Introduction

Lung cancer is the leading cause of cancer-related death worldwide (1). The two major histological types are non-small-cell lung cancer (NSCLC) and small-cell lung cancer (SCLC) (2). Lung squamous cell carcinoma (LUSC), one of the common histological types of NSCLC, remains poor prognosis despite of development in therapeutic strategies (3-5). Meanwhile, intratumoral heterogeneity, which refers to heterogeneity among tumor cells of a single patient, is crucial for the clinical outcome of patients with lung cancer, impacting the curative effect of chemotherapy, radiotherapy and immunotherapy (6,7).

Next-generation sequencing (NGS), a method relying on short reads, has been performed on multiregional
tumors to explore intratumoral genetic heterogeneity (ITGH) in NSCLC (8-10). Previous studies focused more on ITGH involving mutations that distinguish different tumor cells in a single or multiple primary NSCLC (7-9,11). A previous study explored the ITGH based on analysis of single nucleotide variants (SNVs) and copy number variants (CNVs) using whole-genome sequencing (WGS) on primary tumors, metastatic lymph nodes and tumor cells in the pleura (10). Because of the challenge in detecting technology, structural variants (SVs) increasingly appears to have an indispensable but undiscovered role in ITGH (12,13). However, ITGH which manifests uneven distribution of genetic alterations among lung tumor cells in primary tumor and associated metastases is not comprehensively characterized due to the lack of studies focusing on distant metastasis and SVs. Recently, optical mapping, a newly non-sequencing method, shed a light to dig large SVs (14,15).

In this study, we combined optical mapping and WGS to reveal the ITGH in various forms of SNVs, indels and SVs, especially large SVs (>5 kb) within primary tumor and associated metastases in a LUSC patient. We also compared SVs detected by optical mapping and those detected by WGS. Furthermore, after comparing the genes affected by variants with those associated with tumorigenesis and progression, we inferred the functional consequence of distinct genomic alterations among tumor cells within the primary site and paired metastatic sites.

Methods

Tissue collection

Surgical specimens of primary tumor (PT), lymph node metastases (LNM), tumor thrombus in the pulmonary vein (TPV) and adjacent normal lung tissue (at least 2cm away from tumor) were obtained from a patient who diagnosed with pathologically confirmed lung squamous cell carcinoma. This study was approved by the Committee for Ethical Review of Research. Informed consent was obtained.

Whole-genome sequencing

DNA extraction and sequencing: After fragmented by sonication to a size of 350 bp, genomic DNA fragments were end-polished, A-tailed, and ligated with adapter for Illumina sequencing. Then after further PCR amplification and purification, libraries were analyzed for size distribution by Agilent 2100 Bioanalyzer and quantified for concentration (2 nM) by fluorigenic-quantitative PCR (Qubit 2.0). Then DNA libraries were sequenced on Illumina Novaseq 6000 sequencing platform with 30X sequencing depth. 150 bp paired-end reads were generated. Contaminated reads including adaptors, low quality reads and those with more “N” was extracted based on chastity score and quality score.

Variants detection and filtration: Paired-end reads in FastQ format were aligned to the reference human genome (UCSC Genome Browser, version hg19) by Burrows-Wheeler Aligner (BWA) (16). Subsequent BAM files were processed by SAMtools (17), Picard tool (http://picard.sourceforge.net/), and the Genome Analysis Toolkit (GATK) (18) to sort and remove duplication, local realignment, and base quality recalibration.

SNVs and indels detection: Mutect (19) was used to detect the somatic SNVs and indel with tumor-normal paired BAM files. ANNOVAR was used to further annotate for VCF (Variant Call Format) (20). Somatic SNVs were further filtered for analysis of mutational spectrum and signatures with the following criteria: SNVs which has no record in 1000 Genomes project, dbsnp or Berry4000 (Berry Genomics) were filtered (21,22).

SVs detection, filtration and classification: Manta was applied for SVs detection (23), SVs were reported as INS (insertion), DEL (deletion), DUP (duplication), INV (inversion), and BND (further identified as inter-chromosomal translocation). Somatic SVs in PT, LNM and TPV were identified with the data of adjacent normal lung sample as control. ANNOVAR was applied for annotation (20). SVs were filtered if: SVs <50 bp; mapped to the mitochondrial genome or chromosome Y; overlapped with gap region, telomere, centromere or low complexity regions; with MinQUAL, MinGQ, Ploidy, MaxDepth, MaxMQ0Frac and NoPairSupport in VCF FILTER fields; and supported by <2 split reads (SR).

Optical mapping

DNA preparation: High Molecular Weight (HMW) DNA were extracted using Bionano Prep Animal Tissue DNA Isolation Fibrous Tissue Protocol (https://bionanogenomics.com/support-page/animal-tissue-dna-isolation-kit/) from the tissue of frozen PT, LNM and TPV. Firstly, approximately 10 mg of tissue were fixed, disrupted with a rotor-stator, embedded in 2% agarose, and digested.
with proteinase K and RNase. After multiple stabilization and recovery followed by digestion with Agarase (Thermo Fisher) enzyme, HMW DNA were released, cleaned by drop dialysis and homogenized. HMW DNA were quantitated using Qubit dsDNA BR Assay Kit.

Direct labeling: HMW DNA were extracted using Bionano Prep Direct Label and Stain (DLS) Protocol (https://bionanogenomics.com/support-page/dna-labeling-kit-dls/). Firstly, 750 ng HMW DNA were nicked by DLE-1 enzyme, recovered, labeled with fluorophore and stained. Then labeled and stained DNA were quantitated using modified Qubit dsDNA HS (High Sensitivity) Assay Kit. Each labeled sample was added to a BioNano Saphyr Chip (Bionano Genomics) and run on the Bionano Saphyr instrument, targeting 100x human genome coverage. The raw data were filtered by Bionano Access (v1.2.1) with the following criteria: molecule length >150 kb with average label density of 10–25/100 kb.

SVs detection and filtration: De novo assembly of long molecules into genome map and SVs detection by comparing with Hg19 were performed with software Bionano Solve (version 3.2.1). SVs were annotated by Enliven (Berry Genomics). Then SVs were filtered if: for translocation and inversion, (I) confidence value <0.9, (II) breakpoints were located in the chromosome fragile site, (III) breakpoints were located in the segmental region of the chromosome, (IV) breakpoints were within these previously identified SVs (24); For insertion and deletion, (I) confidence value <0.9, (II) length of variation <5 kb, (III) breakpoints were in the gap region of reference genome.

Comparison of SVs from optical mapping and WGS
WGS provide SVs breakpoints (start and end) with base pair resolution, while optical mapping provides only the nearest labeling site to the interval of SVs. We determined whether SVs from optical mapping overlap with SVs from WGS with the following criteria: (I) Deletions, insertions and duplications detected by WGS must overlap with the interval of SVs detected by optical mapping. (II) The breakpoints of Inversions detected by WGS must lie within 500 kb to the interval of SVs detected by optical mapping.

Comparison of SVs from WGS among PT, LNM and TPV
Somatic SVs from WGS in PT, LNM and TPV were classified as shared SVs or private SVs among tumors with the following criteria: SVs has the same breakpoints (start and end), consistent type with SVs in another tumor were identified as identical and classified as shared SVs.

Comparison of SVs from optical mapping among PT, LNM and TPV
SVs from optical mapping in PT, LNM and TPV were classified as shared or private SVs among tumors with the following criteria: SVs have overlapped interval, consistent type with SVs in another tumor were identified as shared SVs. We further filtered the shared SVs in all tumors due to the shared somatic SVs and germline SVs could not be distinguished.

Identification of genes affected by SVs
For variants from WGS, we inferred a gene affected by variants if (I) a protein coding gene is annotated with an exon-annotated deletion, insertion and duplication; (II) the breakpoint (start or end) of inversion or inter-chromosome translocation lies within one or more exon of the genes; (III) the genes carried an nonsynonymous variants (nonsynonymous SNVs or frameshifting indels).

For SVs from optical mapping, we inferred a gene affected by variants if the gene was annotated with an exon-annotated SVs.

Functional consequence analysis
For genes affected by variants, we inferred whether these genes are associated with tumorigenesis and progression based on data of lung cancer driver genes (25-27), pan-cancer driver genes (28), COSMIC (https://cancer.sanger.ac.uk/census) (29), DNA repair genes (30) and hallmark genes of epithelial-mesenchymal transition (EMT) (31-38). Based on the data of The Human Protein Atlas (www.proteinatlas.org) (39-41), we further examined whether RNA expression of these genes correlate with the outcome of lung cancer and its protein expression and classified them as unprognostic, prognostic favorable and prognostic unfavorable genes.

KEGG enrichment
Genes only affected by variants in LNM and TPV were used to KEGG enrichment analysis by The Database for Annotation, Visualization and Integrated Discovery (DIVID) (42) and KOBAS 3.0 (http://kobas.cbi.pku.edu.cn/
Statistical analysis

We used R (version 3.3.3, version 3.6.1) software. “SomaticSignatures”, “ggplot2”, “ggrepel”, “ggthemes” were used in the analyses (43,44).

Results

Patients’ characterization

A 50-year-old East Asian male with 20 pack year history of smoking for 20 years, was diagnosed with lung squamous cell carcinoma with histopathological confirmation (Figure 1). Before systematic treatment, primary tumor (PT) located in the left upper lobe of lung, metastasis of left lower paratracheal (4L) lymph node (LNM) and tumor thrombus of the left Superior pulmonary vein (TPV) were sampled by surgical section. Furthermore, there is no reported family history of lung cancer. No significant difference in Tumor grade heterogeneity among tumor cells in primary and metastatic sites were identified by hematoxylin and eosin staining (Figure 1C, Figure S1).

ITGH in the form of SNVs and indels

To gain an insight into alterations of different mutational characteristics between the primary tumor and the metastases, we performed WGS on PT, LNM, TPV and adjacent normal lung tissue at an average depth of 30X.

A total of 268 nonsynonymous somatic variants (including nonsynonymous SNVs and frameshifting indels) in 252 genes were identified in at least one tumor (Table S1), and 14.2% (38) of these variants were shared between PT and either one of the two metastases (Figure 2 and Figure 3A). Among them, 3 mutations were common in all tumors, while compared with...
Figure 2 Exonic somatic variants identified in PT, LNM and TPV. The exonic somatic variants were classified as shared or private variants. Red color represent genes contain different variants among different tumors. PT, primary tumor; LNM, lymph node metastases; TPV, tumor thrombus in pulmonary vein.
Figure 3 Intratumoral genetic heterogeneity in form of SNVs and indels. (A) The number of exonic somatic variants (SNVs and indels) and nonsynonymous somatic variants in each of tumors. (B) The mutation spectrum of SNVs in PT, LNM and TPV. (C) Mutational signatures of all tumor sample. (D) Two mutational signatures (S1, S2) extracted from all tumors. (E) Cluster analysis of S1, S2 and 30 COSMIC mutational signature based on the cosine similarity. (F) The proportion of S1 and S2 in PT, LNM and TPV. PT, primary tumor; LNM, lymph node metastases; TPV, tumor thrombus in pulmonary vein.

LNM (5), a larger number of mutations (36) in TPV were shared with PT. 17, 15 and 195 mutations were uniquely seen in PT, LNM and TPV, respectively. Specifically, nonsynonymous SNV in TP53 which is one of the most commonly mutated gene in LUCC (45) were only detected in TPV. We further analyzed the mutation spectrum of SNVs (Figure 3A,B,C), trying to identify significant discordance between LNM and TPV. To be specific, we identified that TPV and PT both displayed a predominance of cytosine-adenine (C > A) nucleotide transversions which implied a
correlation with tobacco exposure (46), consistent with the long-term smoking history of this patient. Meanwhile, the LNM exhibited a distinct preponderance of guanine-adenine (G > A) and adenine-guanine (A > G). Moreover, the detailed analysis of mutational signature showed S1 and S2 were extracted (Figure 3D). Compared with the previously known mutational signatures shown in COSMIC (29), S1 had the most similarity with signature 4 likely due to direct damage by mutagens in tobacco, and S2 exhibits the thymine-cytosine (T > C) as same as the signature 5 increased in many cancer types due to tobacco smoking (Figure 3E). Primary tumor and metastasis shared identical mutational signatures, but the proportion is different (Figure 3F). These results demonstrated patient have primary tumor and metastasis in different sites has high ITGH in the form of SNVs and indels.

Comparison of structural variants detected by WGS and optical mapping

We utilized WGS data and performed optical mapping on PT, LNM and TPV at 100X coverage. SVs were called and filtered as presented in Figure 4. There were a mean of 3,617 SVs detected by WGS (3,907, 3,580, and 3,365 in PT, LNM, and TPV, respectively), of which deletions were most commonly detected type of SV (Figure S2). While SVs detected by optical mapping was 1,026 on average (979, 1,118, 980 in PT, LNM, TPV, respectively), Insertions account for the most (Figure S2).

By comparing the SVs detected by WGS and optical mapping, we observed an average of 22.9 percent of SVs detected by optical mapping overlapped with those detected by WGS (25.1%, 21.4% and 22.2% in PT, LNM and TPV, respectively), of which the deletions had similar size (the median size was 6,452 bp, 6,191 bp in optical mapping and WGS) (Figure 5C, Figure S3). The median size of non-overlapping SVs in optical mapping was distinct from the non-overlapping ones detected by WGS (8,875 bp, 143 bp in optical mapping and WGS respectively) (Figure 5C, Figure S3). Specifically, Optical mapping is more capable of detecting large SVs (>5,000 bp) (Figure 5D). Generally, WGS can detect SVs at a high resolution of
base but has many limitations: it depends on a short-read sequencing technique, needs a reference genome, and challenges of computational and bioinformatics algorithms exist. In contrast, optical mapping detects large and complex SVs using high molecular weight (HMW) DNA which are longer, ranging from 0.1 to 2Mb. The results suggested that the combination of WGS and optical mapping used for detecting SVs allows to a more comprehensive understanding of structural variants among tumor cells within different sites and demonstrated optical mapping is more sensitive for detection of large SVs.

**ITGH in the form of SVs**

We did an comparison among PT, LNM and TPV based on SVs detected by WGS and SVs detected by optical mapping, identifying a greater amount of private SVs in TPV (126 from WGS, 83 from optical mapping) than in either PT (4 from WGS, 75 from optical mapping) or LNM (4 from WGS, 118 from optical mapping) (Figure 6A), consistent with the results of SNVs and indels analysis. There was no overlap between private SVs identified by WGS and private SVs identified by optical mapping in each of tumors except TPV (7 private SVs from optical mapping overlapped with 6 private SVs from WGS). Smaller number of SVs in TPV (17 from WGS, 23 from optical mapping) overlapped with SVs of PT than those in LNM (105 from optical mapping). Specifically, 52 SVs from optical mapping undetected in PT were shared between LNM and TPV.

We further explored whether these SVs overlap with genes previously associated with tumorigenesis and progression (Figure 6B). Several private SVs of TPV detected by either WGS or optical mapping were associated with DNA repair genes including APEX2, FANCA, FANCB and RAD9A suggesting that mutations in DNA repair genes may play a role in progression of metastatic lung cancer by generating chromosomal instability. We also identified several EMT associated genes including BASP1, LAMA2, SAT1, SERPINH1 and TIMP1 were affected by SVs only detected in TPV. Completely different with TPV, only CSMD3, a frequently mutated gene in LUSC (47,48) was affected by private SVs of LNM. Loss of CSMD3 was
Figure 6 Intratumoral genetic heterogeneity in form of structural variants. (A) Overlap of structural variants detected by whole-genome sequencing (upper) and optical mapping (lower) among PT, LNM and TPV. (B) Genes associated with tumorigenesis and progression affected by structural variants detected by whole-genome sequencing and optical mapping in PT, LNM and TPV. (C) Genes associated with prognosis of lung cancer affected by structural variants detected by whole-genome sequencing and optical mapping. (Red dotted line represents P value >0.05) (D) KEGG enrichment of genes only affected by metastases-specific structural variants. (Red dotted line represents adjusted P value >0.05). PT, primary tumor; LNM, lymph node metastases; TPV, tumor thrombus in pulmonary vein.

reported to be associated with the proliferation of airway epithelial cells (47) and mutations in CSMD3 is associated with a better prognosis in patients with LUSC (48). Compared with the gene expression and survival data in The Human Protein Atlas (HPA) (39-41), we also identified 21 other genes affected by SVs previously unrecognized as tumor associated genes, of which expression was significantly associated with the prognosis of lung cancer patients (Figure 6C).

Furthermore, to comprehensively understand the functional consequence of genomic alterations only found in tumor cells in metastatic sites, we performed a KEGG enrichment analysis based on genes only affected by SNVs, indels and SVs in metastases (Figure 6D). Specifically, genes involved in the PI3K-Akt pathway which has an important role in tumorigenesis and progression (49), were significantly affected by variants in TPV.

Discussion

SNVs and CNVs detected by next-generation sequencing in multiregional tumors has improved our understanding of ITGH (8-10,46,50), while studies focusing on the analysis of ITGH in the form of SVs among tumor cells in primary and different metastatic sites are limited. Previous studies detected SVs through WGS (51,52). WGS, relying on sequencing by synthesis, is based on short reads. The DNA molecules are fragmented to countless reads and amplified by polymerase chain reaction (PCR), to meet the requirement of the high-throughput. And then we detect the SVs based on the read-pair or SR. That is, WGS detects the SVs on the basis of incomplete structure of DNA, which
may miss some SVs in specific locations of chromosome or those with large size (53). In contrast, the integrity of DNA molecular is crucial for optical mapping to detect the SVs, with specific site labeled HMW DNA and nano-channel imaging system, optical mapping could de novo identify SVs without the bias of PCR amplification. Therefore, optical mapping and WGS could complement mutually.

To our knowledge, our study is the first study applying WGS and optical mapping to multiregional samples of a LUSC patient, aiming to compressively investigate the intratumoral heterogeneity within one patient. We do observe a significant difference in the variants burden between primary tumor and metastases and between metastases in different sites. Like SNVs and indels, SVs play an indispensable role in heterogeneity. Combination of WGS and optical mapping allows us to gain a more comprehensive understanding of structural variants, especially large SVs. Compared with the analysis of SVs detected by WGS, optical mapping were more informative in identifying private SVs for ITGH.

Variants shared between primary tumor and metastases indicate that mutations in primary tumor subclones with metastatic potential accumulated before metastasizing. Among them, mutations shared between TPV and PT which affect genes associated with tumorigenesis and progression, may enable tumor cells in the primary site to metastasize and live in hemato-microenvironment. Tumor cells harbor mutations identified both in PT and TPV may have more capability to metastasize and settle down in lymph node.

Meanwhile, private variants detected in different groups of tumors suggest genetic mutations occurred both before and after metastasis. Mutations unique to LNM or TPV indicate an interaction between tumor cells and microenvironment in metastatic sites. Private variants in TPV, especially those affected genes associated with DNA repair and epithelial-mesenchymal transition (EMT), are much more frequently identified than in PT or LNM. This suggests that tumor cells in hemato-microenvironment bear a higher degree of chromosomal instability and has more potential to act as a metastases relay station between primary tumor and metastases of distant organs, previously observed by Ferronika et al. (54).

It should be noted that the major limitation of our study is that analysis only based on one individual. The main reason is that most LUSC patients received surgery are at early stage and non-metastatic. In clinical practice, metastatic lymph node and tumor thrombus collected from the same patient in this study is rare to obtain by surgical resection. And biopsy sampling of multiple metastatic regions has not been widely accepted due to the potential risks for the prognosis of patients (55). Additionally, previous studies confirmed that analysis in a small number of cases even in one patient could reveal ITGH (6,10,15).

Notwithstanding its limitation, our results do demonstrate the ability of optical mapping in detection of large SVs to make up the deficiency of WGS and reveal that SVs are as crucial in describing ITGH as SNVs and indels.

**Acknowledgments**

We thank the patient to provide the samples for this study; Litao Han and Ben Ma for advice to manuscript. We also thank Lili Tan for excellent technical assistance; Hainan Cheng for bioinformatics analysis.

**Funding:** This work was supported by Ministry of Science and Technology of the People’s Republic of China (2017YFA0505500; 2016YFA0501800), Science and Technology Commission of Shanghai Municipality (19XD1401300).

**Footnote**

**Conflicts of Interest:** All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tlcr-19-401). The authors have no conflicts of interest to declare.

**Ethical Statement:** The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Fudan University Shanghai Cancer Center Institutional Review Board (No. 090977-1) and written informed consent was obtained from all patients.

**Open Access Statement:** This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the
References


Figure S1 Postoperative paraffin section and hematoxylin and eosin (H&E) staining image for PT (A and B), LNM (C) and TPV (D) based on 40–100x magnification. PT, primary tumor; LNM, lymph node metastases; TPV, tumor thrombus in pulmonary vein.
## Table S1

Somatic nonsynonymous SNVs and indels detected in PT, LNM and TPV:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome</th>
<th>Position</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HERC1</td>
<td>20</td>
<td>248737734</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>OR51A4</td>
<td>11</td>
<td>118005739</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>VAMP1</td>
<td>3</td>
<td>32713598</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>CYLD</td>
<td>4</td>
<td>143269952</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>TPV</td>
<td>19</td>
<td>195509939</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>FCRL5</td>
<td>14</td>
<td>1306817</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>TRAM1L1</td>
<td>15</td>
<td>225422494</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>MIOS</td>
<td>22</td>
<td>23040922</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>CRACR2A</td>
<td>13</td>
<td>32647032</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>MUC12</td>
<td>21</td>
<td>150269712</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>IQGAP2</td>
<td>17</td>
<td>33954985</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>SLC37A3</td>
<td>20</td>
<td>32757165</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>C10orf67</td>
<td>10</td>
<td>100644211</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>TPSD1</td>
<td>16</td>
<td>148577938</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>TYK2</td>
<td>1</td>
<td>143269952</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>TPCN2</td>
<td>3</td>
<td>32757165</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>RYR3</td>
<td>1</td>
<td>72341009</td>
<td>Frameshift insertion</td>
<td></td>
</tr>
<tr>
<td>TESMIN</td>
<td>1</td>
<td>72341009</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>ZNF99</td>
<td>13</td>
<td>74363307</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>ADGRA1</td>
<td>3</td>
<td>54745682</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>MAPK11</td>
<td>1</td>
<td>23040922</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
<tr>
<td>ZNF723</td>
<td>12</td>
<td>47870310</td>
<td>Nonsynonymous SNV</td>
<td></td>
</tr>
</tbody>
</table>
Figure S2 The proportions of different types of SVs detected by whole-genome sequencing (left) or optical mapping (right) in PT (upper), LNM (middle) and TPV (lower). PT, primary tumor; LNM, lymph node metastases; TPV, tumor thrombus in pulmonary vein.
Figure S3 The number of different types of structural variants detected by whole-genome sequencing and optical mapping in PT (A) and LNM (C), of which size distribution of deletions in PT (B) and LNM (D). PT, primary tumor; LNM, lymph node metastases.