

### **Translational Relevance**

Oncogene fusions, such as the *ALK*, *RET*, and *ROS1* fusions, have recently been revealed as therapeutic targets in lung adenocarcinoma. We identified multiple druggable oncogene fusions, including those involving the *NRG1*, *ERBB4*, and *BRAF* genes, in invasive mucinous adenocarcinoma (IMA), a malignant type of lung adenocarcinoma. The fusions occurred mutually exclusively with *KRAS* mutations, a common driver oncogene aberration in IMA. These fusions represent potentially clinically relevant targets for treatment of IMAs that lack *KRAS* mutations.

## Abstract

**Purpose:** To identify druggable oncogenic fusions in invasive mucinous adenocarcinoma (IMA) of the lung, a malignant type of lung adenocarcinoma in which *KRAS* mutations frequently occur.

**Experimental Design:** From an IMA cohort of 90 cases, consisting of 56 cases (62%) with *KRAS* mutations and 34 cases without (38%), we conducted whole-transcriptome sequencing of 32 IMAs, including 27 cases without *KRAS* mutations. We used the sequencing data to identify gene fusions, and then performed functional analyses of the fusion gene products.

**Results:** We identified oncogenic fusions that occurred mutually exclusively with *KRAS* mutations: *CD74-NRG1*, *SLC3A2-NRG1*, *EZR-ERBB4*, *TRIM24-BRAF*, and *KIAA1468-RET*. *NRG1* fusions were present in 17.6% (6/34) of *KRAS*-negative IMAs. The *CD74-NRG1* fusion activated HER2:HER3 signaling, whereas the *EZR-ERBB4* and *TRIM24-BRAF* fusions constitutively activated the ERBB4 and BRAF kinases, respectively. Signaling pathway activation and fusion-induced anchorage-independent growth/tumorigenicity of NIH3T3 cells expressing these fusions were suppressed by tyrosine kinase inhibitors approved for clinical use.

**Conclusions:** Oncogenic fusions act as driver mutations in IMAs without *KRAS* mutations, and thus represent promising therapeutic targets for the treatment of such IMAs.

## Introduction

Oncogene fusions have recently been identified as driver mutations and (possible) therapeutic targets in lung adenocarcinoma (LADC), a major histological type of lung cancer (1). Such fusions include *EML4-* or *KIF5B-ALK*, *KIF5B* or *CCDC6-RET*, and *CD74-*, *EZR-*, or *SLC34A2-ROS1* (2-9). These oncogene fusions occur mutually exclusively with one another, and with other targetable oncogene aberrations such as *EGFR*, *KRAS*, *BRAF*, and *HER2* mutations. Therefore, molecular targeted therapy combined with the identification of driver oncogene aberrations represents a powerful and promising approach to personalized treatment of LADC (10, 11).

Invasive mucinous adenocarcinoma (IMA) of the lung is composed predominantly of goblet cells. IMA is morphologically characterized by tall columnar cells with basal nuclei and a pale cytoplasm containing varying amounts of mucin (12, 13). IMAs, which constitute 2–10% of all LADCs in Japan, the USA, and European countries (14-16), are indicated as being more malignant than more common types of LADC, such as acinar or papillary adenocarcinoma. The *KRAS* mutation is the only driver aberration commonly detected in IMAs (in 50–80% of cases). To date, no driver gene aberrations have been detected in *KRAS*-negative IMAs; these aberrations must be identified to facilitate the development of effective treatments for such cancers. Therefore, we performed whole-transcriptome sequencing (RNA sequencing) of IMAs lacking *KRAS* mutations to identify novel chimeric fusion transcripts that represent potential targets for cancer therapy.

## **Materials and Methods**

### **Samples**

Ninety IMAs were identified among consecutive patients with primary adenocarcinoma of the lung who were treated surgically at the National Cancer Center Hospital, Tokyo, Japan, from 1998 to 2013. Histological diagnoses were based on the most recent World Health Organization classification and the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society (IASLC/ATS/ERS) criteria for LADC (13, 17). Total RNA was extracted from grossly dissected, snap-frozen tissue samples using TRIzol (Invitrogen, Carlsbad, CA, USA). The study was approved by the Institutional Review Boards of the participating institutions.

### **RNA sequencing**

RNA sequencing libraries were prepared from 1 or 2  $\mu$ g of total RNA using the mRNA-Seq Sample Prep Kit or TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA, USA). The resultant libraries were subjected to paired-end sequencing of 50 or 75 bp reads on a Genome Analyzer IIx (GAIIx) or HiSeq 2000 (Illumina). Fusion transcripts were detected using the TopHat-Fusion algorithm (18). Experimental conditions for RNA sequencing are described in **Supplementary Table 1**.

### **Examinations of oncogenic properties of fusion products.**

To construct lentiviral vectors for expression of the CD74-NRG1, EZR-ERBB4 and TRIM24-BRAF fusion proteins, full-length cDNAs were amplified from

tumor cDNA by PCR and inserted into pLenti-6/V5-DEST plasmids (Invitrogen). The integrity of each inserted cDNA was verified by Sanger sequencing. Expression of fusion products of the predicted sizes was confirmed by western blot analysis of transiently transfected and virally infected cells (**Supplementary Fig. 1A**). Details of plasmid transfection, viral infection, western blot analysis, and soft agar colony and tumorigenicity assays are described in **Supplementary Materials and Methods**.

### **Results and discussion**

We prepared an IMA cohort of 90 cases consisting of 56 (62%) cases with *KRAS* mutations and 34 (38%) cases without. The 34 *KRAS*-negative cases included two, one, and one cases with *BRAF* mutation, *EGFR* mutation, and *EML4-ALK* fusion, respectively; the remaining 30 were “pan-negative” for representative driver aberrations in LADCs. Thirty-two cases, consisting of 27 pan-negative and 5 *KRAS* mutation-positive cases, were subjected to RNA sequencing (**Supplementary Table 1**). Analysis of  $>2 \times 10^7$  paired-end reads obtained by RNA sequencing and subsequent validation by Sanger sequencing of reverse transcription (RT)-PCR products revealed five novel gene-fusion transcripts detected only in the pan-negative IMAs: *CD74-NRG1*, *SLC3A2-NRG1*, *EZR-ERBB4*, *TRIM24-BRAF*, and *KIAA1468-RET* (**Figs. 1A–B and Table 1**; details in **Supplementary Materials and Methods, Supplementary Fig. 2, and Supplementary Table 2**). RT-PCR screening of these fusions in the remaining 58 IMAs that had not been subjected to RNA sequencing revealed one additional pan-negative case with the *CD74-NRG1* fusion. Thus, the *CD74-NRG1* fusion,

detected in 5/34 (14.7%) of cases negative for *KRAS* mutations, was the most frequent fusion among *KRAS* mutation-negative IMAs. Fusions of *CD74* or *SLC3A2* with *NRG1* were present in 17.6% (6/34) of cases. The five novel fusions were mutually exclusively with one another and were not present in any of the *KRAS* mutation-positive cases (**Table 2**).

Four of the novel fusions, *CD74-NRG1*, *SLC3A2-NRG1*, *EZR-ERBB4*, and *TRIM24-BRAF*, involved rearrangements of genes encoding protein kinases or a ligand of a receptor protein kinase (*NRG1*/neuregulin/heregulin) for which oncogenic rearrangements have not been previously reported in lung cancer (**Supplementary Fig. 3**). The remaining fusion was a novel type involving the *RET* oncogene; fusions with *RET* are observed in 1–2% of LADCs (4, 5, 7, 8, 11). In a screen of 315 LADCs without IMA features from Japanese patients and 144 consecutive LADCs from US patients, all tumors were negative for all of the *NRG1*, *BRAF*, and *ERBB4* fusions, as well as the novel *RET* fusion. Therefore, these fusions might be driver aberrations specific to LADCs with IMA features. The four novel gene fusions were likely to have been caused by inter-chromosomal translocations or paracentric inversion (**Table 1**, **Supplementary Fig. 3**). Consistently, separation of the signals generated by the probes flanking the translocation sites of *NRG1* in fusion-positive tumors was observed upon fluorescence *in situ* hybridization (FISH) analysis of *CD74-NRG1* fusion positive tumors (**Supplementary Fig. 4**). We also confirmed over-expression of *NRG1*, *ERBB4*, and *BRAF* proteins in tumor cells carrying the corresponding fusions by immunohistochemical analysis, using antibodies recognizing polypeptides retained in the fusion proteins; expression of *NRG1*,

ERBB4, and BRAF proteins was also observed in some fusion-negative cases (**Supplementary Fig. 5**). IMAs harboring gene fusions were obtained from both male and female patients, although *NRG1* fusion-positive cases were preferentially from female never smokers (**Table 1**).

The CD74-*NRG1* and SLC3A2-*NRG1* fusion proteins, whose sequences were deduced from RNA sequencing data, contained the CD74 or SLC3A2 transmembrane domain and retained the epidermal growth factor (EGF)-like domain of the *NRG1* protein (*NRG1* III- $\beta$ 3 form) (**Fig. 1A**). The *NRG1* III- $\beta$ 3 protein has a cytosolic N-terminus and a membrane-tethered EGF-like domain, and mediates juxtacrine signals signaling through HER2:HER3 receptors (19). Because parts of CD74 or SLC3A2 replaced the transmembrane domain of wild-type *NRG1* III- $\beta$ 3, we speculated that the membrane-tethered EGF-like domain might activate juxtacrine signaling through HER2:HER3 receptors. In addition, it was also possible that expression of these fusion proteins resulted in the production of soluble *NRG1* protein due to proteolytic cleavage at sites derived from *NRG1* (dashed green lines in **Fig. 1A**), as recently suggested for *NRG1* type III proteins (20, 21). Exposing EFM-19 cells to conditioned media from H1299 human lung cancer cells expressing exogenous CD74-*NRG1* fusion protein resulted in phosphorylation of endogenous ERBB2/HER2 and ERBB3/HER3 proteins, suggesting that autocrine HER2:HER3 signaling was activated by secreted *NRG1* ligands generated from CD74-*NRG1* polypeptides (**Fig. 2A**). Phosphorylation of ERK and AKT, downstream mediators of HER2:HER3, was also elevated. HER2, HER3, and ERK phosphorylation was suppressed by lapatinib and afatinib, FDA-approved TKIs that target HER

kinases (22-24). Together, these observations indicate that *NRG1* fusions activated HER2:HER3 signaling by juxtacrine and/or autocrine mechanisms.

The EZR-ERBB4 fusion protein contained the EZR coiled-coil domain, which functions in protein dimerization, and also retained the full ERBB4 kinase domain (**Fig. 1A**). These features indicated that the EZR-ERBB4 protein is likely to form a homodimer via the coiled-coil domain of EZR, causing aberrant activation of the kinase function of ERBB4, similar to the situation of *EZR-ROS1* fusion (5). Indeed, when the *EZR-ERBB4* cDNA was exogenously expressed in NIH3T3 fibroblasts, tyrosine 1258, located in the activation loop of the ERBB4 kinase site, was phosphorylated in the absence of serum stimulation, indicating that fusion with EZR aberrantly activated the ERBB4 kinase (**Fig. 2B**). Consistent with this, phosphorylation of a downstream mediator, ERK, was also elevated. Phosphorylation of ERBB4 and ERK was suppressed by lapatinib and afatinib, which inhibit ERBB4 protein (22-24).

The TRIM24-BRAF fusion protein retained the BRAF kinase domain but lacked the N-terminal RAS-binding domain responsible for negatively regulating BRAF kinase. These features suggested that the fusion was constitutively active, as in the cases of the *ESRP1-BRAF* and *AGTRAP-BRAF* fusions in other cancers (25). When the *TRIM24-BRAF* cDNA was exogenously expressed in NIH3T3 cells, ERK, a downstream mediator of BRAF, was phosphorylated in the absence of serum stimulation, indicating that fusion with TRIM24 aberrantly activated BRAF kinase (**Fig. 2C**). ERK phosphorylation was suppressed by sorafenib, an FDA-approved drug originally identified as a RAF kinase inhibitor (26), and also by the MEK



inhibitor U0126 (**Fig. 2C**).

Exogenous expression of fusion gene cDNAs induced anchorage-independent growth of NIH3T3 fibroblasts, indicating their transforming activities (**Figs. 2D–F**). This growth was suppressed by the kinase inhibitors that suppressed fusion-induced activation of signal transduction, as described above. NIH3T3 cells expressing *EZR-ERBB4* or *TRIM24-BRAF* fusion cDNA formed tumors in nude mice (**Fig. 3**). Therefore, we concluded that these three fusions function as driver mutations in IMA development. We screened 200 commonly used human lung cancer cell lines, but all were negative for these three fusions (data not shown); thus, the oncogenic properties of these fusions remain unvalidated in human cancer cells.

The results here suggest that the *NRG1*, *ERBB4* and *BRAF* fusions are novel driver mutations involved in the development of IMAs of the lungs (**Fig. 1C**) and potential targets for existing TKIs. The recurrent *NRG1* fusions were especially notable because *NRG1* was previously identified as a regulator of goblet-cell formation in primary cultures of human bronchial epithelial cells (27); therefore, activation of the *NRG1*-mediated signaling pathway(s) might play a part in IMA development by contributing to both cell transformation and acquisition of goblet-cell morphology. In addition to a small fraction of known druggable aberrations (an *ALK* fusion and an *EGFR* mutation), more than 10% (11/90; 12.2%) of IMAs harbored other druggable aberrations targeted by existing kinase inhibitors: these aberrations were represented by fusions involving *NRG1*, *ERBB4*, *BRAF*, or *RET*, or *BRAF* mutations (**Table 2**,

**Fig. 1C).** To facilitate translation of these findings to the cancer clinic, it will be necessary to establish diagnostic methods, particularly using break-apart and fusion FISH methods, capable of detecting these aberrations. Such methods will also help identify additional fusions involving other partner genes and contribute to a greater understanding of the significance of gene fusions in lung carcinogenesis.