Lung Cancer 116 (2018) 7–14

Contents lists available at ScienceDirect

Lung Cancer

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

A cross-sectional study examining the expression of splice variants *K-RAS4A* and *K-RAS4B* in advanced non-small-cell lung cancer patients

Veronica Aran[a,*], Pedro Masson Domingues[a], Fabiane Carvalho de Macedo[b], Carlos Augusto Moreira de Sousa[a], Tatiane Caldas Montella[a], Maria Theresa de Souza Accioly[b], Carlos Gil Ferreira[a,c]

[a] *Clinical Research Division, Brazilian National Cancer Institute (INCA), Rio de Janeiro, Brazil*
[b] *Pathology Division (DIPAT), Brazilian National Cancer Institute (INCA), Rio de Janeiro, Brazil*
[c] *D'Or Institute for Research and Education (IDOR), Rio de Janeiro, Brazil*

ARTICLE INFO

Keywords:
Non-small cell lung cancer
K-RAS4A
K-RAS4B
Splice variants
Adenocarcinoma
Cancer patients

ABSTRACT

Objectives: Mammalian cells differently express 4 RAS isoforms: *H-RAS, N-RAS, K-RAS4A* and *K-RAS4B*, which are important in promoting oncogenic processes when mutated. In lung cancer, the K-RAS isoform is the most frequently altered RAS protein, being also a difficult therapeutic target. Interestingly, there are two *K-RAS* splice variants (*K-RAS4A* and *K-RAS4B*) and little is known about the role of *K-RAS4A*. Most studies targeting K-RAS, or analysing it as a prognostic factor, have not taken into account the two isoforms. Consequently, the in-depth investigation of them is needed.

Methods: The present study analysed 98 specimens from advanced non-small cell lung cancer (NSCLC) adenocarcinoma patients originated from Brazil. The alterations present in *K-RAS* at the DNA level (Sanger sequencing) as well as the expression of the splicing isoforms at the RNA (qRT-PCR) and protein levels (immunohistochemistry analysis), were evaluated. Possible associations between clinicopathological features and the molecular findings were also investigated.

Results: Our results showed that in the non-smoking population, the cancer incidence was higher among women. In contrast, in smokers and former smokers, the incidence was higher among men. Regarding sequencing results, 10.5% of valid samples presented mutations in exon 2, being all wild-type for exon 3, and the most frequently occurring base change was the transversion G → T. Our qRT-PCR and immunohistochemical analysis showed that both, *K-RAS4A* and *K-RAS4B*, were differently expressed in NSCLC tumour samples. For example, tumour specimens showed higher *K-RAS4A* mRNA expression in relation to commercial normal lung control than did *K-RAS4B*. In addition, K-RAS4B protein expression was frequently stronger than K-RAS4A in the patients analysed.

Conclusion: Our results highlight the differential expression of K-RAS4A and K-RAS4B in advanced adenocarcinoma NSCLC patients and underline the need to further clarify the enigma behind their biological significance in various cancer types, including NSCLC.

1. Introduction

Lung cancer is the leading cause of cancer-related deaths in the world. One of the causes of this lethality is the fact that this type of cancer is usually detected in advanced stages, since the symptomatology in the early stages of the disease is not common, resulting in poor prognosis. The American Cancer Society estimated 117,920 new cases among men and 106,470 among women in 2016, with 158,080 deaths in both sexes [1]. In 2017, data from the Brazilian National Cancer Institute (INCA) estimates 17,330 new cases of lung cancer among men and 10,890 among women [2]. Smoking is the prominent cause of this cancer type in most populations, however othercauses have also been attributed, such as air pollution, exposure to occupational and environmental carcinogens [3]. Family history may also be a risk factor, although it is difficult to establish its real impact, since it can also be connected to environmental factors. The implementation of comprehensive tobacco control, and increased awareness of the risks, remains the main form of reduction in the occurrence of this type of neoplasia. In 2012, for example, the prediction was 160,340 deaths for both sexes in Brazil, nevertheless in 2016 the value decreased to 158,080 deaths for both sexes [2].

Lung cancer is a heterogeneous disease being classified as either

* Corresponding author at: Clinical Research Division, Brazilian National Cancer Institute (INCA), Rua André Cavalcanti, 37, 2° Andar, Rio de Janeiro, RJ, 20231-050, Brazil.
E-mail address: varanponte@gmail.com (V. Aran).

https://doi.org/10.1016/j.lungcan.2017.12.005
Received 9 October 2017; Received in revised form 8 December 2017; Accepted 11 December 2017

small cell (SCLC, 15%) or non-small cell (NSCLC, 85%). NSCLC is further classified into adenocarcinoma (40–50%), squamous cell carcinoma (~30%) and large cell carcinoma (~9%) [4]. The concept of personalized medicine has improved both lung cancer diagnosis and treatment helping in therapeutic decisions, which can be evaluated according to specific histologic and genetic characteristics of the patient's tumour. This can determine eligibility for certain types of molecular testing and therapeutic strategies. The discovery of crucial molecular alterations has led to the development of specific inhibitor drugs such as erlotinib and gefitinib which target the epidermal growth factor receptor (EGFR) [5,6] or crizotinib for ALK translocation that results in the *EML4-ALK* oncogene [7]. Nevertheless, other genes are also frequently found mutated in lung cancer, being important candidates for targeted therapy.

Some of the most important examples of proto-oncogenes belong to the family of RAS proteins, which are small GTPases frequently found mutated in different cancer types. There are three RAS genes called *H-, N-*and *K-RAS* that encode 4 highly homologous protein isoforms known as H-RAS, N-RAS, K-RAS4A and K-RAS4B which induce different biological responses [8]. The K-RAS4A and -4B isoforms are the result of the alternate splicing of exon 4 resulting in an additional amino acid in the 4A isoform (189 amino acids) [9]. Interestingly, this alternative fourth exon encodes the hypervariable protein's region responsible for membrane targeting, rendering K-RAS4A and K-RAS4B different (K-RAS4B lacks a site of palmitoylation, whereas K-RAS4A is palmitoylated) [10]. Studies performed in mouse have shown that *K-RAS4A* is dispensable for development, whereas the isoform *4B* is essential [11–13].

K-RAS mutations are present in ~15–30 % NSCLC patients, being adenocarcinoma the most frequently found subtype [14]. Mutations affect identical regions in *K-RAS4A* and *K-RAS4B*, corresponding to codons 12 and 13 and less frequently found at codon 61, making the protein GAP insensitive and constitutively bound to GTP causing the activation of downstream effectors such as RAF, MEK, ERK, PI3K, AKT, MTOR [15]. These effectors are studied as targets for drug development, since RAS itself has been shown to be hard to target. Different RAS are associated with different types of tumours thus demonstrating a certain biological specificity. For example, mutations in *K-RAS* occur frequently in pancreatic, colorectal, and lung cancer, whereas *H-* and *N-RAS* mutations are not common in these tumours. In acute leukemias, *N-RAS* is the most frequently mutated isoform and, in general, *H-RAS* is the least mutated isoform found in tumours [16].

The *K-RAS* splice variants are conserved across mammalian genomes and it was hypothesized that *K-RAS4A* and *K-RAS4B* have distinct physiological activities and different expression patterns depending on the tissue analysed [8,16,17]. Despite the fact that *K-RAS4A* was the first splice variant discovered [18], most studies highlighted *K-RAS4B* as the major variant, leaving *K-RAS4A* under its shadow. Therefore, the main aim of the present study was to investigate, for the first time, the relative expression of the *K-RAS* splice variants transcripts across a group of advanced adenocarcinoma NSCLC patients. We have performed immunohistochemistry analysis and qRT-PCR to quantify the transcript expression (protein and RNA levels, respectively) contribution of each isoform. The results were also compared to expression in normal lung. In addition, we have conducted *K-RAS* sequencing in order to detect the presence or absence of mutation in the samples. Possible links between clinicopathological features and the molecular findings were also investigated.

2. Materials and methods

2.1. Study population

The study included 98 patients diagnosed with advanced adenocarcinoma NSCLC (stages IIIb-IV). Samples were obtained *via* tru-cut lung biopsies and the tissues were formalin-fixed paraffin-embedded. The eligible cases included patients who were diagnosed with primary adenocarcinoma NSCLC between 2003 and 2012 at the Brazilian National Cancer Institute (INCA). Patients' data, including: clinical staging of NSCLC [19], gender, age at diagnosis, tumour histology, ethnical origin, clinical TNM stage, performance status (PS), smoking history, treatments undertaken; were collected from their medical records. In ~12% of patients, none of the molecular analyses could be performed mostly due to an insufficient amount of tissue sample or missing sample.

This study was approved by the local ethical committee (*Comitê de Ética em Pesquisas*- CEP from the Brazilian National Cancer Institute, INCA, Protocol 79/05). The detailed outline of the valid percent patients' characteristics is shown Table 1 of the results Section 3.1.

2.2. DNA and RNA extractions

The tissue slide corresponding to the paraffin-embedded tumour block was analysed by a pathologist. The tumour area was marked and a fragment was digged out from the tissue block using proper stylet and transferred into 1.5 ml tubes. After routine deparaffination of the tissue sections, the AllPrep DNA/RNA FFPE Kit (QIAGEN, Germany) was used to allow the simultaneous purification of genomic DNA and total RNA from the same biological sample as indicated by the manufacturer. The concentration and purity of the RNA samples were assessed using NanoDrop Spectrophotomer (NanoDrop Technologies Inc., Wilmington, DE). The Nanodrop manufacture's protocol was followed to assess both DNA and RNA purity. Samples were considered valid unless they did not present good conditions for analysis, or were missing from the bank.

2.3. Polymerase chain reaction (PCR) and DNA sequencing

For the K-RAS mutation status analysis, the extracted DNA was subjected to conventional PCR for amplification of relevant portions of exons 2 and 3 of the K-RAS gene. The primers utilized were: *K-RAS* exon 2 F (5′-TAAGGCCTGCTGAAAATGACTG -3′), *K-RAS* exon 2 R (5′-AATGGTCCTGCACCAGTAATA-3′), *K-RAS* exon 3 F (5′-GACT GTGTTTCTCCCTTCTCAGGAT-3′), *K-RAS* exon 3 R (5′-TATGGCAA ATACACAAAGAAAGCCC-3′). Each PCR mix contained the forward and reverse primers for each exon (10 µM each), 200 ng of DNA (varying volume), 2 µl $MgSO_4$ (50 mM), 1 µl of 10 mM dNTPmix, 5 µl 10× High fidelity PCR buffer (Invitrogen), 0.5 µl Platinum Taq DNA Polymerase high fidelity (5 U/µl, Invitrogen), plus distilled water up to a total volume of 50 µl. PCR conditions consisted of initial denaturing at 95 °C for 2 min; 40 cycles of 95 °C for 30 s, 57 °C for 30 s, 68 °C for 30 s; a final 1 cycle extension at 68 °C for 10 min. All reactions were performed using the Applied Biosystems Veriti Thermal Cycler equipment (ThermoFisher scientific, USA). The PCR products were analysed in a 2% agarose gel and visualised under ultraviolet light. The samples were considered positive when the expected band was seen. After agarose gel electrophoresis, 40 µl of PCR substrate was purified using the commercial kit illustra™ GFX PCR DNA and Gel Band Purification Kit (GE Healthcare, NJ, USA) following manufacturer's instructions.

The amount of 4 µl of PCR substrate and 3.2 pmol of *K-RAS* oligos were used in subsequent sequencing reactions using the commercial kit BigDye Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems, Massachusetts, USA) according to manufacturer's instructions. PCR products were analysed by a 16-capillary automated sequencer (ABI PRISM® 3100 GeneticAnalyzer/HITACHI, Applied Biosystems, Massachusetts, USA), based on the Sanger method. Chromatograms were analysed and sequences were imported into a sequence analysis software (Mutation Surveyor v3.9, SoftGenetics, USA) using *K-RAS* reference samples for exons 2 and 3 available from Genbank.