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

previously [6,7], we also found mutations in the *CREBBP* (2 cases), *EP300* (2 cases), Ephrin family (19 cases in 10 genes), and FLT family (7 cases in 3 genes) genes with low frequencies (<10%). Therefore, it is absolutely necessary to analyze a larger number of SCLCs to define the prevalence and clinical significance of their mutations. However, it is difficult to collect SCLC samples for molecular analysis, the present results will be highly informative to further define genes targetable for therapy of SCLC patients.

AUTHORS'S DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Employment, Leadership, or Stock Ownership: None

Honoraria: Ryo Nishikawa, EISAI; Rui Yamaguchi and Seiya Imoto; Fujifilum

Consultant or Advisory Role: Ryo Nishikawa, Roche

Speaker's Bureau: Ryo Nishikawa, Chugai

Research Funding, or Patents, Royalties, Other Intellectual Propety: None

AUTHOR CONTRIBUTIONS

Conception and design: Reika Iwakawa, Takashi Kohno, Jun Yokota

Collection and assembly of data: Reika Iwakawa, Yasushi Totoki, Tatsuhiko Shibata, Katsuya Tsuchihara, Sachiyo Mimaki, Koji Tsuta, Yoshitaka Narita, Ryo Nishikawa, Masayuki Noguchi, Curtis C. Harris, Ana I. Robles

Data analysis and interpretation: Reika Iwakawa, Rui Yamaguchi, Seiya Imoto, Satoru Miyano, Hirohiko Totsuka, Teruhiko Yoshida

Manuscript writing: All authors

Final approval of manuscript: All authors

Administrative support: Teruhiko Yoshida

REFERENCES

1. Jemal A, et al. (2011) Global cancer statistics. *CA Cancer J Clin.*, **61**, 69-90.
2. Kalemkerian GP, et al. (2013) Small cell lung cancer. *J Natl Compr Canc Netw.*, **11**, 78-98.
3. van Meerbeeck JP, et al. (2011) Small-cell lung cancer. *Lancet*; **378**, 1741-1755.
4. Pleasance ED, et al. (2010) A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature*, **463**, 184-90.
5. Sos ML, et al. (2012) A framework for identification of actionable cancer genome dependencies in small cell lung cancer. *Proc Natl Acad Sci USA*, **109**, 17034-17039.
6. Peifer M, et al. (2012) Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nat Genet.*, **44**, 1104-1110.
7. Rudin CM, et al. (2012) Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nat Genet.*, **44**, 1111-1116.
8. Iwakawa R, et al. (2013) Genome-wide identification of genes with amplification and/or fusion in small cell lung cancer. *Genes Chromosomes Cancer*; **52**, 802-816.
9. Sobin LH, et al. (2009) International Union Against Cancer: TNM Classification of Malignant Tumours, 7th Ed. West Sussex: Wiley-Blackwell.
10. Travis WD, et al. (2004) World Health Organization classification of tumors: pathology and genetics; tumours of the lung, pleura, thymus and heart. IARC Press: Lyon, France.
11. Phelps RM, et al. (1996) NCI-Navy Medical Oncology Branch cell line data base. *J Cell Biochem Suppl.*, **24**, 32-91.
12. Iwakawa R, et al. (2011) MYC amplification as a prognostic marker of early-stage lung adenocarcinoma identified by whole genome copy number analysis. *Clin Cancer Res.*, **17**, 1481-1489.
13. Shiraishi Y, et al. (2013) An empirical Bayesian framework for somatic mutation detection from cancer genome sequencing data. *Nucleic Acids Res.*, **41**, e89.
14. Lawrence MS, et al. (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*, **499**, 214-218.
15. Kohno T, et al. (2012) KIF5B-RET fusions in lung adenocarcinoma. *Nat Med.*, **18**, 375-377.

16. Li H, et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **16**, 2078–2079.
17. Li, H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, **21**, 2987-2993.
18. Okayama H, et al. (2012) Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer Res.*, **1**, 100-111.
19. Burrell RA, et al. (2013) The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*, **501**, 338-345.
20. Yokota J. (2000) Tumor progression and metastasis. *Carcinogenesis*, **21**, 497-503.
21. Takahashi K, et al. (2007) Clonal and parallel evolution of primary lung cancers and their metastases revealed by molecular dissection of cancer cells. *Clin Cancer Res.*, **13**, 111-120.
22. Vignot S, et al. (2013) Next-generation sequencing reveals high concordance of recurrent somatic alterations between primary tumor and metastases from patients with non-small-cell lung cancer. *J Clin Oncol.*, **31**, 2167-2172.
23. Vakiani E, et al. (2012) Comparative genomic analysis of primary versus metastatic colorectal carcinomas. *J Clin Oncol.*, **30**, 2956-2962.

Figure legends

Figure 1. Numbers of genes frequently mutated in SCLC: Comparison between this study and two previous studies (a, ref 6; b, ref 7). Genes with mutation frequencies ≥10% of cases were selected in each study. Then, similarities and differences of genes selected were compared among three studies.

Figure 2. Expression of mutated alleles in significantly or frequently mutated genes in SCLC. Twenty tumors from 19 patients were analyzed by both exome sequencing and RNA sequencing. Types of mutations detected in tumors by exome sequencing are indicated by colored boxes, and expression of mutated alleles validated by RNA sequencing (RNA-seq validated) is indicated by black open squares.

Figure 3. Number of genes mutated in early and late stages of SCLCs. (A) Comparison of mutated genes between stage I tumors and stage II-IV tumors. (B) Comparison of mutated genes between primary tumors and metastases. Mutation status of three significantly mutated genes and 8 frequently mutated and expressed genes are indicated in the boxes.

Figure 4. Heterogeneity in the accumulated mutations among multiple tumors in a single SCLC patient. The primary tumor (1591T), a hilar lymph node metastasis (1592M), a liver metastasis (1594M), and a para-aortic lymph node metastasis (1595M) from a single patient were analyzed for accumulated mutations by exome sequencing. Numbers in each region indicate the numbers of genes with mutations detected in tumors. Mutation status of one of significantly mutated gene, *TP53*, and three of frequently mutated and expressed genes, *TMEM132D*, *SPTA1* and *FBN3*, are indicated in the boxes.

Table 1. Clinicopathological characteristics of 38 SCLC cases

Characteristic	No
Gender Male/Female	10/28
Age Median (range)	67 (56-89)
Smoking status Ever/Never	33/0 (5 unknown)
Brinkman index Median (range)	1000 (0-2040)
Pathological stage I/II/III/IV	10/8/10/10
Sampling at Surgery/Autopsy	36/2
Treatment before sampling -/+	31/8*

*In one case, tumors were obtained before and after chemotherapy.