Int J Clin Exp Pathol 2019;12(3):740-749
www.ijcep.com /ISSN:1936-2625/IJCEP0087082

*Original Article*

# A detrimental mutation on USP40 unlocks the tumorigenesis in a rare case of lung cancer

Zhi-Hong Xu[1*], Hui Wang[3*], Xiao-Yang Ji[1*], Fei-Xu Zhang[3], Bei-Li Gao[2], Jia-An Hu[1], Jing Zheng[3]

*Departments of [1]Geriatrics, [2]Respiratory Medicine, Ruijin Hospital, Shanghai Jiaotong University, Shanghai, P. R. China; [3]Laboratory of Molecular Neuropharmacology, School of Pharmacy East China University of Science and Technology, Shanghai, P. R. China. [*]Co-first authors.*

Received October 21, 2018; Accepted November 20, 2018; Epub March 1, 2019; Published March 15, 2019

**Abstract:** Lung adenocarcinoma (LUAD) is the most common histologic subtype of lung cancer. Previous research has shown heterogeneity in lung cancer, with the parallel existence of multiple subclones characterized by their own specific mutational landscape. The aim of our study was to gain insight into the evolutionary pattern of lung cancer by investigating the genomic heterogeneity between a nodule and its distant tumor. Luckily, we obtained nodule and tumor samples derived from surgery and a blood sample from a single patient. The samples are very unique, for tissues with the same genetic background from nodules to malignant tumors are rarely available and require precise micro-cutting. In this study, we performed whole-genome sequencing of these two samples, to identify novel candidate driver genes associated with LUAD. The nodule and tumor were found to have common significant ubiquitin-specific protease 40 (USP40) mutations, indicating an important driver role for the gene. Moreover, we also observed the two novel candidate driver genes ASCL5 and CAPNS1 in the LUAD sample. In summary, we pinpoint the predominant mutations in LUAD by WES, highlighting the substantial genetic alterations contributing to LUAD tumorigenesis. This may provide a better understanding of the clonal evolution during tumor development.

**Keywords:** Lung adenocarcinoma tumorigenesis, USP40, subclonal mutations

## Introduction

Lung cancer is one of the most common cancers in humans and is the leading cause of cancer deaths in the USA and worldwide [1]. Early detection and treatment have been demonstrated to be the most effective measures to prevent lung cancer progression. According to the Nomenclature Committee of the Fleischner Society, a lung nodule has been defined as a rounded opacity, which is an approximately round lesion less than 3 cm in diameter and completely surrounded by pulmonary parenchyma. There are no associated abnormalities, and benign and malignant categories are defined by biopsy [2, 3]. Lung cancer screening tests show that the probability of having at least one nodule is 8%-51% and the prevalence of nodules as malignant tumors is 1.1-2%. Therefore, nodules are very common in the clinic and are difficult to control, because the nodule may represent primary lung cancer or other malignant or benign lesions. For example, bronchogenic carcinoma as a cause of solitary nodules has been increasing clinically, especially in the elderly [4]. When a pulmonary nodule becomes "malignant", the nodule often appears as lesions together with a malignant tumor. Therefore, blocking lung nodule progression to tumors is of great importance for the prevention of lung cancer advancement. However, tissues with the same genetic background from nodules to malignant tumors are rarely available and require precise micro-cutting. This accounts for the limited availability of specimens, which is an obstacle to research progress.

Recent advances in technology, including the radiologic, clinical and pathologic aspects primarily, have broadened our understanding and awareness of pulmonary nodules, in determining whether the lesion of pulmonary nodules is benign or malignant. We can understand the direct cause of the malignant carcinogenesis of nodules, due to the changes that have occurred

in the DNA sequence of cancer cell genomes. These changes include point mutations, copy-number changes, and fusion events that affect the function of critical genes regulating cellular growth and survival. Genome sequencing is a powerful way to discover key genes playing causal roles in oncogenesis and to identify driver genomic regions that undergo frequent alteration in human cancers [5, 6]. 'Driver' mutations confer a growth advantage on the cells carrying them and have been positively selected during cancer evolution. The identification of these mutations has greatly accelerated progress in both the understanding of cancer pathogenesis and the identification of novel therapeutic vulnerabilities [7]. By searching for a set of mutations in pulmonary nodules and malignancy samples, we can create a pedigree to describe the evolutionary history of cancer.

Recently, knowledge of the molecular basis of lung cancer has advanced at a rapid pace. The molecularly targeted therapies have dramatically improved treatment for patients whose tumors harbor somatically activated oncogenes [8]. But the evolution of a pulmonary nodule into a malignant tumor has not been characterized or provided any invasion "triggers" in the same genetic background.

Here, we examined a lung cancer patient who harbored pulmonary nodule and malignant tumor samples in his right middle lung lobe. Since these two samples had the same genetic background, they are very suitable for studying the evolution of pulmonary nodule to malignant tumor. We performed whole exon sequencing (WES) and high-throughput screening to identify somatic mutations from the pulmonary nodule and malignant lesions samples from the same patient. These results will greatly accelerate the understanding of the gene alterations that transition the pulmonary nodule into a malignant tumor and the identification of new treatment vulnerability in lung cancer.

## Materials and methods

### *Patients and sample collection*

A 76-year old male with no smoking history was diagnosed with pulmonary tuberculosis from his chest X-ray when he underwent a physical examination in 2000, and further computed tomography (CT) scanning of his chest showed that there was tuberculosis with a diameter about 7.5 mm in the right middle lung lobe. There was no significant change found in the annual chest CT until 6th June, 2008. The chest CT taken in 2008 showed that the right middle lung lobe had slightly increased in size. Its size was about 14.6 × 11 mm, with tiny glitches, adhesive with the adjacent pleural which led to the depression of the pleural (**Figure 1**). In 2009 a lesion with a diameter of 25 mm in the lateral segment of the right middle lung lobe was detected during the operation. Postoperative pathological tissue stained with hematoxylin-eosin demonstrated that the nodule was identified to be atypical adenomatous hyperplasia (AAH) with focal canceration of bronchioloalveolar carcinoma (BAC) (**Figure 2**). During the operation, we obtained two independent samples from this patient. These two samples were pathologically confirmed to be separate pulmonary nodule (PN) and lung adenocarcinoma (LUAD) lesions. This study was conducted under the oversight of the Internal Review and Ethics Boards of Shanghai Ruijin Hospital, and informed consent was obtained from all human subjects. Both the PN and LUAD samples were subjected to pathology review to confirm the pathological diagnosis.

### *Whole-exome sequencing*

Genomic DNA was extracted from the samples mentioned above, and each sample for the sequencing libraries was prepared using the DNA Sample Prep Kit (Illumina), quantified using Qubit (Invitrogen), and examined by agarose gel electrophoresis. Whole-exome capture was performed using ≥ 1.0 μg of the patient's genomic DNA based on the KAPA library preparation (Kapa Biosystems) using the Agilent Sure Select Human All Exon V4 kit and paired-end multiplex sequencing reads were generated for each sample on the Illumina HiSeq 2000 sequencing platform. All samples were obtained at the required average depth of the sequencing from 100 × to 330 ×.

### *Variant calling*

Briefly, the paired-end raw sequence reads for each sample generated from the Illumina HiSeq 2000 were converted from a bcl file to a FastQ file by bcl2fastq (v1.8.4) and mapped to the reference haploid human-genome sequence (UCSC Genome Browser, hg19) with the use of

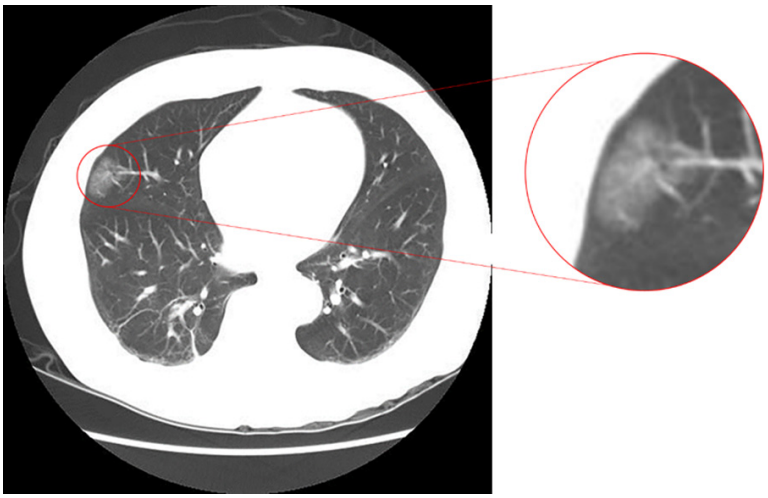

**Figure 1.** In a 76-year-old male patient, a nodule was located in the lateral segment of right middle lung lobe. Its size was about 14.6 × 11 mm, and its edge was surrounded by a fine burr with sunken adjacent pleura. There were small quasi-circular nodules scattered in the two lungs.

the Burrows-Wheeler Aligner (BWA) program with default settings.

GATK Best Practices were used for duplicate removal, indel realignment, and base recalibration to further process the data. Somatic single nucleotide substitutions were detected using MuTect in high confidence mode [9], and small indels were identified using Strelka [10]. The mutational consequences were annotated by ANNOVAR based on UCSC refGene genes. Based on combined consideration of two calling results, filtering was performed using in-house, custom Perl scripts. Filtering was applied according to the following criteria: (i) total read count in tumor DNA ≥ 20; (ii) variant allele frequency (VAF) in the normal samples = 0; and (iii) variants in positions listed in the 1000 Genomes (1000G) database with MAF<0.01 were removed.

*Identification of copy number changes*

We used PyLOH [11] and Sequenza [12] to call CNVs, and nodule and tumor samples were called against the blood sample, respectively. In PyLOH, we preprocessed the reads aliments of paired normal-tumor samples in BAM format to produce the paired count file, and produce a heatmap visualization of the preprocessed segment file and the preprocessed B allele frequency (BAF) in the samples. Then, the log2 copy ratios for the tumor versus normal reads were calculated for each tumor region after adjusting for the total mapped reads in that tumor region and the DNA copy package of Bioconductor provides the functionalities for segmenting the log ratios. In addition, the tumor purity was calculated using the software package PyLOH, which integrates somatic copy number alterations and loss of heterozygosity [11].

Sequenza, a software package that uses paired tumor-normal DNA sequencing data to estimate tumor cellularity and ploidy, improves the detection accuracy of sample purity and absolute copy number variation (CNV) calls. These data set the stage for understanding the functional impact of copy number alterations in our samples.

## Results

*Monoclonal origins of PN and LUAD*

After sequencing (100 × depth for blood and 300 × depth for nodule and tumor samples), one somatic mutation was identified in common in both the nodule and LUAD tissues. 35 mutations were determined in nodule samples, and 38 mutations were found in LUAD samples, indicating that the sampling and preparation satisfied our goal. The allele frequencies are shown in Supplementary Table 1.

*Somatic mutations and the mutation spectrum in human PN tissue sample and a LUAD samples revealed by whole-exome sequencing*

From the WES analysis, we estimated the allele frequencies of the LUAD and PN somatic mutations, respectively (**Figure 3**). We identified 39 somatic mutations in the LUAD sample, and a set of 30 genes was found with significantly higher frequencies of non-synonymous single nucleotide variants (SNVs) (77.0%), indicating a gain/loss of function for tumorigenesis. This may suggest that they were candidate tumor genes. Among the 11 possible mutation classes, the G>A transition dominated the mutation

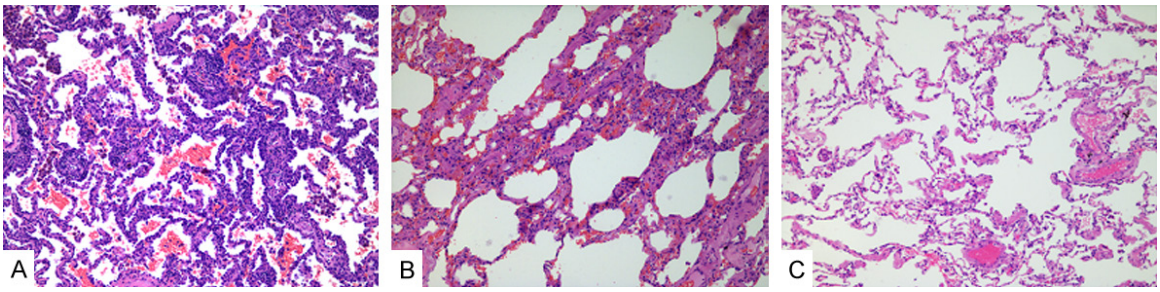


**Figure 2.** A: The tumor consists predominantly of lepidic growth without invasion (× 100). B: The adjacent non-tumor lung tissue showed proliferated fibers in the interalveolar septum (× 100). C: Normal lung tissue adjacent to the tumor (× 100).

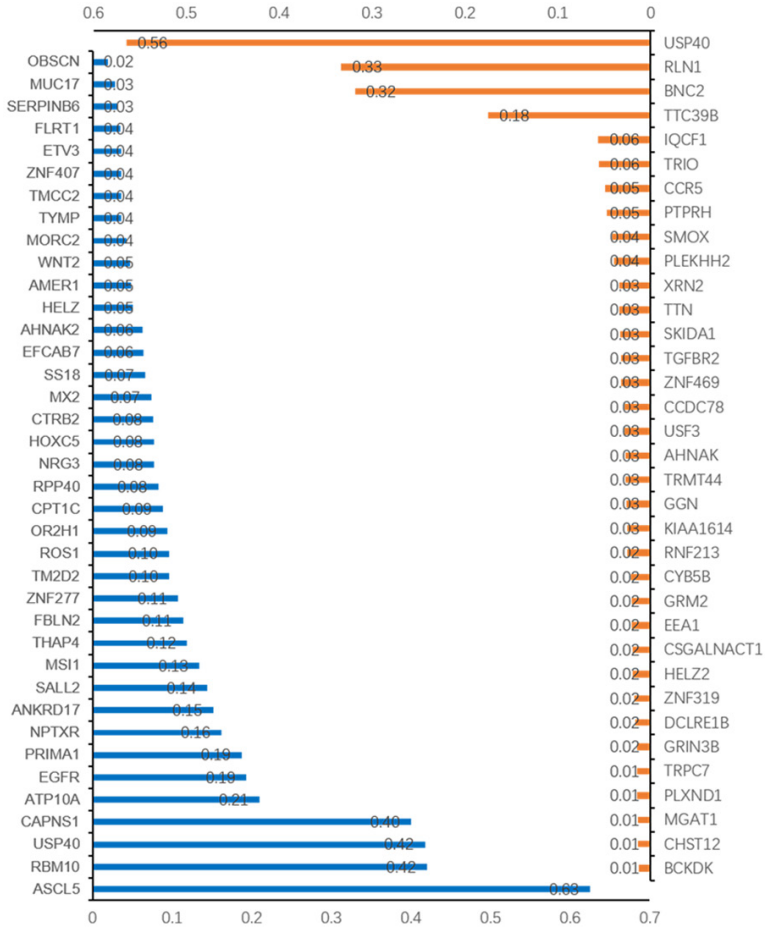


**Figure 3.** A bar chart showing the mutated genes in LUAD (blue) and PN (orange), respectively. In the X-axis the percentage of observed mutation frequency is represented. In the Y-axis the genes are listed.

spectra, with LUAD and PN accounting for 53.8% and 36.4% of the total mutations respectively. The LUAD contained more mutations than the PN, indicating more malignant biological behavior. The relative fractions of the mutation spectra in each sample are presented in **Figure 3**. The data describe the heterogeneity between the PN and LUAD samples.

*Somatic copy number variants (CNVs)*

CNV is one of the most important somatic aberrations, which were derived from a PN sample and a LUAD sample. The analysis of chromosomal segments deletion reveals homology between the PN and LUAD samples (**Figure 4**). 46 Common segments spanning 165 Mb were deleted in LUAD, while 16 Common segments spanning 2 Mb were deleted in PN. These segments covered 569 genes and 102 genes, respectively. We determined that 7 Shared segments (1 Mb) were deleted in both PN and LUAD, suggesting that the PN and LUAD samples were homologous. Then, we analyzed the amplified LOH segments to distinguish the evolutionary direction of the tumors from the nodules. In PN, 55 common segments spanning 97 Mb were amplified, and 13 common segments spanning 75 Mb were LOH. These segments covered 481 genes and 557 genes, respectively. In LUAD, 54 common segments spanning 110 Mb were amplified, and 113 common segments spanning 498 Mb were LOH. These segments cover

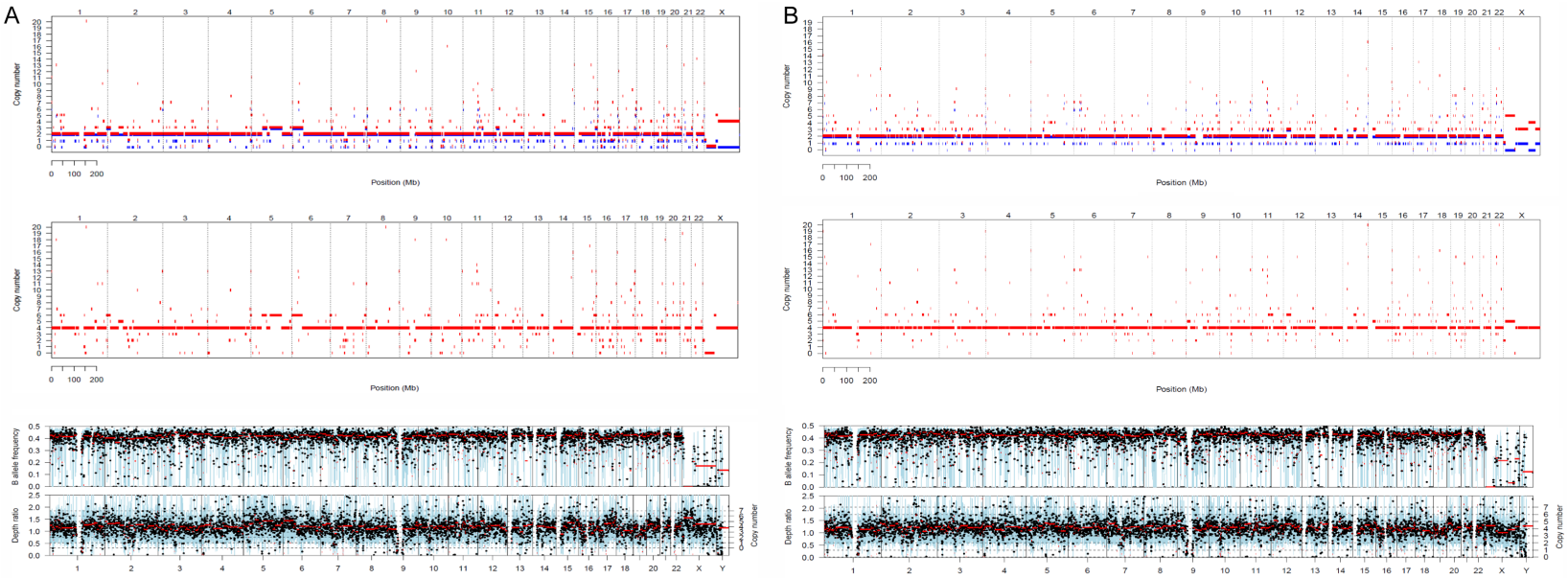


**Figure 4.** To demonstrate the patterns of LRR and BAF in regions with copy number changes, we plotted these values on chromosomes from LUAD and PN.

**Table 1.** The analysis of CNV the amplified/LOH/deleted shared segments in both the PN and LUAD samples

| | Chromosome | Start.pos | End.pos | Cytoband |
|---|---|---|---|---|
| Amplification | 4 | 100867977 | 101344012 | q23-q24 |
| | 5 | 376994 | 491826 | p15.33 |
| | 8 | 106013469 | 143310815 | q22.3-q24.3 |
| | 8 | 143381810 | 144457678 | q24.3 |
| | 9 | 68431137 | 68781905 | q13-q21.11 |
| | 11 | 64981522 | 65268726 | q13.1 |
| | 11 | 65349063 | 65561732 | q13.1 |
| | 11 | 65562257 | 66003139 | q13.1- q13.2 |
| | 11 | 67352256 | 67434010 | q13.2 |
| | 14 | 106877618 | 107083574 | q32.33 |
| | 17 | 62541865 | 63185658 | q23.3-q24.1 |
| | 17 | 73732025 | 106877556 | q25.1-q25.3 |
| Deletion | 1 | 12853403 | 13747815 | p36.21 |
| | 3 | 129811323 | 129830064 | q22.1 |
| | 4 | 9217205 | 9369758 | p16.1 |
| | 9 | 141071438 | 141121531 | q34.3 |
| | 16 | 90167905 | 90233687 | q24.3 |
| | 17 | 60342311 | 60351603 | q23.2 |
| | X | 49162087 | 49365125 | p11.23 |
| LOH | 1 | 146423054 | 146501334 | q21.1 |
| | 1 | 153588537 | 154561648 | q21.3 |
| | 2 | 131935207 | 132124541 | q21.1 |
| | 3 | 33060230 | 33618944 | p22.3 |
| | 8 | 86556026 | 86839873 | q21.2 |
| | X | 61687294 | 134571642 | q11.1-q26.3 |

460 genes and 1,411 genes, respectively. The segments were considered a "shared segment" when amplification or LOH was found in the samples from the PN and LUAD samples. We determined that 12 shared segments (74 Mb) were amplified and 6 Shared segments (74 Mb) were LOH in both PN and LUAD (**Table 1**).

*A comparison of VAFs for PN and LUAD*

We determined VAFs for the gr-oup of infrequently mutated genes at PN and LUAD. The PN sample carried 4 gene mutations in the primary clone, including USP40, RLN1, BNC2, and TTC39B. In the LUAD samples, 4 non-synonymous mutations in the genes, including ASCL5, RBM10, USP40, and CAPNS1, were identified, indicating that those mutations may happen in the early stage of tumorigenesis. The USP40 mutation was discovered in both PN and LUAD, supporting a common evolutionary origin. Besides, ATP10A, EGFR, and PRIMA1 nonsynonymous mutations were considered as the subclones, which may contain true, albeit infrequent drivers of tumorigenesis that, together with the primary clones, constitute candidate genes. Of additional interest is the fact that EGFR mutation was present in the tumor subclone (VAF = 0.19), suggesting that EGFR gene mutation is not a primary clone in tumor evolution.

Phylogenetic tree analysis based on somatic mutation and copy number variation was used to study the clonal relationships among the different regions of PN and LUAD. To reveal the evolutionary relationship between the PN and LUAD samples, we speculated that there are several possible scenarios for this evolution (see in **Figure 5**).

*The functional genes in phylogenetic tree*

Then, we characterized functional pathways affected by these genes. The USP40 and CAPNS1 gene related to 'cysteine-type peptidase activity' were enriched in the primary clone, and the USP40 gene belongs to a large family of cysteine proteases which are deubiquitinating enzymes that specifically recognize and remove ubiquitin from proteins [13]. CAPN4 (CAPNS1) can enhance the invasion ability of lung cancer cells by upregulating the expression of matrix metalloproteinase 2 (MMP2) [14]. The ASCL5, RBM10, and EGFR genes related to the 'regulation of transcription from RNA polymerase II promoter' were enriched. The ASCL5 gene serves as a potential oncogene during lung cancer development, and it has a gain of function mutation [15]. Recent results indicate that RBM10 is an alternative splicing regulator [16, 17] that modulates alternative splicing of the notch regulator gene NUMB, and the overexpression of RBM10 could inhibit lung adenocarcinoma cell malignant behaviors in vitro. The heterozygous deletion of ATP10A in mice causes diet-induced obesity, and we suggested it was a gain of function

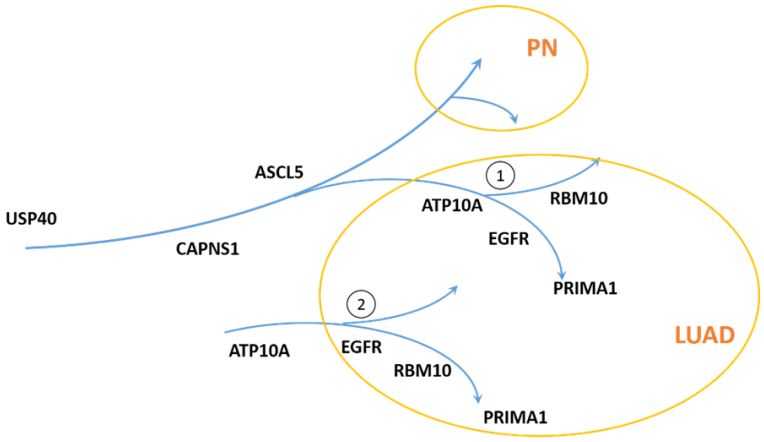


**Figure 5.** Phylogenetic tree analysis of PN evolution to LUAD using somatic mutations and copy number aberrations.

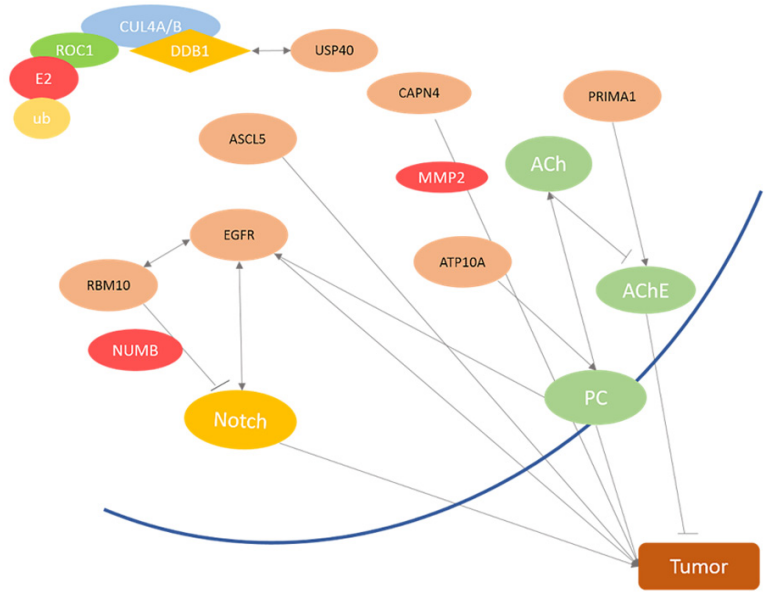


**Figure 6.** The relationships between gene mutations revealed an evolutionary process of nodules into tumors.

mutation [18]. PRIMA1 is a transmembrane protein that encodes a proline-rich transmembrane protein that efficiently transforms secreted acetylcholinesterase (AChE) into an enzyme anchored on the outer cell surface [19]. Epidermal growth factor receptor (EGFR) is a protein on the surface of cells. Research has proven that the EGFR group of growth factors are key molecules that promote lung cancer generation and propagation [20] and EGFR mutations also affect diverse growth factor signaling pathways in lung neoplasms, which ultimately lead to aggressive lung carcinogenesis and metastasis [21].

## Discussion

The incidence of indeterminate pulmonary nodules has risen constantly over the past few years. PN in an at-risk population is an alerting signal of possible lung cancer, so blocking lung nodule progression to tumors is of great importance for the prevention of lung cancer advancement. Without the knowledge of the key alterations that drive PN into LUAD, to predict the risk of developing lung cancer is a daunting task. In the era of precision medicine, molecular testing in lung cancer has led to remarkable improvements in response and survival. However, the study of oncogenic mutations in the evolution of malignant nodules in lung adenocarcinomas has not had much success because of the lack of sufficient samples with the same genetic background. Luckily, we found two independent PN and LUAD samples in one single patient. These samples seem to be particularly suitable to investigate the primary driving factors that determine the progression from PN to LUAD.

From these samples, we identified some genes that may play potential roles in tumors and the relationships between gene mutations in tumor evolution (**Figure 6**). USP40 mutations were found with high VAF in both PN and LUAD, suggesting that USP40 plays a crucial role in nodules and tumors. USP40 is a deubiquitinating enzyme (DUB) widely expressed in various human tissues [22]. It belongs to a large family of cysteine proteases that function as deubiquitinating enzymes, which counteract the ubiquitylation of selected target proteins through the removal of ubiquitin from these molecules. Ubiquitination, a reversible post-translational modification, is involved in many biological processes such as cell cycle progression, DNA-repair, organelle biogenesis,

vesicular trafficking, transcriptional activation, signal transduction, and intracellular proteolysis. There is increasing evidence that altered DUB function is associated with the pathogenesis of multiple tumors [23]. For example, USP1 expression correlates with the initial steps of transformation in gastric cancer [24], whereas USP2 is overexpressed in ovarian and prostate carcinomas [25]. USP7 overexpression in prostate cancer has also been associated with tumor aggressiveness [26]. High levels of USP17 have been identified in primary lung, colon, esophagus, and cervix tumor biopsies [27]. In addition, the association between several USPs such as USP 39 [28], USP 37 [29], USP4 [30], USP 36 [31], USP33 [32], and USP22 [33] with lung cancers has been studied. However, USP40 hasn't been reported to be related to lung cancer, meaning the novelty of our research. Researchers have found that USP40 and DDB1 are interacting proteins, and DDB1 primarily functions as a core component of the CUL4A- and CUL4B-based E3 ubiquitin ligase complexes [34, 35]. The dynamic balance between E3 ligase and DUB activity within a cell regulates key cellular activities including apoptosis, cell division, and cell signaling [36]. The USP40 mutation affects the reversal of protein ubiquitination mediated by the E3 ligase, therefore undermining the stability or activity of the target protein.

In addition, CAPNS1 and ASCL5 are also the major clonal mutations in LUAD. That make sense because CAPNS1 enhances the invasion ability of lung cancer cells by upregulating the expression of matrix metalloproteinase 2 (MMP2), and plays an important role in the development of malignant tumors [14]. The study demonstrated that alterations in the expression of ASCL genes may affect cellular behavior such as cell proliferation, thereby initiating tumor development. However, few studies determine a potential oncogenic role for ASCL5 in lung cancer. Only a recent one suggested a correlation between elevated ASCL5 expression and lung cancer development. Besides, the other four driver subclonal mutations, including RBM10, ATP10A, PRIMA1 and EGFR, may also play a role in lung cancer.

The results of our study revealed an evolutionary process of nodules into tumors. To our knowledge, few studies show genetic changes of nodules into tumors by using high-throughput sequencing. Here, we identified a special evolving pattern of tumors and the exact genes that might be important for LUAD tumorigenesis and progression. However, a single case study has some limitations. It remains unclear whether these conclusions can be extended to other patients. Therefore, further studies are needed to validate and explore the roles of mutated genes in the development of this kind of tumor so that it may provide an avenue to guide LUAD precision medicine therapy and prognosis.

## Acknowledgements

This study was supported by Grands of Shanghai Municipal Commission of Health, Family Planning for Key Discipline Establishment and Charity Projects (201840083), the Fundamental Research Funds for the Central Universities and Shanghai Construction Fund for Important Weak Disciplines (2015ZB0503).

## Disclosure of conflict of interest

None.

**Address correspondence to:** Jing Zheng, Laboratory of Molecular Neuropharmacology, School of Pharmacy East China University of Science and Technology, 130 Meilong Road, Shanghai 200-37, P. R. China. Tel: 86-21-64253032; E-mail: zhengjing@ecust.edu.cn; Jia-An Hu, Department of Geriatrics, Ruijin Hospital, Shanghai Jiaotong University, Shanghai, P. R. China. Tel: 86-21-64370045; E-mail: hujiaaan@163.com

## References

[1] Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T and Thun MJ. Cancer statistics, 2008. CA Cancer J Clin 2008; 58: 71-96.

[2] Ost D, Fein AM and Feinsilver SH. Clinical practice. The solitary pulmonary nodule. N Engl J Med 2003; 348: 2535-2542.

[3] Tuddenham WJ. Glossary of terms for thoracic radiology: recommendations of the nomenclature committee of the fleischner society. AJR Am J Roentgenol 1984; 143: 509-517.

[4] Tan BB, Flaherty KR, Kazerooni EA and Iannettoni MD. The solitary pulmonary nodule. Univ Mich Med Cent J 2003; 123: 54.

[5] Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N and Stratton MR. A census of human cancer genes. Nat Rev Cancer 2004; 4: 177-183.

[6] Stuart D and Sellers WR. Linking somatic genetic alterations in cancer to therapeutics. Curr Opin Cell Biol 2009; 21: 304-310.
[7] Stratton MR, Campbell PJ and Futreal PA. The cancer genome. Nature 2009; 458: 719.
[8] Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature 2014; 511: 543-550.
[9] Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES and Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013; 31: 213-219.
[10] Saunders CT, Wong WS, Swamy S, Becq J, Murray LJ and Cheetham RK. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. Bioinformatics 2012; 28: 1811-1817.
[11] Li Y and Xie X. Deconvolving tumor purity and ploidy by integrating copy number alterations and loss of heterozygosity. Bioinformatics 2014; 31: 2121-2129.
[12] Molenaar JJ, Koster J, Zwijnenburg DA, van Sluis P, Valentijn LJ, van der Ploeg I, Hamdi M, van Nes J, Westerman BA, van Arkel J, Ebus ME, Haneveld F, Lakeman A, Schild L, Molenaar P, Stroeken P, van Noesel MM, Ora I, Santo EE, Caron HN, Westerhout EM, Versteeg R. Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. Nature 2012; 483: 589-593.
[13] Takagi H, Nishibori Y, Katayama K, Katada T, Takahashi S, Kiuchi Z, Takahashi SI, Kamei H, Kawakami H and Akimoto Y. USP40 gene knockdown disrupts glomerular permeability in zebrafish. Am J Physiol Renal Physiol 2017; 312: F702-F715.
[14] Gu J, Xu FK, Zhao GY, Lu CL, Lin ZW, Ding JY and Ge D. Capn4 promotes non-small cell lung cancer progression via upregulation of matrix metalloproteinase 2. Med Oncol 2015; 32: 51.
[15] Wang CY, Shahi P, Huang JT, Phan NN, Sun Z, Lin YC, Lai MD and Werb Z. Systematic analysis of the achaete-scute complex-like gene signature in clinical cancer patients. Mol Clin Oncol 2017; 6: 7-18.
[16] Wang Y, Gogol-Döring A, Hao H, Fröhler S, Ma Y, Jens M, Maaskola J, Murakawa Y, Quedenau C and Landthaler M. Integrative analysis revealed the molecular mechanism underlying RBM10-mediated splicing regulation. EMBO Mol Med 2013; 5: 1431-1442.
[17] Inoue A, Yamamoto N, Kimura M, Nishio K, Yamane H and Nakajima K. RBM10 regulates alternative splicing. FEBS Lett 2014; 588: 942-947.
[18] Dhar MS, Sommardahl CS, Kirkland T, Nelson S, Donnell R, Johnson DK and Castellani LW. Mice heterozygous for Atp10c, a putative amphipath, represent a novel model of obesity and type 2 diabetes. J Nutr 2004; 134: 799.
[19] Hildebrand MS, Tankard R, Gazina EV, Damiano JA, Lawrence KM, Dahl HH, Regan BM, Shearer AE, Smith RJ, Marini C, Guerrini R, Labate A, Gambardella A, Tinuper P, Lichetta L, Baldassari S, Bisulli F, Pippucci T, Scheffer IE, Reid CA, Petrou S, Bahlo M, Berkovic SF. PRIMA1 mutation: a new cause of nocturnal frontal lobe epilepsy. Ann Clin Transl Neurol 2015; 2: 821-830.
[20] Carcereny E, Morán T, Capdevila L, Cros S, Vilà L, de Los Llanos Gil M, Remón J and Rosell R. The epidermal growth factor receptor (EGRF) in lung cancer. Transl Respir Med 2015; 3: 1.
[21] Hodkinson PS, Mackinnon A and Sethi T. Targeting growth factors in lung cancer. Chest 2008; 133: 1209-1216.
[22] Quesada VC, Díaz-Perales A, Gutiérrez-Fernández A, Garabaya C, Cal S, López-Otín C. Cloning and enzymatic analysis of 22 novel human ubiquitin-specific proteases. Biochem Biophys Res Commun 2004; 314: 54-62.
[23] Fraile JM, Quesada V, Rodríguez D, Freije JM and López-otín C. Deubiquitinases in cancer: new functions and therapeutic options. Oncogene 2012; 31: 2373-88.
[24] Luise C, Capra M, Donzelli M, Mazzarol G, Jodice MG, Nuciforo P, Viale G, Fiore PP and Confalonieri S. An atlas of altered expression of deubiquitinating enzymes in human cancer. PLoS One 2011; 6: e15891.
[25] Yang Y, Hou JQ, Qu LY, Wang GQ, Ju HW, Zhao ZW, Yu ZH and Yang HJ. [Differential expression of USP2, USP14 and UBE4A between ovarian serous cystadenocarcinoma and adjacent normal tissues]. Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi 2007; 23: 504-506.
[26] Song MS, Salmena L, Carracedo A, Egia A, Lo-Coco F, Teruya-Feldstein J and Pandolfi PP. The deubiquitinylation and localization of PTEN are regulated by a HAUSP-PML network. Nature 2008; 455: 813-817.
[27] Mcfarlane C, Kelvin AA, de la Vega M, Govender U, Scott CJ, Burrows JF and Johnston JA. The deubiquitinating enzyme USP17 is highly expressed in tumor biopsies, is cell cycle regulated, and is required for G1-S progression. Cancer Res 2010; 70: 3329-39.
[28] Lin Z, Xiong L and Lin Q. Ubiquitin-specific protease 39 is overexpressed in human lung cancer and promotes tumor cell proliferation in vitro. Mol Cell Biochem 2016; 422: 97-107.
[29] Pan J, Deng Q, Jiang C, Wang X, Niu T, Li H, Chen T, Jin J, Pan W and Cai X. USP37 directly deubiquitinates and stabilizes c-Myc in lung cancer. Oncogene 2014; 34: 3957-3967.
[30] Su JH, Lee HW, Kim HR, Hong L, Chang HS, Yun SI, Dong HL, Kim DH, Kim KK and Joo KM. Ubiquitin-specific protease 4 controls meta-

static potential through β-catenin stabilization in brain metastatic lung adenocarcinoma. Sci Rep 2016; 6: 21596.
[31] Sun XX, He X, Yin L, Komada M, Sears RC and Dai MS. The nucleolar ubiquitin-specific protease USP36 deubiquitinates and stabilizes c-Myc. Proc Natl Acad Sci U S A 2015; 112: 3734-3739.
[32] Wen P, Kong R, Liu J, Zhu L, Chen X, Li X, Nie Y, Wu K and Wu JY. USP33, a new player in lung cancer, mediates Slit-Robo signaling. Protein Cell 2014; 5: 704-713.
[33] Ning J, Zhang J, Liu W, Lang Y, Xue Y and Xu S. Overexpression of ubiquitin-specific protease 22 predicts poor survival in patients with early-stage non-small cell lung cancer. Eur J Histochem 2012; 56: e46.
[34] Chen X, Zhang Y, Douglas L and Zhou P. UV-damaged DNA-binding proteins are targets of CUL-4A-mediated ubiquitination and degradation. J Biol Chem 2001; 276: 48175-48182.
[35] Yan H, Bi L, Wang Y, Zhang X, Hou Z, Wang Q, Snijders AM and Mao JH. Integrative analysis of multi-omics data reveals distinct impacts of DDB1-CUL4 associated factors in human lung adenocarcinomas. Sci Rep 2017; 7: 333.
[36] Macdonald C, Winistorfer S, Pope RM, Wright ME and Piper RC. Enzyme reversal to explore the function of yeast E3 ubiquitin-ligases. Traffic 2017; 18: 465-484.

**Supplementary Table 1.** The analysis of CNV the amplified/LOH/deleted shared segments in both PN and LUAD sample

| No. | Mutated site | MutGene | VAF | Class | Frq in TCGA | Star | Recurrent | CN | B | A | LOH |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2_234428301_234428301_C_T | USP40 | 56.45% | 3 | 0.011146 | 0 | 1 | 4 | 2 | 2 | 0 |
| 2 | 9_5339544_5339544_G_T | RLN1 | 33.33% | 2 | 0.002322 | 0 | 0 | 5 | 3 | 2 | 0 |
| 3 | 9_16436544_16436544_A_C | BNC2 | 31.85% | 2 | 0.018731 | 0 | 0 | 4 | 2 | 2 | 0 |
| 4 | 9_15188018_15188018_T_C | TTC39B | 17.50% | 2 | 0.006037 | 0 | 0 | 5 | 3 | 2 | 0 |
| 6 | 3_51937003_51937003_C_T | IQCF1 | 5.63% | 1 | 0.005573 | 0 | 1 | 4 | 2 | 2 | 0 |
| 7 | 5_14508083_14508083_T_C | TRIO | 5.56% | 1 | 0.027245 | 0 | 0 | 4 | 2 | 2 | 0 |
| 8 | 3_46415061_46415061_G_A | CCR5 | 4.91% | 1 | 0.005418 | 0 | 0 | 4 | 2 | 2 | 0 |
| 9 | 19_55716742_55716742_C_T | PTPRH | 4.72% | 1 | 0.016409 | 0 | 0 | 4 | 2 | 2 | 0 |
| 10 | 20_4162736_4162736_G_T | SMOX | 4.13% | 1 | 0.006656 | 0 | 0 | 4 | 2 | 2 | 0 |
| 11 | 2_43927538_43927538_C_A | PLEKHH2 | 3.93% | 1 | 0.018885 | 0 | 1 | 4 | 2 | 2 | 0 |
| 12 | 20_21327129_21327129_C_T | XRN2 | 3.43% | 1 | 0.010836 | 0 | 0 | 4 | 2 | 2 | 0 |
| 13 | 2_179449186_179449186_G_A | TTN | 3.35% | 1 |  | 0 | 0 | 4 | 2 | 2 | 0 |
| 14 | 10_21805036_21805036_G_C | SKIDA1 | 3.30% | 1 |  | 0 | 0 | 4 | 2 | 2 | 0 |
| 15 | 3_30713619_30713619_C_T | TGFBR2 | 3.22% | 1 | 0.015325 | 1 | 0 | 4 | 2 | 2 | 0 |
| 16 | 16_88496531_88496531_C_G | ZNF469 | 3.21% | 1 |  | 0 | 0 | 4 | 2 | 2 | 0 |
| 17 | 16_774386_774386_G_A | CCDC78 | 2.83% | 1 | 0.00356 | 0 | 0 | 4 | 2 | 2 | 0 |
| 18 | 3_113375683_113375683_C_G | USF3 | 2.82% | 1 |  | 0 | 0 | 4 | 2 | 2 | 0 |
| 19 | 11_62296535_62296535_T_C | AHNAK | 2.75% | 1 | 0.048762 | 0 | 0 | 4 | 2 | 2 | 0 |
| 20 | 4_8469917_8469917_C_T | TRMT44 | 2.68% | 1 | 0.001238 | 0 | 0 | 4 | 2 | 2 | 0 |
| 21 | 19_38877255_38877255_G_T | GGN | 2.57% | 1 | 0.006502 | 0 | 0 | 4 | 2 | 2 | 0 |
| 22 | 1_180905294_180905294_C_T | KIAA1614 | 2.52% | 1 | 0.012074 | 0 | 1 | 4 | 2 | 2 | 0 |
| 23 | 17_78350159_78350159_T_C | RNF213 | 2.49% | 1 | 0.035294 | 0 | 0 | 4 | 2 | 2 | 0 |
| 24 | 16_69458591_69458591_C_T | CYB5B | 2.11% | 1 |  | 0 | 0 | 4 | 2 | 2 | 0 |
| 25 | 3_51746780_51746780_G_A | GRM2 | 2.02% | 1 | 0.012074 | 0 | 0 | 4 | 2 | 2 | 0 |
| 26 | 12_93171826_93171826_G_A | EEA1 | 2.02% | 1 | 0.013003 | 0 | 0 | 4 | 3 | 1 | 0 |
| 27 | 8_19363306_19363306_G_A | CSGALNACT1 | 1.93% | 1 | 0.008204 | 0 | 0 | 4 | 2 | 2 | 0 |
| 28 | 20_62191974_62191974_G_A | HELZ2 | 1.90% | 1 |  | 0 | 0 | 4 | 2 | 2 | 0 |
| 29 | 16_58030741_58030741_G_A | ZNF319 | 1.78% | 1 | 0.005108 | 0 | 0 | 4 | 2 | 2 | 0 |
| 30 | 1_114454062_114454062_G_A | DCLRE1B | 1.67% | 1 | 0.006347 | 0 | 0 | 4 | 2 | 2 | 0 |
| 31 | 19_1005322_1005322_C_T | GRIN3B | 1.51% | 1 | 0.008669 | 0 | 0 | 4 | 2 | 2 | 0 |
| 32 | 5_135692493_135692493_C_T | TRPC7 | 1.47% | 1 | 0.016873 | 0 | 1 | 4 | 2 | 2 | 0 |
| 33 | 3_129289655_129289655_C_T | PLXND1 | 1.46% | 1 | 0.016718 | 0 | 0 | 5 | 3 | 2 | 0 |
| 34 | 5_180219451_180219451_G_A | MGAT1 | 1.40% | 1 | 0.004334 | 0 | 0 | 4 | 2 | 2 | 0 |
| 35 | 7_2472555_2472555_C_T | CHST12 | 1.39% | 1 | 0.004954 | 0 | 0 | 6 | 3 | 3 | 0 |
| 36 | 16_31120842_31120842_G_A | BCKDK | 1.31% | 1 | 0.004799 | 0 | 0 | 4 | 2 | 2 | 0 |