


ORIGINAL ARTICLE

Carcinogen-induced tumors in *SFN*-transgenic mice harbor a characteristic mutation spectrum of human lung adenocarcinoma

Yunjung Kim¹ | Aya Shiba-Ishii¹  | Karina Ramirez² | Masafumi Muratani³ |
Noriaki Sakamoto¹ | Tatsuo Iijima⁴ | Masayuki Noguchi¹

¹Department of Pathology, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan

²Ph.D. Program in Human Biology, School of Integrative and Global Majors, University of Tsukuba, Tsukuba, Japan

³Department of Genome Biology, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan

⁴Department of Pathology, Ibaraki Prefectural Central Hospital, Kasama, Japan

Correspondence

Aya Shiba-Ishii, Department of Pathology, Faculty of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575, Japan.

Email: aya_shiba@md.tsukuba.ac.jp

Abstract

The landscape of genetic alterations in disease models such as transgenic mice or mice with carcinogen-induced tumors has provided a huge amount of information that has shed light on the process of tumorigenesis in human non-small-cell lung cancer (NSCLC). We have previously identified stratifin (*SFN*) as a potent oncogene, and generated *SFN*-transgenic (Tg-SPC-*SFN*^{+/-}) mice, which express human *SFN* (h*SFN*) only in the lung. Here, we have found that carcinogen nicotine-derived nitrosaminoketone (NNK)-induced tumors developing in Tg-SPC-*SFN*^{+/-} mice show a similar histology to human lung adenocarcinoma and exhibit high h*SFN* expression. In order to compare the genetic characteristics of Tg-SPC-*SFN*^{+/-} tumors and human lung adenocarcinoma, the former were subjected to whole-exome sequencing. Interestingly, Tg-SPC-*SFN*^{+/-} tumors showed the distinct distribution of exonic mutations and high number of mutated genes and transversion. Moreover, Tg-SPC-*SFN*^{+/-} tumors showed 73 genes that were commonly detected in more than 2 tumors, mutations of which were also found in human lung adenocarcinoma. The expression levels of some of these genes were significantly associated with the clinical outcome of lung adenocarcinoma patients. Additionally, mutated genes in Tg-SPC-*SFN*^{+/-} tumors were closely associated with key canonical pathways such as PI3K/AKT signaling and apoptosis signaling. These results suggest that *SFN* overexpression is a universal abnormality in human lung adenocarcinogenesis and Tg-SPC-*SFN*^{+/-} tumors recapitulate key features of major human lung adenocarcinoma. Therefore, Tg-SPC-*SFN*^{+/-} mice provide a useful model for clarifying the molecular mechanism underlying lung adenocarcinogenesis.

KEYWORDS

lung adenocarcinoma, mutational signature, nicotine-derived nitrosaminoketone, stratifin, transgenic mice

1 | INTRODUCTION

Lung cancer is the most frequent cause of cancer-related mortality worldwide^{1,2} due to its unsatisfactory cure rate and poor prognosis, and is divided into 2 major histological subtypes: small-cell lung cancer and non-small-cell lung cancer (NSCLC).³ Lung adenocarcinoma, the predominant type of NSCLC, is considered to show a multistep process of carcinogenesis. Atypical adenomatous hyperplasia progresses to adenocarcinoma in situ, and subsequently to invasive adenocarcinoma with specific molecular alterations that are responsible for both the initiation and maintenance of tumorigenesis.⁴

Previously, we examined the gene expression profiles of early stage lung adenocarcinomas⁵ and identified stratifin (SFN, 14-3-3 sigma), which showed significantly higher expression in small but invasive lung adenocarcinoma (tumor size less than 3 cm in diameter) than in adenocarcinoma in situ or normal lung.^{5,6} Overexpression of SFN is due to DNA hypomethylation in its promoter region^{7,8} and is associated with poor patient outcome in lung adenocarcinoma.⁹ We have found that suppression of SFN reduces cell proliferation in vitro^{9,10} and tumor formation or metastasis in vivo,¹¹ suggesting that SFN facilitates tumor progression. Consistently, SFN-transgenic mice (Tg-SPC-SFN^{+/-}), which express human SFN (hSFN) only in the lung under control of the SPC promoter, a lung-specific enhancer, develop lung tumors spontaneously,¹¹ implying that SFN has oncogenic potential in the context of lung adenocarcinogenesis. Additionally, overexpression of SFN dysregulates signal transduction associated with cellular proliferation, the cell cycle, apoptosis, and protein trafficking by interacting with phosphorylated serine or threonine in the binding partner proteins.^{9,10,12}

Genetic imbalances of DNA or mRNA sequences in the cancer genome trigger various genetic alterations during the process of tumorigenesis. Unlike the situation in early stage lung adenocarcinoma, many studies have revealed comprehensive somatic mutations in advanced lung adenocarcinomas using whole-exome sequencing (WES),¹³ which has high accuracy, high sensitivity, and cost effectiveness for identifying driver mutations^{14,15} and can predict the efficacy of anticancer agents such as molecular targeting drugs and immune checkpoint inhibitors.¹⁶ Transgenic mice or mice with carcinogen-induced tumors can be used to assess potential molecular alterations related to human tumorigenesis. Transgenic mouse models of oncogenetic activation show frequent aneuploidy and copy number alteration, but few somatic single-nucleotide variants.¹⁷ However, carcinogens such as methyl-nitrosourea and nicotine-derived nitrosaminoketone (NNK) included in tobacco smoke initiate tumorigenesis through induction of numerous deleterious mutations at the loci of oncogenes such as *KRAS*^{17,18} or tumor suppressor genes such as *TP53*.^{19,20} In this regard, we have been able to recapitulate the early stage of human lung adenocarcinoma using Tg-SPC-SFN^{+/-} mice treated with NNK to trigger tumorigenesis. This has clarified the genetic profile of lung tumors developing in Tg-SPC-SFN^{+/-} mice, thus helping to identify the gene alterations contributing to human lung adenocarcinogenesis.

In the present study, we investigated the genetic landscape of NNK-induced lung adenocarcinomas in a Tg-SPC-SFN^{+/-} background. Paired samples of tumor tissue and normal lung were subjected to WES, and mutated genes associated with SFN overexpression in lung adenocarcinoma were examined. Here, we emphasize the importance of SFN overexpression in lung adenocarcinogenesis and show that Tg-SPC-SFN^{+/-} mice provide a valuable model for clarifying early genetic alterations in human lung adenocarcinoma.

2 | MATERIALS AND METHODS

2.1 | SFN transgenic mouse model and pathological analysis

We generated the Tg-SPC-SFN^{+/-} strain from ICR mice that had been purchased from (Charles River Laboratories, Tsukuba, Japan), as described previously.¹¹ Wild-type (WT) and Tg-SPC-SFN^{+/-} mice were given a single i.p. dose of 4 mg NNK (Toronto Research Chemicals, North York, ON, Canada) dissolved in saline to induce pulmonary tumors. Twenty weeks after NNK treatment, all of the mice were killed, and the induced tumors were resected from the lungs. The resected tumors were partially fixed with formalin for histological analysis including hematoxylin and eosin (H&E) staining and immunohistochemistry for hSFN, as described previously,¹¹ and also frozen in liquid nitrogen for WES and stored at -80°C until DNA extraction. All animal experiments were approved by the Institutional Animal Experiment Committee of the University of Tsukuba (Tsukuba, Japan).

2.2 | Library preparation and WES

The WES libraries were prepared with the NEBNext DNA Library Prep Master Mix Set (New England Biolabs, Ipswich, MA, USA) and a SeqCap EZ Developer Library kit (NimbleGen; Roche, Pleasanton, CA, USA) from 0.5 µg of genomic DNA. The quality of the DNA library was evaluated using Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Sequencing of the multiplexed library using a NextSeq 500 sequencer (Illumina, San Diego, CA, USA) was carried out and the data were imported to CLC Genomics Workbench software (version 10.1.1; Qiagen, Hilden, Germany). Sequences were mapped to the GRCm38/mm10 mouse reference and annotated using Ensemble, Hanava, or Ensemble and Hanava merged transcript database. (<http://www.ensembl.org/index.html>)

Possible single-nucleotide polymorphisms or possible ICR drift mutations were excluded from the data as follows: (a) single base substitutions common to each normal lung sample were removed from datasets of tumor samples, (b) single base substitutions found in normal lung samples from WT or Tg-SPC-SFN^{+/-} mice treated with saline were also removed from each dataset. In addition, we undertook filtering based on the following criteria: single nucleotide variant type with nonsynonymous mutation; coverage of reads 10% or higher; frequency 10% or higher; variant at least 5 bases away from

the end of the read; and an average variant quality score of 20 or more.

2.3 | Bioinformatics analysis

Mutations were first categorized into 6 types according to the sequence on the coding strand (C:G>A:T, C:G>G:C, C:G>T:A, T:A>A:T, T:A>C:G, or T:A>G:C). Here, we used pyrimidines to denote each group such as C>A, C>G, C>T, T>A, T>C, and T>G, and counted the number of mutations detected in tumor samples from WT or Tg-SPC-SFN^{+/-} mice. They were then further subdivided into 16 possible pairs of 5' and 3' nucleotide neighbors. The trinucleotide bases were verified using the University of California Santa Cruz Cancer Genomics Browser (<https://genome.ucsc.edu/>). We compared mutational signatures among WT, Tg-SPC-SFN^{+/-}, and also human cancers, which we obtained from the Catalogue of Somatic Mutations in Cancer (COSMIC) database (<https://cancer.sanger.ac.uk/cosmic>).

To compare mutated gene lists among 4 tumors that had developed in Tg-SPC-SFN^{+/-} mice, 4-way Venn diagrams were produced using the online Venn diagram plotting tool (<http://bioinfo.cnb.csic.es/tools/venny/>), Venny 2.1.0. To identify significantly mutated genes and canonical pathways induced by overexpression of hSFN, we used Ingenuity Pathway Analysis (IPA; Qiagen) or Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis through DAVID Bioinformatics Resources 6.8. (<https://david.ncifcrf.gov/>)

2.4 | Expression analysis of mutated genes in human lung adenocarcinoma

Mutation profiles, gene expression profiles, and clinical data for 482 cases of lung adenocarcinoma from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>) were used to explore mutation frequency and its correlation with patient survival. The percentage of lung adenocarcinoma patients having a mutation

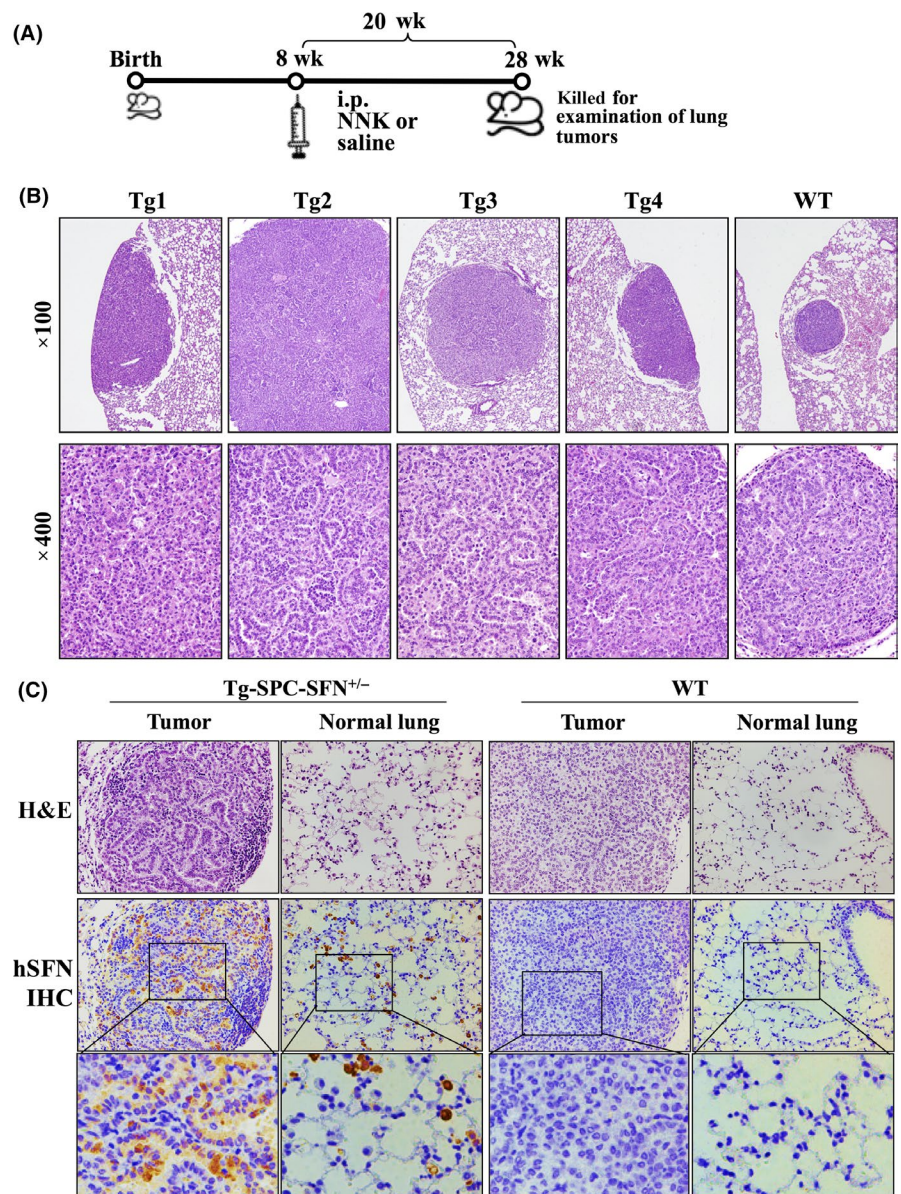


FIGURE 1 Histological analysis of tumors that developed in Tg-SPC-SFN^{+/-} mice. A, Schematic overview for experimental design of nicotine-derived nitrosaminoketone (NNK)-induced lung tumorigenesis in Tg-SPC-SFN^{+/-} and WT mice. At 8 wk after birth, all mice were given saline or NNK by i.p. injection. The mice were observed for 20 wk after administration with NNK or saline and killed at 28 wk of age to collect lung tumors. B, All tumors that developed in Tg-SPC-SFN^{+/-} or WT mice were subjected to H&E staining to examine their histology. Tg-SPC-SFN^{+/-} tumors were more differentiated adenoma than WT tumor, which resembles the histology of human lung adenocarcinoma. C, H&E and human SFN (hSFN) immunohistochemistry (IHC) images of representative paired samples of normal lung and tumor tissue in Tg-SPC-SFN^{+/-} and WT mice. hSFN was highly expressed in tumors, type II alveolar epithelial cells, and basal cells in the bronchial epithelium of the normal lung in Tg-SPC-SFN^{+/-} mice

in each gene was calculated as: number of cases where each gene is mutated / total number of cases tested for simple somatic mutations. Survival plots for each gene were examined using the Kaplan-Meier method through OncoLnc (<http://www.oncolnc.org>), a webtool for investigating the association between patient outcome and mRNA expression data for lung adenocarcinoma obtained from TCGA.²¹

2.5 | Statistical analysis

For the significant canonical pathways associated with mutated genes in WT or Tg-SPC-SFN^{+/-} tumors, we used the Benjamini-Hochberg method and adjusted the Fisher exact *P* value threshold to <.05. Additionally, we used the log-rank *P* value for significance of the survival rate between low and high gene expression.

3 | RESULTS

3.1 | Histological examination of tumors attributed to NNK treatment in Tg-SPC-SFN^{+/-} mice

To investigate the unique function of hSFN on tumorigenesis specifically in the lung, we generated transgenic mice (Tg-SPC-SFN^{+/-} mice) with overexpression of hSFN rather than mouse SFN (mSFN). As hSFN and mSFN show 98% homology in amino acid sequences (Figure S1A), we supposed that overexpression of hSFN might show no toxicity against mice, and needed to distinguish hSFN and mSFN to understand whether the resulting tumorigenesis is due to hSFN overexpression or endogenous mSFN expression.¹¹ To examine tumorigenic activity of hSFN, we i.p. injected NNK or saline as a control into Tg-SPC-SFN^{+/-} and WT mice (Figure 1A). Twenty weeks after the treatment, we killed all mice and collected the lungs and developing tumors, which were divided into 2 parts: histological examinations and WES. First, we undertook H&E staining to examine the histology of tumors that developed in Tg-SPC-SFN^{+/-} and WT mice. Although 10 tumors developed in Tg-SPC-SFN^{+/-} mice

and we selected 4 tumors that were sufficient size to examine WES, most WT tumors were small and detected in the central area of the lung when we undertook histological examinations. We obtained fresh material from only one WT tumor to be applied to WES. Interestingly, tumors that developed in Tg-SPC-SFN^{+/-} mice showed papillary or lepidic differentiation, which are major histologic subtypes of human lung adenocarcinoma, whereas those in WT mice showed the tendency to have poorer differentiated histology (Figure 1B,C). Moreover, by immunohistochemistry, we confirmed that tumors developing in Tg-SPC-SFN^{+/-} mice showed positivity for hSFN (Figure 1C). These results indicated that the transgenic mice had been generated successfully with appropriate gene recombination. As overexpression of SFN is a common event facilitating tumor progression in most human invasive adenocarcinomas,^{5,9} we speculated that genetic alterations in tumors that developed in Tg-SPC-SFN^{+/-} mice might mimic those in human lung adenocarcinoma.

3.2 | Whole-exome sequencing reveals exonic mutational signatures in Tg-SPC-SFN^{+/-} tumors

We undertook WES using paired samples of normal lung and tumor tissue described in Table 1 to identify genetic alterations due to overexpression of hSFN (Tables S1 and S2). After excluding single-nucleotide polymorphisms, the mean number of somatic mutations in 4 samples of Tg-SPC-SFN^{+/-} tumors was 373.8 (Figure S1B). Among nonsynonymous somatic mutations, 7.3% of them in average were nonsense mutations in Tg-SPC-SFN^{+/-} tumors, which included several tumor suppressor genes, such as *Banp*, *Smarcc1*, *Epha3*, *Ccdc154*, and *Ing5* (Figure S1C). Therefore, we expect that these tumor suppressor genes harboring nonsense mutations might produce incomplete forms and possibly induce tumor progression.

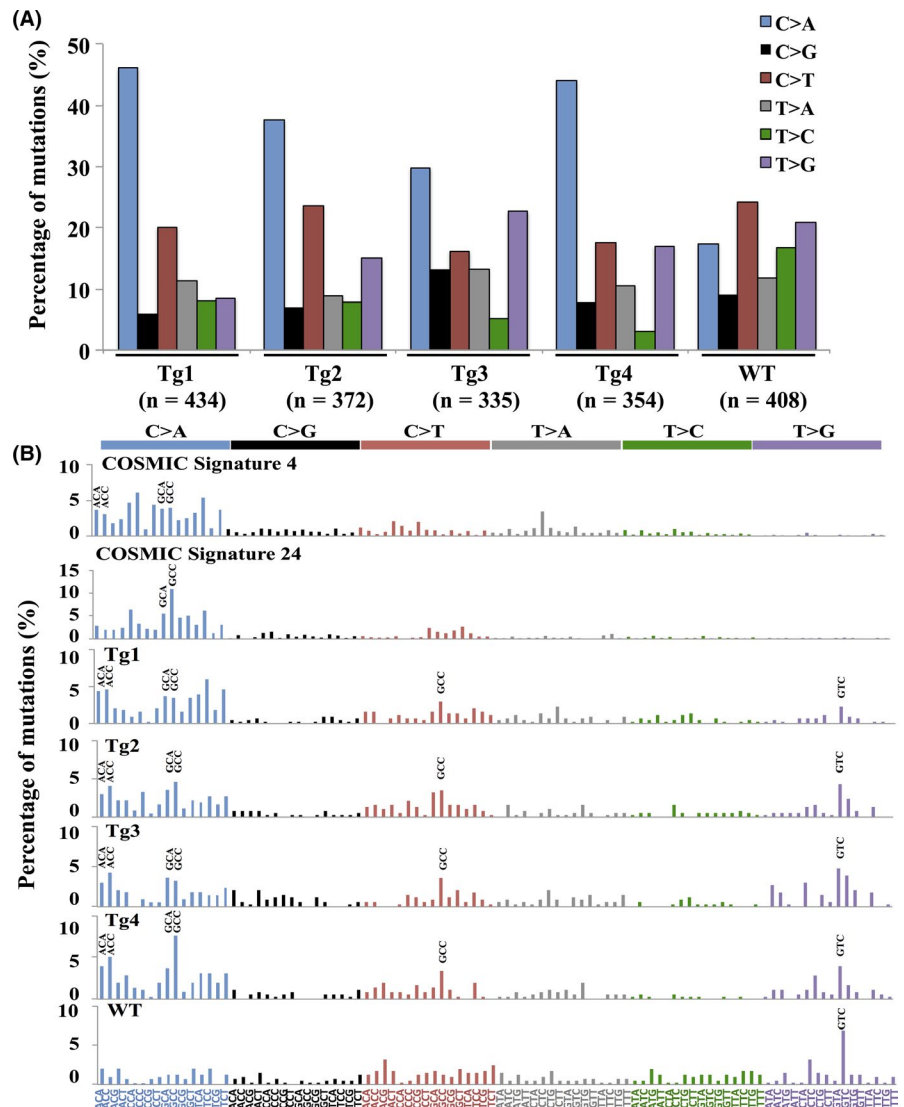
Moreover, according to previous reports, frequent somatic mutation of T>C or T>G in mice is caused by spontaneous point

	Sex	Treatment	Strain	Tissue type	Genotype
WT-CON	Male	Saline	ICR	Normal	Wild type
Tg-CON	Male	Saline	ICR	Normal	Tg-SPC-SFN ^{+/-}
WT	Male	NNK	ICR	Normal	Tg-SPC-SFN ^{+/-}
				Adenoma	
Tg1	Male	NNK	ICR	Normal	Tg-SPC-SFN ^{+/-}
				Adenoma	
Tg2	Male	NNK	ICR	Normal	Tg-SPC-SFN ^{+/-}
				Adenoma	
Tg3	Male	NNK	ICR	Normal	Tg-SPC-SFN ^{+/-}
				Adenoma	
Tg4	Female	NNK	ICR	Normal	Tg-SPC-SFN ^{+/-}
				Adenoma	

TABLE 1 Details of samples used for whole-exome sequencing analysis

Abbreviations: CON, control; NNK, nicotine-derived nitrosaminoketone; Tg, SFN-transgenic mouse; WT, wild-type mouse

FIGURE 2 Distribution of nucleotide alterations in Tg-SPC-SFN^{+/-} tumors and comparison with the human cancer mutational signature in the COSMIC database. A, Six classes of somatic single nucleotide mutations are indicated on the horizontal axis and the percentages of each mutation type are displayed on the vertical axis. The numbers below the chart indicate the mutation numbers detected in each tumor. B, Each signature is displayed according to the 96-substitution classification defined by the substitution class and sequence context immediately 5' and 3' to the mutated base. The frequency of flanking base pairs in Tg-SPC-SFN^{+/-} tumors is similar to COSMIC signatures 4 and 24, which are closely associated with smoking and aflatoxin in human cancer, respectively



mutations, unlike the situation in humans.²² Consistently, as shown in Figure 2A, WT tumor showed a high frequency of T>C and T>G. However, Tg-SPC-SFN^{+/-} tumors showed a high frequency of C>A (Figure 2A), a feature that is known to be associated with smoking in humans,²³ suggesting that Tg-SPC-SFN^{+/-} tumors might have a close resemblance to human tumors.

We also examined the mutational signatures of Tg-SPC-SFN^{+/-} and WT tumor. On the basis of the COSMIC database, the signatures of Tg-SPC-SFN^{+/-} tumors were similar to the COSMIC signatures 4 and 24, which show a high frequency of C>A.²⁴ Consistent with our results, TCGA also indicated that signatures 4 and 24 were frequently observed in human lung adenocarcinoma and showed a high frequency of C>A and C>T substitutions (Figure S2A,B). The Tg-SPC-SFN^{+/-} tumors typically had a high frequency of trinucleotide variants such as ACA, ACC, GCA, and GCC in C>A type and GCC in C>T type (Figure 2B). However, GTC showed the highest percentage of mutations in T>G, both in WT tumor and Tg-SPC-SFN^{+/-} tumors, indicating that GTC might be associated with NNK treatment or spontaneous mutations in mice (Figure 2B).

3.3 | Exonic mutational profiles of Tg-SPC-SFN^{+/-} tumors

Next, we analyzed associations between the number of mutations and *SFN* gene expression in lung adenocarcinoma patients in TCGA and lung cancer cell lines in the COSMIC database. The patients with high mutation number showed significantly higher expression of *SFN* than those with low mutation number (Figure S3A). Similar to lung adenocarcinoma patients, lung cancer cell lines with high mutation number, such as HCC827, A549, H1975, Calu-3, H1755, and H1573, had significantly higher expression of *SFN* than those with low mutation number, such as COR-L105 and Calu-6 (Figure S3B). Interestingly, the cell lines with high number of mutations, such as A549, H1755, and H1573 showed high frequency of C>A and C>T substitutions, a similar mutational signature of Tg-SPC-SFN^{+/-} tumors (Figures 2B and S3C).

Based on these results, we hypothesized that Tg-SPC-SFN^{+/-} tumors might be associated with high mutation burden, which is a critical indicator of the efficacy of immune checkpoint inhibitors.^{16,25}

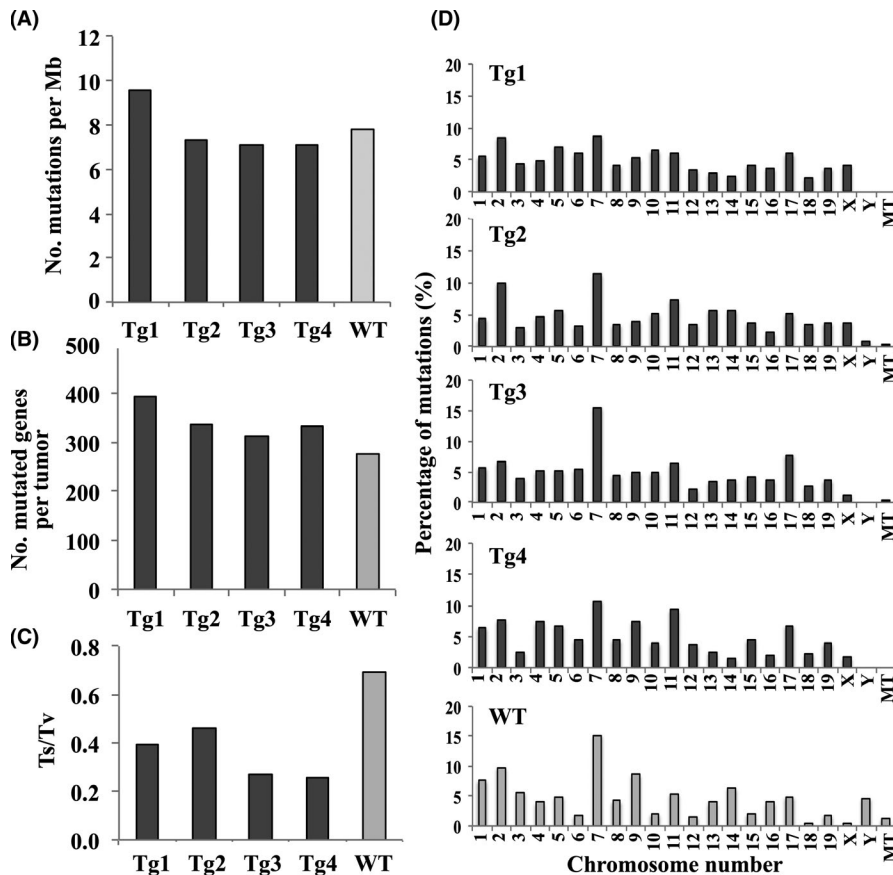


FIGURE 3 Distinct mutational spectrum patterns in Tg-SPC-SFN^{+/-} tumors. A, Exonic mutation frequencies were calculated as the number of mutations detected per million bases sequenced, which were estimated from the average coverage and base coverage data from whole-exome sequencing. B, Mutated genes were counted in each Tg-SPC-SFN^{+/-} and WT tumor. C, Ratios of transitions (Ts; C>T, T>C) and transversions (Tv; C>A, C>G, T>A, T>G) were calculated for Tg-SPC-SFN^{+/-} and WT tumors. D, Mutation frequency was counted according to the gene chromosome location in each Tg-SPC-SFN^{+/-} and WT tumor. Values indicate the percentage number of mutated genes among the total number of mutated genes detected in the each chromosome

Similar to the mutation burden in human lung cancer, which ranges from 0.1 to 100 nonsynonymous mutations per Mb,²⁶ both Tg-SPC-SFN^{+/-} and WT tumor showed approximately 7.8 per Mb (Figure 3A). However, the number of mutated genes was higher in all Tg-SPC-SFN^{+/-} tumors than WT tumor, indicating that the former had mutations in various genes whereas WT tumor tended to have multipoint mutations in particular genes (Figure 3B). Additionally, the transition to transversion (Ts/Tv) ratio is known to be widely variable across tumor types depending on the mutagens involved or the mechanisms responsible for DNA repair.²⁶ As the mean Ts/Tv ratio for Tg-SPC-SFN^{+/-} tumors was 0.34, being similar to that for human lung adenocarcinoma (LUAD Ts/Tv = 0.44), Tg-SPC-SFN^{+/-} tumors might have characteristics comparable to those of human lung adenocarcinoma (Figures 3C and S2C).

Next, we analyzed the mutation frequency in each chromosome. The Tg-SPC-SFN^{+/-} and WT tumors showed the highest frequency of mutated genes in chromosome 7, including the loci of *Blm*, *Herc2*, *Nlrp2*, and *Smg1* in Tg-SPC-SFN^{+/-} tumors (Figure 3D and Table S3).

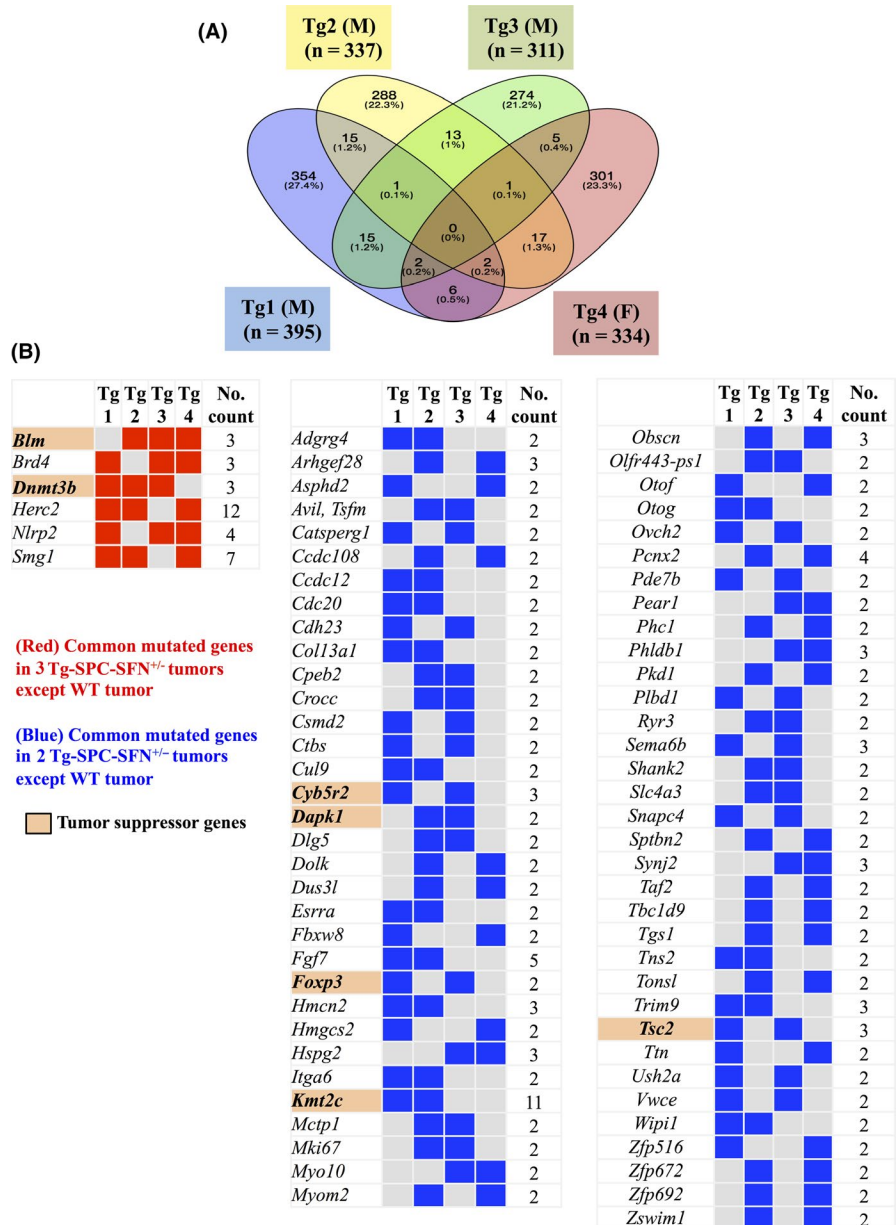
3.4 | Frequently mutated genes in Tg-SPC-SFN^{+/-} tumors

To identify common mutated genes among 4 Tg-SPC-SFN^{+/-} tumors, which were predicted to be associated with tumorigenesis, we drew Venn diagrams using single-nucleotide variants detected by WES (Figure 4A). We detected 73 genes that were commonly

detected in 2 or more samples of Tg-SPC-SFN^{+/-} tumors but not in WT tumor (Figure 4B and Table S4). Consistent with the previous study, which examined somatic mutations in NSCLC,^{18,27} we found that *Blm*, *Brd4*, *Dnmt3b*, *Herc2*, *Nlrp2*, and *Smg1* were common mutated genes in 3 samples of Tg-SPC-SFN^{+/-} tumors, whereas 67 other genes, including *Cdc20*, *Cdh23*, *Cul9*, *Fgf7*, *Itga6*, *Kmt2c*, *Mki67*, *Tbc1d9*, and *Tsc2*, were present in two samples of Tg-SPC-SFN^{+/-} tumors (Figure 4B). Notably, 8 genes contained identical mutation sites among Tg-SPC-SFN^{+/-} tumors, including G235E in *Blm*, I108V and R124C in *Smg1*, I92V in *Ctbs*, S283G in *Cyb5r2*, S153A in *Fgf7*, D665E/H in *Kmt2c*, H366R in *Ovch2*, and L575I in *Tsc2* (Table S2). However, according to TCGA, these mutated genes do not have mutational hot spots in human cancers. Similar to the previous study, in which lung tumors were induced in mice by exposure to carcinogens such as urethane,¹⁷ we detected *Kras* mutation (Q61R) only in WT tumor, indicating that *Kras* mutation could work as a driver mutation contributing to WT lung tumorigenesis (Table S2).

Moreover, Tg-SPC-SFN^{+/-} tumors showed accumulation of passenger gene and tumor suppressor gene mutations such as *Blm*, *Dnmt3b*, *Cyb5r2*, *Dapk1*, *Foxp3*, *Kmt2c*, and *Tsc2* rather than a specific driver oncogene mutation, such as *Kras* mutation in WT tumor (Figure 4B and Table S2). From these results, we expect that overexpression of *SFN* would be enough to induce tumor initiation and progression without any driver mutations, such as *Kras* mutation.

FIGURE 4 Frequently mutated genes identified in Tg-SPC-SFN^{+/-} tumors by whole-exome sequencing. A, Venn diagrams were used to identify commonly mutated genes among 4 Tg-SPC-SFN^{+/-} tumors. Numbers in each region indicate the numbers of mutated genes detected in Tg-SPC-SFN^{+/-} tumors. Percentages in each region indicate the percentage number of genes among the total number of mutated genes detected in the 4 Tg-SPC-SFN^{+/-} tumors. M, male; F, female; Tg, SFN-transgenic mouse. B, Somatic genes commonly mutated among Tg-SPC-SFN^{+/-} tumors were selected. The counted number indicates the mutation frequency of each gene detected in Tg-SPC-SFN^{+/-} tumors by whole-exome sequencing. Each column represents an individual tumor, and each row denotes a gene. All mutations were nonsynonymous missense mutations



3.5 | Impact of mutated genes in mice and human lung adenocarcinoma

To clarify distinct signaling pathways related to tumorigenesis in Tg-SPC-SFN^{+/-} and WT mice, we carried out pathway analyses for exonic mutations using the KEGG pathway and IPA software. First, we selected particular signaling pathways related to cell growth or cancer, such as PI3K/AKT signaling, ERK/MAPK signaling, STAT3 signaling, the DNA damage response and repair system, cell cycle signaling, p53 signaling, PTEN signaling, and 14-3-3-mediated signaling, and counted the number of mutated genes categorized into those pathways in each tumor using the KEGG pathway and IPA software. This revealed that all Tg-SPC-SFN^{+/-} tumors had a high percentage of mutated genes related to PI3K/AKT signaling, such as *Fgf7*, *Iiga6*, *Synj2*, and *Tsc2* and the apoptosis signaling pathway (Figure 5A and Table S5).

We then collected all mutated genes in 4 Tg-SPC-SFN^{+/-} tumors and investigated pathways that were significantly enriched using the KEGG pathway and IPA software. The KEGG pathway analysis revealed that mutated genes in Tg-SPC-SFN^{+/-} tumors were significantly associated with pathways in cancer, ubiquitinated proteolysis, small-cell lung cancer, the PI3K-AKT signaling pathway, and transcriptional misregulation in cancer (Table S6). In contrast, IPA software analysis indicated that Tg-SPC-SFN^{+/-} tumors harbored gene mutations that were significantly correlated with canonical pathways such as 14-3-3-mediated signaling, apoptosis signaling, PTEN signaling, and DNA double-strand break repair by homologous recombination (Table S6). Taken together, overexpression of SFN in the lung appeared to induce gene mutations mainly involved in PI3K/AKT signaling and apoptosis signaling pathways facilitating tumorigenesis.

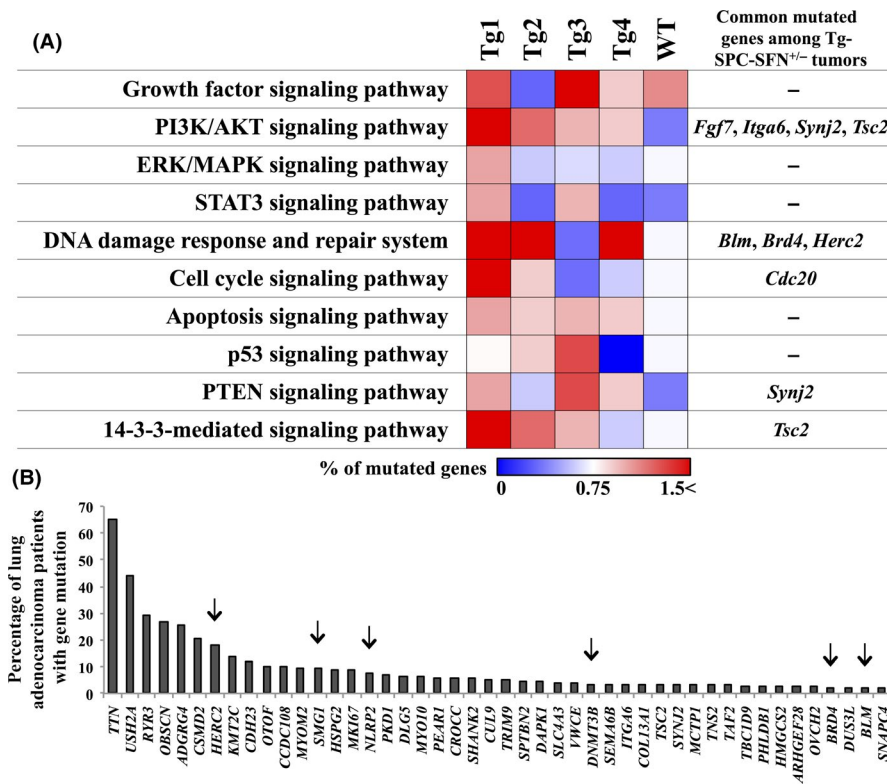


FIGURE 5 Pathway analysis for mutated genes in Tg-SPC-SFN^{+/-} tumors and frequencies of commonly mutated genes in lung adenocarcinoma patients. A, IPA and KEGG pathway analysis were used to investigate the molecular functions of each of the mutated genes in Tg-SPC-SFN^{+/-} and WT tumors. The heatmap indicates the percentages of mutated genes among the total number of mutated genes detected in the each Tg-SPC-SFN^{+/-} and WT tumor and was drawn using Morpheus, a webtool (<https://software.broadinstitute.org/morpheus/>). Commonly mutated genes among Tg-SPC-SFN^{+/-} tumors, which are included in each pathway, are indicated in the right column. B, Percentages of lung adenocarcinoma patients harboring each mutation commonly found in Tg-SPC-SFN^{+/-} tumors are displayed in the bar graph. TCGA was used for this analysis. Arrows denote commonly mutated genes in 3 Tg-SPC-SFN^{+/-} tumors apart from WT tumor

Next, using TCGA, we investigated the mutation frequency of 73 genes that were commonly mutated in Tg-SPC-SFN^{+/-} tumors in human lung adenocarcinomas. Most of the top-ranked genes were not well known, and likely passenger mutations that do not directly confer a selective tumor growth advantage. Among them, titin (*TTN*), a so-called giant gene with 363 coding exons in the human genome and tending to have frequent mutations because of random DNA repair error,^{28,29} was the top ranked gene, harbored by 314 of 487 lung adenocarcinoma patients. Moreover, *HERC2*, *SMG1*, and *NLRP2*, which were detected in 3 samples of Tg-SPC-SFN^{+/-} tumors, also had a high ranking for mutation frequency in human lung adenocarcinoma (Figure 5B and Table S7).

In order to clarify the impact of these gene alterations on prognosis, overall survival was compared between patients showing high expression and those showing low expression using TCGA (Table S8). This revealed that patients with high expression of *BLM*, *DNMT3B*, *MKI67*, *ITGA6*, or *CDC20* had a significantly poorer outcome than those with low expression of the genes (Figure S4A). In contrast, patients with high expression of *NLRP2*, *CUL9*, *TBC1D9*, or *CDH23* had a significantly more favorable outcome relative to those with low expression of these genes (Figure S4B).

In summary, overexpression of *SFN* might frequently induce a large number of mutations, most of which are involved in PI3K/AKT

signaling and apoptosis signaling, and altered expression of those genes affects patient outcome.

4 | DISCUSSION

Recently, the genomic alterations driving lung adenocarcinogenesis have been investigated to clarify the molecular mechanisms involved, discover novel drug targets, and screen candidate patients for clinical trials using next-generation sequencing.^{27,30} Extensive molecular analysis has shown that normal cells progressively acquire a variety of genetic imbalances that lead to tumorigenesis.³¹ Tumorigenesis is mediated by a combination of events including upregulation of oncogenes sustaining proliferative signaling, downregulation of tumor suppressor genes, and also resistance to apoptotic cell death, thus enabling replicative immortality.³²

In the present study, by analyzing tumors that developed in Tg-SPC-SFN^{+/-} mice after exposure to NNK, we showed that *SFN* had potent oncogenic functions causing distinct genetic alterations during the process of tumorigenesis. As Tg-SPC-SFN^{+/-} mice without NNK treatment developed few but histologically similarly tumors to NNK-induced tumors in Tg-SPC-SFN^{+/-} mice,¹¹ we consider that overexpression of hSFN in the lung in ICR mice robustly induces initiation

of tumor formation, similar to NNK exposure. Moreover, all NNK-induced tumors in Tg-SPC-SFN^{+/-} mice have high similarity to most human lung adenocarcinomas in terms of histology³³ and genetic profiles, including a high frequency of passenger mutations, a high frequency of transversion, a large mutation burden, and smoking-like signatures (Figures 1-3, S2, and S3). Moreover, tumors that developed in Tg-SPC-SFN^{+/-} mice did not contain *Kras* mutation, known as a driver mutation for murine lung cancer.³⁴ Therefore, we believe that this mouse model closely mimics actual human adenocarcinoma.

We revealed that NNK induced tumor formation accompanied by frequent C>T transition substitution in both Tg-SPC-SFN^{+/-} and WT mice (Figure 2A), indicating that the high frequency of this substitution is caused by a defective DNA repair process attributed to NNK.³⁵ Consistent with our results, NNK is known to be a strong tobacco-derived carcinogen and undergoes metabolism to DNA adducts, which then bind to DNA, resulting in mutation of *KRAS* and *STK11*, and thus lung tumor formation in various animal models.^{36,37} Unlike tobacco smoke, which contains more than 70 carcinogenic compounds such as benzo[a]pyrene, which predominantly induces C>A transversion mutation, treatment of rats with NNK leads to frequent C>T transition mutation caused by O⁶-mGua, which is a highly promutagenic adduct induced by NNK and can be removed by DNA repair proteins such as O-6-methylguanine-DNA methyltransferase.^{35,38}

In contrast, mutated genes in Tg-SPC-SFN^{+/-} tumors had a high frequency of C>A transversion (Figure 2B). This substitution might be due to failure of the transcription-coupled nucleotide excision repair machinery, similar to human mutational signatures that are correlated with smoking history.³⁹ Such mutational signatures can indicate the cause of somatic mutations such as DNA replication infidelity, defective DNA repair processes, and mutagen exposure occurring throughout life.⁴⁰ Among 30 mutational signatures in the COSMIC database, signature 4 is the most closely related to lung cancers associated with smoking,^{23,39} and signature 24 is linked to liver cancer and exposure to aflatoxin, a carcinogen produced by fungi in nature,³⁹ both of which show similarity to Tg-SPC-SFN^{+/-} tumors.

Through interaction with the target proteins, 14-3-3 proteins including SFN alter the activity, modifications, and intracellular localization of ligands.⁴¹⁻⁴³ In particular, SFN regulates DNA repair factors such as DNA-PKcs, when DNA is damaged by chemotherapy or radiotherapy in human cancer.⁴⁴ Based on the previous report, we assume that the high number of gene mutations in Tg-SPC-SFN^{+/-} tumors might be due to binding with unknown factor(s), which is associated with the DNA repair system. SFN might inhibit DNA repair-related factors, resulting in the accumulation of numerous gene mutations. Indeed, we found some mutated genes involved in the DNA damage response and repair system such as *Blm*, *Brd4*, *Herc2*, *Brca1*, *Pole*, *Ddb1*, *Msh6*, and *Xrcc1* only in Tg-SPC-SFN^{+/-} tumors (Table S5), suggesting that these alterations could also indirectly accelerate dysfunction of the DNA repair system and thus these mutated genes might cause additional mutation again. High tumor mutation burden is known to induce high neoantigen expression,⁴⁵ which enhances tumor growth, invasion, and metastasis by evasion

from immune attack.⁴⁶ It implies that Tg-SPC-SFN^{+/-} might serve as a useful tool to evaluate the efficacy of immune checkpoint inhibitors.

Additionally, our previous pull-down assay and LC-MS/MS analysis has identified SFN binding proteins in lung adenocarcinoma cells including cell cycle-related proteins such as S-phase kinase tumor progression (SKP1),¹⁰ and proliferation related proteins such as serine/threonine-protein kinase A-Raf (ARAF)⁴⁷ and ubiquitin-specific protein 8 (USP8), which particularly enhances stability of receptor tyrosine kinase such as epidermal growth factor receptor and MET.⁹ We speculate that, as SFN controls those proteins, tumor cells might pass through the cell cycle checkpoint even though they recognize a large number of mutations due to insufficient DNA repair.⁴⁸ Consequently, SFN might induce a large number of mutations, which are predominantly involved in PI3K-AKT signaling and apoptosis signaling (Figure 5A), and those mutations might boost canonical signaling for adenocarcinogenesis.

Moreover, we revealed that Tg-SPC-SFN^{+/-} tumors harbored mutations of *Blm*, *Dnmt3b*, *Mki67*, *Itga6*, *Cdc20*, *Nlrp2*, *Cul9*, *Tbc1d9*, and *Cdh23* that were also detectable in lung adenocarcinoma patients, and that the expression of these genes was significantly correlated with overall survival (Figure S4). However, the mutational hot spots and oncogenic activity of these mutated genes in lung cancers have not yet been clarified. Although there is a possibility that these mutated genes might be involved in lung adenocarcinogenesis triggered by SFN, further verification and functional analysis using human lung adenocarcinoma specimens will be needed in order to understand the overall molecular mechanisms involved. Additionally, as WES has limitations to detect in particular genetic alterations such as mutations, deletions, and insertions, we are planning to undertake whole genome amplification-sequencing or RNA-sequencing to investigate the amplification or expression of genes for further study.

Furthermore, we examined the mutational landscape in both Tg-SPC-SFN^{+/-} tumors and WT tumor. Compared to the Tg-SPC-SFN^{+/-} tumors, WT tumor showed poorer differentiated histology, higher frequency of T>C and T>G, and contained a lower number of mutated genes. Additionally, only WT tumor contained *Kras* mutation. These differences between Tg-SPC-SFN^{+/-} and WT tumors are very interesting. Unfortunately, in the current study, we could examine only one WT tumor, as the tumorigenic frequency is very low in ICR mice by NNK treatment. Large-scale examinations should be planned in the future to make clear the characteristics of Tg-SPC-SFN^{+/-} tumors.

In summary, we characterized tumors that developed in Tg-SPC-SFN^{+/-} mice and found distinct exonic alterations, remarkably mimicking the features of human lung adenocarcinoma. We believe that SFN overexpression is critical to human lung adenocarcinogenesis and that our mouse model, created by both genetic engineering and carcinogen exposure, would be a valuable resource for preclinical examinations of therapy for lung adenocarcinoma.

DISCLOSURE

The authors have no financial conflicts of interest related to this manuscript.

ORCID

Aya Shiba-Ishii  <https://orcid.org/0000-0001-6639-4135>

REFERENCES

- McGuire S. *World Cancer Report 2014*. Geneva, Switzerland: World Health Organization, International Agency for Research on Cancer, WHO Press, 2015. *Adv Nutr*. 2016;7:418-419.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018;68:7-30.
- Travis WD, Brambilla E, Nicholson AG, et al. The 2015 world health organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J Thorac Oncol*. 2015;10:1243-1260.
- Noguchi M. Stepwise progression of pulmonary adenocarcinoma—clinical and molecular implications. *Cancer Metastasis Rev*. 2010;29:15-21.
- Shiba-Ishii A, Kano J, Morishita Y, Sato Y, Minami Y, Noguchi M. High expression of stratifin is a universal abnormality during the course of malignant progression of early-stage lung adenocarcinoma. *Int J Cancer*. 2011;129:2445-2453.
- Qi W, Liu X, Qiao D, Martinez JD. Isoform-specific expression of 14-3-3 proteins in human lung cancer tissues. *Int J Cancer*. 2005;113:359-363.
- Shiba-Ishii A, Noguchi M. Aberrant stratifin overexpression is regulated by tumor-associated CpG demethylation in lung adenocarcinoma. *Am J Pathol*. 2012;180:1653-1662.
- Husni RE, Shiba-Ishii A, Nakagawa T, et al. DNA hypomethylation-related overexpression of SFN, GORASP2 and ZYG11A is a novel prognostic biomarker for early stage lung adenocarcinoma. *Oncotarget*. 2019;10:1625-1636.
- Kim Y, Shiba-Ishii A, Nakagawa T, et al. Stratifin regulates stabilization of receptor tyrosine kinases via interaction with ubiquitin-specific protease 8 in lung adenocarcinoma. *Oncogene*. 2018;40:5387-5402.
- Shiba-Ishii A, Hong J, Hirokawa T, et al. Stratifin inhibits SCFFBW7 formation and blocks ubiquitination of oncoproteins during the course of lung adenocarcinogenesis. *Clin Cancer Res*. 2019;25:2809-2820.
- Shiba-Ishii A, Kim Y, Shiozawa T, et al. Stratifin accelerates progression of lung adenocarcinoma at an early stage. *Mol Cancer*. 2015;14:142.
- Li Z, Liu JY, Zhang JT. 14-3-3sigma, the double-edged sword of human cancers. *Am J Transl Res*. 2009;1:326-340.
- Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511:543-550.
- Rabbani B, Tekin M, Mahdieh N. The promise of whole-exome sequencing in medical genetics. *J Hum Genet*. 2014;59:5-15.
- Schwarze K, Buchanan J, Taylor JC, Wordsworth S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet Med*. 2018;20:1122-1130.
- Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015;348:124-128.
- Westcott PM, Halliwill KD, To MD, et al. The mutational landscapes of genetic and chemical models of Kras-driven lung cancer. *Nature*. 2015;517:489-492.
- McFadden DG, Politi K, Bhutkar A, et al. Mutational landscape of EGFR-, MYC-, and Kras-driven genetically engineered mouse models of lung adenocarcinoma. *Proc Natl Acad Sci USA*. 2016;113:E6409-e17.
- Shen Y, Zhang S, Huang X, Chen K, Shen J, Wang Z. Involvement of p53 mutation and mismatch repair proteins dysregulation in NNK-induced malignant transformation of human bronchial epithelial cells. *Biomed Res Int*. 2014;2014:920275.
- Pfeifer GP, Denissenko MF, Olivier M, Tretyakova N, Hecht SS, Hainaut P. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene*. 2002;21:7435-7451.
- Anaya J. OncoLnc: linking TCGA survival data to mRNAs, miRNAs, and lncRNAs. *PeerJ Comput Sci*. 2016;2:e67.
- Milholland B, Dong X, Zhang L, Hao X, Suh Y, Vijg J. Differences between germline and somatic mutation rates in humans and mice. *Nat Commun*. 2017;8:15183.
- Govindan R, Ding L, Griffith M, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell*. 2012;150:1121-1134.
- Alexandrov LB, Ju YS, Haase K, et al. Mutational signatures associated with tobacco smoking in human cancer. *Science*. 2016;354:618-622.
- Hellmann MD, Nathanson T, Rizvi H, et al. Genomic features of response to combination immunotherapy in patients with advanced non-small-cell lung cancer. *Cancer Cell*. 2018;33:843-852.e4.
- Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499:214-218.
- Wang D, Pham NA, Tong J, et al. Molecular heterogeneity of non-small cell lung carcinoma patient-derived xenografts closely reflect their primary tumors. *Int J Cancer*. 2017;140:662-673.
- Tan H, Bao J, Zhou X. Genome-wide mutational spectra analysis reveals significant cancer-specific heterogeneity. *Sci Rep*. 2015;5:12566.
- Savarese M, Sarparanta J, Vihola A, Udd B, Hackman P. Increasing role of titin mutations in neuromuscular disorders. *J Neuromuscul Dis*. 2016;3:293-308.
- Devarakonda S, Morgensztern D, Govindan R. Genomic alterations in lung adenocarcinoma. *Lancet Oncol*. 2015;16:e342-e351.
- Kadara H, Scheet P, Wistuba II, Spira AE. Early events in the molecular pathogenesis of lung cancer. *Cancer Prev Res (Phila)*. 2016;9:518-527.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144:646-674.
- Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/American thoracic society/European respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol*. 2011;6:244-285.
- Kadota K, Yeh YC, D'Angelo SP, et al. Associations between mutations and histologic patterns of mucin in lung adenocarcinoma: invasive mucinous pattern and extracellular mucin are associated with KRAS mutation. *Am J Surg Pathol*. 2014;38:1118-1127.
- Xue J, Yang S, Seng S. Mechanisms of cancer induction by tobacco-specific NNK and NNN. *Cancers (Basel)*. 2014;6:1138-1156.
- Ge GZ, Xu TR, Chen C. Tobacco carcinogen NNK-induced lung cancer animal models and associated carcinogenic mechanisms. *Acta Biochim Biophys Sin (Shanghai)*. 2015;47:477-487.
- Ding L, Getz G, Wheeler DA, et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature*. 2008;455:1069-1075.
- Jansen JG, de Groot AJ, van Teijlingen CM, Tates AD, Vrieling H, van Zeeland AA. Induction of hprt gene mutations in splenic T-lymphocytes from the rat exposed in vivo to DNA methylating agents is correlated with formation of O6-methylguanine in bone marrow and not in the spleen. *Carcinogenesis*. 1996;17:2183-2191.
- Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500:415-421.
- Alexandrov LB, Stratton MR. Mutational signatures: the patterns of somatic mutations hidden in cancer genomes. *Curr Opin Genet Dev*. 2014;24:52-60.
- Pennington KL, Chan TY, Torres MP, Andersen JL. The dynamic and stress-adaptive signaling hub of 14-3-3: emerging mechanisms of

- regulation and context-dependent protein-protein interactions. *Oncogene*. 2018;37:5587-5604.
42. Aghazadeh Y, Papadopoulos V. The role of the 14-3-3 protein family in health, disease, and drug development. *Drug Discov Today*. 2016;21:278-287.
43. Dougherty MK, Morrison DK. Unlocking the code of 14-3-3. *J Cell Sci*. 2004;117:1875-1884.
44. Chen Y, Li Z, Dong Z, et al. 14-3-3sigma contributes to radioresistance by regulating DNA repair and cell cycle via PARP1 and CHK2. *Mol Cancer Res*. 2017;15:418-428.
45. Miller A, Asmann Y, Cattaneo L, et al. High somatic mutation and neoantigen burden are correlated with decreased progression-free survival in multiple myeloma. *Blood Cancer J*. 2017;7:e612.
46. Lee CH, Yelensky R, Jooss K, Chan TA. Update on tumor neoantigens and their utility: why it is good to be different. *Trends Immunol*. 2018;39:536-548.
47. Mooz J, Oberoi-Khanuja TK, Harms GS, et al. Dimerization of the kinase ARAF promotes MAPK pathway activation and cell migration. *Sci Signal*. 2014;7:ra73.
48. Jeggo PA, Lobrich M. Contribution of DNA repair and cell cycle checkpoint arrest to the maintenance of genomic stability. *DNA Repair (Amst)*. 2006;5:1192-1198.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Kim Y, Shiba-Ishii A, Ramirez K, et al. Carcinogen-induced tumors in *SFN*-transgenic mice harbor a characteristic mutation spectrum of human lung adenocarcinoma. *Cancer Sci*. 2019; 110: 2431-2441. <https://doi.org/10.1111/cas.14081>