In most ongoing cancer exome studies, normal tissue counterparts have been sequenced in parallel with cancer tissue [15-19]. This is assumed to be necessary because germline variants must be excluded from the full set of SNVs to detect the somatic SNVs that are unique to cancers. However, the sequencing of normal tissue counterparts increases the cost and time of the analysis. Also, in some cases, it is difficult to obtain normal tissue counterparts. In addition, it remains unclear how accurately germline SNVs can be excluded using normal tissue exomes. To conservatively exclude germline SNVs, their sequence depths and accuracies may need to be greater than those that are obtained from the cancer exomes.

In this study, we generated and analyzed 97 cancer exomes from Japanese lung adenocarcinoma patients. We also demonstrate that somatic SNVs can be enriched to a level that is sufficient for further statistical analyses even in the absence of the sequencing of normal tissue counterparts. To separate the germline from the somatic SNVs, we first compared the variation patterns between a cancer exome with the 96 other patients' normal tissue exomes. We also attempted to conduct a similar mutual comparison solely utilizing cancer exomes, without the consideration of exomes of normal tissue counterparts. It is true that if we completely omitted normal tissue sequencing, we would tentatively disregard of somatic mutations that occurs at exactly the same genomic position in multiple cancers. However, recent papers have elucidated that such shared SNVs are very rare [15,20-22]. Moreover, many of these recursively mutations have been registered in the cancer somatic mutation databases such as Sanger COSMIC [23,24], and those recurrent SNVs can be recovered by follow-up studies partially using the data from the normal tissues. To understand the unique nature of each cancer, a statistical analysis of the distinct SNVs is presumed to be essential in addition to the analysis of the common SNVs.

In this study, we demonstrate that it is possible to identify the first candidates for cancer-related genes and pathways, even without the sequencing of a normal tissue counterpart. We show that this approach is useful not only to reduce the cost of the sequencing but also to improve the fidelity of the data. It should be also useful for analyzing old archive samples, for which normal tissue counterparts are not always available. Here, we describe a practical and cost-effective method to expedite cancer exome sequencing.

Results and Discussion

Characterization of SNVs using the 97 exome dataset

Firstly, we generated and analyzed whole-exome sequences from 97 Japanese lung adenocarcinoma patients. Exome data were collected from both cancer and normal-tissue counterparts, separated by laser capture microdissection. We purified the exonic DNA (exomes) and generated 76-base paired-end reads using the illumina GAIIx platform. Approximately 30 million mapped sequences were obtained from each sample, providing 74× coverage of the target regions; 93% of the target regions had 5× coverage (Figure S1 in File **S1**). Burrows-Wheeler Aligner (BWA) [25] and the Genome Analysis Toolkit (GATK) [26,27] were used to identify SNVs (Figure S2 in File **S1**). Only SNVs that were detected in cancer tissues and showed no evidence of variation in normal tissues were selected for further analysis.

The obtained dataset was used to characterize the cancer-specific mutation patterns (Table S3 in File **S1**). We calculated the enrichment of the SNVs within particular genes, protein domains, functional categories, and pathways. We searched for genes with somatic SNVs significantly enriched in Japanese lung adenocarcinoma. As shown in Table S4 in File **S1**, several genes were identified as significantly mutated. In particular, we searched for domains that are enriched with SNVs and harbor known cancer-related mutations in the COSMIC database. In total, 11 genes were identified ($P < 0.02$, Table **1**). For example, the Dbl homology (DH) domain of PREX1 gene [28] was enriched with SNVs ($P = 0.00071$). However, in the PREX2 gene [29], the Pleckstrin homology (PH) domain was enriched with SNVs ($P = 0.011$) (Figure 1A and B). Both the PREX1 and the PREX2 genes activate the exchange of GDP to GTP for the Rho family of GTPases and the DH/PH domains are indispensable for nucleotide exchange of GTPases and its regulation [30-32]. In addition, we analyzed the expression patterns of these genes using a cancer gene expression database, GeneLogic (Figure S3 in File **S1**). Expression levels of PREX1 and PREX2 were not enhanced in lung adenocarcinoma but were enhanced in wide variety of cancers, which is partly indicated in previous studies [33]. The SNVs in the PREX1 and PREX2 genes, which were concentrated at its pivotal signaling domains, might enhance activities in these genes, and thereby functionally mimics the increased expressions of this gene in some different types of cancers. The cancer-related gene candidates identified from this dataset are listed in Table **1**.

Similarly, pathway enrichment analyses using the KEGG database [34] also detected several putative cancer-related pathways. The identified pathways are listed in Table **2**. Interestingly, the endometrial cancer pathway [35] was detected in this enrichment analysis ($P = 3.1e\text{-}15$, Figure **2A**). This pathway includes major cancer-related pathways, for example, the MAPK signaling pathway and the PI3K/AKT pathway. For this pathway, we compared mutation patterns between our Japanese data and those of the previous study of lung adenocarcinoma in Caucasians [21]. We found that the SNVs in the EGFR gene were four times more frequent in the Japanese population than among Caucasian populations (Figure **2B**, left panel). EGFR mutations were frequently occurring in non-smoker, female and Asian patients of lung adenocarcinoma [36], which is a molecular target of anti-cancer drug, *gefitinib* [20,37,38]. Conversely, KRAS mutations, which are also well-known cancer-related mutations [39], were more than four times frequent among Caucasians (Figure **2B**, center panel). However not all mutational patterns are different between populations. For instance, TP53 harbored mutations in both datasets with similar frequency (Figure **2B**, right panel).

Ambiguity in SNV identification of normal tissue counterparts

In the aforementioned analysis, we discriminated germline variants using the normal tissue counterparts. A number of

Table 1. List of the identified possible cancer-related genes.

		Number of SNVs		
Gene	Domain	Domain	Gene	P-value*
EGFR†	IPR001245:Serine-threonine/tyrosine-protein kinase	34	37	4.4e-21
KRAS†	IPR001806:Ras GTPase	6	7	8.0e-6
TNN	IPR003961:Fibronectin, type III	4	5	5.2e-5
TP53†	IPR008967:p53-like transcription factor, DNA-binding	20	23	9.5e-5
PREX1	IPR000219:Dbl homology (DH) domain	4	5	0.00071
DNAH7	IPR004273:Dynein heavy chain	5	7	0.0025
FSTL5	IPR011044:Quinoprotein amine dehydrogenase, beta chain-like	7	7	0.0043
NRXN3	IPR008985:Concanavalin A-like lectin/ glucanase	5	7	0.0063
PREX2	IPR001849:Pleckstrin homology	3	7	0.011
FER1L6	IPR008973:C2 calcium/lipid-binding domain, CaLB	3	6	0.013
COL22A	IPR008985:Concanavalin A-like lectin/ glucanase	3	6	0.015

* $P < 0.02$

† Reported in the Cancer Gene Census [11]. Note that the genes atop the list are previously reported to be associated with this cancer type, while most of them are novel possible cancer-related genes.

doi: 10.1371/journal.pone.0073484.t001

SNVs initially identified as somatic were also found to be present in normal tissues, thus, were false positive calls under the validations by visual inspection of the mapped sequences and Sanger sequencing. To examine the cause of this problem, we inspected the errors in randomly selected 26 cancers and their normal tissues. On average in each cancer, twenty-five percent of somatic SNV candidates were found to be false positive (Figure **3**). In these cases, the sequence coverage and quality of the normal counterpart were not sufficient. Indeed, the sequences supporting each SNV and these qualities were significantly diverged between the cancer and normal tissues. Although we increased the total number of reads in the normal tissues, it was difficult in practice to cover all of the genomic positions (Figure S4 in File **S1**). A summary of the germline SNV validations is shown in Table S5 in File **S1**.

However, we noticed that some were correctly identified as germline SNVs in external reference exomes. Twenty-five exomes allowed us to exclude eight false positive calls in each cancer. This raised the possibility that the SNVs from the other patients may be used as surrogates to increase the depth and quality of the sequencing.

Excluding germline SNVs by considering mutual overlaps of other persons' exomes

To further test this possibility, we examined whether cancer exome analyses would be possible without sequencing of the normal tissue counterpart of each cancer. First, we evaluated the extent to which the germline SNVs could be discriminated using external exomes. For this purpose, we used the 97 paired cancer-normal exome datasets for the validation dataset. We found that we could detect 54% of the germline SNVs by using the 96 normal tissue exomes from the external reference (Figure **4A**). We further expanded the filtration dataset using the externally available 73 Japanese exome data and 48 in-house Japanese exome datasets. Altogether, we were able to remove 64% of the germline SNVs, using a total of 217 Japanese exome datasets from other individuals, without sequencing each cancer's normal counterpart (Figure **4A**). The extrapolation of the graph also indicated that 1,350 and 2,000 samples would be required to remove 90% and 95% of the germline SNVs, respectively. We expect that such a sample size will be available in near future considering current rapid expansion of the exome analysis.

We further evaluated if the same filtration could be done by solely using cancer exomes. We obtained essentially the same results (Figure S5 in File **S1**). Obvious caveat of this approach is that this would disregard about 3% of somatic SNVs recurrently occurring (Figure S5 in File **S1**, blue). However, as aforementioned, we found that those recurrent SNVs were very rare [15,19] and most of them were derived from dubious somatic SNVs, which were overlooked in the normal tissues. We also consider that most of those recurrent SNVs, if any, can be analyzed separately by sequencing a limited number of normal tissues.

Filtering out germline SNVs by considering mutual overlaps for different ethnic groups and for rare SNPs

We examined whether SNVs in other ethnic backgrounds could be used as external datasets for the filtration. We obtained exome data from individuals of various ethnic backgrounds from the 1000 Genome Project. We used these exome datasets to exclude the germline SNVs that were identified in the Japanese cancers. We found that the discriminative power was significantly lower compared with exomes from Japanese populations. Therefore, these datasets were not suitable for this purpose (Figure **4B**). We also examined and found that the exomes in each ethnic group were useful to discriminate the germline SNVs in the corresponding group (Figure S6, S7 and Table S6 in File **S1**).

We, then, examined to what extent minor germline variants could be covered with this approach in the Japanese population. We evaluated the sensitivity of the filtration process for the SNVs in the 97 cancers (Figure S8 in File **S1**). We found that 88% of the germline SNVs occurring in more than five percent of the 97 exomes could be detected using the 73 external Japanese datasets. For the SNVs occurring in 1% of the 97 cancers, 19% could be excluded.

Using the crude dataset to characterize cancer related SNVs and pathways

Taken together, with 217 Japanese exomes used for filtration, 36% of the germline SNVs remained unfiltered. Nevertheless, we considered that it may be still possible to use the crude SNV dataset as a first approximation for identifying and analyzing cancer-related genes and pathway candidates. To validate this idea, we compared the results of enrichment analyses between the crude dataset and the refined somatic

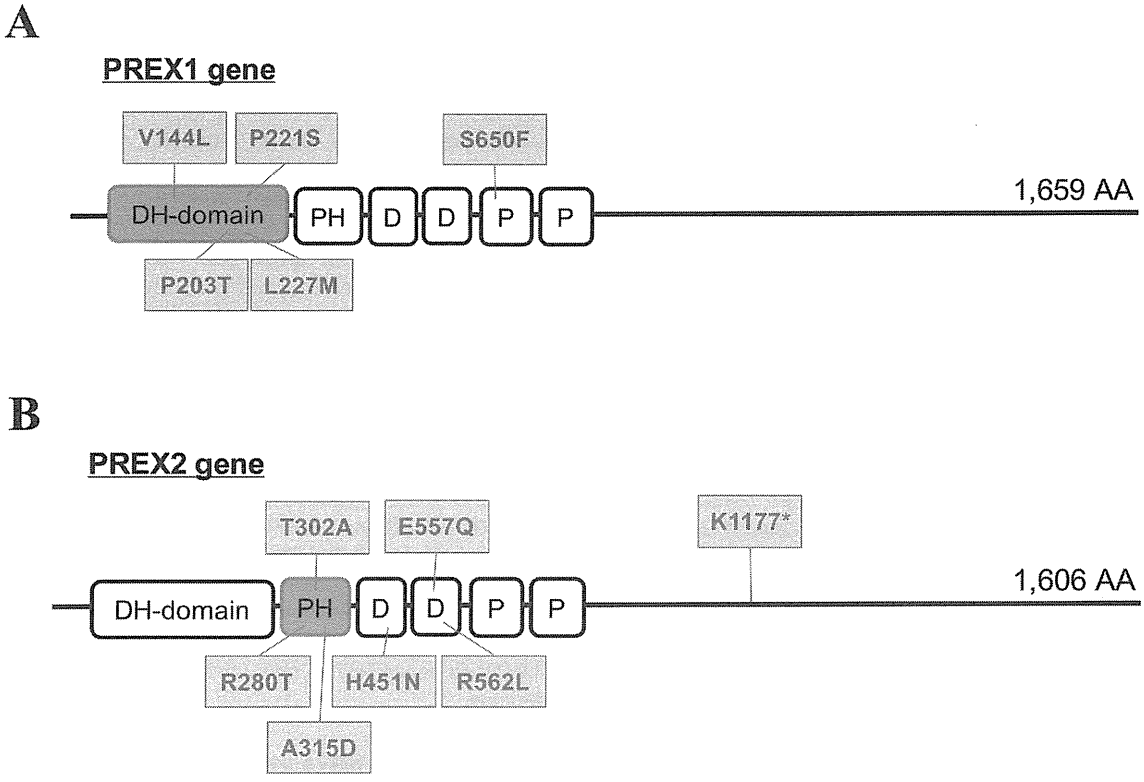

Fig. 1

Figure 1. Identification and characterization of the putative cancer-related genes using 97 cancer exomes. SNVs in the PREX1 (**A**) and PREX2 (**B**) genes are represented in the boxes. The protein domains in which the enrichments of the SNVs were statistically significant are represented in orange boxes (also see Materials and Method). DH-domain: Dbl homology (DH) domain; PH: Pleckstrin homology domain; D: DEP domain; P: PDZ/DHR/GLGF.

doi: 10.1371/journal.pone.0073484.g001

SNV datasets, which were generated from the paired cancer-normal exomes.

Most of the putative cancer-related genes and pathways that were identified from the refined dataset were also present in the crude dataset (Tables S7 and S8 in File **S1**). The example of the TNN gene, which was reported as a marker of tumor stroma [40-42], is shown in Figure S9 in File **S1**. In this case, even with the germline SNVs, which were unfiltered in the crude dataset (indicated by black in Figure S9 in File **S1**), the enrichment of somatic SNVs in this domain was statistically significant. In total, nine genes which identified as possessing cancer-related SNVs from the refined dataset were also detected in the crude dataset. On the other hand, two genes from the refined dataset were not represented in the crude dataset. In the pathway analysis, we identified 26 cancer-related pathways which were identified from the refined dataset. In addition, 19 pathways were also represented in the crude dataset as well as the refined dataset. The overlap between the datasets is summarized in Table **3**. It should be noted that statistically enrichment analyses were possible even at the current coverage of the filter dataset. With the expanded external dataset, it would be more practical to subject the candidates to the results of Sanger sequencing validations as well as removing remaining germline SNVs.

Identification of prognosis related genes by using the crude dataset

As one of the most important objectives of the cancer exome studies, we investigated whether mutations affecting cancer prognoses can be identified by using crude dataset (Table S9 and Figure S10 in File **S1**). In the Kaplan-Meier analysis, seven patients who carried SNVs in the ATM gene (Figure **5A**)

Table 2. List of the identified possible cancer-related pathways.

KEGG ID	Pathway definition	Number of cancers with SNVs	P-value*
hsa05213	Endometrial cancer	72	3.1e-15
hsa04320	Dorso-ventral axis formation	48	4.4e-15
hsa05219	Bladder cancer	62	4.9e-14
hsa05223	Non-small cell lung cancer	66	7.1e-12
hsa05214	Glioma	70	6.5e-11
hsa05218	Melanoma	70	1.3e-9
hsa05212	Pancreatic cancer	68	6.9e-9
hsa05215	Prostate cancer	71	4.3e-7
hsa05216	Thyroid cancer	36	1.1e-6
hsa04520	Adherens junction	59	3.7e-6
hsa05210	Colorectal cancer	53	1.8e-5
hsa04012	ErbB signaling pathway	64	2.6e-5
hsa05120	Epithelial cell signaling in *Helicobacter pylori* infection	53	4.8e-5
hsa04540	Gap junction	60	0.00024
hsa04912	GnRH signaling pathway	61	0.0011
hsa05217	Basal cell carcinoma	41	0.0020
hsa05222	Small cell lung cancer	52	0.0069
hsa05220	Chronic myeloid leukemia	46	0.010
hsa05160	Hepatitis C	67	0.012
hsa05014	Amyotrophic lateral sclerosis (ALS)	36	0.014
hsa04977	Vitamin digestion and absorption	20	0.015
hsa05416	Viral myocarditis	40	0.028
hsa04512	ECM-receptor interaction	47	0.034
hsa02010	ABC transporters	29	0.035
hsa04510	Focal adhesion	78	0.037
hsa05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	40	0.039

* $P < 0.05$

doi: 10.1371/journal.pone.0073484.t002

showed statistically significant poor prognoses (P = 9.6e-6, Figure **5B**). Three SNVs in the ATM gene were significantly enriched in the the phosphatidylinositol 3-/4-kinase catalytic domain (P = 0.014). ATM senses DNA damage and phosphorylates TP53, which, in turn, invokes various cellular responses, such as DNA repair, growth arrest and apoptosis, and collectively prevents cancer progression (Figure S11 in File **S1**) [43,44].

We also examined whether other frequently mutated genes were associated with better or worse prognoses. We found that patients with PAPPA2 mutations showed prolonged survival times (P = 0.026, Figure 5C and D). PAPPA2 proteolyzes IGFBP5 [45,46], which is an inhibitory factor for IGFs [47]. Mutations in the PAPPA2 gene may result in the accumulation of IGFBP5, and the resulting decrease in IGF signaling may impair the proliferation of cancer cells [48]. Again, it should be noted that for both the ATM and PAPPA2 genes, the statistical significance of the prognostic difference persisted both before (black line) and after (red line) the remaining germline mutations were removed, which was validated by Sanger sequencing (Figure 5B, D and Table S10 in File **S1**).

Conclusions

We have identified and characterized the SNVs in lung adenocarcinoma in a Japanese population. Further biological evaluations of the discovered SNVs will be described elsewhere. In particular, information of transcriptome and epigenome should be important for further analyses of cancer genomes, as they would shed new lights on the cancer biology (Table **S1**) [49]. In this study, we also presented a useful approach for the analysis of cancer exomes, without the need to sequence the normal tissue counterpart. We believe that the approach not only lowers the barriers in cost, time and data fidelity in the exome analysis, but also enables exome analysis of archive samples, for which normal tissue counterparts are not always available.

Materials and Methods

Ethics statement

All of the samples were collected by following the protocol (and written informed consent) which were approved by Ethical Committee in National Cancer Center, Japan (Correspondence to: Katsuya Tsuchihara; ktsuchih@east.ncc.go.jp).

Case selection and DNA preparation

All of the tissue materials were obtained from Japanese lung adenocarcinoma patients with the appropriate informed consent. Surgically resected primary lung adenocarcinoma samples with lengthwise dimensions in excess of 3 cm were selected. Data on the 52 patients who had relapses and other clinical information about the 97 cases are shown in Table S11 in File **S1**. All 97 cancer and normal tissues were extracted from methanol-fixed samples by laser capture microdissection. DNA purification was performed using an EZ1 Advanced XL Robotic workstation with EZ1 DNA Tissue Kits (Qiagen).

Whole-exome sequencing

Using 1 μg of isolated DNA, we prepared exome-sequencing libraries using the SureSelect Target Enrichment System (Agilent Technologies) according to the manufacturer's protocol. The captured DNA was sequenced by the illumina Genome Analyzer IIx platform (Illumina), yielding 76-base paired-end reads.

Somatic SNV detection

The methods that were used to detect the SNVs, including BWA, SAMtools [50] and GATK, are shown in Figure S2 in File S1. Using data from NCBI dbSNP build 132 and one Japanese genome [51], major germline SNVs were excluded. In addition, rare germline SNVs were discarded using 97 exomes from normal tissue counterparts, 73 Japanese exomes provided from the 1000 Genomes Project (the phase1 exome data, 20110521) and 48 in-house Japanese exomes. We also validated a portion of the SNV datasets by the Sanger

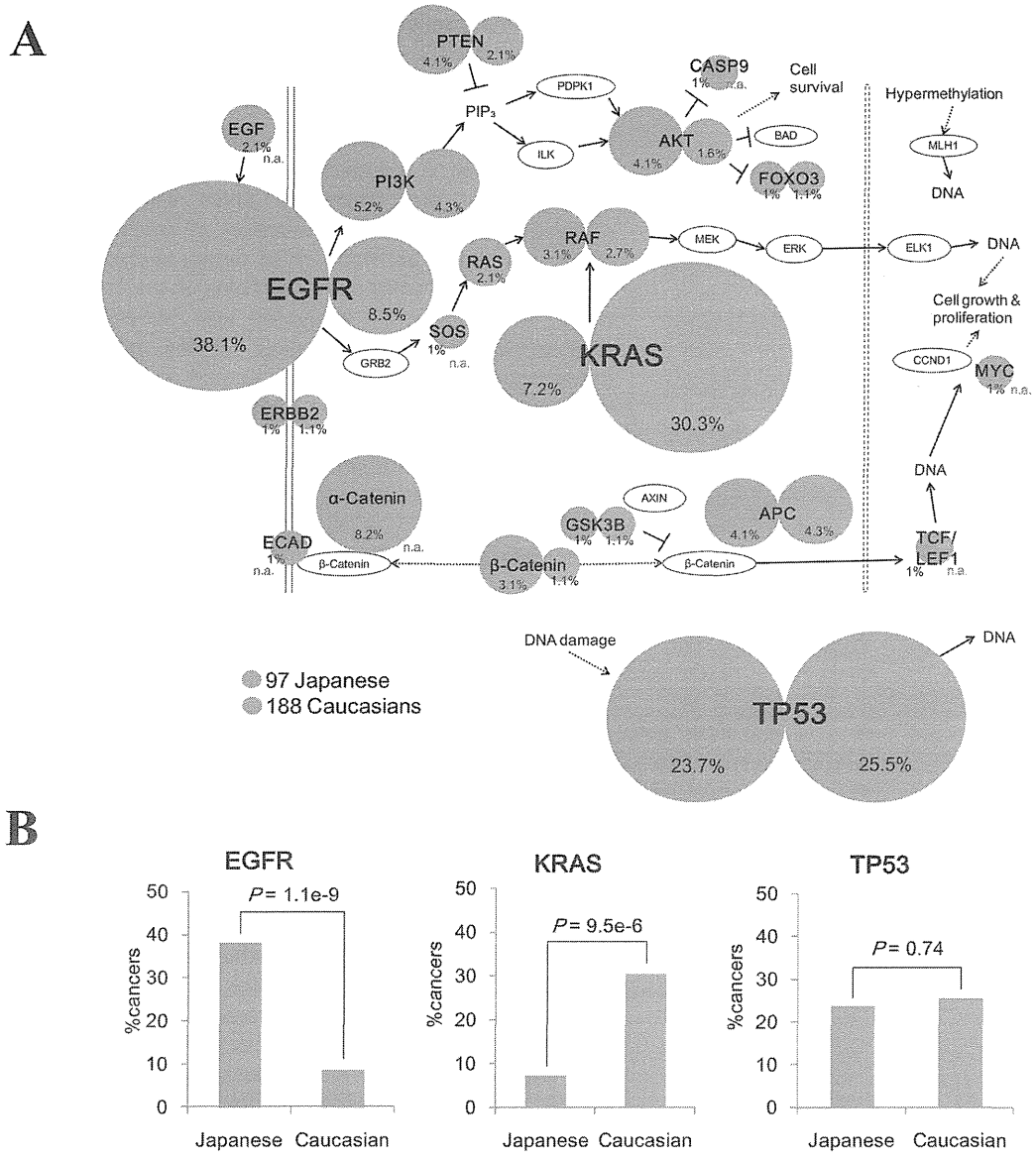

Fig. 2

Figure 2. The EGFR/Ras pathways in Japanese and Caucasian populations. (**A**) Mutation patterns in the endometrial cancer pathway that was detected in the enrichment analysis are shown. The size of the circle represents the population of the cancers harboring the SNVs in the corresponding gene (percentage is also shown in the margin). SNVs in this study and the external dataset in Caucasian populations are shown in red and blue circles, respectively. n.a.: mutation frequencies were not available. (**B**) Comparison of mutation ratio of EGFR, KRAS and TP53 genes among both datasets. The p-values were calculated by two-sample test for equality of proportions.

doi: 10.1371/journal.pone.0073484.g002

sequencing of cancer tissues and their normal tissue counterparts (Figure S12 in File **S1**).

Identification of highly mutated genes

We detected genes which were significantly enriched with SNVs by calculating the expected number of cancers with SNVs in the gene. The length of total CDS regions was represented in N (approximately 30.8 M bases). When one patient harbored total of m SNVs, the probability that the patient harbors SNVs in the gene t (length: n) was calculated as P:

$$P_{m,t,n} = 1 - \left(1 - \frac{m}{N}\right)^n$$

The sum of P in 97 cancers was represented in the expected number of cancers with SNVs in the gene t. The p-values of the

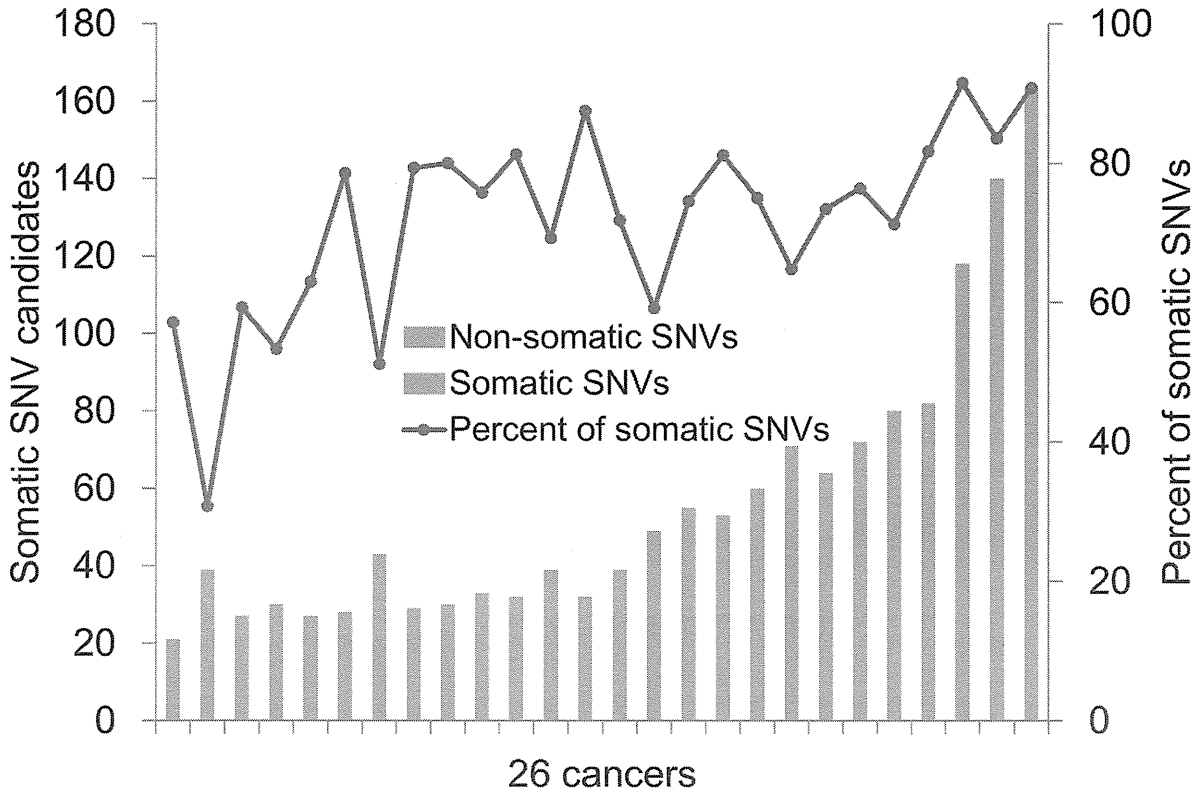

Fig. 3

Figure 3. Fidelity of the germline SNV detection in cancer exome analysis. Somatic SNV candidates were identified by using 26 cancer exomes and each normal counterpart. Correct somatic SNVs and false positives were shown in pink and blue bars, respectively. The 26 cancers used for the analysis were sorted by the increasing total number of SNVs (x-axis).
doi: 10.1371/journal.pone.0073484.g003

observed number were calculated by the Poisson probability function using R ppois.

Statistical approach to enrichment analyses

To examine the enrichment of mutations in functional protein domains, we mapped the SNVs to domains using InterProScan [52] and assigned them to the Catalogue of Somatic Mutations in Cancer (COSMIC). We analyzed the enrichment of the SNVs in the same domains as the mutations that were provided by the COSMIC. The p-values for the observed mutations in these domains were calculated using their hypergeometric distributions (R phyper). Briefly, the domains in which the SNVs were enriched statistically significantly than the expected number of SNVs in the given length of the domain were selected. For estimating the expected number, the total number of the SNVs belonging to the gene was divided by the gene length. For this analysis, we used genes harboring five or more SNVs in the coding region and three or more SNVs in the domain.

We assigned SNVs to pathways as described by the Kyoto Encyclopedia of Genes and Genomes (KEGG) and calculated the enrichments of the SNVs in the pathways. The mutation rate M represented the ratio of the average number of mutated genes to the total number of genes (17,175) that were used in our study. The expected value for the number of cancers with SNVs in pathway t was designated λ and calculated from the mutation rate M and the number of genes in the pathway n as follows:

$$\lambda_{t,n} = \left\{1-(1-M)^n\right\} \times 97$$

The p-value for the observed number of cancers with SNVs in pathway t was calculated by the Poisson probability function using R ppois.

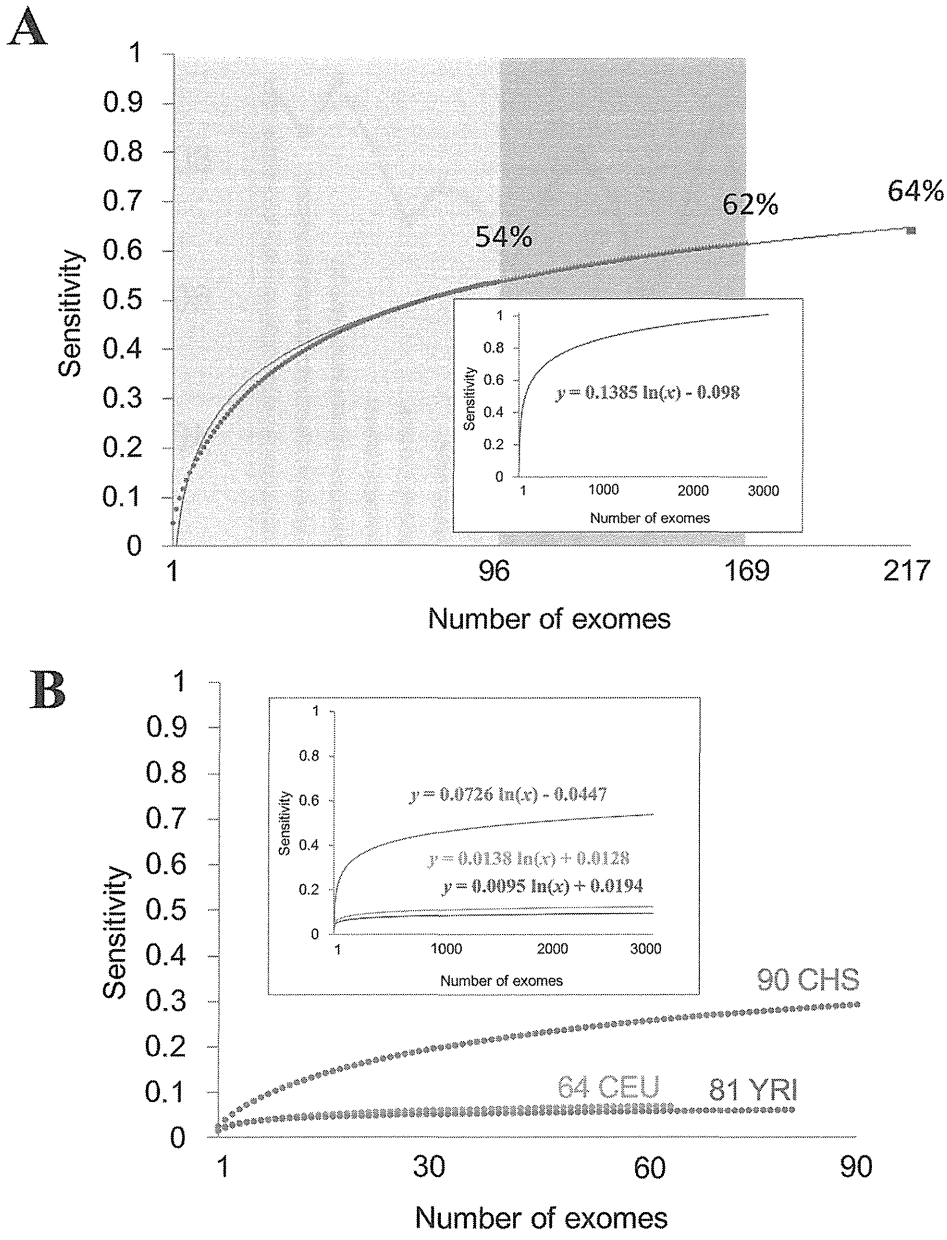

Fig. 4

Figure 4. Discriminative powers of detecting germline SNVs using external references. (**A**) The power of detecting germline SNVs considering mutual overlap between other Japanese individuals. Sensitivity represents the proportion of germline SNVs correctly detected. The datasets used to exclude the germline SNVs are shown on the x axis. The inset represents the extrapolation of the graph. Fitting curve of the graph is also shown. (**B**) Discriminative powers of three different ethnic groups for the germline SNVs in 97 Japanese cancers. Sensitivities for detecting germline SNVs are shown by the following colors; green: Chinese; purple: Yoruba; orange: Caucasian.

doi: 10.1371/journal.pone.0073484.g004

Estimate of discriminative power for exclusion of germline SNVs by considering mutual overlaps

We estimated the discriminative power for the exclusion of germline SNVs by considering those from other non-cancerous exomes. Germline SNVs from 97 paired tumor-normal exomes were used as reference datasets. Up to 217 samples (96 normal tissue exomes from others and 121 additional Japanese exomes) were randomly selected, and their sensitivities and specificities for detecting the germline SNVs were detected by taking the averages of either all of the combinations or a subset of approximately 10,000 combinations. We also estimated the discriminative power with

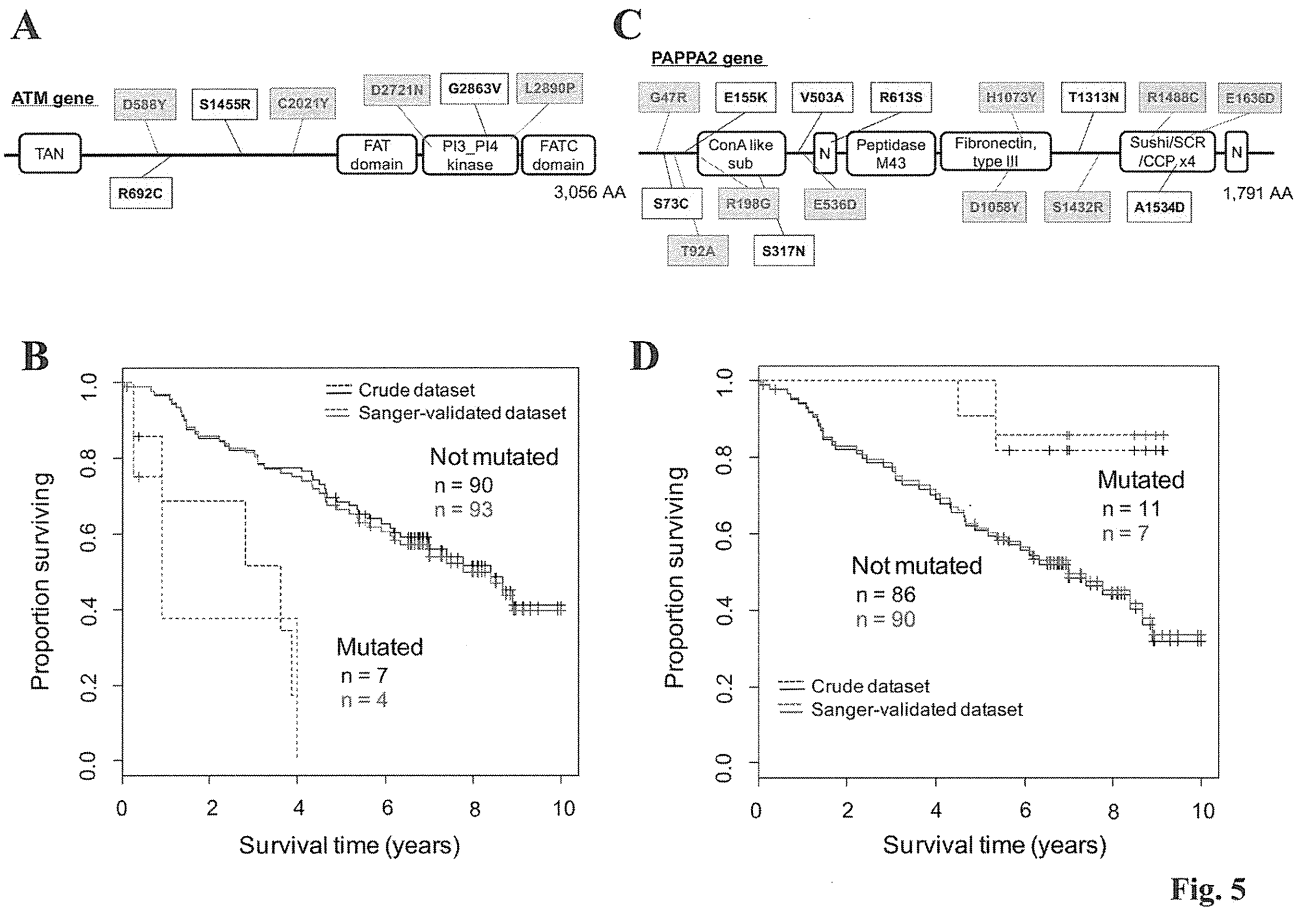

Fig. 5

Figure 5. Identification of the putative prognosis-related genes. (**A**) SNVs in the ATM gene. The SNVs that were identified in the initial screening and those remaining after the Sanger sequencing validation of the normal-tissue counterpart were shown in black and red, respectively. TAN: Telomere-length maintenance and DNA damage repair; PI3_PI4 kinase: Phosphatidylinositol 3-/4-kinase, catalytic. (**B**) Survival analysis of patients with and without ATM SNVs. The datasets before and after the Sanger sequencing validation are represented by black and red lines, respectively. Statistical significance was calculated using a log-rank test ($P < 0.05$). Note that the survival differences for individuals with SNVs in the non-Sanger-validated dataset were significant before the Sanger validation. (**C**, **D**) Results of a similar analysis as that described in A and B for the PAPPA2 gene. In this case, the patients with the SNVs showed better prognoses. ConA like sub: Concanavalin A-like lectin/glucanase, subgroup; N: Notch dimain; Peptidase M43: Peptidase M43, pregnancy-associated plasma-A.

doi: 10.1371/journal.pone.0073484.g005

Table 3. Comparison of the results in the enrichment analyses between the crude and refined dataset.

	Number of identified genes/pathways		
	Crude*	Refined†	Overlap‡
Genes	16	11	9
Pathways	23	26	19

* Identified using the crude dataset.

† Identified using the refined dataset.

‡ Significant in both crude and refined datasets.

doi: 10.1371/journal.pone.0073484.t003

data from the 1000 Genomes Project for four ethnic groups (73 JPT, 90 CHS, 81 YRI and 64 CEU) using similar trials. Whole-exome sequences (the phase1 exome data, 20110521) were obtained from the ftp site in the 1000 Genomes Project.

Kaplan-Meier curves

The Kaplan-Meier method was used to test the relations of the observed mutations to survival time, and calculations were performed using the R software package. Changes in survival rates that were correlated with SNVs were examined using the log-rank test (R survdiff).

Data access

Full raw datasets will be shared with researchers upon request. The information of somatic mutations at the respective genomic coordinates has been provided in Table **S2**.

Supporting Information

File S1. Figures S1 to S12 and Tables S3 to S11 are included.
(PDF)

Table S1. The comparison of our dataset with the other different study. We provided the comparison of our dataset with the genes identified in the other different study with transcriptome and epigenome data in lung cancers.
(XLSX)

Table S2. The list of somatic mutations identified from the refined dataset. All mutations described in this table are somatic and non-synonymous mutations.
(XLSX)

Acknowledgements

We thank W. Makałowski for constructive comments and suggestions for this manuscript. We are also grateful to F. Todokoro and F. Iguchi for their technical assistance with the data processing and N. Ogasawara for the Sanger sequencing validation.

Author Contributions

Conceived and designed the experiments: KT YS HE KG SS. Performed the experiments: SM YY AK KM MS. Analyzed the data: AS YS KT. Contributed reagents/materials/analysis tools: KG KT. Wrote the manuscript: AS KT YS.

References

1. The 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. Nature 467: 1061-1073. doi:10.1038/nature09534. PubMed: 20981092.
2. The 1000 Genomes Project Consortium (2012) An integrated map of genetic variation from 1,092 human genomes. Nature 491: 56-65. doi: 10.1038/nature11632. PubMed: 23128226.
3. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR et al. (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 106: 19096-19101. doi:10.1073/pnas.0910672106. PubMed: 19861545.
4. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW et al. (2009) Targeted capture and massively parallel sequencing of 12 human exomes. Nature 461: 272-276. doi:10.1038/nature08250. PubMed: 19684571.
5. Clark MJ, Chen R, Lam HY, Karczewski KJ, Chen R et al. (2011) Performance comparison of exome DNA sequencing technologies. Nat Biotechnol 29: 908-914. doi:10.1038/nbt.1975. PubMed: 21947028.
6. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ et al. (2011) Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 12: 745-755. doi:10.1038/nrg3031. PubMed: 21946919.
7. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK et al. (2010) Exome sequencing identifies the cause of a mendelian disorder. Nat Genet 42: 30-35. doi:10.1038/ng.499. PubMed: 19915526.
8. Biesecker LG (2010) Exome sequencing makes medical genomics a reality. Nat Genet 42: 13-14. doi:10.1038/ng0110-13. PubMed: 20037612.
9. Louis-Dit-Picard H, Barc J, Trujillano D, Miserey-Lenkei S, Bouatia-Naji N et al. (2012) KLHL3 mutations cause familial hyperkalemic hypertension by impairing ion transport in the distal nephron. Nat Genet.
10. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L et al. (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29: 308-311. doi:10.1093/nar/29.1.308. PubMed: 11125122.
11. Futreal PA, Coin L, Marshall M, Down T, Hubbard T et al. (2004) A census of human cancer genes. Nat Rev Cancer 4: 177-183. doi: 10.1038/nrc1299. PubMed: 14993899.
12. International Cancer Genome Consortium (2010) International network of cancer genome projects. Nature 464: 993-998. doi:10.1038/nature08987. PubMed: 20393554.
13. Totoki Y, Tatsuno K, Yamamoto S, Arai Y, Hosoda F et al. (2011) High-resolution characterization of a hepatocellular carcinoma genome. Nat Genet 43: 464-469. doi:10.1038/ng.804. PubMed: 21499249.
14. Jones DT, Jäger N, Kool M, Zichner T, Hutter B et al. (2012) Dissecting the genomic complexity underlying medulloblastoma. Nature 488: 100-105. doi:10.1038/nature11284. PubMed: 22832583.
15. The Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. Nature 474: 609-615. doi: 10.1038/nature10166. PubMed: 21720365.
16. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C et al. (2012) The landscape of cancer genes and mutational processes in breast cancer. Nature 486: 400-404. PubMed: 22722201.
17. The Cancer Genome Atlas Research Network (2012) Comprehensive genomic characterization of squamous cell lung cancers. Nature.
18. The Cancer Genome Atlas Research Network (2012) Comprehensive molecular characterization of human colon and rectal cancer. Nature 487: 330-337. doi:10.1038/nature11252. PubMed: 22810696.
19. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ et al. (2012) Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 150: 1107-1120. doi:10.1016/j.cell.2012.08.029. PubMed: 22980975.
20. Davies H, Hunter C, Smith R, Stephens P, Greenman C et al. (2005) Somatic mutations of the protein kinase gene family in human lung cancer. Cancer Res 65: 7591-7595. PubMed: 16140923.
21. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD et al. (2008) Somatic mutations affect key pathways in lung adenocarcinoma. Nature 455: 1069-1075. doi:10.1038/nature07423. PubMed: 18948947.
22. Kan Z, Jaiswal BS, Stinson J, Janakiraman V, Bhatt D et al. (2010) Diverse somatic mutation patterns and pathway alterations in human cancers. Nature 466: 869-873. doi:10.1038/nature09208. PubMed: 20668451.
23. Forbes SA, Bhamra G, Bamford S, Dawson E, Kok C et al. (2008) The Catalogue of Somatic Mutations in Cancer (COSMIC). Curr Protoc Hum Genet Chapter 10: Unit 10.11: PubMed: 18428421
24. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY et al. (2011) COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res 39: D945-D950. doi: 10.1093/nar/gkq929. PubMed: 20952405.
25. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754-1760. doi: 10.1093/bioinformatics/btp324. PubMed: 19451168.
26. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20: 1297-1303. doi:10.1101/gr.107524.110. PubMed: 20644199.
27. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43: 491-498. doi: 10.1038/ng.806. PubMed: 21478889.
28. Welch HC, Coadwell WJ, Ellson CD, Ferguson GJ, Andrews SR et al. (2002) P-Rex1, a PtdIns(3,4,5)P3- and Gbetagamma-regulated guanine-nucleotide exchange factor for Rac. Cell 108: 809-821. doi: 10.1016/S0092-8674(02)00663-3. PubMed: 11955434.

29. Rosenfeldt H, Vázquez-Prado J, Gutkind JS (2004) P-REX2, a novel PI-3-kinase sensitive Rac exchange factor. FEBS Lett 572: 167-171. doi:10.1016/j.febslet.2004.06.097. PubMed: 15304342.
30. Das B, Shu X, Day GJ, Han J, Krishna UM et al. (2000) Control of intramolecular interactions between the pleckstrin homology and Dbl homology domains of Vav and Sos1 regulates Rac binding. J Biol Chem 275: 15074-15081. doi:10.1074/jbc.M907269199. PubMed: 10748082.
31. Hill K, Krugmann S, Andrews SR, Coadwell WJ, Finan P et al. (2005) Regulation of P-Rex1 by phosphatidylinositol (3,4,5)-trisphosphate and Gbetagamma subunits. J Biol Chem 280: 4166-4173. PubMed: 15545267.
32. Chhatriwala MK, Betts L, Worthylake DK, Sondek J (2007) The DH and PH domains of Trio coordinately engage Rho GTPases for their efficient activation. J Mol Biol 368: 1307-1320. doi:10.1016/j.jmb.2007.02.060. PubMed: 17391702.
33. Sosa MS, Lopez-Haber C, Yang C, Wang H, Lemmon MA et al. (2010) Identification of the Rac-GEF P-Rex1 as an essential mediator of ErbB signaling in breast cancer. Mol Cell 40: 877-892. doi:10.1016/j.molcel.2010.11.029. PubMed: 21172654.
34. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M (2010) KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res 38: D355-D360. doi:10.1093/nar/gkp896. PubMed: 19880382.
35. Hecht JL, Mutter GL (2006) Molecular and pathologic aspects of endometrial carcinogenesis. J Clin Oncol 24: 4783-4791. doi:10.1200/JCO.2006.06.7173. PubMed: 17028294.
36. Li C, Fang R, Sun Y, Han X, Li F et al. (2011) Spectrum of oncogenic driver mutations in lung adenocarcinomas from East Asian never smokers. PLOS ONE 6: e28204. doi:10.1371/journal.pone.0028204. PubMed: 22140546.
37. Shigematsu H, Lin L, Takahashi T, Nomura M, Suzuki M et al. (2005) Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 97: 339-346. doi:10.1093/jnci/dji055. PubMed: 15741570.
38. Sharma SV, Bell DW, Settleman J, Haber DA (2007) Epidermal growth factor receptor mutations in lung cancer. Nat Rev Cancer 7: 169-181. doi:10.1038/nrc2088. PubMed: 17318210.
39. Bos JL (1989) ras oncogenes in human cancer: a review. Cancer Res 49: 4682-4689. PubMed: 2547513.
40. Degen M, Brellier F, Kain R, Ruiz C, Terracciano L et al. (2007) Tenascin-W is a novel marker for activated tumor stroma in low-grade human breast cancer and influences cell behavior. Cancer Res 67: 9169-9179. doi:10.1158/0008-5472.CAN-07-0666. PubMed: 17909022.
41. Degen M, Brellier F, Schenk S, Driscoll R, Zaman K et al. (2008) Tenascin-W, a new marker of cancer stroma, is elevated in sera of colon and breast cancer patients. Int J Cancer 122: 2454-2461. doi:10.1002/ijc.23417. PubMed: 18306355.
42. Brellier F, Martina E, Degen M, Heuzé-Vourc'h N, Petit A et al. (2012) Tenascin-W is a better cancer biomarker than tenascin-C for most human solid tumors. BMC Clin Pathol 12: 14. doi:10.1186/1472-6890-12-14. PubMed: 22947174.
43. Dasika GK, Lin SC, Zhao S, Sung P, Tomkinson A et al. (1999) DNA damage-induced cell cycle checkpoints and DNA strand break repair in development and tumorigenesis. Oncogene 18: 7883-7899. PubMed: 10630641.
44. Bartkova J, Horejsi Z, Koed K, Kramer A, Tort F et al. (2005) DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature 434: 864-870. doi:10.1038/nature03482. PubMed: 15829956.
45. Yan X, Baxter RC, Firth SM (2010) Involvement of pregnancy-associated plasma protein-A2 in insulin-like growth factor (IGF) binding protein-5 proteolysis during pregnancy: a potential mechanism for increasing IGF bioavailability. J Clin Endocrinol Metab 95: 1412-1420. doi:10.1210/jc.2009-2277. PubMed: 20103653.
46. Overgaard MT, Boldt HB, Laursen LS, Sottrup-Jensen L, Conover CA et al. (2001) Pregnancy-associated plasma protein-A2 (PAPP-A2), a novel insulin-like growth factor-binding protein-5 proteinase. J Biol Chem 276: 21849-21853. doi:10.1074/jbc.M102191200. PubMed: 11264294.
47. Shimasaki S, Shimonaka M, Zhang HP, Ling N (1991) Identification of five different insulin-like growth factor binding proteins (IGFBPs) from adult rat serum and molecular cloning of a novel IGFBP-5 in rat and human. J Biol Chem 266: 10646-10653. PubMed: 1709938.
48. Su Y, Wagner ER, Luo Q, Huang J, Chen L et al. (2011) Insulin-like growth factor binding protein 5 suppresses tumor growth and metastasis of human osteosarcoma. Oncogene 30: 3907-3917. doi:10.1038/onc.2011.97. PubMed: 21460855.
49. Huang T, Jiang M, Kong X, Cai YD (2012) Dysfunctions associated with methylation, microRNA expression and gene expression in lung cancer. PLOS ONE 7: e43441. doi:10.1371/journal.pone.0043441. PubMed: 22912875.
50. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079. doi:10.1093/bioinformatics/btp352. PubMed: 19505943.
51. Fujimoto A, Nakagawa H, Hosono N, Nakano K, Abe T et al. (2010) Whole-genome sequencing and comprehensive variant analysis of a Japanese individual using massively parallel sequencing. Nat Genet 42: 931-936. doi:10.1038/ng.691. PubMed: 20972442.
52. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N et al. (2005) InterProScan: protein domains identifier. Nucleic Acids Res 33: W116-W120. doi:10.1093/nar/gni118. PubMed: 15980438.

Pathology International

Pathology International 2013; **63**: 77–84 doi:10.1111/pin.12033

Original Article

Clinicopathological characteristics of EGFR mutated adenosquamous carcinoma of the lung

Toshihiro Shiozawa,[1,3] Genichiro Ishii,[1] Koichi Goto,[3] Kanji Nagai,[4] Sachiyo Mimaki,[2] Shotaro Ono,[4] Seiji Niho,[3] Satoshi Fujii,[1] Yuichiro Ohe,[3] Katsuya Tsuchihara[2] and Atsushi Ochiai[1]

[1]Pathology Division, and [2]Cancer Physiology Project, Research Center for Innovative Oncology, and [3]Divisions of Thoracic Oncology and [4]Thoracic Surgery, National Cancer Center Hospital East, Kashiwa, Chiba, Japan

Adenosquamous carcinoma of the lung (Ad-Sq) is an uncommon subtype with poor prognosis. We analyzed the clinicopathological characteristics of Ad-Sq, focusing the correlation between Epidermal Growth Factor Receptor (EGFR) mutation and clinicopathological factors. A total of 67 cases were selected from September 1992 to May 2011. EGFR mutational analysis ($n = 59$) was performed by direct sequence. We also performed immunohistochemical staining for EGFR mutated cases using the two mutation-specific antibodies for deletion and L858R. Postoperative 3-year survival rate of Ad-Sq was 58.7%, statistically worse in comparison with adenocarcinoma (58.7% vs. 78.1%, $P = 0.038$). Twenty-four percent (14/59) were positive for EGFR mutations. Patients who had never been smokers and who were lymphatic permeation positive were seen more frequently in the mutation positive group ($P = 0.035$, 0.027, respectively). Moreover, the EGFR mutated group tended to have a more positive prognosis than negative. Focusing on the pathological features, the lepidic growth pattern was more frequently seen in the positive group ($P = 0.018$). Immunoreactivity for the DEL-specific and L858-specific antibody were observed in both adenocarcinoma and squamous cell carcinoma components. Our study demonstrated that EGFR mutated Ad-Sq had similar clinicopathological features as EGFR mutated adenocarcinoma.

Key words: EGFR mutation, immunohistochemistry, lepidic growth pattern, lung adenosquamous carcinoma

Correspondence: Genichiro Ishii, MD, PhD, Pathology Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, Kahiwa, 277-0882 Chiba, Japan. Email: gishii@east.ncc.go.jp; Atsushi Ochiai, MD, PhD, Pathology Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, Kahiwa, 277-0882 Chiba, Japan. Email: aochiai@east.ncc.go.jp

Received 27 July 2012. Accepted for publication 3 January 2013.

Adenosquamous carcinoma of the lung (Ad-Sq), which is defined as a carcinoma containing components of both squamous cell carcinoma (Sq) and adenocarcinoma (Ad) and with each component comprising at least 10% of the tumor, is an uncommon subtype, accounting for 0.4 to 4% of all non-small cell lung cancer cases.[1]

Because of its rarity, the clinicopathological aspects of Ad-Sq have not received much attention, and the pathogenesis of Ad-Sq remains unknown. In general, the prognosis of Ad-Sq was reportedly worse than that of Ad and Sq.[2,3]

Epidermal growth factor receptor tyrosine kinase inhibitors (EGFR-TKIs), such as gefitinib and erlotinib, are therapeutic options widely used for the treatment of Ad patients. The responders to EGFR-TKI have somatic mutations in the EGFR tyrosine kinase domain, including deletions in exon19 and a point mutation at codon 858 in exon21 (L858R). For these patients, EGFR-TKI has a high response rate of approximately 70% and results in better prognosis.[4–7]

However, the usefulness of EGFR-TKI treatment is limited to Ad, and few studies have analyzed the effectiveness of EGFR-TKI for Ad-Sq. Until recently, EGFR mutation in Ad-Sq has been mentioned in only a few reports. Sasaki *et al.* reported that 4 (15%) of 26 Ad-Sq patients were positive for EGFR mutations, which is a lower frequency than that for Ad.[8] On the other hand, Kang *et al.* reported that 11 (44%) of 25 Ad-Sq patients were positive for EGFR mutations among a Korean population.[9]

In the present study, we selected 67 Ad-Sq cases and investigated the clinicopathological characteristics of Ad-Sq. Focusing on the EGFR mutation in Ad-Sq, we analyzed 59 Ad-Sq cases with special reference to the correlation between mutation status and clinicopathological factors. Additionally, we performed immunohistochemical staining for the EGFR-mutated cases using two mutation-specific antibodies for a deletion and L858R to detect the distribution of mutated cells in each component.

MATERIALS AND METHODS

Case selection

During the period from September 1992 to May2011, 3953 patients with lung cancer underwent surgical resection at the National Cancer Center Hospital East in Kashiwa, Japan. A total of 67 cases (2.6%) of Ad-Sq were selected. The clinicopathological characteristics of the Ad-Sq cases were examined.

Clinical information

All available clinical information was obtained from the clinical records. The records were reviewed for age, sex, smoking index, tumor markers, pathological stage, and duration of follow up. Two of the 67 patients underwent a wedge resection, 4 underwent a partial resection, 57 underwent a lobectomy, and 4 underwent a pneumonectomy. In addition, we collected clinicopathological data about 2628 Ad and 1258 Sq patients who underwent surgical resection, for the purpose of comparing with Ad-Sq.

Pathological information

In each case, the resected tissue had been fixed in 10% formalin or absolute methyl alcohol and embedded in paraffin. The tumors had been cut into 5–10 mm slices, and serial 4-μm sections had been stained with hematoxylin and eosin. The Alcian blue-periodic acid Schiff (AB-PAS) method was used to visualize cytoplasmic mucin production, and the Victoria van-Gieson (VVG) method was used to visualize elastic fibers. Two pathologists (T.S. and G.I.) reviewed all the slides through a multiheaded microscope. The tumors were classified according to the criteria of the current histological classification adopted by the World Health Organization.[1] The Ad component was characterized as tumor cells showing an acinar, papillary, bronchioalveolar, or solid with mucin growth pattern or that were positive for PAS Alcian Blue staining.[1,10,11] The Sq component was defined as tumor cells with keratinization, pearl formation, and/or intercellular bridges.[1] If each of the components with these features did not account for a minimum of 10% of the entire tumor, the case was excluded from the present study. As a result, we identified 67 (2.6%) cases of pathologically diagnosed Ad-Sq (Fig. 1).

EGFR mutational analysis using a molecular technique

From the FFPE tissue blocks, three slices of 10-μm-thick unstained sections were cut. Tissue areas in which the tumor cells occupied more than 70% of the area were macroscopically dissected and genomic DNA was isolated using the QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany). The *EGFR* exon-19 and 21 fragments were amplified according to a previously described method, with some modifications. Briefly, each 50-μL PCR cocktail contained 50 ng of genomic DNA, 1.5 mM of magnesium chloride, 200 mM of deoxynucleotide triphosphates, 0.2 mM of PCR primers and 2.5 U of HotStarTaqw DNA polymerase (Qiagen). The PCR conditions were as follows: 1 cycle at 95°C for 15 min; 40 cycles at 95°C for 30 s, 65°C for 30 s and 72°C for 1 min; and 1 cycle at 72°C for 10 min. The primer sequences were CCAGATCACTGGGCAGCATGTGGCACC and AGCA GGGTCTAGAGCAGAGCAGCTGCC for exon 19, and

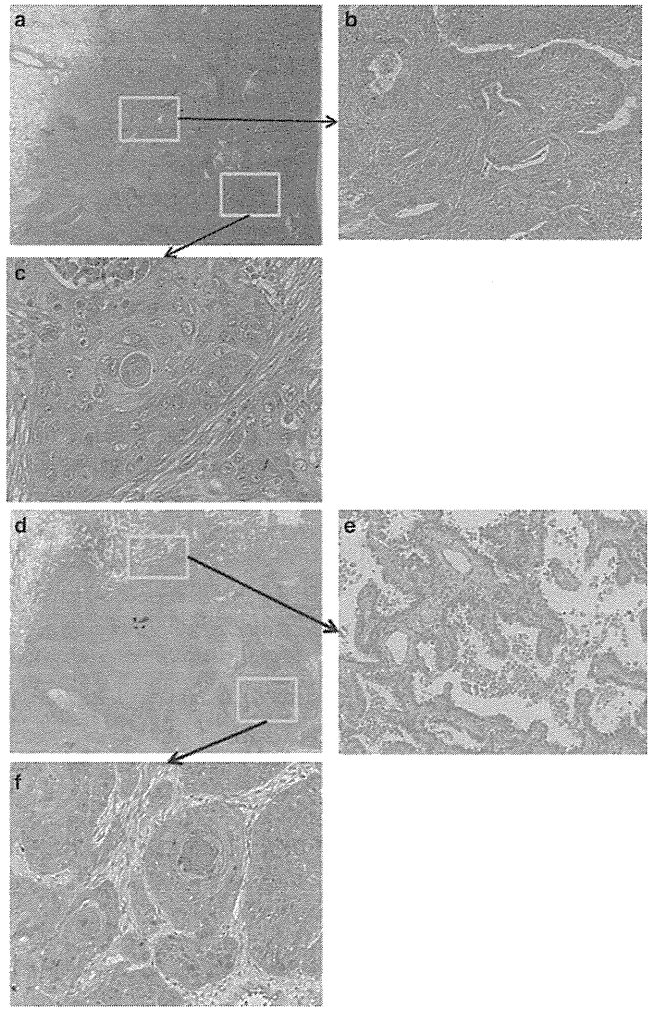

Figure 1 Two typical examples of adenosquamous carcinoma. (a) Low-power field showing tumor sections (H&E). Both adenocarcinoma and squamous cell carcinoma components are visible in the tumor sections. (b) Adenocarcinoma component with acinar pattern (H&E x100). (c) Squamous cell component with keratin pearl formation (H&E). (d) Low-power field showing tumor sections (H&E). (e) Adenocarcinoma component with lepidic growth pattern (H&E, x100). (f) Squamous cell component with keratin pearl formation (H&E).

TCAGAGCCTGGCATGAACATGACCCTG and GGTCCC TGGTGTCAGGAAAATGCTGG for exon 21.

The PCR products were purified using ExoSAP-IT (Affymetrix, Santa Clara, CA, USA), and the amplicon size and amount were confirmed using DNA agarose gel electrophoresis. The purified PCR products were sequenced directly using the same primers as those used for the PCR. The BigDye Terminator v3.1 Cycle Sequencing Kit and 3500 Genetic Analyzer (Life Technologies, Carlsbad, CA, USA) were used according to the manufacturer's instructions. Analyses of the DNA sequences were performed using Sequencing Analysis Software v3.4.1 (Life Technologies).

Immunohistochemistry

After the detection of the EGFR mutational status using direct sequencing, immunohistochemical staining was performed in the EGFR-mutated cases. For the immunohistochemical staining, 5-μm-thick sections were deparaffinized. Positive control was taken along with the immunohistochemical analysis. Antigen retrieval was performed at 95°C for 45 min in Target Retrieval Solution, pH 9.0, x10 (Dako, Glostrup, Denmark). The primary antibodies that we used were a monoclonal antibody against human EGFR with the DEL (E746-A750del) mutation (1:75, clone 6B6; Cell Signaling Technology) and a monoclonal antibody against human EGFR with the L858R mutation (1:75, clone 43B2; Cell Signaling Technology). The antibodies were diluted in Signal Stain (Cell Signaling Technology), and the slides were incubated at 4°C overnight. After three washes in TBS for 5 min, the slides were incubated for 30 min at room temperature with anti-rabbit polymer-HRP secondary antibody (Dako). The immunoreactions were detected using 3,3-diaminobenzidine, followed by counterstaining with hematoxylin.

Statistical analysis

The statistical analysis was performed using SPSS 12.0 for Windows. Chi-square tests were used for categorical variables, and a P value < 0.05 was regarded as statistically significant.

RESULT

Patient characteristics

The clinical and pathological characteristics of Ad-Sq are summarized in Table 1. Among the 67 Ad-Sq patients, 48 were men and 19 were women. The median age at surgery was 70 years (range, 45–84 years). None of them received EGFR-TKI therapy.

Table 1 Clinicopathological characteristics of the patients with adenosquamous carcinoma (Adsq)

	Ad-Sq $n = 67$	Ad $n = 2628$	P value	Sq $n = 1258$	P value
Gender					
Male/female	48/19	1373/1255	0.001	1139/119	<0.001
Age, years	70 (45–84)	66 (22–92)	0.002	68 (31–88)	
<70 years	32	1731		689	0.261
≧70 years	35	897		569	
Smoking History					
Current or former	43	1508	0.266	1215	<0.001
Never	24	1120		43	
CEA					
<5.0 ng/dl	23	1687	<	848	<0.001
≧5.0 ng/dl	44	941	0.001	410	
SCC					
<1.5 ng/ml	40	2111	<0.001	474	0.001
≧1.5 ng/ml	27	517		784	
Pathological stage					
I/II-IV	32/35	1617/1011	0.022	582/676	0.810
p-T					
T1/T2-4	13/54	1421/1207	<0.001	344/914	0.153
p-N					
N0/N1-3	41/26	1737/891	0.403	671/587	0.209
Vascular invasion					
Positive/negative	50/17	1015/1613	<0.001	721/537	0.004
Lymphatic permeation					
Positive/negative	30/37	699/1929	<0.001	383/875	0.001
Pleural involvement					
Positive/negative	30/37	767/1861	0.005	481/777	0.283

When the clinicopathological factors of Ad-Sq were compared with those of Ad and Sq, Ad-Sq was more likely to occur in men and older patients than Ad ($P = 0.001$, 0.002, respectively). On the other hand, Ad-Sq was less likely to occur in men and smokers than Sq ($P < 0.001$). T factor and pleural involvement were also seen in Ad-Sq more frequently than in Ad ($P < 0.001$). However, no statistically significant difference was seen between Ad-Sq and Sq with regard to the T factor and pleural involvement. The Ad-Sq was significantly correlated with vascular invasion ($P < 0.001$ vs. Ad, $P = 0.004$ vs. Sq) and lymphatic permeation ($P < 0.001$ vs. Ad, $P = 0.001$ vs. Sq).

Figure 2 shows the Kaplan–Meier survival curves for the three histological subtypes. For the 59 Ad-Sq patients with follow-up data, the postoperative 3-year survival rate was 58.7%. A statistically significant difference was recognized when compared with Ad (58.7% vs. 78.1%, $P = 0.038$), but not when compared with Sq (58.7% vs. 66.3% $P = 0.103$).

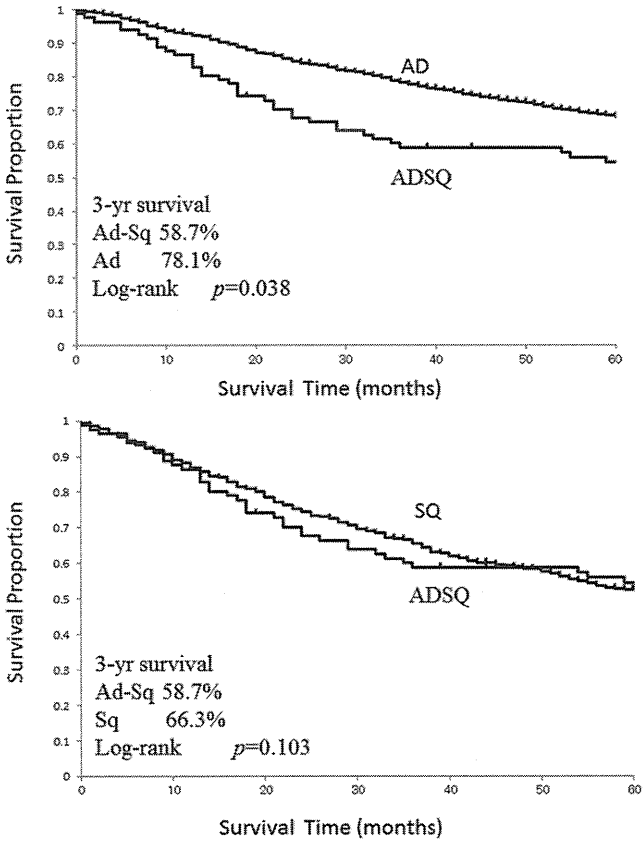

Figure 2 Three-year survival curves after surgery, as calculated using the Kaplan–Meier method. The Ad-Sq group had a significantly poorer survival than the Ad group (log-rank test, $P = 0.038$). However, no significant difference compared with the Sq group was seen (log-rank test, $P = 0.103$).

EGFR mutational status of Ad-Sq

The EGFR mutational status was measured in 59 patients using direct sequencing. A total of 14 patients were positive for EGFR mutations (24%). DEL in exon19 was detected in nine cases. The details were as follows: del E746-A750 was detected in six cases, and the other three cases were del L747-T751, del E746-A753, and del S752-I759, respectively. A point mutation at codon 858 (L858R) in exon21 was detected in three cases, and L861Q, which is associated with an increased sensitivity to EGFR-TKIs, was detected in two cases.[12,13]

Correlation between EGFR mutations and the clinicopathological factors of Ad-Sq

To reveal the clinical and pathological features of Ad-Sq with EGFR mutation, we assessed the relationship between EGFR mutations and the clinicopathological factors (Table 2). Patients who had never been smokers and with lymphatic permeation were more frequently seen in the mutation-positive group than in the mutation-negative group ($P = 0.035$, 0.027 respectively). The 3-year survival rate of the mutation-positive group tended to be better than that of

Table 2 Correlation between the clinicopathological factor and EGFR mutation status

	EGFR-mutation(+) $n = 14$	EGFR-mutation(–) $n = 45$	*P* value
Gender			
Male/female	9/5	35/10	0.311
Age, years	70 (45–832)	70 (45–84)	0.966
<70 years	6	19	
≧70 years	8	26	
Smoking History			
Current or former	6	33	0.035
Never	8	12	
CEA			
<5.0 ng/dl	5	15	0.869
≧5.0 ng/dl	9	30	
SCC			
<1.5 ng/ml	11	25	0.123
≧1.5 ng/ml	3	20	
Pathological stage			
I/II-IV	5/9	23/22	0.314
p-T			
T1/T2-4	2/12	11/34	0.423
p-N			
N0/N1-3	8/6	16/29	0.151
Vascular invasion			
positive/negative	13/1	33/12	0.124
Lymphatic permeation			
Positive/negative	7/7	9/36	0.027
Pleural involvement			
Positive/negative	7/7	19/26	0.609

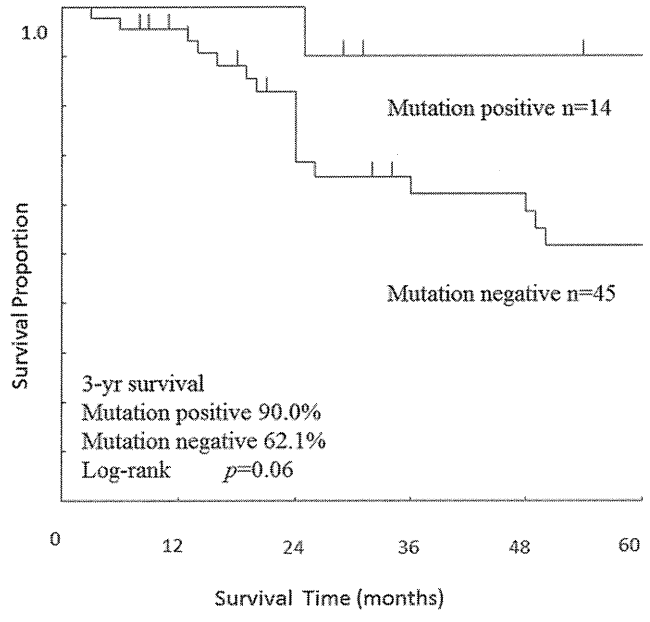

Figure 3 EGFR mutational status effects on survival, as calculated using the Kaplan–Meier method. The EGFR-mutation-positive group tended to have a better survival than the negative one (log-rank test, $P = 0.06$).

Table 3 Correlation between morphological characteristics and EGFR mutation

	EGFR mutation positive $n = 14$	EGFR mutation negative $n = 45$	*P* value
Adenocarcinoma component			
Predominant subtype			
BAC	3	4	0.816
Papillary	4	9	
Acinar	7	29	
Solid	0	3	
Lepidic growth pattern			
Positive/negative	10/4	16/29	0.018
Squamous cell carcinoma component grade of differentiation			
Well/moderate	13	43	0.543
Poorly	1	2	

the mutation-negative group (90.0% vs. 62.8% respectively, $P = 0.06$; Fig. 3).

We compared the pathological findings of both groups (Table 3). No significant difference in the proportions of predominant subtypes in the Ad component was seen between the mutation-positive and negative groups. On the other hand, a lepidic growth pattern was more frequently seen in the mutation-positive group ($P = 0.018$) (Fig. 1d–f). An acinar pattern was the most frequent subtype of the Ad component in both the mutation-positive and negative groups. No statistically significant difference in the grade of differentiation was seen for the Sq component.

Immunohistochemical staining

We performed immunohistochemical staining for EGFR-mutated cases. A DEL-specific antibody was used for the cases that were positive for a DEL in exon19, while an L858R-specific antibody was used for the cases that were positive for L858R. The clinicopathological and immunohistochemical results for the 14 EGFR-mutated cases are summarized in Table 4. Immunoreactivity for the DEL-specific antibody was observed in three of the six del E746-A750 cases and the one del E746-A753 case (identified using direct sequencing). Neither the L747-T751-positive nor the S752-I759-positive case exhibited immunoreactivity for the DEL-specific antibody. Although all the L858R mutated cases (100%) were positive for the L858R-specific antibody, neither of the L861Q-mutated cases exhibited a positive immunoreaction. As shown in Table 4 and Supplemental Table, a positive outcome for the DEL-specific antibody was more likely to occur in women and non-smokers. In contrast, a positive outcome for the L858R-specific antibody was more likely to occur in men and smokers. There were not any other clinicopathological differences every mutation. In all seven cases that were positive for mutation-specific immunostaining, the same immunohistochemical reactions were detected in both the Ad component and the Sq component (Fig. 4a–d).

DISCUSSION

In this study, we analyzed the clinicopathological characteristics of Ad-Sq, focusing on the correlation between EGFR mutations and clinicopathological factors. This is the first study to demonstrate that EGFR-mutated Ad-Sq had clinicopathological features, similar to those of EGFR-mutated adenocarcinoma, specifically to never having been a smoker, better prognosis, and pathologically lepidic growth pattern. It was suggested that Ad-Sq of the lung could be subclassified into distinct subtypes with different pathogenetic mechanisms according to the EGFR mutation status.

Until recently, a few studies have analyzed EGFR mutations in Ad-Sq.[8,9,14–17] However, we analyzed 67 cases of Ad-Sq, the largest number of cases among similar studies published to date. Moreover, we performed a clinicopathological analysis using not only a molecular method, but also immunohistochemistry with two mutation-specific antibodies.

In our study, the frequency of the EGFR mutation in Ad-Sq was 24% (14/59). In the previous studies, Sasaki *et al.* reported that 4 of 26 (15%) Ad-Sq were positive for an EGFR mutation, and Toyooka *et al.* reported that 3 of 11 (27%) cases were positive, similar to the results of the present study.[8,15] On the other hand, a discrepancy was observed with regard to the frequency of EGFR mutation when

Table 4 EGFR mutation positive case ($n = 14$)

	Age	Sex	Smoking	p-stage	Mutation	IHC-stain Ad	IHC -stain Sq
1	84	F	–	IB	Del E746-A750	–	–
2	75	M	–	IIB	Del L747-T751	–	–
3	58	M	–	IB	Del E746-A750	+	+
4	71	F	–	IB	Del E746-A750	–	–
5	66	F	–	IIIA	Del E746-A753	+	+
6	45	M	+	IIB	Del E746-A750	–	–
7	57	M	–	IA	Del E746-A750	+	+
8	73	M	+	IIA	Del S752-I759	–	–
9	77	F	–	IIB	Del E746-A750	+	+
10	70	F	–	IIIA	L858R	+	+
11	60	M	+	IIA	L858R	+	+
12	66	M	+	IIA	L858R	+	+
13	70	M	+	IIA	L861Q	–	–
14	70	M	+	IA	L861Q	–	–

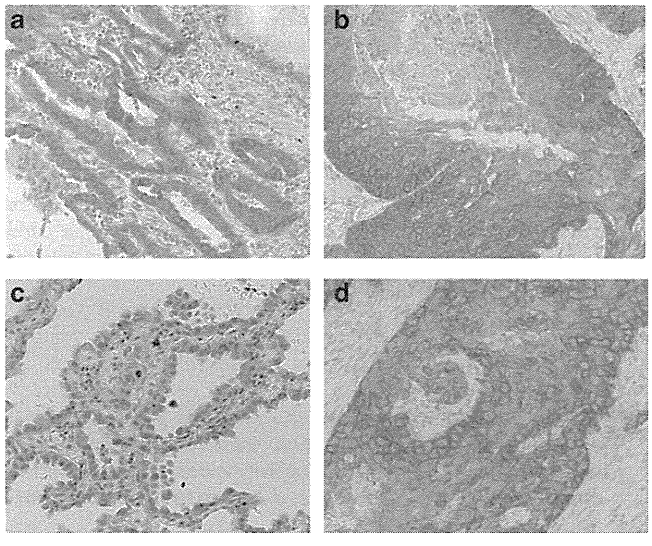

Figure 4 Immunohistochemical staining for EGFR-mutation-specific antibody. (a, b) Both the adenocarcinoma component with an (a) acinar pattern and (b) the squamous cell carcinoma component were positive for the L858R-specific antibody. (c,d) The adenocarcinoma component with (c) a lepidic growth pattern and (d) the squamous cell carcinoma component were positive for the DEL-specific antibody.

compared with results obtained in neighboring countries. Kang *et al.* reported that 44% (11/25) of Ad-Sq were positive for an EGFR mutation in a Korean population, and Xiao *et al.* reported a frequency of approximately 38% (21/55) in a Chinese population.[9,17] This discrepancy might be attributable to differences in the detection methods used for the mutational analysis; alternatively, genetic differences between Japanese populations and others may also affect the frequency of mutation.

The sensitivity of the DEL-specific antibody used in our study was inferior to that of the L858R-specific antibody, as previously reported.[18–22] Kozu *et al.* reported that the sensitivities of DEL-specific and L858R-specific antibodies were 42.2% and 75.6%, respectively.[18] Immunohistochemical staining revealed that the EGFR mutation occurred not only in the Ad component, but also in the Sq component. Staining intensity in both components was similar, as shown in Fig. 4 and Table 4. Several studies using microdissection analyses have demonstrated identical EGFR mutations in both components, but few immunohistochemical analyses have been performed. Recently, Miyamae *et al.* demonstrated that identical EGFR mutations were observed in the Ad and Sq components using a DEL-specific antibody, suggesting that Ad-Sq to be generated from monoclonal origin.[23] Rekhtman *et al.* described that squamous cell carcinoma with EGFR mutation had a component of adenocarcinoma using immunohistochemistry.[24] These findings may suggest that squamous cell carcinoma componemt was originated from adenocarcinoma somponent in Ad-Sq. In contrast, Kanazawa *et al.* described the monoclonal transition from squamous cell carcinoma to adenocarcinoma.[25] So, it is difficult at this time to conclude the origin of Ad-Sq.

Interestingly, a lepidic growth pattern was more frequent in the EGFR-mutated group than in the negative group. Blons

et al. reported that EGFR mutations in Ad were more prevalent among cases with a lepidic growth pattern.[26] In this study, we detected the same correlation between the EGFR mutation status and the lepidic growth pattern in Ad-Sq as well as Ad. Sasaki *et al.* reported that the EGFR mutational status of Ad-Sq was significantly correlated with the smoking history and gender.[8] Kang *et al.* also reported that EGFR mutations in Ad-Sq were more frequent among people who had never been smokers and women.[9] The current study revealed a similar correlation with smoking, but we did not find a significant correlation with sex.

A Kaplan–Meier survival curve showed that the survival of the EGFR-mutated cases was better when compared with the negative cases, even though a statistically significant difference was absent. Although a similar tendency was observed in Ad cases,[27,28] no previous studies have compared the survival of EGFR-positive Ad-Sq cases with those of EGFR-negative cases. In the future, large-scale, multi-institutional studies will be needed to verify the results.

In this study, we found several common features between EGFR-mutated Ad-Sq cases and those with Ad. Therefore, we speculated that a subtype with clinicopathological features resembling those of Ad existed among the Ad-Sq tumors with an EGFR mutation. Based on our analysis, Ad-Sq developing in a patient who had never been a smoker and pathologically exhibiting a lepidic growth pattern in the Ad component was likely to be positive for an EGFR mutation. EGFR-TKIs, such as gefitinib and erlotinib, may be effective therapeutic options for patients with EGFR-mutated Ad-Sq as well as those with EGFR-mutated Ad. Mitsudomi *et al.* have reported excellent clinical response to gefitinib in a patient with EGFR mutated Ad-Sq.[29] Shukuya *et al.* reported that efficacy of EGFR-TKIs for non-adenocarcinoma with EGFR mutation was different from that of pure adenocarcinoma.[30] So, we consider that further large-scale clinical trials of EGFR-TKI for adenosquamous carcinoma patients are needed, as is validation of an optimal therapeutic strategy.

As shown in Fig. 4, some cancer cells showed positive reaction for mutation specific antibodies. Although biological significance of this nuclear staining pattern is not clarified, further examination will be required.

In conclusion, we demonstrated that EGFR-mutated Ad-Sq exhibited characteristics similar to those of EGFR-mutated Ad in that they were more likely to occur in patients who had never been smokers and more likely to exhibit a lepidic growth component pathologically. Moreover, patients with EGFR-mutated Ad-Sq tend to have a better outcome, compared with the EGFR-mutation-negative group. So we found the correlation between EGFR mutation and the prognosis of adenosquamous carcinoma of lungs, similar to that of adenocarcinomas. In particular, Ad-Sq cases that carry an EGFR mutation might respond to EGFR-TKI therapy.

REFERENCES

1 Travis WD, Branbilla E, Muller-Hermelink HK. *World Health Organization Classification of Tumours, Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart.* Lyon: IARC, 2004.
2 Filosso PL, Ruffini F, Asioli S *et al.* Adenosquamous lung carcinomas: A histologic subtype with poor prognosis. *Lung Cancer* 2011; **74**: 25–9.
3 Cooke DT, Nguyen DV, Yang Y *et al.* Survival comparison of adenosquamous, squamous cell, and adenocarcinoma of the lung after lobectomy. *Ann Thorac Surg* 2010; **90**: 943–8.
4 Paez JG, Jänne PA, Lee JC *et al.* EGFR mutations in lung cancer: Correlation with clinical response to gefitinib therapy. *Science* 2004; **304**: 1497–500.
5 Lynch TJ, Bell DW, Sordella R *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004; **350**: 2129–39.
6 Pao W, Miller VA. Epidermal growth factor receptor mutations, smallmolecule kinase inhibitors, and non-small-cell lung cancer: Current knowledge and future directions. *J Clin Oncol* 2005; **16**: 2256–68.
7 Mitsudomi T, Kosaka T, Yatabe Y *et al.* Biological and clinical implications of EGFR mutations in lung cancer. *Int J Clin Oncol* 2006; **11**: 190–8.
8 Sasaki H, Endo K, Fujii Y *et al.* Mutation of epidermal growth factor receptor gene in adenosquamous carcinoma of the lung. *Lung Cancer* 2007; **55**: 129–30.
9 Kang SM, Kang HJ, Shin JH *et al.* Identical epidermal growth factor receptor mutations in adenocarcinomatous and squamous cell carcinomatous components of adenosquamous carcinoma of the lung. *Cancer* 2007; **109**: 581–7.
10 Sobin L, Wittekind C. *TNM Classification of Malignant Tumors.* New York: Wiley and Sons, 2002.
11 Tsuta K, Ishii G, Nitadori J *et al.* Comparison of the immunophenotypes of signet-ring cell carcinoma, solid adenocarcinoma with mucin production, and mucinous bronchioloalveolar carcinoma of the lung characterized by the presence of cytoplasmic mucin. *J Pathol* 2006; **209**: 78–87.
12 Shigematsu H, Lin L, Gazdar AF *et al.* Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst* 2005; **97**: 339–46.
13 Kancha RK, Peschel C, Duyster J *et al.* The epidermal growth factor receptor-L861Q mutation increases kinase activity without leading to enhanced sensitivity toward epidermal growth factor receptor kinase inhibitors. *J Thorac Oncol* 2011; **6**: 387–92.
14 Ohtsuka K, Ohnishi H, Fujiwara M *et al.* Abnormalities of epidermal growth factor receptor in lung squamous-cell carcinomas, adenosquamous carcinomas, and large-cell carcinomas: Tyrosine kinase domain mutations are not rare in tumors with an adenocarcinoma component. *Cancer* 2007; **109**: 741–50.
15 Toyooka S, Yatabe Y, Tomii K *et al.* Mutations of epidermal growth factor receptor and K-ras genes in adenosquamous carcinoma of the lung. *Int J Cancer* 2006; **118**: 1588–90.
16 Tochigi N, Dacic S, Nikiforova M *et al.* Adenosquamous carcinoma of the lung: A microdissection study of KRAS and EGFR mutational and amplification status in a western patient population. *Am J Clin Pathol* 2011; **135**: 783–9.
17 Jia XL, Chen G *et al.* EGFR and KRAS mutations in Chinese patients with adenosquamous carcinoma of the lung. *Lung Cancer* 2011 May 17(Epub) **74**: 396–400.

18 Kozu Y, Tsuta K, Kohno T *et al.* The usefulness of mutation-specific antibodies in detecting epidermal growth factor receptor mutations and in predicting response to tyrosine kinase inhibitor therapy in lung adenocarcinoma. *Lung Cancer* 2011; **73**: 45–50.
19 Yu J, Kane S, Wu J *et al.* Mutation-specific antibodies for the detection of EGFR mutations in non-small-cell lung cancer. *Clin Cancer Res* 2009; **15**: 3023–8.
20 Li D, Chen S, Hu C *et al.* Immunohistochemistry with mutation specific antibodies detecting the status of EGFR mutations in non-small-cell lung cancer. *Mod Pathol* 2010; **23**: 407A.
21 Kato Y, Peled N, Wynes MW *et al.* Novel epidermal growth factor receptor mutation-specific antibodies for non-small cell lung cancer: Immunohistochemistry as a possible screening method for epidermal growth factor receptor mutations. *J Thorac Oncol* 2010; **5**: 1551–8.
22 Kitamura A, Hosoda W, Yatabe Y *et al.* Immunohistochemical detection of EGFR mutation using mutation-specific antibodies in lung cancer. *Clin Cancer Res* 2010; **16**: 3349–55.
23 Miyamae Y, Shimizu K, Hirato J *et al.* Significance of epidermal growth factor receptor gene mutations in squamous cell lung carcinoma. *Oncol Rep* 2011; **25**: 921–8.
24 Rekhtman N, Paul K, Marc L *et al.* Clarifying the spectrum of driver oncogene mutations in biomarker-verified squamous cell carcinoma of lung: Lack of EGFR/KRAS and preence of PIK3CA/ALT1 mutaions. *Clin Cancer Res* 2012; **18**: 1167–76.
25 Kanazawa H, Ebina M, Nukiwa T *et al.* Transition frm squamous cell carcinoma to adenocarcinoma in adenosquamous carcinoma of the lung. *Am J Pathol* 2000; **156**: 1289–98.
26 Blons H, Cote JF, Le CD *et al.* Epidermal growth factor receptor mutation in lung cancer are linked to bronchioloalveolar differentiation. *Am J Surg Pathol* 2006; **30**: 1309–15.
27 Fukuoka M, Wu YL, Mok TS *et al.* Biomarker analyses and final overall survival results from a phase III, randomized, open-label, first-line study of gefitinib versus carboplatin/paclitaxel in clinically selected patients with advanced non-small-cell lung cancer in Asia (IPASS). *J Clin Oncol* 2011; **29**: 2866–74.
28 Sonobe M, Kobayashi M, Date H *et al.* Impact of KRAS and EGFR Gene Mutations on Recurrence and Survival in Patients with Surgically Resected Lung Adenocarcinomas. *Ann Surg Oncol* 2011; May 24; **19**: 347–353.
29 Mitsudomi T, Kosaka T, Yatabe Y *et al.* Mutation of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence. *J Clin Oncol* 2005; **23**: 2513–20.
30 Shukuya T, Takahashi T, Yamamoto N *et al.* Efficacy of gefitinib for non-adenocarcinoma non-small cell lung cancer patients harboring epidermal growth factor receptor mutations: A pooled analysis of previous reports. *Cancer Sci* 2011; **102**: 1032–7.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. Clinicopathological characteristics of the patients with EGFR mutated adenosquamous carcinoma ($n = 14$)

Cancer Science
The official journal of the Japanese Cancer Association

Long-term results of concurrent chemoradiotherapy using cisplatin and vinorelbine for stage III non-small-cell lung cancer

Hidehito Horinouchi,[1,6] Ikuo Sekine,[1] Minako Sumi,[2] Kazumasa Noda,[3] Koichi Goto,[4] Kiyoshi Mori[5] and Tomohide Tamura[1]

[1]Division of Internal Medicine and Thoracic Oncology, [2]Division of Radiation Oncology, National Cancer Center Hospital, Tokyo; [3]Department of Thoracic Oncology, Kanagawa Cancer Center, Yokohama; [4]Division of Thoracic Oncology, National Cancer Center Hospital East, Kashiwa; [5]Department of Medical Oncology, Division of Thoracic Oncology, Tochigi Cancer Center, Utsunomiya, Japan

(Received March 20, 2012/Revised September 9, 2012/Accepted September 13, 2012/Accepted manuscript online September 24, 2012/Article first published online November 8, 2012)

Concurrent chemoradiotherapy is the standard treatment for unresectable stage III non-small cell lung cancer (NSCLC). The long-term feasibility and efficacy of vinorelbine and cisplatin with concurrent thoracic radiotherapy were investigated. Eighteen patients received cisplatin (80 mg/m^2) on day 1 and vinorelbine (20 mg/m^2 in level 1, and 25 mg/m^2 in level 2) on days 1 and 8 every 4 weeks for four cycles in a phase I trial. Ninety-three patients received the same chemotherapy regimen except for the fixed vinorelbine (20 mg/m^2) dosage and consolidation therapy with docetaxel (60 mg/m^2, every 3 weeks). The thoracic radiotherapy consisted of a single dose of 2 Gy once daily to a total dose of 60 Gy. A total of 111 patients were analyzed in the present study: male/female, 91/20; median age, 60 years; stage IIIA/IIIB, 50/61; and squamous/non-squamous histology, 26/85. The 3-, 5-, and 7-year overall survival rates (95% CI) were 43.2% (33.9–52.2), 25.2% (17.6–33.5), and 23.2% (15.8–31.4), respectively. The median progression-free survival and median survival time (95% CI) were 13.5 (10.1–16.7) months and 30.0 (24.3–38.8) months, respectively. Four patients (4%) experienced Grade 5 pulmonary toxicities from 4.4 to 9.4 months after the start of treatment. In conclusion, approximately 15% of patients with unresectable stage III NSCLC could be cured with chemoradiotherapy without severe late toxicities after 10 months of follow-up. Although based on the data from highly selected population participated in phase I and phase II trial, this analysis would strengthen and confirm the previous reports concerning concurrent chemoradiotherapy with third generation cytotoxic agents. (*Cancer Sci* 2013; 104: 93–97)

Stage III locally advanced non-small cell lung cancer (NSCLC) accounts for 25–30% of all lung cancer cases.[1,2] Because of the equal frequency of local and distant recurrences, the combination of systemic chemotherapy and thoracic radiotherapy has been established as a standard of care for patients with stage III NSCLC.[3] Concurrent chemoradiotherapy is superior to a sequential approach, as shown by phase III trials in stage III NSCLC.[4,5]

Ohe *et al.*[6] reported the long-term follow-up analysis of concurrent chemoradiotherapy with former generation chemotherapy agents (median survival time 16.1 months, and 7-year overall survival rate 12.0%). Few researchers, however, have reported follow-up data of longer than 5 years after concurrent chemoradiotherapy with third-generation chemotherapy. The long-term safety and efficacy of vinorelbine and cisplatin with concurrent thoracic radiotherapy were investigated.

Materials and Methods

Study selection. Two previous studies were included in this analysis. One was a phase I study of concurrent thoracic radiotherapy with cisplatin plus vinorelbine, and the other evaluated docetaxel consolidation therapy following concurrent chemoradiotherapy.[7,8] These studies were approved by the institutional review board at each institution. Written, informed consent was obtained from all participating patients.

Patient selection. The two studies had similar eligibility criteria. They were: histologically or cytologically proven NSCLC; unresectable stage IIIA or IIIB disease; no previous treatment; measurable disease; tumor within an estimated irradiation field no larger than half the hemithorax; age between 20 years and 74 years; Eastern Cooperative Oncology Group (ECOG) performance status 0 or 1; and adequate organ function, including bone marrow, liver, kidney, and lung. Patients were diagnosed to have unresectable disease based on a consensus of thoracic oncologists including surgeons in each institution. The exclusion criteria were reported in previous papers.[7,8]

Treatment schedule. In the phase I study, treatment consisted of chemotherapy with four cycles of cisplatin and vinorelbine (20 mg/m^2 in level 1, and 25 mg/m^2 in level 2) and concurrent thoracic radiotherapy (see below). In the other study, treatment consisted of a chemoradiotherapy portion with three cycles of cisplatin and vinorelbine followed by a consolidation portion with three cycles of docetaxel. Cisplatin (80 mg/m^2) was administered every 4 weeks by intravenous infusion for 60 min with 2500–3000 mL of fluid for hydration. Vinorelbine 20 mg/m^2 diluted in 50 mL of normal saline was administered intravenously on days 1 and 8 every 4 weeks. All patients received prophylactic antiemetic therapy consisting of a $5HT_3$-antagonist and a steroid. In the docetaxel (60 mg/m^2, every 3 weeks) consolidation trial, consolidation therapy was started sequentially in patients whose general condition was acceptable. Follow-up computed tomographies after chemoradiotherapy were scheduled as follows; every 2–4 months during the 1 year, every 6 months in the 2 and 3 years, and every 1 year thereafter.

Thoracic radiotherapy was delivered with megavoltage equipment ($\geq$6 MV) using anterior/posterior opposed fields up to 40 Gy in 20 fractions, including the primary tumor, the metastatic lymph nodes, and the regional nodes. A booster dose of 20 Gy in 10 fractions was given to the primary tumor

[6]To whom correspondence should be addressed.
E-mail: hhorinou@ncc.go.jp

doi: 10.1111/cas.12028

and the metastatic lymph nodes for a total dose of 60 Gy using bilateral oblique fields. Computed tomography (CT) scan-based treatment planning was used in all patients. The clinical target volume (CTV) for the primary tumor was defined as the gross tumor volume (GTV) plus 1 cm taking into account subclinical extension. CTV and GTV for the metastatic nodes (>1 cm in the shortest dimension) were the same. Regional nodes, excluding the contralateral hilar and supraclavicular nodes, were included in the CTV, but the lower mediastinal nodes were included only if the primary tumor was located in the lower lobe of the lung. The planning target volumes for the primary tumor, the metastatic lymph nodes, and regional nodes were determined as CTVs plus 0.5–1.0-cm margins laterally and 1.0–2.0-cm margins craniocaudally, taking into account setup variations and internal organ motion. Lung heterogeneity corrections were not used.

Toxicity assessment. Toxicities were graded according to the National Cancer Institute (NCI) Common Toxicity Criteria version 2.0 issued in 1998, and late toxicities associated with thoracic radiotherapy were graded according to the Radiation Therapy Oncology Group/European Organization for Research and Treatment of Cancer late radiation morbidity scoring scheme.[9] Late toxicities were defined as those that occurred or persisted 90 days after completion of radiotherapy. The detailed methods of treatment modification due to toxicity were reported in previous papers.[7,8]

Response evaluation. In the phase I trial, the objective tumor response was evaluated according to the World Health Organization (WHO) criteria issued in 1979.[10] The Response Evaluation Criteria in Solid Tumors were used to evaluate objective tumor response in the docetaxel consolidation trial.[11] Local recurrences were defined as tumor progression in the primary site and in the hilar, mediastinal, and supraclavicular lymph nodes after a partial or complete response; regional recurrence was defined as the development of malignant pleural and pericardial effusions; and distant recurrence was defined as the appearance of distant metastases.

Statistical analyses. Progression-free and overall survival times were estimated by the Kaplan–Meier method, and confidence intervals (CIs) were based on Greenwood's formula.[12] Progression-free survival time was measured from the date of registration to the date of disease progression, death (from any cause), or the last follow-up. Overall survival time was measured from the date of registration to the date of death (from any cause) or to the last follow-up. Patients who were lost to follow-up without an event were censored at the date of their last known follow-up. A CI for response rate (RR) was calculated using methods for exact binomial CIs. To investigate the association between survival and factors related to patient characteristics, the Cox regression model was used. Potential factors investigated were as follows: age (in 10-year increments), sex, body weight loss ($\leq 5.0\%$ *vs* $\geq 5.1\%$), histology (squamous cell carcinoma versus non-squamous cell carcinoma), T factor (T1/2 *vs* T3/4), N factor (N0-2 *vs* N 3), and stage (IIIA *vs* IIIB). The STATA 10 for Windows software package (StataCorp LP, College Station, TX, USA) was used for statistical analyses.

Results

Characteristics of the patients. From October 1999 to June 2003, 13 patients were registered at dose level 1 and five at dose level 2 of the phase I study, and 93 patients were enrolled in the docetaxel consolidation trial. Thus, a total of 111 patients were analyzed in the present study. The participants' characteristics were as follows (Table 1): male/female 91/20; median age (range) 60 (31–74) years; body weight loss $\leq 5.0\%/\geq 5.1\%$ 96/14; stage IIIA/IIIB 50/61; and squamous/non-squamous histology 26/85.

Table 1. Patients' characteristics

	Clinical trial		
	Phase I trial†	DTX consolidation‡	Total
Number of patients	18	93	111
Age (years)			
Median	58.5	60	60
Range	48–69	31–74	31–74
Sex			
Male	15	76	91
Female	3	17	20
Performance status			
0	4	32	36
1	14	51	65
Unknown	0	10	10
Body weight loss (minus, %)			
0	11	72	83
0.1-5.0	4	9	13
5.1–	3	11	14
Unknown	0	1	1
Clinical stage			
IIIA	9	41	50
IIIB	9	52	61
N factor			
N0	0	6	6
N1	0	3	3
N2	11	58	69
N3	7	26	33
T factor			
T1	1	18	19
T2	6	31	37
T3	7	13	20
T4	4	30	34
Unknown	0	1	1
Histology			
Adenocarcinoma	14	57	71
Squamous cell carcinoma	3	23	26
Adenosquamous	1	0	1
Large cell carcinoma	0	6	6
NOS§	0	6	6
Others	0	1	1

†The phase I study of concurrent thoracic radiotherapy with cisplatin plus vinorelbine. ‡The docetaxel consolidation therapy following concurrent chemoradiotherapy study. §Non-small cell lung cancer not otherwise specified.

Treatment delivery. Full cycles (four in the phase I trial, three in the docetaxel consolidation trial) of cisplatin and vinorelbine and the full dose (60 Gy) of thoracic radiotherapy were administered in 94 (85%) and 102 (92%) patients, respectively. The delay in radiotherapy was less than 5 days in 74 (67%) patients. In the docetaxel consolidation trial, 59 (63%) patients could enter the consolidation phase, and only 34 (37%) patients completed three cycles of docetaxel chemotherapy, mainly because of toxicities. Of 91 patients with relapses, 27 (30%) received gefitinib as salvage treatments.

Objective tumor response and survival. The objective response rate was 82.0% (95% CI, 74.5–89.1). The 3-, 5-, and 7-year progression-free and overall survival rates (95% CI) were 21.0% (13.9–29.1), 15.7% (9.5–23.4), 14.4% (8.4–22.0), and 43.2% (33.9–52.2), 25.2% (17.6–33.5), and 23.1% (15.7–31.4), respectively (Fig. 1). The median progression-free survival and median survival time (95% CI) were 13.4 (9.8–16.4)

doi: 10.1111/cas.12028