



OPEN ACCESS

EDITED BY

Hongtao Xiao,
University of Electronic Science and
Technology of China, China

REVIEWED BY

Kai Qin,
Huazhong University of Science and
Technology, China
Imran Tipu,
University of Management and
Technology, Pakistan

*CORRESPONDENCE

Xinwen Zhong,
xwzhong@cmu.edu.cn
Tian Li,
fmmult@foxmail.com

[†]These authors share first authorship

SPECIALTY SECTION

This article was submitted to
Pharmacology of Anti-Cancer Drugs,
a section of the journal
Frontiers in Pharmacology

RECEIVED 16 May 2022

ACCEPTED 28 July 2022

PUBLISHED 26 August 2022

CITATION

Yao Y, Guan X, Bao G, Liang J, Li T and
Zhong X (2022), Whole-exome
sequencing and bioinformatics analysis
of a case of non-alpha-fetoprotein-
elevated lung
hepatoid adenocarcinoma.
Front. Pharmacol. 13:945038.
doi: 10.3389/fphar.2022.945038

COPYRIGHT

© 2022 Yao, Guan, Bao, Liang, Li and
Zhong. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Whole-exome sequencing and bioinformatics analysis of a case of non-alpha-fetoprotein-elevated lung hepatoid adenocarcinoma

Yao Yao^{1†}, Xiaojiao Guan^{2†}, Guangyao Bao¹, Jie Liang¹, Tian Li^{3*} and Xinwen Zhong^{1*}

¹Department of Thoracic Surgery, First Affiliated Hospital, China Medical University, Shenyang, China,

²Department of Pathology, Second Affiliated Hospital, China Medical University, Shenyang, China,

³School of Basic Medicine, Fourth Military Medical University, Xi'an, China

Hepatoid adenocarcinoma of the lung (HAL) is an exceptionally rare malignant tumor with prominent hepatocellular carcinoma (HCC)-like characteristics in organs or tissues outside the liver, while there is no tumor in the liver. Most HAL cases have various degrees of serum alpha-fetoprotein (AFP) levels and exhibit a similar origin and clonal evolution process to HCC. We studied a case of HAL without elevating the AFP level by performing whole-exome sequencing (WES) and bioinformatics analyses after surgical resection. Our results showed mutations in two driver genes, NLRP3 and PBX1, and we identified HNRNPR, TP73, CFAP57, COL11A1, RUSC1, SLC6A9, DISC1, NBPF26, and OR10K1 as potential driver mutation genes in HAL. In addition, 76 significantly mutated genes (SMG) were identified after the statistical test of each mutation type on genes.

KEYWORDS

hepatoid adenocarcinoma of the lung, whole-exome sequencing, alpha-fetoprotein, lung cancer, clinical research

Introduction

Hepatoid adenocarcinoma (HAC), an exceptionally rare malignancy with prominent HCC-like characteristics in extra-hepatic organs or tissues, has been discovered in various organs, including the ovary (Mazouz et al., 2015), lung (Lin et al., 2013; Su et al., 2013; Shi et al., 2019), uterus (Kato et al., 2007), pancreas (Gardiner et al., 1992), and other sites. HAL occurs in approximately 5% of HAC cases (Haninger et al., 2014), and HAL was first reported by Ishikura et al. (1990). Although this rare disease has been identified for more than 30 years, the origin of HAC is still controversial to this date. During embryonic development, some organs like the liver, lung, stomach, and esophagus are diversified from the primitive foregut. Several abnormal differentiation processes may contribute to the development of HAC in these regions (Nakashima et al., 1987).

HAL cases are extremely rare, with less than 60 cases reported in the literature in the last 40 years (Hou et al., 2021). Usually, HAL patients exhibit chest and back pain, shoulder shooting pain, cough and breathlessness, shortness of breath, nausea, and weight loss (Sun et al., 2016). These symptoms are nonspecific and basically the same as those of other types of lung cancer, thus making accurate and timely diagnosis difficult. The main antidiastole for HAL is HCC metastases and non-small cell lung cancer. Collecting data on the patient's medical history and clinical presentation is critical to the diagnosis as HCC often occurs in cirrhosis associated with hepatitis virus infection and alcohol abuse. Meanwhile, it is insufficient to use computed tomography (CT) images alone for diagnosis; cytomorphological and immunohistochemical techniques are necessary to confirm the diagnosis (Terracciano et al., 2003). According to the literature, most HAL tumors occur mainly in the upper lobes, usually near the large blood vessels in the chest wall or mediastinal pleura, while a small proportion of HAL tumors occur in the other lobes. The incidence of HAL is much higher in middle-aged and older male smokers than in women. Because of the nonspecific symptoms and difficulty in diagnosis, most patients are diagnosed with HAL after the disease has progressed to an intermediate or advanced stage (Zhuansun et al., 2021).

A significant increase of alpha-fetoprotein (AFP) in tissues or serum is common in most HAL cases (Ishikura et al., 1990; Grossman et al., 2016). By comparison, HAL without elevated AFP levels is more challenging to diagnose and treat owing to the scarcity of clinical cases and lack of evidence-based medicine (Sun et al., 2016).

The prognosis of HAL is worse than that of other types of lung tumors; the 5-year survival rate is only 8.0%, and the 2-year survival rate is 35.3% (Hou et al., 2021). At present, the understanding of HAL is still insufficient, and the etiology and underlying mechanism of such patients are still unclear; so far, there is no unified diagnostic standard, treatment, or prevention strategy to guide clinical practice. Surgical resection, chemotherapy, and/or radiotherapy are the most common treatments for this disease (Qian et al., 2016), and recently, molecular targeted therapy has shown great promise in treating HAL (Khozin et al., 2012). Therefore, this project provides a new basis for molecular diagnosis and targeted therapy of HAL through whole-exome sequencing (WES) and bioinformatics analysis.

Materials and methods

Sample

One patient, who was diagnosed with HAL in the First Affiliated Hospital of China Medical University, was enrolled

in the study. Tumor and paracancerous tissues were rapidly frozen during surgery and saved in liquid nitrogen.

DNA extraction and sequencing

DNA fragmentation was performed by NEBNext dsDNA Fragmentase (NEB, Ipswich, MA, United States), and the DNA fragments were then subjected to end-repair. In the next step, these DNA fragments were hydrolyzed and mixed with RNA library decoys and streptavidin-coated magnetic beads, followed by bead capturing, washing the beads, and RNA digestion, and finally, these DNA fragments were amplified to 150 bp. The cDNA/DNA/small RNA libraries were sequenced on the Illumina sequencing platform X Ten system (Illumina, San Diego, CA, United States) by Gene Denovo Biotechnology Co., Ltd. (Guangzhou, China).

Quality control

Quality trimming is a very important step to guarantee high confidence in variant calling. Fastp (Chen et al., 2018) was used to complete the quality trimming. The filtering criteria were: 1) unidentified nucleotides (N) in the reads >10%; 2) the number of bases >50% of the reads with a phred quality score ≤ 20 ; 3) reads aligned to the barcode adapter. If the reads met one of these three criteria, the reads would be cleared.

Variant identification and annotation

For mapping the valid sequencing reads to the reference human genome (GRCh38) and to identify SNVs and indels, the Burrows–Wheeler Aligner (BWA) (Li and Durbin, 2009) was used. Duplicate reads were marked and realigned using the Picard suite. The Genome Analysis Toolkit (GATK) (McKenna et al., 2010) Unified Genotyper was used to find variant sites on the genome, including SNV and small fragments of indels, and the base quality score recalibration. The software tool, ANNOVAR (Wang et al., 2010), was used to align and annotate SNVs or indels to the following databases: 1000 Genomes Project, HAMAP, ESP6500, dbSNP, and Kaviar.

Mutation deleteriousness prediction

SIFT (Ng and Henikoff, 2003), PolyPhen-2 (Adzhubei et al., 2010), MutationTaster (Schwarz et al., 2010), and PROVEAN (Choi et al., 2012) were used to predict whether SNVs and indels lead to changes in protein structure and function.

Driver gene identification

Driver mutation genes were identified by MuSiC (Dees et al., 2012) and OncodriveCLUST (Tamborero et al., 2013a). Comprehensively considering the preference of driver mutations in the locus distribution to form mutation clusters, we used synonymous mutations to construct a background mutation rate model with the characteristics of an unbiased distribution to find possible driver mutations by OncodriveCLUST software. MuSiC software was used to find genes with higher mutation frequency in tumor samples than control samples, and convolution tests were performed for each mutation type. The cutoff of MuSiC is FDR, 0.01 and that of OncodriveCLUST is FDR, 0.05.

Gene functional enrichment analyses

We used the clusterProfiler package on R- 4.1.2 to map the significantly mutated genes to each term of the GO database (<http://www.geneontology.org/>) and each pathway of the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and calculated the number of genes in each term and pathway, so as to obtain a list of genes with a certain GO function and pathway of KEGG. A hypergeometric test was then applied to find GO terms and KEGG pathways that were significantly enriched in genes compared to the whole-genome background. Modified Fisher's exact test was applied, and $p < 0.05$ was considered a statistically significant difference.

Structural variation identification and analyses

SV types include translocations, inversions, and insertion events, and SVs were determined using CREST (1.0) software (Wang et al., 2011). Copy number variants (CNVs) were identified using VarScan2 (Kadalayil et al., 2015) with the following parameters: Phred base quality ≥ 20 and minimum coverage ≥ 20 . Recurrent somatic CNA identification was conducted by GISTIC (Mermel et al., 2011), and CNVs were classified using Control-FREEC (10.4) (Boeva et al., 2012).

Results

Case presentation

The 63-year-old male patient was admitted to the First Affiliated Hospital of China Medical University on 5 September 2018, with intermittent dyspnea and occasional chest pain. The chest-CT scan showed a 7.5*5.5 cm² mass in the left upper lung with local pleural wrinkling and slightly enlarged

lymph nodes in the mediastinum (Figure 1). Before the operation, no metastasis was found in brain CT or bone nuclide scan, while serum carcinoma embryonic antigen (CEA) and neuron-specific enolase (NSE) were slightly elevated, and AFP was in the normal range. On 19 September 2018, the patient underwent left upper lobe resection and mediastinal lymph node dissection. Referring to the preoperative abdominal B-ultrasound examination and postoperative abdominal B-ultrasound examination, 3 months after the surgery, there was no obvious abnormality in the liver, and the patient's status was average with no history of hepatitis, alcohol abuse, exposure to radiation, or toxins. Meanwhile, he had no symptoms of liver cirrhosis such as spider-burst, liver palms, or hepatjugular reflex. We measured the patient's plasma AFP levels 4 days before surgery, 1 day after surgery, and 3 months after surgery. The results of AFP were 11.02 ng/ml, 13.11 ng/ml, and 9.21 ng/ml, respectively. Given the patient's symptoms, personal history, physical examination, preoperative examination, and immunohistochemical staining results of postoperative pathology, the patient was diagnosed with HAL. Meanwhile, there was no metastatic lesion in the mediastinal lymph nodes or adjacent tissue or organs; thus, the postoperative staging is T4N0M0. The patient recovered well after surgery and was discharged from the hospital one week later. Subsequently, a 300 gene-panel genomic testing was conducted, which contained targeted drug-related genes and chemotherapy drug-related genes. Nevertheless, no targeted drug-matched gene mutation was found in the testing. The patient did not receive radiotherapy or chemotherapy after surgery and died of an accidental cerebral hemorrhage six months later.

Histological and immunohistochemical analysis

Hematoxylin-eosin (HE) staining of tumor tissue suggested that the typical cell clusters were flaky, trabecular, and densely arranged; the nuclei were large and deeply stained, and the nucleoplasm ratio was imbalanced. Additionally, alveolar expansion with fracture and fusion was found in the lung tissue (Figure 2). Immunohistochemistry staining demonstrated hepatocyte (+), glypican-3 (+), cytokeratin (CK) (+), Ki-67 (30% +), P40 (-), P63 (-), thyroid transcription factor-1 (TTF-1) (-), and napsin-A (-) (Figure 3).

Whole-exome sequencing and analysis of somatic mutations

To investigate the mutation in this case of HAL, we performed whole-exome sequencing of the tumor tissue. After sequencing and clean read filtering, 122 somatic SNVs and 3 somatic indels were detected using MuTect (Cibulskis et al.,

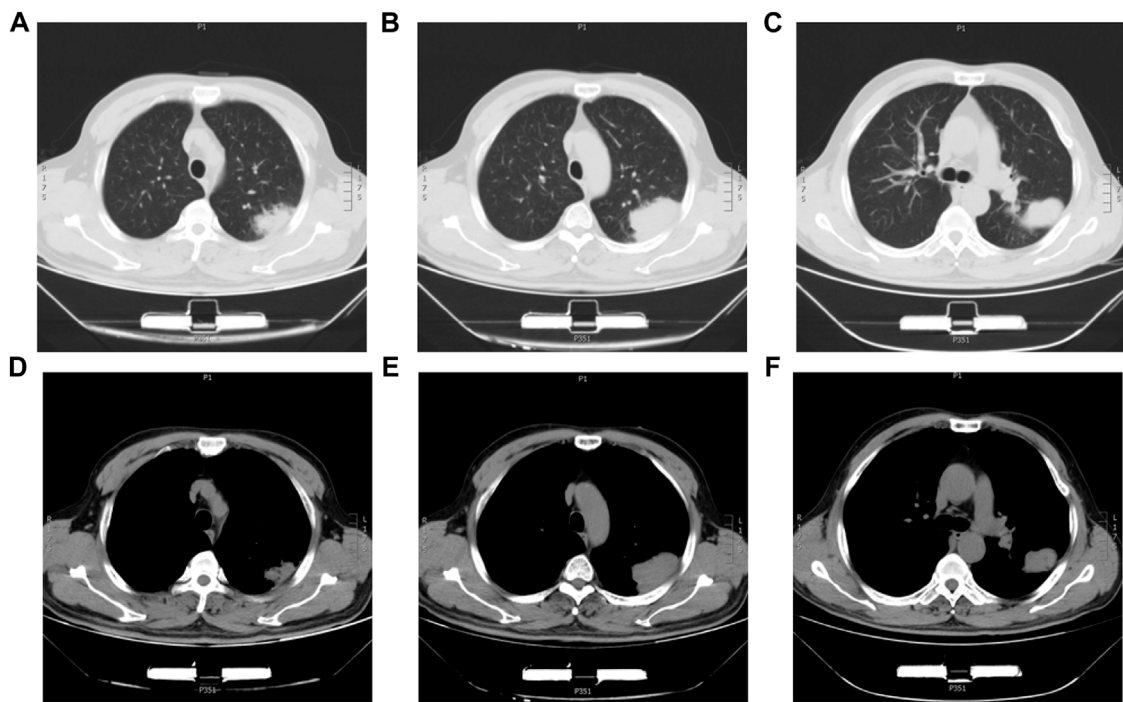


FIGURE 1

Chest CT images of the HAL patient. (A,B,C) CT images of the lung window with a lung mass in the left upper lobe. (D,E,F) CT images of the mediastinum window with a lung mass in the left upper lobe.

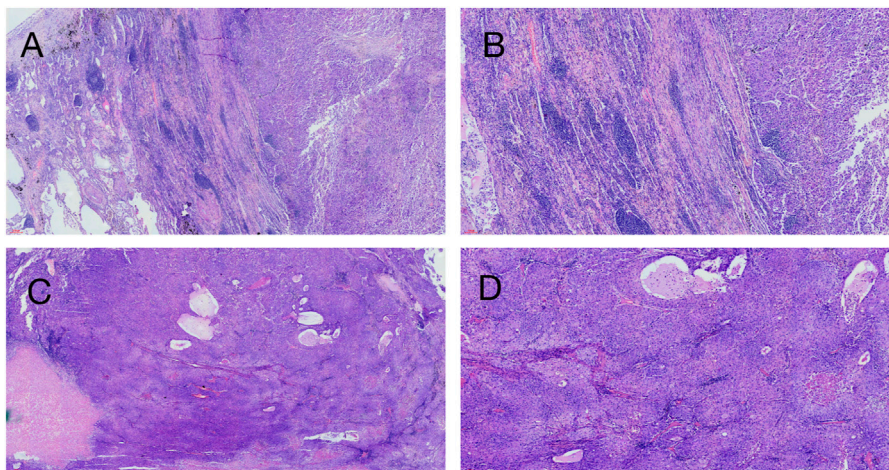


FIGURE 2

Hematoxylin-eosin staining on patient's tumor tissue. (A) ($\times 40$, $200 \mu\text{m}$), (B) ($\times 100$, $100 \mu\text{m}$) Left side of the picture is the lung tissue, and alveolar fusion and inflammatory cell infiltration can be seen; the right side is a large tumor tissue, and fibrous tissue proliferation and inflammatory cell infiltration can be seen around the tumor tissue. (C) ($\times 40$, $300 \mu\text{m}$), (D) ($\times 100$, $100 \mu\text{m}$) Tumor tissue is distributed in the form of beams and nests, with abundant cytoplasm, similar to hepatocellular carcinoma.

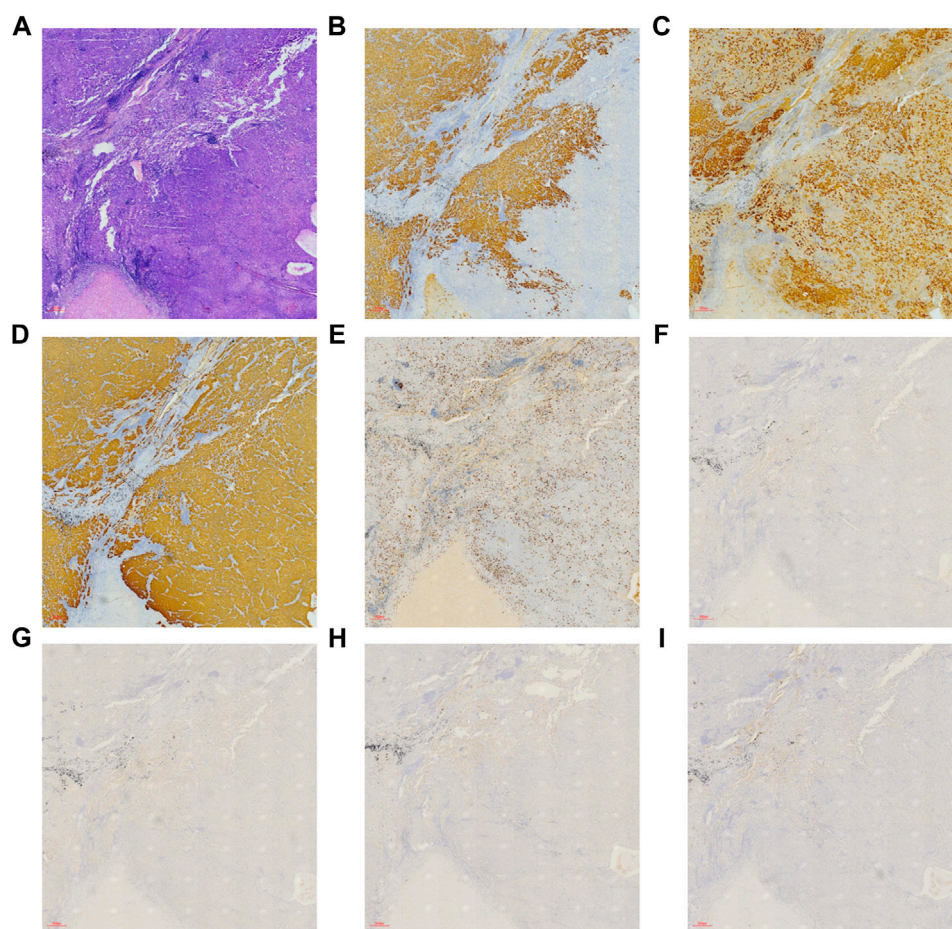


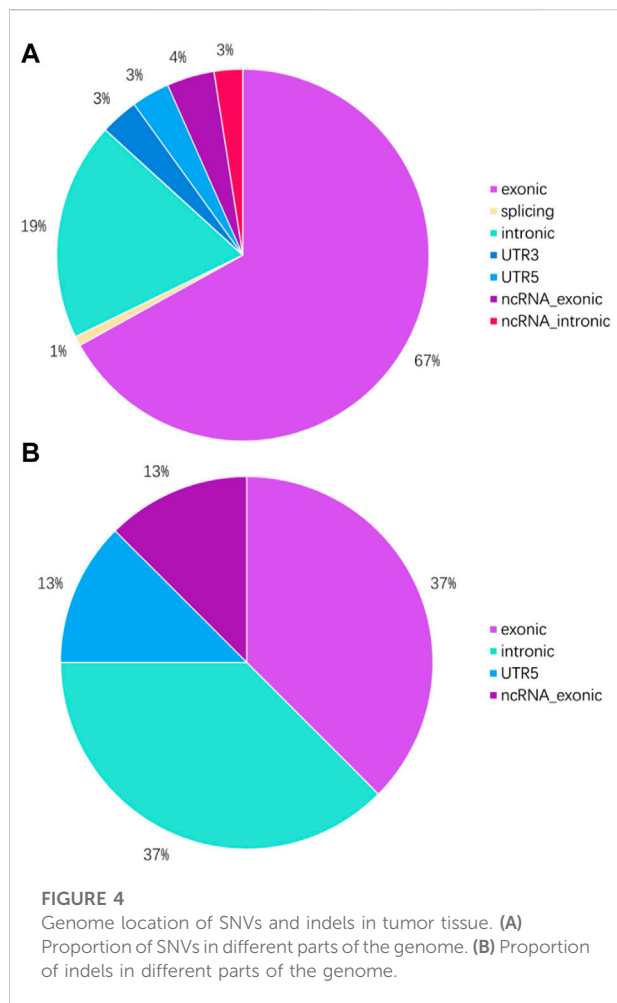
FIGURE 3

Immunohistochemical and hematoxylin-eosin staining on patient's tumor tissue ($\times 40$, $300 \mu\text{m}$). (A) HE staining, (B) hepatocyte (+), (C) glypican-3 (+), (D) CK (+), (E) Ki-67 (30%+), (F) P40 (-), (G) P63 (-), (H) TTF-1 (-), and (I) napsin-A (-).

2013), most of which were situated in the exonic parts of the genome (Figure 4). Also, we used VarScan2 (28) to filtrate the somatic CNVs. Correlation analysis revealed that most somatic mutations clustered in chromosome 1 (Figure 5). Among the somatic mutations detected in the tumor samples, two alternations were found in the coding region of NLRP3, including one synonymous mutation (c.1299C>A) and one nonsynonymous mutation (c.1300 C>A). Also, one nonsynonymous mutation (c.731C>T) was detected in the exonic region of PBX1, after searching reported driver gene databases, including Cancer Gene Census, MDG125 (Vogelstein et al., 2013), SMG127 (Kandoth et al., 2013), and CDG291 (Tamborero et al., 2013b). The synonymous mutation (c.1299C>A) in NLRP3 was archived in the COSMIC database from a lung adenocarcinoma tumor sample (TCGA-50-5941-01), and the nonsynonymous mutations of neither NLRP3 nor PBX1 were archived in dbSNP, 1,000 Genomes Project, ExAC database, and COSMIC. So, we annotated the mutations with

MutationTaster, SIFT, PolyPhen2, and PROVEAN to predict whether these mutations are detrimental to protein function. The PolyPhen2 predicted the variant (c.1300C>A) to be benign with a score of 0.016 and the score of the variant (c.731C>T) was 0.997, so it was determined to be probably damaging. Meanwhile, PROVEAN predicted the variant (c.1300C>A) to be neutral with a score of -1.75 and the variant (c.731C>T) to be deleterious with a score of -3.69 . Although the NLRP3 variants (c.1300C>A and c.1299C>A) and the PBX1 variant (c.731C>T) have been identified, there are no recommended therapies for diseases caused by these genetic mutations at present.

To further verify those cancer-related mutations, we used MuSiC(25) to seek significantly mutated genes in cancer samples, and simultaneously, all the somatic alternations were statistically tested by the convolution test (Figure 6). GO and KEGG interaction analyses were conducted among the significantly mutated genes, which are shown in Figure 7. The results



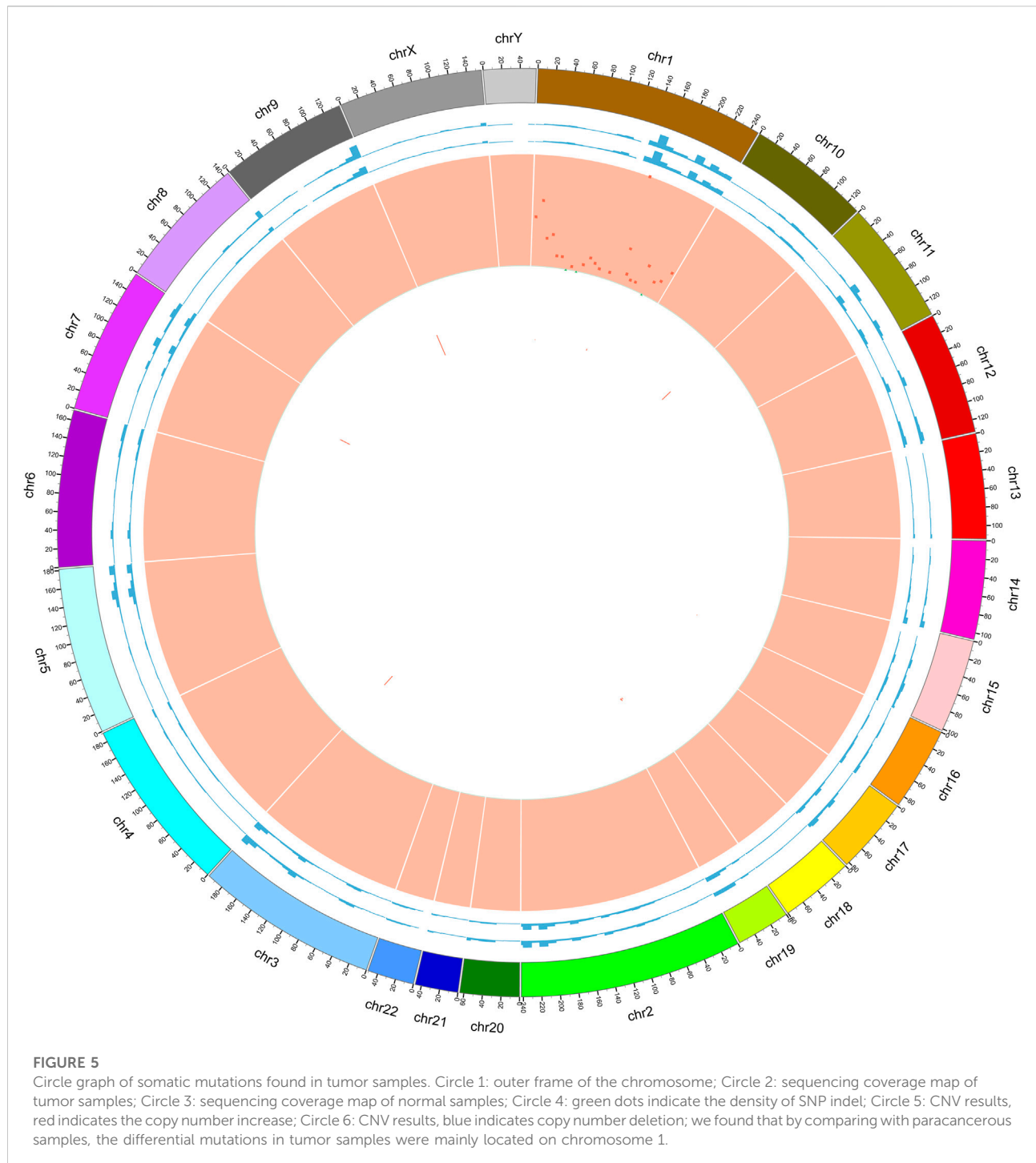
showed that, for molecular function, these significantly mutated genes (SMGs) mainly participated in ceramide transporter activity, sphingolipid transporter activity, intermembrane lipid transfer activity, antioxidant activity, ceramide binding, and peroxidase activity (all p -adjusted <0.05). Additionally, “amyotrophic lateral sclerosis (ALS)” (p -adjusted = 0.012624) and “axon guidance” (p -adjusted = 0.032708) are the KEGG pathways that were mainly enriched by SMGs.

Discussion

Interpretation

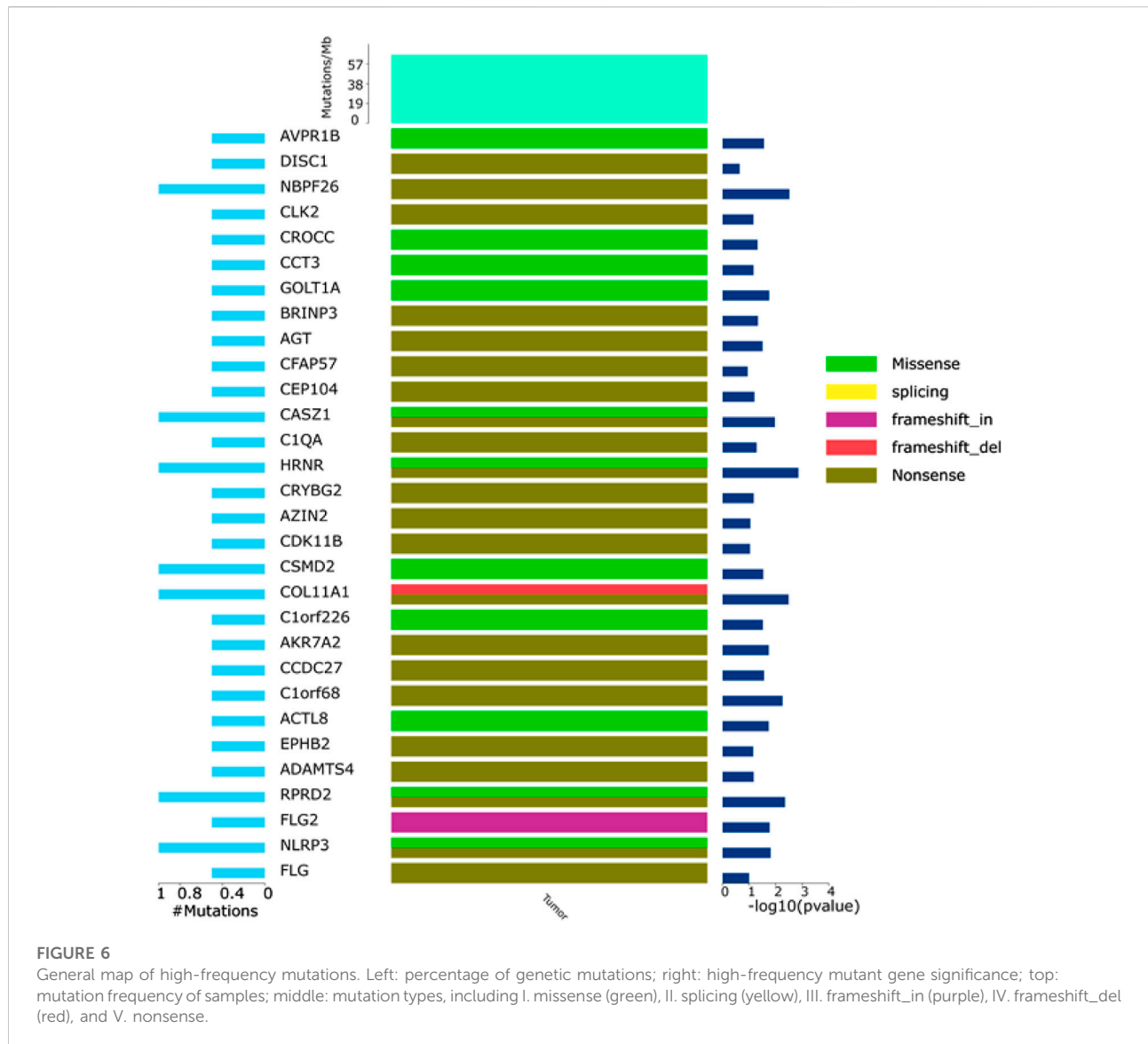
The incidence of HAL is extremely low, accounting for only about 5% of HAC. Among them, smoking men above middle age are the most commonly affected individuals, and most patients have elevated AFP levels. In recent years, some patients without elevated AFP levels have occasionally appeared. Most of the lesions in patients with primary HAL appear in the upper lobe of the lung,

resulting in the symptoms and imaging findings of the patients being very similar to those of other types of lung cancer, but HE and immunohistochemical staining of the tumor pathology showed HCC-like features (Zhuansun et al., 2021). Therefore, to avoid misdiagnosis, CT examination and HE and immunohistochemical staining should be conducted together to help in the diagnosis (Terracciano et al., 2003). Ishikura et al. (1990) proposed two diagnostic criteria for HAL at that time, namely, typical acinar or papillary adenocarcinoma, most of which were like HCC, and either the expression of AFP was positive or there was a significant increase of AFP in the serum content. However, the aforementioned clinical diagnostic criteria are imperfect, and occasional cases of HAL without elevated AFP had been reported over the next few years. In the first case of HAL proposed by Haninger et al. (2014), there was no increase in AFP, and the pathological manifestations were tissue neuroendocrine carcinoma and signet ring cells. Therefore, Haninger redefined the criteria for HAL diagnosis. First, the tumor can be a typical acinar or papillary adenocarcinoma, signet ring cell carcinoma, or containing neuroendocrine function. Second, if it expresses markers of liver differentiation, the abnormality in the patient should be diagnosed as HAL, no matter the AFP level is elevated or not. A thorough review of the published cases showed that patients with early diagnosis of HAL had longer postoperative survival, and the current treatment for HAL is the same as that for other lung cancers; radiotherapy and chemotherapy are effective treatments for patients with advanced HAL (Zhuansun et al., 2021). However, owing to the lack of analysis of gene mutations in HAL cases, the progress of new gene-targeted therapy has been limited. Li et al. (2019) conducted whole-genome sequencing (WGS) in a HAL patient with elevated AFP and detected a FAT1 driver gene mutation. FAT1 is one of the common mutations in malignant tumors, which inhibits tumor growth through the activation signal of Hippo signaling (Li and Guan, 2021) and promotes tumor migration by inducing actin polymerization (Katoh, 2012; Coudray et al., 2018). However, there are no effective drugs against FAT1 mutation. Chen et al. (2019) reported a case of HAL harboring an EGFR mutation and being responsive to tyrosine kinase inhibitor (TKI) therapy (Harrison et al., 2020). Also, Chen et al. (2020) reported a case of HAL with a KRAS driver mutation, but due to the lack of effective drugs against KRAS mutations, the patient experienced fifth-grade pneumonia and died after 6 months of anti-PD-1 treatment. Sun et al. (2022) performed next-generation sequencing (NGS) on three primary HAL cases and one HAS-metastatic HAL case and detected TP53 mutations in all four patients, although TP53 is a particularly common driver mutation. There have been several tumor studies targeting p53 inactivation (Chen et al., 2021; Liu et al., 2022), but there are no effective therapeutic agents for TP53 mutations currently, with only three compounds having reached phase III clinical trials: the mutant p53-reactivating drugs, APR-246 and COTI-2, and the MDM2 antagonist idasanutlin (Duffy et al., 2022). Meanwhile, CDK8, CKDN2A, CSF1R, EPHA5, PKHD1, SMARCA4, and



STK11 were detected as high-frequency mutations with a mutation rate of 50%. The expression of AFP in the immunohistochemistry of one out of four patients was also negative, and the gene mutations detected by sequencing were TP53, CDKN2A, STK11, and INPP4B, which were different from our sequencing results. In our study, only two nonsynonymous mutations, variant (c.1300C>A) in

NLRP3 and variant (c.731C>T) in PBX1, were detected. NLRP3, the sensor component of the NLRP3 inflammasome, plays a crucial role in innate immunity and inflammation (Nakanishi et al., 2017). Inflammasomes can also induce pyroptosis, an inflammatory form of programmed cell death (Mitoma et al., 2013; Mishra et al., 2022). Hamarsheh summarized the therapeutic potential of targeting

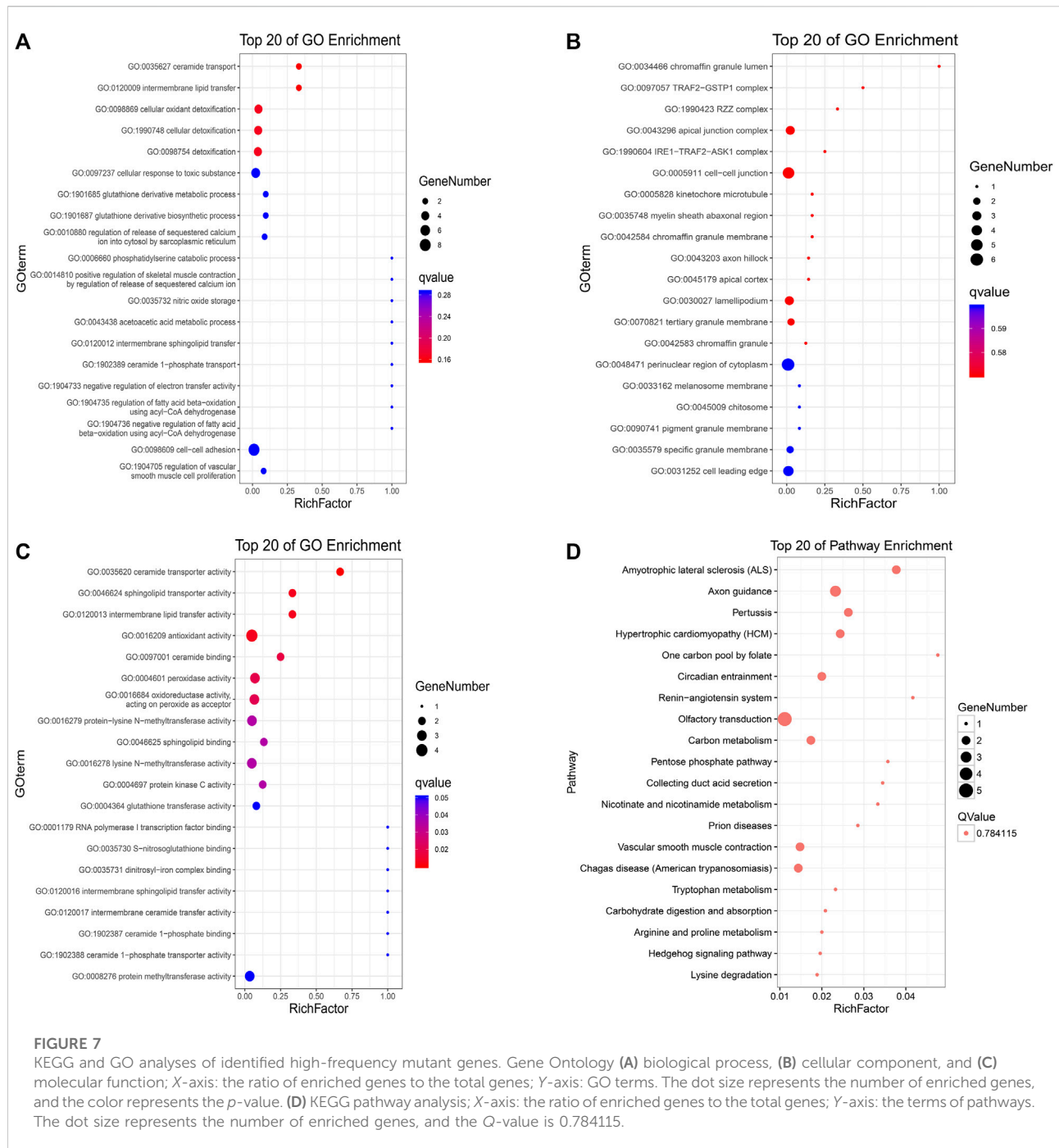


NLRP3 inflammasome in cancer (Hamarsheh and Zeiser, 2020). Moreover, some small-molecule compounds have anti-inflammatory effects that can be designed for drugs, such as thalidomide (Franks et al., 2004) and VX-765 (Wannamaker et al., 2007). PBX1 belongs to the PBX homeobox family of transcriptional factors. The *E2A-PBX1* fusion gene caused by t (McKenna et al., 2010; Mazouz et al., 2015) (q23; p13) can be found in many types of cancers, which makes it an attractive potential target for the design of new drugs (Mo et al., 2013).

Given the similar pathological features of HAL, hepatocellular carcinoma (HCC), and lung adenocarcinoma (LUAD), there may be similar disease progression among them. We compared the gene mutation information of NLRP3 and PBX1 in HCC (data from TCGA, PanCancer Atlas with 372 samples) and LUAD (data from TCGA,

PanCancer Atlas with 566 samples) through the cBioPortal database (<https://www.cbioportal.org/>) and analyzed the relationship between HAL-related driver gene mutations. NLRP3 is altered in 11% of 566 cases in the mutation data from TCGA-LUAD samples and 2.2% of 366 cases in the mutation data from TCGA-HCC samples, and PBX1 is altered in 1.8% of 566 cases in the mutation data from TCGA-LUAD samples and 1.6% of 372 cases in the mutation data from TCGA-HCC samples. After comparing the co-expression information between mutated NLRP3, PBX1, FAT1, EGFR, KRAS, and TP53 in HCC and LUAD tissues through the cBioPortal database, we found that the samples with both NLRP3 and TP53 mutations were the most in LUAD (Figure 8).

Despite these studies showing that some gene mutations lead to the occurrence and development of HAL, its potential



mechanism needs to be explored. Comprehensively considering the preference of driver mutations in the locus distribution to form mutation clusters, we used synonymous mutations to construct a background mutation rate model with the characteristics of an unbiased distribution to find possible driver mutations using OncodriveCLUST software. We identified HNRNPR, TP73, CFAP57, COL11A1, RUSC1,

SLC6A9, DISC1, NBPF26, and OR10K1 as potential driver mutation genes in HAL.

HNRNPR was originally identified as a component of the heterogeneous nuclear ribonucleoproteins (hnRNPs) family, which plays a vital role in many aspects of (pre)mRNA processing (Dreyfuss et al., 2002). HNRNPR expression is altered in many malignancies, suggesting that it is associated

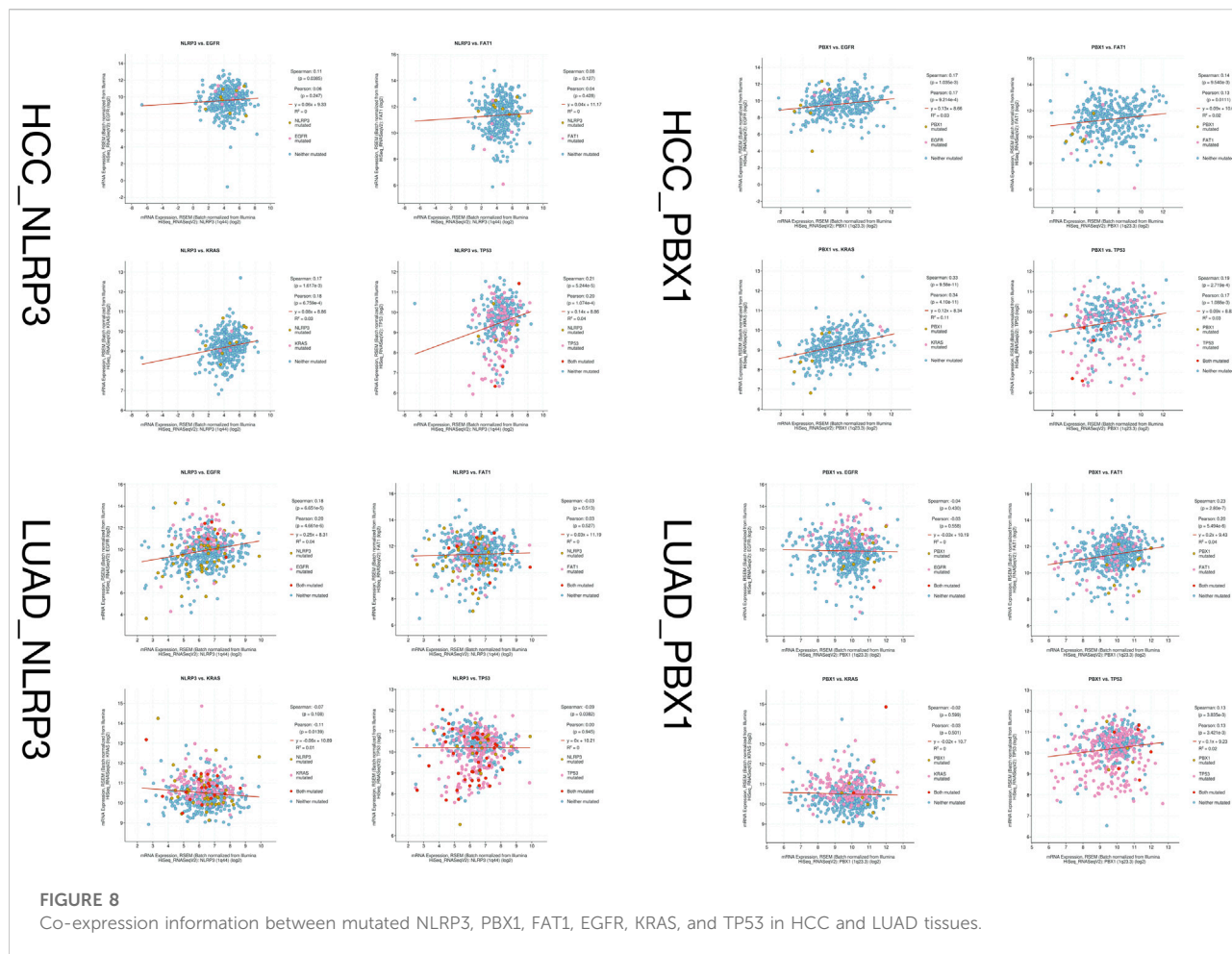


FIGURE 8 Co-expression information between mutated NLRP3, PBX1, FAT1, EGFR, KRAS, and TP53 in HCC and LUAD tissues.

with tumor formation (Pereira et al., 2017). The characteristics of HNRNPR in neurodevelopment have also been reported in many literature works (Dombert et al., 2014), which explain the occurrence of neuroendocrine differentiation in tumors of HAL. TP73 is a member of the TP53 family, mutations of which are found in many tumors and are associated with prognosis. TP73 can transform multiple variants with oncogenic and tumor-suppressive functions (Jost et al., 1997). Collagen type XI $\alpha 1$ (COL11A1), a minor fibrillar collagen, which plays a crucial role in cell proliferation, migration, and tumorigenesis of many malignancies, may be a valuable diagnostic marker for lung cancer. Smad signaling is activated by the overexpression of COL11A1 and promotes cell proliferation and migration (Shen et al., 2016). DISC1, an oncogene, activates the GSK3 β / β -catenin signaling to promote NSCLC growth, so it could be designed to be a novel anti-NSCLC therapeutic target (Wang et al., 2017), which means that HAL has the same mutation gene and differentiation process as NSCLC does.

In this study, a comprehensive interaction analysis of GO and KEGG was performed among 76 significantly mutated

genes to gain a more specific understanding of tumor genome mutation enrichment and biological functions. GO-cellular components revealed that these high-frequency mutated genes were mainly related to cell, cell part, and organelle. The GO-biological process showed these genes were related to cellular process and single-organism process, and GO-molecular function showed that these genes were related to binding function. Additionally, KEGG pathway enrichment revealed that the mutant genes mainly participated in “amyotrophic lateral sclerosis (ALS)” and “axon guidance” These findings suggest that these pathways that mutant gene aggregation enriched might play a vital role in the tumorigenesis process.

Limitation and strength

This study has several limitations, especially the small size of the sample due to the rarity of HAL, and second, there were some false-negative results caused by the inadequate sequencing coverage inevitably, despite the sequencing depth and

coverage of WES being much more improved than the Sanger sequencing. Finally, there is heterogeneity in different parts of the tumor, and the tissue obtained from the biopsy is so finite that it is impossible to cover the whole tumor genome. However, it was the first time that the WES was conducted on a non-AFP elevated HAL patient to explore the potential mechanisms of tumorigenesis. Meanwhile, the gene mutations are valuable to be the prospective candidates for gene-targeted therapy.

Further clinical guidance

Cancer is caused by abnormal cell division and differentiation due to genetic mutation (Heng and Heng, 2022); therefore, inhibiting mutations in genes related to tumor formation is the most fundamental approach for treating cancer, and with the increasing maturity of next-generation sequencing technology, more and more gene-specific research studies are being conducted in cancer diagnosis and treatment. In this study, our goal was to identify rare deleterious variants and biological pathways of HAL mutant gene enrichment that will guide future genetic and functional studies on the pathogenesis of this rare malignancy.

In the past, cancer was classified according to its morphological characteristics. To date, understanding of cancer has reached the molecular level, and based on next-generation sequencing technology, individualized precision medicine is becoming more widespread. Therefore, it is necessary to recommend targeted drugs for patients with different genetic backgrounds.

Conclusion

In summary, the rarity of HAL making large-scale trials difficult to organize, we used WES to reveal molecular patterns of HAL without an elevated AFP level. As we know, this is the first time that NLRP3 and PBX1 mutations were found in HAL without an elevated AFP level, which suggests the underlying mechanisms of tumorigenesis. Clinical treatment of HAL should not be based only on the treatments of common lung cancer or HCC; individualized and precise treatment should be implemented for its unique driver gene mutations, and our research reveals the potential mechanism of its development and provides a new potential target for treating this rare tumor.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and

accession number(s) can be found at: <https://www.ncbi.nlm.nih.gov/>, SUB11172444.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of China Medical University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

Design of the research: XZ and YY; acquisition, analysis, and interpretation of data: YY and GB; drafting the manuscript: YY and TL; revision of the manuscript for important intellectual content: YY, JL, GB, and XZ.

Funding

This study was supported by the Jieping Wu Medical Foundation Special Fund for Clinical Research (320.6750.2020-17-7).

Acknowledgments

The authors thank Guangzhou Gene Denovo Biotechnology Co., Ltd. For assisting in sequencing and bioinformatics analyses.

Conflict of interest

The 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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., et al. (2010). A method and server for predicting damaging missense mutations. *Nat. Methods* 7 (4), 248–249. doi:10.1038/nmeth0410-248
- Boeva, V., Popova, T., Bleakley, K., Chiche, P., Cappel, J., Schleiermacher, G., et al. (2012). Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* 28 (3), 423–425. doi:10.1093/bioinformatics/btr670
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34 (17), i884–i890. doi:10.1093/bioinformatics/bty560
- Chen, H. F., Wang, W. X., Li, X. L., Xu, C. W., Du, K. Q., Zhu, Y. C., et al. (2019). Hepatoid adenocarcinoma of the lung with EGFR mutation and the response to tyrosine kinase inhibitors. *J. Thorac. Oncol.* 14 (10), e217–e219. doi:10.1016/j.jtho.2019.04.032
- Chen, L., Han, X., Gao, Y., Zhao, Q., Wang, Y., Jiang, Y., et al. (2020). Anti-PD-1 therapy achieved disease control after multiline chemotherapy in unresectable KRAS-positive hepatoid lung adenocarcinoma: a case report and literature review. *Oncol. Targets. Ther.* 13, 4359–4364. doi:10.2147/OTT.S248226
- Chen, S., Thorne, R. F., Zhang, X. D., Wu, M., and Liu, L. (2021). Non-coding RNAs, guardians of the p53 galaxy. *Semin. Cancer Biol.* 75, 72–83. doi:10.1016/j.semcancer.2020.09.002
- Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., and Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels. *PLoS One* 7 (10), e46688. doi:10.1371/journal.pone.0046688
- Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., et al. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* 31 (3), 213–219. doi:10.1038/nbt.2514
- Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., et al. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* 24 (10), 1559–1567. doi:10.1038/s41591-018-0177-5
- Dees, N. D., Zhang, Q., Kandoth, C., Wendl, M. C., Schierding, W., Koboldt, D. C., et al. (2012). MuSiC: identifying mutational significance in cancer genomes. *Genome Res.* 22 (8), 1589–1598. doi:10.1101/gr.134635.111
- Dombert, B., Sivadasan, R., Simon, C. M., Jablonka, S., and Sendtner, M. (2014). Presynaptic localization of Smn and hnRNP R in axon terminals of embryonic and postnatal mouse motoneurons. *PLoS One* 9 (10), e110846. doi:10.1371/journal.pone.0110846
- Dreyfuss, G., Kim, V. N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.* 3 (3), 195–205. doi:10.1038/nrm760
- Duffy, M. J., Synnott, N. C., O'Grady, S., and Crown, J. (2022). Targeting p53 for the treatment of cancer. *Semin. Cancer Biol.* 79, 58–67. doi:10.1016/j.semcancer.2020.07.005
- Franks, M. E., Macpherson, G. R., and Figg, W. D. (2004). Thalidomide. *Lancet* 363 (9423), 1802–1811. doi:10.1016/S0140-6736(04)16308-3
- Gardiner, G. W., Lajoie, G., and Keith, R. (1992). Hepatoid adenocarcinoma of the papilla of Vater. *Histopathology* 20 (6), 541–544. doi:10.1111/j.1365-2559.1992.tb01044.x
- Grossman, K., Beasley, M. B., and Braman, S. S. (2016). Hepatoid adenocarcinoma of the lung: Review of a rare form of lung cancer. *Respir. Med.* 119, 175–179. doi:10.1016/j.rmed.2016.09.003
- Hamarsheh, S., and Zeiser, R. (2020). NLRP3 inflammasome activation in cancer: a double-edged sword. *Front. Immunol.* 11, 1444. doi:10.3389/fimmu.2020.01444
- Haninger, D. M., Kloecker, G. H., Bousamra Ii, M., Nowacki, M. R., and Slone, S. P. (2014). Hepatoid adenocarcinoma of the lung: report of five cases and review of the literature. *Mod. Pathol.* 27 (4), 535–542. doi:10.1038/modpathol.2013.170
- Harrison, P. T., Vyse, S., and Huang, P. H. (2020). Rare epidermal growth factor receptor (EGFR) mutations in non-small cell lung cancer. *Semin. Cancer Biol.* 61, 167–179. doi:10.1016/j.semcancer.2019.09.015
- Heng, J., and Heng, H. H. (2022). Genome chaos: Creating new genomic information essential for cancer macroevolution. *Semin. Cancer Biol.* 81, 160–175. doi:10.1016/j.semcancer.2020.11.003
- Hou, Z., Xie, J., Zhang, L., Dai, G., Chen, Y., and He, L. (2021). Hepatoid adenocarcinoma of the lung: a systematic review of the literature from 1981 to 2020. *Front. Oncol.* 11, 702216. doi:10.3389/fonc.2021.702216
- Ishikura, H., Kanda, M., Ito, M., Nosaka, K., and Mizuno, K. (1990). Hepatoid adenocarcinoma: a distinctive histological subtype of alpha-fetoprotein-producing lung carcinoma. *Virchows Arch. A Pathol. Anat. Histopathol.* 417 (1), 73–80. doi:10.1007/BF01600112
- Jost, C. A., Marin, M. C., and Kaelin, W. G. (1997). p73 is a simian [correction of human] p53-related protein that can induce apoptosis. *Nature* 389 (6647), 191–194. doi:10.1038/38298
- Kadalayil, L., Rafiq, S., Rose-Zerilli, M. J., Pengelly, R. J., Parker, H., Oscier, D., et al. (2015). Exome sequence read depth methods for identifying copy number changes. *Brief. Bioinform.* 16 (3), 380–392. doi:10.1093/bib/bbu027
- Kandoth, C., McLellan, M. D., Vandin, F., Ye, K., Niu, B., Lu, C., et al. (2013). Mutational landscape and significance across 12 major cancer types. *Nature* 502 (7471), 333–339. doi:10.1038/nature12634
- Kato, K., Suzuka, K., Osaki, T., Itami, M., and Tanaka, N. (2007). Primary hepatoid adenocarcinoma of the uterine cervix. *Int. J. Gynecol. Cancer* 17 (5), 1150–1154. doi:10.1111/j.1525-1438.2007.00901.x
- Kato, M. (2012). Function and cancer genomics of FAT family genes (review). *Int. J. Oncol.* 41 (6), 1913–1918. doi:10.3892/ijo.2012.1669
- Khozin, S., Roth, M. J., Rajan, A., Smith, K., Thomas, A., Berman, A., et al. (2012). Hepatoid carcinoma of the lung with anaplastic lymphoma kinase gene rearrangement. *J. Thorac. Oncol.* 7 (11), e29–31. doi:10.1097/JTO.0b013e3182697a23
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25 (14), 1754–1760. doi:10.1093/bioinformatics/btp324
- Li, F. L., and Guan, K. L. (2021). The two sides of Hippo pathway in cancer. *Semin. Cancer Biol.* doi:10.1016/j.semcancer.2021.07.006
- Li, J., Qi, H., Xu, B., Zhao, J., Gao, H., Ma, X., et al. (2019). Genomic profiles of a patient of pulmonary hepatoid adenocarcinoma with high AFP level: a case report. *Front. Oncol.* 9, 1360. doi:10.3389/fonc.2019.01360
- Lin, S.-F., Hsu, W.-H., and Chou, T.-Y. (2013). Primary pulmonary hepatoid carcinoma: report of a case and review of the literature. *Kaohsiung J. Med. Sci.* 29 (9), 512–516. doi:10.1016/j.kjms.2013.01.007
- Liu, C., Moten, A., Ma, Z., and Lin, H. K. (2022). The foundational framework of tumors: Gametogenesis, p53, and cancer. *Semin. Cancer Biol.* 81, 193–205. doi:10.1016/j.semcancer.2021.04.018
- Mazouz, A., Amaadour, L., Ameurtesse, H., Chban, L., Amarti, A., Kettani, F., et al. (2015). Primary hepatoid carcinoma of the ovary: a case report. *Pan Afr. Med. J.* 20, 93. doi:10.11604/pamj.2015.20.93.5953
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20 (9), 1297–1303. doi:10.1101/gr.107524.110
- Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhi, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 12 (4), R41. doi:10.1186/gb-2011-12-4-r41
- Mishra, S. R., Mahapatra, K. K., Behera, B. P., Bhol, C. S., Prahara, P. P., Panigrahi, D. P., et al. (2022). Inflammasomes in cancer: effect of epigenetic and autophagic modulations. *Semin. Cancer Biol.* 83, 399–412. doi:10.1016/j.semcancer.2020.09.013
- Mitoma, H., Hanabuchi, S., Kim, T., Bao, M., Zhang, Z., Sugimoto, N., et al. (2013). The DHX33 RNA helicase senses cytosolic RNA and activates the NLRP3 inflammasome. *Immunity* 39 (1), 123–135. doi:10.1016/j.immuni.2013.07.001
- Mo, M. L., Chen, Z., Zhou, H. M., Li, H., Hirata, T., Jablons, D. M., et al. (2013). Detection of E2A-PBX1 fusion transcripts in human non-small-cell lung cancer. *J. Exp. Clin. Cancer Res.* 32 (1), 29. doi:10.1186/1756-9966-32-29
- Nakanishi, H., Kawashima, Y., Kurima, K., Chae, J. J., Ross, A. M., Pinto-Patarroyo, G., et al. (2017). NLRP3 mutation and cochlear autoinflammation cause syndromic and nonsyndromic hearing loss DFNA34 responsive to anakinra therapy. *Proc. Natl. Acad. Sci. U. S. A.* 114 (37), E7766–e75. doi:10.1073/pnas.1702946114
- Nakashima, N., Fukatsu, T., Nagasaka, T., Sobue, M., and Takeuchi, J. (1987). The frequency and histology of hepatic tissue in germ cell tumors. *Am. J. Surg. Pathol.* 11 (9), 682–692. doi:10.1097/0000478-198709000-00004
- Ng, P. C., and Henikoff, S. (2003). SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 31 (13), 3812–3814. doi:10.1093/nar/gkg509
- Pereira, B., Billaud, M., and Almeida, R. (2017). RNA-binding proteins in cancer: old players and new actors. *Trends Cancer* 3 (7), 506–528. doi:10.1016/j.trecan.2017.05.003

- Qian, G. Q., Yin, F. Y., Li, G. X., and Chu, J. G. (2016). Hepatoid adenocarcinoma of the lung. *QJM* 109 (9), 619–620. doi:10.1093/qjmed/hcw107
- Schwarz, J. M., Rödelberger, C., Schuelke, M., and Seelow, D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* 7 (8), 575–576. doi:10.1038/nmeth0810-575
- Shen, L., Yang, M., Lin, Q., Zhang, Z., Zhu, B., and Miao, C. (2016). COL11A1 is overexpressed in recurrent non-small cell lung cancer and promotes cell proliferation, migration, invasion and drug resistance. *Oncol. Rep.* 36 (2), 877–885. doi:10.3892/or.2016.4869
- Shi, Y-F., Lu, J-G., Yang, Q-M., Duan, J., Lei, Y-M., Zhao, W., et al. (2019). Primary hepatoid adenocarcinoma of the lung in yungui plateau, China: a case report. *World J. Clin. Cases* 7 (13), 1711–1716. doi:10.12998/wjcc.v7.i13.1711
- Su, J-S., Chen, Y-T., Wang, R-C., Wu, C-Y., Lee, S-W., and Lee, T-Y. (2013). Clinicopathological characteristics in the differential diagnosis of hepatoid adenocarcinoma: a literature review. *World J. Gastroenterol.* 19 (3), 321–327. doi:10.3748/wjg.v19.i3.321
- Sun, J. N., Zhang, B. L., Li, L. K., Yu, H. Y., and Wang, B. (2016). Hepatoid adenocarcinoma of the lung without production of α -fetoprotein: a case report and review of the literature. *Oncol. Lett.* 12 (1), 189–194. doi:10.3892/ol.2016.4559
- Sun, H., Li, X., Zhang, J., and Liu, Y. (2022). Clinicopathological features and genomic profiles of hepatoid adenocarcinoma of the lung: report of four cases. *Pathol. Res. Pract.* 229, 153652. doi:10.1016/j.prp.2021.153652
- Tamborero, D., Gonzalez-Perez, A., and Lopez-Bigas, N. (2013). OncodriveCLUST: exploiting the positional clustering of somatic mutations to identify cancer genes. *Bioinformatics* 29 (18), 2238–2244. doi:10.1093/bioinformatics/btt395
- Tamborero, D., Gonzalez-Perez, A., Perez-Llamas, C., Deu-Pons, J., Kandoth, C., Reimand, J., et al. (2013). Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci. Rep.* 3, 2650. doi:10.1038/srep02650
- Terracciano, L. M., Glatz, K., Mhawech, P., Vasei, M., Lehmann, F. S., Vecchione, R., et al. (2003). Hepatoid adenocarcinoma with liver metastasis mimicking hepatocellular carcinoma: an immunohistochemical and molecular study of eight cases. *Am. J. Surg. Pathol.* 27 (10), 1302–1312. doi:10.1097/0000478-200310000-00002
- Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., and Kinzler, K. W. (2013). Cancer genome landscapes. *Science* 339 (6127), 1546–1558. doi:10.1126/science.1235122
- Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38 (16), e164. doi:10.1093/nar/gkq603
- Wang, J., Mullighan, C. G., Easton, J., Roberts, S., Heatley, S. L., Ma, J., et al. (2011). CREST maps somatic structural variation in cancer genomes with base-pair resolution. *Nat. Methods* 8 (8), 652–654. doi:10.1038/nmeth.1628
- Wang, S., Chen, Y. Y., Li, Y. P., Gu, J., Gu, S. D., Shi, H., et al. (2017). DISC1 overexpression promotes non-small cell lung cancer cell proliferation. *Oncotarget* 8 (39), 65199–65210. doi:10.18632/oncotarget.18055
- Wannamaker, W., Davies, R., Namchuk, M., Pollard, J., Ford, P., Ku, G., et al. (2007). (S)-1-((S)-2-([1-(4-amino-3-chloro-phenyl)-methanoyl]-amino)-3, 3-dimethyl-butanoyl)-pyrrolidine-2-carboxylic acid ((2R, 3S)-2-ethoxy-5-oxo-tetrahydro-furan-3-yl)-amide (VX-765), an orally available selective interleukin (IL)-converting enzyme/caspase-1 inhibitor, exhibits potent anti-inflammatory activities by inhibiting the release of IL-1 β and IL-18. *J. Pharmacol. Exp. Ther.* 321 (2), 509–516. doi:10.1124/jpet.106.111344
- Zhuansun, Y., Bian, L., Zhao, Z., Du, Y., Chen, R., Lin, L., et al. (2021). Clinical characteristics of Hepatoid adenocarcinoma of the lung: four case reports and literature review. *Cancer Treat. Res. Commun.* 29, 100474. doi:10.1016/j.ctarc.2021.100474