様式第１９

学 会 等 発 表 実 績

委託業務題目「ゲノム医療実施体制の開発と試行的・実証的臨床研究【神経・筋疾患及び知的障害領域】」
機関名　国立成育医療研究センター研究所

１．学会等における口頭・ポスター発表

発表した成果（発表題目、口頭・ポスター発表の別）	発表者氏名	発表した場所（学会等名）	発表した時期	国内・外の別
なし				

２．学会誌・雑誌等における論文掲載

掲載した論文（発表題目）	発表者氏名	発表した場所（学会誌・雑誌等	発表した時期	国内・外の別
Targeted Next-Generation Sequencing Effectively Analyzed the Cystic Fibrosis Transmembrane Conductance Regulator Gene in Pancreatitis.	1. Nakano E, Masamune A, Niihori T, Kume K, Hamada S, Aoki Y, Matsubara Y, Shimosegawa T	Dig Dis Sci. 2014		国外
Molecular basis of non-syndromic hypospadias: systematic mutation screening and genome-wide copy-number analysis of 62 patients.	2. Kon M, Suzuki E, Dung VC, Hasegawa Y, Mitsui T, Muroya K, Ueoka K, Igarashi N, Nagasaki K, Oto Y, Hamajima T, Yoshino K, Igarashi M, Kato-Fukui Y, Nakabayashi K, Hayashi K, Hata K, Matsubara Y, Moriya K, Ogata T, Nonomura K, Fukami M	Hum Reprod	30:499-506, 2015	国外
Two novel gain-of-function mutations of STAT1 responsible for chronic mucocutaneous candidiasis disease: Impaired production of IL-17A and IL-22, and the Presence of anti-IL-17F autoantibody	3. Yamazaki Y, Yamada M, Kawai T, Morio T, Onodera M, Ueki M, Watanabe N, Takada H, Takezaki S, Chida N, Kobayashi I, Ariga T	J. Immunology	193: 4880-4887, 2014.	国外
Interstitial lung disease with multiple microgranulomas in chronic granulomatous disease.	4. Kawai T, Watanabe N, Yokoyama M, Nakazawa Y, Goto F, Uchiyama T, Higuchi M, Maekawa T, Tamura E, Nagasaka S, Hojo M, Onodera M	J Clin Immunol	34: 933-940, 2014.	国外
Palivizumab use in Japanese infants and children with immunocompromised conditions	5. Mori M, Onodera M, Morimoto A, Kosaka Y, Morio T, Notario GF Sharma S, Saji T	Pediatr Infect Dis J	33: 1183-1185, 2014.	国外

Basiliximab treatment for steroid-resistant rejection in pediatric patients with acute liver failure after liver transplantation.	6. Shigeta T, Sakamoto S, Uchida H, Sasaki K, Hamano K, Kanazawa H, Fukuda A, Kawai T, Onodera M, Nakazawa A, Kasahara M:	Pediatrics Transplantation	18: 860-867, 2014	国外
The natural history of MPS I: global perspectives from the MPS I Registry.	7. Beck M, Arn P, Giugliani R, Muenzer J, Okuyama T, Taylor J, Fallet S	Genet Med.	2014 Mar 27	国外
Orphanet Overcoming the barriers to diagnosis of Morquio A syndrome.	8. Bhattacharya K, Balasubramaniam S, Choy Y, Fietz M, Fu A, Jin D, Kim OH, Kosuga M, Kwun Y, Inwood A, Lin HY, McGill J, Mendelsohn NJ, Okuyama T, Samion H, Tan A, Tanaka A, Thamkunanon V, Toh TH, Yang AD, Lin SP	J Rare Dis	2014 Nov 30;9(1):192	国外

（注１）発表者氏名は、連名による発表の場合には、筆頭者を先頭にして全員を記載すること。
（注２）本様式はexcel形式にて作成し、甲が求める場合は別途電子データを納入すること。

様式第１９

学　会　等　発　表　実　績

委託業務題目「ゲノム医療実施に係る患者等の意思決定支援と情報の管理手法の開発」

機関名 大阪大学大学院医学系研究科

１．学会等における口頭・ポスター発表

発表した成果（発表題目、口頭・ポスター発表の別）	発表者氏名	発表した場所（学会等名）	発表した時期	国内・外の別
Genomic Data Sharing for the Benefit of Patients and Citizens - How Can We Promote it in Asia?(口頭)	Kato, K.	Human Genome Meeting 2015, Kuara Lumpur Convention Centre, Malaysia.	2015/3/14-17	国外

２．学会誌・雑誌等における論文掲載

掲載した論文（発表題目）	発表者氏名	発表した場所（学会誌・雑誌等名）	発表した時期	国内・外の別
なし				

（注１）発表者氏名は、連名による発表の場合には、筆頭者を先頭にして全員を記載する
（注２）本様式はexcel形式にて作成し、甲が求める場合は別途電子データを納入すること

様式第１９

学　会　等　発　表　実　績

委託業務題目「ゲノム医療従事者の育成プログラム開発」
機関名　栃木県立がんセンター

１．学会等における口頭・ポスター発表

発表した成果（発表題目、口頭・ポスター発表の別）	発表者氏名	発表した場所（学会等名）	発表した時期	国内・外の別
遺伝性腫瘍の発症前診断とサーヘ〝イランスおよひ〝予防的介入.	菅野 康吉、斎藤 伸哉、高橋 雅博、松岡 千咲、青木 幸恵、牧島 恵子、田中屋 宏爾、中島 健、牛尼 美年子、吉田 輝彦	日本人類遺伝学会第59回大会、日本遺伝子診療学会第21回大会	平成26年11月21日（東京）	国内
リンチ症候群の拾い上け〝とその後の内鏡検査によるサーベイランスの実際.	中島 健、関根 茂樹、中島 好美、坂本 琢、松本 美野里、松田 尚久、菅野 康吉、牛尼 美年子、吉田 輝彦、斎藤 豊	日本人類遺伝学会第59回大会、日本遺伝子診療学会第21回大会	平成26年11月21日（東京）	国内
BRCA1/2 遺伝子変異保持者に対するリスク低減卵巣卵管切除術.	平沢 晃、増田 健太、赤羽 智子、片岡 史夫、冨永 英一郎、阪埜 浩司、進 伸幸、菅野 康吉、小崎 健次郎、青木 大輔	日本人類遺伝学会第59回大会、日本遺伝子診療学会第21回大会	平成26年11月21日（東京）	国内
網膜芽細胞腫の早期発見と遺伝子検査の意義.	鈴木 茂伸、吉田 輝彦、牛尼 美年子、菅野 康吉	日本人類遺伝学会第59回大会、日本遺伝子診療学会第21回大会	平成26年11月21日（東京）	国内
がん家系症候群（リンチ症候群）の沿革と周縁および現代的意義.	菅野康吉	第１回リンチ症候群研究会	平成26年11月29日	国内
菅野康吉：家族性腫瘍の診療と研究2015 -発症前診断、サーベイランス、予防的介入- .	菅野康吉	第20回信州遺伝子診療研究会	平成27年1月30日	国内

２．学会誌・雑誌等における論文掲載

掲載した論文（発表題目）	発表者氏名	発表した場所（学会誌・雑誌等名）	発表した時期	国内・外の別
Relationship between smoking and multiple colorectal cancers in patients with Japanese Lynch syndrome: a cross-sectional study conducted by the Japanese Society for Cancer of the Colon and Rectum.	Tanakaya K, Furukawa Y, Nakamura Y, Hirata K, Tomita N, Tamura K, Sugano K, Ishioka C, Yoshida T, Ishida H, Watanabe T, Sugihara K.	Jpn J Clin Oncol. 45:307-10, 2015.	2015年	国内
Haplotype defined by the MLH1-93G/A polymorphism is associated with MLH1 promoter hypermethylation in sporadic colorectal cancers.	Miyakura Y, Tahara M, Lefor AT, Yasuda Y, Sugano K.	BMC Res Notes. 7:835, 2015.	2015年	国外
Comparison of clinical features between suspected familial colorectal cancer type X and Lynch syndrome in Japanese patients with colorectal cancer: a cross-sectional study conducted by the Japanese Society for Cancer of the Colon and Rectum.	Yamaguchi T, Furukawa Y, Nakamura Y, Matsubara N, Ishikawa H, Arai M, Tomita N, Tamura K, Sugano K, Ishioka C, Yoshida T, Moriya Y, Ishida H, Watanabe T, Sugihara K.	Jpn J Clin Oncol. 45:153-9, 2015.	2015年	国内
家族性腫瘍コーディネーター・家族性腫瘍カウンセラー(FCC)制度の歩みと今後.	菅野康吉	家族性腫瘍 15: 20-24, 2015.	2015年	国内

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

obtained mostly from Japanese in this study, while those were from Europeans/Americans in the previous studies. Therefore, it was also possible that such differences could be due to the ethnic and geographic differences of patients. Similarities in mutation frequencies and diversities in spectra of mutated genes among three studies indicated that genes commonly and frequently mutated in SCLC could be further selected by combining these data. Therefore, we further selected genes mutated in ≥10% of 95 SCLC cases, and also in ≥10% of 38 cases in this study and of 27 and 30 cases in two previous studies, respectively. A total of 38 genes were identified as being frequently mutated (≥10%) in all three studies as well as in a total of 95 cases (Figure 1 and Supplementary Table S3). There was a low variability in the mutation frequencies of the 38 genes among three studies, and *TP53* and *RB1* were the first and second most frequently mutated genes in all studies. Thus, we concluded that, in addition to *TP53* and *RB1*, there are 36 genes that are frequently mutated in SCLCs irrespective of the geographic and ethnic differences of patients. *PTEN* was not included in the 38 genes, because of low frequencies of mutations (<10%) in two other studies.

Mutated Genes with Expression

MutSigCV analysis identified only three genes as being significantly mutated in SCLC. However, it is possible that there are additional genes whose mutations are also involved in SCLC development among the 36 genes frequently mutated in SCLCs. Indeed, four of the 36 genes, *TMEM132D*, *HCN1*, *SPHKAP* and *COL11A1*, were identified as being significantly mutated in SCLCs by Peifer et al. [6], and two of the 36 genes, *COL22A1* and *TMEM132D*, were also identified by Rudin et al. [7]. Expression of the mutated allele is an important factor to support the significance of mutations in cancer development and to consider target therapies against mutated gene products. We previously performed whole transcriptome sequencing in 19 cases analyzed in this study [8]. Therefore, transcripts from mutated alleles were searched for in the sequencing data of 36 frequently mutated genes in addition to *TP53*, *RB1* and *PTEN* (Figure 2). Expression of mutated alleles in the *TP53* and *RB1* genes was validated in several SCLC cases. *PTEN* was mutated only in one case, and expression of the mutated allele was not detected in this tumor. In addition, mutated alleles were expressed in eight of the 36 genes, *TMEM132D*,

SPTA1, *VPS13B*, *CSMD2*, *ANK2*, *ASTN1*, *ASPM* and *FBN3*, in at least one of the 19 tumors (Figure 2).

Mutated Genes in Stage I SCLCs Before Chemotherapy/Radiotherapy

Driver mutations essential for SCLC development and its maintenance should be detected in early stage tumors before chemotherapy/radiotherapy. Therefore, we next analyzed the timing of mutation occurrence by considering the pathological stages and the effect of chemotherapy/radiotherapy. Average numbers of mutated genes tended to increase with stage progression (Supplementary Figure S5), and significant difference in the number of mutated genes was observed between stage I tumors and stage II-IV tumors (p=0.048 by the Student's t-test). The mean number of mutated genes in stage I tumors was 177.7, while that in stage II-IV tumors was 245.9 (Figure 3A). The result was consistent with the concept of sequential accumulation of somatic mutations during tumor progression [19-21]. A total of 727 genes were mutated commonly in both stage I and stage II-IV tumors, while other 821 genes were mutated only in stage I tumors. It was noted that 4,121 (72.7%) of the 5,669 genes identified in this study were mutated only in stage II-IV tumors but not in stage I tumors. This result strongly indicates that most of mutations detected in SCLC cells have occurred during tumor progression, thus, may not be necessary for tumor development, although the small number of early stage tumors may influence this assumption. *TP53* and *RB1* were mutated in any progression stages of tumors, supporting their importance in tumor development and its maintenance. On the other hand, *PTEN* mutations were detected only in stage II-IV tumors. Among eight frequently mutated and expressed genes, four genes, *TMEM132D*, *SPTA1*, *VPS13B*, *ASPM*, were mutated in both stage I and stage II-IV tumors, while the remaining four genes, *CSMD2*, *ANK2*, *ASTN1*, *FBN3*, were mutated only in stage II-IV tumors but not in stage I tumors. Therefore, although the frequency of their mutations were not significantly different between stage I tumors and stage II-IV tumors, the *TP53*, *RB1*, *TMEM132D*, *SPTA1*, *VPS13B* and *ASPM* genes could play more important roles in SCLC development than other genes.

We further investigated whether the 5,669 genes were mutated before or after chemotherapy/radiotherapy (Supplementary Figure S6). In case 10, mutation data in two tumors obtained before and after chemotherapy were analyzed separately. Importantly, 722 genes were mutated only in tumors after

chemotherapy/radiotherapy, and 38 of them showed significantly higher mutation frequencies in tumors after therapy than in tumors before therapy (data not shown). Therefore, it is possible that these mutations have occurred after therapy. However, none of the eight genes frequently mutated and expressed in SCLCs and the *TP53*, *RB1* and *PTEN* genes showed higher mutation frequencies in tumors after therapy than in tumors before therapy. In particular, mutations in the *VPS13B* and *ASPM* genes were detected only in tumors before therapy and not in tumors after therapy. Thus, none of the 11 genes was considered to be secondary mutated after or during chemotherapy/radiotherapy.

Genes Mutated in Both Primary Tumors and Metastases

Driver mutations essential for SCLC development and its maintenance should be detected in both primary tumors and metastases. Both primary tumors and metastases were analyzed in three cases; therefore, mutation data in primary tumors and metastases in these cases were analyzed separately. The mean number of mutated genes in 28 primary tumors was 194.1, while that in 13 metastases was 298.0. Although the difference was not statistically significant ($p=0.073$ by the Student's *t*-test), metastases tended to carry more mutations than primary tumors (Figure 3B). This result was also consistent with the concept of sequential accumulation of somatic mutations during tumor progression [19-21]. Although 1,622 of the 5,669 genes were mutated only in the metastases, seven genes frequently mutated and expressed in SCLCs and the *TP53*, *RB1* and *PTEN* genes were commonly mutated in both primary tumors and metastases. In contrast, *ASPN* mutations were detected only in primary tumors but not in metastases suggesting their possibility of being passengers.

Genetic Heterogeneity among Multiple Tumors in the Same Patients

We next compared the accumulated genetic alterations among multiple tumors in four cases. The results strongly support the concept of clonal and parallel evolution of primary tumors and metastases [19-21]. Namely, common and unique mutations were detected between primary tumors and metastases, and also between metastases of different organs (Figure 4; Supplementary Figure S7). In particular, in the analysis of the primary tumor and three different metastases from a single patient, drastic heterogeneity in the accumulated mutations was observed among the

four tumors (Figure 4). The *TP53*, *TMED132D* and *SPTA1* genes, which were defined as being frequently mutated and expressed in SCLCs, were mutated in all the tumors. In contrast, a mutation of another gene frequently mutated and expressed in SCLCs, *FBN3*, was detected only in a hilar lymph node metastasis but not in the primary tumor, a liver metastasis, and a para-aortic lymph node metastasis. Therefore, *FBN3* mutation could have occurred during tumor progression after *TP53*, *TMED132D* and *SPTA1* mutations.

DISCUSSION

MutSigCV analysis revealed that *TP53*, *RB1* and *PTEN* were significantly mutated genes in SCLC. These genes are well-known tumor suppressors involved in the development of various cancers, including SCLC. Therefore, the present results further support that these genes are involved in SCLC development. However, this analysis failed to identify various other genes previously identified as being significantly mutated in SCLCs [6,7]. Therefore, we further selected genes frequently mutated and expressed in SCLCs and investigated their clinical significance by analyzing their mutation status in various stages of SCLCs. As in the case of *TP53* and *RB1*, the *TMEM132D, SPTA1 and VPS13B* genes were mutated in any stage of tumor progression and their mutated alleles were expressed in tumors. Recent comparative genomic analyses of primary and metastatic non-small cell lung cancers and colorectal cancers revealed high concordance rates of driver alterations and low concordance rates of likely passenger alterations [22,23]. Therefore, in addition to *TP53*, *RB1* and *PTEN, TMEM132D, SPTA1 and VPS13B* could be also mutated as drivers in SCLC development. In particular, *TMED132D* was considered as being a significantly mutated gene in the two previous studies [6,7]. Since mutated products from the *TMEM132D, SPTA1 and VPS13B* genes are often expressed in SCLCs, those products could be novel therapeutic targets in SCLC patients.

In this study, *TP53* mutations were not detected in 8 of the 38 cases (21.2%), and *RB1* mutations were not detected in 10 of the 38 cases (26.3%). The results indicate the presence of a noble subset of SCLCs without *TP53* mutations and/or *RB1* mutations. Therefore, it is very important to clarify whether the *TP53* and *RB1* genes are inactivated or not inactivated in tumors without detectable mutations by exome sequencing. We previously performed SNP array analysis in 19 of the 38

SCLC cases analyzed in this study [8]. Therefore, the presence/absence of loss of heterozygosity (LOH) at the *TP53* and *RB1* loci were examined in tumors without *TP53* and/or *RB1* mutations (Supplementary Table S6). LOH of the *TP53* locus was detected in 4 of 5 cases without *TP53* mutations, and LOH of the *RB1* locus was detected in 3 of 4 cases without *RB1* mutations. Therefore, it is likely that undetectable *TP53* and *RB1* mutations are present in the cases without their mutations. However, it is also possible that there is still a noble small subset of SCLCs without mutations of the *TP53* and/or *RB1* genes. Further molecular analyses, such as whole genome sequencing and methylation array analysis in combination with mRNA expression array analysis, will be able to reveal the presence/absence of a noble subset of SCLCs.

Similar studies were also performed to predict whether the *TMEM132D, SPTA1 and VPS13B* genes function as oncogenes or tumor suppressor genes. LOH, loss of wild type allele as well as nonsense mutation were found in the *TMEM132D* gene, while missense mutations without allelic losses or with chromosomal gain were common in the *SPTA1* and *VPS13B* genes. Therefore, the *TMEM132D* gene is likely to function as a tumor suppressor gene, whereas the *SPTA1* and *VPS13B* genes are likely to function as oncogenes. However, since involvement of the *TMEM132D, SPTA1 and VPS13B* genes in cancer development is unknown at present, functional studies will be necessary to elucidate the pathogenic significance of their mutations in SCLC development. The *TMEM132D* (transmembrane protein 132D) gene encodes a single-pass transmembrane protein, the *SPTA1* (spectrin, alpha, erythrocytic 1) gene encodes an actin crosslinking and molecular scaffold protein that links the plasma membrane to the actin cytoskeleton, spectrin, and the *VPS13B* (vascular protein sorting 13 homolog B) gene also encodes a potential transmembrane protein (NCBI databases: http://www.ncbi.nlm.nih.gov). The results indicate that mutated proteins are expressed on the membrane of SCLC cells. Therefore, although functional roles of these mutated gene products are presently unknown, the products could be appropriate targets of immunotherapy in SCLC patients.

Our study has several limitations. The number of cases analyzed in this study was small; therefore, we mainly analyzed the status of genes mutated in ≥10% of SCLCs. For the identification of targetable mutations in SCLCs, we should also analyze the status of genes with low frequency of mutations in SCLCs. As reported