Ⅳ．研究成果の刊行物・別刷

OXFORD UNIVERSITY PRESS | Carcinogenesis

Expression and Clinical Significance of Genes Frequently Mutated in Small Cell Lung Cancers Defined by Whole Exome/RNA Sequencing

Journal:	*Carcinogenesis*
Manuscript ID:	CARCIN-2014-00970.R1
Manuscript Type:	Original Manuscript
Date Submitted by the Author:	02-Mar-2015
Complete List of Authors:	Iwakawa, Reika; National Cancer Center Research Institute, Division of Genome Biology Kohno, Takashi; National Cancer Center Research Institute, Division of Genome Biology Totoki, Yasushi; National Cancer Center Research Institute, Divisions of Cancer Genomics Shibata, Tatsuhiro; National Cancer Center Research Institute, Divisions of Cancer Genomics Tsuchihara, Katsuya; National Cancer Center, Exploratory Oncology Research and Clinical Trial Center Mimaki, Sachiyo; National Cancer Center, Exploratory Oncology Research and Clinical Trial Center Tsuta, Koji; National Cancer Center Hospital, Pathology Division Natita, Yoshitaka; National Cancer Center Hospital, Department of Neurosurgery and Neuro-Oncology Nishikawa, Ryo; International Medical Center, Saitama Medical University, Department of Neuro-Oncology/Neurosurgery Noguchi, Masayuki; Faculty of Medicine, University of Tsukuba, Department of Pathology Harris, Curtis; National Cancer Institute, Laboratory of Human Carcinogenesis Roble, Ana I.; National Cancer Institute, Laboratory of Human Carcinogenesis Yamaguchi, Rui; Institute of Medical Science, University of Tokyo, Human Genome Center Imoto, Seiya; Institute of Medical Science, University of Tokyo, Human Genome Center Miyano, Satoru; Institute of Medical Science, University of Tokyo, Human Genome Center Totsuka, Hirohiko; Hitachi Government and Public Corporation System Engineering Ltd, Bioinformatics Group Yoshida, Teruhiko; National Cancer Center Research Institute, Division of Genetics Yokota, Jun; National Cancer Center Research Institute, Division of Genome Biology

Keywords:	Small cell lung cancer, whole-exome sequencing, whole RNA sequencing

Expression and Clinical Significance of Genes Frequently Mutated in Small Cell Lung Cancers Defined by Whole Exome/RNA Sequencing

Reika Iwakawa[1], Takashi Kohno[1,2], Yasushi Totoki[3], Tatsuhiro Shibata[3], Katsuya Tsuchihara[2], Sachiyo Mimaki[2], Koji Tsuta[4], Yoshitaka Narita[5], Ryo Nishikawa[6], Masayuki Noguchi[7], Curtis C. Harris[8], Ana I. Robles[8], Rui Yamaguchi[9], Seiya Imoto[9], Satoru Miyano[9], Hirohiko Totsuka[10], Teruhiko Yoshida[11], Jun Yokota[1,12]

[1]Division of Genome Biology, [3]Division of Cancer Genomics, [11]Division of Genetics, National Cancer Center Research Institute, Tokyo, Japan
[2]Exploratory Oncology Research and Clinical Trial Center, National Cancer Center, Tokyo, Japan
[4]Pathology Division, [5]Department of Neurosurgery and Neuro-Oncology, National Cancer Center Hospital, Tokyo, Japan
[6]Department of Neuro-Oncology/Neurosurgery, International Medical Center, Saitama Medical University, Saitama, Japan
[7]Department of Pathology, Faculty of Medicine, University of Tsukuba, Ibaraki, Japan
[8]Laboratory of Human Carcinogenesis, National Cancer Institute, National Institutes of Health, MD, USA
[9]Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan
[10]Bioinfomatics Group, Research and Development Center, Solution Division 4, Hitachi Government and Public Corporation System Engineering Ltd, Tokyo, Japan
[12]Cancer Genome Biology Group, Institute of Predictive and Personalized Medicine of Cancer, Barcelona, Spain.

Supported in part by Grants-in-Aid from the Ministry of Health, Labor and Welfare for the 3rd-term Comprehensive 10-year Strategy for Cancer Control and from the Ministry of Education, Culture, Sports, Science and Technology for Scientific Research on Innovative Areas, Japan. R. I. is an awardee of Research Resident Fellowship from the Foundation for Promotion of Cancer Research for the 3rd Term Comprehensive 10-Year Strategy for Cancer Control, Japan.

Corresponding Author: Jun Yokota, Division of Genome Biology, National Cancer Center Research Institute, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan. e-mail: jyokota@ncc.go.jp

Running Title: Mutation and Expression Profiles of Small Cell Lung Cancer

Keywords: Small cell lung cancer, whole-exome sequencing, whole RNA sequencing

Appropriate Category: Research Articles - Cancer Cell Biology

Brief Description: Whole-Exome/RNA sequencing of 38 SCLC cases identified five genes, with expression of their mutated alleles, as being frequently mutated in any progression stage of SCLC; therefore, their mutated gene products will be promising targets for therapy of SCLC patients

ABSTRACT

Small cell lung cancer (SCLC) is the most aggressive type of lung cancer. Only 15% of SCLC patients survive beyond 2 years after diagnosis. Therefore, for the improvement of patients' outcome in this disease, it is necessary to identify genetic alterations applicable as therapeutic targets in SCLC cells. The purpose of this study is the identification of genes frequently mutated and expressed in SCLCs that will be targetable for therapy of SCLC patients. Exome sequencing was performed in 28 primary tumors and 16 metastatic tumors from 38 patients with SCLCs. Expression of mutant alleles was verified in 19 cases by RNA sequencing. *TP53*, *RB1*, and *PTEN* were identified as being significantly mutated genes. Additional 36 genes were identified as being frequently (≥10%) mutated in SCLCs by combining the results of this study and two recent studies. Mutated alleles were expressed in 8 of the 36 genes, *TMEM132D*, *SPTA1*, *VPS13B*, *CSMD2*, *ANK2*, *ASTN1*, *ASPM* and *FBN3*. In particular, the *TMEM132D*, *SPTA1* and *VPS13B* genes were commonly mutated in both early and late stage tumors, primary tumors and metastases, and tumors before and after chemotherapy, as in the case of the *TP53* and *RB1* genes. In addition to *TP53*, *RB1* and *PTEN*, *TMEM132D*, *SPTA1* and *VPS13B* could be also involved in SCLC development, with the products from their mutated alleles being potential therapeutic targets in SCLC patients.

Summary

We identified the *TMEM132D*, *SPTA1* and *VPS13B* genes as frequently mutated and expressed in any progression stage of SCLC, in addition to known tumor suppressor genes, *TP53*, *RB1* and *PTEN*. These genes could be involved in SCLC development, and their mutated gene products will be promising therapeutic targets in SCLC patients.

INTRODUCTION

Small cell lung cancer (SCLC) is the most aggressive type of lung cancer with extremely poor prognosis [1-3]. Therefore, it is critical to develop novel therapies for the improvement of prognosis. Up to the present, extensive genome-wide molecular analyses have been performed in SCLC [4-8]. However, targetable genetic alterations have not been defined. This could be due to several reasons. First, mutation rates in SCLCs are very high. Second, except for the frequent inactivation of the *TP53* and *RB1* genes, the spectra of genes with high mutation frequencies are considerably different among studies. Third, only a small number of clinical SCLC cases have been analyzed to date, due to the small proportion of surgical cases. Fourth, no known druggable mutations have been identified to date. Therefore, we still do not have enough information to select genes with significant clinical importance and with druggable mutations in SCLC.

We previously performed the whole-genome copy number analysis and RNA sequencing of 58 and 42 SCLCs, respectively, and revealed the importance of amplification of *MYC* family genes and several other genes [8]. In this study, we attempted to obtain further information on genes frequently mutated in SCLC through the analysis of 38 SCLC cases. The study confirmed high mutation rates in SCLC, and identified *TP53*, *RB1* and *PTEN* as significantly mutated genes in SCLCs. Furthermore, by a comparative study of our results with those of two previous studies [6,7], additional 36 genes were identified as being frequently (≥10%) mutated in SCLCs. Therefore, expression of mutated alleles in these genes was validated by whole transcriptome sequencing, and their mutation statuses were investigated in early and late stages of SCLCs, in primary tumors and metastases, and in tumors before and after chemotherapy. Mutated alleles were expressed only in 8 of the 36 genes, and sets of genes mutated in tumors were drastically different between early and late stages SCLCs, between primary tumors and metastases, and also between tumors before and after chemotherapy. We discussed the possible roles of significantly and frequently mutated genes in SCLC development, and defined several therapeutically targetable genes in SCLCs.

PATIENTS AND METHODS

Tissues

Forty-four tumors and the corresponding non-cancerous tissues were obtained from 38 patients with SCLC at surgery or autopsy from 1985 to 2010 at the National Cancer Center Hospital, Tokyo, Japan; Saitama Medical University, Saitama, Japan; University of Tsukuba, Ibaraki, Japan; and hospitals in the Metropolitan Baltimore area in the U.S, and kept frozen until DNA and RNA extraction. Clinicopathological characteristics and sample information are summarized in Table 1 and Supplementary Table S1, respectively. The tumors were histologically diagnosed according to the 2004 WHO classification and pathologically staged according to the tumor-node-metastasis classification of malignant tumors [9,10]. Immunohistochemical staining data for neuroendocrine markers, synaptophysin, chromogranine and neural cell adhesion molecule (NCAM), were available in most cases; thus, were added in Supplementary Table S1. Primary tumors were obtained from 25 cases, and metastases were obtained from 9 cases. Both primary tumors and metastases were obtained in cases 10 and 22. Two metastases were obtained in case 20, while the primary tumors and metastases to three different organs were obtained in case 21. Thirty cases were not treated by chemotherapy/radiotherapy, while the other 7 cases were treated before sampling. In case 10, the primary tumor was obtained before chemotherapy, and the metastasis was obtained after chemotherapy. Genomic DNA was extracted from frozen samples with a QIAamp DNA mini kit (Qiagen, Hilden, Germany). Total RNA was also extracted from frozen samples using TRIzol reagent (Invitrogen, Carlsbad, CA), purified by an RNeasy kit (Qiagen), and reverse-transcribed to cDNA by using the SuperScript III First-Strand Synthesis System (Invitrogen) with random hexamers. This study was performed under the approval of the Institutional Review Board of the National Cancer Center, Tokyo, Japan.

Exome Sequencing and Data Processing

Methods of exome sequencing and data processing are summarized in Supplementary Figure S1. One microgram of genomic DNA was fragmented using the Covaris S220. After ligation of the paired-end (PE) adapter including index, a fraction of 300-350 bp was gel-purified and amplified with PCR using the TruSeq

DNA Sample Prep Kit (Illumina). Exome capture was performed using the TruSeq Exome Enrichment Kit (Illumina) according to the manufacturer's protocol. The resulting libraries were subjected to the PE sequencing of 100-bp reads on the Genome Analyzer IIx (GAIIx) or HiSeq 2000 (Illumina). Mutation calling was TMEM132Dperformed using the EBCall algorithm [13]. Germline variations represented in dbSNP Build 131 or 1000 Genome project (Nov. 2010) and synonymous variations were filtered out. Variants present in the tumor with *p*-values < 0.01 by Fisher's exact test as compared with those in the matched normal tissue were predicted to be somatic mutations. Mean of target coverage in 44 tumors and 38 normal tissues was 97.3 (31.4-170.1) (Supplementary Figure S2). MutSigCV analysis [14] was performed using expression profile data of 19 tumors analyzed by exome sequencing.

RNA Sequencing and Data Processing

RNA sequencing was performed as described previously [8,15]. Expression of mutated alleles detected by exome sequencing of the corresponding tumors was validated by a Bayesian inference method using the samtools and bcftools software [16,17].

Microarray Experiments and Data Processing

Two micrograms of total RNA were labeled using a 5X MEGAscript T7 Kit (Ambion, Inc., Austin, Texas, USA) and analyzed by U133Plus2.0 arrays (Affymetrix). Data was processed by the MAS5 algorithm as described previously [18].

RESULTS

Significantly Mutated Genes

Exome sequencing of 44 tumors from 38 SCLC patients identified 9,279 protein-altering somatic mutations. In 4 cases with multiple tumors analyzed, common and unique mutations were detected between primary tumors and metastases and also between metastases of different organs. Therefore, total numbers of different somatic mutations among the multiple tumors in the same patients were considered as the number of somatic mutations in each case. Thirty-eight cases had an average of 244.2 protein-altering somatic mutations (19-1023)

with the mean rate of 7.4 mutations (0.6-30.8) per megabase in 5,669 genes (Supplementary Figure S3). Genes with protein altering mutations are listed with their nucleotide sequence changes in Supplementary Table S2. G/C>T/A transversions occurred most frequently (Supplementary Figure S4), consistent with the association of tobacco smoking with the occurrence of G/C>T/A transversions. Several software programs are now available to distinguish driver mutations from passenger mutations. We applied the MutSigCV method [14] to identify significantly mutated genes among the 5,669 genes by integrating the expression profile data. Then, the *TP53*, *RB1* and *PTEN* genes were identified as being significantly mutated in SCLCs. *TP53* was mutated in 30 cases (78.9%), *RB1* in 28 cases (73.7%), and *PTEN* in 5 cases (13.2%) (Supplementary Table S3). Types of *TP53*, *RB1* and *PTEN* mutations are summarized in Supplementary Table S4. Previously, Peifer et al [6]. identified 22 significantly mutated genes, and Rudin et al [7]. identified another set of 22 significantly mutated genes in SCLCs. Although the criteria of "significantly mutated gene" are different among three studies, both *TP53* and *RB1* were identified in all studies, while *PTEN* was identified only in this study (Supplementary Table S5). *PTEN* was mutated in 2/27 (7.4%) and 2/30 (6.7%) in those studies, respectively (Supplementary Table S3). Therefore, although *PTEN* mutations could be significant, its mutation frequency is much lower than those of *TP53* and *RB1* in SCLCs.

Frequently Mutated Genes

A total of 263 genes were mutated in ≥10% of the 38 cases (Figure 1). However, the power to detect frequently mutated genes could be small because of a small sample size in this study and a high background mutation frequency in SCLC. To increase the power of defining genes frequently mutated in SCLC, we further selected frequently mutated genes (≥10%) from two previous studies for whole exome sequencing of SCLC, in which the results of 27 and 30 SCLC cases were reported, respectively [6,7]. Similarly high mutation rates of 7.4 and 5.5 per megabase, respectively, were reported in those studies, justifying a combined analysis. A total of 331 genes and 230 genes, respectively, were identified as being mutated in ≥10% of SCLC cases in those studies (Figure 1). Sets of frequently mutated genes were considerably different among three studies possibly due to the difference in analytical methods used and the small number of cases analyzed in each study, and also due to the high mutation rates in SCLCs. SCLC samples were