

# The Genomic Characteristics of RET Fusion Positive Tumors in Chinese Non-Small Cell Lung Cancer (NSCLC) Patients

**Guowu Wu** (✉ [guowuwugd@163.com](mailto:guowuwugd@163.com))

Meizhou Peoples's Hospital(Huangtang Hospital) <https://orcid.org/0000-0002-2022-3258>

**Longhua Guo**

Meizhou's Hospital(Huangtang Hospital)

**Yinfang Gu**

Meizhou's Hospital(Huangtang Hospital)

**Tanxiao Huang**

HaploX Biotechnology

**Ming Liu**

HaploX Biotechnology

**Xiaofang Zou**

Meizhou People's hospital (Huangtang Hospital)

**Bo Yang**

HaploX Biotechnology

**Ping Huang**

Meizhou Peoples' Hospital(Huangtang Hospital)

**Chunling Wen**

Meizhou People's Hospital(Huangtang Hospital)

**Lilan Yi**

Meizhou People's Hospital

**Wenting Liao**

HaploX Biotechnology

**Dongdong Zhao**

Meizhou People's Hospital

**Junlin Zhu**

Meizhou People's Hospital

**Xiaoni Zhang**

HaploX Biotechnonlogy

**Yuanyuan Liu**

HaploX Biotechnonlogy

**Yan Yin**

HaploX Biotechnology

**Shifu Chen**

HaploX Biotechnology

---

## Research Article

**Keywords:** RET, Non-Small Cell Lung Cancer, Next Generation Sequencing

**Posted Date:** September 28th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-655553/v2>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Journal of Cancer Research and Clinical Oncology on February 26th, 2022. See the published version at <https://doi.org/10.1007/s00432-022-03959-6>.

# Abstract

**Background:** Approximately 1-2% of non-small cell lung cancer (NSCLC) patients harbor RET (rearranged during transfection) fusions. The oncogenic RET fusions could lead to constitutive kinase activation and oncogenesis.

**Methods:** 1746 Chinese NSCLC patients were enrolled from in this study. Tumor tissues were collected during surgery, and were formalin fixed, paraffin-embedded (FFPE) and archived. Peripheral blood (PB) samples were also collected from each patient as control. In addition, we selected 17 of them for cfDNA NGS testing and 14 tumor samples for immunohistochemistry testing using PD-L1 rabbit monoclonal antibody, clones 28-8 (Abcam, Cambridge, UK).

**Results:** Of the 1746 NSCLC cases, RET rearrangements were identified in 25 cases (1.43 %) with locally advanced or metastatic NSCLC, of which 20 (80%) were female. We found that 14 out of 25 patients had a KIF5B-RET fusion, with KIF5B exon15-RET exon12, KIF5B exon23-RET exon12, and KIF5B exon24-RET exon11 detected in 14, 3, and 1 patients respectively. We also identified one novel RET fusion partner PLCE1 and 4 intergenic-breakpoint fusions.

**Conclusion:** In this study, using the hybrid capture based NGS (the next generation sequencing) techniques, we revealed the genomic profiling for the patients with RET fusion-positive NSCLC. To the best of our knowledge, this is the first study that exhibited the detailed breakpoints of Chinese NSCLC patients with RET rearrangement, and we found a novel new partner PLCE1. The results provided genomic information for patients with RET fusion which is significant for personalized clinical management in the era of precision medicine.

## Background

It was not until 2012 that the RET fusions were identified as oncogenic drivers in NSCLC [1]. Approximately 1–2% of NSCLC patients harbor RET fusions [2], and they tend to be young, non-smokers, and adenocarcinomas[3]. The RET proto-oncogene encodes a transmembrane receptor tyrosine kinase belonging to members of the glial cell line-derived neurotrophic factor (GDNF) family. The RET signaling is modulated by ligands and could activate multiple downstream pathways, such as RAS/MAPK/ERK, PI3K/AKT and JAK/STAT, which are essential for cellular differentiation and proliferation. However, the oncogenic RET fusions could lead to constitutive kinase activation and oncogenesis [4–6]. RET fusions are caused by chromosomal rearrangement, which fuses the 3' coding regions for the RET kinase domain on chromosome 10 with a 5' upstream partner gene containing coiled-coil or LIS1 homology domains [7–9]. The most common partner genes are KIF5B, CCDC6 and NCOA4, which originate from intrachromosomal rearrangements, but there are also some relatively infrequent interchromosomal partner genes, such as TRIM33, TRIM24, MYO5C, EPHA5, CLIP1, and so on [10, 11].

The advent of precision medicine has revolutionized the therapeutic landscape of NSCLC, and targeted therapies have been investigated in many clinical studies against patients with RET fusion-positive

NSCLC, which includes multi-kinase inhibitors (MKIs) and selective RET inhibitors [12]. MKIs target not only RET, but also other kinases, such as VEGFR2, KIT, BRAF, etc. This probably leads to decreased effectiveness against RET and off-target side effects. Several MKIs that were approved by FDA for cancer therapy on other purposes (such as sunitinib, sorafenib, vandetanib, cabozantinib, regorafenib, lenvatinib, and alectinib) all showed modest clinical activity and had side effects for patients with RET fusion-positive NSCLC[13–18]. Besides, MKI resistance is almost inevitable. The common resistant mechanisms are secondary MET gene mutation and downstream signaling pathway activation [19, 20]. These limitations in MKIs prompt the development of selective RET inhibitors. On May 8, 2020, the FDA granted accelerating approval to selpercatinib for adult patients with metastatic RET fusion-positive NSCLC. This was the first targeted therapy approved for RET fusion-positive NSCLC [21]. The other selective RET inhibitor, pralsetinib, was also granted as Breakthrough Therapy designation by the FDA for advanced NSCLC with RET fusions after progression on platinum chemotherapy. Both of them showed higher response rates and tolerability, but drug resistance were also inevitable for them [22, 23].

The development of hybrid capture-based NGS techniques brings great convenience in revealing the genomic profiling for cancer patients. It's becoming more and more available in clinical cancer treatment. It not only could be used to identify the RET fusion events, but also allows us to investigate the co-occurring genomic alterations in the same assay, which may be related to prognosis or therapeutic response. In this study, we sequenced 1746 NSCLC patients by hybrid capture-based NGS techniques, and eventually identified 25 RET fusion-positive cases.

## Methods

### Patients and samples

In this study, 1746 Chinese NSCLC patients were prospectively enrolled. Tumor tissues and peripheral blood (PBL) samples were collected for each patient. Tumor tissues were formalin fixed, paraffin-embedded (FFPE) while PB samples were tested as control. In addition, we selected 17 of them for cfDNA NGS testing and 14 tumor samples for immunohistochemistry testing using PD-L1 rabbit monoclonal antibody, clones 28 – 8 (Abcam, Cambridge, UK).

### FFPE DNA extraction

DNA samples from tumor tissues were extracted using QIAamp DNA FFPE tissue kit (Qiagen) after paraffin-embedded. DNA samples from PBL were extracted using the RelaxGene Blood DNA system (Tiangen Biotech Co., Ltd., Beijing, China). Quantification of all the DNA samples were conducted by both the Qubit 2.0 fluorometer and the Qubit dsDNA HS Assay kit (Thermo Fisher Scientific, Inc., Waltham, MA, USA).

### Plasma isolation and cfDNA extraction.

Blood samples from patients were collected in tubes containing EDTA and centrifuged at 1600 g for 10 min at 4°C within 2 hours of collection. The peripheral blood lymphocyte (PBL) debris was stored at -20°C

until further use. The supernatants were further centrifuged at 10,000 g for 10 min at 4°C, and plasma was harvested and stored at -80°C until further use. DNA from PBLs was extracted using RelaxGene Blood DNA System (TianGen Biotech Co., Ltd., Beijing, China), and cell free DNA (cfDNA) was extracted from at least 2 mL plasma using QIAamp Circulating Nucleic Acid kit (QIAGEN) following the manufacturers' instructions, respectively. Extracted DNA was then quantified by Qubit 2.0 (Thermo Fisher Scientific, Inc., Waltham, MA, USA), according to manufacturer's instructions.

### **FFPE and genomic DNA Library construction and sequencing**

100ng of FFPE DNA and genomic DNA from PBLs for each patient was sheared by the dsDNA Fragmentase (New England BioLabs, Inc., Ipswich, MA, USA). And then size of 150–250 bp were selected using Ampure XP beads (Beckman Coulter, Inc., Brea, CA, USA). Library construction was conducted using the KAPA Library Preparation kit (Kapa Biosystems, Inc., Wilmington, MA, USA). And then the concentration assessment of the library was performed using the e Qubit dsDNA HS Assay kit, while the fragment length was acquired on a 4200 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA). Targeted capture was performed using a custom set of biotinylated DNA probes (HapOncoCDx panel) which contain 464 cancer-related genes encompassing 1.31 Mb (Roche NimbleGen). The hybridization of the amplified sample libraries and the SeqCap EZ Library was used according to the manufacturer's protocol for 16–20h at 47°C. After hybrid selection, the captured DNA fragments were amplified with 12 to 14 cycles of PCR using 1× KAPA HiFi Hot Start Ready Mix and Post-LM-PCR Oligos in two separate 50µL reactions. The reactions were then pooled and purified by Agencourt AMPure XP beads. DNA sequencing was then performed using Illumina Novaseq 6000 system with an average depth at 2000X.

### **CfDNA Library construction and sequencing**

Library construction was conducted using cfDNA with the KAPA Library Preparation kit (Kapa Biosystems, Inc., Wilmington, MA, USA). Agencourt AMPure XP beads (Beckman Coulter, Inc., Brea, CA, USA) were applied for cleanup steps. DNA fragments were purified using the Qubit dsDNA HS Assay kit and the concentration was evaluated by the Qubit 2.0 fluorometer. Then end repair and 3'-end A-tailing were conducted. Ligation was performed at 20°C for 15min. Single-step size selection was achieved by 50 µL (1×) of PEG/NaCl SPRI Solution buffer. Then the ligated fragments were amplified in 1× KAPA HiFi Hot Start Ready Mix with Pre-LM-PCR Oligos in 50µL reactions, thereafter PCR were performed with 7–12 cycles depending on the quantity of input DNA. Qubit 2.0 Fluorometer and Qubit dsDNA HS Assay kit were applied again to evaluate the library's purity and concentration. Fragment length was detected on a 4200 Bioanalyzer using DNA 1000 Kit (Agilent).

Targeted capture was conducted using customized HapOncoCDx panel. The hybridization of the amplified sample libraries and the SeqCap EZ Library was used according to the manufacturer's protocol for 16–20h at 47°C. After hybrid selection, the captured DNA fragments were amplified in PCR reaction using 1× KAPA HiFi Hot Start Ready Mix and Post-LM-PCR Oligos with 12–14 cycles. The reactions were then purified by Agencourt AMPure XP beads. Multiplexed libraries were denatured by Tris-HCl and diluted by

0.2N NaOH according to the manufacturer's protocol (Illumina). Then the libraries were sequenced using 150-bp paired-end runs on an Illumina NovaSeq 6000 system (Illumina).

## Data analysis and variant calling

Raw data were pre-processed by fastp with version 0.18.0 (<https://github.com/OpenGene/fastp>) [24]. Then clean reads were aligned to hg19 genome (GRch37) using maximal exact matches algorithm of Burrows-Wheeler Aligner [25]. Duplicate reads were removed by Gencore version 0.12.0(<https://github.com/OpenGene/gencore>) [26]. After applying Samtools version 0.1.19(<http://www.htslib.org/>) [27], pileup files were generated with mapping quality  $\geq 60$ . VarScan2 with version 2.3.8 (<http://varscan.sourceforge.net/>) [28] were applied to call somatic variants [minimum read depth = 20; variant allele frequency (VAF) threshold  $\geq 0.01$ ; somatic-P-value  $\leq 0.01$ ; strand-filter = 1; other parameters, default]. Copy number variation were detected using CNV kit with version 0.9.3(<https://github.com/etal/cnvkit>) [29], while structural variation were calculated by GeneFuse with version v0.6.1 (<https://github.com/OpenGene/GeneFuse>) [30]. Microsatellite instability (MSI) status were determined by VisualMSI(<https://github.com/OpenGene/VisualMSI>)[31]. Maftools were used for visualizing somatic variant analysis [32]. Data which met the following criteria were chosen for subsequent analysis: the ratio of remaining data filtered by fastq in raw data is  $\geq 85\%$ ; the proportion of Q30 bases is  $\geq 85\%$ ; the ratio of reads on the reference genome is  $\geq 85\%$ ; target region coverage  $\geq 98\%$ ; average sequencing depth in tissues is  $\geq 500\times$ ; average sequencing depth in blood cfDNA is  $\geq 1000\times$ . The called somatic variants need to meet the following criteria: the read depth at a position is  $\geq 100\times$ ; the variant allele frequency (VAF) is  $\geq 2\%$  for tissue DNA and  $\geq 0.2\%$  for cfDNA from blood; somatic-P  $\leq 0.01$ ; strand filter = 1. Allele frequencies were calculated for Q30 bases. For cfDNA, somatic variant calls (SNV or indel) present at least on 5 unique reads, at least 1 on each strand, and less than 0.5% mutant allelic frequency in the paired normal sample (PBLs) were retained. A manual visual inspection step was used to further remove artifactual changes by GenomeBrowse [33].

## Results

### Sample collection and patient characteristics

Of the 1746 NSCLC cases, RET rearrangements were identified in 25 cases (1.43 %) with locally advanced or metastatic NSCLC, of which 20 (80%) were female. The 25 patients have mean age of 53.5 ranging from 27 to 78. NGS testing was performed on their 25 pairs of FFPE tumor tissue and PBL samples. In addition, we performed cfDNA NGS testing on 17 of them to check their RET gene status in cfDNA sample. All the samples passed the histology quality control (HQC) and yielded sufficient DNA for NGS.

### Identification of RET rearrangements using targeted sequencing

In this study, we designed probes to cover the intron 6, 7, 8, 9, 10, 11 of RET as well as introns of some well-known RET fusion partners to identify RET rearrangement of the DNA from patients' FFPE samples. We identified RET rearrangements and analyzed the corresponding breakpoints for these patients. The statistical summary of the rearrangement events was presented in Table 1 and in Fig. 1. The breakpoint

distribution in RET was shown in Fig. 2. The results showed that 14 out of 25 patients had an KIF5B-RET fusion, with KIF5B exon15-RET exon12, KIF5B exon23-RET exon12, and KIF5B exon24-RET exon11 detected in 14, 3, and 1 patient respectively. It also detected one novel RET fusion partner PLCE1 and 4 intergenic-breakpoint fusions.

Table 1  
Fusion patterns of RET

Fusion Type	Counts	percent(%)
KIF5B-exon15-RET-exon12	14	56
CCDC6-exon1-RET-exon12	3	12
KIF5B-exon23-RET-exon12	1	4
KIF5B-exon24-RET-exon11	1	4
PLCE1-exon20-RET-exon11	1	4
RET-exon11-CCDC6-exon3	1	4
Other	4	16

### Mutational profiles of RET fusion-positive NSCLC patients

Genomic alterations were detected in 24 (n = 24/25, 96%) samples with a total of 113 alterations including nonsynonymous mutations and splicing mutations. The top 20 alterations were listed in Figure 3A. The mutation landscapes of RET fusion-positive NSCLC patients were highly heterogeneous. The median TMB was 2.4 mut/Mb with a range between 0 to 8.4 mut/Mb, which is similar to the TMB value of TCGA NSCLC cohort[35].

Besides, a heatmap was created to illustrate the somatic mutations detected in the tumor tissues of the patients (Figure 3A). TP53 was the top altered (n = 10, 42%), followed by SETD2 (n = 4, 17%), CSMD3 (n = 3, 12%), and PTEN (n = 3, 12%). Other genomic alterations with low frequencies were ATM (n = 2, 6%), CACNA1C (n = 2, 8%), CIC (n = 2, 8%), CTCF (n = 2, 8%), DOT1L (n = 2, 8%), FANCA (n = 2, 8%), FANCG (n = 2, 8%), LRP1B (n = 2, 8%), MAP2K4 (n = 2, 8%), NOTCH1 (n = 2, 8%), PRKCI (n = 2, 8%), PTPRT (n = 2, 8%), RB1 (n = 2, 8%), SMAD4 (n = 2, 8%), and SUZ12 (n = 2, 8%). Alterations in DICER1 were identified in one sample (n = 1, 4%). Moreover, the results were compared with the MSK-IMPACT study [32], from which we extracted 30 RET fusion positive cases that yielded 81 mutations. Overall, the results of our study and MSK-IMPACT research were highly consistent, both of which showed that TP53 and SETD2 were the most frequently altered genes (Fig. 3B).

Then the mutational signatures were further studied. It was observed that C>T transition happened most frequently, followed by C>G transversions (Fig. 4). This pattern is consistent with COSMIC signature 84

according to website(<https://cancer.sanger.ac.uk/signatures/sbs/sbs84/>) that had been found in some cancer samples.

Different driver gene mutations demonstrated inter-tumor heterogeneity. TP53 mutations in exon 4-8 were observed, and the TP53 mutation sites on the peptide sequence were elaborately portrayed in a lollipop plot. (Fig. 5).

### **Copy number aberrations of RET fusion-positive NSCLC patients**

Somatic copy number alterations were detected in 11 (n = 11/25, 44%) samples. A total of 22 alterations were discovered, including gain and loss (Fig. 6). CDK4 were the most commonly amplified gene (n = 3/11, 27%). Loss of copy number was observed in FGFR3 with highest frequency (n = 4/11, 36%).

### **PD-L1 expression and Microsatellite Instability (MSI) status of RET fusion-positive NSCLC patients**

High ( $\geq 50\%$ ), intermediate (1–49%), and negative ( $< 1\%$ ) PD-L1 expression was observed in 0/14 (0%), 8/14 (57%), and 6/14(43%) cases, respectively.

MSI status were evaluated in 20 cases. They were all microsatellite stable (MSS).

### **Concordance in tumor DNA(tDNA) and plasma DNA(ctDNA) sample pairs of RET fusion-positive NSCLC patients**

A total of 17 tDNA and ctDNA sample pairs were analyzed. 9 patients are found to have the same breakpoint of RET fusions in both tDNA and ctDNA samples, indicating 52.9% RET fusion could be detected in ctDNA. A total of 111 mutations (snv and indel) were identified, including 90 in tDNA and 64 in plasma ctDNA, and 43 concordant mutations in both tDNA and plasma ctDNA. Seven sample pairs (7/17, 41.2%) had concordant mutations in both tDNA and plasma ctDNA, and the average variant frequency in these plasma ctDNA samples was 8.17%.

## **Discussion**

In this study, we identified RET rearrangement events in 25 Chinese NSCLC patients by hybrid capture based NGS. Consistent with other reports, the most common RET fusion partner was KIF5B and CCDC6, and the breakpoints in the genome mainly located in the intron 11 of RET, intron 15 of KIF5B, and intron 1 of CCDC6. Amazingly, we discovered a new RET fusion partner PLCE1. Besides, we identified 4 intergenic-breakpoint fusions in 4 cases. A study by Weihua Li reported that intergenic-breakpoint fusions might also generate functional fusion transcripts [36], so additional validation testing such as RNA-seq or IHC was required for these patients to guide treatment. At the same time, we checked the concordance of RET fusions between tDNA and cfDNA for the same patient. In most cases, they harbored the same breakpoint, this clearly proved that the fragments of DNA harboring RET fusions were derived from tDNA. Therefore, cfDNA is an excellent alternative material for patients who have difficulty in obtaining tumor tissues. These results also implied that NGS-based assessment for RET fusions had the advantages of

detecting unknown RET fusion partners and identifying the same breakpoints as the traditional diagnostic testing, such as FISH and IHC.

At the same time, we characterized the co-occurring genomic alterations of these RET fusion-positive patients. The results were consistent with the TCGA cohort, in terms of the relatively higher frequency of TP53 mutations, fewer co-mutations, and lower TMB compared to RET fusion-negative NSCLC patients. Moreover, we analyzed the copy number alterations in the genome of the patients. Besides the genes with frequent copy number amplification, such as CDK4, we also discovered some genes with frequent copy number loss, such as FGFR3. This information was important for guiding optimal clinical treatment.

## **Conclusion**

In conclusion, we successfully detected the RET fusion events in 25 Chinese NSCLC patients using our customized HapOncoCDx panel. In addition, we also explored the genomic mutational landscapes of the patients. This is the first study that explored the detailed breakpoints of Chinese NSCLC patients with RET rearrangement, and we discovered a novel new partner PLCE1. The results provided genomic information for patients with RET fusion which is significant for personalized clinical management in the era of precision medicine.

## **List Of Abbreviations**

Non-small cell lung cancer (NSCLC)

Multi-kinase inhibitors (MKIs)

Formalin-fixed paraffin-embedded (FFPE)

peripheral blood lymphocyte (PBL)

Cell free DNA(cfDNA)

Microsatellite Instability (MSI)

## **Declarations**

### **Ethics approval and consent to participate**

The study was approved by the ethics committee of The Meizhou People's Hospital (Huangtang Hospital) and complied with Good Clinical Practices, the principles of the Declaration of Helsinki and all applicable regulatory requirements. All patients provided written informed consent prior to any study-specific procedures.

### **Consent to publish**

All authors agree to the publication of the article.

### **Availability of data and materials**

The datasets for this manuscript are not publicly available due to data privacy. Requests to access the datasets should be directed to the corresponding authors Guowu Wu([guowuwugd@163.com](mailto:guowuwugd@163.com)), Longhua Guo([1851258807@qq.com](mailto:1851258807@qq.com)), and Shifu Chen([chen@haplox.com](mailto:chen@haplox.com)).

### **Competing interests**

TH, ML, BY, WL, XZ, YL, YY and SC were employed by the company HaploX Biotechnology Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### **Funding**

This work was supported by the project of Medical Science and Technology Research Foundation of Guangdong Province (Grant No. B2021118), the project of Shenzhen Science and Technology Innovation Committee Technical Research Project (Grant No. KQJSCX20180330124428928 and JSGG20180508152646606), the project of Development and Reform Commission of Shenzhen Municipality (Grant No. XMHT20190104006 and XMHT20200104013).

### **Authors' Contributions**

GW and LG contributed equally to this study. GW, LG, and SC designed the study. YG, XZ, PH, CW, LY, DZ, JZ, XZ, YL carried out the sequencing experiment and collected data. TH, ML, WL, and BY performed the bioinformatics analysis. GW, LG, and TH wrote this manuscript. YY and SC revised this manuscript. GW, LG and SC supervised the study.

### **Acknowledgements**

None.

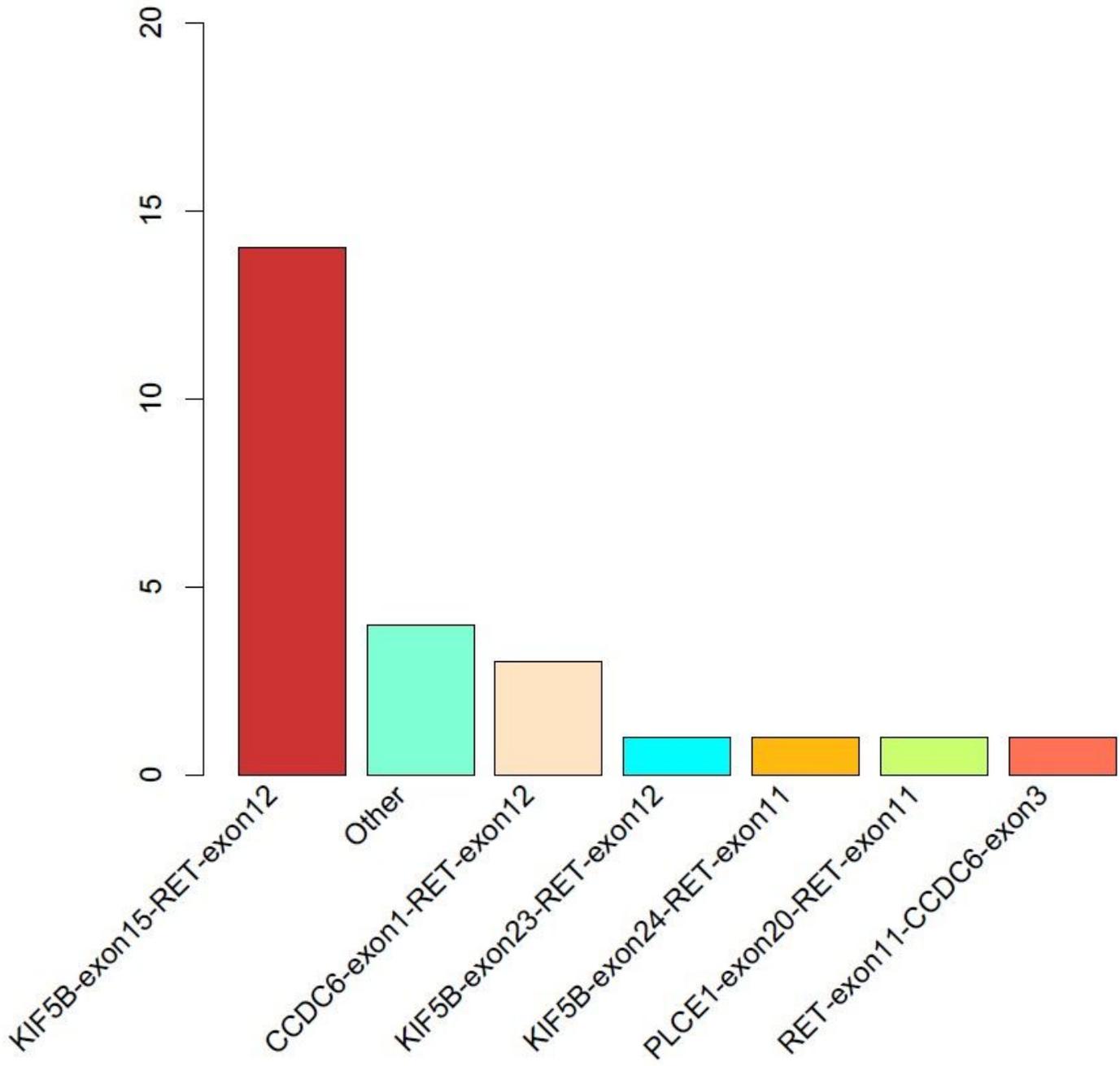
## **References**

1. Ju YS, Lee WC, Shin JY, et al. A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Res.* 2012;22(3):436–445.
2. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* 2014;511(7511):543–550.
3. Wang R, Hu H, Pan Y, et al. RET fusions define a unique molecular and clinicopathologic subtype of non-small-cell lung cancer. *J Clin Oncol* 2012;30:4352-9.
4. Worby CA, Vega QC, Zhao Y, et al. Glial cell line-derived neurotrophic factor signals through the RET receptor and activates mitogen-activated protein kinase. *J Biol Chem*1996;271:23619-22.

5. Qian Y, Chai S, Liang Z, et al. KIF5B-RET fusion kinase promotes cell growth by multilevel activation of STAT3 in lung cancer. *Mol Cancer* 2014;13:176.
6. Trupp M, Scott R, Whittemore SR, et al. Ret-dependent and -independent mechanisms of glial cell line-derived neurotrophic factor signaling in neuronal cells. *J Biol Chem* 1999;274:20885-94.
7. Kohno T, Ichikawa H, Totoki Y, et al. KIF5B-RET fusions in lung adenocarcinoma. *Nat Med* 2012;18:375-7.
8. Lipson D, Capelletti M, Yelensky R, et al. Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. *Nat Med* 2012;18:382-4.
9. Ju YS, Lee WC, Shin JY, et al. A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Res* 2012;22:436-45.
10. Chao BH, Briesewitz R, Villalona-Calero MA. RET fusion genes in non-small-cell lung cancer. *J Clin Oncol*. 2012;30(35):4439–4441.
11. Ferrara R, Auger N, Auclin E, Besse B. Clinical and translational implications of RET rearrangements in non-small cell lung cancer. *J Thorac Oncol*. 2018;13(1):27–45.
12. Choudhury NJ, Drilon A. Decade in review: a new era for RET-rearranged lung cancers. *Transl Lung Cancer Res* 2020;9(6):2571-2580. doi: 10.21037/tlcr-20-346
13. Drilon A, Hu ZI, Lai GGY, et al. Targeting RET-driven cancers: lessons from evolving preclinical and clinical landscapes. *Nat Rev Clin Oncol* 2018;15:150.
14. Gautschi O, Milia J, Filleron T, et al. Targeting RET in Patients With RET-Rearranged Lung Cancers: Results From the Global, Multicenter RET Registry. *J Clin Oncol* 2017;35:1403-10.
15. Drilon A, Rekhtman N, Arcila M, et al. Cabozantinib in patients with advanced RET-rearranged non-small-cell lung cancer: an open-label, single-centre, phase 2, single-arm trial. *Lancet Oncol* 2016;17:1653-60.
16. Yoh K, Seto T, Satouchi M, et al. Vandetanib in patients with previously treated RET-rearranged advanced non-small-cell lung cancer (LURET): an open-label, multicentre phase 2 trial. *Lancet Respir Med* 2017;5:42-50.
17. Drilon A, Fu S, Patel MR, et al. A Phase I/Ib Trial of the VEGFR-Sparing Multikinase RET Inhibitor RXDX-105. *Cancer Discov* 2019;9:384-95.
18. Ribeiro M, Alessi JVM, Oliveira LJC, et al. Alectinib activity in chemotherapy-refractory metastatic RET-rearranged non-small cell lung carcinomas: A case series. *Lung Cancer* 2020;139:9-12.
19. Nakaoku T, Kohno T, Araki M, et al. A secondary RET mutation in the activation loop conferring resistance to vandetanib. *Nat Commun* 2018;9:625.
20. Nelson-Taylor SK, Le AT, Yoo M, et al. Resistance to RET-Inhibition in RET-Rearranged NSCLC Is Mediated By Reactivation of RAS/MAPK Signaling. *Mol Cancer Ther* 2017;16:1623-33.
21. Markham A. Selpercatinib: first approval[J]. *Drugs*, 2020, 80(11): 1119-1124. Materials and methods.
22. Velcheti V, Bauer T, Subbiah V, et al. OA 12.07 LOXO-292, a Potent, Highly Selective RET Inhibitor, in MKI Resistant RET Fusion-Positive Lung Cancer Patients with and without Brain Metastases. *J*

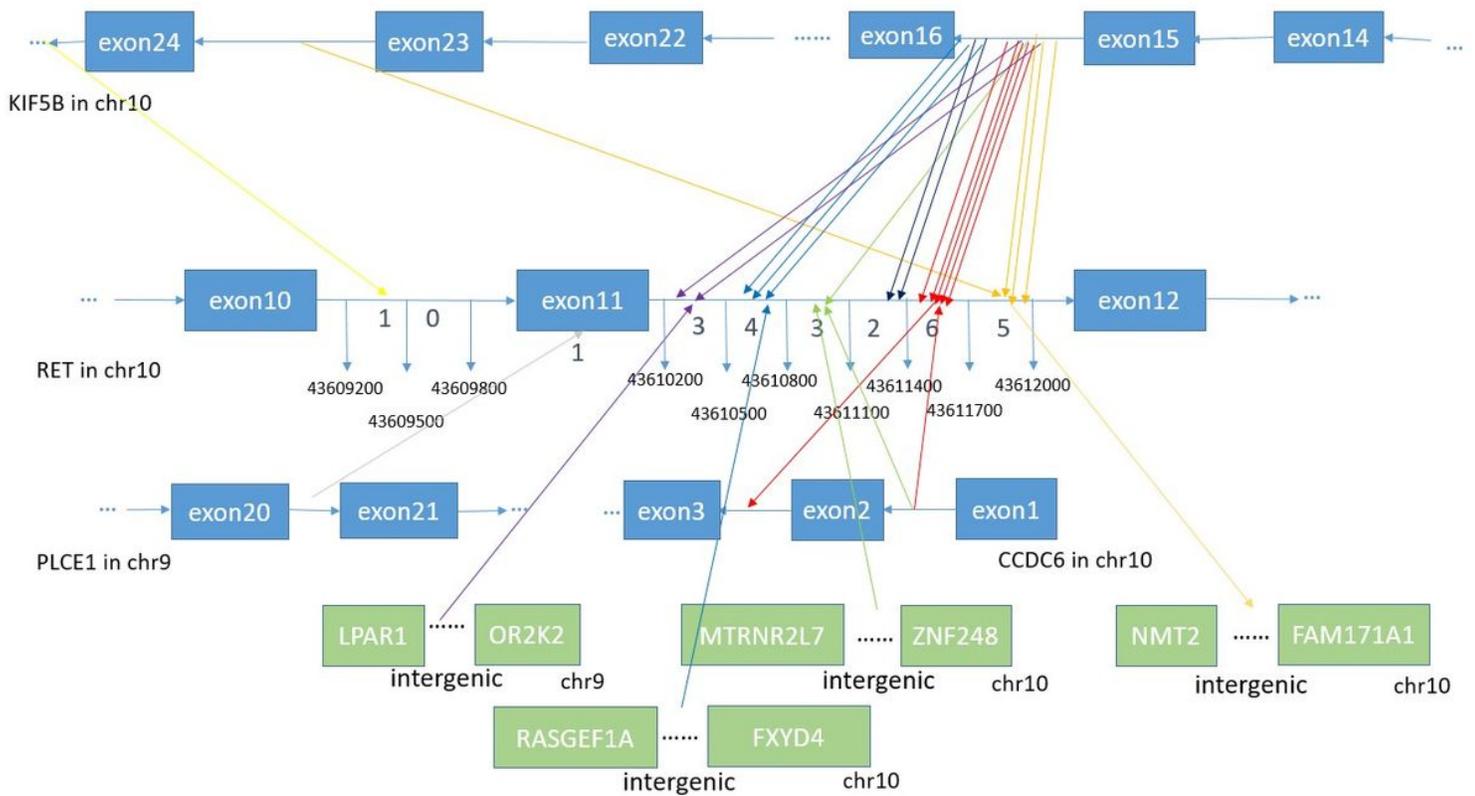
- Thorac Oncol 2017;12:S1778.
23. Subbiah V, Gainor JF, Rahal R, et al. Precision Targeted Therapy with BLU-667 for RET-Driven Cancers. *Cancer Discov* 2018;8:836-49.
  24. Shifu Chen, Yanqing Zhou, Yaru Chen, Jia Gu; fastp: an ultra-fast all-in-one FASTQ preprocessor, *Bioinformatics*, Volume 34, Issue 17, 1 September 2018, Pages i884–i890, <https://doi.org/10.1093/bioinformatics/bty560>
  25. Li H and Durbin R: Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26: 589-595, 2010.
  26. gencore: an efficient tool to generate consensus reads for error suppressing and duplicate removing of NGS data. Shifu Chen, Yanqing Zhou, Yaru Chen, Tanxiao Huang, Wenting Liao, Yun Xu, Zhihua Liu, Jia Gu *bioRxiv* 501502; doi: <https://doi.org/10.1101/501502>.
  27. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G and Durbin R; 1000 Genome Project Data Processing Subgroup: The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25 2078-2079, 2009.
  28. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller C, Mardis ER, Ding L and Wilson RK: VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 22: 568-576, 2012.
  29. Talevich, E., Shain, A.H., Botton, T., & Bastian, B.C. (2014). CNVkit: Genome-wide copy number detection and visualization from targeted sequencing. *PLOS Computational Biology* 12(4):e1004873.
  30. Shifu Chen, Ming Liu, Tanxiao Huang, Wenting Liao, Mingyan Xu and Jia Gu. GeneFuse: detection and visualization of target gene fusions from DNA sequencing data. *International Journal of Biological Sciences*, 2018; 14(8): 843-848. doi: 10.7150/ijbs.24626.
  31. Chen, S., Qu, H., Yang, B., Huang, T., Zhang, X., Liu, Y., ... & Gu, J. (2019). Detect and visualize tumor microsatellite instability status from next-generation sequencing data by simulating PCR techniques. *Journal of Clinical Oncology* 37, no. 15\_suppl.e13052.
  32. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. 2018. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Research* PMID: 30341162.
  33. Golden Helix GenomeBrowse® visualization tool (Version 2.x) [Software]. Bozeman, MT: Golden Helix, Inc. Available from <http://www.goldenhelix.com>.
  34. Zehir A, Benayed R, et al: Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med*. 2017 Jun;23(6):703-713. doi: 10.1038/nm.4333.
  35. Chalmers, Z.R., Connelly, C.F., Fabrizio, D. et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med* 9, 34 (2017). <https://doi.org/10.1186/s13073-017-0424-2>.
  36. Li W, Liu Y, Li W, et al. Intergenic breakpoints identified by DNA sequencing confound targetable kinase fusion detection in NSCLC[J]. *Journal of Thoracic Oncology*, 2020, 15(7): 1223-1231.

# Figures



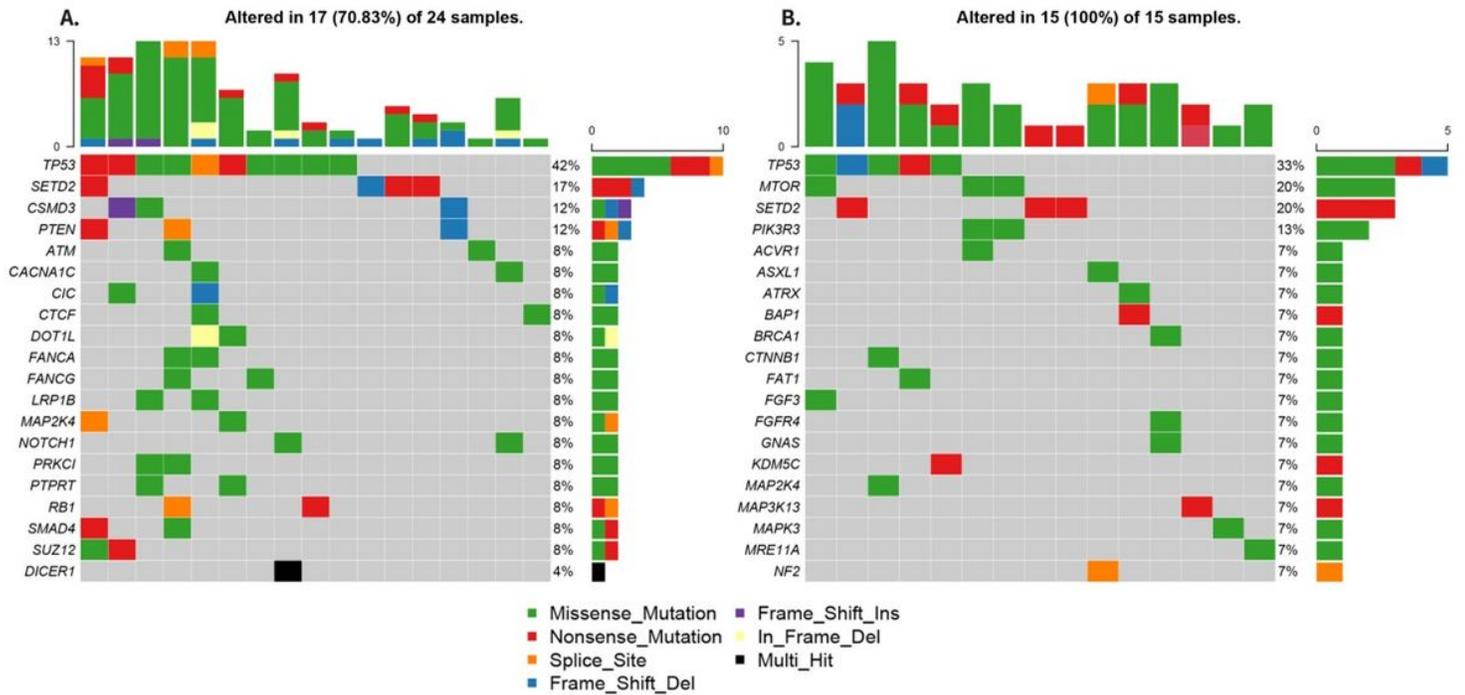
**Figure 1**

The statistics of different RET rearrangement forms. The distribution of each RET fusion pattern identified in 25 NSCLC patients are shown in the barchart.



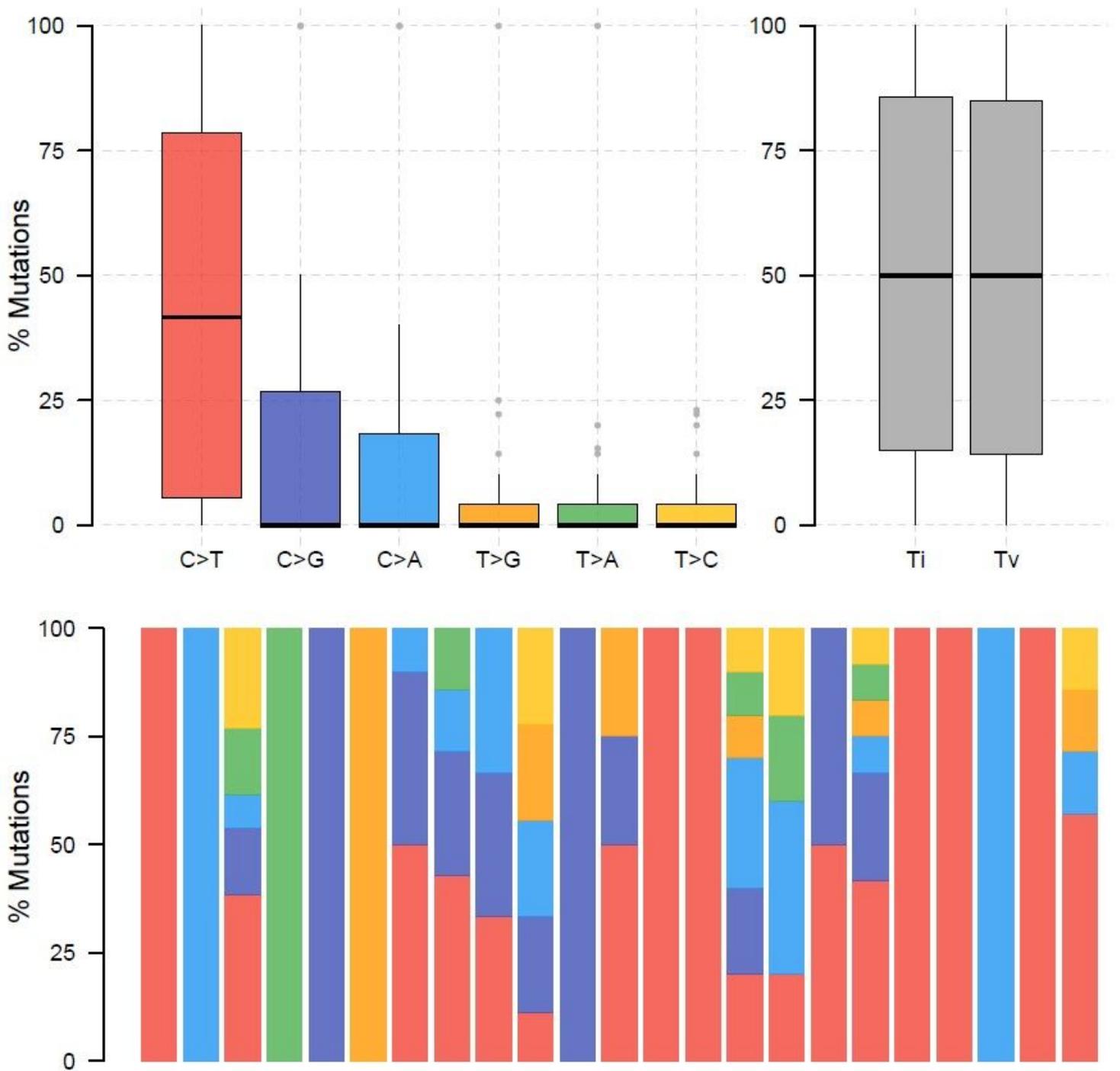
**Figure 2**

The breakpoint distribution in RET and the corresponding fusion partners. Each fusion event was represented with an arrowed line. The breakpoints of RET in GRch37 was shown in the middle panel with the fused exons of RET fusion partners on the top or bottom panels. The sequences of KIF5B and CCDC6 are exhibited reversely (from right to left), while the sequences of RET and PLCE1 are represented in the forward direction. The genomic region of RET between 43609200 and 43609800 in intron 10 as well as region between 43610200 and 43612000 in intron 11 was divided into regions every 300bp. Breakpoint positions in RET locate between 43609200 and 43609500 with a yellow arrow, between 43610200 and 43610500 with a purple arrow, between 43610500 and 43610800 with a blue arrow, between 43610800 and 43611100 with an green arrow, between 43611100 and 43611400 with a black arrow, between 43611400 and 43611700 with a red arrow, between 4361700 and 43612000 with an orange arrow,.



**Figure 3**

Mutational profiles of RET fusion-positive NSCLC patients. A. This is the oncoprint for the top 20 genes of the somatic SNVs and Indels of the 25 patients in our study. Somatic alterations included missense, nonsense, frameshift indel, in-frame indel, splice site, translation start site, and multi\_Hit mutations. The genes are ranked by the frequency of mutations across all samples. B. This is the oncoprint for the top 20 genes of the somatic SNVs and Indels of the 15 patients from the MSK-IMPACT study[34]. Somatic alterations included missense, nonsense, frameshift indel, in-frame indel, and splice-site mutations. The genes are ranked by the frequency of mutations across all samples.



**Figure 4**

Mutational signatures of RET fusion positive NSCLC patients. SNPs are summarized into transitions and transversions. Statistical data were visualized as a boxplot presenting overall distribution of six kinds of different conversions (top) as well as a stacked barplot presenting the fraction of different kinds of conversions in each sample (bottom).

TP53 : [Somatic Mutation Rate: 41.67%]  
NM\_000546

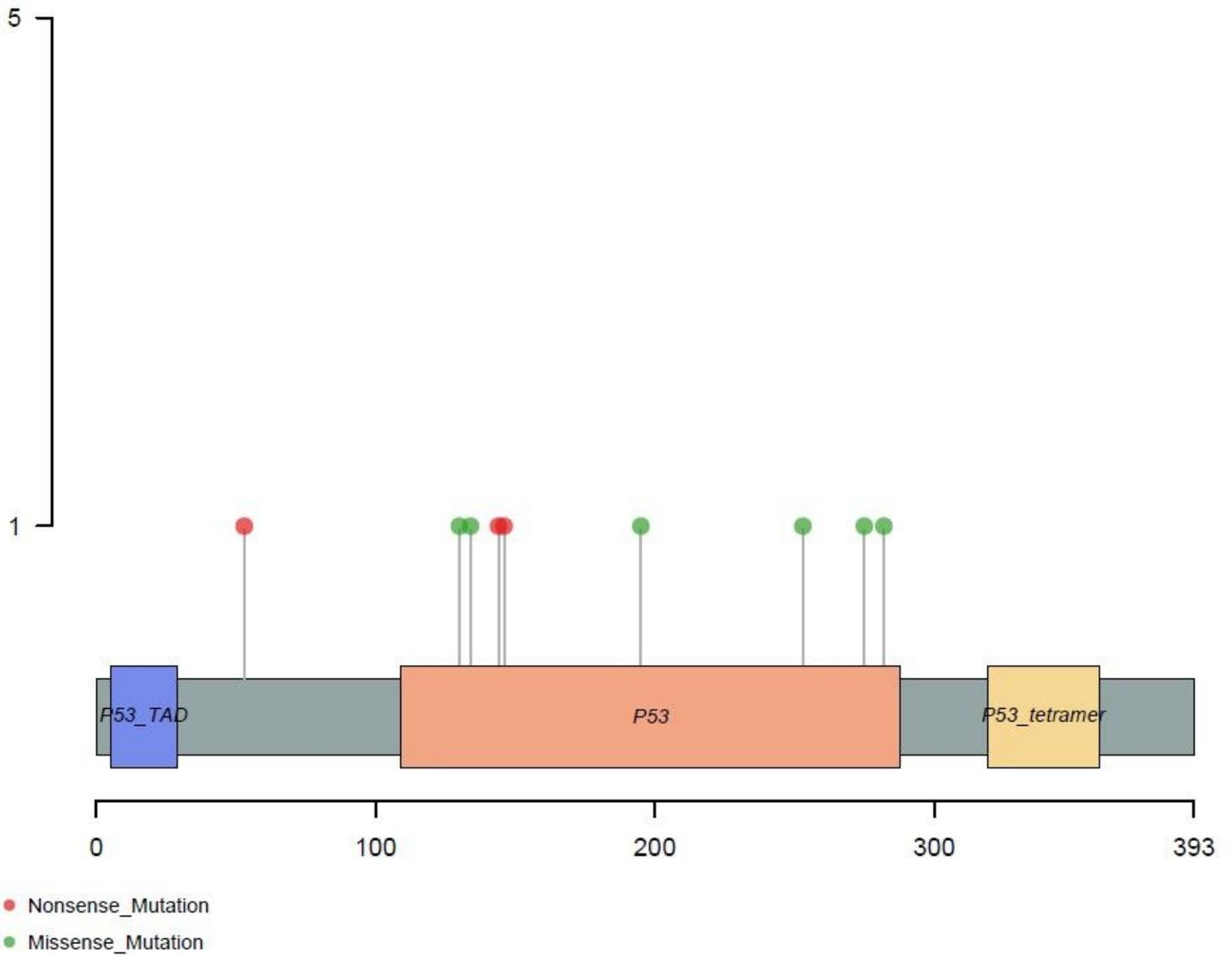
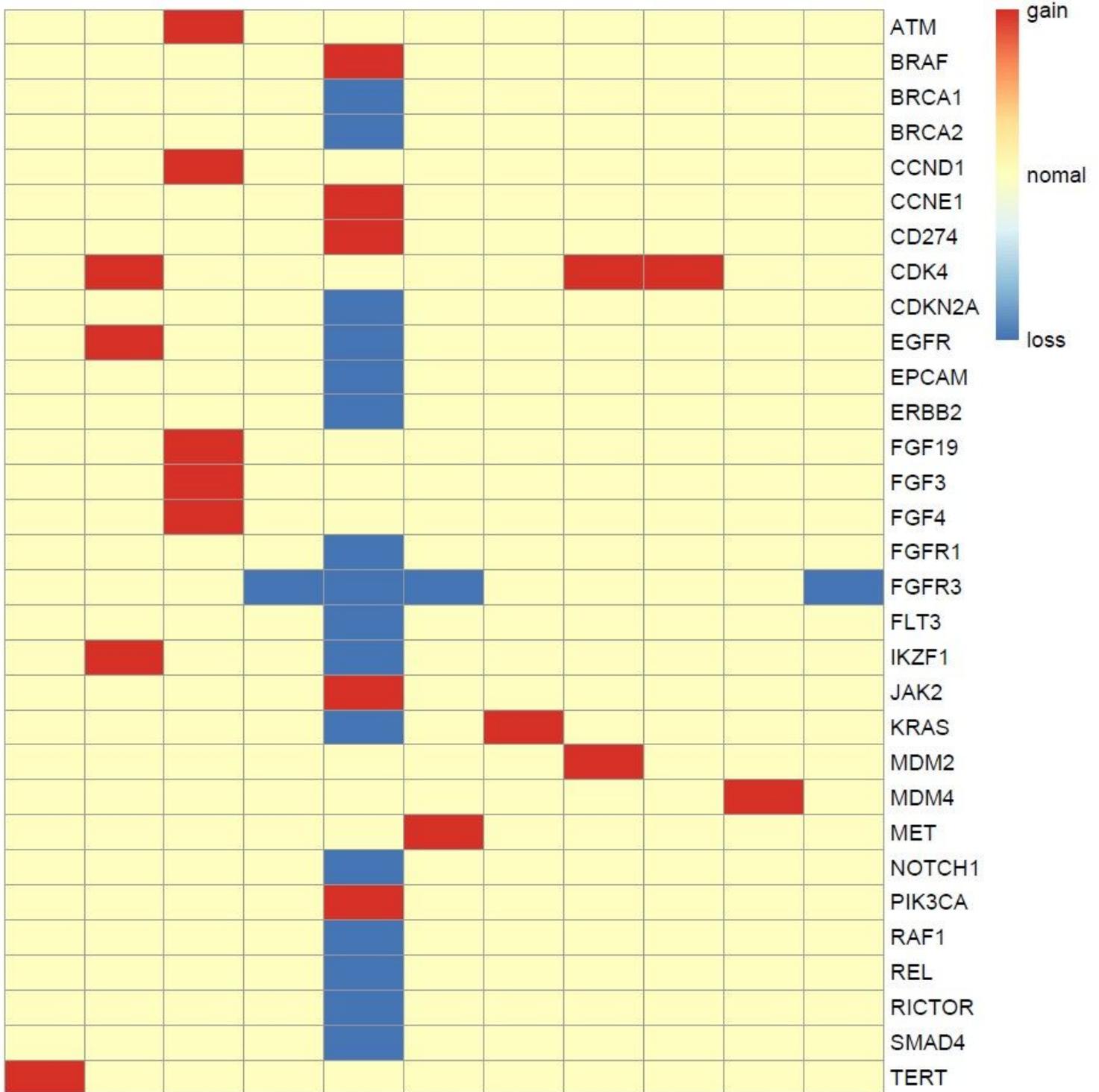


Figure 5

Protein variants resulted from TP53 mutations. The lollipop plot showed the protein variants caused by TP53 mutations, which were considered to be mutational hot-spots.



**Figure 6**

Copy number aberrations in 11 RET fusion-positive NSCLC patients. The y-axis represented the names of the aberrant genes, while x-axis represents individual patient. The types of copy number aberrations, including gain, normal and loss were indicated by red, yellow, and blue respectively.