Lung Cancer 114 (2017) 90–95

Contents lists available at ScienceDirect

Lung Cancer

journal homepage: www.elsevier.com/locate/lungcan

Research paper

Clinical characteristics and sequence complexity of anaplastic lymphoma kinase gene fusions in Chinese lung cancer patients

Hong-xia Tian[1], Xu-chao Zhang[1], Jin-ji Yang, Wei-bang Guo, Zhi-hong Chen, Zhen Wang, Yi-long Wu*

Medical Research Center, Guangdong Lung Cancer Institute, Guangdong General Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, China

ARTICLE INFO

Keywords:
NSCLC
SCLC
ALK
EFGR
Targeted molecular therapy
Variants
Diversity

ABSTRACT

Objectives: To investigate the clinical characteristics of anaplastic lymphoma kinase (*ALK*) rearrangements and sequence complexity of the *ALK* fusion gene in Chinese lung cancer patients.

Methods: We prospectively screened *ALK* rearrangements in 1474 lung cancer specimens, including 1387 cases of non-small cell lung cancer (NSCLC), 54 cases of small cell lung cancer (SCLC), and 33 cases of cancer with lung metastasis from other organs by both standard polymerase chain reaction (PCR) and rapid amplification of cDNA ends (RACE)-coupled PCR. Fifteen cases of *ALK*-positive RACE-coupled PCR products were transformed into *Escherichia coli* for molecular cloning and sequenced for complexity analysis.

Results: The overall frequency of *ALK* rearrangements was 5.1% (71/1387) in NSCLC. In 71 positive cases, the coexistence of epidermal growth factor receptor (*EGFR*) and *ALK* variations was found in 6 cases (8.5%), and the coexistence of different *ALK* variants was found in 2 cases (2.8%) (1 case with variants 1 and 9; the other case with variants 3 and 2) by PCR analysis. Furthermore, through sequence cloning analysis of 15 cases of non-selective *ALK*-positive samples, two cases with variants 1 and 3 harbored the coexistence of three subtypes (variant 1 subtypes: E13; A20, E13del63; A20 and E7E12E13; A20 and variant 3 subtypes: E6; A20, E6ins33; A20 and E3E6; A20). Variant 3a and 3b subtypes were always coexistent and had the same proportion of *ALK* variant 3 rearrangements. *ALK* rearrangement was associated with young age, female gender, never-smokers, those with adenocarcinoma, advanced stage, and *EGFR* mutations. No *ALK* fusion was detected in 54 cases of SCLC or 33 cases of cancer with lung metastasis from other organs.

Conclusions: The identification of novel *ALK* variants, the coexistence of *EGFR* mutations and *ALK* fusions, the coexistence of *ALK* variants, and the coexistence of subtypes reveal the diversity and sequence complexity of *ALK* fusions.

1. Introduction

Anaplastic lymphoma kinase (*ALK*) gene rearrangement was first identified fused to nucleophosmin (*NPM1*) in anaplastic large-cell lymphoma [1]. The *ALK* fusion gene was defined as a new driver gene after epidermal growth factor receptor (*EGFR*) mutations in non-small cell lung cancer (NSCLC), since individuals harboring an *ALK* fusion gene benefited from *ALK* tyrosine kinase inhibitors (TKIs) [2–4]. However, their responses varied widely in terms of magnitude and duration for each patient (95% confidence interval [CI], 67–81% and 8.3–13.9 months, respectively) in the first-line setting [4]. It was thus critical to investigate the potential mechanisms leading to the heterogeneity of responses.

Fusion partners of the *ALK* gene vary widely, and include echinoderm microtubule-associated protein-like 4 (*EML4*), kinesin family member 5 B (*KIF5B*) [5], kinesin light chain 1 (*KLC1*) [6], among others. The *EML4* gene is the predominant 5′ partner in *ALK* fusion-positive lung cancers, and chromosomal breakpoints within the *EML4* gene often occur at different sites. However, breakpoints within the *ALK* gene are more conservative and are mostly reported within intron

Abbreviations: ALK, anaplastic lymphoma kinase; DCTN1, dynactin 1; EGFR, epidermal growth factor receptor; EML4, microtubule-associated protein-like 4; FISH, fluorescence in situ hybridization; GLCI, Guangdong Lung Cancer Institute; HELP, hydrophobic EMAP-like protein; IHC, immunohistochemistry; IPTG, isopropyl β -D-1-thiogalactopyranoside; KIF5B, kinesin family member 5B; KLC1, kinesin light chain 1; LB, Luria-Bertani; NSCLC, non-small cell lung cancer; NPM1, nucleophosmin; PCR, polymerase chain reaction; RT, reverse transcription; RACE, rapid amplification of cDNA ends; SCLC, small cell lung cancer; TKIs, tyrosine kinase inhibitors; X-gal, 5-bromo-4-chloro-3-indolyl β -D-galactoside

* Corresponding author.
E-mail address: syylwu@live.cn (Y.-l. Wu).

[1] Hongxia Tian and Xuchao Zhang contributed equally to this work.

https://doi.org/10.1016/j.lungcan.2017.11.001
Received 5 September 2017; Received in revised form 15 October 2017; Accepted 1 November 2017

19 or exon 20 [7–10]. Patients with the *EGFR* 19 deletion achieve longer progression-free survival from *EGFR*-TKIs than *EGFR* L858R [11], and the *EGFR* 19 deletion is associated with longer overall survival after first-line afatinib treatment [12]. This phenomenon inspired us to investigate the complexity of *ALK* variants in Chinese patients with lung cancer.

2. Materials and methods

2.1. Patient specimens

A total of 1474 patients (95.5% Han, 0.5% ethnic minorities) with lung cancer, including 1387 cases of NSCLC, 54 cases of small cell lung cancer (SCLC), and 33 cases of cancer with lung metastasis from other organs (e.g., cervical, breast, and thyroid cancer), diagnosed between January 2011 and February 2013 were prospectively enrolled. All specimens were obtained from the Guangdong Lung Cancer Institute (GLCI) biological sample bank, which was approved by the local ethics committee of Guangdong General Hospital (No.GDREC2013013). Informed consent was obtained from each person for subsequent *ALK* analyses before any biopsy was performed. All major histologic subtypes of NSCLC are represented in Table 1.

2.2. Study design

This prospective study included 1474 patients with lung cancer (Fig. 1). All specimens were subjected to an *ALK* rearrangement assay by commonwith the standard reverse transcription polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE)-coupled PCR and sequenced. The incidence of *ALK* fusion and the frequency of all variants were analyzed, and the clinical characteristics were evaluated in NSCLC. Furthermore, 15 cases of *ALK*-positive variants underwent molecular cloning for sequence complexity analysis.

Table 1
Summary of patient demographics and clinicopathological profiles of patients with NSCLC with *ALK* rearrangement.

Characteristic	Total	*ALK* status		
		Pos., no. (%)	Neg., no. (%)	P-value
No. of patients with NSCLC	1387	71 (5.1)	1316 (94.9)	
Age (years)				
Median	59	48	60	< 0.001
Range	23–89	23–74	24–89	
≤60	750 (54.1)	62	688	
> 60	637 (45.9)	9	628	
Sex				
Male	872 (62.9)	34	838	0.007
Female	515 (37.1)	37	478	
Smoking history				
Never-smoker	778 (56.1)	54	724	0.001
Ever-smoker	609 (43.9)	17	592	
Pathology type				
Adenocarcinoma	1100 (79.3)	68	1032	< 0.001*
Squamous	209 (15.1)	2	207	
Other	78 (5.6)	1	77	
Stage				
I–IIIa	557 (40.2)	18	539	0.009
IIIb–IV	830 (59.8)	53	777	
EGFR status				
Wild-type	897 (64.7)	65	832	< 0.001
Mutated	490 (35.3)	6	484	
Specimen type				
Surgical biopsy	807 (58.2)	37	770	0.287
Percutaneous biopsy	580 (41.8)	34	546	

Abbreviations: NSCLC, non-small cell lung cancer; ALK, anaplastic lymphoma kinase; EGFR, epidermal growth factor receptor.

* Upon comparing the adenocarcinoma type with squamous and others.

2.3. Nucleic acid extraction

All samples were assessed pathologically using hematoxylin and eosin staining, Then, we removed the nomal tissue from-semi frozen samples with a shape blade to enrich the tumor cells to ensure almostly more than 50%, while we operated at a low temperature to protect the RNA from degradation. Total RNA and DNA were extracted from lung tissue samples using the RNeasy® Kit (QIAGEN, Valencia, CA, USA) and the DNeasy® Kit (QIAGEN), respectively. Nucleic acid quality was assessed by 1% gel electrophoresis and quantified on a NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Extracted RNA and DNA were stored at −20 °C until use.

2.4. Reverse transcription

An *ALK* gene-specific primer was used to reverse transcribe RNA into cDNA. The primer sequence was as follows: 5′-TTCAGGCAGCGT GTTCACAGCCA-3′. Reverse transcription was performed according to the manufacturer's recommended protocol for the TaKaRa RNA PCR (AMV) Kit, Ver. 3.0 (Takara, Shiga, Japan). Briefly, the RT reaction consisted of 2 μL $MgCl_2$, 1 μL RT buffer (10×), 3.75 μL H_2O, 1 μL dNTPs (10 mM), 0.25 μL RNase inhibitor, 0.5 μL AMV (avian myeloblastosis virus) enzyme, 0.5 μL gene-specific primer (12.5 μM), and 1 μL RNA (< 500 ng total RNA). The reaction was incubated at 42 °C for 30 min, 99 °C for 5 min, and 5 °C for 5 min.

2.5. Common PCR for ALK fusion analysis

The primer sequences used for multiplex PCR were as follows: 5′-TACCAGTGCTGTCTCAATTGCAGG-3′ (forward) and 5′-accaggaaaca gctatgaccTCTTGCCAGCAAAGCAGTAGTTGG-3′ (reverse). The lowercase letters represent bases used for sequencing. Common PCR was performed according to the manufacturer's instructions for the HS PCR Kit (TaKaRa, Dalian, China). Briefly, the reaction consisted of 5 μL PCR buffer (10×), 4 μL dNTPs (10 mM), 0.25 μL TaKaRa Ex Taq™ HS, 1.0 μL forward and reverse primers (12.5 μM), 1 μL RT product, and 38.75 μL H_2O. The reaction conditions were as follows: 94 °C for 10 min; 35 cycles of 94 °C for 40 s, 60 °C for 40 s, and 72 °C for 90 s; and 72 °C for 10 min. The PCR products were then visualized by 2% agarose gel electrophoresis.

2.6. RACE-coupled PCR for ALK fusion analysis

This methodology was based on the principles of 5′ RACE PCR. After reverse transcription, cDNAs were purified using the High Pure PCR Product Purification Kit (Roche, Basel, Switzerland) according to the manufacturer's protocol. Purified cDNAs were further subjected to polycytidine tailing. Then PCRs were performed twice to amplify target cDNA fragments spanning exon 20 of *ALK* and upstream sequences that may contain transcript sequences for any genes fused to *ALK*. The detailed processes and primers used were described previously [13].

2.7. Molecular cloning

The purified RACE PCR products of 15 cases of non-selective *EML4-ALK*-positive tissues, including 6 cases of the V1 variant, 3 cases of the V2 variant, and 6 cases of the V3 variant, were cloned into the T vector according to the pGEM®-T Easy Vector T Vector (Promega, Madison, WI, USA) protocol. The ligation reaction comprised 5 μL ligation buffer (2×), 1 μL T4 DNA ligase, 1 μL T Vector, and 3 μL PCR products. The cocktail was gently mixed and incubated for 1 h at room temperature, followed by transformation into *Escherichia coli* (TaKaRa) according to the manufacturer's protocol. Luria–Bertani (LB) solid medium was