS mut 40%
CD74-NRG1 27%
15 Invasive Mucinous Adenocarcinomas

C

CD74-NRG1

NRG1 III-β3

D

Intracellular
Extracellular

% of Max

0 10^2 10^3 10^4

0 10 10^2 10^3 10^4

0 10 10^2 10^3 10^4

anti-CD74

anti-NRG1
CD74-NRG1 fusions in lung adenocarcinoma
Lynnette Fernandez-Cuesta, Dennis Plenker, Hirotaka Osada, et al.

Cancer Discovery  Published OnlineFirst January 27, 2014.

Updated version  Access the most recent version of this article at:
doi:10.1158/2159-8290.CD-13-0633

Supplementary Material  Access the most recent supplemental material at:
http://cancerdiscovery.aacrjournals.org/content/suppl/2014/01/27/2159-8290.CD-13-0633.DC1

Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.